

UNIVERSITÉ PARIS-SUD  
ÉCOLE DOCTORALE : SCIENCES DU VÉGÉTAL

Laboratoire de génétique végétale du Moulon

**DISCIPLINE : génétique**

**THÈSE DE DOCTORAT**

Soutenue le 15 janvier 2014

par

**Pierre RIVIÈRE**

---

**Méthodologie de la sélection décentralisée et  
participative : un exemple sur le blé tendre.**

---

**Composition du jury :**

Directrice de thèse	Isabelle GOLDRINGER	Directrice de recherche (INRA)
Rapporteurs	Pierre M. STASSART	Professeur (Université de Liège)
	Magali SAN CRISTOBAL	Directrice de Recherche (INRA)
Examineurs	Jane LECOMTE	Professeur (Université Paris-Sud)
	Jacques DAVID	Professeur (SupAgro Montpellier)
	Pierre-Henri GOUYON	Professeur (MNHN, Science Po, AgroParisTech, ENS)
Membre invité	Jean-François BERTHELLOT	Paysan-boulangier (Réseau Se- mences Paysannes)





UNIVERSITÉ PARIS-SUD  
ÉCOLE DOCTORALE : SCIENCES DU VÉGÉTAL

Laboratoire de génétique végétale du Moulon

**DISCIPLINE : génétique**

**THÈSE DE DOCTORAT**

Soutenue le 15 janvier 2014

par

**Pierre RIVIÈRE**

---

**Méthodologie de la sélection décentralisée et  
participative : un exemple sur le blé tendre.**

Création de variétés-populations adaptées à la diversité des systèmes agroécologiques et gestion dynamique de la biodiversité cultivée dans des fermes en réseau.

---

**Composition du jury :**

Directrice de thèse	Isabelle GOLDRINGER	Directrice de recherche (INRA)
Rapporteurs	Pierre M. STASSART	Professeur (Université de Liège)
	Magali SAN CRISTOBAL	Directrice de Recherche (INRA)
Examineurs	Jane LECOMTE	Professeur (Université Paris-Sud)
	Jacques DAVID	Professeur (SupAgro Montpellier)
	Pierre-Henri GOUYON	Professeur (MNHN, Science Po, AgroParisTech, ENS)
Membre invité	Jean-François BERTHELLOT	Paysan-boulangier (Réseau Semences Paysannes)



« *Une pratique scientifique qui omet de se mettre elle-même en question  
ne sait pas à proprement parler ce qu'elle fait.* »

Bruno Latour, 1992.



Ferme du Moulon, Gif sur Yvette, le 14 février 2014



## **Avant-propos**

Rédiger en anglais un travail qui vise à briser les frontières entre les laboratoires de recherche et les fermes serait un comble! Dans cette thèse, j'ai voulu que tout le corps du texte soit en français. Dans les annexes, des articles en anglais vont plus dans les détails mais ne pas les lire ne doit pas perturber la compréhension du texte. J'espère ainsi que toutes les personnes avec lesquelles j'ai collaboré de près ou de loin dans ce magnifique projet : paysannes, paysans, animatrices, animateurs, chercheuses, chercheurs, boulangères, boulangers, meunières, meuniers, artistes, poètes, photographes, journalistes, cuisinières, cuisiniers, médecins, citoyennes, citoyens ... pourront lire ce travail sans avoir à surmonter l'obstacle de la langue.





# Remerciements

**Avertissement** : il y a 46 fois le mot « merci » (dont celui là !) dans ces remerciements. Je n'ai pas trouvé de synonyme adéquate.

J'ai découvert la sélection participative lors de mon master à Wageningen. J'avais enfin trouvé un moyen d'utiliser toutes les connaissances que j'avais acquises lors de mes études pour les mettre au service d'une cause qui me parlait.

Rentré en France, à Montpellier, j'ai eu la chance de tomber sur une annonce de stage proposée par l'équipe d'Isabelle Goldringer. Ni une ni deux, je saute sur l'occasion ! Mail puis entretien téléphonique : je suis pris ! J'ai ensuite rencontré Isabelle au Moulon où nous nous sommes mis d'accord pour définir un sujet de thèse. Sujet qui a été accepté par l'école doctorale Science du Végétal puis par la région Ile de France qui apporte le financement de la thèse à travers le DIM ASTREA : merci à eux. Tant qu'on est dans le financement, j'en profite pour remercier l'Europe, qui, avec SOLIBAM, a financé le fonctionnement de cette thèse.

Un grand merci, Isabelle, de m'avoir pris en stage puis de m'avoir fait confiance pour cette thèse. C'est toi qui m'a ouvert la porte de la sélection participative, c'est toi qui m'a permis de rentrer dans ce monde fascinant. J'ai pu me reposer sur les solides bases que tu avais créées avec les paysans dès 2005. Merci ensuite de m'avoir laissé une grande liberté. J'ai pu faire ce que je souhaitais, j'ai pu tester mes idées, aller creuser certains points qui me tenaient plus à cœur.

Tu m'as montré que l'on pouvait faire de la science tout en étant citoyen, que l'on se devait de réfléchir à notre responsabilité sociale dans le cadre de nos recherches, qu'une

réflexion politique et éthique était primordiale dans le métier de chercheur.

Ces années de thèse ont été extraordinaires pour moi autant d'un point de vue scientifique qu'humain.

Merci aux membres de mon comité de thèse : Olivier David, Patrick de Kochko, François Papy, Salvatore Ceccarelli, Elise Demeulenaere, Laurent Hazard et Yannick de Oliveira.

Merci Olivier de m'avoir enseigné les statistiques bayésiennes. Merci pour ta grande disponibilité et ta gentillesse.

Merci Laurent de m'avoir accueilli aux journées jeunes chercheurs du SAD. Merci pour les discussions que l'on a eu ensemble qui m'ont aidé à prendre du recul sur mon action dans les différentes « arènes » ! Tu as éveillé ma curiosité sur les concepts de sciences sociales et je t'en suis très reconnaissant ! C'est grâce à cela que j'ai pu écrire le chapitre 4 de cette thèse.

Merci Elise d'avoir relu mon chapitre 4 (un week-end !) et d'avoir partagé tes réflexions sur les questions de bases de données.

Merci Patrick pour les discussions qui m'ont permis de prendre un peu plus de recul sur le projet.

Merci Yannick de m'avoir enseigné l'art de la gestion des données.

Merci à Guy Kastler pour ses conférences : tu m'as permis de mieux comprendre le monde si complexe de la législation ! Merci également pour tes analyses des enjeux plus globaux autour des semences : cela m'a énormément aidé à y voir plus clair !

Merci à Anne-Charlotte Möy d'avoir passé beaucoup de temps à m'expliquer les aspects liés à la juridiction des semences.

Merci à mon équipe d'accueil : DEAP. Merci Jérôme pour les discussions que l'on a eu sur le projet et sur l'impact de la bière sur le potentiel d'innovation ! Merci Nathalie et Sophie pour les road-trips en France



L'équipe (de choc) DEAP en 2011 : Nathalie Galic, Jérôme Enjalbert, Isabelle Goldringer, Mathieu Thomas, Abdul Rheman Khan, Pierre Rivière, Sophie Pin, Stéphanie Thépot.

chez les paysans à faire les mesures. Il restera de nombreux souvenirs : le routier en Bourgogne, le camping, les piques-niques, les soirées pizzas à la Tour du Pin, ... Merci Sophie

d'avoir assuré l'animation avec le réseau lors de ma dernière année de thèse. Merci pour ton sens de l'organisation, devenu maintenant légendaire, et ta bonne humeur ! Merci également Mathieu pour les échanges que l'on a eu sur ma thèse et la tienne ! Merci Stéphanie et Abdul pour votre aide inestimable sur la biologie moléculaire ! Vous m'avez donné des marqueurs qui fonctionnaient ainsi que de nombreux conseils pour les analyses : merci ! Merci Julie Dawson de m'avoir encadré lors de mon master 2 dans l'équipe et pour les nombreuses discussions que l'on a eu sur le projet ! Et bien sûr un grand merci à Carine Remoué qui, avec Jérémy Pansier, ont extrait les ADNs de nos blés paysans : c'était un travail monstrueux.

Merci Bubar et Olivier Langella pour vos tuyaux sur `linux`, `bash` et `condor`. J'en profite pour remercier les communautés  $\text{\LaTeX}$ , R et `linux` 🐧 !

Merci à Charlotte Urien, ma coloc de jour du confortable<sup>i</sup> et silencieux<sup>ii</sup> algéco n°2 ! Merci, bien sûr, aux M&Ms qui ont rythmé nos dures journées de travail !

Merci aux stagiaires qui m'ont aidé et accompagné dans cette thèse : Guillemette Fruchaud, Cathy Bouffartigue, Florian Gery, Darkawi Madi. Votre aide a été précieuse notamment pour les mesures dans les champs ! Merci aussi à toutes les personnes qui nous ont aidés dans les fermes : Lucie et Élise<sup>iii</sup>, Magali, les potes de Cathy, les stagiaires des assos, les gens de passages dans les fermes.

Merci aux animatrices et animateurs des associations pour votre aide sur le terrain et les discussions toujours très enrichissantes sur le projet (et pas seulement !).

---

i. Sauf l'hiver et l'été

ii. Au nettoyage des feuilles près

iii. La sélection participative, c'est aussi une histoire de famille !



Nath et Sophie dans les champs qui posent les étiquettes sur les blés.



Sophie qui mesure la hauteur des blés.



Un immense merci aux paysannes et aux paysans ! Merci pour vos accueils toujours très chaleureux, les discussions, les repas, les apéros, les échanges nombreux que l'on a eu. Je ne vais pas décrire ici tous les très bons souvenirs que j'ai eu en votre compagnie : ils ont été si nombreux !

Merci à Jacques et Catherine Carpentier pour les relectures de ma thèse. Vous n'avez pas fui en voyant ces covariances hostiles : merci !



Lucie et Elise au boulot pour aider leur frère !



Mag qui note avec précaution les mesures !



Sarah et Cathy qui préparent les étiquettes qui seront collées sur les blés.



Darkawi qui laisse son ordi pour noter les seed lot !

Merci à mon jury (et oui, je rédige cette partie une fois la soutenance passée !). Merci d'avoir pris le temps de lire ma thèse et de vos retours. J'étais très heureux que Jean-François Berthelot, paysan-boulangier, ait pu siéger à côté des professeurs d'université et autres directeurs de recherches. C'était un beau symbole que de faire asseoir à la même table deux mondes qui ne se croisent que trop rarement : paysans et chercheurs.

La soutenance restera pour moi un moment inoubliable. Merci à ma famille et à celle de Mag d'avoir géré le pot qui a suivi !

Merci à la jeunesse moulonienne pour les soirées et autres barbucs !

Merci à toutes les personnes que j'ai croisées pendant ma thèse : paysannes, paysans, animatrices, animateurs, chercheuses, chercheurs, boulangères, boulangers, meunières, meuniers, artistes, poètes, photographes, journalistes, cuisinières, cuisiniers, médecins, citoyennes, citoyens ... Les discussions toujours riches m'ont énormément aidé à mieux appréhender mon sujet et à structurer ma pensée.

Merci aux amis qui m'ont aidé à décompresser et à faire évoluer mes réflexions.

Un grand merci à Magali !!! 😊. Merci d'avoir supporté mes sauts d'humeur parce-que mes scripts R ne fonctionnaient pas ou parce-que j'en avais marre de la thèse !

Enfin, un merci tout particulier à mes parents. Vous avez toujours assuré pour me permettre de faire ces études : je vous ai baladé de Caen à Bordeaux, de Wageningen à Montpellier. Vous m'avez toujours soutenu dans mes choix. MERCI !



## Résumé

L'agroécologie a des attentes fortes en termes de recherche sur les concepts et sur les méthodes de sélection et de gestion de la diversité. En effet, la prise en compte de la diversité des environnements et des pratiques spécifiques à chaque système agroécologique appelle des approches plus décentralisées, c'est à dire une meilleure prise en compte des interactions génotype  $\times$  environnement, et associant les savoirs empiriques des praticiens aux connaissances scientifiques.

L'objectif de cette thèse est de développer une méthodologie de la sélection participative pour le blé tendre, basée sur la décentralisation et la co-construction entre paysans, associations du Réseau Semences Paysannes et chercheurs de l'équipe DEAP de l'INRA du Moulon. Cette approche vise à créer des variétés-populations adaptées aux environnements et aux pratiques des paysans, à développer des méthodes et des outils opérationnels pour la gestion et la sélection de la biodiversité cultivée à la ferme et à renforcer l'apprentissage et l'autonomie des paysans en matière de sélection.

Ce travail a contribué à atteindre ces objectifs, notamment les deux derniers, en se basant sur une approche interdisciplinaire dans les champs de la statistique, génétique quantitative, génétique des populations, bio-informatique et sociologie.

Les objectifs ont été en partie atteints avec la création de méthodes et d'outils : fiches de suivi de la culture, base de données permettant de gérer les relations entre lots de semences dans un large réseau d'acteurs, dispositifs expérimentaux à la ferme et méthodes statistiques permettant de prendre en compte le déséquilibre des essais à la ferme, en profitant du large réseau d'expérimentation, afin de réaliser des comparaisons de moyennes dans les fermes et d'analyser les interactions génotype  $\times$  environnement dans

le réseau de fermes, programme informatique qui permet de créer un dossier avec des résultats personnalisés pour chaque paysan, livret technique sur la sélection participative.

Une analyse moléculaire et phénotypique montre que le programme de sélection participative s'insère dans la gestion *in-situ* des ressources génétiques par la création de nouvelles variétés-populations et par leur gestion dans un large réseau de fermes regroupant une grande diversité d'environnements et de pratiques.

Les paysans ont renforcé leur apprentissage et leur autonomie. Ils étaient acteurs du programme et leurs pratiques ont évolué avec, par exemple, un nombre croissant de populations évaluées sur le réseau, de sélections et d'échanges de semences entre paysans au fur et à mesure que le projet avançait.

Deux innovations émergent de ce projet : organisationnelle et génétique. L'innovation organisationnelle correspond au mode de fonctionnement du projet, basé sur la décentralisation de la sélection directement dans les fermes et la co-construction entre les acteurs et leur mise en réseau. Les innovations génétiques découlent de ce mode d'organisation : une large diversité génétique et phénotypique a été créée, est maintenue et évaluée dans les fermes et de nouvelles variétés populations, issues du programme, sont utilisées par des paysans dans leurs fermes.

La méthodologie que nous avons développée est évolutive, flexible et adaptable. Elle place les paysans au coeur de la sélection et de la gestion des ressources génétiques. Ce projet participe à la transition vers un nouveau système semencier adapté à l'agroécologie.



## Liste des abréviations

<b>ABI</b>	Atelier de Bio-informatique et d'Informatique
<b>ABS</b>	Access and Benefits Sharing
<b>ACP</b>	Analyse en Composantes Principales
<b>ADPIC</b>	Accord sur les Aspects des Droits de Propriété Intellectuelle qui touchent au Commerce. TRIPS en anglais : Agreement on Trade-Related Aspects of Intellectual Property Rights.
<b>ANOVA</b>	Analyse de variance
<b>AMOVA</b>	Analyse de variance sur des données moléculaires
<b>AMMI</b>	Additive Main Effect and Multiplicative Interaction
<b>ATM</b>	Accord de Transfert de Matériel
<b>ARDEAR</b>	Association pour le Développement de l'Emploi Agricole et Rural
<b>BIC</b>	Bayesian Information Criterion
<b>CBD</b>	Convention sur la Diversité Biologique
<b>CCP</b>	Composite Cross Population
<b>CETAB</b>	Centre d'Etude et Terre d'Accueil des Blés
<b>CGIAR</b>	Consultative Group for International Agricultural Research
<b>CNRS</b>	Centre National de la Recherche Scientifique
<b>COV</b>	Certificat d'Obtention Végétale
<b>CRB</b>	Centre de Ressources Biologiques

<b>CTPS</b>	Comité Technique Permanent de la Sélection
<b>CVO</b>	Cotisation Volontaire Obligatoire
<b>DAPC</b>	Analyse Discriminante sur les Composantes Principales
<b>DEAP</b>	Diversité Évolution et Adaptation des Populations
<b>DHS</b>	Distinction, Homogénéité et Stabilité
<b>DIM ASTREA</b>	Domaines d'Intérêt Majeur, Agrosociétés, Territoires, Ecologie, Alimentation
<b>FAO</b>	Food and Agricultural Organisation : l'Organisation des Nations Unies pour l'alimentation et l'agriculture.
<b>FMI</b>	Fond Monétaire International
<b>FNSEA</b>	Fédération Nationale des Syndicats d'Exploitants Agricoles
<b>FSOV</b>	Fonds de Soutien à l'Obtention Végétale
<b>GIEC</b>	Groupement d'Experts Intergouvernemental sur l'Évolution du Climat
<b>GNIS</b>	Groupement National Interprofessionnel des Semences et plants
<b>GPL</b>	General Public Licence
<b>IAASTD</b>	évaluation internationale des sciences et des technologies agricoles pour le développement : International Assessment of Agricultural Science and Technology for Development en anglais
<b>IFOAM</b>	organisation internationale des mouvements de l'agriculture biologique. International Federation of Organic Agriculture Movements en anglais.
<b>INRA</b>	Institut National de la Recherche Agronomique
<b>IUPGR</b>	International Undertaking on Plant and Genetic Resource
<b>ITAB</b>	Institut Technique de l'Agriculture Biologique
<b>LLSD</b>	distance dernière feuille base de l'épi : last leaf spike distance en anglais.
<b>MCMC</b>	Markov Chain Monte Carlo
<b>OCDE</b>	Organisation de Coopération et de Développement
<b>OCVV</b>	Office Communautaire des Variétés Végétales
<b>OGM</b>	Organisme Génétiquement Modifié
<b>OMC</b>	Organisation Mondiale du Commerce
<b>OMPI</b>	Organisation Mondiale de la Propriété Intellectuelle

<b>ONG</b>	Organisation Non Gouvernementale
<b>PASMA</b>	Plan d'Action du Sommet Mondial de l'Alimentation
<b>PICRI</b>	Partenariats Institutions-Citoyens pour la Recherche et l'Innovation
<b>PVPA</b>	Plant Variety Protection Act : Acte sur la protection des variétés de plantes.
<b>RSP</b>	Réseau Semences Paysannes
<b>SNP</b>	Single Nucleotide Polymorphism
<b>SOLIBAM</b>	Strategies for Organic and low-input Integrated Breeding Management
<b>SPEAL</b>	Sélection Participative d'Espèces Annuelles et Ligneuses adaptées à des pratiques agroécologiques
<b>TIRPAA</b>	Traité International sur les Ressources Phytogénétiques pour l'Alimentation et l'Agriculture
<b>UE</b>	Union Européenne
<b>UPOV</b>	Union Internationales pour la Protection des Obtentions Végétales
<b>VAT</b>	Valeur Agronomique et Technologique
<b>VATE</b>	Valeur Agronomique, Technologique et Environnementale





# Sommaire

Avant-propos	9
Remerciements	11
Résumé	17
Liste des abréviations	19
Sommaire	24
1 Système semencier et souveraineté alimentaire	25
1.1 Souveraineté alimentaire, agroécologie et semences	26
1.2 Stratégies de gestion des ressources génétiques et de sélection variétale pour l'agroécologie	39
1.3 Participation des acteurs pour la sélection et la gestion des ressources génétiques	55
1.4 Gouvernance des semences et des ressources génétiques	68
1.5 Contexte et objectifs du projet de sélection décentralisée et participative sur le blé tendre en France	80
2 Analyse phénotypique du programme de sélection participative	87
2.1 Dispositifs expérimentaux dans les fermes	88
2.2 Analyse phénotypique dans les fermes	100
2.3 Analyse des interactions $G \times E$	134
3 Evaluation phénotypique et moléculaire du programme de sélection participative	145
3.1 Matériels et méthodes	146
3.2 Résultats	157

3.3	Discussion	185
4	Description et analyse de la méthodologie mise en place	197
4.1	Organisations sociales dans lesquelles s'inscrivent les acteurs	198
4.2	Co-construction de la méthodologie	200
4.3	Innovations organisationnelles et génétiques	230
4.4	Evolution des pratiques des acteurs	236
5	Transition vers la création d'un système semencier adapté à l'agroécologie	249
5.1	Résultats principaux de la thèse et propositions au collectif de travail	250
5.2	Propositions pour une transition au niveau local et régional	258
5.3	Propositions au niveau national pour accompagner la transition	272
	Conclusion	279
	Bibliographie	281
	Table des matières	301



# 1 . Système semencier et souveraineté alimentaire

Souveraineté alimentaire, agroécologie et semences

Stratégies de gestion des ressources génétiques et de sélection variétale pour l'agroécologie

Participation des acteurs pour la sélection et la gestion des ressources génétiques

Gouvernance des semences et des ressources génétiques

Contexte et objectifs du projet de sélection décentralisée et participative sur le blé tendre en France

La souveraineté alimentaire est la condition incontournable de l'accès pour chaque être humain à une nourriture saine et en quantité suffisante. Les semences sont le premier maillon d'une chaîne qui aboutira à la nourriture. Produire, gérer, utiliser, gouverner les semences de façon efficace et durable est donc un enjeu majeur. Ce premier chapitre explicite en cinq parties, différents éléments nécessaires à la contextualisation du projet de thèse.

Tout d'abord, la partie 1.1 fait le point sur quelques textes internationaux qui régissent l'agriculture et présente les alternatives possibles pour produire la nourriture tout en préservant les ressources naturelles. L'agroécologie est proposée comme une solution.

Puis, la partie 1.2 présente les stratégies de gestion des ressources génétiques et de sélection variétale dans des systèmes agroécologiques.

Ensuite, la partie 1.3 présente des approches participatives permettant d'associer les connaissances et les savoirs locaux avec les savoirs scientifiques formels dans la gestion des ressources génétiques et l'amélioration des plantes.

La partie 1.4 analyse les modes de gouvernance actuellement en place dans la gestion des ressources génétiques et dans la création des semences.

Enfin, la partie 1.5 pose le contexte et les objectifs d'un programme de sélection décentralisée et participative sur le blé tendre en France et la manière dont ce travail de thèse s'y insère.

## 1.1 Souveraineté alimentaire, agroécologie et semences

### 1.1.1 Du concept de sécurité alimentaire à celui de souveraineté alimentaire

Lors du sommet mondial de l'alimentation à la Food and Agricultural Organisation (FAO) en novembre 1996, la Déclaration de Rome pose les bases d'une concertation mondiale sur les problèmes de l'alimentation dans le monde<sup>87</sup>. Les parties contractantes s'engagent à mettre en oeuvre et soutenir le Plan d'Action du Sommet Mondial de l'Alimentation (PASMA). Ce plan définit la sécurité alimentaire :

*« La sécurité alimentaire existe lorsque tous les êtres humains ont, à tout moment, un accès physique et économique à une nourriture suffisante, saine et nutritive leur permettant de satisfaire leurs besoins énergétiques et leurs préférences alimentaires pour mener une vie saine et active. »*

et les délégations s'engagent à éradiquer « *la faim dans tous les pays* ». Elles reconnaissent que la paix, la stabilité politique, la démocratie, les droits de l'homme, les libertés fondamentales, la réduction de la pauvreté, la participation égalitaire des hommes et des femmes, la préservation rationnelle et durable de l'environnement et des ressources sont essentiels pour atteindre cet objectif. Pour parvenir à la sécurité alimentaire, la Déclaration de Rome préconise d'augmenter la production alimentaire, d'aménager les ressources naturelles, d'éliminer les modes non viables de consommation et de stabiliser la population mondiale.

La Déclaration met en lumière également que : « *La nourriture ne doit pas être utilisée comme instrument de pression politique et économique.* ». Elle insiste sur « *l'importance de la coopération et de la solidarité internationales* ». Elle souligne que le commerce et les ressources financières doivent être adaptés pour mettre en oeuvre des politiques de sécurité alimentaire durable.

La Déclaration du forum des Organisations Non Gouvernementales (ONG) à Rome en 1996 rejoint le PASMA et pointe « *certaines causes fondamentales de l'insécurité alimentaire* » :

1. La globalisation de l'économie mondiale livrée aux multinationales et au modèle de surconsommation qui amène à la « *destruction des économies rurales et à la faillite de l'agriculture familiale* ».



2. L'agriculture industrialisée qui détruit « *l'exploitation traditionnelle et empoisonne la planète et tous les êtres vivants* »
3. Les exportations subventionnées qui accroissent « *l'instabilité du marché au détriment des agriculteurs familiaux* ».
4. Le rôle du Fonds Monétaire International (FMI) et de la Banque Mondiale qui obligent « *les agriculteurs familiaux et les personnes vulnérables* » à payer « *le prix de l'ajustement structurel et du remboursement de la dette* ».

Pour répondre à ces difficultés, les ONG proposent un modèle alternatif « *basé sur la décentralisation* » qui redistribuerait la richesse et le pouvoir et permettrait de parvenir à la sécurité alimentaire et à la protection des écosystèmes qui rendent la vie possible sur terre.

La Déclaration souligne « *six éléments clés de ce modèle alternatif* » : aider la capacité à produire des petits producteurs, inverser la concentration des richesses et du pouvoir, développer l'agroécologie, améliorer la capacité des états et des gouvernements à garantir la sécurité alimentaire, renforcer la participation des organisations populaires et des ONG à tous les niveaux et faire garantir le droit à l'alimentation par le droit international.

C'est de cette Déclaration que le concept de souveraineté alimentaire émerge. Les questions liées à la gouvernance et à l'autonomie s'ajoutent aux questions liées à la sécurité alimentaire. La dimension politique du contrôle des richesses et du pouvoir par les populations est au centre du concept de souveraineté. Comment une population maîtrise-t-elle son alimentation ? Son accès au marché, à la ressource, à la nourriture ? En d'autres termes, comme le souligne la coordination européenne Via Campesina<sup>i</sup> dans un communiqué de 2009 lors du Sommet Mondial sur la Sécurité Alimentaire à la FAO : « *La construction de la souveraineté alimentaire repose sur la démocratisation de la prise de décision.* ».

### 1.1.2 Etat des lieux et menaces pour la sécurité alimentaire dans le monde

Selon la FAO<sup>96</sup>, en 2012, 870 millions de personnes souffrent de la faim dans le monde dont 850 millions dans les pays en voie de développement. La principale piste pour enrayer cette situation est que « *la croissance économique soit doublée d'une action volontariste*

---

i. La Via Campesina est, d'après leur site internet, « *un mouvement international qui rassemble des millions de paysannes et de paysans, de petits et de moyens producteurs, de sans terre, de femmes et de jeunes du monde rural, d'indigènes, de migrants et de travailleurs agricoles ... Elle défend l'agriculture durable de petite échelle comme moyen de promouvoir la justice sociale et la dignité.* ».

*et déterminante des pouvoirs publics qui, pour être efficaces [...] doivent s'appuyer sur un système amélioré de gouvernance, reposant sur la transparence, la participation, l'obligation de rendre compte, l'état de droit et le respect des droits de l'homme ».*

La Déclaration de Rome cite comme levier l'augmentation de la production alimentaire. D'ici 2050, une augmentation de 70% de la production de nourriture sera nécessaire pour nourrir une population mondiale qui va dépasser les 9 milliards d'individus<sup>92</sup>.

Dans les pays du nord, le système agricole intensif mis en place n'est pas durable. L'impact sur la sécurité alimentaire est direct et négatif<sup>92;95;133;162;193</sup>. Sont dangereusement réduits, les bienfaits rendus par les agrosystèmes à l'humanité :

1. équilibre des cycles des nutriments, de la formation des sols, de la pollinisation, du contrôle biologique,
2. régulation du climat et des maladies,
3. approvisionnement en nourriture, eau, bois,
4. et par là même, accès à la culture, à l'esthétisme, à l'éducation.

Ces services éco-systémiques sont malmenés et la biodiversité recule<sup>162</sup>. La production de nourriture est basée sur une dépendance accrue aux intrants tels que les fertilisants chimiques, les produits phyto-sanitaires (herbicides, pesticides, fongicides), à l'irrigation qui provoque salinisation, érosion et abaissement des nappes phréatiques ; et également aux antibiotiques utilisés pour les élevages<sup>92</sup>.

Par ailleurs, pour la société, ce modèle agricole a un coût élevé pour sauvegarder la santé des hommes, des végétaux, des animaux et plus généralement de tout l'écosystème<sup>21;193</sup>.

A ces problèmes liés à l'agriculture non durable s'ajoute le changement climatique qui accentue l'insécurité alimentaire. D'après le Groupement d'Experts Intergouvernemental sur l'Évolution du Climat (GIEC)<sup>217</sup>, l'agriculture est responsable de 10 à 12 % des émissions de gaz à effet de serre dans le monde. Les pays de l'Organisation de Coopération et de Développement (OCDE) sont responsables d'environ 75% des émissions à cause d'une agriculture intensive basée sur le pétrole (mécanisation, transport, production de nitrate et de produits phytosanitaires) et sur l'élevage intensif (méthane dégagé par le bétail notamment). Le changement climatique influe sur les périodes de reproduction animale et végétale, la durée des saisons de croissance, la répartition des espèces, la densité des populations, la pression des maladies. La biodiversité diminue et les systèmes biologiques sont déstabilisés<sup>226</sup>. Le changement climatique engendre une irrégularité et généralement une

baisse des rendements agricoles<sup>158;217</sup>. Par exemple, en 2012 aux États-Unis, une vague de sécheresse a provoqué une perte de 20 milliards de dollars et une grave tension sur le marché mondial liée à l'augmentation du prix des céréales<sup>184;241</sup>. Cette agriculture mondialisée et sa financiarisation ont été en partie responsables des émeutes de la faim de 2007-2008 notamment en Afrique<sup>96</sup>.

La non-durabilité des systèmes agricoles, la perte des services éco-systémiques, le changement climatique, l'évolution des modes de consommation, sont autant de points clés à étudier pour résoudre le problème de la souveraineté alimentaire. La géopolitique joue évidemment un rôle fondamental dans ce contexte<sup>30</sup>. Il n'en est pas question dans la suite de ce travail qui aborde les problèmes de la souveraineté alimentaire principalement sous l'angle de l'agronomie.

### 1.1.3 Les alternatives : agroécologie et participation des acteurs

#### 1.1.3.1 L'agriculture doit changer de paradigme

Il y a un consensus au niveau des rapports internationaux : pour atteindre un système agricole qui permette, à la fois, de produire assez de nourriture saine et nutritive, qui s'adapte au changement climatique, qui respecte l'environnement, qui soit durable et fasse participer les différents acteurs aux processus de décisions : l'agriculture doit changer de paradigme<sup>91;92;133;162;240</sup>.

Pour réaliser un tel objectif, l'évaluation internationale des sciences et des technologies agricoles pour le développement (IAASTD) - un consortium de 400 experts d'origines et de disciplines très différentes qui ont travaillé pendant quatre ans - appelle à une réorientation des sciences agronomiques vers une approche plus holistique des systèmes<sup>83;133</sup>. Il s'agit de prendre en compte la complexité, la diversité, l'aspect pluri-factoriel et multi-fonctionnel des systèmes agricoles. Selon le rapport de l'IAASTD, les dimensions sociales, culturelles, économiques et écologiques liées aux services éco-systémiques et aussi la diversité des processus d'innovation technologique doivent associer connaissances et savoirs locaux avec les savoirs scientifiques formels.<sup>93;133;162</sup>

L'IAASTD souligne l'importance de l'accès aux ressources, notamment l'eau et la terre, du rapprochement entre science et société et souligne la nécessité de faire évoluer les modes de gouvernance, afin de renforcer la participation des différents acteurs du monde agricole aux processus de décision et d'évaluation.

Le rapport promeut le développement de l'agroécologie et des recherches interdiscipli-

naires<sup>83</sup>. Ce point est corroboré par d'autres rapports de la FAO<sup>90;92</sup>.

### 1.1.3.2 Les principes de l'agroécologie

Gliessman, cité par Tomich et al.<sup>236</sup>, définit l'agroécologie comme « *la science d'appliquer des concepts et des principes de l'écologie à la construction et la gestion de systèmes alimentaires durables* ». Elle développe « *des agrosystèmes biodiversifiés capables de subvenir à leur propre fonctionnement* » pour parvenir à une agriculture durable qui ne dégrade pas les ressources<sup>7;8</sup>.

L'agroécologie repose sur l'optimisation de plusieurs processus<sup>7;221;236;246</sup> :

1. Le maintien de l'activité biologique du sol grâce à la couverture des sols, la gestion de la matière organique, l'augmentation du recyclage de la biomasse. La pratique de l'agroécologie entraîne une meilleure gestion des ressources en eau et une meilleure utilisation de l'énergie (radiation solaire notamment).
2. L'optimisation du cycle des nutriments (azote, potasse, phosphore) afin d'en augmenter la disponibilité et d'en équilibrer la balance.
3. Les interactions biologiques bénéfiques entre les composantes de l'agrobiodiversité<sup>ii</sup> doivent être augmentées afin de promouvoir les processus écologiques et les services éco-systémiques : pollinisation, contrôle biologique, cycle des nutriments, ...
4. La diversité des espèces et la diversité génétique au sein de l'agrosystème de la parcelle au paysage doivent être importantes dans le temps et dans l'espace.

Des pratiques telles que les mélanges variétaux, l'agroforesterie, les rotations de cultures, les couverts végétaux, la permaculture, permettent d'optimiser ces procédés. Ils permettent de rendre l'agrosystème plus stable<sup>iii</sup> dans le temps et plus résilient<sup>iv</sup>.

Au delà de l'agronomie et de l'écologie, l'agroécologie promeut une compréhension du fonctionnement des composants de l'agrosystème dans son environnement. Un travail interdisciplinaire liant agronomie, écologie, sociologie, anthropologie, droit de l'environnement et économie est alors nécessaire<sup>7;54;106</sup>. Travail où s'expriment diversité des points de vues et diversité des perspectives et des méthodes<sup>32</sup>. Alors que l'agronomie travaille traditionnellement à l'échelle de la parcelle ou du paysage, les sciences sociales et envi-

---

ii. L'agrobiodiversité correspond à une fraction de la biodiversité qui est utilisée directement ou indirectement pour la nourriture et l'agriculture.

iii. La stabilité est la faculté de l'agrosystème de se comporter en moyenne de la même manière au cours du temps.

iv. La résilience est la capacité de l'agrosystème à retrouver son point d'équilibre après une perturbation<sup>103</sup>.

ronnementales réfléchissent à l'échelle du paysage, du département, de la région, de la nation, du continent ou du monde<sup>236</sup>. Deux approches différentes, deux paradigmes, s'opposent : réductionnisme et holisme<sup>246</sup>. Le réductionnisme, héritage des sciences du XIX<sup>ème</sup> siècle qui voulait discipliner le vivant en l'enfermant dans des « boîtes séparées »<sup>22</sup>, vise à comprendre le monde en l'étudiant à l'échelle des parcelles. A l'inverse, l'holisme vise à comprendre le monde comme un tout, c'est à dire de façon systémique. Dans ce cadre holistique, l'agrosystème est étudié dans son ensemble. L'agroécologie propose de faire travailler ensemble des champs scientifiques très différents. Réductionnisme et holisme, deux visions complémentaires qui permettent d'appréhender l'agrosystème dans sa plus grande dimension.

Le terme d'agroécologie désigne à la fois une pratique agronomique, une discipline scientifique et un mouvement politique ou social. Ce mouvement de l'agroécologie est de fait associé, même implicitement, à une réflexion critique du système conventionnel de recherche, du productivisme lié au monde de l'agro-industrie et des biotechnologies<sup>32;246;253</sup>.

L'agroécologie est un concept d'action scientifique, une force qui mobilise le changement social par une réflexion sur des propositions concrètes<sup>221</sup>. Elle repense le lien entre l'action et la science, c'est le cas de la recherche participative par exemple. Elle souligne l'importance des savoirs locaux et s'emploie à sauvegarder l'autonomie des producteurs<sup>221</sup>. Enfin, l'agroécologie est une force de propositions à destination des décideurs. Elle permet d'impulser la mise en œuvre de politiques adaptées<sup>7;60;221</sup>.

L'agriculture biologique peut être considérée comme une forme de l'agroécologie. Elle utilise certaines pratiques de gestion des agrosystèmes basées sur l'agroécologie. D'après l'organisation internationale des mouvements de l'agriculture biologique (IFOAM), l'agriculture biologique repose sur quatre principes qui sont<sup>134</sup> :

1. la santé des plantes, des animaux, des sols, des humains, de la planète comme un tout,
2. l'écologie, c'est à dire le respect des cycles des écosystèmes dont il faut assurer la durabilité,
3. la justice, l'agriculture biologique doit construire des relations justes dans un environnement commun et partagé par tous,
4. et la protection de l'environnement pour assurer la santé et le bien être des générations présentes et futures,

chacun, élément d'un tout cohérent.

Cependant, le cahier des charges de l'agriculture biologique a évolué avec une réglementation européenne moins stricte sur certains points, notamment sur l'élevage<sup>102</sup>. De plus, dans certains cas, l'agriculture biologique reproduit le système productiviste conventionnel et de fait, s'éloigne des principes d'ordre socio-économique défendus par l'agroécologie<sup>223;246</sup>.

#### 1.1.3.2.1 Les spécificités agronomiques d'un mode de culture en agroécologie

Afin d'étudier au mieux les agrosystèmes, il est important de comprendre leurs spécificités, notamment agronomiques. Une meilleure compréhension de l'environnement, des itinéraires techniques et des systèmes de culture<sup>v</sup> dans lesquelles les plantes évoluent peut permettre de mettre en place des stratégies de gestion adaptées aux spécificités des agrosystèmes.

L'environnement, pris en compte dans sa complexité, comprend l'environnement biophysique, les techniques culturales, les acteurs, les besoins du marché, la réglementation et enfin les attentes de la société ; ces paramètres sont à la fois socio-économiques et agroécologiques<sup>70</sup>. L'agriculture biologique partage avec l'agroécologie de nombreux principes dans sa conception agronomique et écologique de l'agrosystème. Il ne sera ici question que de cette double dimension agronomique et écologique. Un état des lieux exhaustif des composants agronomiques d'un agrosystème n'a pas pu être envisagé ici. L'aspect socio-économique est traité dans la partie 1.3 avec la question de la participation des acteurs.

#### 1.1.3.2.2 Le champ cultivé dans son environnement

L'agrosystème a trois composants : le sol, le peuplement végétal cultivé et l'atmosphère<sup>205</sup> (Figure 1.1)<sup>vi</sup>. Les services éco-systémiques se retrouvent dans ces trois composants et participent à la régulation de l'agrosystème<sup>162;195</sup>. Il y a en permanence de nombreuses interactions entre les composantes physiques, chimiques et biologiques d'un agrosystème. Cela engendre une grande complexité qu'il est difficile de maîtriser<sup>73</sup>.

Contrairement à l'agriculture biologique, l'agriculture conventionnelle est basée sur

---

v. D'après Doré et Meynard<sup>73</sup>, qui citent Sebillotte, l'itinéraire technique est une « combinaisons logique et ordonnée des techniques mises en oeuvre sur une parcelle en vue d'en obtenir une production », le système de culture est « l'ensemble des modalités techniques mises en oeuvre sur des parcelles [...] ». Chaque système de culture se définit par la nature des cultures [...] et les itinéraires techniques [...] »

vi. Il manque évidemment le paysan au centre de ces composants ! Merci à P-H. Gouyon d'en avoir fait la remarque lors de la soutenance.

l'utilisation d'intrants venant de l'extérieur de la ferme, de la région ou du pays : eau, engrais chimiques, produits phytosanitaires, semences, . . . qui réduisent les hétérogénéités au sein des différents agrosystèmes<sup>49</sup>.

D'après Lammerts Van Bueren et Myers<sup>151</sup>, les principales différences d'itinéraires techniques entre l'agriculture conventionnelle et l'agriculture biologique concernent : la biodiversité, la fertilisation, les rotations, la protection des plantes, la gestion des adventices, le traitement des semences et le travail du sol (Figure 1.1). Pour résumer, l'agriculture biologique ne recherche pas le meilleur rendement, mais un rendement stable sans intrants chimiques.

### 1.1.3.2.3 Le sol

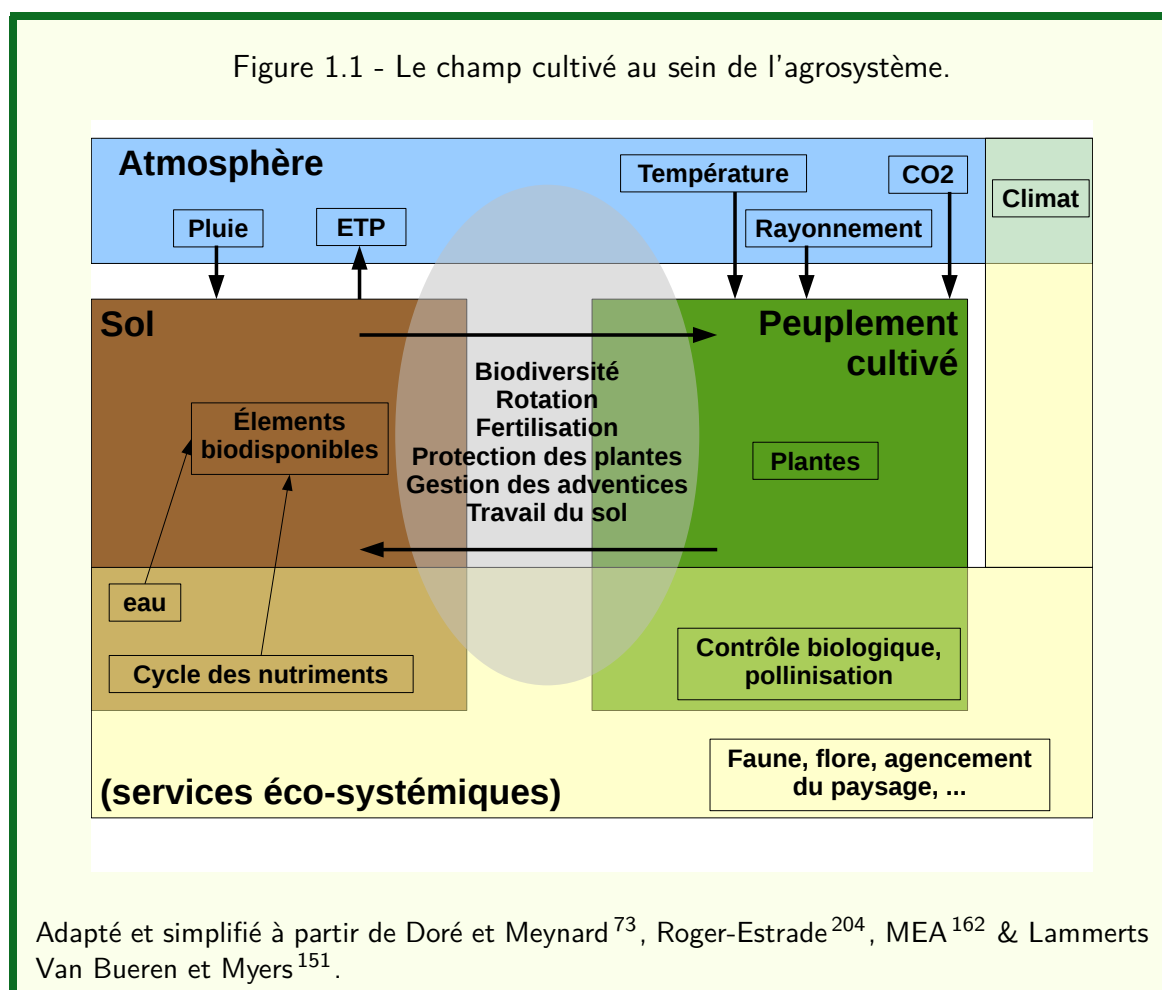
Le sol est une interface d'échange pour les plantes. En agriculture biologique la bonne santé du sol est primordiale. Le sol doit être riche en eau et minéraux (azote, phosphore et potassium) pour assurer la croissance des plantes. Alors qu'un système conventionnel pourra ajouter des minéraux chimiques qui sont solubles et facilement utilisables par la culture au moment clé de son développement, en agriculture biologique cette bio-disponibilité passe par la gestion du sol. A titre d'exemple, pour le blé tendre il est recommandé en agriculture conventionnelle de procéder à trois apports fractionnés d'unité d'azote<sup>vii</sup>. Ces apports dépendent des reliquats d'azote de la culture précédente qui seront libérés via la minéralisation. En général, les apports recommandés sont autour de 40 unités au tallage, 100 unités au stade épi 1 cm et 40 unités fin montaison. Cela permet de s'adapter au mieux aux besoins de la culture<sup>14</sup>. Si l'azote ne manque pas en agriculture biologique, il présente néanmoins le problème de ne pas être disponible au bon moment du cycle de la culture. Les motifs de cette non-disponibilité de l'azote peuvent être les températures faibles au printemps qui retardent la minéralisation, qui ne libère pas assez d'azote pour la culture. C'est souvent le cas à la montaison chez le blé, qui correspond au deuxième apport préconisé en agriculture conventionnelle. Les conditions optimales de la minéralisation arrivent souvent en fin de cycle, c'est à dire à un moment où les besoins sont moindres et la température élevée (troisième apport préconisé en agriculture conventionnelle).

L'absorption des nutriments lorsqu'ils sont difficilement disponibles, notamment l'azote en début de stade de la culture, est un point important dans les systèmes agroécologiques<sup>163</sup>. Pour améliorer cette disponibilité, il faut travailler sur le système racinaire, les

---

vii. Une unité d'azote équivaut à un kilogramme d'azote par hectare.

racines doivent être plus profondes, plus denses, et explorer un volume de sol plus grand. Les interactions plantes micro-organismes dans la rhizosphère<sup>viii</sup>, pourraient contribuer à l'absorption des nutriments, notamment du phosphore<sup>116</sup>. Il peut être intéressant de caler le cycle de la plante afin que ses plus forts besoins correspondent au moment où l'azote est le plus disponible. Les variétés tardives profitent mieux de l'azote qui sera disponible au moment de la montaison<sup>163</sup>. D'un autre côté, dans des conditions de déficit hydrique précoce, les variétés tardives sont plus susceptibles d'échauder pendant le remplissage du grain<sup>112</sup>.



La diversité des systèmes de culture et des itinéraires techniques influence la disponibilité des nutriments, la structure du sol et ses interactions avec les plantes. Les systèmes en polyculture-élevage sont les plus équilibrés en terme de gestion de la fertilisation des sols. Les produits organiques issus de l'élevage apportent du carbone et de l'azote au sol<sup>182;239</sup>.

viii. la fine couche autour des racines



Et les cultures fourragères favorisent la fertilité des sols en apportant de l'azote atmosphérique dans le système. Elles cassent le cycle des adventices, ravageurs ou maladies et structurent le sol grâce à leurs systèmes racinaires importants<sup>104</sup>. Sans élevage, il faut intégrer des légumineuses dans les rotations, des fourragères ou des engrais verts<sup>104;182</sup>. Dans ce cas, les rotations sont très variables selon les conditions pédo-climatiques, les pratiques (notamment l'irrigation ou non), la réglementation, l'environnement économique, les surfaces, la main d'oeuvre et le matériel disponible (Figure 1.2)<sup>139</sup>. Cette diversité s'accroît lorsque l'on considère les systèmes de culture avec de l'élevage.

#### 1.1.3.2.4 Le peuplement cultivé

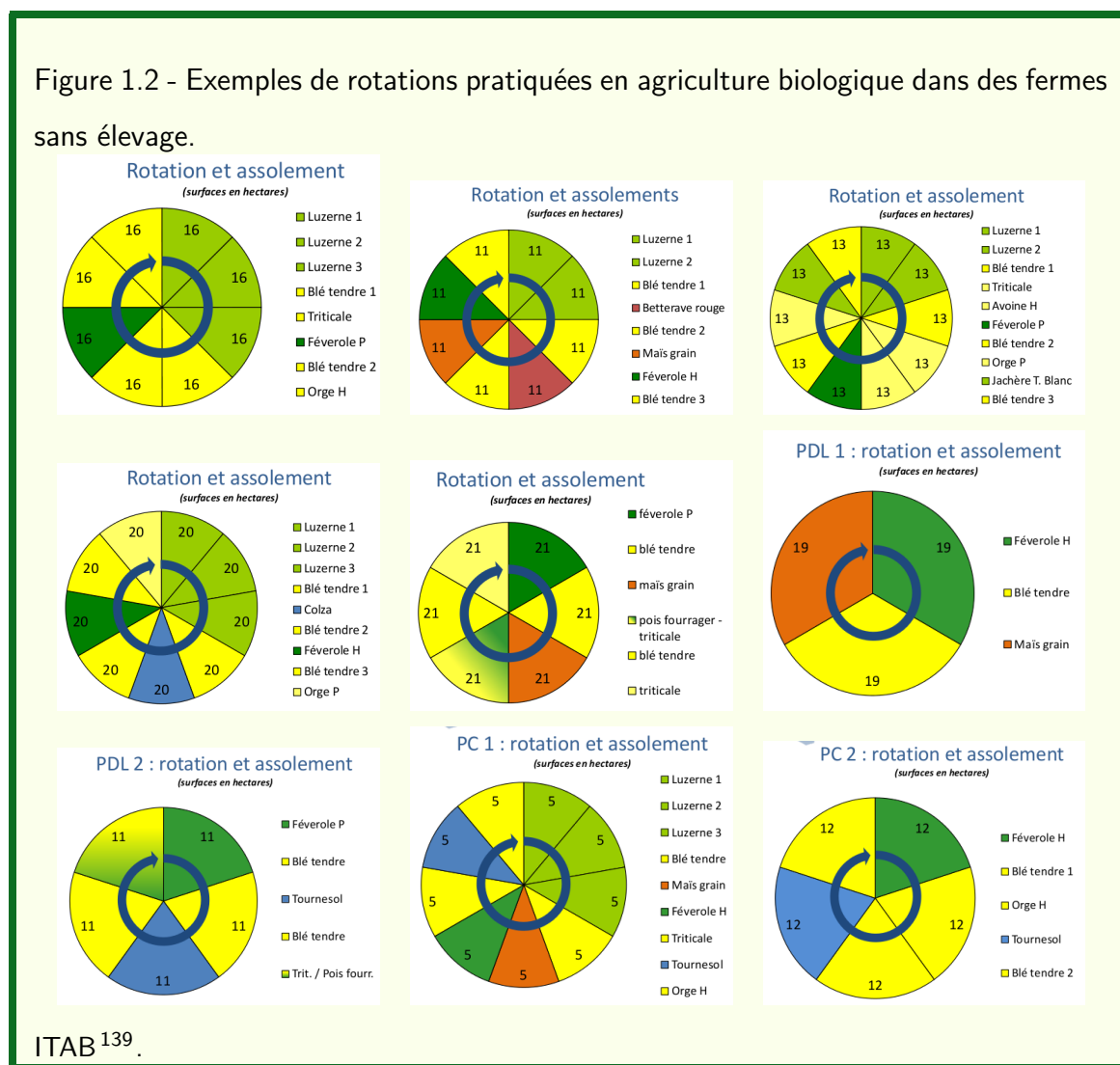
Le peuplement cultivé est soumis à différents stress d'origine biotique comme les maladies fongiques, les insectes et abiotique comme la sécheresse, la salinisation. Par exemple, pour le blé tendre, dans le cas des stress biotiques, deux catégories de maladies peuvent attaquer les plantes : les maladies du pied et des épis comme la fusariose et les maladies des feuilles et des épis comme les rouilles, la septoriose ou l'oidium (Annexe G). Les ravageurs que sont les pucerons et les taupins attaquent le blé tendre du semis jusqu'à la montaison ; plus tard les cécidomyies colonisent les épis (Annexe E). Les stress biotiques sont très variables et peu prévisibles : maladies et les ravageurs évoluent dans le temps et l'espace.

Les stress abiotiques, liés au gel, à la chaleur, au manque d'eau, à l'excès d'eau, peuvent être amplifiés par la présence d'adventices<sup>112</sup>. Dans un lieu donné, ils sont moins variables : les sécheresses annuelles, les hivers toujours rudes, ... Néanmoins, le changement climatique induit une augmentation de la variabilité de ces événements<sup>226</sup>, ce qui remet en question la relative stabilité de cette deuxième catégorie.

Pour lutter contre ces stress, l'agriculture conventionnelle utilise, en plus du choix variétal et parfois de l'irrigation, les produits phytosanitaires, herbicides, pesticides, insecticides issus de la chimie de synthèse. En agriculture biologique, ces produits sont interdits<sup>1</sup> et une solution basée sur la prise en compte des différentes composantes de l'agrosystème est privilégiée.

Une utilisation optimale de la diversité génétique<sup>255</sup> et des stratégies de gestion réfléchies : rotations des espèces, désherbage mécanique, densité de semis, pyramidage de gènes, mosaïques de variété dans un paysage, mélanges de variétés ...<sup>140;161</sup> permettent de s'adapter à ces stress qui évoluent dans le temps et l'espace.

Figure 1.2 - Exemples de rotations pratiquées en agriculture biologique dans des fermes sans élevage.



Hoad et al.<sup>130</sup> proposent de sélectionner des variétés capables de concurrencer les adventices pour l'accès en ressources en minéraux, eau et lumière. Les critères les plus importants sont une bonne implantation de la culture, un recouvrement du sol précoce, un fort pouvoir de tallage (pour les céréales à paille) et la hauteur (mais attention à la verse). Contre les maladies et les ravageurs, Döring et al.<sup>77</sup> proposent quatre stratégies de lutte :

1. utiliser des variétés résistantes,
2. renforcer la qualité du sol,
3. affaiblir la maladie et les pathogènes en capturant les insectes et en utilisant des produits phytosanitaires comme le cuivre, en gérant un environnement de culture capable d'accueillir les prédateurs naturels,
4. éviter la maladie par des itinéraires techniques qui jouent sur les rotations, la nutrition des plantes.

Alors que les trois derniers points sont liés à la gestion agronomique de l'agrosystème, le premier relève de la sélection variétale. La sélection peut se faire sur le critère de la maladie comme la résistance, la résistance induite, la tolérance, l'optimisation de la diversité intra-variétale et intra-parcelle (partie 1.2.3). Elle peut également porter sur la santé des plantes en général et sur l'architecture des plantes<sup>77</sup>.

### 1.1.3.3 La participation des acteurs

L'agroécologie est un mode d'action ancré dans le réel qui impacte et qui transforme le monde dans lequel il se situe<sup>152;212</sup>. Au delà des principes originels de l'agroécologie qui sont liés à la dimension agronomique des agrosystèmes, Stassart et al.<sup>222</sup> ajoutent des principes méthodologiques et socio-économiques. Les dispositifs de recherche participatives font partie de ces principes.

La recherche participative peut se définir comme un partenariat entre la recherche académique et la société civile.<sup>169</sup> D'après Diamond<sup>71</sup>, la société civile peut se définir comme *« le domaine de la vie sociale organisée qui se fonde sur le volontariat, la spontanéité, une autosuffisance, l'autonomie vis-à-vis de l'État, qui est lié par un ordre légal ou un ensemble de règles communes. Elle [...] implique des citoyens qui agissent collectivement dans un espace public pour exprimer leurs intérêts, leurs passions et leurs idées, échanger des informations, atteindre des buts communs, interpeler les pouvoirs publics et demander des comptes aux représentants de l'État. »*. Alors que l'État organise son action par l'intermédiaire des administrations, la société civile s'incarne dans des organes multiples comme les syndicats, les associations, les ONG, les réseaux informels qui participent à définir l'intérêt général<sup>18</sup>. La société civile est aussi l'expression couramment employée pour désigner les participants des forums sociaux et des mouvements altermondialistes<sup>48</sup>.

Comme dans tous les cas, dans le cas de l'agroécologie, il est difficile d'induire des changements sans le consentement des parties prenantes : praticiens, consommateurs ou plus largement citoyens. Ceux-ci peuvent accepter ou rejeter, ignorer ou adopter les innovations proposées par des experts<sup>222</sup>. La société civile a des attentes, des demandes, des visions et une force de proposition qui doivent être intégrées dans la construction de systèmes alimentaires alternatifs basés sur les concepts de l'agroécologie<sup>106;234;252</sup>. C'est de que souligne l'engagement 3 du PASMA :

*« nous poursuivrons des politiques et méthodes participatives et durables de développement alimentaire, agricole, [...] qui sont essentielles pour assurer des approvisionnements alimentaires adéquats et fiables au niveau des ménages ainsi qu'à l'échelon national, régional et mondial, [...] »*

En participant, les citoyens jouent un rôle dans l'établissement d'un modèle de société respectueux de l'environnement, sain et durable dans l'intérêt de tous, y compris des générations futures. Des groupes de la société civile se sont organisés sur ces questions. Citons, entre autres, les Incroyables Comestibles<sup>136</sup>, le mouvement Slow Food<sup>215</sup>, les Associations pour le Maintien d'une Agriculture Paysanne<sup>10</sup>. Ces mouvements illustrent le désir des citoyens de se réapproprier leur nourriture. La participation des citoyens, des scientifiques, des producteurs, des consommateurs, dans une construction scientifique rigoureuse permet de réaliser des changements dans des environnements complexes<sup>106;222</sup>. Associer savoirs savants et savoir-faire traditionnels permet de transformer des problématiques, de modifier des méthodologies et de contribuer à atteindre les résultats que la société civile souhaite<sup>222</sup>. C'est ce que reconnaît l'IAASTD :

*« la connaissance est coproduite par les chercheurs, les agriculteurs et tous les acteurs agricoles au sens large (forestiers, pêcheurs, éleveurs, pasteurs), mais aussi la société civile, les organisations et les administrations publiques. »*

Parmi les acteurs de ces actions participatives, les paysans jouent un rôle clé : ils produisent la nourriture. La mise en place de systèmes alimentaires durables n'est efficace que s'il est porté par les paysans eux mêmes. C'est leur participation qui permet de mieux cerner leurs besoins mais également les enjeux sociétaux et culturels liés à l'introduction de l'agroécologie<sup>7;234</sup>. La participation des paysans dans la gouvernance du système agricole est indispensable, les institutions jouant alors le rôle de facilitateur<sup>8</sup>.

#### **1.1.4 Semences et ressources génétiques au sein de l'agrobiodiversité**

L'agroécologie repose sur de nombreux principes parmi lesquels la diversification génétique dans l'espace et dans temps et la valorisation de l'agrobiodiversité<sup>222</sup>. L'agrobiodiversité correspond à une fraction de la biodiversité qui est utilisée directement ou indirectement pour la nourriture et l'agriculture. Elle concerne les ressources génétiques,

la biodiversité associée aux services éco-systémiques qui sont responsables de la régulation et du bon fonctionnement de l'agrosystème, et les savoirs et pratiques associés à cette diversité<sup>89;257</sup>. Ces savoirs et ces pratiques sont valorisés par la participation des différents acteurs : paysans, chercheurs, institutions, ONG, ... dans la gestion de l'agrobiodiversité. Cette démarche est encouragée par la FAO<sup>92</sup>.

Une meilleure utilisation et une meilleure gestion de l'agrobiodiversité permettent de répondre aux demandes d'une agriculture durable, respectueuse de l'environnement et capable de s'adapter aux changements climatiques et sociétaux<sup>92;94;103;107;211;234</sup>.

Un des objectifs pour l'agrobiodiversité est d'introduire ou de réintroduire de la diversité dans les agrosystèmes et de contribuer à une utilisation et une gestion durable des ressources génétiques<sup>126;211</sup>. Une partie importante de ces ressources génétiques, comme les semences, est directement utilisée dans les champs par un grand nombre de paysans avec les « variétés de pays »<sup>ix</sup><sup>94</sup>. Elles sont également la base de tout programme de sélection pour créer des « variétés modernes »<sup>x</sup><sup>160</sup>. Ce sont ces variétés, « variétés de pays » et « variétés modernes », qui produisent essentiellement la nourriture dans le monde<sup>234;257</sup>. De nombreuses autres productions en sont issues, par exemple les médicaments<sup>234</sup>.

\*\*\*

La partie 1.1 a montré que l'agroécologie peut être une solution pour répondre au problème de la souveraineté alimentaire. Au sein de l'agroécologie, l'agrobiodiversité a un rôle majeur. La partie 1.2 présente quelles méthodes peuvent être adoptées pour la gestion des ressources génétiques et la sélection variétale en agroécologie.

## **1.2 Stratégies de gestion des ressources génétiques et de sélection variétale pour l'agroécologie**

### **1.2.1 La gestion des ressources génétiques**

Dans tout programme de sélection, qu'il concerne des systèmes agroécologiques ou conventionnels, la problématique pour créer de nouvelles variétés reste la même : il faut créer de la diversité ou mobiliser la diversité existante pour pouvoir innover. A part les approches de mutagenèse et de type Organisme Génétiquement Modifié (OGM), qui ne

---

ix. On appelle variétés de pays les variétés issues de la sélection naturelle et de la sélection des paysans. Ces variétés sont très hétérogènes. Il en est question dans les parties 1.2, 1.3 et 1.4.

x. On appelle variétés modernes les variétés créées par les sélectionneurs qui sont disponibles sur les catalogues officiels. Ces variétés sont très homogènes. Il en est question dans les parties 1.2, 1.3 et 1.4.

seront pas abordés ici, le seul moyen de développer de nouvelles variétés est de croiser, de mélanger, de sélectionner dans la diversité des ressources génétiques.

Le système semencier « conventionnel », comme en Europe, produit de la nourriture dans les champs d'un côté et conserve les ressources génétiques dans des banques de semences (centres de ressources génétiques) de l'autre côté. Il s'agit d'une gestion *ex-situ*. Cette compartimentation des tâches, hérité d'un système fordiste<sup>xi</sup>, s'oppose à l'approche en réseau de l'agroécologie qui produit la nourriture et gère les ressources génétiques en même temps ; c'est la gestion à la ferme, qui est une forme de gestion *in-situ*.

#### 1.2.1.1 La gestion *ex-situ*

A partir des années 1960, la communauté internationale par le biais de la FAO, mesure l'importance des ressources génétiques pour l'alimentation mondiale et décide de mettre en oeuvre des stratégies de conservation de ces ressources<sup>25</sup>. Bien que les gestions *ex-situ* et *in-situ* soient également reconnues, les moyens financiers et humains sont principalement déployés pour la gestion *ex-situ*. Cette gestion *ex-situ*, qui consiste à conserver les ressources génétiques en dehors des conditions environnementales dans lesquelles elles se sont développées, c'est à dire en dehors du champ, représente actuellement le moyen le plus utilisé pour conserver les ressources génétiques dans le monde<sup>94</sup>.

Les ressources génétiques se répartissent entre matériels de recherche ou de sélection (21%), variétés modernes (17%), variétés de pays (44%) et espèces sauvages (18%)<sup>94</sup>. Les céréales représentent 45% des ressources génétiques conservées dans les banques de graines dans le monde. Certaines espèces orphelines, comme le chêne ou l'érable, sont mal représentées<sup>94</sup>.

En 2010, plus de 1 750 centres de ressources génétiques sont répertoriés dans le monde. Parmi ces centres, 130 contiennent plus de 10 000 accessions<sup>xii</sup> chacun (Figure 1.3). Au total, plus de 7,4 millions d'accessions sont stockées parmi lesquelles 25 à 30% sont des accessions originales, les autres étant des copies de sauvegarde de ces accessions que les centres s'échangent entre eux<sup>94</sup>. Les centres les plus importants sont les centres historiques construits par le Consultative Group for International Agricultural Research (CGIAR) dans les années 70, viennent ensuite les centres gérés par les gouvernements. Les universités, les jardins botaniques, les Organisation Non Gouvernementale (ONG), les entreprises gèrent également des centres de stockage<sup>94;171</sup>. Au delà des copies que s'échangent les

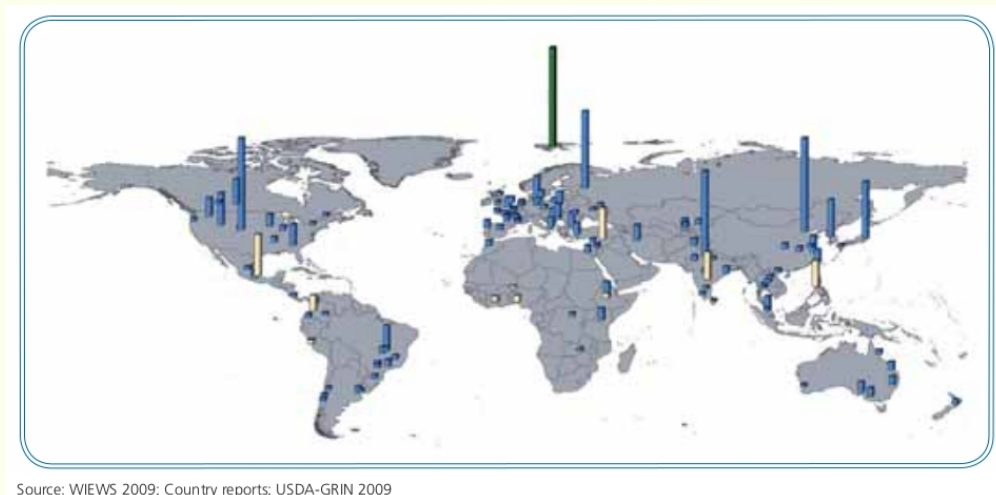
xi. Organisation du travail hiérarchisée et à la chaîne

xii. Une accession est un lot unique de semences défini par son histoire (lieu et année de collecte par exemple), ainsi que par une description phénotypique voire moléculaire.

centres entre eux, un filet de sécurité a été créé à Svalbard par le gouvernement norvégien. Ce site situé au pôle Nord, contient 412 000 accessions dupliquées provenant de centres du monde entier<sup>94</sup>.

En plus de conserver, les centres décrivent aussi et recensent l'histoire et les savoirs associés à ces ressources<sup>194;232</sup>. La gestion *ex-situ* est liée à un coûteux savoir faire technologique (chambre froide, base de données, cryo-conservation, culture *in vitro*), ce qui explique qu'elle soit prise en charge par les organismes de recherche ou les états. Les pays qui n'ont pas les moyens financiers de gérer les collections sont en difficulté devant ces contraintes. Cela engendre des risques de pertes des accessions<sup>94</sup>.

Figure 1.3 - Répartition géographique des banques de semences avec plus de 10 000 accessions.



Source: WIEWS 2009; Country reports; USDA-GRIN 2009

En bleu : les banques de graines nationales et regionales.  
 En beige : les centres CGIAR.  
 En vert : le centre de Svalbard.  
 D'après<sup>94</sup>.

La conservation *ex-situ* consiste en une gestion statique des ressources génétiques : elle vise à stopper l'évolution des échantillons conservés en chambre froide. Cette méthode peut conduire à ce que les plantes ne soient plus adaptées à leur environnement après quelques années de stockage. Par exemple, dans un environnement donné, les pathogènes continuent d'évoluer et développent de nouvelles virulences. Les plantes stockées en dehors de leur environnement n'étant pas en contact avec les pathogènes, ne pourront pas

évoluer, et pourront se trouver inadaptées à leur sortie<sup>128</sup>. De plus, les processus d'échantillonnage lors des collectes et lors de la régénération des échantillons peut conduire à certains « goulots d'étranglement »<sup>xiii</sup> induisant nécessairement une perte d'une partie de la diversité<sup>229</sup>. Les accessions stockées dans les centres de ressources génétiques ne peuvent pas être représentatives de toute la diversité présente dans les environnements d'origine<sup>9</sup>, notamment de la diversité au sein des variétés de pays<sup>229</sup>.

### 1.2.1.2 La gestion *in-situ*

C'est lors de la Convention sur la Diversité Biologique (CBD) réunie à Rio en 1992 que les instances internationales ont reconnu l'importance de la conservation *in-situ* pour les espèces sauvages mais aussi pour les espèces cultivées<sup>94;160</sup>. On passe alors du concept de « ressources génétiques » à celui de « biodiversité cultivée », d'un stock de gènes ou d'allèles à un flux de processus dynamiques qui maintiennent et créent de la variabilité génétique. C'est un véritable changement de paradigme entre une vision fixiste et une vision évolutionniste du vivant<sup>25</sup>.

La diversité est en perpétuel mouvement, régie par les forces évolutives que sont la sélection, la migration, la dérive et la mutation. Il convient donc de parler de gestion dynamique dans le cadre de la gestion *in-situ*<sup>122</sup>. Cette gestion dynamique *in-situ* a deux rôles : conserver la diversité génétique et les processus évolutifs qui en sont à l'origine et créer un réservoir d'individus qui peuvent être utilisés dans des programmes de sélection<sup>123</sup>. La gestion dynamique peut se conduire dans des stations expérimentales ou de recherche ou à la ferme dans les champs des paysans<sup>80</sup>.

La gestion dynamique se base sur la théorie des méta-populations. Une méta-population est un ensemble de sous-populations dynamiques qui occupent un nombre de sites variables dans l'espace et le temps, et sont connectées, ou non, entre elles par des échanges de gènes. Ces échanges peuvent venir du pollen ou d'échanges physiques de graines. Chaque sous-population évolue en réponse aux conditions environnementales auxquelles elle est soumise<sup>183</sup> (Figure 1.4). Dans ce cadre, la migration est un point clé de la gestion dynamique. Ces échanges diminuent en effet la probabilité d'extinction par accumulation de fardeau génétique et pertes d'allèles<sup>xiv</sup> favorables<sup>80;127</sup>. Ils sont une pratique courante chez les paysans, qui se regroupent le plus souvent en réseau<sup>191;230</sup>. Dans les stations expérimentales, l'animation par un intervenant est indispensable pour favoriser les échanges

---

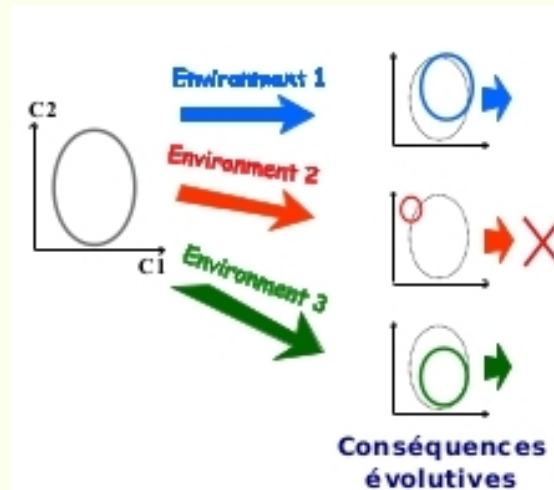
xiii. Réduction aléatoire ou pilotée du nombre de plantes ou de grains prélevés pour constituer un échantillon.

xiv. Un allèle est une version d'un gène à un locus donné.



de semences.

Figure 1.4 - Fonctionnement d'une méta-population dans le contexte de la gestion dynamique.



La population de départ (grise) est cultivée dans différents environnements et est soumise à différentes forces qui la font évoluer, se différencier (populations bleue et verte) ou s'éteindre (population rouge). Chaque population conserve une partie de la diversité de départ, voire crée une nouvelle diversité.

La taille est proportionnelle à la diversité.

Schéma repris de Thomas<sup>229</sup>.

Une expérience de gestion dynamique de la diversité dans des stations conduite par le laboratoire de Génétique Végétale du Moulon a débuté en France en 1984. Plusieurs populations à base génétique très large ont été cultivées dans différentes stations réparties en France. Rapidement (cinq à dix générations), les populations ont évolué selon les lieux de culture quant à la hauteur, la précocité et la résistance aux maladies<sup>80;153;187;188</sup>. Chaque population s'est adaptée localement en perdant de la diversité. Cependant, dans le réseau expérimental au total, la diversité a été conservée<sup>80</sup>. Ce mode de gestion a également permis de détecter de nouveaux allèles et de nouvelles combinaisons alléliques associées à la précocité qui n'existaient pas dans les populations de départ<sup>202</sup>.

La gestion dynamique en station a des limites : elle repose souvent sur un faible nombre d'environnements correspondant aux conditions relativement standardisées des stations expérimentales. Il peut être intéressant de développer un tel processus directement sur les fermes<sup>9</sup>. Ce qui existe encore dans les pays du Sud : les paysans cultivent des variétés

locales de pays et jouent un grand rôle dans la gestion des espèces cultivées<sup>94</sup>. Dans les pays où l'agriculture est principalement industrielle, de plus en plus de paysans organisés en associations se mettent à cultiver des variétés de pays délaissées<sup>64;177;230</sup>. Cette gestion dynamique à la ferme est liée aux structures sociales et culturelles des communautés et des associations de paysans<sup>191</sup>. Par exemple, dans leur étude sur la variété de pays de blé *Rouge-de-Bordeaux*, gérée par des paysans du Réseau Semences Paysannes (RSP) en France, Thomas et al.<sup>231</sup> ont montré que les pratiques des paysans ont un impact sur la diversité des populations de blé. La circulation des semences ainsi que les pratiques de gestion des semences sont les principaux facteurs de structuration de la diversité. La diversité conservée à la ferme apparaît complémentaire de celle conservée en banques de graines. Ces deux stratégies, *ex-situ* et *in-situ*, sont complémentaires<sup>65;80;194</sup>.

Au delà des effets de l'environnement, la gestion à la ferme est directement liée aux contextes socio-économiques locaux. Cette gestion décentralisée, en faisant intervenir de nombreux acteurs, maximise le maintien et la création de variabilité génétique. De plus, cette diversité, adaptée localement, pourra servir de base pour des programmes de sélection décentralisés (partie 1.2.2) et participatifs (partie 1.3.3)<sup>80;123</sup>.

Enfin, au niveau international, l'article 9.1 du Traité International sur les Ressources Phytogénétiques pour l'Alimentation et l'Agriculture (TIRPAA) (On revient sur ce traité dans la partie 1.4.) reconnaît « *l'énorme contribution que les communautés locales et autochtones ainsi que les agriculteurs [...], ont apportée et continueront d'apporter à la conservation et à la mise en valeur des ressources phytogénétiques [...]* ». Ce dernier point est repris avec l'article 5.1.c qui invite à « *encourager ou soutenir, selon qu'il convient, les efforts des agriculteurs et des communautés locales pour gérer et conserver à la ferme leurs ressources phytogénétiques pour l'alimentation et l'agriculture.* ».

### 1.2.2 La sélection décentralisée

Les agrosystèmes sont d'une grande complexité et d'une grande diversité (partie 1.1.3.2.1). L'enjeu de la sélection variétale en agroécologie est l'adaptation à cette diversité d'agrosystèmes. L'amélioration des plantes repose sur la sélection de caractères qui vont être transmis à leur descendance. Un caractère phénotypique ( $P$ ) est la somme d'effets génétiques ( $G$ ) et environnementaux ( $E$ ) :  $P = G + E$ . Si les effets génétiques et environnementaux sont indépendants, alors on a  $var(P) = var(G) + var(E)$ , avec  $var$  la variance. La variance génétique peut être divisée en variance génétique « additive » et en variance génétique « non additive ». C'est la variation génétique « additive » de la

variation phénotypique observée qui est transmise à la descendance alors que la variation génétique « non additive » et la variation due à l'environnement ne sont pas transmises. Pour quantifier cette part due à la génétique, le sélectionneur calcule l'héritabilité d'un caractère. L'héritabilité au sens large ( $h_{sl}^2$ ) est la part de variation phénotypique qui est due à la génétique (variance génétique totale). L'héritabilité au sens strict ( $h_{ss}^2$ ) est la part de variation phénotypique qui est due à la variance génétique additive ( $\sigma_A$ , celle qui est transmise à la descendance). On a donc  $0 \leq h_{ss}^2 \leq h_{sl}^2 \leq 1$ .

Dans un contexte où l'on souhaite sélectionner dans un environnement donné, l'héritabilité au sens large s'écrit telle que :

$$h_{sl}^2 = \frac{var(G)}{var(G) + var(E)}$$

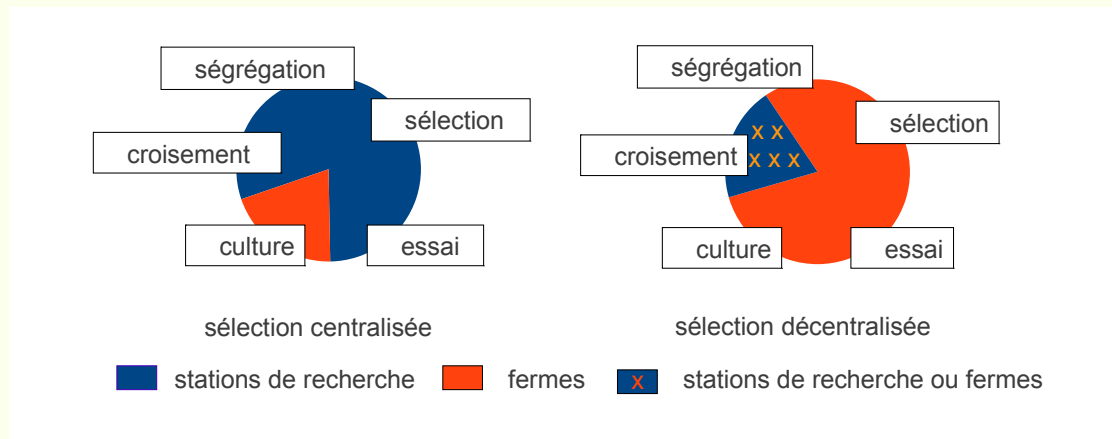
Afin d'estimer au mieux  $var(G)$ , il faut contrôler les conditions de l'environnement afin de minimiser  $var(E)$ . Cela est possible en utilisant un dispositif expérimental orthogonal, c'est à dire sans confusions d'effets entre la partie génétique et la partie environnementale de la variation, avec des répétitions. Plus il y aura de répétitions et plus l'héritabilité sera bien estimée et la sélection fiable<sup>111</sup>. Cela est illustré par le réseau d'essai du Comité Technique Permanent de la Sélection (CTPS) en France qui évalue les variétés en vue de leur inscription au catalogue officiel : les variétés sont semées dans deux à quatre blocs<sup>52</sup>.

Lorsque l'on raisonne dans un réseau multi-essais, le caractère phénotypique  $P$  est la somme d'effets génétiques ( $G$ ), environnementaux ( $E$ ) et de l'interaction entre ces deux derniers ( $G \times E$ ) :  $P = G + E + G \times E$ . L'héritabilité au sens large s'écrit alors telle que :

$$h_{sl}^2 = \frac{var(G)}{var(G) + var(E) + var(G \times E)}$$

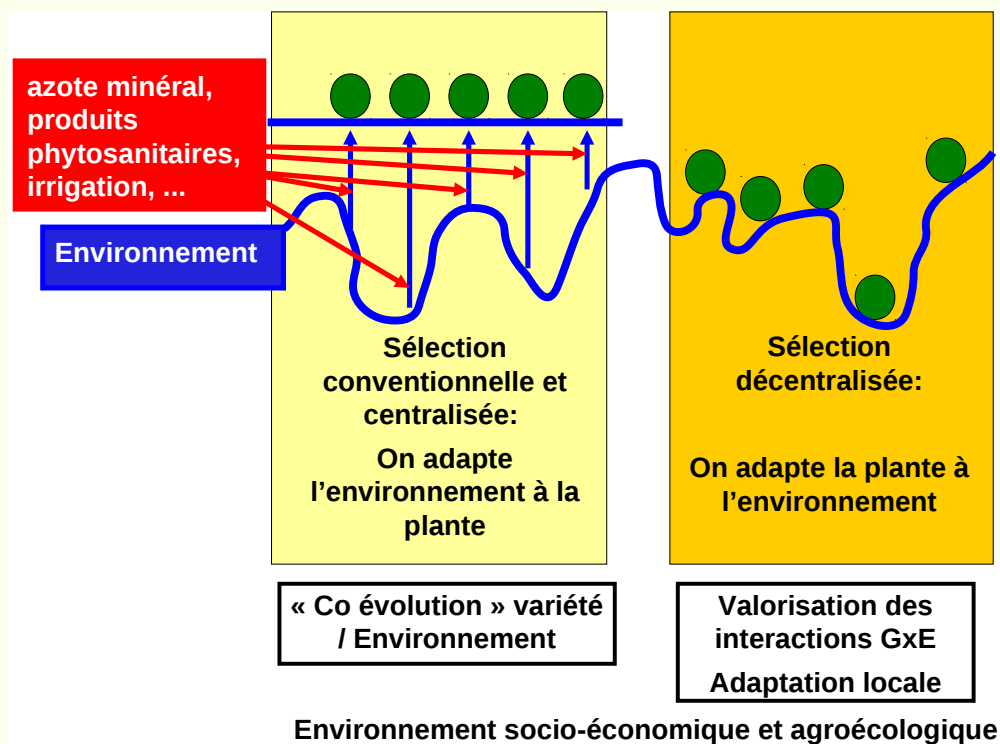
L'héritabilité sur le réseau est donc inférieur ou égale à l'héritabilité sur un site donné du simple fait de la présence d'une variance due à l'interaction ( $G \times E$ ) qui est supérieure à zéro. C'est ce terme au dénominateur qui va être pris en compte différemment dans les deux modalités de sélection : la sélection centralisée et la sélection décentralisée. Dans la sélection centralisée, les différentes étapes du processus de sélection, de la création de variabilité aux essais, se font à la station de recherche. Dans la sélection décentralisée, les différentes étapes du processus sont directement faites dans les environnements cibles, c'est à dire dans les environnements où seront effectivement cultivées les variétés (Figure 1.5). Ces environnements cibles peuvent aller de la région climatique jusqu'à la ferme.

Figure 1.5 - Différences entre sélection centralisée (gauche) et la sélection décentralisée (droite).



Inspiré de Ceccarelli et Grando<sup>39</sup>.

Figure 1.6 - Représentation des différences entre une sélection centralisée et une sélection décentralisée.



L'environnement est représenté par la ligne bleue.

Les ronds verts représentent les variétés.

Les flèches rouges représentent les apports d'intrants qui lissent l'environnement.

La sélection centralisée fait l'hypothèse que l'environnement où est faite la sélection n'est pas différent des environnements où seront cultivées les plantes : on parle de sélection indirecte. Ce type de sélection est utilisé en agriculture conventionnelle pour deux raisons. La première est que les entreprises semencières n'ont pas les moyens financiers, humains ni logistiques de sélectionner dans des environnements spécifiques représentant des marchés trop petits<sup>68</sup>. La deuxième est que l'agriculture conventionnelle utilise des intrants qui ont tendance à homogénéiser les environnements et à créer des conditions proches de celles de la station de recherche (Figure 1.6)<sup>49;57</sup>. Par exemple, le réseau d'évaluation du CTPS pratique, selon les spécificités de la céréale étudiée, sur une ou deux zones d'expérimentation (nord et sud), un nombre d'essais allant de sept à 21 pendant deux années. Les conditions de ces essais sont celles de l'agriculture conventionnelle avec utilisation de produits phytosanitaires et apport d'azote minéral<sup>52</sup>. La sélection centralisée peut aussi être utilisée pour développer des variétés en agriculture biologique. C'est le cas pour le blé tendre en France où la sélection se fait à Rennes et l'évaluation dans 16 sites dans le quart nord-ouest de la France<sup>206</sup>.

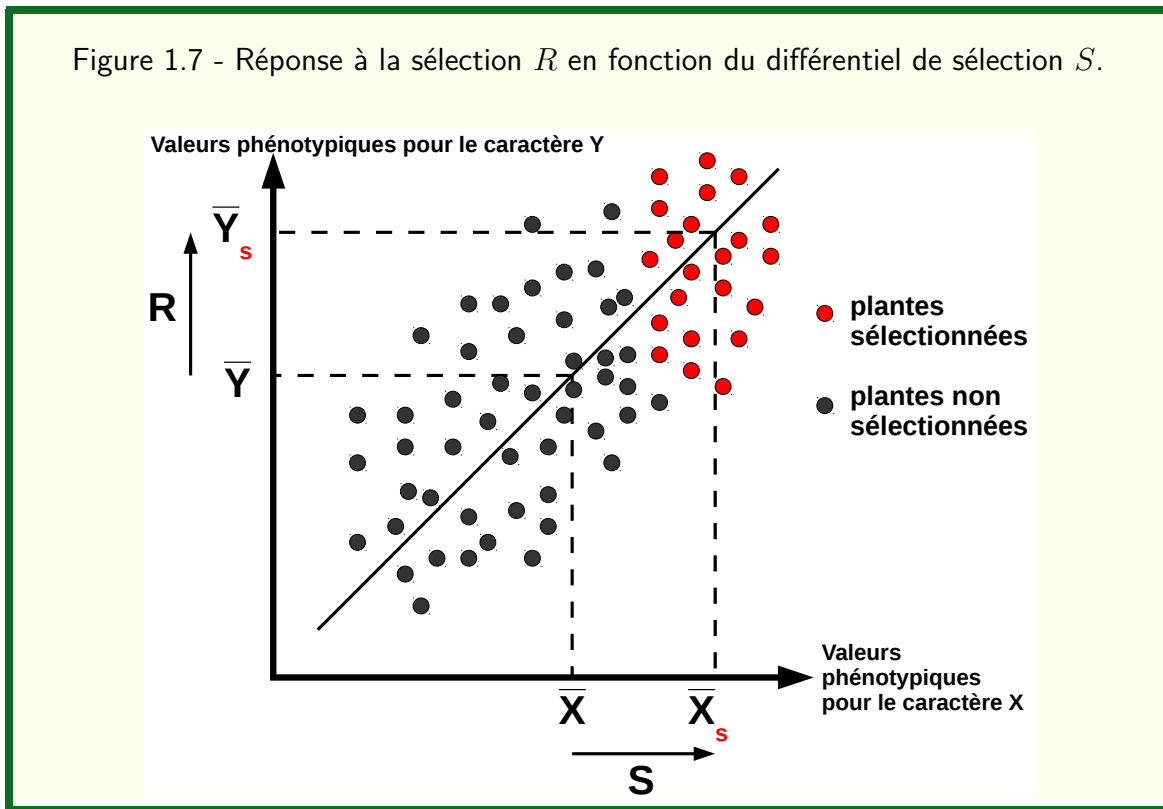
La sélection indirecte peut se concevoir comme suit :

Soit  $S$ , le différentiel de sélection : c'est la différence de moyenne entre l'échantillon des individus sélectionnés ( $\bar{X}_S$ ) et la moyenne d'un échantillon d'individus non sélectionnés pris au hasard dans la population de départ ( $\bar{X}$ ) :  $S = \bar{X}_S - \bar{X}$  (Figure 1.7).

$S$  dépend uniquement de la proportion d'individus sélectionnés et de l'écart type phénotypique du caractère dans la population  $\sigma_P(X)$ . Pour une même proportion d'individus sélectionnés, la valeur de  $S$  sera différente selon la variance du caractère. Avec une variance forte,  $S$  sera plus important qu'avec une variance faible.

Pour pouvoir comparer des valeurs de  $S$  pour différents caractères ou différentes populations, il est intéressant de corriger  $S$  par la variance de la population. Pour ce faire, on utilise l'intensité de sélection  $i$  telle que  $i = \frac{S}{\sigma_P}$ ,  $i$  est en unité d'écart type. Il vient alors  $S = i \times \sigma_P(X)$ . Plus le taux de sélection, c'est à dire la proportion d'individus sélectionnée, est faible et plus l'intensité de sélection est forte<sup>111</sup>.

Soit  $R$ , la réponse à la sélection : c'est la différence de moyenne entre un échantillon d'individus pris au hasard parmi les descendants de la population sélectionnée ( $\bar{Y}_S$ ) et un échantillon d'individus pris au hasard parmi les descendants des individus non sélectionnés ( $\bar{Y}$ ) :  $R = \bar{Y}_S - \bar{Y}$  (Figure 1.7).



On pose  $R_{Y|X}$ , la réponse à la sélection dans l'environnement  $y$  pour le caractère mesuré  $Y$  sachant que la sélection a été faite dans l'environnement  $x$  pour le même caractère mesuré dans cet environnement ( $X$ ).  $X$  et  $Y$  sont les mesures phénotypiques du même caractère mesuré dans deux environnements,  $x$  et  $y$ . Pour prédire  $R_{Y|X}$ , on écrit la prédiction linéaire de  $\hat{Y}$  en fonction de  $X$  :

$$\hat{Y} = \bar{Y} + b_{XY}(X - \bar{X})$$

Pour la valeur des individus sélectionnés, on a :

$$\hat{Y}_S = \bar{Y} + b_{XY}(\bar{X}_S - \bar{X})$$

D'où

$$\hat{Y}_S - \bar{Y} = b_{XY}(\bar{X}_S - \bar{X})$$

Ce qui revient à

$$R_{Y|X} = b_{XY} \times S$$

Avec  $b_{XY}$  le coefficient de régression de la droite.

Sachant que  $b_{XY} = \frac{cov(X,Y)}{\sigma_P^2(X)}$  et que  $S = i \times \sigma_P(X)$ , il vient

$$R_{Y|X} = \frac{cov(X,Y)}{\sigma_P^2(X)} \times i \times \sigma_P(X) = \frac{cov(X,Y)}{\sigma_P(X)} \times i$$

Soit  $r_A(X,Y)$ , la corrélation génétique additive entre les individus mesurés dans l'environnement  $x$  et les individus mesurés dans l'environnement  $y$ . Sachant que  $r_A(X,Y) = \frac{cov(X,Y)}{\sigma_A(X) \times \sigma_A(Y)}$ , il vient  $cov(X,Y) = r_A(X,Y) \times \sigma_A(X) \times \sigma_A(Y)$ , avec  $\sigma_A(X)$ , la variance génétique additive de  $X$  (dans l'environnement  $x$ ) et  $\sigma_A(Y)$ , la variance génétique additive de  $Y$  (dans l'environnement  $y$ ). On a alors,

$$R_{Y|X} = \frac{i \times r_A(X,Y) \times \sigma_A(X) \times \sigma_A(Y)}{\sigma_P(X)}$$

Sachant que  $\sigma_A(X) = h_{ss}(X) \times \sigma_P(X)$ , avec  $h_{ss}(X)$  l'héritabilité au sens strict du caractère  $X$  dans l'environnement  $x$  et que  $\sigma_A(Y) = h_{ss}(Y) \times \sigma_P(Y)$ , avec  $h_{ss}(Y)$  l'héritabilité au sens strict du caractère  $Y$  dans l'environnement  $y$ , il vient

$$R_{Y|X} = i \times r_A(X,Y) \times h_{ss}(X) \times h_{ss}(Y) \times \sigma_P(Y)$$

Si les environnements  $x$  et  $y$  sont les mêmes, les caractères mesurés  $X$  et  $Y$  aussi, alors la réponse est  $R_Y = i \times h_{ss}^2(Y) \times \sigma_P(Y)$ .

Il y a trois leviers pour augmenter  $R_{Y|X}$ <sup>15</sup> :

1. augmenter  $i$  en ne sélectionnant que quelques plantes dans une diversité large,
2. augmenter la corrélation génétique additive entre les caractères mesurés dans les deux environnements ( $r_A(X,Y)$ ), cela revient à ce que l'environnement où la sélection a lieu soit le plus proche possible de l'environnement cible,
3. et augmenter l'héritabilité au sens strict dans l'environnement de sélection ( $h_{ss}(X)$ ).

Si on veut comparer l'efficacité de la sélection directe de  $Y$  ( $R_Y$ ) et l'efficacité de la sélection indirecte ( $R_{Y|X}$ ), il faut étudier le ratio  $\frac{R_{Y|X}}{R_Y}$ .

$$\begin{aligned} \frac{R_{Y|X}}{R_Y} &= \frac{i \times r_A(X,Y) \times h_{ss}(X) \times h_{ss}(Y) \times \sigma_P(Y)}{i \times h_{ss}^2(Y) \times \sigma_P(Y)} \\ &= r_A(X,Y) \times \frac{h_{ss}(X)}{h_{ss}(Y)} \end{aligned} \quad (1.1)$$

Plus le ratio  $\frac{R_{Y|X}}{R_Y}$  sera grand et plus la sélection indirecte sera efficace par rapport à la sélection directe. Cela sera le cas si  $h_{ss}(X) \gg h_{ss}(Y)$  et que la corrélation génétique additive entre les caractères mesurés dans les deux environnements est élevée.

La corrélation génétique entre les milieux difficiles et les milieux favorables est souvent très faible, voire négative, ce qui rend dans la réalité la sélection indirecte souvent peu efficace<sup>15;33;256</sup>. Par exemple, Murphy et al.<sup>173</sup> montrent sur le blé des corrélations génétiques modérées voire faibles entre une conduite en agriculture conventionnelle et une conduite en agriculture biologique. D'autres exemples de ce type sont donnés dans Wolfe et al.<sup>256</sup> et Messmer et al.<sup>163</sup>. Cela confirme les travaux de Falconer<sup>85</sup> qui montre que d'un point de vue théorique, la sélection directe est presque toujours plus efficace que la sélection indirecte. A l'inverse, Annicchiarico et al.<sup>12</sup> montrent des corrélations génétiques et des héritabilités fortes favorisant l'idée de sélection indirecte.

Les stations de recherche ne représentent pas la grande diversité ni l'hétérogénéité des agrosystèmes basés sur l'agroécologie. De plus, les conditions au sein d'un même champ en agriculture biologique peuvent être hétérogènes<sup>6</sup>. Dans ce cas, les variations phénotypiques observées au sein d'un champ peuvent être, en grande partie, dues à des variations environnementales. Ces variations peuvent engendrer alors une baisse de l'héritabilité et donc une diminution de la réponse à la sélection. Cependant certaines études ont montré que l'héritabilité n'est pas toujours plus faible dans les milieux à forte variance environnementale et/ou stressant<sup>37;173</sup>. Dans certains cas, l'héritabilité dépend plus de la variabilité génétique ( $G$ ) et de l'adaptation d'un génotype à un lieu ( $G \times E$ ) que de l'environnement ( $E$ ) *stricto sensu*<sup>37</sup>.

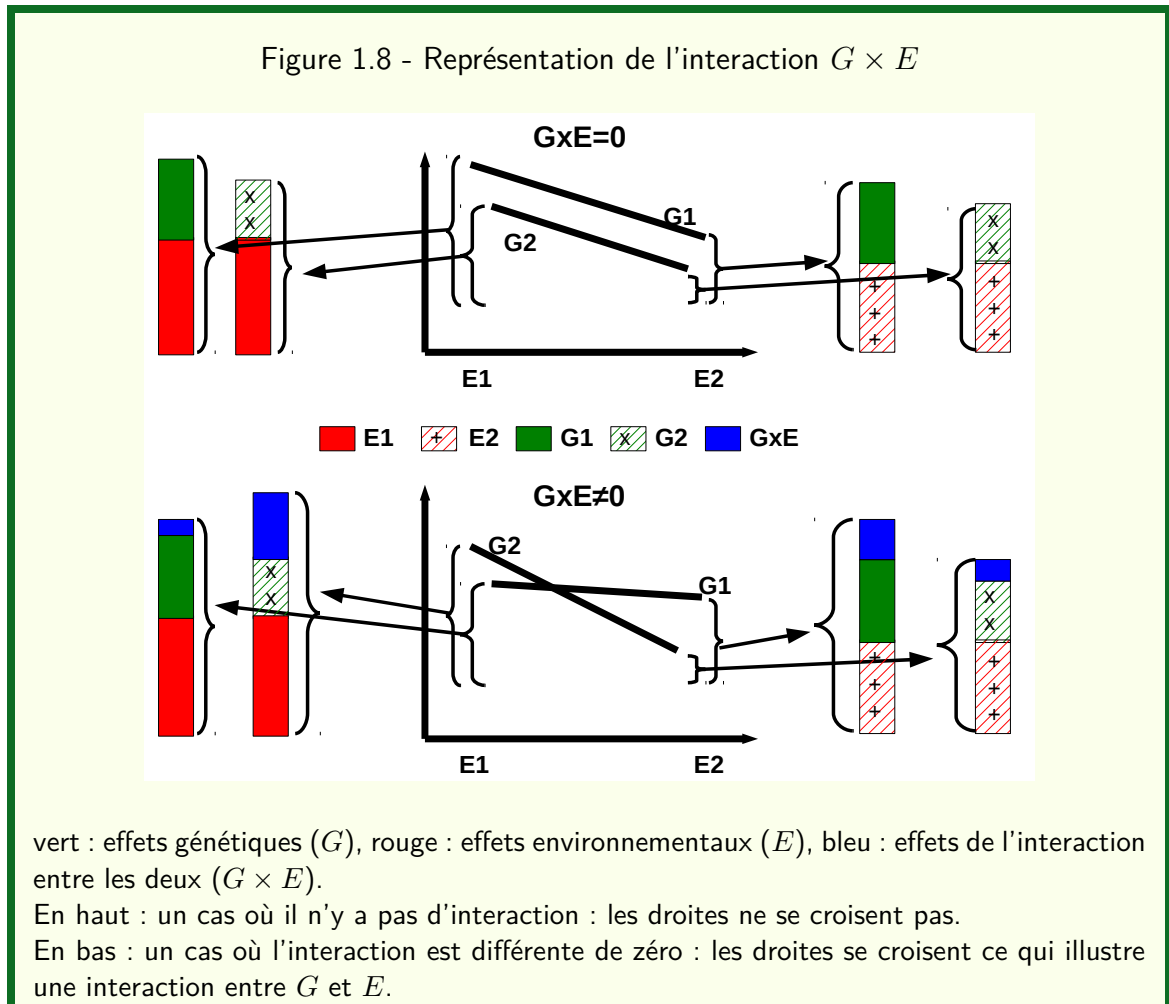
Un moyen de coller au mieux à la diversité des agrosystèmes est de décentraliser la sélection. La sélection décentralisée évalue et sélectionne les variétés directement dans les environnements cibles : on parle de sélection directe (Figure 1.6). Ce mode de sélection permet de mieux prendre en compte les interactions  $G \times E$  importantes dans les milieux contrastés comme en agriculture biologique<sup>39;40;57;114;259</sup>.

Il y a deux stratégies pour prendre en compte les interactions  $G \times E$ <sup>40</sup>. La première est de minimiser ces interactions et de sélectionner des variétés adaptées largement. Dans ce cas, le sélectionneur cherche des rendements stables ou une résistance aux maladies efficace sur la plus grande gamme d'environnements possibles : les mega-environnements<sup>113;199</sup>; La deuxième est de valoriser ces interactions. La maximisation des interactions positives permet de sélectionner des variétés pour une adaptation spécifique<sup>43;173;216;249</sup>.

Le bas de la Figure 1.8 représente un cas où une variété ( $G2$ ) est bien adaptée à un environnement en particulier ( $E1$ ) et médiocrement dans un autre ( $E2$ ) (cas d'interaction croisée). A l'inverse, une autre variété ( $G1$ ) est plus stable et se comporte en moyenne mieux sur les environnements ( $E1$  et  $E2$ ). Dans ce cas, il peut être intéressant de conserver



la variété adaptée à son environnement. La sélection décentralisée est efficace lorsqu'il y a des interactions  $G \times E$  importantes; elle vise alors à identifier les meilleurs génotypes dans chaque environnement et non un génotype moyen universel<sup>6;37</sup>.



### 1.2.3 La structure génétique des variétés

Décentraliser la sélection est la première étape pour sélectionner des variétés adaptées à la diversité des agrosystèmes basés sur l'agroécologie. Mais que la sélection soit centralisée ou non, on peut développer différentes structures variétales. La structure génétique des variétés peut être homogène, c'est à dire avec une variabilité intra-variétale nulle (lignées pures<sup>xv</sup>, hybrides F1<sup>xvi</sup>, clones) ou hétérogène, c'est à dire avec une variabilité intra-

xv. Une lignée pure est une variété issue d'autofécondation où toutes les plantes sont identiques. Les caractères ne bougent pas d'une génération à l'autre.

xvi. Un hybride F1 est le résultat d'un croisement entre deux plantes homozygotes, il correspond à la première génération où toutes les plantes sont identiques. Un hybride est 100% hétérozygote, de ce fait,

variétale importante (mélange, variété-population).

Les variétés hétérogènes participent à la préservation de la biodiversité *in-situ* (partie 1.2.1). En outre, elles ont un potentiel pour s'adapter à la diversité des agrosystèmes et peuvent leur conférer stabilité et résilience et contribuer aux services écosystémiques<sup>125;142;179;186</sup>. Ces caractéristiques répondent au point 6 de la Déclaration de Rome selon lequel « *on peut réduire l'instabilité saisonnière et inter-annuelle néfaste des approvisionnements alimentaires. Les progrès doivent notamment viser à réduire le plus possible la vulnérabilité aux fluctuations climatiques et aux ravageurs et maladies et leurs conséquences.* »

### 1.2.3.1 La variabilité pour s'adapter

Les variétés hétérogènes (nommées « populations » dans la suite de ce texte), représentent un réservoir de diversité qui évolue au gré des forces évolutives qui vont les faire s'adapter. L'adaptation génétique est possible si les populations sont hétérogènes pour des caractères liés à la valeur sélective : par exemple le nombre de grains, le poids des grains ou le taux de germination, et que ces caractères sont héréditaires. De plus, la plasticité phénotypique, qui est la capacité d'une population à exprimer un phénotype différent dans différents environnements, va également avoir un rôle dans l'adaptation. La plasticité correspond aux effets liés à l'environnement et aux interactions  $G \times E$  s'il y a plusieurs populations dans plusieurs sites.

Dans ces mécanismes, on retrouve les notions d'héritabilité et d'interaction  $G \times E$  qui sont des éléments clés de la sélection décentralisée (partie 1.2.2). Ce mode de sélection permet de valoriser les adaptations locales qui se mettent en place dans des populations<sup>122</sup>.

La sélection végétale évolutive<sup>xvii</sup> permet l'adaptation des populations à un agrosystème. Elle repose sur la culture d'une population avec une diversité très large dans les environnements cibles<sup>227</sup>. Au fur et à mesure des générations, les individus avec la meilleure valeur sélective se reproduisent plus et leurs gènes augmentent en fréquence dans la population. On parle de sélection darwinienne : la composition génétique de la population va évoluer et changer au cours du temps. Plus la diversité de départ sera importante, plus il y aura de chance de trouver des individus adaptés et plus la sélection sera efficace<sup>78</sup>. Goldringer et al.<sup>123</sup> montrent que les conditions climatiques influencent beaucoup la différenciation de populations de blé génétiquement hétérogènes, par exemple les

---

les caractères vont changer d'une génération à l'autre.

xvii. la formule en anglais « evolutionary breeding » est couramment utilisée.

populations cultivées dans le sud de la France deviennent plus précoces que celles du nord. De même, l'étude de la résistance à l'oïdium du blé dans ces mêmes populations montre un changement de fréquence des gènes de résistance spécifique après huit et dix ans de culture<sup>153;187;188</sup>. D'autres exemples de ce type sont présentés par Dawson et Goldringer<sup>56</sup>.

Cependant, la sélection naturelle, au sein des populations d'espèces cultivées, ne favorise pas toujours les caractères agronomiques ni les qualités qui présentent de l'intérêt pour les paysans<sup>256</sup>. Par exemple, la compétition entre individus pour la lumière qui favorise les plantes hautes peut diminuer la production en grain de la population<sup>122;198</sup>. Les plantes hautes peuvent être intéressantes pour concurrencer les adventices, mais elles sont plus susceptibles de verser. Dans ce cas, la valeur sélective individuelle est antagoniste de la valeur sélective de la population. Il est alors recommandé de modérer les interactions compétitives entre les plantes.

Un programme de sélection peut être conduit pour faire évoluer de façon plus dirigée les populations dans le sens souhaité. Les variétés conservées à la ferme et adaptées localement peuvent constituer la base à partir de laquelle on croisera, mélangera, sélectionnera (sélection massale intra-population, sélection inter-population)<sup>56;122</sup>. Cette sélection peut mener à des lignées pures ou bien à des populations moins hétérogènes qui conservent une certaine variabilité intra-variétale<sup>192</sup>.

### 1.2.3.2 La variabilité pour conférer stabilité et résilience

La diversité intra-variétale présente de nombreux avantages pour contribuer à la stabilité et la résilience du peuplement cultivé. La variabilité génétique peut être une réponse à la stochasticité<sup>xviii</sup> inter-annuelle des conditions environnementales en jouant l'effet de tampon et en augmentant la stabilité et la résilience des systèmes aux stress biotiques<sup>97;98</sup> et abiotiques liés au climat<sup>200</sup>.

La diversité permet d'augmenter la probabilité qu'une partie de la population soit adaptée aux conditions environnementales qui peuvent être changeantes<sup>78</sup>. Au sein d'une population, si par exemple les génotypes n'ont pas les mêmes besoins en termes de lumière ou d'accès aux nutriments de par leur système racinaire superficiel ou profond, alors ils n'utiliseront pas les mêmes ressources<sup>78</sup>. Les génotypes différents vont se compléter et au sein de la population, il n'y aura pas de compétition entre eux mais une possible coopération. Par exemple, un génotype donné peut communiquer avec un autre par l'intermédiaire de messagers chimiques volatiles et induire des réponses plus fortes aux agresseurs<sup>72;180</sup>.

---

xviii. définit un caractère aléatoire.

L'hétérogénéité des populations peut leur permettre d'associer des résistances quantitatives et qualitatives différentes qui augmentent la durabilité des résistances et préviennent l'évolution rapide des populations de pathogènes et d'insectes<sup>97;237</sup>. Un exemple célèbre est celui de Zhu et al.<sup>260</sup> qui montrent que différentes variétés de riz cultivées en association ont un rendement 89% plus important et sont plus résistantes aux principales maladies fongiques (94% en moins de sévérité) que lorsqu'une variété est cultivée seule. La sélection de variétés résistantes, souvent reposant sur des gènes majeurs, ou l'utilisation de fongicide sont des solutions peu durables car les résistances aux gènes et aux molécules des produits phytosanitaires sont rapidement contournées<sup>255</sup>. Au delà d'une meilleure résistance aux maladies, les mélanges ont un rendement équivalent, voire supérieur<sup>150</sup> et peuvent diminuer l'effet de certains stress abiotiques comme le froid<sup>255</sup>. Enfin, la qualité des mélanges est maintenue voire augmentée par une meilleure valorisation de l'azote par exemple<sup>167</sup>.

Les mélanges multilignée partagent un fond génétique proche, ce qui diminue leur adaptabilité à l'évolution des stress<sup>98</sup>. Un mélange de deux lignées pures comportera deux haplotypes<sup>xix</sup>. Cette diversité pourra augmenter si le mélange est resemé mais le potentiel adaptatif restera restreint aux allèles présents dans les lignées de départ dans un premier temps. Une population hétérogène aura un grand nombre d'haplotypes et donc une diversité plus importante<sup>229</sup>.

Au delà de la structure de la variété en elle-même, la répartition des variétés dans le paysage est importante. Une répartition raisonnée des variétés, notamment liée à la diversité des gènes de résistance, contribuera à la stabilité et la résilience de l'agrosystème<sup>189</sup>.

En plus de son rôle direct sur le peuplement cultivé, la variabilité intra-population a un rôle indirect dans l'agrosystème en contribuant à certains services éco-systémiques<sup>125;142</sup>.

Chateil et al.<sup>47</sup> montrent qu'un mélange de populations hétérogènes génétiquement cultivés en agriculture biologique accueille plus de biodiversité sauvage qu'une variété lignée pure. La structure hétérogène du peuplement augmente le nombre d'arthropodes (collemboles, araignées et carabes), sans doute en relation avec la fourniture d'un habitat plus varié. Cette biodiversité pourrait en retour contribuer à augmenter les régulations biologiques dans l'agrosystème.

Une méta analyse de la bibliographie montre une tendance selon laquelle les systèmes

---

xix. Un haplotype correspond à une succession d'allèles à l'ensemble des locus considérés. Si une plante est homozygote, elle aura un haplotype. Si une plante est hétérozygote, elle aura au moins deux haplotypes.

en agriculture biologique ont un impact positif sur la richesse et l'abondance des espèces<sup>20</sup> et par là même sur les services éco-systémiques associés qui sont indispensables au bon fonctionnement de l'agrosystème<sup>16;162</sup>.

\*\*\*

La partie 1.2 a présenté les stratégies spécifiques à la gestion des ressources génétiques et à la sélection variétale en agroécologie. La partie 1.3 présente comment différents acteurs de la société peuvent participer à cette gestion et à cette sélection de la diversité cultivée.

### **1.3 Participation des acteurs pour la sélection et la gestion des ressources génétiques**

Il y a deux aspects importants pour parvenir à développer des semences adaptées aux systèmes basés sur l'agroécologie : le lieu de la sélection (décentralisation ou non) et l'organisation (degré de participation des acteurs)<sup>68</sup>.

Dans cette partie, nous illustrons l'évolution de la place des paysans dans la sélection variétale et la gestion des ressources génétiques du néolithique jusqu'à aujourd'hui en prenant le cas de la France. Cette évolution est liée à l'organisation de la sélection et de la gestion des ressources génétiques.

#### **1.3.1 Les paysans sélectionneurs et gestionnaires des ressources génétiques depuis le néolithique**

L'agriculture est apparue il y a environ 12 000 ans à l'époque néolithique. La domestication du blé a débuté dans le croissant fertile au Moyen Orient, il y a environ 10 000 ans. Les croisements entre espèces sauvages et domestiquées, la sélection naturelle et la sélection artificielle des individus les plus adaptés, la dérive<sup>xx</sup>, ont conduit à une forte structuration de la diversité génétique des blés<sup>119</sup>.

Les principales caractéristiques de la domestication du blé sont l'augmentation de la grosseur et du nombre des grains, la perte de la caducité des grains (rachis solide) qui facilite la récolte et le passage de grains vêtus à des grains nus<sup>55</sup>. La domestication et la

---

xx. échantillonnage au hasard des gamètes lors de la reproduction, c'est l'effet d'échantillonnage qui provoque des goulots d'étranglement (voir note page 42).

sélection sont des processus dynamiques dans le temps et l'espace, une co-évolution lente entre l'homme, les animaux et les plantes<sup>196</sup>.

Les paysans ont été les principaux acteurs de ces processus de sélection. En effet, les échanges de semences, la gestion, la sélection, les migrations des hommes, notamment du Moyen Orient jusqu'en Europe, ont permis de créer une grande diversité de variétés de pays, adaptées à des contextes culturels et pédo-climatiques très différents. Ces variétés de pays sont définies par Camacho Villa et al.<sup>34</sup> comme des populations dynamiques qui ont une histoire, une identité distincte. Elles ont été améliorées par les paysans, génétiquement diversifiées, adaptées localement et associées à des systèmes de culture traditionnels.

### 1.3.2 L'institutionnalisation de la sélection et de la gestion des ressources génétiques et son impact sur la diversité cultivée

Le travail des paysans depuis 10 000 ans a offert à l'humanité des ressources extraordinaires. Ce travail a subi plusieurs ruptures, plusieurs « révolutions agricoles »<sup>190</sup>. L'une d'elle a eu lieu au milieu du XIX<sup>ème</sup> siècle, dans les pays du nord, où l'ère industrielle a induit l'uniformisation des procédés de fabrication qui a influencé jusqu'à l'agriculture. On passe d'une vision selon laquelle une espèce est évolutive à une vision où une espèce fixée et homogène peut trouver sa place dans des processus industriels standardisés<sup>22</sup>.

La sélection variétale va suivre l'évolution de l'agriculture industrialisée, basée sur de forts intrants. On parle plutôt d'une « co-évolution » entre la sélection des plantes et les procédés de culture<sup>190</sup> : les variétés répondent aux modes de culture et les modes de culture répondent aux variétés. C'est la société Vilmorin en France au début du XIX<sup>ème</sup> siècle qui amorce cette professionnalisation de la sélection. Elle sera suivie par d'autres sélectionneurs (tel Paul Bormans qui créa la variété de blé tendre *Ile-de-France* (Figure 1.9), qui contribueront à la création de variétés élites. La sélection cherche à minimiser l'influence de l'environnement et a pour objectif le « progrès génétique ». La redécouverte des lois de Mendel, associée à la naissance de la biométrie (travaux de Pearson sur la régression parents/descendants et de Fisher sur le modèle infinitésimal utilisé en génétique quantitative) accompagne cette professionnalisation de la sélection. La maîtrise de l'homogénéité génétique liée à un raisonnement scientifique rigoureux basé sur la biométrie est synonyme d'efficacité économique. Un génotype donné est assimilé à un comportement standardisé qui permet de répondre au marché et de vendre un produit stable<sup>22</sup>.

Les populations hétérogènes, gérées et cultivées *in-situ* par les paysans, sont devenues des ressources génétiques qui ont été collectées et stockées dans le cadre d'une gestion

*ex-situ* (partie 1.2.1). Les ressources génétiques vont constituer un réservoir pour créer ces variétés élités homogènes. Cavanagh et al.<sup>36</sup> illustrent parfaitement ce point dans leur étude sur 2 994 accessions de blé tendre représentatives de la diversité mondiale : 134 variétés de pays représentent près de 99% des allèles présents dans les 2 860 variétés modernes étudiées.

La France met en place, dès les années 40 une politique de pilotage de l'amélioration des plantes basée sur les critères Distinction, Homogénéité et Stabilité (DHS) en rendant obligatoire l'inscription des variétés au catalogue officiel des variétés. Cette politique est portée par Jean Bustarret. Aucune place n'est laissée au rôle des paysans dans la création et la gestion de la diversité cultivée<sup>25</sup>. Au niveau international, cette politique servira de base pour la création de l'Union Internationales pour la Protection des Obtentions Végétales (UPOV) (partie 1.4). Cette institution défend un système semencier basé sur la division des tâches, un mode de fonctionnement vertical, hiérarchisé. Cette homogénéisation des variétés est accompagnée d'une structuration de la filière semences appuyée par la création en 1942 du Comité Technique Permanent de la Sélection (CTPS), suivie en 1946 de la création de l'Institut National de la Recherche Agronomique (INRA) et d'autres institutions (Figure 1.10).

Cette structuration de la sélection et de la gestion des ressources génétiques engendre un bouleversement dans le rôle des différents acteurs. Le métier de sélectionneur s'officialise avec la création de l'INRA et d'entreprises privées stimulées par les opportunités qu'offre la propriété intellectuelle. Le droit national et le droit international évoluent pour répondre à cette réorganisation (partie 1.4). Les paysans ne sont plus alors que de simples usagers de l'innovation, des producteurs de grains qui ne sont plus intégrés dans le processus de sélection ni de conservation des ressources génétiques<sup>25</sup>.

Cette évolution de la place des acteurs a un impact sur la sélection et sur la gestion des ressources génétiques. Avec le catalogue et la professionnalisation de la sélection, la diversité intra-variétale des variétés cultivées diminue au cours du temps. Les variétés de pays, les plus hétérogènes, laissent leur place aux premières variétés issues de sélection beaucoup plus homogènes qui disparaissent pour laisser les variétés modernes, dont la variabilité intra-variétale est quasi nulle, envahir le paysage français<sup>117</sup>.

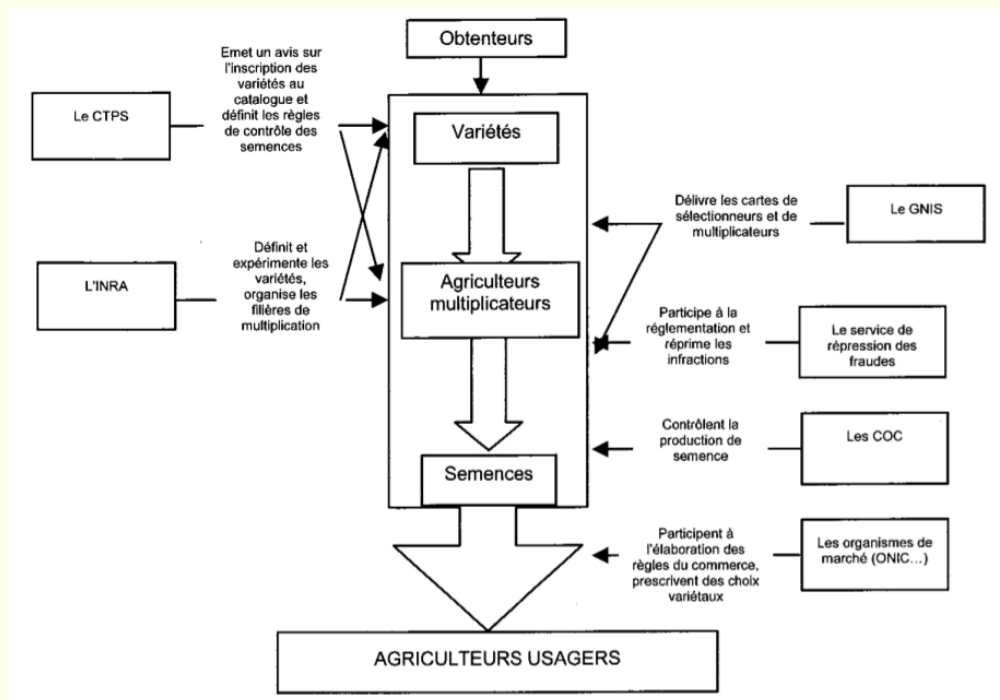
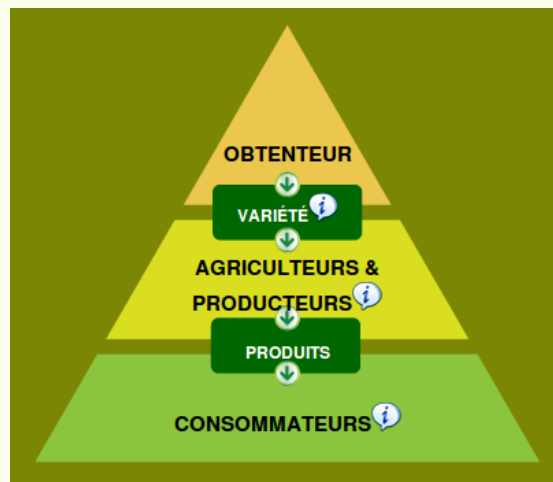


Figure 1.9 - Paul Bormans, sélectionneur à Méré dans les Yvelines en France, effectuant des croisements en 1927 à gauche. A droite, il prend des notes accompagné d'une collègue en 1949. Crédit : Sophie Richard, fille de P. Bormans.





Figure 1.10 - La filière des semences vue par l'UPOV<sup>243</sup> en haut et en France<sup>24</sup> en bas.



De nombreuses études montrent une perte de diversité entre les variétés populations et les variétés modernes<sup>36;199;207;208</sup>. Globalement dans le monde, la biodiversité cultivée a diminué : 75% de la diversité a été perdue entre 1900 et 2000<sup>88</sup>. Bonneuil et al.<sup>26</sup> proposent un indicateur de la biodiversité cultivée dans les territoires : *Ht\**. Il révèle, au cours du XX<sup>ème</sup> siècle, trois étapes dans l'évolution de la diversité cultivée<sup>117;119</sup>.

La première de 1912 aux années 70 : les variétés de pays disparaissent des territoires et sont remplacées par des lignées modernes homogènes génétiquement. Les variétés modernes, homogènes génétiquement, ne représentent pas la diversité des variétés de pays

(notamment à cause de l'introduction des gènes de nanisme). De plus, quelques variétés sont prépondérantes sur le territoire, cela engendre une diminution de moitié de  $Ht^*$ . Cette tendance se retrouve dans d'autres pays où les variétés modernes ont remplacé les variétés de pays locales. Par exemple, les variétés de riz au Sri Lanka sont passées de 2000 en 1959, à 100 dans les années 1990. Dans les années 1990, aux États-Unis, 50% des blés cultivés représentent seulement neuf variétés ; 75% des pommes de terre représentent seulement quatre variétés<sup>234</sup>.

La deuxième période couvre les années 70 et 80 où  $Ht^*$  augmente grâce à un effort de création des sélectionneurs (privés et publics) à partir de ressources génétiques plus diversifiées.

La troisième période débute à la fin des années 80,  $Ht^*$  stagne puis diminue dans les années 2000. Les cinq premières variétés de blé tendre cultivées en France représentaient 75% des surfaces en 1980 et entre 40 et 50% depuis les années 2000. Cependant, ces variétés les plus fréquemment cultivées deviennent de plus en plus proches génétiquement<sup>119</sup>.

### 1.3.3 La réappropriation de la sélection et de la gestion des ressources génétiques par la société civile en France

#### 1.3.3.1 L'émergence de la société civile sur la question des semences

Pour répondre à la diminution de la biodiversité cultivée et à l'uniformisation de la sélection variétale, la société civile s'organise. En France, sur la question des semences, la société civile est représentée par des paysans, des jardiniers, des consommateurs, des artisans semenciers et d'autres associés mobilisés autour de la conservation plus ou moins patrimoniale des variétés locales. Le terme « paysan » n'est pas anodin. Le paysan est issu d'un « pays », d'un terroir. Il incarne une tradition de partage de pratiques et de valeurs particulières. Ces paysans, acteurs de la société civile, remettent en question le système semencier actuel basé sur le productivisme et la propriété intellectuelle sur le vivant<sup>63</sup>. Ils souhaitent se réapproprier la sélection et la gestion des ressources génétiques par la mise en place de systèmes alternatifs de gestion collective. *In fine*, l'objectif est d'atteindre la souveraineté alimentaire dont les semences sont « *le premier maillon* »<sup>64</sup>. Ce mouvement défend l'intérêt général et non des intérêts individuels. C'est une initiative citoyenne qui se démarque aussi bien de l'action publique que de l'économie marchande<sup>xxi</sup>.

Ces acteurs de la société civile sont le plus souvent regroupés en associations. En 2003, est créé en France le Réseau Semences Paysannes (RSP) qui regroupe en 2013 plus de

---

xxi. L'économie marchande correspond aux échanges qui font intervenir de l'argent.

70 associations qui promeuvent la culture de la biodiversité. Par la coordination et la consolidation des initiatives locales, le RSP travaille à la promotion de modes de gestion collective et de modes de protection des semences paysannes. Le RSP s'efforce d'obtenir la caractérisation scientifique et la reconnaissance juridique des pratiques paysannes de production et d'échange de semences et de plants<sup>210</sup>. Le RSP accompagne ses adhérents dans la réappropriation des choix et des savoir-faire liés aux semences. Les paysans sont en attente de nouvelles variétés adaptées à leurs pratiques agroécologiques. En effet, les variétés inscrites au catalogue des variétés officielles sont rarement adoptées par les paysans car peu, ou pas, adaptées à leurs modes de culture. Un besoin de créer de nouvelles variétés émerge<sup>39;63;175</sup>. Le RSP illustre un mouvement plus global. par exemple, en 2009, Osman et Chable<sup>185</sup> dénombrent 68 initiatives différentes dans 17 pays Européens.

Dans la suite de ce chapitre, il ne sera question que des paysans. Les autres acteurs de la société civile, comme les consommateurs, les artisans semenciers et les jardiniers, bien qu'interagissant en permanence avec les paysans et entre eux, ne seront pas mentionnés.

### 1.3.3.2 La sélection décentralisée et participative

En plus d'être décentralisée, la sélection pour être efficace doit être participative<sup>43</sup>. La sélection décentralisée et participative (abrégée désormais dans ce texte « sélection participative ») fait partie de la recherche participative. Elle peut se définir comme une collaboration entre un partenaire scientifique académique, un laboratoire de recherche, un institut technique, et un partenaire de la société civile : les paysans, regroupés en associations par exemple. Les chercheurs apportent leur savoir théorique en génétique, statistique, agronomie ... les paysans apportent leurs connaissances concrètes sur l'environnement, les pratiques, les cultures dont ils ont l'expérience dans leur ferme et guident les recherches selon leurs besoins<sup>23;57</sup>.

La sélection participative engendre un partage des bénéfices<sup>211</sup> : cette collaboration étroite entre généticiens, agronomes, statisticiens, paysans, sociologues, ... produit des connaissances, des savoirs, à la fois pour le chercheur et pour le paysan. L'intégration des paysans dans un processus de sélection permet de réduire l'écart entre chercheur et bénéficiaire de la recherche et ainsi de favoriser l'autonomie du paysan<sup>171</sup>.

Les paysans sont acteurs du processus de recherche et interviennent dans les étapes majeures de la sélection : définition des objectifs, création de diversité, observations de cette diversité, sélection, diffusion. Les paysans deviennent des « *co-chercheurs* »<sup>xxii 218</sup>.

---

xxii. et les chercheurs des co-paysans?!?

La sélection par l'environnement combinée à la sélection par le paysan peut développer de nouvelles variétés. Ces variétés seront adaptées aux conditions pédo-climatiques, aux pratiques des paysans et au contexte socio-économique dans lequel elles s'insèrent<sup>192</sup>.

La sélection participative diffère de la sélection conventionnelle sur trois points principaux<sup>40;43</sup> :

1. L'évaluation de la valeur agronomique puisque la sélection se fait à la ferme plutôt que dans une station de recherche ;
2. les sélections se font dans des conditions contrastées afin de caractériser et d'utiliser au mieux les interactions  $G \times E$  ;
3. les critères de sélection sont choisis conjointement par le paysan et le chercheur, le sélectionneur dans certains cas.

Les points 1. et 2. sont liés à la décentralisation de la sélection (partie 1.2.2) et le point 3. à la participation.

Morris et Bellon<sup>171</sup> présentent cinq modèles de sélection où les étapes de la participation des paysans varient (Tableau 1.1). Ces modèles vont de la sélection menée uniquement par les paysans (modèle 1) à la sélection menée uniquement par les sélectionneurs (modèle 5). Entre ces deux extrêmes, on trouve des situations où la collaboration entre paysans et chercheurs a lieu à chaque étape du processus (modèle 2) ou à quelques étapes seulement (modèles 3 et 4). Le modèle 4 est connu sous le nom d' « évaluation variétale participative »<sup>xxiii</sup>. Évidemment, cette classification ne représente pas la totalité des situations rencontrées sur le terrain. Aucun programme de sélection participative ne ressemblera à un autre.

Le rôle des différents acteurs est très divers selon les projets et les contextes. Pour chacune des étapes, le degré de participation peut varier<sup>218</sup> :

1. Les chercheurs consultent les paysans pour faire la recherche ;
2. les tâches sont réparties entre chercheurs et paysans ;
3. les chercheurs accompagnent les paysans qui gèrent le programme de bout en bout.

Ces degrés de participation sont variables (on trouve plus d'informations sur ces points dans Morris et Bellon<sup>171</sup>).

Il existe de nombreux exemples de sélection participative dans la littérature scientifique (Tableau 1.2). La grande majorité des programmes ont lieu dans les pays du Sud, mais récemment des programmes se sont développés dans les pays du Nord. Les chercheurs

xxiii. Participatory varietal selection en Anglais

pionniers dans la sélection participative sont l'équipe de L. Sperling sur le haricot en Afrique de l'est, de J. Witcombe sur le riz au Népal ou encore de S. Ceccarelli au Moyen-orient sur la sélection de l'orge<sup>23</sup>.

Tableau 1.1 - Place des paysans (P) et des chercheurs (C) dans différents modèles de programme de sélection. Les chercheurs peuvent être des sélectionneurs. D'après<sup>171</sup>.

modèle	(1)	(2)	(3)	(4)	(5)
choix de la variabilité de départ	P	P & C	P & C	C	C
pré-sélection des caractères importants <sup>a</sup>	P	P & C	C	C	C
développement des variétés	P	P & C	C	C	C
évaluation des variétés	P	P & C	P & C	P & C	C

<sup>a</sup>. pre-breeding en Anglais

Au niveau des textes internationaux, la sélection participative est recommandée par le rapport de l'IAASTD<sup>133</sup>. L'article 6 du Traité International sur les Ressources Phyto-génétiques pour l'Alimentation et l'Agriculture (TIRPAA) demande des politiques agricoles encourageant l'utilisation de la diversité biologique (6.2.a.). Cet article recommande également que les agriculteurs participent à la sélection de variétés adaptées à leurs agro-systèmes (article 6.2.c), qu'ils participent à la gestion et la conservation des ressources génétiques (article 6.2.f), qu'ils aient accès à une large diversité génétique (article 6.2.d). Enfin des efforts sont demandés pour faire davantage de recherches sur la gestion de la diversité et sur la création de variétés adaptées aux principes écologiques (article 6.2.b.).

### 1.3.3.3 La gestion des ressources génétiques par les paysans

A travers le monde, les paysans cultivent et sélectionnent une multitude de variétés locales (modèle 1 du Tableau 1.1). Ils participent ainsi à la gestion des ressources génétiques et leur rôle est de plus en plus reconnu au niveau international ainsi que français (partie 1.2.1). En Europe, des paysans cultivent des variétés de pays qui évoluent et s'adaptent à leurs pratiques<sup>58;59;213;231</sup>.

Au delà de la création variétale, la sélection participative contribue également à cette gestion des ressources génétiques<sup>146;219;247</sup>. Augmenter le nombre de programmes de sé-

lection participative augmentera la diversité des espèces en apportant une diversité qui n'est pas présente dans les banques de semences *ex-situ*<sup>6:256</sup>.

\*\*\*

La partie 1.3 a présenté comment les différents acteurs peuvent participer à la gestion des ressources génétiques et à l'amélioration des plantes dans le contexte de l'agroécologie. La participation des acteurs de la société civile dans la gestion des ressources génétiques et la création variétale doit s'insérer dans un monde principalement dominé par les sélectionneurs professionnels. La partie 1.4 fait le point sur la gouvernance actuelle dans le monde des semences.

Tableau 1.2 - Quelques exemples de programmes de sélection participative sur différentes espèces dans le monde.

Tableau 1.2 – (partie 1/3)

Espèce	Pays	Résultats marquants	Références
Orge	Syrie	Les paysans ont une sélection aussi efficace ou même plus que les sélectionneurs dans leurs champs : la réponse à la sélection sur le rendement et la biomasse est supérieure ou égale. Les sélectionneurs sont plus efficaces dans les stations expérimentales que les paysans.	42;43
Orge	Jordanie	Dans leurs fermes, les paysans sélectionnent plus efficacement que les sélectionneurs.	110
Orge	Maroc, Syrie, Tunisie	Les critères de sélection diffèrent entre sélectionneurs et paysans. Les variétés sélectionnées à la station par les sélectionneurs ne sont pas les mêmes que celles sélectionnées par les paysans dans leurs fermes.	41
Haricot	Nicaragua	Deux variétés sont créées, par les paysans et les chercheurs, résistantes au virus à mosaïque du haricot.	5
Haricots	Rwanda	Les variétés sélectionnées par les paysans, notamment des femmes, ont souvent un meilleur rendement et sont plus souvent adoptées que celles sélectionnées par le sélectionneur.	219

Tableau 1.2 – (partie 2/3)

Espèce	Pays	Résultats marquants	Références
Mais	Chine	Développement de variétés hybrides par les paysans et les chercheurs.	157
Mais	Mexique, Honduras	La sélection des paysans valorise plus les environnements très éloignés des conditions de la station de recherche.	216
Mais	Portugal	Les paysans sélectionnent des épis gros et larges.	170
Blé	Népal	La sélection des paysans sur des critères qualitatifs se rapproche de la sélection des sélectionneurs sur des critères quantitatifs. L'évaluation sur plusieurs années est importante : les paysans ne sélectionnent pas les mêmes variétés suivant les années.	228
Choux	France	Nouvelles variétés créées par les paysans à la suite d'une collaboration avec les chercheurs.	45
Fève	Allemagne	Deux variétés sélectionnées par les paysans ont été adoptées.	115
Pomme de terre	Ouganda	Les paysans participant au projet de sélection participative adoptent 37 fois plus les variétés développées dans le cadre du projet que les autres paysans qui n'ont pas participé. Cette adoption passe à 6,7 pour les variétés issues d'évaluation variétale participative.	149



Tableau 1.2 – (partie 3/3)

Espèce	Pays	Résultats marquants	Références
Riz	Bangladesh, Népal	Deux variétés sélectionnées par les paysans ont été adoptées dans deux régions contrastées.	145
Riz	Inde	La sélection des paysans a développé 11 variétés qui ont été adoptées.	250
Sorgho	Burkina Faso	Les paysans ont une vision plus variée des caractères que les sélectionneurs, ils peuvent sélectionner des caractères spécifiques tout en poursuivant des objectifs agronomiques tels que l'adaptation au milieu.	251

## 1.4 Gouvernance des semences et des ressources génétiques

D'après Baron<sup>18</sup> ; « le terme de gouvernance vise la prise en compte des interactions entre les dynamiques transnationales, les logiques nationales et les initiatives locales, tout en considérant la porosité croissante des frontières entre les sphères du public et du privé. ». La gouvernance des semences regroupe ces différents niveaux d'interaction :

1. Les « dynamiques transnationales » sont représentées par l'Union Internationale pour la Protection des Obtentions Végétales (UPOV) (en lien avec l'Accord sur les Aspects des Droits de Propriété Intellectuelle qui touchent au Commerce (ADPIC) et l'Organisation Mondiale du Commerce (OMC)); le Traité International sur les Ressources Phytogénétiques pour l'Alimentation et l'Agriculture (TIRPAA), initié et géré par la FAO.
2. Les « logiques nationales » concernent, pour la commercialisation des semences, le catalogue d'inscription des variétés qui sont régies par les lois nationales et par là, les enjeux liés à chaque pays. Par exemple, dans le domaine des semences, la France a un chiffre d'affaires de 2,7 milliards d'Euros et est le 1<sup>er</sup> exportateur de semences dans le monde. Limagrain, entreprise française, est la 4<sup>ème</sup> multinationale mondiale avec 3.8% de part de marché mondial<sup>75</sup>. Pour le blé, la France est le cinquième exportateur mondial.
3. Les « initiatives locales » sont représentées par la société civile décrite dans la partie 1.3.
4. La « porosité croissante des frontières entre les sphères du public et du privé » est représentée par les partenariats de la recherche avec l'industrie (projet génoplane, investissements d'avenir) qui sont courants pour l'amélioration des plantes<sup>246</sup>. Mais elle recouvre également - moins fréquemment - des partenariats entre les institutions publiques et la société civile via des associations, par exemple les Partenariats Institutions-Citoyens pour la Recherche et l'Innovation (PICRI) en Ile de France.

Cette gouvernance peut être apolitique. Mais certains considèrent qu'elle redistribue les cartes du pouvoir dans la société<sup>18</sup>. Dans le cas des semences, les enjeux sont considérables<sup>75;211</sup> : Dans quel but le contrôle des semences est-il instauré ? Qui contrôle les semences ? Qui y a accès ? Comment ?

Celui qui contrôle les semences, contrôle ce qui est cultivé, et finalement l'accès à la nourriture. Les questions de propriété intellectuelle sur les semences sont au coeur de l'accès aux ressources.

Cette partie présente rapidement les différents niveaux qui interviennent dans la gouvernance des semences : d'abord le droit des semenciers et le système de propriété intellectuelle, ensuite le droit des paysans, puis l'accès aux ressources génétiques et le partage des bénéfices et enfin la commercialisation des semences, envisagée seulement dans le cas de l'Europe et la France.

#### 1.4.1 Le droit des semenciers et le système de propriété intellectuelle

Le besoin de mettre en place de la propriété intellectuelle sur les variétés est apparu avec le métier de sélectionneur, dans le but de rémunérer cette activité créatrice (partie 1.3). Un système réglementaire a été mis en place pour définir et encadrer les droits de propriété intellectuelle. Pour protéger une variété, il y a deux grandes approches défendues d'un côté par l'Union Européenne (UE) et de l'autre par les États-Unis : un système *sui generis* et le système des brevets. *sui generis* est une expression du droit latin, signifiant « de son propre genre ». Cette expression qualifie une situation de droit inédite dont la singularité prévient tout classement dans une catégorie déjà répertoriée et nécessite de créer des textes spécifiques<sup>172</sup>. Le vivant n'est pas fixé et est reproductible facilement, la protection nécessite donc des textes spécifiques. Le Tableau 1.3 présente les grandes dates de l'évolution de la propriété intellectuelle sur les semences.

##### 1.4.1.1 Le système des brevets

Les États-Unis sont les premiers à mettre en place un outil juridique sur la propriété intellectuelle sur le vivant avec l'adoption en 1930 du Plant Patent Act, littéralement l'acte sur le brevet des plantes. Il s'applique pour les plantes ornementales et les plantes à reproduction végétative, clonale. Le Patent Act, qui régit tous les brevets aux États-Unis, peut aussi être utilisé pour les inventions d'ordre biotechnologique. Comme pour tout brevet, pour avoir accès à la ressource brevetée, il faut en demander l'autorisation au détenteur du brevet qui réclame en général un paiement.

Dans l'UE, la directive 98/44 interdit les brevets sur les plantes ou les animaux. Cette directive réaffirme le droit des sélectionneurs à avoir accès librement aux ressources génétiques. En revanche l'UE autorise les brevets concernant les OGM en tant que technologie brevetée sur un gène et des processus non biologiques. Un brevet tombe dans le domaine public 20 ans après avoir été enregistré<sup>172</sup>.

Tableau 1.3 - Principaux événements concernant la mise en place des juridictions sur la propriété intellectuelle sur les semences. M : Monde, USA : États-Unis, UE : Union Européenne, FR : France

Tableau 1.3 – (partie 1/2)

Lieu	Date	Événements	Description
USA	1930	Adoption du Plant Patent Act	Il s'applique pour les plantes ornementales et les plantes à reproduction végétative (clonale).
M	1957	Conférence de Paris	Mise en place du système <i>sui generis</i> inspiré du système DHS du catalogue français des variétés officielles.
M	1961	Convention UPOV	
M	1968	Mise en place de la convention UPOV	création du Certificat d'Obtention Végétale (COV), basé sur les critères DHS. Pas de référence au droit des agriculteurs, il n'est question que du droit des sélectionneurs. Le « privilège » des paysans est implicite. Les paysans peuvent resemer leurs semences, et les échanger entre eux.
USA	1970	Plant Variety Protection Act	réglementation en harmonie avec l'UPOV. Les paysans peuvent garder, resemer les semences, ils peuvent également vendre leurs semences à d'autres paysans.
M	1991	Révision UPOV	Le sélectionneur est propriétaire de la semence récoltée. Les pays décident s'ils autorisent ou non, « avec des limites raisonnables et tant que les intérêts légitimes des sélectionneurs sont conservés », que les paysans puissent resemer leurs semences. Les échanges de semences entre paysans sont interdits.

Tableau 1.3 – (partie 2/2)

Lieu	Date	Événements	Description
M	1994	Signature de l'ADPIC	Pour adhérer à l'OMC, il faut reconnaître l'ADPIC.
USA	1994	Ratification de la convention de l'UPOV (1991 Act).	Pour vendre leurs semences, les paysans doivent demander l'autorisation aux semenciers. Les paysans peuvent toujours resemer leurs propres semences.
UE	1994	Application du règlement 2100/94.	Ce règlement rallie la convention UPOV de 1991. Les paysans peuvent resemer en échange d'une contribution financière.
M	1996	Mise en place de l'ADPIC	Les pays signataires doivent prévoir une protection des variétés végétales soit par brevets, soit par un système <i>sui generis</i> ou par une combinaison des deux (article 27.3(b)).
UE	1998	Directive 98/44	Interdiction des brevets sur les plantes ou les animaux. Le sélectionneur a un accès libre aux ressources génétiques.
M	1998	Les nouveaux membres de l'UPOV doivent adopter le 1991 UPOV Act.	
FR	2011	Loi pour rejoindre l'UPOV 91	La Fédération Nationale des Syndicats d'Exploitants Agricoles (FNSEA) défend la loi alors que la Confédération Paysanne et la Coordination Rurale la rejettent. Le décret n'est toujours pas appliqué le 11 mars 2014.

### 1.4.1.2 Le système *sui generis*

#### 1.4.1.2.1 Le système UPOV

C'est avec la convention de Paris en 1957 que les sélectionneurs commencent à se mobiliser pour trouver une alternative au brevet. Cette convention pose les bases d'un système qui sera ensuite géré par l'Union Internationale pour la Protection des Obtentions Végétales (UPOV), créée en 1961 à Paris. L'outil *sui generis* créé est le Certificat d'Obtention Végétale (COV). Il est largement inspiré du système DHS du catalogue français des variétés officielles. La convention UPOV de 1961 sera adoptée en 1968. Les trois points clés de la convention sont les suivants<sup>211;242</sup> :

1. Pour être protégée, c'est à dire obtenir un COV, la variété créée doit être nouvelle ou ne doit pas être notoirement connue. Elle doit être distincte, homogène et stable (critères DHS).
2. Chaque variété décrite doit avoir un nom unique.
3. Le COV fait la différence entre la variété comme une invention utilisée dans un but commercial et la variété comme une ressource libre d'accès.

Pour obtenir un COV, la variété doit être définie par des caractères précis qui, selon les espèces, correspondent à des variétés assez ou très homogènes (lignée pures et hybrides). Le COV permet de protéger la création variétale sans être un frein à l'innovation. En effet, les sélectionneurs peuvent utiliser gratuitement et sans autorisation une variété ayant un COV comme ressource pour leurs programmes de sélection.

La majorité des pays (En Europe, au Canada, en Amérique Latine, ...) basent leur protection variétale sur le système de l'UPOV<sup>211</sup>. Ces pays doivent mettre en place une législation adaptée pour être en accord avec l'UPOV. Par exemple, dans l'UE, c'est l'Office Communautaire des Variétés Végétales (OCVV) qui délivre les COV. Un COV est valable dans les seuls pays où il a été reconnu. Il est valable 25 ans pour toutes les espèces sauf les espèces pérennes pour lesquelles il est valable 30 ans. Les variétés appartiennent ensuite au domaine public<sup>172</sup>.

Dans la convention, il n'y a pas de référence aux droits des paysans, il n'est question que du droit des sélectionneurs. L'UPOV ne reconnaît pas le rôle des paysans comme innovateurs et détenteurs de pratiques et de savoirs essentiels pour le développement de biodiversité agricole. Pour l'UPOV, seuls les sélectionneurs peuvent innover<sup>211</sup>. Étant donné qu'aucun interdit n'est formulé, il est reconnu implicitement que les paysans peuvent ressemer leurs semences, et les échanger entre eux<sup>211</sup>. En 1978 une révision de l'UPOV garde la même direction que la convention adoptée en 1968. C'est la version de référence

à comparer à la révision de 1991.

En parallèle, les États-Unis publient le Plant Variety Protection Act (PVPA) en 1970 pour se mettre en harmonie avec l'UPOV. Les paysans peuvent garder, ressemer les semences, ils peuvent également vendre leurs semences à d'autres paysans. Cette loi ne s'applique qu'aux espèces qui se reproduisent sexuellement. Pour les autres, le recours au brevet est en vigueur.

Dans la révision de 1991 de la convention UPOV, le sélectionneur est propriétaire de la semence récoltée et même du produit final. Les pays décident s'ils autorisent ou non, « *avec des limites raisonnables et tant que les intérêts légitimes des sélectionneurs sont conservés* », les paysans à ressemer leur récolte. Les échanges sont interdits. C'est un tournant majeur pour le droit des paysans à gérer leur semence.

C'est dans cet esprit que la France a mis en place en 2001 la Cotisation Volontaire Obligatoire (CVO) sur le blé tendre pour rémunérer les sélectionneurs. Les paysans sont exonérés de cette taxe s'ils peuvent fournir une facture des semences ou que les variétés récoltées ne sont pas soumises à un droit de propriété intellectuelle, comme c'est le cas pour les variétés du domaine public. L'argent de la CVO est géré par le Fonds de Soutien à l'Obtention Végétale (FSOV) qui le redistribue par des appels d'offre à des recherches impliquant les entreprises de sélection.

Le sélectionneur est défini comme une personne qui sélectionne et développe une variété. En France, il doit s'inscrire auprès du Groupement National Interprofessionnel des Semences et plants (GNIS) pour pouvoir exercer. La sélection massale des paysans n'est pas reconnue ni dans la convention UPOV ni dans le droit français<sup>50</sup>.

En 1994, l'UE applique cette nouvelle version dans le règlement 2100/94. Elle va encore plus loin en interdisant de ressemer les semences de ferme pour la majorité des espèces et les espèces hybrides. Pour 21 espèces, le paysan peut ressemer en échange d'une contribution financière. Les petits paysans - définis comme « petit » si ils récoltent, dans le cas des céréales, moins de 92 tonnes - sont exonérés de cette mesure.

La France élabore des lois dès 1970 pour rentrer dans le cadre UPOV, en allant plus loin sur les critères DHS et les notions de nouveauté. La législation française s'harmonise avec le règlement européen 2100/94 à la suite du vote de la loi 2011-1843 du 8 décembre 2011 relative au COV<sup>50</sup>. La France ratifie l'acte de 1991 le 27 avril 2012.

#### 1.4.1.2.2 L'OMC et l'UPOV

L'ADPIC, signé en 1994 lors des accords de Marrakech, est devenu l'un des piliers de l'OMC. Pour faire partie de l'OMC, il faut adhérer à l'ADPIC, ce qui signifie :

1. qu'il faut breveter les inventions dans tous les champs des technologies (auparavant, certains pays comme le Brésil ont exclu des brevets les produits pharmaceutiques, alimentation et produits chimiques),
2. que les états doivent mettre en place des lois pour faire respecter les accords de l'ADPIC,
3. que tous les pays au sein de l'OMC sont traités de la même manière en matière de propriété intellectuelle.

Selon l'article 27.3(b) de l'ADPIC, les pays doivent obligatoirement prévoir une protection des variétés de plantes soit par brevet, soit par un système *sui generis* ou par une combinaison des deux.

Pour adopter un système *sui generis*, il n'est pas dit explicitement qu'il faut rentrer dans l'UPOV. Cependant beaucoup de pays ont décidé de suivre la convention UPOV. Les autres s'en étant inspiré. Néanmoins, suivre les conventions UPOV facilite l'acceptation par l'ADPIC. De plus, l'UPOV et l'Organisation Mondiale de la Propriété Intellectuelle (OMPI) offrent une assistance technique et administrative pour la mise en place des lois *sui generis*. Ce service est rare dans bon nombre de pays<sup>211</sup>.

#### 1.4.1.3 Différences et interactions entre le système *sui generis* et le système des brevets

Le COV et le brevet sont très différents. Avec le premier, l'accès aux ressources est libre et gratuit et avec le second, il est payant. Aux États-Unis, les systèmes de protection se chevauchent (Patent Act, PVPA et Plant Patents Act). C'est la jurisprudence qui permet de fixer les limites.

Rien n'empêche une double protection des variétés par COV et brevet. C'est le cas pour bon nombre de variétés qui ont utilisé un brevet dans le processus de sélection<sup>50</sup>. Cette double protection permet d'un côté d'être propriétaire des produits issus de la récolte grâce au COV et de l'autre côté de suivre à la trace les variétés pour pouvoir lutter contre la contrefaçon grâce au brevet.

### 1.4.2 Le droit des paysans

Dans la Convention sur la Convention sur la Diversité Biologique (CBD) signée à Rio en 1992, la question du partage des bénéfices issus des ressources génétiques est abordée,



mais à aucun moment celle du droit des paysans<sup>211</sup>. Le droit des paysans apparaît dans un texte : le Traité International sur les Ressources Phytogénétiques pour l'Alimentation et l'Agriculture (TIRPAA). Ce traité porte sur le partage des bénéfices et l'accès aux ressources génétiques, mais aussi sur le droit des paysans (partie 1.4.3.2).

L'introduction du TIRPAA reconnaît les droits :

« [...] d'utiliser, d'échanger et de vendre des semences de ferme [...] et de participer à la prise de décision concernant l'utilisation des ressources phytogénétiques pour l'alimentation et l'agriculture ainsi qu'au partage juste et équitable des avantages en découlant sont un élément fondamental de la concrétisation des droits des agriculteurs [...] »

L'article 6 du TIRPAA demande un système juridique adapté (6.1. et 6.2.g). Cependant, les pays restent libres d'appliquer le traité selon leurs propres contraintes à travers leurs lois nationales. Les articles 9.2., 9.3. et 7.1. vont également dans ce sens. On pense aux contraintes liées à l'UPOV ou aux brevets qui ne sont pas compatibles avec les objectifs du traité. Il est mentionné que les droits de propriété intellectuelle sur les ressources génétiques s'appliquent en conformité aux droits internationaux et nationaux (article 12.3.f).

Les organisations paysannes s'insurgent contre le système UPOV et le système des brevets. Lors du sommet mondial de l'alimentation à Rome à la FAO en 1996, la Déclaration des ONG indique que « *les ressources génétiques sont essentielles pour la sécurité alimentaire et ne peuvent faire l'objet des droits de propriété intellectuelle.* ».

### 1.4.3 L'accès aux ressources génétiques et le partage des bénéfices

En 1989, pour la première fois avec l'International Undertaking on Plant and Genetic Resource (IUPGR), la FAO reconnaît le droit des paysans au même titre que le droit des sélectionneurs. Ce traité qui considère alors les ressources génétiques comme un « *héritage commun de l'espèce humaine* »<sup>211</sup>, a été le point de départ d'autres traités concernant les ressources génétiques et les droits des paysans.

#### 1.4.3.1 Convention sur la Diversité Biologique

La Convention sur la Diversité Biologique (CBD) à Rio en 1992 a été signée par 157 pays. En 2012, ils sont 193. Le traité a trois objectifs :

1. La conservation de la biodiversité.

2. L'utilisation durable de la biodiversité.
3. Le partage équitable et juste, élaboré par discussion entre les parties et par accords bilatéraux, des bénéfices de l'utilisation des ressources génétiques.

Le traité reconnaît les états souverains sur leurs ressources génétiques : ce sont les lois nationales qui régulent l'accès aux ressources génétiques. Ce qui contredit l'idée que les ressources génétiques sont un « *héritage commun de l'espèce humaine* ». Pour certains auteurs, la CBD a permis de légitimer la propriété intellectuelle sur les ressources génétiques<sup>22;25</sup>. En effet, selon eux, pour faire un partage juste des ressources, il faut leur attribuer une valeur. Les états sont détenteurs des ressources s'ils sont considérés comme « *centre d'origine* » de la ressource ou « *centre de la diversité* » de cette ressource. Mais il n'y a pas de définition claire de ces termes<sup>211</sup>.

Les échanges de ressources se font par accords bilatéraux entre les deux pays qui doivent signer un Access and Benefits Sharing (ABS). L'ABS permet aux parties de se mettre d'accord sur le partage des bénéfices. Les ABS sont parfois difficiles à mettre en place, ce qui n'a pas permis, aux personnes ou institutions intéressées, d'avoir un accès total aux ressources. En effet, il est important de définir à qui appartiennent les ressources ce qui est très compliqué quand la ressource est gérée par un réseau d'acteurs et de communautés. Le traité ne prend pas en compte les processus biologiques et sociologiques qui construisent l'agrobiodiversité et les savoirs associés<sup>211</sup>.

#### 1.4.3.2 Traité International sur les Ressources Phytogénétiques pour l'Alimentation et l'Agriculture

La résolution 3 de la CBD approuvée à Nairobi en 1992 reconnaît le besoin de trouver des solutions pour les ressources génétiques concernant l'alimentation et l'agriculture, et plus particulièrement l'accès aux ressources *ex-situ* et le droit des paysans. En 1993, la FAO prend le dossier en main pour créer un traité issu du IUPGR qui puisse être en harmonie avec la CBD. C'est à la suite d'un consensus entre les pays, qu'est signé le Traité International sur les Ressources Phytogénétiques pour l'Alimentation et l'Agriculture (TIRPAA) en 2001. Il rentre en vigueur en 2004.

Alors que la CBD a un système d'échange bilatéral, le TIRPAA installe un système d'échange multilatéral. Il est censé permettre l'échange des ressources et aussi les avantages liés à ces ressources, de manière rapide, juste, transparente et équitable. Les échanges se font entre les pays et relèvent des législations nationales. Les ressources génétiques dans le système multilatéral ne sont pas soumises à un droit de propriété intellectuelle et sont

accessibles à tous<sup>172</sup>. De même, le domaine public regroupe les variétés qui ne relèvent pas du droit de propriété intellectuelle. C'est à dire les variétés notoirement connues et qui n'ont pas eu de COV ou alors des variétés qui ont perdu leur titre de COV<sup>172</sup>.

Comme pour la CBD, les états sont souverains sur leurs ressources et décident de ce qu'ils veulent mettre dans le système multilatéral (article 10). Les 64 espèces qui rentrent dans ce système multilatéral sont mentionnées dans l'annexe 1 du traité : elles couvrent 80% de la nourriture mondiale<sup>211</sup>. Les accessions de l'annexe 1 concernent plus de 1 240 banques de semences *ex-situ* parmi lesquelles 51% sont conservées dans 800 banques de graines qui ont signé le TIRPAA et 13% dans les centres internationaux du CGIAR (article 11 et FAO<sup>94</sup>). Pour les espèces qui ne font pas partie de l'annexe 1, c'est sous l'égide de la CBD et des accords bilatéraux que se font les échanges.

L'accès aux ressources au sein du système multilatéral se fait par un Accord de Transfert de Matériel (ATM). Les informations liées aux accessions sont dans le passeport qui les accompagne et sont accessibles si les informations ne sont pas confidentielles et sous réserve des législations nationales (article 12.3.c). Cependant, ce qui est considéré comme confidentiel ou non n'est pas clairement défini<sup>211</sup>. Par ailleurs, les données détenues par les centres de ressources génétiques des pays signataires du traité ne sont pas facilement accessibles<sup>94;194;232</sup>. Enfin, 45% de ces accessions sont conservées dans seulement sept pays, ce qui pointe la nécessité d'un système d'échange performant<sup>94</sup>. En 2008, ces ressources ont été majoritairement utilisées par les organisations publiques (28% du Sud et 25% du Nord), suivies par les centres de ressources génétiques (31%) et le secteur privé (16%).

Les bénéficiaires du système multilatéral ne peuvent pas revendiquer un droit de propriété intellectuelle sur les ressources (article 12.3.c). Cependant, l'UE, qui a rejoint le traité en 2004, énonce que toute innovation à partir des ressources du système multilatéral pourra être soumise à un droit de propriété intellectuelle. En effet, un COV pourra être déposé dès lors que la variété est nouvelle et n'est pas notoirement connue ; même si elle a été créée à partir de ressources génétiques du système multilatéral. Le COV rend la ressource disponible sans restrictions pour de futures innovations ; ceci en accord avec l'article 13 qui demande aux parties de partager les avantages du système multilatéral de façon juste et équitable. Cependant, toute restriction pour un paysan de ressemer sa récolte, demandé par la convention UPOV de 1991, est une entrave au partage des bénéfices établi par le TIRPAA<sup>50</sup>.

Si le produit final issu des ressources n'est pas « *disponible sans restriction* » à la

recherche ou à la sélection, alors le détenteur du produit doit payer une redevance à un fonds dédié de partage des bénéfices en pourcentage des ventes du produit. Si le produit issu des ressources est accessible, une participation au fonds est volontaire, mais pas obligatoire (article 13.2.d.ii.). En d'autres termes, avec un brevet sur une variété, le sélectionneur doit payer une redevance au fonds du TIRPAA. Les pays qui utilisent les brevets n'ont pas ratifié le TIRPAA sauf l'Australie<sup>211</sup>. L'impact de ce fonds de partage des bénéfices est donc limité<sup>25</sup>. De plus, le terme « *disponible sans restriction* » n'est pas clair : comment l'interpréter à propos des variétés protégées via l'UPOV ou des variétés qui ne sont pas réutilisables pour la sélection comme les hybrides ou les plantes CMS<sup>xxiv</sup> ? Une meilleure formulation permettrait d'être plus clair sur le partage des bénéfices<sup>211</sup>.

Bien que les parlementaires français aient approuvé à l'unanimité le Traité International sur les Ressources Phytogénétiques pour l'Alimentation et l'Agriculture (TIRPAA) en 2005, celui-ci n'a toujours pas été traduit dans la législation nationale.

#### 1.4.4 La commercialisation des semences en Europe et en France

Les états étant souverains, ce sont les législations nationales, c'est à dire les lois de commercialisation des semences, qui permettent de trouver un équilibre entre droits des sélectionneurs (UPOV) et droits des paysans (TIRPAA). Comme nous l'avons vu précédemment, ces deux textes sont parfois en conflit.

La commercialisation des semences est un point clé, car c'est elle, concrètement, qui permet de définir au sein d'un état les conditions d'accès au marché. La commercialisation des semences se fait à travers les catalogues officiels des variétés. Ces catalogues sont, en premier lieu, pensés pour les semences utilisées par l'agriculture conventionnelle<sup>211</sup>, c'est à dire les semences qui sont DHS, avec des droits de propriété intellectuelle et qui correspondent à une certaine Valeur Agronomique et Technologique (VAT).

##### 1.4.4.1 Le catalogue officiel européen

Chaque pays de l'UE doit posséder un catalogue national. Le catalogue officiel européen est ensuite la somme de tous les catalogues nationaux. Une variété inscrite au catalogue français sera automatiquement sur le catalogue européen et donc accessible à un paysan allemand par exemple ; même si cette variété n'est pas sur le catalogue allemand. En 2002, les directives de l'UE définissent des bases communes pour tous les catalogues nationaux ; notamment les critères DHS et VAT. Les critères sont ajustés par

---

xxiv. Cytoplasmic Male Sterility : technique permettant de contrôler les croisements en rendant le mâle stérile

les états et sont plus ou moins stricts. Toutes ces directives sont actuellement en cours de discussion au niveau de la commission européenne dans le plan « better regulation » qui vise à réformer le système en place.

#### 1.4.4.2 Le catalogue de conservation

La directive 2008/62 de l'UE souhaite assurer la conservation *in-situ* et l'utilisation durable des ressources génétiques et ainsi prévenir l'érosion génétique. Elle met pour cela en place le catalogue des variétés de conservation.

Une variété de pays est définie par : « *un jeu de populations ou de clones d'une espèce végétale qui est naturellement adaptée aux conditions environnementales de sa région.* ». La définition de ce catalogue ne prend pas en compte le rôle des paysans dans la gestion de ces variétés. Ce catalogue de « *variété de conservation* » ou « *sans valeur intrinsèque* » spécifie notamment que la variété doit être à 90% uniforme, les variétés ne peuvent être produites que dans leur région d'origine, la quantité de grains est limitée.

Ces directives européennes doivent être transposées en droit national pour être mises en oeuvre dans chaque pays. Les pays ont toujours la possibilité d'adapter la loi selon leurs propres contextes locaux. Ce catalogue fait avancer le droit de cultiver des variétés non DHS et non protégées par un COV. Cependant, à cause de ses critères trop stricts, peu de variétés de conservation ont été inscrites<sup>50</sup>.

#### 1.4.4.3 Le catalogue officiel français

En France, le décret 81-605 définit les semences et les plants comme « *des végétaux ou parties de végétaux de toute nature destinés à la production ou à la multiplication et concerne la production en vue de la commercialisation.* ». Ce même décret mentionne que les variétés doivent être DHS et VAT pour être inscrites au catalogue<sup>xxv</sup>. Cette inscription permet la vente et les échanges « *en vue d'une exploitation commerciale* ». La VAT devient Valeur Agronomique, Technologique et Environnementale (VATE) en 2009, le « E » signifiant « *environnementale* » et est censée prendre mieux en compte les interactions génotype × environnement<sup>159</sup>.

Si les semences sont vendues ou échangées dans le cadre d'une exploitation non commerciale, les variétés ne doivent pas forcément être inscrites. Dans ce cas, les échanges doivent se faire en petites quantités (article 1-3 du décret); mais la quantité n'est pas définie clairement<sup>172</sup>. Un cadre d'exploitation non commerciale correspond à l'autoconsommation, la conservation, la recherche et la sélection<sup>172</sup>. Un paysan qui reçoit une petite

---

xxv. Dans le cas des céréales. La VAT n'est pas appliquée aux espèces fruitières et potagères.

quantité de semences pour les multiplier et qui ensuite vend les produits issus de ces semences est dans la légalité; cette multiplication est faite dans le cadre d'un travail de recherche et de sélection.

Le décret 81-605 définit également la notion de club : « *la fourniture de semences pour une prestation de service en vue de la transformation ou du conditionnement est exclue de la commercialisation.* ». On retrouve ces clubs avec le pain Jacquet de Limagrain ou dans des associations comme Koal Kozh<sup>172</sup>. Comme il n'y a pas de transfert de propriété au sein du club, il n'y a pas de commercialisation des semences, donc pas d'acte de commercialisation. Il existe enfin des dérogations pour les semences de mélange fourrager qui ne rentrent pas dans les standards DHS du catalogue<sup>172</sup>.

Tout cela serait remis en cause avec la loi 2011-1843 relative aux COV qui supprime la restriction « *en vue d'une exploitation non commerciale* »<sup>50</sup>. C'est une menace directe sur les activités énumérées précédemment. De plus, tous les producteurs de semences devraient déclarer leurs activités, afin que les semenciers puissent mettre en place un système type CVO sur certaines espèces<sup>50</sup>. En février 2014, de nombreux décrets de cette loi n'ont toujours pas été adoptés.

\*\*\*

Les parties 1.1, 1.2, 1.3 et 1.4 ont montré l'importance de l'agroécologie pour atteindre la souveraineté alimentaire. Au sein de l'agroécologie, la diversité génétique, la décentralisation et la participation des acteurs dans la gestion des ressources génétiques et l'amélioration variétale sont centrales. Au delà des agrosystèmes, l'environnement socio-économique et juridique sont importants à prendre en compte. La partie 1.5 présente le contexte et le projet de sélection décentralisée et participative pour des systèmes agroécologiques en France, dans lequel mon travail de thèse s'est inséré.

## **1.5 Contexte et objectifs du projet de sélection décentralisée et participative sur le blé tendre en France**

### **1.5.1 Le contexte français**

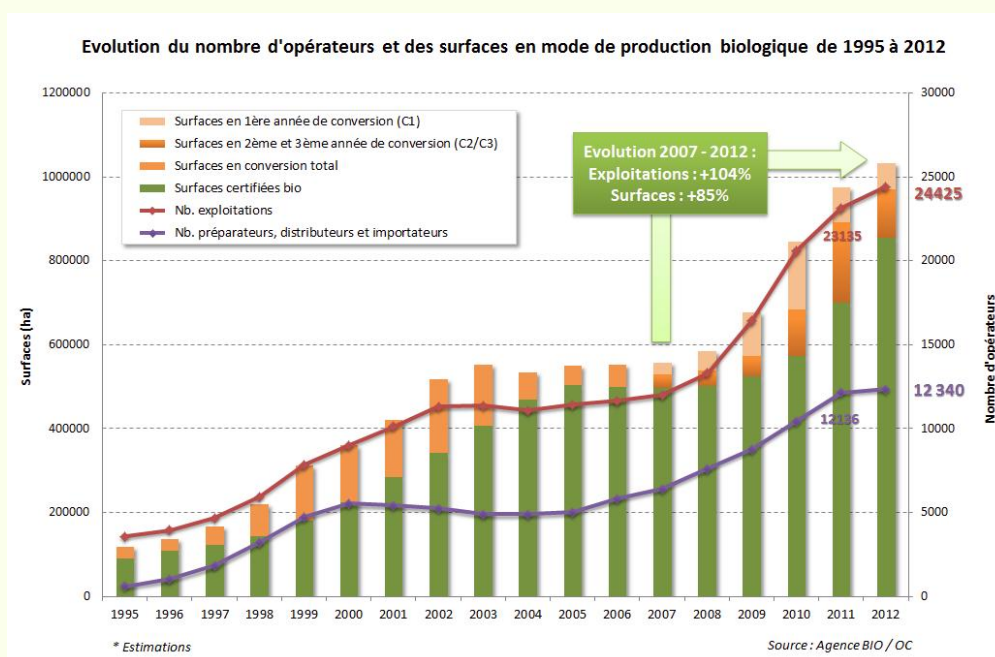
La filière semence en France a un poids économique très important, est très organisée et est gérée par plusieurs institutions qui contrôlent la mise sur le marché des semences (partie 1.4).

Dans ce contexte, la France a débuté en 2010 le projet BreedWheat. Il dure neuf ans et a pour but de créer de nouvelles variétés de blés pour l'agriculture de demain. Ce projet est une collaboration entre partenaires publics et privés. Il représente un investissement total de 34 millions d'euros, dont une contribution publique de 9 millions dans le cadre du programme Investissement d'avenir<sup>138</sup>. De même, pour le maïs, le projet Amaizing, collaboration entre partenaires publics et privés, a débuté en 2011. Il est prévu pour une durée de huit ans et représente un investissement de 27,5 millions d'euros, dont 9 financés par le programme Investissement d'avenir<sup>137</sup>.

Dans ces deux projets, aucun fonds n'est prévu pour la recherche variétale en agroécologie. Néanmoins, la nouvelle loi d'avenir agricole fait la part belle à l'agroécologie « *au coeur de pratiques innovantes* »<sup>154</sup>. L'appel d'offre du CASDAR 2013 « *mobilisation collective pour l'agro-écologie* » et la labellisation de l'Institut Technique de l'Agriculture Biologique (ITAB) comme institut technique agréé en 2012 illustrent cette dynamique.

D'après l'Agence Bio, en 2012, l'agriculture biologique en France représente 4,7% des exploitations agricoles et 3,7% de la surface agricole. Ces deux chiffres sont en constante augmentation depuis les années 90 (Figure 1.11).

Figure 1.11 - Evolution du nombre d'opérateurs et des surfaces en mode de production biologique de 1995 à 2012.



D'après l'Agence Bio.

En 2011, le centre d'études et de prospective du Ministère de l'agriculture identifie huit priorités de recherche-développement. Il plaide entre autres pour un effort d'innovation sur le plan de la génétique afin de créer des variétés et des races adaptées à l'agriculture biologique<sup>164</sup>. Les variétés modernes ne sont pas adaptées aux conditions de l'agriculture biologique et il y a peu de variétés disponibles<sup>56</sup>. Cependant, une dynamique est en train de se créer. En effet, l'ITAB évalue depuis 2004, dans un réseau d'essais, des variétés de blé tendre en conditions biologiques<sup>141</sup>. A la suite d'une évaluation en 2010 et 2011 à travers ce réseau d'essais, deux nouvelles variétés de blé tendre de l'INRA de Rennes (équipe de B. Rolland) ont été inscrites au catalogue : *Hendrix* et *Skerzzo*<sup>206</sup>. L'avenir dira si ces variétés supplanteront *Renan* et *Pirénéo* qui restent les variétés les plus cultivées en France avec respectivement 16 et 10% des surfaces cultivées en bio<sup>105</sup>.

Il existe une grande diversité de courants au sein de l'agriculture biologique en France (Tableau 1.4).

« Nature & Progrès » et « Demeter » sont deux labels historiques issus des premiers regroupements de producteurs et de consommateurs soucieux de la qualité et la production de leur nourriture. En 1985, l'agriculture biologique s'est institutionnalisée avec l'arrivée du label « AB » soumis à une réglementation moins stricte que Nature & Progrès. Le label AB s'est ensuite aligné sur le label européen, qui est encore moins strict. Ce dernier est adapté à une agriculture biologique « industrielle » loin des valeurs sociales et solidaires défendues par Nature & Progrès. Pour répondre à cette évolution du cahier des charges, un bon nombre d'agriculteurs s'organisent pour créer la marque collective « Biocoherence » qui se rapproche des valeurs du label AB.

L'agriculture biologique et l'agroécologie ne sont pas des notions avec des limites rigides et fixées. Il existe autant d'agriculture biologique et d'agroécologie que d'agrosystèmes, d'environnements socio-économiques et d'acteurs. Le développement de semences adaptées à l'agriculture biologique doit prendre en compte ces différentes facettes.

Les acteurs de la société civile dans les semences (partie 1.3) sont proches des valeurs de Nature & Progrès et de Demeter, voire Biocoherence pour certains. Ce sont les valeurs de l'agriculture paysanne. L'agriculture paysanne, défendue par la Confédération Paysanne<sup>xxvi</sup>, se définit comme « *une agriculture productrice de biens marchands et non marchands dans l'intérêt de la société* ». Elle fait apparaître six thèmes : la répartition,

---



xxvi. La Confédération Paysanne est un syndicat agricole « pour une agriculture paysanne et la défense de ses travailleurs ».



le travail avec la nature, l'autonomie, la qualité des produits, le développement local et la transmissibilité<sup>63;84</sup>.

Parmi ces acteurs de la société civile, certains paysans du Réseau Semences Paysannes (RSP) souhaitent développer des variétés « paysannes » adaptées à leurs pratiques. Jean François Berthelot (*JFB*), paysan-boulangier près d'Agen, a réalisé dans sa ferme avec l'aide de l'équipe de recherche Diversité Évolution et Adaptation des Populations (DEAP), dirigée par Isabelle Goldringer de l'INRA du Moulon, 90 croisements en 2005 pour créer de nouvelles populations de blé tendre. *JFB* a choisi les parents qui sont des variétés de pays, des variétés anciennes et quelques variétés modernes de Suisse et d'Allemagne créées pour l'agriculture biologique. L'objectif de *JFB* était d'associer les qualités agronomiques, résistance à la verse, concurrence vis à vis des adventices, aux qualités gustatives et nutritionnelles du pain. Ces nouvelles populations hétérogènes ont été multipliées dans sa ferme entre 2006 et 2008. En 2008, ces populations ont été distribuées au sein du RSP. 20 à 25 paysans, répartis dans toute la France, participent au projet de sélection participative depuis 2008. Cette thèse, démarrée en octobre 2010, s'inscrit dans la continuité de ce travail.

Tableau 1.4 - Les différents mouvements de l'agriculture biologique en France.

Logo	Description
	Le label européen est en vigueur depuis le 1 <sup>er</sup> juillet 2010, il est moins strict que l'ancienne législation française notamment sur l'élevage et les OGM. Des traces d'OGM sont admises en dessous de 0.9%, les animaux ne doivent pas forcément manger des produits de la ferme, la mixité bio/non bio sur une ferme est autorisée. Ce label est adapté à l'agriculture biologique « industrielle ». Le contrôle est fait par des organismes extérieurs type Ecocert ou Qualité France.
	Créé en 1985, le logo AB est la propriété du ministère de l'agriculture. Il a perdu son sens depuis la création du label européen. Il est aujourd'hui synonyme du logo européen.
	Le label Biocohérence est plus exigeant que le label européen. Il revient aux critères du label AB avant son changement pour le modèle européen. Notamment interdiction de la coexistence bio/non-bio, production des aliments pour les animaux en majorité sur la ferme, interdiction totale de trace d'OGM, nécessité de 100% d'ingrédients bio dans un produit transformé, restriction des traitements sanitaires sur les animaux. Le contrôle est fait par des organismes extérieurs type Ecocert ou Qualité France.
	Le label Nature & Progrès va plus loin que le label Biocohérence sur certains points, notamment la préservation de l'environnement extérieur à la ferme afin de protéger les ressources collectives. Le contrôle est fait par les adhérents (contrôle participatif).
	Le label Demeter est spécifique aux fermes pratiquant l'agriculture biodynamique. Le label s'appuie sur le label européen mais apporte d'autres principes basés sur les énergies et la philosophie de Rudolf Steiner. Le contrôle est fait par des organismes extérieurs type Ecocert ou Qualité France.

D'après INC<sup>135</sup>.

### 1.5.2 Objectifs du travail de thèse

Ce travail a été effectué dans l'équipe DEAP au sein du laboratoire de Génétique Végétale de l'INRA du Moulon. Le projet Européen Strategies for Organic and low-input Integrated Breeding Management (SOLIBAM) a financé le fonctionnement du projet. La région Ile de France, à travers le Domaines d'Intérêt Majeur, Agrosociété, Territoires, Ecologie, Alimentation (DIM ASTREA), a financé ma bourse de thèse pendant trois ans.

L'objectif de cette recherche est, d'une part, de mettre au point une méthodologie de la sélection décentralisée et participative en agroécologie, dans le contexte français, qui permette de créer des variétés adaptées à la diversité des agrosystèmes, la diversité des contextes socio-économiques et la diversité des acteurs. D'autre part, elle vise à mettre en place des outils adaptés pour la réalisation d'un programme de sélection participative dans le cadre d'une recherche-action. L'action consiste en la co-construction du programme de sélection entre les associations membres du RSP, les paysans et les chercheurs. Les objectifs du projet de sélection participative sont divisés en trois points :

1. la création des variétés-populations adaptées aux conditions locales, aux pratiques des paysans,
2. le développement de méthodes et d'outils opérationnels pour la gestion de la biodiversité cultivée et la sélection à la ferme,
3. le renforcement de l'apprentissage et de l'autonomie des paysans en matière de gestion et de sélection des semences.

Ma thèse se concentre sur les deux derniers objectifs. Pour aborder ces points, la présentation de mon travail de thèse est divisée en cinq chapitres, le chapitre 1 étant consacré à l'introduction.

Le chapitre 2 analyse le programme de sélection participative. Il traite des dispositifs expérimentaux mis en place, de l'analyse statistique des données dans les fermes (comparaisons de moyennes, sélections paysannes) et de l'analyse statistique des interactions  $G \times E$  dans le réseau de ferme.

Le chapitre 3 évalue l'évolution phénotypique et moléculaire au sein du programme de sélection participative et étudie sa contribution à la gestion des ressources génétiques.

Le chapitre 4 propose une description et une analyse de la méthodologie mise en place pour répondre aux attentes des différents acteurs.

Le chapitre 5 discute les différents résultats obtenus et analyse comment notre projet pourrait participer à la transition vers la création d'un système semenciers adapté à

l'agroécologie.

## 2 . Analyse phénotypique du programme de sélection participative

Dispositifs expérimentaux dans les fermes

Analyse phénotypique dans les fermes

Analyse des interactions  $G \times E$

Ce chapitre analyse les données phénotypiques issues du programme de sélection participative.

Dans le projet de sélection participative, les paysans sont libres de semer les populations qu'ils souhaitent dans leur ferme. L'unique contrainte est de semer, au moins, un témoin répété dans un ou plusieurs blocs.

Cette flexibilité engendre des dispositifs expérimentaux très déséquilibrés : dans chaque ferme, il y a peu de degrés de liberté pour la résiduelle<sup>i</sup> et il y a de nombreuses combinaisons (ferme  $\times$  population) manquantes sur le réseau. Pour pallier ce déséquilibre, deux modèles statistiques bayésiens hiérarchiques ont été développés. Ils permettent de comparer les populations dans une ferme et d'étudier les interactions population  $\times$  environnement.

Une première partie décrit les fermes, les dispositifs expérimentaux et les mesures phénotypiques réalisées. Une deuxième partie traite de l'analyse phénotypique dans les fermes. La comparaison des populations, la réponse à la sélection inter- et intra-population sont étudiées. Une troisième partie porte sur l'analyse des interactions population  $\times$  environnement ( $G \times E$ ), en présentant dans un premier temps la démarche, puis le modèle finalement retenu et les résultats.

---

i. Cela revient à dire qu'il y a peu de répétitions dans les fermes

## 2.1 Dispositifs expérimentaux dans les fermes

Les dispositifs expérimentaux utilisés dans les analyses statistiques de cette partie sont présentés ici. La saison 2012-2013 n'a pas été prise en compte. Le chapitre 4 présente l'élaboration et l'évolution des dispositifs expérimentaux et des mesures dans le processus de co-construction du projet.

### 2.1.1 Fermes participant au projet

Toutes les fermes participant au projet pratiquent l'Agriculture Biologique. Elles sont réparties dans toute la France et sont adhérentes au RSP à travers une association locale ou à titre personnel (Figure 2.1 et Tableau 2.1).

Figure 2.1 - Répartition des fermes participant au programme pour la saison 2011-2012.



En rouge : fermes régionales.  
 En bleu : fermes satellites.  
 En jaune : associations du RSP.  
 En vert : l'INRA du Moulon.



Tableau 2.1 - Caractéristiques des fermes participant au projet de sélection participative.

Table 2.1 – (partie 1/3)

ferme	commune	longitude	latitude	altitude	associations	mm	temp	année de participation au projet
ALP	44380	47.67	-2.24	11	Triptolème	NA	NA	2010 2011
ANB	05260	45.03	6.26	1171	ARDEAR R-A	61.4 +/- 38.9	11.0 +/- 7.2	2011 2012 2013
BEN								2013
BER	21260	47.56	5.27	297	Graines de Noé	66.9 +/- 35.0	11.1 +/- 6.7	2009 2010 2011 2012
BRE	21110	47.37	5.11	218	Graines de Noé	66.9 +/- 35.0	11.1 +/- 6.7	2010 2011
CHD	38690	45.83	5.40	487	ARDEAR R-A	70.0 +/- 41.7	13.1 +/- 7.2	2009 2010 2011 2012 2013
CHP	47600	44.21	0.37	72	CETAB	51.4 +/- 33.5	13.6 +/- 6.2	2009 2012 2013
EUK	21220	47.18	5.03	221	Graines de Noé	66.9 +/- 35.0	11.1 +/- 6.7	2012
FLM	49080	47.71	-0.60	24	Triptolème	55.2 +/- 33.8	12.8 +/- 6.3	2009 2010 2011 2012 2013
FLP	81170	44.21	1.92	269	CETAB	53.1 +/- 32.3	14.3 +/- 6.5	2009 2010 2011 2012 2013
FRC	30320	44.2	4.50	68	Touselles	58.8 +/- 47.6	15.2 +/- 6.6	2009 2010 2011 2012 2013

temp : température mensuelle moyenne pour les années étudiées +/- écart type; mm : précipitation mensuelle moyenne pour les années étudiées +/- écart type. NA : données non disponibles L'annexe A détaille sur le climat des différentes fermes.

L'année de participation correspond à l'année de récolte.

RSP : adhérent au RSP à titre individuel; ARDEAR R-A : Association pour le Développement de l'Emploi Agricole et Rural en Rhône-Alpes; CETAB : Centre d'Etude et Terre d'Accueil des Blés.

ARDEAR R-A, CETAB, Triptolème, Kerna ün Shoma et Graines de Noé ont un animateur, les autres associations non.

Table 2.1 – (partie 2/3)

ferme	commune	longitude	latitude	altitude	associations	mm	temp	année de participation au projet
GIM	04340	44.74	6.54	1218	ARDEAR R-A	61.4 +/- 38.9	11.0 +/- 7.2	2010 2011 2012
HEC	82240	44.39	1.65	165	CETAB	53.1 +/- 32.3	14.3 +/- 6.5	2009 2010 2011 2012
JAB	47400	44.46	0.40	47	CETAB	60.6 +/- 37.1	13.6 +/- 5.9	2009 2010 2011 2013
JFB	47130	44.4	0.42	107	CETAB	52.2 +/- 33.2	13.2 +/- 6	2006 2007 2008 2009 2010 2011 2012
JJG	32550	43.76	0.49	239	CETAB	NA	NA	2009 2010
JMC	24140	45.04	0.60	144	CETAB	68.0 +/- 39.0	12.8 +/- 6.1	2009 2010 2011 2012 2013
JPB					CETAB	NA	NA	2009 2010
JSG	63360	46.2	3.09	318	RSP	49.1 +/- 29.9	11.8 +/- 6.4	2010 2011 2012 2013
JUB	56230	47.73	-2.44	73	Triptolème	70.6 +/- 46.3	11.2 +/- 4.7	2012
LUD	38940	45.48	5.18	489	ARDEAR R-A	70.0 +/- 41.7	13.1 +/- 7.2	2010 2011
MAV	3739 (B)	49.87	5.49	255	RSP			2010 2011 2012
MPH	40190	43.99	-0.29	102	CETAB	73.7 +/- 51.8	13.7 +/- 5.9	2009 2010 2011 2012 2013
OLM	1555 (CH)	47.04	6.98	639	RSP	NA	NA	2010 2012
OLR	95710	49.43	1.71	123	RSP	46.5 +/- 23.4	11.27 +/- 5.9	2009 2010 2011 2013
PAJ	67000	48.58	7.74	152	Kerna ùn Sohma	NA	NA	2012
PIS	64800	43.2	-0.29	143	Pétanielle	NA	NA	2010
PHG	47440	44.47	0.65	87	CETAB	NA	NA	2009



Table 2.1 – (partie 3/3)

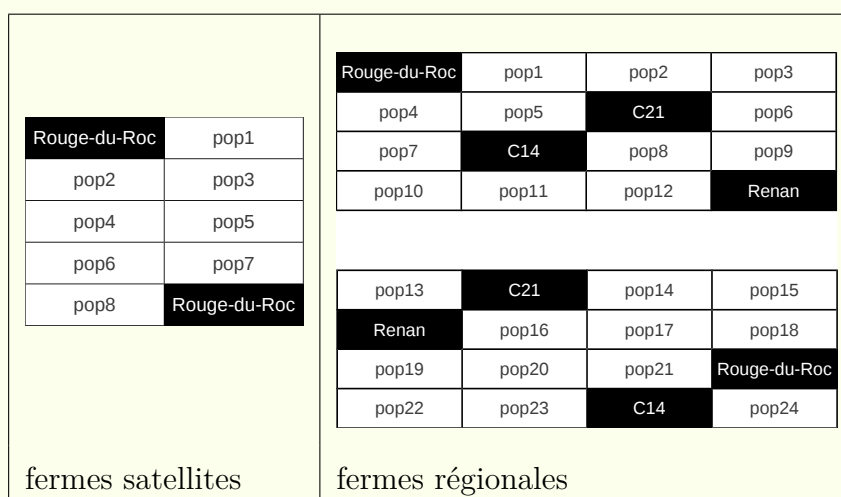
ferme	commune	longitude	latitude	altitude	associations	mm	temp	année de participation au projet
RAB	74380	46.46	6.35	501	ARDEAR R-A	82.8 +/- 47.9	11.5 +/- 6.6	2009 2010 2011 2012 2013
RAL	47260	44.47	0.44	99	CETAB	60.5 +/- 37.1	13.6 +/- 5.9	2010 2011 2012
RIH	67120	48.81	7.66	157	Kerna ùn Shoma	53.5 +/- 33.0	11.2 +/- 6.9	2011 2012 2013
ROW	92500	49.26	2	70	RSP	48.9 +/- 25.8	11.1 +/- 5.9	2010 2011 2012 2013
STP	21440	47.44	4.74	508	Graines de Noé	66.9 +/- 35.0	11.1 +/- 6.7	2012 2013
THG	67330	48.82	7.39	241	Kerna ùn Shoma	53.5 +/- 33	11.2 +/- 6.9	2012
VIC	44190	47.08	-1.28	38	Triptolème	NA	NA	2009

### 2.1.2 Dispositifs mis en place dans les fermes

*JFB* a réalisé dans sa ferme, avec l'aide de l'équipe de recherche DEAP de l'INRA du Moulon, près de 90 croisements en 2005 pour créer de nouvelles populations de blé tendre. *JFB* a choisi les parents qui sont des variétés de pays, des variétés anciennes et quelques variétés modernes de Suisse et d'Allemagne créées pour l'agriculture biologique. Ces nouvelles populations hétérogènes ont été multipliées dans sa ferme entre 2006 et 2008. En 2008, ces populations ont été distribuées au sein du RSP dans différentes fermes réparties dans toute la France. Les fermes ont été divisées en deux types selon le dispositif expérimental qu'elles accueillait : les fermes régionales et les fermes satellites (Figure 2.2).

La variété-population de blé tendre *Rouge-du-Roc* est une sélection paysanne, faite par *JFB*, au sein de la variété *Rouge-de-Bordeaux*. Elle a été utilisée comme témoin répété dans toutes les fermes. Le dispositif expérimental des fermes régionales comportait deux blocs avec quatre populations ou variétés présentes dans chaque bloc : *C14* et *C21* issus de croisements de 2006, *Renan* qui est largement utilisé en agriculture biologique et *Rouge-du-Roc* (Figure 2.2). Dans les fermes satellites, *Rouge-du-Roc* était répété deux fois. Les fermes régionales accueillait entre 20 et 35 populations. Les fermes satellites accueillait entre cinq et treize populations (Tableau 2.2). En plus des témoins, les paysans ont choisi les populations qu'ils souhaitaient semer. Ce dispositif a été mis en place à partir de 2010.

Figure 2.2 - Dispositif dans les fermes satellites à gauche et les fermes régionales à droite.



Les témoins sont sur fond noir.

Le dispositif expérimental est le résultat d'un compromis entre une maximisation du nombre de populations testées dans le terrain alloué et la possibilité d'estimer la variance environnementale dans chaque ferme pour pouvoir différencier les populations. C'est l'équipe de recherche qui propose un plan de semis en fonction de la place disponible et des populations choisies par les paysans. Les plans de la Figure 2.2 sont un exemple. Une grande diversité de dispositifs a été mise en place. Néanmoins, les dispositifs conservaient, au moins, un témoin et un ou plusieurs blocs ; quelques exemples de dispositifs expérimentaux sont présentés en annexes B.

Les micro-parcelles étaient de tailles variables selon les fermes. La majorité des micro-parcelles (80%) mesurait entre 5 et 10 m<sup>2</sup>. Certaines parcelles mesuraient 1 m<sup>2</sup> et pour une ferme, une année, les parcelles ont été réduites à une ligne.

Les micro-parcelles ont directement été incluses dans les champs des paysans. De ce fait, les populations étaient soumises aux pratiques agronomiques des paysans et étaient moins vulnérables aux gibiers et aux oiseaux.

Dans la suite, un environnement est défini comme la combinaison d'une ferme et d'une année. Par exemple, l'environnement CHD2011 correspond à la ferme CHD pour la récolte 2011.

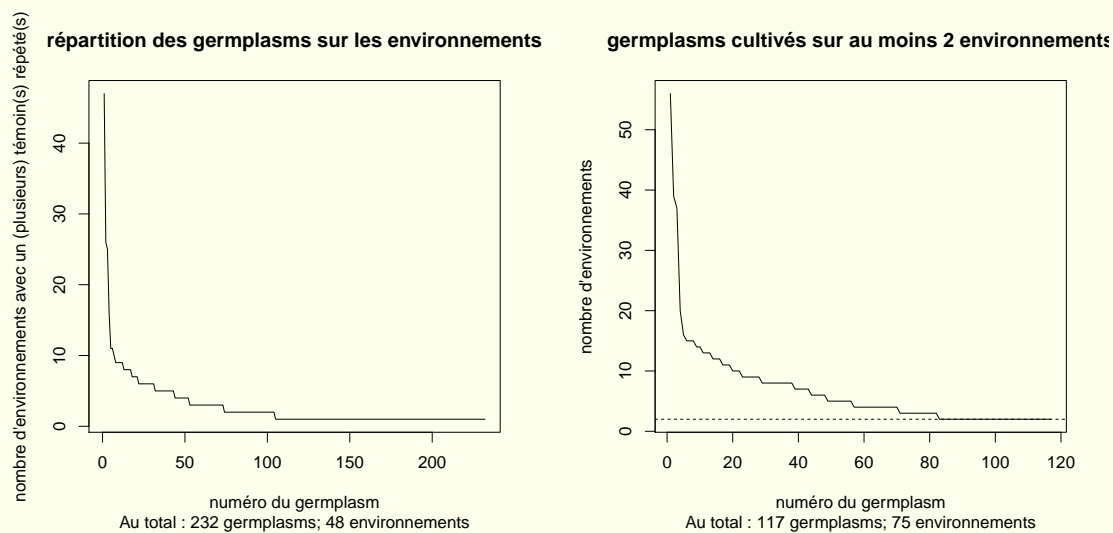
Un germplasm est défini comme une entité génétique regroupant des plantes issues d'un même processus de création, partageant une même origine : un croisement, une variété de pays, une variété ancienne, un mélange, une variété moderne, une variété issue d'une sélection massale dans les types précédemment cités. Une grande diversité de germplasm est cultivée dans le réseau depuis le début du projet (Tableau 2.3). On définit une population comme un germplasm dans une ferme, une année donnée avec des pratiques données. Nous avons appelé « lot de semences » une population dans un espace du champ donné. Par exemple, une population répétée deux fois dans un essai créera deux lots de semences distincts.

Au total, en ne considérant que les lots de semences dont nous avons reçu des données, 254 germplasms ont été cultivés sur 75 environnements (ferme × année).

Sur ces 75 environnements, 48 ont suivi le dispositif avec les témoins, ce qui correspond à 232 germplasms. La répartition des germplasms dans les fermes était très hétérogène (Figure 2.3). 117 germplasms étaient présents dans au moins deux environnements (Figure 2.3). Dans une matrice regroupant toutes les combinaisons germplasm × environnement possibles, il y a plus de 95% de données manquantes.

La proportion de fermes régionales et de fermes satellites a changé au cours du temps (Tableau 2.2). Cinq fermes avec une structure en deux blocs, avec deux témoins répétés au minimum et peu de populations ont été considérées comme des fermes régionales.

Figure 2.3 - Répartition des germplasms dans les environnements.



Un germplasm est représenté par un numéro.

A gauche, le nombre d'environnements (combinaison ferme  $\times$  année) avec, au moins, un témoin répété sur chaque environnement.

A droite, le nombre d'environnements où est présente chaque population du réseau (il n'y a pas forcément de témoins répétés).

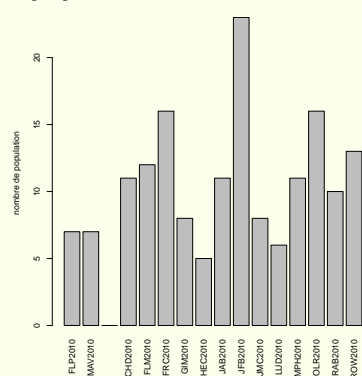
La ligne pointillée représente deux environnements.

Tableau 2.2 - Nombre de fermes régionales (FR) et satellites (FS) pendant les saisons 2009-2010, 2010-2011 et 2011-2012 qui ont envoyé des échantillons d'épis au Moulon ; nombre de fermes qui avaient au moins un témoin répété et nombre de populations cultivées sur chaque ferme avec au moins un témoin répété (témoins compris).

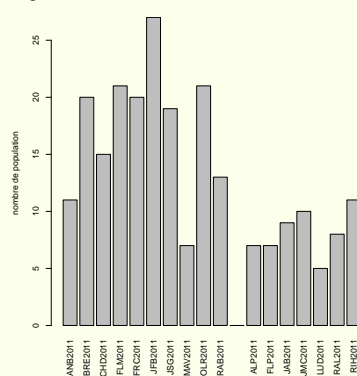
année	type de ferme	nombre de ferme au total	nombre de ferme avec un ou plusieurs témoins répétés
2010	FR	2	2
	FS	18	13
2011	FR	10	10
	FS	8	7
2012	FR	9	9
	FS	13	7

nombre de population cultivée pour chaque FR (gauche) et FS (droite).

2010



2011



2012

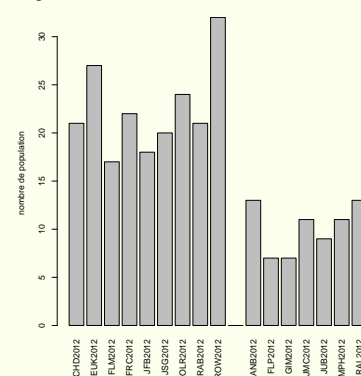


Tableau 2.3 - Proportion et type de germplasm cultivés sur le réseau entre 2008 et 2012.

Type de germplasm	Définition	nb	%
Populations dérivées de croisements de <i>JFB</i>	Création dans la ferme de <i>JFB</i> de croisements entre des variétés de pays, des variétés anciennes et des variétés modernes créées pour l'agriculture biologique.	90	12,9%
Populations dérivées de croisements réalisés par d'autres paysans	Les paysans ont la possibilité de créer leur propre population à partir de parents qu'ils choisissent. Les croisements ont été faits soit dans leur ferme soit à la station de recherche du Moulon.	91	13,1%
Variétés de pays	Variétés qui ont été cultivées avant 1850 et qui ont une forte variabilité intra-variétale.	75	10,8%
Variétés anciennes	Variétés qui ont été sélectionnées dans des variétés de pays après 1850. Comme ces variétés ont été resemées dans les fermes chaque année, elles ont aussi une forte variabilité intra-variétale.	16	2,3%
Mélanges	Mélange entre tous les types possibles.	19	2,7%
Variétés modernes	Variétés modernes sélectionnées pour l'agriculture biologique en France, en Suisse ou en Allemagne.	22	3,2%
Variétés sélectionnées	Variétés issues de sélection massive intra-population faite par un paysan sur un des types précédemment cités.	382	55%

nb : Nombre de type. % pourcentage de type par rapport au total. Les parents des populations issues des croisements sont en annexe C.

### 2.1.3 Variables mesurées

Le tableau 2.4 recense les observations faites par les paysans à quatre étapes du cycle du blé : semis, sortie de l'hiver, épiaison, maturité. Ce tableau est la dernière version des mesures qui ont été prises. Les mesures ont évolué au cours du temps au fur et à mesure de la co-construction du projet (chapitre 4). Les paysans ont réalisé eux-mêmes ces observations sur leurs essais du semis à maturité. Parfois un animateur ou une personne de l'équipe de recherche pouvait aider le paysan dans les observations, mais c'était très rare.

Tableau 2.4 - Mesures réalisées par les paysans au cours du cycle du blé.

Étape du cycle	variable	annexe
semis	nom et taille de la parcelle, topographie, type de sol, battance, drainage, labour, date de semis, méthode de semis, nombre de grains, levée, précédent cultural, informations sur le climat	E
sortie d'hiver	note globale, note de reprise, densité de peuplement, port au tallage, attitude des feuilles, commentaires, informations sur le climat	F
épiaison	note globale, tallage, vigueur, couleur, adventices, maladies, épiaison	G
maturité	note globale, biomasse, courbure et couleur de l'épi, hétérogénéité de la population, barbes, verse, hauteur, date de récolte, poids récolté	H

Les annexes présentent les fiches qui sont envoyées aux paysans ainsi que les notices qui expliquent les méthodes.

Des mesures à maturité et post-récolte ont été réalisées par l'équipe de recherche au champ et en salle technique. Ces mesures varient des fermes régionales aux fermes satellites (Tableau 2.5).

Il y a deux types de mesures : globales et individuelles. Les mesures individuelles concernent 25 épis et/ou plantes pris au hasard dans les populations. Comme il est impossible d'aller dans toutes les fermes, l'équipe de recherche a fait des mesures individuelles

dans les champs sur 25 plantes par population uniquement dans les fermes régionales. Quand cela a été possible, les épis ont été marqués par des étiquettes numérotées pour faire le lien entre les plantes mesurées au champ et les épis mesurés en salle technique. Toutes les fermes, régionales et satellites, ont envoyé au Moulon un échantillon de 50 épis pris au hasard pour chaque population. Des mesures individuelles ont été prises dans la salle technique du Moulon sur 25 épis. Les mesures globales ont été une mesure sur les 25 épis mélangés. Tous les épis ont été battus et les grains conservés en stock au Moulon. Les protocoles utilisés par l'équipe de recherche pour organiser les mesures sont en annexe I.



Tableau 2.5 - Variables mesurées par l'équipe de recherche sur les épis ou sur les plantes et méthode associée.

variable	FR	FS	méthode
couleur <sup>i,1</sup>	X	X	couleur de l'épi de 0 (blanc) à 2 (rouge ou très foncé). Les valeurs possibles sont 0, 0.5, 1, 1.5 et 2.
barbes <sup>i,1</sup>	X	X	présence de barbes sur les épis de 0 (pas de barbes) à 2 (très barbus). Les valeurs possibles sont 0, 0.5, 1, 1.5 et 2.
courbe <sup>i,1</sup>	X	X	courbe de l'épi à maturité de 0 (épi droit) à 2 (épi faisant une rotation à 180°). Les valeurs possibles sont 0, 0.5, 1, 1.5 et 2.
hauteur <sup>g,i,1</sup>	X		hauteur de la plante du sol au sommet de l'épi (sans les barbes) en mm.
distance dernière feuille base de l'épi (LLSD) <sup>i,1</sup>	X		distance de la feuille drapeau à la base de l'épi en mm.
longueur de l'épi <sup>i,1</sup>	X		longueur de l'épi en mm.
nb d'épillets <sup>i,2</sup>	X		nombre d'épillets.
nb d'épillets stériles <sup>i,2</sup>	X		somme d'épillets stériles sur un épi.
poids de l'épi <sup>i,2</sup>	X	X	poids de l'épi en g.
protéine <sup>g,2</sup>	X	X	% de protéine dans le grain mesuré avec la technologie NIRS (FOSS NIRSystem 6500 : spectrométrie infrarouge) à l'INRA de Clermont Ferrand sur grain entier.
pmg <sup>g,2</sup>	X	X	poids de mille grains en g.

FR : fermes régionales.

FS : fermes satellites.

<sup>g</sup> : mesure globale.

<sup>i</sup> : mesure individuelle sur 25 plantes.

<sup>1</sup> : mesure à la ferme.

<sup>2</sup> : mesure à la station de recherche sur un échantillon d'épi envoyé par les paysans.

## 2.2 Analyse phénotypique dans les fermes

Dans un premier temps, le modèle statistique bayésien hiérarchique prenant en compte le déséquilibre des données est décrit. L'article en anglais « Modélisation hiérarchique de la variance pour des essais déséquilibrés : application à des expérimentations flexibles de sélection décentralisée et participative » détaille la méthode statistique et la replace dans le contexte de la sélection participative (annexe K). Ensuite, nous étudions le comportement des populations dans les fermes et les comparons entre elles. Dans un troisième temps, nous étudions la sélection inter-population. Enfin, nous présentons la réponse à la sélection massale intra-population en génération précoce par un paysan et également dans les fermes, par plusieurs paysans.

### 2.2.1 Méthodes

#### 2.2.1.1 Comparaison de moyennes dans les fermes

La comparaison de moyennes dans chaque ferme est un élément pour la sélection inter-population. Les statistiques peuvent être classées en deux catégories principales : les statistiques fréquentistes et les statistiques bayésiennes. L'annexe N décrit rapidement les fondements des statistiques bayésiennes. Plus de détails sur la méthode présentée ci-après sont disponibles dans l'article en annexe K. Tout ce qui suit a été codé en fonctions dans R.

##### 2.2.1.1.1 Le modèle bayésien hiérarchique

Dans notre projet, les populations dans les fermes étaient peu répétées. Lorsque l'on avait 25 individus par population, ces 25 mesures ont dû être considérées comme de « fausses » répétitions. En effet, il y avait 25 valeurs pour décrire une population mais ce sont 25 individus qui étaient échantillonnés dans une même micro-parcelle.

Le modèle que nous avons développé utilise les données moyennes. Toutefois, ces moyennes nous ont donné une caractérisation plus précise de la population que ne l'aurait fait une donnée individuelle. En conséquence, il y avait peu de degrés de liberté pour la résiduelle ce qui ne permettait pas une comparaison de moyenne robuste.

Les fermes régionales, avec quatre témoins répétés deux fois, avaient trois degrés de liberté à la résiduelle. Les fermes satellites, avec un témoin répété deux fois, avaient un degré de liberté à la résiduelle. Cela est peu sachant qu'au moins 30 degrés de liberté sont recommandés<sup>38</sup>.

Le modèle bayésien nous a permis d'utiliser l'information issue des expériences sur

toutes les fermes pour améliorer la connaissance que l'on avait de la variance résiduelle sur une ferme. Pour estimer les variances environnementales par ferme et faire les comparaisons de moyennes, on a utilisé le modèle 2.1 suivant :

$$Y_{ijk} = \mu_{ij} + \beta_{jk} + \varepsilon_{ijk} \quad \varepsilon_{ijk} \sim N(0, \sigma_j^2) \quad (2.1)$$

Où  $Y_{ijk}$  était la valeur phénotypique moyenne pour une variable donnée pour la population  $i$ , l'environnement  $j$  et le bloc  $k$ ;  $\mu_{ij}$  était l'effet de la population  $i$  dans l'environnement  $j$ ;  $\beta_{jk}$  était l'effet lié au bloc  $k$  dans l'environnement  $j$ ;  $\varepsilon_{ijk}$  était ce qui n'était pas expliqué par le modèle sur l'environnement  $j$ .  $\varepsilon_{ijk}$  était issu d'une loi normale centrée sur 0 et de variance  $\sigma_j^2$ .

Ce modèle a permis d'estimer une variance environnementale par ferme. On a fait l'hypothèse que les variances environnementales de chaque ferme étaient issues d'une loi gamma inverse. Avec  $\tau_j = \sigma_j^{-2}$  on a :

$$\tau_j \sim \text{Gamma}(\nu, \rho)$$

La loi gamma inverse est bornée à 0 (cohérent avec la définition de la variance qui est toujours positive) et peut avoir des formes asymétriques. D'un point de vue agronomique, c'est cohérent avec l'idée que la diversité des fermes en agriculture biologique est grande et donc que la variance au sein de ces fermes est prise dans une distribution asymétrique. D'un point de vue statistique, la loi gamma inverse a des propriétés de conjugaison qui facilitent la convergence de l'algorithme MCMC<sup>ii</sup>. L'idée du modèle était d'utiliser toute l'information qui était disponible dans le réseau de fermes pour mieux estimer la loi des variances résiduelles de chaque ferme. Nous avons fait l'hypothèse que la variance des témoins était représentative de la variance des autres populations.

Comme il n'y avait pas beaucoup de répétitions, il était plus raisonnable de profiter de toute l'information des données pour estimer la loi gamma, on a donc cherché à estimer les hyperparamètres  $\nu$  et  $\rho$  tel que :

$$\nu \sim \text{Uniforme}(2, 10) \quad \rho \sim \text{Gamma}(10^{-6}, 10^{-6})$$

On a défini la loi uniforme des  $\nu$  pour que les variances de la loi gamma inverse soient définies<sup>203</sup>. La loi sur les  $\rho$  était sans a priori et laissait un large de choix pour les valeurs que pouvaient prendre les paramètres.

---

ii. plus de détails en annexe N.

On a fait également l'hypothèse que :

$$\mu_{ij} \sim N(\mu_{.j}, 10^6) \quad \beta_{jk} \sim N(0, 10^6)$$

Les valeurs phénotypiques mesurées sur les populations suivaient une loi par ferme, la moyenne de cette loi était la valeur moyenne des populations dans la ferme ( $\mu_{.j}$ ). La loi sur les  $\beta$  a été choisie avec une grande variance pour laisser un large choix pour les valeurs que pouvaient prendre les paramètres.

A partir des distributions des paramètres a priori (les priors) et des données, les distributions a posteriori des paramètres (les posteriors) ont été estimées avec 20 000 points issus de 100 000 simulations MCMC<sup>iii</sup>.

#### 2.2.1.1.2 Les comparaisons de moyennes

La comparaison des populations dans une ferme a été basée sur la distribution des  $\mu_{ij}$ . J'ai calculé le nombre de simulations MCMC qu'avaient en commun les deux distributions que l'on souhaitait comparer. Cette proportion correspond à la probabilité que les deux distributions soient égales. 20 000 points ont été utilisés pour calculer cette probabilité. Les probabilités que deux populations soient identiques ont été stockées dans une matrice  $M$  de dimension  $(I - 1, I - 1)$ , avec  $I$  le nombre de population dans l'environnement  $j$ . Il y avait autant de matrices  $M$  que d'environnements.

En théorie, les probabilités estimées devraient décroître au fur et à mesure que la différence entre les estimations des paramètres augmentaient. Prenons un exemple. Sur une ferme, on observe  $\mu_{i,j} < \mu_{i',j} < \mu_{i'',j}$ . Soit  $H_0'$  et  $H_0''$ , les hypothèses nulles telles que :

$$H_0' : \mu_{i,j} \geq \mu_{i',j} \quad \text{et} \quad H_0'' : \mu_{i,j} > \mu_{i'',j}$$

En théorie, on doit avoir  $Pr\{H_0'|y\} \geq Pr\{H_0''|y\}$ , avec  $y$ , les observations. En effet, sachant que  $\mu_{i,j} \leq \mu_{i',j} \leq \mu_{i'',j}$ , la probabilité que  $\mu_{i,j} \leq \mu_{i'',j}$  est plus forte que celle de  $\mu_{i,j} \leq \mu_{i',j}$ . Cela revient à dire que l'on s'attend à ce que la probabilité que  $H_0''$  soit vraie soit plus faible que la probabilité que  $H_0'$  soit vraie.

A partir de la matrice  $M$ , des groupes de populations ont été créés avec un seuil donné. Dans notre cas, on a pris comme risque individuel  $\alpha = 0.1/I$  (avec  $I$  le nombre de populations dans l'environnement). Pour un environnement avec 20 populations,  $\alpha =$

---

iii. plus de détails sur ces méthodes en annexe N.

0.005. Pour un environnement avec 10 populations,  $\alpha = 0.01$ . Cela se rapproche d'une correction de Bonferroni moins stricte.

Pour chaque ferme, la valeur du risque global  $\alpha = 0.1$  a été prise par défaut. Cependant, si avec ce seuil, il n'y avait qu'un seul groupe de populations dans la ferme, alors nous avons utilisé la valeur de  $\alpha$  la plus petite, telle que deux groupes puissent être détectés dans l'essai. Cela signifie que, pour chaque ferme, nous avons obtenu au moins deux groupes avec une valeur d'erreur globale  $\alpha$  donnée, en nous limitant à  $\alpha < 0,5$  (afin que le risque de se tromper reste inférieur à 0,5). Seuls 9% des couples environnement  $\times$  variable n'ont pas permis de différencier au moins deux groupes.

### 2.2.1.1.3 Sélection inter-population

Chaque année, un paysan peut choisir de ressemer ou non des populations l'année suivante. On parle de sélection inter-population.

Afin d'analyser ce choix en fonction des classements que nous avons faits entre populations, nous avons regardé à quel(s) groupe(s) de significativité appartiennent les populations qui n'ont pas été gardées pour le semis de l'année suivante (sélection inter-population négative). Les groupes de significativité ont été déterminés selon la méthode présentée précédemment et avec une valeur de risque global  $\alpha$  permettant de discriminer au moins deux groupes par ferme.

Pour chaque variable, sur chaque ferme, une valeur a été donnée à chaque population correspondant au(x) groupe(s) de significativité auquel elle était assignée. Pour toutes les variables, les groupes ont été ordonnés dans l'ordre croissant des moyennes : le groupe « a » correspondait aux moyennes les plus faibles. De fait, toutes les fermes qui n'avaient pas de témoin, ou qui ne nous ont pas donné l'information, n'ont pas été analysées.

Au total, 40 populations issues des récoltes 2010, 2011 et 2012 sur dix fermes ont été analysées. Les 40 populations choisies pour l'analyse ont été sélectionnées sur la base d'informations fiables qui permettaient d'affirmer que ces populations étaient issues d'une sélection inter-population négative. En effet, certaines populations étaient gardées pour être ressemées plus tard, des populations n'ont pas été ressemées par exemple pour cause d'accident de récolte. Tous ces cas n'étaient pas représentatifs de sélections délibérées.

Nous avons attribué à chaque groupe de significativité une valeur. Le groupe « a » a eu la valeur 0, le groupe « b » a eu la valeur 1, ..., le groupe « z » a eu la valeur 26. La valeur assignée à chaque population était égale à la moyenne de la valeur des groupes auquel elle appartenait. Par exemple, si la population appartenait aux groupes « a » et

« b », sa valeur était  $(0 + 1)/2 = 0.5$  ; si la population appartenait aux groupes « b » et « c », sa valeur était  $(1 + 2)/2 = 1.5$ .

Cette valeur a été divisée par le nombre de groupe total - 1 (car le premier groupe avait une valeur de 0) pour donner un score. Les scores, obtenus ainsi, sont compris entre 0 et 1. Un score de 0 signifie que la population appartenait à un groupe avec la moyenne la plus faible pour le caractère considéré. Un score de 1 signifie que la population appartenait à un groupe avec la moyenne la plus forte pour le caractère considéré.

### 2.2.1.2 Sélection massale intra-population

#### 2.2.1.2.1 Réponse à la sélection massale précoce par un paysan

Au départ du projet en 2008, *JFB* a réalisé des sélections massales au sein des populations *F2* issues des croisements. Les descendances *F3* issues de sélections et de vrac en *F2* pour 35 populations ont été évaluées au Moulon en trois répétitions en 2009-2010. L'article en anglais « Réponse à la sélection massale paysanne en génération précoce dans des populations de blé tendre issues de croisements » en annexe J, présente ces analyses.

#### 2.2.1.2.2 Réponse à la sélection massale de différents paysans sur leurs fermes

##### Dispositif expérimental dans les fermes

Les premiers résultats sur les sélections massales en génération précoce de *JFB* ont donné une indication sur la possibilité des paysans à faire de la sélection. Ces résultats obtenus par le paysan « expert » du réseau ont encouragé les autres paysans à réaliser eux aussi des sélections massales. L'équipe de recherche a proposé aux paysans qui le souhaitent d'étudier la réponse à leur sélection. Pour cela, les paysans ont envoyé leur bouquet de sélection (autant d'épis qu'ils le souhaitent). Nous avons ensuite mesuré 25 épis représentatifs du bouquet (s'il y en a moins, tous les épis du bouquet) ainsi que 25 épis dans la population non sélectionnée (vrac), correspondante. Cela a permis de calculer le différentiel de sélection *S*. L'année suivante, la réponse à la sélection a été calculée sur des échantillons de 25 épis pris au hasard dans les descendances des sélectionnés et du vrac.

Les dispositifs expérimentaux n'ont pas changé et les fermes sont restées dans la catégorie régionale ou satellite (partie 2.1.2). Cependant, il a été proposé que les descendances issues des bouquets de sélection, et celles issues du vrac, soient semées côte à côte (Figure 2.4). Les protocoles utilisés lors des récoltes sont en annexe I.1.

L'analyse s'est déroulée en deux temps :

1. Calcul du différentiel de sélection *S* l'année de la sélection.

2. Calcul de la réponse à la sélection  $R$  l'année suivante.

Figure 2.4 - Exemple d'un dispositif expérimental dans une ferme satellite qui souhaitait étudier la réponse à la sélection.

Rouge-du-Roc	pop1
pop2	pop2#S
pop4	pop5
pop6#S	pop6
pop7	Rouge-du-Roc

*pop2* représente la descendance issue du vrac (version non sélectionnée) et *pop2#S* représente la version sélectionnée; idem pour *pop6*.

Pour chaque population, le taux de sélection a été estimé par le rapport entre :

1. le nombre de plantes sélectionnées, on a fait l'hypothèse que chaque épi sélectionné provenait d'une plante différente,
2. et le nombre de plantes sur la parcelle que l'on a obtenu en multipliant la densité de semis (obtenue par l'intermédiaire de la fiche sortie d'hiver) par la surface des micro-parcelles (relevée avec la fiche semis).

Comme la densité était souvent donnée dans une fourchette, le taux de sélection l'a été également. A cause du peu de précision du taux de sélection, nous n'avons pas calculé l'intensité de sélection.

#### Calcul du différentiel de sélection $S$

Le différentiel de sélection a été estimé pour chaque couple : bouquet de sélection/vrac par le modèle suivant :

$$Y_{ij} = \mu + v_i + \varepsilon_{ij} \quad (2.2)$$

Avec  $Y_{ij}$  la valeur phénotypique de l'individu  $j$  de la version  $i$  pour une variable donnée,  $\mu$  la moyenne générale,  $v_i$  l'effet de la version  $i$  (sélectionnée ou non),  $\varepsilon_{ij}$  la résiduelle.

Ce modèle est équivalent à un test de Student sur les deux moyennes. Comme le bouquet et le vrac étaient issus de la même parcelle, on a fait l'hypothèse que la variance micro-environnementale était la même pour les deux versions. Les différences de variance entre le bouquet de sélection et le vrac ont été testées par un test de Fisher avec la fonction `var.test` de R.

Pour les variables quantitatives sans données individuelles par échantillon, on ne pouvait pas faire de comparaisons de moyennes car il n'y avait pas de variance dans l'échantillon. Pour les variables qualitatives (barbe, couleur et courbe), un test du chi2 a été réalisé.

Pour les valeurs de  $S$  obtenues pour différents paysans et différentes populations pour une même variable, nous avons réalisé le test des signes pour voir si les valeurs de  $S$  étaient plutôt positives ou négatives<sup>53</sup>. Soit  $H_0 : p(+) = p(-) = 0.5$  : il y a autant de différences positives ( $p(+)$ ) et négatives ( $p(-)$ ). Quand  $H_0$  est vraie, le nombre de différences positives (ou négatives) suit une loi binomiale de paramètre  $p = \frac{1}{2}$  et  $n$ , le nombre total de différence. On compare le nombre de signes attendus sous  $H_0$ ,  $\frac{n}{2}$ , avec le nombre de signes réels. On a alors :  $P\{X \geq x\} = (\frac{1}{2})^n \sum_{i=0}^x \binom{n}{i}$  avec  $x$  le plus petit des nombres de différences, + ou -.  $H_0$  est rejetée quand  $P\{X \geq x\} \geq \frac{\alpha}{2}$ , avec  $\alpha = 0.05$ . Ce test a été appliqué pour chaque variable, sur les différentiels de sélection des différents paysans pour toutes les populations mélangées.

### Calcul de la réponse à la sélection $R$

Pour la réponse à la sélection, les mêmes analyses que pour le différentiel de sélection ont été faites.

Pour comparer la moyenne des descendants issus de la sélection et la moyenne des descendants issus du vrac, on a utilisé le modèle 2.1 présenté précédemment (approche bayésienne sur l'ensemble des environnements), la  $p$ -value étant dans la matrice  $M$ . Dans ce cas, on prend un risque individuel d'erreur de 5%. Si la matrice  $M$  n'a pas été calculée (fermes sans témoin ou variables qui n'ont pas été mesurées sur les témoins), alors le modèle 2.2 a été utilisé.

### Étude des couples $S$ et $R$

Chaque couple  $S$  et  $R$  a été analysé. Les héritabilités réalisées ont été calculées. Pour une variable donnée, nous avons considéré comme nul une réponse ou un différentiel de sélection compris entre  $[-\frac{0,1}{\sigma_p}, \frac{0,1}{\sigma_p}]$ , avec  $\sigma_p$ , l'écart type de la variable dans le jeu de données total.



## 2.2.1.2.3 Théorie

Soit  $\phi_A$ , le coefficient de simple parenté ; c'est la probabilité pour qu'à un locus donné, un gène<sup>iv</sup> pris au hasard chez un individu soit identique à un gène pris au hasard chez l'autre, c'est à dire que ces deux gènes dérivent par descendance d'un même gène ancêtre.

Soit  $\phi_D$ , le coefficient de double parenté ; c'est la probabilité que les individus  $X$  et  $Y$  aient reçus, via leurs parents respectifs, le même couple de gènes, c'est à dire que l'on tire le même couple de gène dans  $X$  et  $Y$ .

La covariance entre les parents  $P$  et les descendants  $O$  peut s'écrire :

$$\text{cov}(O, P) = 2\phi_{A(O,P)}\sigma_A^2 + \phi_{D(O,P)}\sigma_D^2 \quad (2.3)$$

En population panmictique<sup>v</sup>,  $\phi_{A(O,P)} = \frac{1}{4}$  et  $\phi_{D(O,P)} = 0$  ;

Il vient

$$\text{cov}(O, P) = 2 \times \frac{1}{4} \times \sigma_A^2 + 0 \times \sigma_D^2 = \frac{1}{2} \times \sigma_A^2 \quad (2.4)$$

Nous reprenons ici le formalisme de la partie 1.2.2, qui étudiait la sélection indirecte. La réponse des descendants  $O$  suite à la sélection dans les parents  $P$ , mesurée sur le même caractère dans le même environnement, peut s'écrire :

$$R = b_{OP} \times S, \text{ avec } b_{OP} = \frac{\text{cov}(O, P)}{\text{var}(P)} = \frac{\frac{1}{2}\sigma_A^2}{\text{var}(P)}$$

Si l'on sélectionne sur les deux parents on peut considérer la régression des descendants en fonction des parents moyens  $P1$  et  $P2$ . Sachant que la variance  $\text{var}(P)$  correspond à la variance de la moyenne des parents  $P1$  et  $P2$ , la variance phénotypique observée s'écrit :

$$\text{var}(P) = \text{var}\left(\frac{P1 + P2}{2}\right) = \frac{1}{4}(\text{var}(P1) + \text{var}(P2)) = \frac{1}{4}(2 \times \sigma_P^2) = \frac{1}{2}(\sigma_P^2)$$

D'où

$$b_{OP} = \frac{\text{cov}(O, P)}{\text{var}(P)} = \frac{\frac{1}{2}\sigma_A^2}{\frac{1}{2}\sigma_P^2} = \frac{\sigma_A^2}{\sigma_P^2} = h_{ss}^2$$

Avec  $h_{ss}^2$ , l'héritabilité au sens strict.

Dans un environnement donné, pour une même variabilité phénotypique observée  $\sigma_P^2$ , l'héritabilité au sens strict va dépendre de la valeur de  $\text{cov}(O, P)$ .

Dans une population panmictique,  $\text{cov}(O, P) = \frac{1}{2}\sigma_A^2$ .

iv. Le gène ici est au sens de la génétique des population, il correspond à un allèle.

v. Une population panmictique est une population où les individus se reproduisent au hasard.

Dans une population consanguine, comme c'est le cas chez le blé autogame, la covariance entre deux individus apparentés  $X$  et  $Y$  est augmentée du fait qu'un individu consanguin va avoir plus de chance de transmettre des gènes identiques. Cette covariance peut s'écrire  $cov(X, Y) = c_{tgg'}$  avec  $t$  la génération du plus proche ancêtre commun,  $g$  la génération de  $X$  et  $g'$  la génération de  $Y$ <sup>118</sup> :

$$c_{tgg'} = (1 + F_t)\sigma_A^2 + \frac{(F_g + F_{g'} + 2F_t)}{2}\sigma_{AD_0} + \frac{(1 - F_g)(1 - F_{g'})}{(1 - F_t)}\sigma_D^2 + \left[ F_t + \frac{(F_g - F_t)(F_{g'} - F_t)}{2(1 - F_t)} \right] \sigma_{D_0}^2 + F_t \left[ \frac{(1 - F_g)(1 - F_{g'})}{(1 - F_t)} \right] \sum_l D_{0,l}^2 \quad (2.5)$$

Avec

$F_g, F_{g'}, F_t$  les coefficients de consanguinité des individus  $X, Y$  et de leur ancêtre commun,

$\sigma_A^2$  la variance génétique additive,

$\sigma_{AD_0}$  la covariance entre les effets d'additivité et de dominance à l'état homozygote,

$\sigma_D^2$  la variance génétique de dominance,

$\sigma_{D_0}^2$  la variance des résidus de dominance entre gènes identiques,

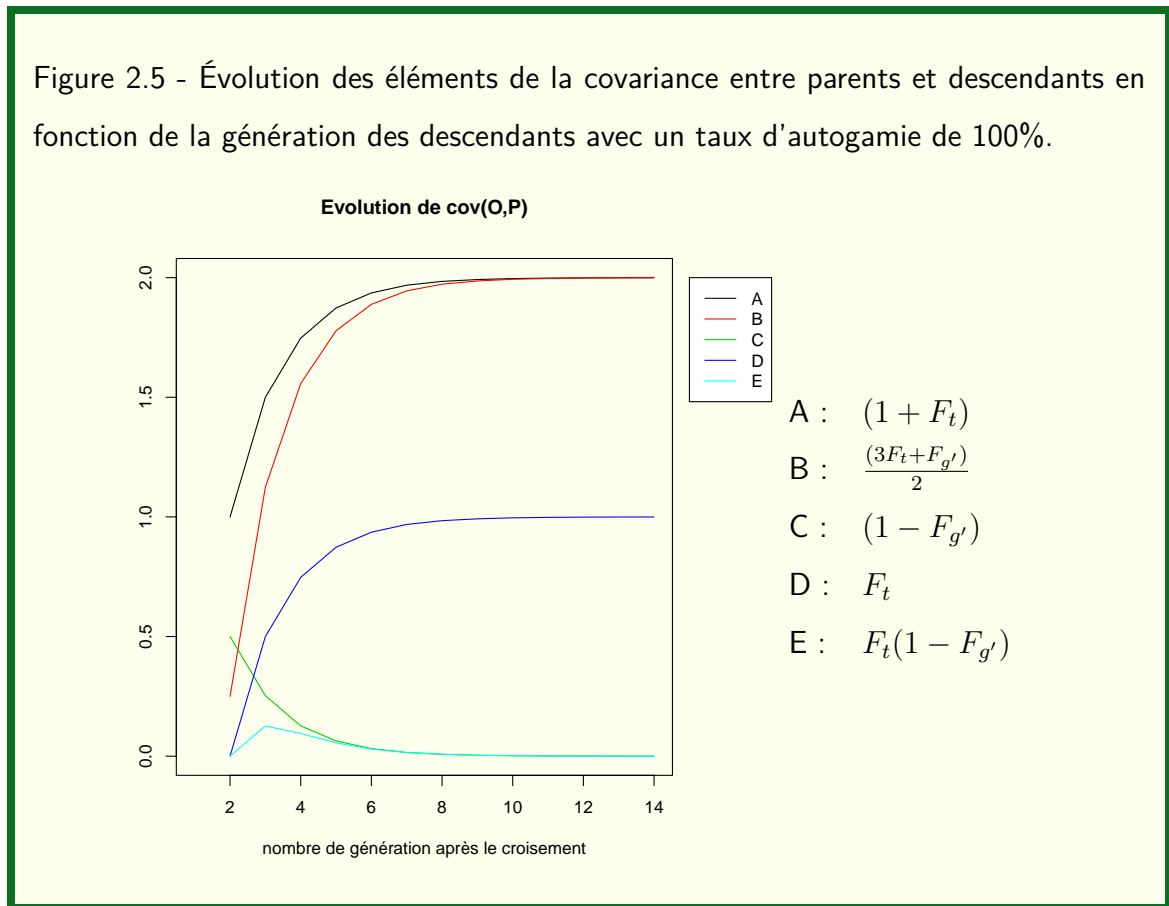
$\sum_l D_{0,l}^2$  la somme des carrés de la dépression de consanguinité pour le locus  $l$ .

L'expression de la covariance parent-descendant, dans le cas d'une autofécondation, peut s'écrire  $c_{ttg'}$  car  $g' = t$  la génération des descendants et  $t = g$  la génération des parents. L'expression de  $c_{ttg'}$  devient :

$$c_{ttg'} = (1 + F_t)\sigma_A^2 + \frac{(3F_t + F_{g'})}{2}\sigma_{AD_0} + (1 - F_{g'})\sigma_D^2 + F_t\sigma_{D_0}^2 + F_t(1 - F_{g'}) \sum_l D_{0,l}^2 \quad (2.6)$$

Les éléments de  $cov(O, P) = c_{ttg'}$  dépendent du coefficient de consanguinité  $F_t$  des parents et  $F_{g'}$  des descendants et évoluent au fil des générations (Figure 2.5).

Figure 2.5 - Évolution des éléments de la covariance entre parents et descendants en fonction de la génération des descendants avec un taux d'autogamie de 100%.



Au cours des générations d'autofécondation, la consanguinité augmente,  $F_t$  et  $F_{g'}$  tendent vers 1 et la covariance entre les parents et les descendants est beaucoup plus forte qu'en panmixie.

Lorsque l'homozygotie est totale, on a  $F_{g'} = F_t = 1$ . La covariance entre parents et descendants s'exprime de la façon suivante :

$$cov(O, P) = 2\sigma_A^2 + 2\sigma_{AD_O} + \sigma_{D_O}^2 \quad (2.7)$$

Tout cela néglige les effets d'épistasie, c'est à dire les interactions entre les gènes à des locus différents. Notons toutefois que les effets d'épistasie additive  $\times$  additive sont également mieux transmis par autofécondation qu'en panmixie<sup>120</sup>.

Plus d'informations sur les formules utilisées peuvent être trouvées en annexe O.

## 2.2.2 Résultats

### 2.2.2.1 Modèle Bayésien hiérarchique

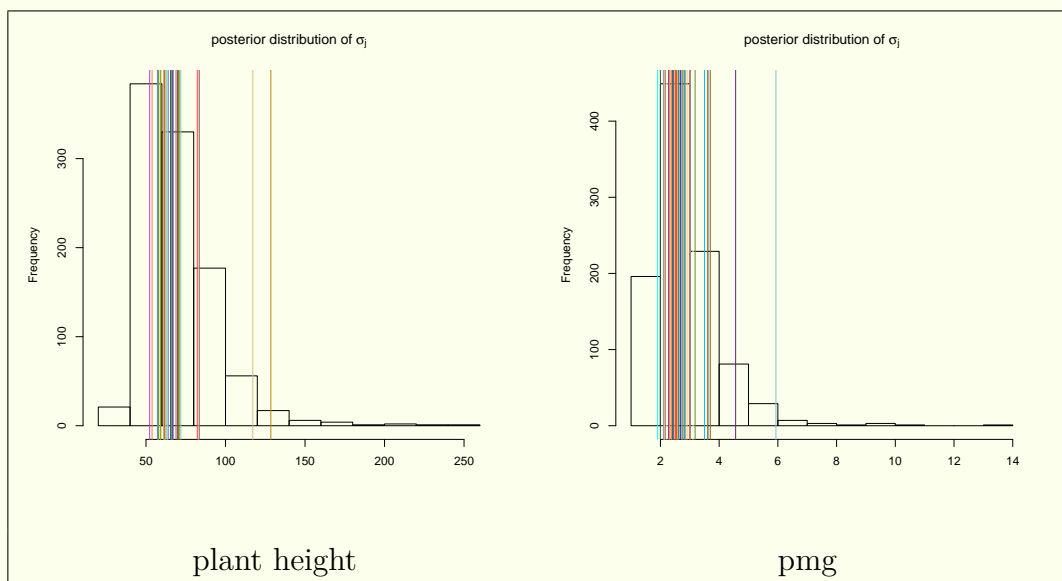
L'article en annexe K détaille le modèle et ses sorties.

Pour chaque variable, les variances résiduelles de chaque environnement ont été estimées à partir de la loi commune des variances environnementales (Figure 2.6). Nous avons observé que les variances résiduelles estimées ( $\sigma_j^2$ ) apparaissaient assez regroupées avec

quelques environnements où elles étaient nettement supérieur.

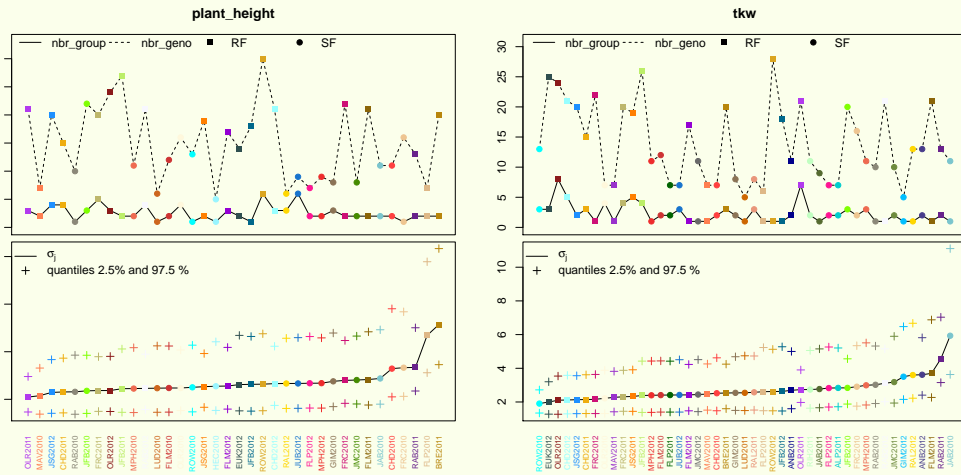
Pour chaque environnement, une fois la valeur de  $\sigma_j$  estimée, les moyennes des populations ont été comparées afin de détecter des groupes de populations significativement différentes (Figure 2.7). Les résultats indiquent que plus la différence entre deux moyennes était grande, plus la probabilité que ces moyennes soient égales diminuait (Figure 2.8). Ce résultat est cohérent avec l'estimation des probabilités.

Figure 2.6 - Distribution des  $\sigma_j(\sqrt{(1/\tau_j)})$  à partir des valeurs de  $\nu$  et  $\rho$  pour la hauteur à gauche et le pmg à droite.



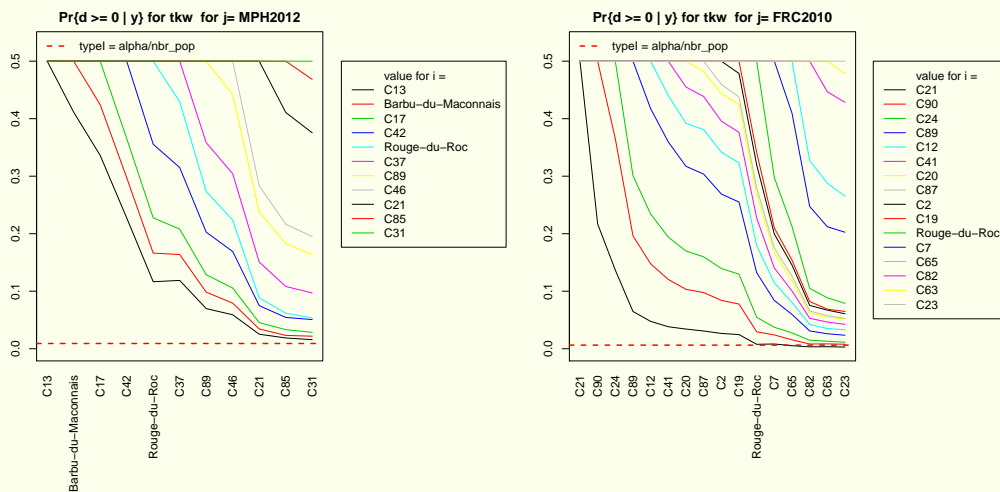
Les barres verticales sont les valeurs observées pour chaque environnement  $j$ .

Figure 2.7 - Nombre de groupes détecté (en haut) et  $\sigma_j$  estimé (en bas), avec les intervalles de confiance, pour chaque environnement pour la hauteur à gauche et le pmg à droite.



RF : fermes régionales ; SF : fermes satellites.

Figure 2.8 - Probabilités que les moyennes deux à deux soient égales pour les environnements MPH2012 et FRC2010 pour le pmg.



Les moyennes sont classées dans l'ordre croissant.

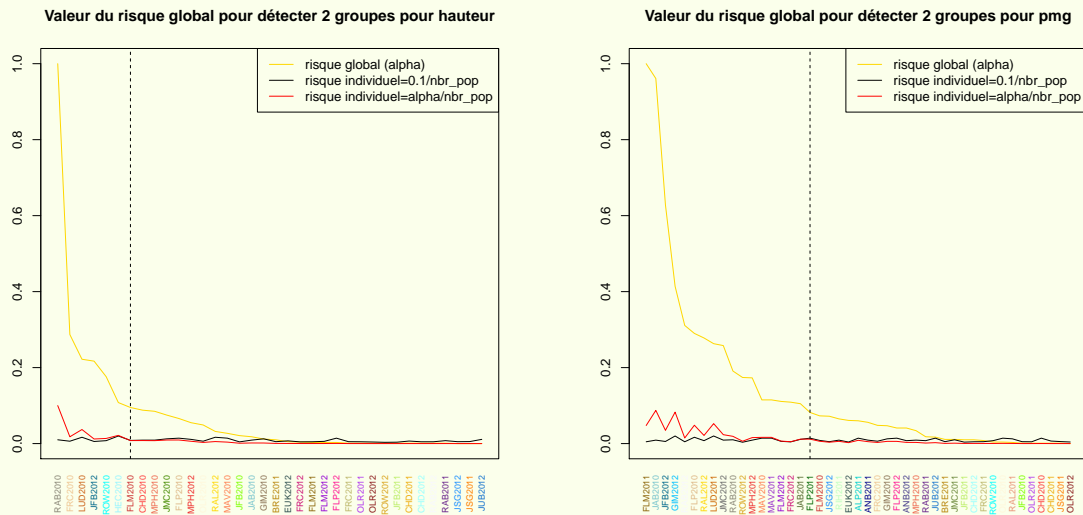
Toutes variables confondues, la méthode a permis de détecter deux groupes avec un risque  $\alpha \leq 0.1$  dans 63% des environnements. Pour les 37% restant, afin de distinguer deux groupes, il a fallu augmenter le risque global. Ainsi, augmenter le risque entre 10% et 20% a permis de distinguer deux groupes dans 15% d'environnements supplémentaires. Augmenter le risque entre 20% et 50% a permis de distinguer deux groupes dans 13% d'environnements supplémentaires. Au delà de 50%, il y a plus de 50% de chance de se tromper, il était donc préférable de garder un seul groupe. Pour certaines variables comme pour LLSD, la méthode a permis de détecter de nombreux groupes alors que pour d'autres, comme le nombre d'épillets stériles ou le taux de protéine, nous n'avons pu détecter que peu de groupes (Tableau 2.6).

La Figure 2.9 illustre la valeur nécessaire de risque global pour détecter deux groupes pour la hauteur et le poids de mille grains. Prenons comme exemple la hauteur. Pour distinguer deux groupes dans les six environnements à gauche de la ligne pointillée, il fallait des valeurs de  $\alpha$  supérieures à 10%. Au contraire, pour les environnements à droite de la ligne pointillée, il fallait des valeurs de  $\alpha$  inférieur à 10%. Cela signifie qu'avec une valeur de  $\alpha=10\%$ , la méthode a permis de distinguer, sur ces environnements, au moins deux groupes.

Tableau 2.6 - Nombre d'environnements avec, au moins, deux groupes pour différentes valeurs du risque global  $\alpha$ .

	$\alpha \leq 0.1$	$0.1 < \alpha \leq 0.2$	$0.2 < \alpha \leq 0.3$	$0.3 < \alpha \leq 0.4$	$0.4 < \alpha \leq 0.5$	$0.5 < \alpha$
couleur	20	2	0	0	1	1
barbes	17	0	0	0	0	0
courbe	19	6	1	1	1	5
nb d'épillets stériles	10	10	5	2	2	6
hauteur	27	2	3	0	0	1
LLSD	22	0	0	0	0	0
longueur de l'épi	17	5	3	0	0	1
nb d'épillets	16	2	1	0	0	2
poids de l'épi	15	9	3	5	3	9
pmg	27	8	4	1	1	3
protéine	15	5	2	1	1	1
<b>pourcentage</b>	<b>63</b>	<b>15</b>	<b>7</b>	<b>3</b>	<b>3</b>	<b>9</b>

Figure 2.9 - Évolution du risque global et individuel pour qu'il y ait deux groupes de populations significativement différentes pour la hauteur et le pmg.



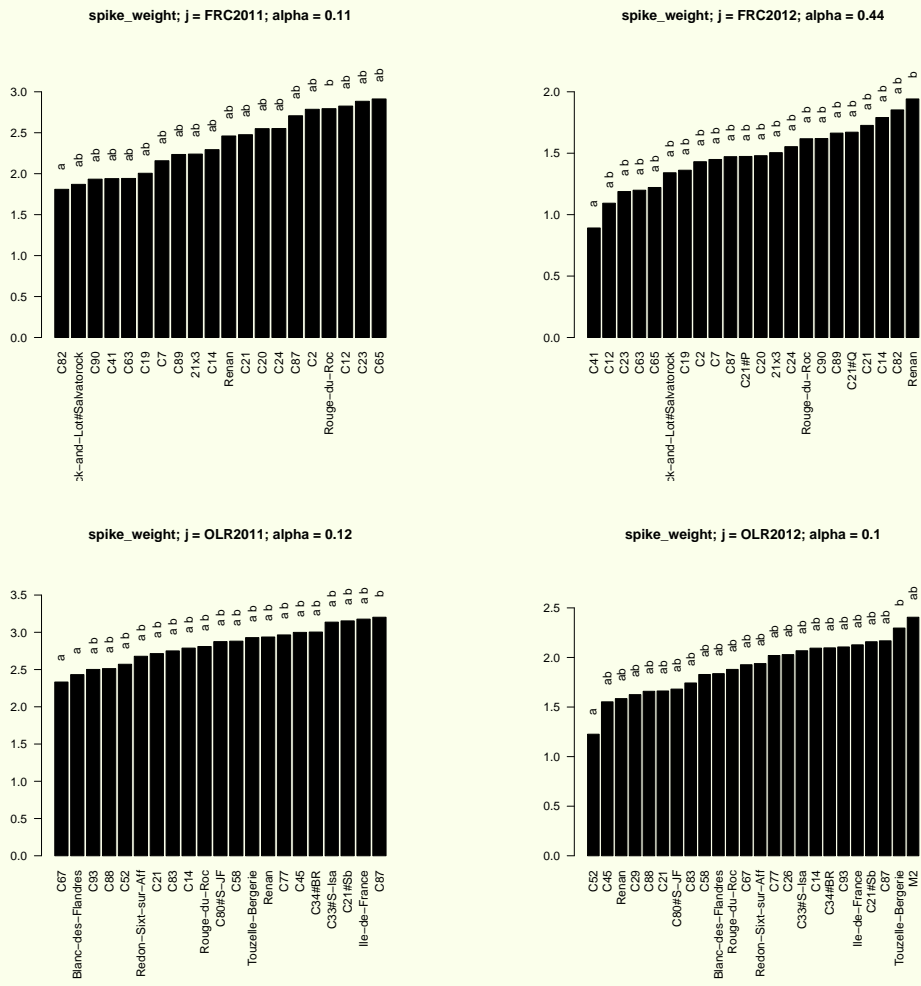
La ligne verticale en pointillé indique la limite entre les environnements qui présentent deux groupes de populations significativement différentes avec le risque global  $\alpha=0.1$ . Les environnements sont en abscisse.

### 2.2.2.2 Comportement des populations dans les fermes

Pour chaque variable et chaque environnement (ferme  $\times$  année), une représentation graphique des comparaisons des moyennes des populations a été donnée aux paysans. Les populations qui n'étaient pas différentes, partagent la même lettre. L'erreur globale  $\alpha$  retenue, pour avoir au moins deux groupes, est indiquée sur le graphique (Figure 2.10 & 2.11). Il n'y a que deux groupes de populations pour chaque exemple avec un risque global qui varie. Certaines populations, comme *Rouge-du-Roc* chez *OLR* ont présenté une teneur en protéine élevée et stable sur les deux années. A l'inverse, *Ile-de-France* a montré un changement de comportement suivant l'année (Figure 2.11). Ce type de résultats permet au paysan de sélectionner entre les populations.

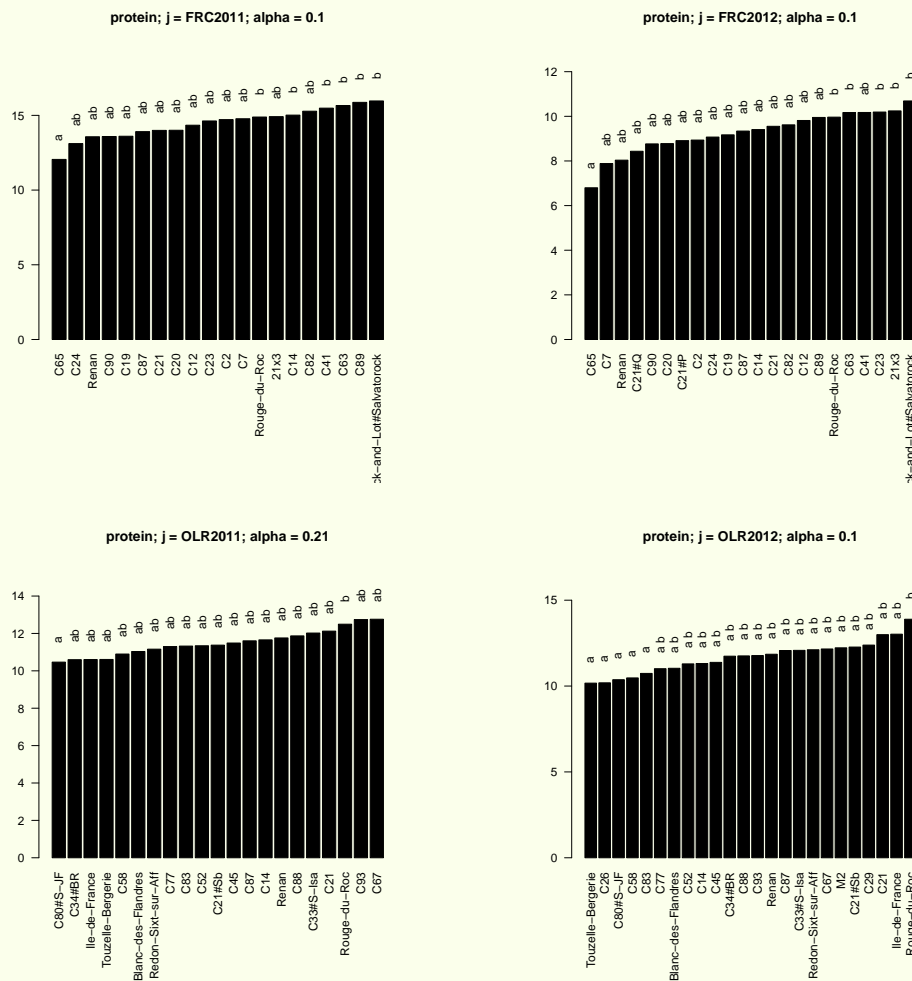


Figure 2.10 - Comparaison de moyennes pour *FRC2011*, *FRC2012*, *OLR2011* et *OLR2012* pour le poids de l'épi.



Le risque global  $\alpha$  est noté dans le titre.

Figure 2.11 - Comparaison de moyennes pour  $FRC_{2011}$ ,  $FRC_{2012}$ ,  $OLR_{2011}$  et  $OLR_{2012}$  pour le taux de protéine.



Le risque global  $\alpha$  est noté dans le titre.

### 2.2.2.3 Sélection inter-population

Afin d'analyser le choix des paysans d'éliminer certaines populations, j'ai recherché dans quel groupe de significativité se trouvaient les populations qui n'étaient pas gardées pour le semis de l'année suivante (sélection inter-négative). Ces populations ont été trouvées distribuées dans tous les groupes (Tableau 2.7). Cependant, nous pouvons observer que les populations qui n'ont pas été ressemées l'année suivante se trouvaient plutôt dans les groupes correspondant aux valeurs les plus faibles sauf pour la hauteur, et dans une moindre mesure LLSD et la longueur de l'épi pour lesquelles les populations qui ont été éliminées se trouvaient dans les groupes correspondant aux valeurs les plus élevées.

Deux exemples sont donnés dans la Figure 2.12. Pour la hauteur, à part une population, toutes les populations écartées avaient un score supérieur à 0,5. Parmi elles, 15 populations faisaient partie des plus grandes. On observe une tendance inverse pour le taux de protéine.

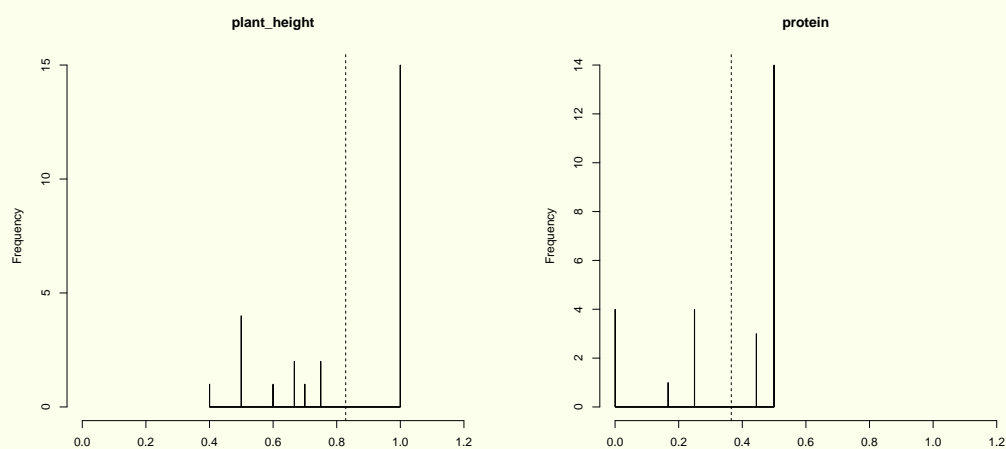
Tableau 2.7 - Pour chaque variable, moyenne, écarts-types et coefficient de variation des scores assignés pour chaque population.

	moyenne	écart type	coefficient de variation
couleur	0.45	0.30	0.66
barbes	0.34	0.35	1.05
courbes	0.46	0.20	0.44
nb d'épillets stérile	0.43	0.33	0.78
hauteur	0.83	0.22	0.26
LLSD	0.54	0.24	0.45
longueur de l'épi	0.56	0.22	0.40
nb d'épillets	0.44	0.28	0.65
poids de l'épi	0.38	0.30	0.77
pmg	0.33	0.32	0.97
protéine	0.37	0.19	0.52

0 : le lot de semences appartient au groupe de significativité avec la moyenne la plus petite.

1 : le lot de semences appartient au groupe de significativité avec la moyenne la plus grande.

Figure 2.12 - Distribution des scores des populations non semées l'année suivante pour la récolte 2010, 2011 et 2012 pour la hauteur à gauche et le taux de protéine à droite.



0 : le lot de semences appartient au groupe de significativité avec la moyenne la plus petite.  
 1 : le lot de semences appartient au groupe de significativité avec la moyenne la plus grande.  
 La ligne pointillée représente la moyenne.

#### 2.2.2.4 Sélection massale intra-population

##### 2.2.2.4.1 Cas particulier de la sélection massale d'un paysan en génération précoce

Les résultats de cette étude sont développés en annexe J. Pour chaque variable étudiée, la sélection s'est montré efficace pour certaines populations  $F3$  dérivées des  $F2$ . 24 populations sur 38 répondent significativement sur au moins un caractère, quatre populations répondent sur cinq caractères ou plus. Pour des caractères d'intérêt comme le poids de mille grains ou le poids du grain par épi, quand la réponse était significative, elle était toujours positive. Pour la plupart des caractères étudiés, la variance génétique entre les familles a augmenté après la sélection alors que la variance génétique à l'intérieur des familles diminuait.

##### 2.2.2.4.2 Réponse à la sélection dans les fermes par les paysans

###### Études des différentiels de sélection

Les données sur 183 bouquets de sélection, et sur les échantillons de vrac correspondants, ont été analysées.

Les taux de sélection variaient entre 2.5 pour 1000 et 12.5 pour 1000 exceptés deux cas à 20 pour 1000 du fait de très petites parcelles. Les bouquets de sélection contenaient tous autour de 25 épis, sauf cinq cas avec cinq épis et un cas avec 175 épis (issus de plusieurs populations).

Dans 60% des cas (15/26) des couples année  $\times$  variable, toutes populations confondues, les signes de  $S$  étaient significativement différents (Tableau 2.8 et Figure 2.13). Le différentiel de sélection était souvent positif pour le poids de l'épi (significativement pour 2010, 2011 et 2012), le nombre d'épillets, la longueur de l'épi et la courbe de l'épi (significativement pour 2010 et 2011). Le différentiel de sélection était souvent négatif pour le nombre d'épillets stériles (significativement pour 2010, 2011 et 2012). Nous n'avons pas observé de tendance pour les barbes et la couleur. Pour le poids de mille grains et le taux de protéine, les tendances variaient d'une année à l'autre : le différentiel de sélection était plus souvent positif en 2011 et jamais significatif en 2012 (Tableau 2.8). Pour le poids de l'épi, le nombre d'épillets et le nombre d'épillets stériles, la majorité des bouquets de sélection étaient significativement différents du vrac.

La sélection n'a pas d'effet clair sur l'évolution des variances qui peuvent augmenter ou diminuer selon les années et les variables.

#### Étude de la réponse à la sélection

J'ai pu analyser la réponse à la sélection sur les descendances de 37 bouquets de sélection.

Pour calculer la réponse à la sélection, le modèle bayésien a été utilisé dans 58% des cas, sinon un modèle simple comme pour le différentiel de sélection a été utilisé.

Dans 9% des cas (2/22) des couples année  $\times$  variable, toutes populations confondues, les signes de  $R$  étaient significativement différents (au seuil de 5%) pour la couleur et le nombre d'épillets : la couleur tendait à devenir plus foncée et le nombre d'épillets à augmenter. Sans être significatif, LLSD, la hauteur, le taux de protéine ont diminué plus souvent alors que la courbure, la longueur et le poids de l'épi augmentaient plus souvent.

Les variances pouvaient diminuer ou augmenter selon les cas (Tableau 2.9 et Figure 2.14). Certaines variances pour la hauteur n'ont pas été mesurées car nous n'avions qu'une donnée moyenne.

#### Étude des couples $S$ et $R$

Dans 40% des cas, les valeurs de  $R$  et de  $S$  étaient cohérentes car du même signe ou nulles (Tableau 2.10). Les descendances évoluent dans le sens de la sélection de l'année précédente. Dans ce cas, les héritabilités variaient entre 0,06 et des valeurs supérieures à 1.

Dans 11,3% des cas, quelque soit la valeur de  $S$ , les réponses étaient nulles (Tableau 2.10). La sélection exercée n'est suivie d'aucune réponse l'année suivante. Dans ce cas, les héritabilités étaient très faibles, nulles, ou négatives du fait que le seuil pour considérer

$R$  nul était de + ou  $- 0,1$  divisé par  $\sigma_p$  (Tableau 2.10).

Nous avons observé 36,1% des cas correspondant à des situations, en théorie, impossibles : les signes de  $S$  et de  $R$  étaient contraires (Tableau 2.10). Les descendances évoluent dans un sens opposé à celui de la sélection.

Dans 12,6% des cas, des valeurs de  $S$  nulles étaient suivies de valeurs de  $R$  non nulles (Tableau 2.10).

Figure 2.13 - Différentiel de sélection  $S$  en 2011 pour le poids de l'épi.

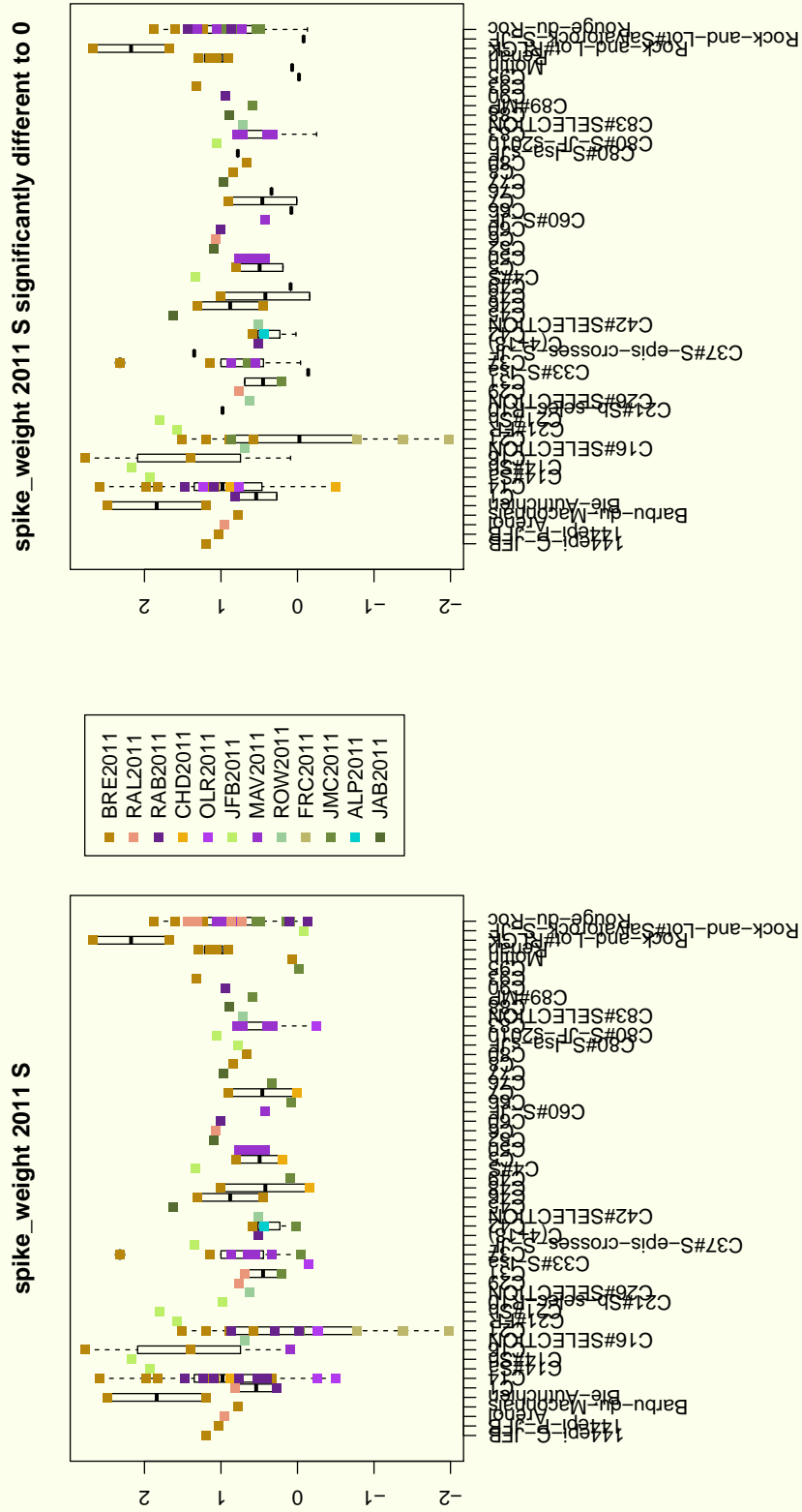


Figure 2.14 - Réponse à la sélection  $R$  en 2012 pour le poids de l'épi.

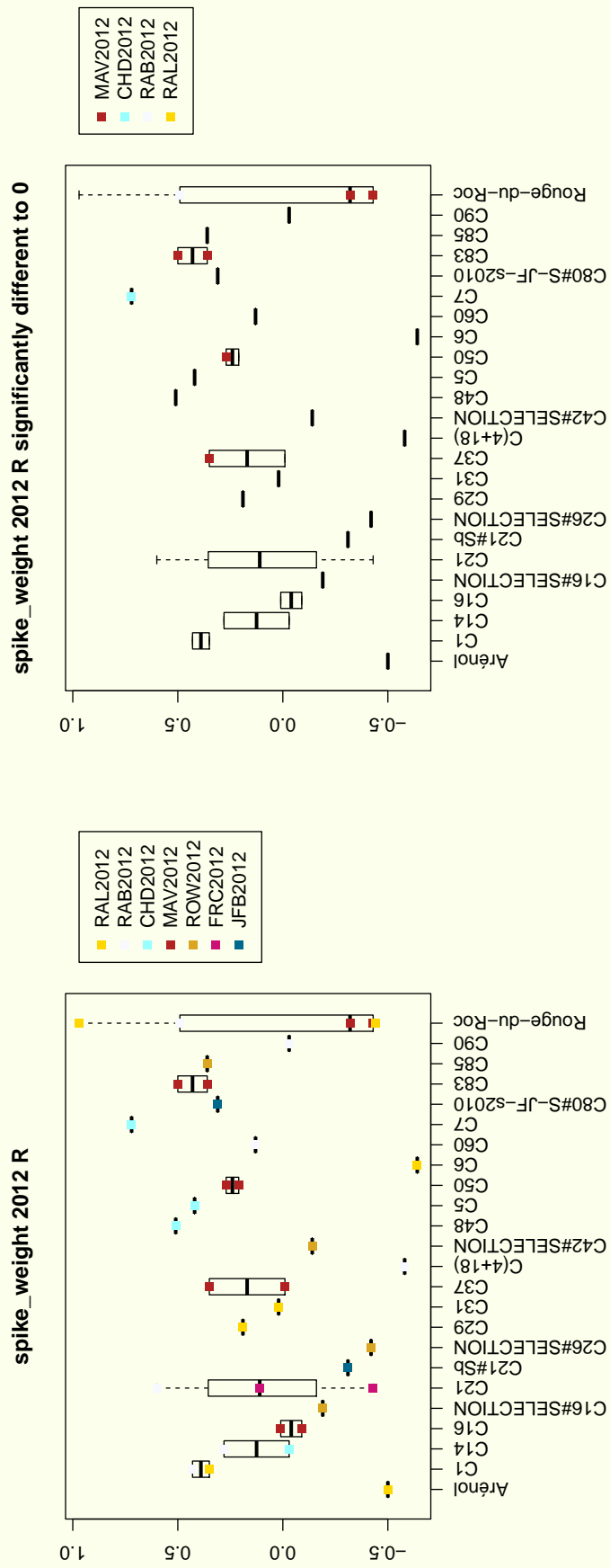




Tableau 2.8 - Pour chaque année, nombre de populations avec les différentiels de sélection  $S$  positifs ( $S+$ ), négatifs ( $S-$ ) ainsi que le nombre de fois où la variance phénotypique augmente ( $var+$ ) ou diminue ( $var-$ ).

Tableau 2.8 – (partie 1/2)

	2010		2011		2012	
	$S$	$var+$ $var-$	$S$	$var+$ $var-$	$S$	$var+$ $var-$
barbes +	<b>22(9)</b>	13(2) 9(8)	67(18)	29(4) 38(27)	26(8)	15(1) 11(4)
barbes -	<b>6(4)</b>	1 5(3)	61(13)	10(5) 51(45)	21(6)	8(1) 13(8)
couleur +	18(2)	11(1) 7(4)	79(39)	33(9) 46(29)	26(16)	13(4) 13(7)
couleur -	10(3)	4(2) 6(4)	48(19)	14(7) 34(22)	21(17)	7(1) 14(7)
courbe +	<b>22(14)</b>	6(2) 16(8)	<b>93(42)</b>	41(4) 52(13)	21(8)	12 9(1)
courbe -	<b>10(2)</b>	6(1) 4(2)	<b>35(11)</b>	12 23(12)	26(12)	10 16(7)
nb d'épillets +	<b>19(2)</b>	9(1) 10(1)	<b>68(51)</b>	36(6) 32(2)	8(3)	2 6(1)
nb d'épillets -	<b>8(0)</b>	6 2	<b>11(7)</b>	8 3	4(2)	2 2

ALP : 1, BRE : 24, CHD : 5, FRC : 2, CHD : 5, EUK : 13, FRC : 2, JMC :  
 JAB : 4, JFB : 10, JMC : 11, MAV :  
 12, JUB : 2, PAJ : 2, RAB : 9, RIH :  
 10, OLR : 6, RAB : 12, RAL : 7, 3, STP : 7  
 ROW : 6

sélectionneur JFB : 17, ROW : 10

Tableau 2.8 – (partie 2/2)

	2010		2011		2012	
	<i>S</i>	<i>var+</i> <i>var-</i>	<i>S</i>	<i>var+</i> <i>var-</i>	<i>S</i>	<i>var+</i> <i>var-</i>
nb d'épillets stériles +	<b>7(1)</b>	5 2	<b>19(6)</b>	8(4) 11	<b>7(0)</b>	6(1) 1
nb d'épillets stériles -	<b>20(6)</b>	7(1) 13(7)	<b>60(30)</b>	37(11) 23(4)	<b>40(19)</b>	6(1) 34(12)
longueur de l'épi +	<b>22(5)</b>	8(1) 14(2)	<b>62(41)</b>	28(3) 34(5)	5(3)	1 4(1)
longueur de l'épi -	<b>6(1)</b>	0 6	<b>17(9)</b>	8(1) 9(1)	3(1)	1 2
poids de l'épi +	<b>22(12)</b>	10 12	<b>111(90)</b>	67(13) 44(4)	<b>39(26)</b>	11(2) 28(3)
poids de l'épi -	<b>5(1)</b>	1 4	<b>17(5)</b>	5 12	<b>8(3)</b>	0 8(2)
pmg +	7(0)		<b>99(0)</b>		30(0)	
pmg -	6(0)		<b>33(0)</b>		35(0)	
protéine +			<b>47(0)</b>		39(0)	
protéine -			<b>23(0)</b>		25(0)	

Les nombres entre parenthèses correspondent au nombre de fois où l'évolution de *S*, *var+* ou *var-*, était significative.

sélectionneur : paysan qui a sélectionné suivi du nombre de sélections.

La différence dans le nombre de populations d'une variable à l'autre est due à des problèmes au champ ou dans la salle technique créant des données manquantes.

**chiffre en gras** : le nombre de populations avec un *S* positif est significativement différent du nombre de populations avec un *S* négatif.

Tableau 2.9 - Pour chaque année, nombre de populations avec les réponses à la sélection  $R$  positifs ( $R+$ ), négatifs ( $R-$ ) ainsi que le nombre de fois où la variance phénotypique augmente ( $var+$ ) ou diminue ( $var-$ ).

	2011			2012		
	$R$	$var+$	$var-$	$R$	$var+$	$var-$
	sélectionneur JFB : 1, ROW : 3			CHD : 4, FRC : 2, JFB : 2, MAV : 5, RAB : 7, RAL : 7, ROW : 5		
barbes +	1(1)		1(1)	17(6)	8(1)	9(6)
barbes -				20(8)	6	14(10)
couleur +				<b>26(13)</b>	15(5)	11(4)
couleur -	1(1)		1(1)	<b>10(5)</b>	5(2)	5(2)
courbe +	1(1)	1(1)		22(7)	15(2)	7
courbe -				14(8)	2	12(4)
nb d'épillets +				<b>10(4)</b>	5	5
nb d'épillets -	1(0)		1	<b>3(0)</b>	3	0
nb d'épillets stériles +	1(0)	1		17(3)	11(3)	6(1)
nb d'épillets stériles -				19(6)	9(1)	10(3)
LLSD +				4(1)	1	3(3)
LLSD -	1(0)	1		11(1)	5(1)	6(1)
longueur de l'épi +				10(2)	7(2)	3(1)
longueur de l'épi -	1(0)	1		5(1)	3	2
poids de l'épi +	1(0)	1		21(6)	14(3)	7(1)
poids de l'épi -				15(2)	8(3)	7
pmg +	3(0)			17(0)		
pmg -	1(0)			19(0)		
protéine +				9(0)		
protéine -	4(0)			14(0)		
hauteur +				7(1)	4	2(1)
hauteur -	1(0)	1		13(1)	7	2

Les nombres entre parenthèses correspondent au nombre de fois où l'évolution de  $R$ ,  $var+$  ou  $var-$ , était significative.

sélectionneur : paysan qui a sélectionné suivi du nombre de sélections.

La différence dans le nombre de populations d'une variable à l'autre est due à des problèmes au champ ou dans la salle technique créant des données manquantes.

**chiffre en gras** : le nombre de populations avec un  $R$  positif est significativement différent du nombre de populations avec un  $R$  négatif.

Tableau 2.10 - Répartition des différents cas de couples  $S$  et  $R$ .

Tableau - 2.10 (partie 1/2)

	$S < 0$ et $R < 0$ et $S > 0$ et $R > 0$		$S < 0$ ou $S > 0$ et $R = 0$		$S > 0$ et $R < 0$ ou $S < 0$ et $R > 0$		$S = 0$ et $R < 0$ ou $R > 0$		$S = 0$ et $R = 0$	
	nb	h2	nb	h2	nb		nb		nb	
barbes	8	0.29; 0.36; 0.59; 0.71; 1.17; 1.88; 5.57; 9.14	7	-0.02; -0.01; 0; 0.04; 0.04; 0.5; 0.67	8		7		5	
couleur	6	0.52; 0.77; 0.83; 1.47; 1.68; 1.81	5	0; 0; 0.11; 0.27; 0.39	10		8		5	
courbe	4	0.98; 1.09; 1.33; 1.52	9	-0.58; -0.58; -0.56; -0.4; - 0.05; 0; 0; 0.11; 0.4	8		10		4	
nb d'épilletts	9	0.06; 0.11; 0.3; 0.48; 0.57; 0.58; 0.81; 0.95; 2.49	0	0	4		0		0	
nb d'épilletts stériles	6	0.18; 0.2; 0.77; 0.8; 1.75; 1.92	0	0	9		0		0	
longueur de l'épi	6	0.1; 0.16; 0.32; 0.47; 1.08; 2.07	0	0	9		0		0	

-Inf : - l'infini, correspond à un cas où  $R = 0$ .Ne sont représentées que les données où le  $S$  et le  $R$  correspondant étaient disponibles.

Tableau - 2.10 (partie 2/2)

	$S < 0$ et $R < 0$ et $S > 0$ et $R > 0$		$S < 0$ ou $S > 0$ et $R = 0$		$S > 0$ et $R < 0$ ou $S < 0$ et $R > 0$		$S = 0$ et $R < 0$ ou $R > 0$		$S = 0$ et $R = 0$	
	nb	h2	nb	h2	nb		nb		nb	
poids de l'épi	14	0.13; 0.25; 0.25; 0.3; 0.31; 0.38; 0.42; 0.43; 0.46; 0.54; 1.33; 1.56; 1.59; 2.21	5	-0.08; -0.04; -0.03; -0.02; 0.03	12		2		1	
pmg	17	0.07; 0.07; 0.16; 0.42; 0.43; 0.46; 0.6; 0.9; 1.13; 1.26; 3.19; 3.45; 4; 5.29; 6.1; 6.13; 8.1	0	0	21		0		0	
protéine	7	0.11; 0.2; 0.36; 0.42; 0.74; 0.87; 2.64	0	0	2		2		0	
pourcentage	33,5		11,3		36,1		12,6		6,5	

### 2.2.3 Discussion

La première partie de la discussion traite de l'aspect méthodologique des analyses. Une deuxième partie traite des résultats du programme de sélection participative.

#### 2.2.3.1 Méthode statistique utilisée et perspectives pour les programmes de sélection participative

L'article en annexe K détaille la méthode statistique et la replace dans les programmes de sélection participative. Notre approche se caractérise par trois aspects.

1. Les paysans ne sont pas contraints par les répétitions et le nombre de populations à tester, chacun peut donc participer selon ses moyens.
2. Les paysans ne sont pas contraints par les choix des autres paysans du réseau. Ils ont une grande liberté dans le choix des populations qu'ils vont semer : une grande diversité de populations peut être évaluée sur le réseau d'essais.
3. La méthode permet d'estimer les effets génétiques sur chaque ferme. Cela permet aux paysans d'avoir accès à une information sur la valeur spécifique de chaque population sur leur ferme, qui regroupe l'effet du génotype et de l'interaction  $G \times E$ . Cette information leur permet de sélectionner parmi les populations.

La modélisation des variances environnementales par une approche bayésienne hiérarchique donne des résultats satisfaisants à condition que les environnements partagent une structure d'essais commune (témoins répétés avec ou sans blocs) et que le nombre d'environnements soit important (au moins 20, selon les variables) (annexe K).

Pour chaque variable et chaque environnement, la méthode fournit pour les populations, une probabilité d'appartenance à des groupes. Plus la différence entre deux populations augmente et plus la probabilité d'appartenance au même groupe diminue, ce qui est cohérent (Figure 2.8).

Cependant, il arrive, dans de rares cas, que la probabilité augmente alors que la différence entre les populations augmente. C'est dû à la méthode statistique utilisée. En effet, en statistique bayésienne, tous les paramètres sont aléatoires (annexe N) et leurs effets sont estimés par les médianes des distributions des posteriors. La médiane est utilisée car elle est plus robuste aux données extrêmes. Dans la suite du texte, nous parlerons de moyennes, plutôt que de médianes, pour simplifier la compréhension. La probabilité que deux moyennes soient égales correspond au nombre de simulation MCMC en commun entre les distributions des deux paramètres. Une distribution avec une grande variance aura donc plus de chance d'être égale à une autre distribution qu'une distribution avec

une variance faible. C'est ce qui est observé pour le poids de l'épi chez FRC2011 où *Rouge-du-Roc* (groupe « *b* ») est entre *C2* et *C12* (tous les deux, groupe « *a* »). Soit  $Pr$ , la probabilité que les deux populations soient égales et  $\bar{X}$  la moyenne de la population  $X$ . Dans ce cas particulier, on a  $\bar{C82} < \bar{C2} < \text{Rouge-}\bar{d}\bar{u}\text{-Roc} < \bar{C12}$  et  $Pr\{C82 = C2\} = 0.054$ ,  $Pr\{C82 = \text{Rouge-du-Roc}\} = 0.030$ ,  $Pr\{C82 = C12\} = 0.038$ . Ceci n'est pas cohérent, la théorie voudrait  $Pr\{C82 = \text{Rouge-du-Roc}\} > Pr\{C82 = C12\}$ . La distribution des valeurs de la population *Rouge-du-Roc* a une variance faible alors que la distribution de la population *C12* a une variance importante. De ce fait, *C82* présente plus de recouvrement avec *C12* qu'avec *Rouge-du-Roc*.

Toutes variables confondues, la méthode permet de différencier au moins deux groupes avec un risque global  $\alpha$  de 0.1 dans 63% des cas. On monte à 78% des cas avec un risque global  $\alpha$  de 0.2. La méthode permet donc, avec un risque modéré, de discriminer les populations sur chaque ferme sur la base des mesures phénotypiques.

Certaines variables comme le poids de l'épi, le poids de mille grains ou les protéines ont des effets génétiques avec des variances larges. Notre modèle arrive moins bien à discriminer des populations pour ces variables que pour d'autres comme la hauteur ou LLSD qui ont des variances plus petites. Les variables avec une plus large variance, sont plus sensibles à la variation environnementale intra-ferme. Plus la variance résiduelle est importante et plus la précision de l'estimation des effets population sur les fermes est faible.

Augmenter le nombre de répétitions permet d'être plus précis, de mieux estimer l'héritabilité et donc, de mieux discriminer les populations. Si l'on veut gagner en précision, il faut sans doute ajouter des témoins répétés pour mieux contrôler l'hétérogénéité environnementale intra-ferme et mieux estimer les variances associées. L'idéal pour tester cette hypothèse serait de simuler des données pour comprendre un peu mieux les limites et les potentialités du modèle.

La compétition des plantes entre micro-parcelles peut entraîner une surestimation de la variance génétique<sup>121</sup>. Une surestimation de la variance génétique engendre une surestimation de l'héritabilité c'est à dire de la discrimination des populations, le risque global  $\alpha$  serait donc sous-estimé. Sur le terrain, des différences dues à la compétition pour les ressources azotées (passage d'un motoculteur dans les allées qui a favorisé la minéralisation de l'azote) ont parfois été observées entre les plantes d'une même parcelle. Dans de tels cas, les mesures ont été réalisées sur les plantes centrales dans chaque parcelle

pour éviter ce biais.

On pourrait aller plus loin dans les analyses et étudier plus en détails les données issues des fiches de notation. Les notes globales pourraient être analysées de la même manière que les notes de barbe, de couleur ou de courbe. Même chose pour les notes de verse des variétés. Malheureusement, peu de ces notes sont disponibles. Le chapitre 4 revient sur ce point.

Les comparaisons de moyennes permettent également d'avoir accès à l'effet année en comparant le comportement des populations d'une année sur l'autre.

### 2.2.3.2 Résultats du programme de sélection participative

#### 2.2.3.2.1 La sélection inter-population

Pour étudier la sélection inter-population, je me suis intéressé aux groupes de significativité auxquels appartenaient les populations non semées l'année suivante (sélection inter-population négative).

Ces résultats doivent être replacés dans le contexte. Lors de ces sélections, les paysans n'avaient pas accès aux comparaisons de moyennes. En effet, la méthode a été développée à la fin de la thèse. De plus, les populations étudiées ne sont pas forcément représentatives des sélections de tous les paysans.

Toutefois, il y a une cohérence entre les choix des paysans et les résultats de cette analyse. Il y a une nette tendance à ce que les populations non semées l'année suivante aient présenté en moyenne une valeur faible pour les caractères d'intérêt pour les paysans : protéine, poids de mille grains, poids de l'épi, nombre d'épillets. De même, les populations écartées étaient en général plus hautes, ce qui est cohérent avec le souhait des paysans d'avoir des blés qui résistent à la verse. Ces résultats sont encourageants et l'information apportée par les comparaisons de moyennes devrait les conforter.

#### 2.2.3.2.2 La sélection intra-population

Pour optimiser la sélection, il faut affiner le modèle de description de la variabilité génétique<sup>118</sup>. Pour mieux comprendre la réponse à la sélection au cours des différentes générations, on pourrait estimer les différentes composantes de l'expression de la covariance entre les parents et les descendants ( $cov(O, P)$ ). Dans l'expression 2.5, il y a 5 éléments à estimer : la variance génétique additive ( $\sigma_A^2$ ), la covariance entre les effets d'additivité et de dominance à l'état homozygote ( $\sigma_{AD_o}$ ), la variance génétique de dominance ( $\sigma_D^2$ ), la variance des résidus de dominance entre gènes identiques ( $\sigma_{D_o}^2$ ) et la somme des carrés de la dépression de consanguinité ( $\sum_l D_{0,l}^2$ ). Pour résoudre ce système à cinq inconnues, il



faut cinq équations. Chaque équation correspondant aux covariances entre différentes générations d'apparentés pour un germplasm donné. Nous aurions les données pour résoudre ces équations mais nous ne les avons pas résolues du fait que l'estimation des variances est souvent peu précise et que cela apporterait peu d'information au programme<sup>118</sup>.

L'expression 2.5 de  $cov(O, P)$  ne prend pas en compte les effets d'épistasie, c'est-à-dire les interactions entre gènes situés à des locus différents. Chez les espèces autogames, l'épistasie additive  $\times$  additive contribue à une part non négligeable de la variance transmise à la descendance<sup>120</sup>. Chez les plantes autogames le déséquilibre de liaison est important du fait de l'autofécondation<sup>46</sup> et les parents transmettent à leur descendance des régions entières de chromosomes qui ne seront pas cassés par la recombinaison.

Selon les caractères, la part de variance génétique due aux effets d'épistasie peut varier. Par exemple, chez des espèces autogames telles que le blé ou l'orge, le rendement, la précocité d'épiaison ou la résistance à l'oïdium montrent une forte variance d'épistasie additive  $\times$  additive<sup>120</sup>. Cela est corroboré par le déterminisme moléculaire de ces caractères qui dépendent des interactions entre gènes<sup>118;129;153;202</sup>. A l'inverse, la hauteur, déterminée par des gènes majeurs à effets forts notamment liés au nanisme, montre peu d'effet d'épistasie<sup>120</sup>. La variance génétique transmise à la génération suivante comporte donc une composante de variance additive et une composante de variance d'épistasie additive  $\times$  additive. Plus les générations seront avancées et plus la part d'épistasie additive  $\times$  additive va croître, du simple fait du taux de recombinaison efficace qui diminue l'hétérozygotie. Dans notre cas, les sélections sont faites sur des générations avancées (F4 ou F5), en négligeant les 5% de croisements<sup>81</sup>, les plantes sont très homozygotes ( $F_{IS} = 0.87$  et  $0.94$ ). Pour une sélection en F4 ou F5, sans épistasie, on s'attend à ce que les parents transmettent autour de  $1.8\sigma_A + 1.8\sigma_{AD_O} + 0.9\sigma_{D_O}^2$  à leur descendance.

Ainsi, si la covariance parent-descendant est quasi maximale à ces générations, les limites de la réponse à la sélection vont émerger du manque de variabilité génétique présente pour le caractère soumis à sélection dans la population ainsi que du faible contrôle des conditions environnementales qui diminue l'héritabilité du caractère.

Pour la sélection massale intra-population, il est difficile de limiter la variabilité environnementale intra-micro-parcelles et son impact. De plus, cette variance résiduelle risque d'augmenter à cause de la compétition entre les plantes<sup>99</sup>, et également à cause de la compétition entre les plantes et les adventices qui ne sont pas rares en agriculture biologique. Il est important de sélectionner des épis dans des zones très différentes de la micro-parcelle afin de minimiser les corrélations micro-environnementales entre les épis choisis, en effet,

l'absence de corrélation entre les plantes et l'environnement est une condition nécessaire pour que les covariances de type génotype-environnement s'éliminent dans l'expression de  $cov(O, P)$  (annexe O).

Après les résultats de la sélection de *JFB* en génération précoce, il était intéressant d'élargir les analyses et de regarder l'effet de la sélection massale paysanne effectuée par plusieurs paysans dans différentes fermes. Les objectifs ici sont de caractériser la manière dont les paysans sélectionnent, si différents paysans sélectionnent dans le même sens (évolution des moyennes), de voir si chacun sélectionne des épis plutôt différents entre eux ou plutôt homogènes (évolution des variances).

Dans 11,3% des cas, des valeurs de  $S$  non nulles, étaient associées à des valeurs de  $R$  nulles. La variation sélectionnée dans ces populations n'a pas été transmise aux descendants. Elle devait donc être principalement environnementale : l'héritabilité au sens strict était quasi nulle. Dans 33,5% des cas, la réponse  $R$  allait dans le même sens que le différentiel de sélection  $S$ , mais toutes les variations n'étaient pas significatives. Dans ce dernier cas, certaines héritabilités ont été estimées supérieures à 1.

Dans 36,1% des cas le différentiel de sélection  $S$  et la réponse  $R$  n'avaient pas le même signe, c'est à dire que la pression de sélection s'est exercée pour faire varier le caractère dans un sens et qu'il a finalement répondu dans l'autre sens. De plus, dans 12,6% des cas, une valeur de  $S$  nulle était associée à une valeur de  $R$  non nulle.

Plusieurs explications sont vraisemblables pour expliquer ces deux derniers points. Tout d'abord, les estimations de  $S$  ou de  $R$  ne sont pas correctes, c'est à dire que les différences de moyennes entre la version vrac et la version sélectionnée sont mal estimées. Cela peut être dû au fait que les épis du vrac ne sont pas représentatifs de la variabilité de la parcelle. Lors des réunions avec les paysans, cette interrogation est revenue souvent. En effet, l'idéal est de mesurer au hasard les épis du brin maître. En pratique, il est compliqué de repérer le brin maître et le hasard est parfois influencé ... Pour l'année 2012-2013, 50 épis ont été mesurés plutôt que 25 afin de tester si les moyennes seraient plus représentatives et les réponses plus souvent cohérentes avec les différentiels de sélection.

Une autre hypothèse est que  $S$  soit dû à de la variance environnementale principalement. La réponse observée dans la descendance peut alors être une réponse corrélative liée à la sélection et l'évolution d'un autre caractère. Une approche multicritère serait alors plus judicieuse.

Enfin, il est possible que des interactions génotype  $\times$  année, au sein des populations,

soient présentes, ce qui pourrait limiter la réponse à la sélection. Par exemple, si la sélection se fait une année sèche, les individus qui se comportent bien dans ces conditions vont présenter de plus gros épis, être plus vigoureux et donc être sélectionnés. Si l'année suivante est une année humide, les descendants issus du bouquet de sélection risquent de se comporter moins bien pour certains caractères de production et de vigueur. Il est donc important de ne pas appliquer un taux de sélection trop fort une année donnée.

Pour la majorité des caractères, le différentiel de sélection  $S$  va dans le même sens pour tous les paysans à quelques exceptions près. Les paysans choisissent dans les populations les épis les plus gros, les plus longs, courbés et de couleur plus rouge, avec beaucoup d'épillets fertiles. En discutant avec les paysans, la courbure de l'épi est un critère signe de maturité et de « bonne santé » de l'épi. De plus, les paysans apprécient les blés plutôt rouges car ils pensent qu'il y a un lien entre la couleur, la composition du grain, la qualité nutritionnelle ou le goût, ainsi que pour des questions d'esthétisme dans les champs. Un différentiel plutôt positif s'observe également pour le poids de mille grains et le taux de protéine. Cela est remarquable car ces deux derniers critères ne sont pas visibles au champ. Les variances phénotypiques des bouquets restent importantes, ce qui peut être relié au fait que les paysans ont sélectionné autour de 25 épis ou plus à chaque fois.

Sachant le différentiel de sélection, il est intéressant de regarder si la direction de la sélection par le bouquet ( $S$ ) se retrouve la génération suivante dans la réponse ( $R$ ).

Parmi les 183 bouquets de sélection, nous n'avons pu analyser que 37 réponses à la génération suivante. Ce faible nombre est dû au fait que des sélections ont été mélangées pour créer de nouvelles populations et que certains paysans n'ont pas suivi les plans de semis pour étudier la réponse, soit par manque d'espace (dans ce cas, les bouquets sont semés sans le vrac correspondant) ou par manque d'intérêt pour cette expérience. Néanmoins, à la suite des réunions sur les semis de 2012 et aux discussions, de nombreux paysans ont été motivés pour mener l'expérience en 2013. Toutefois ces données n'ont pu être analysées dans le temps imparti pour la thèse.

Pour deux caractères seulement, le nombre de populations avec une réponse positive s'est révélé significativement différente du nombre de populations avec une réponse négative : le nombre d'épillets et la couleur des épis. Le nombre d'épillets a augmenté dans les descendances issues de la sélection comparées aux descendances non issues de la sélection, ce qui est cohérent avec les bouquets de 2011 qui vont dans ce sens également. De même, la couleur rouge augmente dans les populations issues de sélection. De plus, les populations issues de sélection ont tendance à avoir des plantes plus courtes, une teneur

en protéine faible, des épis plus longs et plus lourds.

58% des valeurs de  $R$  sont été estimées avec le modèle bayésien hiérarchique alors que 42% l'ont été avec un test de Student. Le modèle bayésien est peut-être trop strict, il sur-estime la variance qu'il pourrait y avoir entre deux parcelles placées côte à côte. De la même façon, il sous-estime les différences entre populations. A l'inverse, le test de Student fait l'hypothèse qu'il n'y a pas de variation environnementale entre les deux parcelles, ce qui peut sur-estimer les différences entre populations.

## 2.3 Analyse des interactions $G \times E$

Lors d'une analyse génotype  $\times$  environnement ( $G \times E$ ), on peut chercher à savoir quel génotype est le meilleur tous environnements confondus, quel génotype est le meilleur pour un environnement spécifique, quels génotypes sont stables sur l'ensemble des environnements<sup>147</sup>.

Dans cette partie, je présente tout d'abord la démarche utilisée pour étudier l'interaction  $G \times E$ . Des modèles classiques de l'analyse de l'interaction  $G \times E$  ont été testés sur des jeux de données équilibrés : Additive Main Effect and Multiplicative Interaction (AMMI), GGE et EGE<sup>11;113;114;259</sup>. Si ces méthodes ont donné des résultats satisfaisants sur des sous-jeux de données équilibrées, elles ne fonctionnaient pas, ou du moins donnaient des résultats peu cohérents, sur les données complètes, c'est à dire très déséquilibrées avec 95% des combinaisons environnement  $\times$  germplasm manquantes.

Dans une première partie, les méthodes sont présentées ainsi que les problèmes rencontrés. Dans une deuxième partie, afin d'utiliser toutes les données, un modèle plus parcimonieux a été utilisé : le modèle de Finlay Wilkinson<sup>100</sup>. Les paramètres de ce modèle sont estimés par une approche bayésienne hiérarchique du même type que pour les comparaisons de moyennes (partie 2.2).

### 2.3.1 Méthodes

#### 2.3.1.1 Travail préliminaire sur les données

Que ce soit sur des données équilibrées ou non, dans les fermes régionales, les données ont été corrigées pour l'effet bloc. Pour les fermes satellites, les données n'ont pas été corrigées car il n'y avait pas de bloc.

Dans certains sous-jeux de données équilibrés, toutes les combinaisons  $G \times E$  existent et il n'y a plus de confusion d'effets. Sur de tels jeux de données, on peut appliquer les méthodes « classiques » d'analyse de l'interaction parmi laquelle la méthode AMMI.

### 2.3.1.1.1 modèle AMMI

Le modèle AMMI repose sur deux analyses :

(i) **une ANOVA** avec le modèle suivant :

$$Y_{ijk} = \mu + \alpha_i + \theta_j + \eta_i\nu_j + \varepsilon_{ijk}; \quad \varepsilon_{ijk} \sim \mathcal{N}(0, \sigma^2) \quad (2.8)$$

Avec,  $Y_{ijk}$  la valeur phénotypique de l'individu  $k$  dans le germplasm  $i$  dans l'environnement  $j$ ,  $\mu$  la moyenne générale,  $\alpha_i$  l'effet lié au germplasm  $i$ ,  $\theta_j$  l'effet lié à l'environnement  $j$ ,  $\eta_i\nu_j$  l'interaction germplasm  $\times$  environnement,  $\varepsilon_{ijk}$  la résiduelle.

(ii) **une ACP** qui étudie plus spécifiquement les interactions  $G \times E$ . Cela revient à poser

$$\eta_i\nu_j = \sum_n^N \lambda_n \gamma_{in} \omega_{jn}$$

On peut aussi écrire :

$$\eta_i\nu_j = \sum_n^N (\sqrt{\lambda_n} \gamma_{in}) (\sqrt{\lambda_n} \omega_{jn})$$

Avec,  $\eta_i\nu_j$  l'interaction du germplasm  $i$  avec l'environnement  $j$ ,  $N$  le nombre de dimension (composante de l'Analyse en Composantes Principales (ACP)) qui a comme valeur maximale le nombre d'environnement,  $\lambda_n$  la valeur propre pour la composante  $n$ ,  $\gamma_{in}$  le vecteur propre (le score) pour le germplasm  $i$  pour la composante  $n$ ,  $\omega_{jn}$  le vecteur propre (le score) pour l'environnement  $j$  pour la composante  $n$ .

On parle alors de données doubles-centrées sur l'environnement  $E$  et la génétique  $G$ . L'ACP permet de décomposer l'interaction et de décrire sa structure. Cela permet, pour un nombre de composantes faibles (une, deux, voire trois) de représenter les valeurs prises par  $G$  et  $E$  dans la nouvelle base, en gardant à l'esprit que le produit des deux valeurs permet d'estimer l'interaction dans la limite de l'explication des axes de l'ACP. Si le premier axe ( $n = 1$ ) explique 100% de la variabilité, alors le produit  $(\sqrt{\lambda_1} \gamma_{i1})(\sqrt{\lambda_1} \omega_{j1})$  est égal à l'effet de l'interaction prédit par le modèle 2.8 (en ayant pour contraintes que la somme des effets est égale à 0, contrainte sur l'Analyse de variance (ANOVA) et sur la matrice d'interaction). La valeur propre  $\lambda_n$  correspond à la somme des carrés capturés

par la dimension  $n$ <sup>113</sup>. Dans l'ACP sur la matrice d'interaction, les environnements sont les variables et les variétés sont les individus. L'ACP sur la matrice d'interaction permet de voir quelles variétés sont les plus stables (contribuent le moins à l'interaction  $G \times E$ ), et quelles variétés interagissent le plus avec quel environnement. On peut également voir quels environnements se ressemblent en terme d'interaction.

Les écovalences de Wricke<sup>258</sup> estiment la part de variation de l'interaction prise soit par un germplasm soit par un environnement. C'est un indicateur de stabilité : une écovalence faible signifie une interaction faible, donc plus de stabilité. Un génotype stable signifie qu'il interagit peu, la valeur phénotypique observée du génotype, dans un environnement donné, sera la somme de l'effet génétique et de l'effet lié à l'environnement dans lequel il se trouve. L'écovalence des germplasm  $i$  s'écrit  $W_i = \sum_i^n (\eta_i \nu_j)^2$ . Un germplasm qui présente peu d'interaction avec les environnements aura une écovalence proche de 0. L'écovalence des environnements  $j$  s'écrit  $W_j = \sum_j^n (\eta_i \nu_j)^2$ . Un environnement qui présente peu d'interaction avec les variétés aura une écovalence proche de 0. Les écovalences sont représentées en fonction des effets moyens (i.e. moyenne brute à la moyenne générale  $\mu$  près) par germplasm et par environnement.

On peut faire un graphique qui analyse un peu plus les effets génétiques, c'est à dire  $\alpha_i$ , et  $\gamma_{i1}$  les valeurs de l'interaction expliquée par  $G$  dans la première dimension de l'ACP<sup>147</sup>.

**Plus généralement, le modèle AMMI peut s'écrire :**

$$Y_{ijk} = \mu + \alpha_i + \theta_j + \sum_n^N \lambda_n \gamma_{in} \omega_{jn} + \varepsilon_{ijk}; \quad \varepsilon_{ijk} \sim \mathcal{N}(0, \sigma^2) \quad (2.9)$$

Avec,  $Y_{ijk}$  la valeur phénotypique de l'individu  $k$  du germplasm  $i$  dans l'environnement  $j$  pour une variable donnée,  $\mu$  la moyenne générale,  $\alpha_i$  l'effet lié au germplasm  $i$ ,  $\theta_j$  l'effet lié à l'environnement  $j$ ,  $N$  le nombre de dimension (composante de l'ACP),  $\lambda_n$  la valeur propre pour la composante  $n$ ,  $\gamma_{in}$  le vecteur propre (le score) pour le germplasm  $i$  pour la composante  $n$ ,  $\omega_{jn}$  le vecteur propre (le score) pour l'environnement  $j$  pour la composante  $n$ .  $\varepsilon_{ijk}$  la résiduelle.

### 2.3.1.1.2 modèle GGE

Le modèle GGE est le suivant :

$$Y_{ijk} = \mu + \theta_j + \sum_n^N \lambda_n \gamma_{in} \omega_{jn} + \varepsilon_{ijk}; \quad \varepsilon_{ijk} \sim \mathcal{N}(0, \sigma^2)$$

Où  $Y_{ijk}$  est la valeur phénotypique pour une variable donnée;  $\mu$  est la moyenne générale;  $\alpha_i$  est l'effet lié à l'environnement  $j$ ;  $N$  est le nombre de dimension (composante de l'ACP);  $\lambda_n$  est la valeur propre pour la composante  $n$ ;  $\gamma_{in}$  est le vecteur propre (le score) pour le germplasm  $i$  pour la composante  $n$ ;  $\omega_{jn}$  est le vecteur propre (le score) pour l'environnement  $j$  pour la composante  $n$ ;  $\varepsilon_{ijk}$  est la résiduelle.

Il est du même type que la modèle AMMI sauf que la matrice sur laquelle est faite l'ACP est centrée sur l'environnement : les effets germplasm et interactions sont confondus.

Ce modèle permet une analyse un peu différente de celle d'AMMI et permet de trouver les environnements où les germplasm se comportent le mieux (le « which won where »<sup>vi</sup> de Gauch et al.<sup>114</sup>).

### 2.3.1.1.3 Le modèle EGE

Le modèle EGE est le suivant :

$$Y_{ijk} = \mu + \alpha_i + \sum_n^N \lambda_n \gamma_{in} \omega_{jn} + \varepsilon_{ijk}; \quad \varepsilon_{ijk} \sim \mathcal{N}(0, \sigma^2) \quad (2.10)$$

Où  $Y_{ijk}$  est la valeur phénotypique pour une variable donnée;  $\mu$  est la moyenne générale;  $\alpha_i$  est l'effet lié au germplasm  $i$ ;  $N$  est le nombre de dimensions (composantes de l'ACP);  $\lambda_n$  est la valeur propre pour la composante  $n$ ;  $\gamma_{in}$  est le vecteur propre (le score) pour le germplasm  $i$  pour la composante  $n$ ;  $\omega_{jn}$  est le vecteur propre (le score) pour l'environnement  $j$  pour la composante  $n$ ;  $\varepsilon_{ijk}$  est la résiduelle.

Il est du même type que le modèle AMMI sauf que la matrice sur laquelle est faite l'ACP est centrée sur le germplasm : les effets environnements et interactions sont confondus.

Ce modèle permet de regrouper les environnements similaires.

### 2.3.1.1.4 Problèmes rencontrés

Sur les données déséquilibrées, que cela soit pour les analyses AMMI, GGE ou EGE, le problème vient du déséquilibre du jeu de données et son analyse. Ces modèles travaillent sur les matrices d'interactions. Lorsqu'une combinaison génotype  $\times$  environnement n'est pas présente, dans le jeu de données, il faut l'estimer.

Les données manquantes doivent être remplacées pour que l'ACP puisse fonctionner. Pour cela, trois solutions ont été testées :

1. remplacer les données manquantes par la moyenne de colonnes, qui correspondent

---

vi. « qui gagne où »

- aux environnements ;
2. estimer les données manquantes par cross validation avec la fonction `imputePCA` du package `MissMDA` ;
  3. remplacer les données manquantes par 0, faisant l'hypothèse qu'en l'absence de données, l'interaction est nulle.

Ces méthodes reviennent à analyser une matrice d'interaction, où un nombre important des données (95% dans notre cas) ont été estimées. Les analyses sont donc basées sur un jeu de données qui n'a que 5% de données « réelles ».

De plus, le déséquilibre induit des confusions d'effet et peut amener à des interprétations fausses sur la contribution des germplasm ou des environnements à l'interaction. J'ai testé les différents modèles présentés précédemment sur des jeux de données équilibrés, qui correspondaient aux fermes régionales et aux témoins. Avec de telles analyses, beaucoup d'informations étaient perdues. De même, seuls les paysans, dont les fermes se trouvaient « dans le bon jeu de données » pouvaient avoir accès à une information fiable (si les sous-jeux de données sont représentatifs de toutes les données disponibles ...).

Au vu du déséquilibre des données, nous avons choisi d'utiliser le modèle de Finlay Wilkinson<sup>176</sup>. Ce modèle permet d'estimer des effets germplasm, fermes et une sensibilité à l'interaction pour chaque germplasm. Un tel modèle ne va pas dans les détails de l'interaction comme le ferait un modèle de type AMMI<sup>114</sup>. Cependant, il est plus parcimonieux et donc moins sensible au déséquilibre des données et permet de retirer beaucoup d'informations. Comme précédemment dans la partie 2.2, nous utilisons un modèle bayésien hiérarchique pour estimer les différents paramètres<sup>203</sup>.

### 2.3.1.2 L'analyse de l'interaction avec le modèle bayésien hiérarchique

L'article en anglais « Modélisation hiérarchique des interactions génotype × environnement pour des expérimentations flexibles de sélection décentralisée et participative » détaille la méthode statistique et la replace dans le contexte de la sélection participative (annexe L).

#### 2.3.1.2.1 Le modèle de Finlay Wilkinson

$Y_{ij}$ , la valeur phénotypique moyenne pour une variable donnée pour le germplasm  $i$  et l'environnement  $j$ , est modélisée par :

$$Y_{ij} = \alpha_i + \theta_j + \eta_i \theta_j + \varepsilon_{ij} \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_{ij}^2) \quad (2.11)$$



Où,  $\alpha_i$ , l'effet lié au germplasm  $i$ ;  $\theta_j$ , l'effet lié à l'environnement (combinaison ferme  $\times$  année)  $j$ ;  $\eta_i\theta_j$ , l'interaction germplasm  $\times$  environnement;  $\varepsilon_{ij}$ , la résiduelle, qui correspond à l'interaction qui n'est pas expliquée par  $\eta_i\nu_j$ .

On peut aussi écrire le modèle 2.11 :

$$Y_{ij} = \alpha_i + (1 + \eta_i)\theta_j + \varepsilon_{ij} \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_e^2)$$

Il vient,

$$Y_{ij} = \alpha_i + \beta_i\theta_j + \varepsilon_{ij}; \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_e^2) \quad (2.12)$$

Où  $\beta_i = (1 + \eta_i)$  est la sensibilité du germplasm  $i$  aux environnements.  $\beta$  varie autour de 1. Nous avons utilisé le modèle 2.12, qui est celui de Finlay Wilkinson<sup>100</sup>. Plus  $\beta_i$  est proche de 1 et plus le germplasm  $i$  est stable.

Au vu du déséquilibre des données et de leur grand nombre, nous avons résolu ce modèle avec une approche hiérarchique bayésienne.

Chaque paramètre est issu d'une distribution a priori :

$$\alpha_i \sim \mathcal{N}(\mu, \sigma_v^2); \quad \beta_i \sim \mathcal{N}(1, \sigma_s^2); \quad \theta_j \sim \mathcal{N}(0, \sigma_l^2); \quad \sigma_e^{-2} \sim \mathcal{Gamma}(10^{-6}, 10^{-6})$$

La moyenne des  $\beta_i$  est de 1, c'est à dire que la somme des interactions est nulle. Alors qu'un prior vague a été utilisé pour  $\sigma_e$ , nous avons utilisé des priors hiérarchiques pour les variances de  $\alpha$ ,  $\beta$ ,  $\theta$ , ainsi que la moyenne de  $\alpha$  ( $\mu$ ). Des priors vagues sont choisis pour les hyperparamètres  $\mu$ ,  $\sigma_v^2$ ,  $\sigma_s^2$  et  $\sigma_l^2$  :

$$\mu \sim \mathcal{N}(\nu, \nu^2); \quad \sigma_v \sim \mathcal{Uniforme}(0, \nu); \quad \sigma_s \sim \mathcal{Uniforme}(0, 1); \quad \sigma_l \sim \mathcal{Uniforme}(0, \nu)$$

où  $\nu$  est la moyenne arithmétique des données :  $\nu = \sum_{ij} Y_{ij}/n$  où  $n$  est le nombre d'observations. La distribution des variances est bornée à 0, car la variance est positive par définition.

Le modèle 2.12 permet d'estimer pour chaque germplasm  $\alpha_i$  et  $\beta_i$  ainsi que pour chaque environnement,  $\theta_j$ .

Les distributions des posteriors sont estimées de la même manière que dans le modèle 2.1 qui a été utilisé dans partie 2.2.1.1.1. Les valeurs initiales des chaines MCMC sont prises au hasard pour les paramètres sauf pour  $\mu$ ;  $\sigma_v$  et  $\sigma_l$  dont les valeurs initiales ont été estimées à la suite du modèle additif (c'est à dire avec  $\beta = 1$ ).

Pour chaque paramètre  $\alpha_i$ ,  $\beta_i$  ou  $\theta_j$ , les différences significatives entre les paramètres ont été estimées de la même manière qu'avec le modèle 2.1 dans la partie 2.2.1.1.1. Les probabilités que les paramètres soient égaux deux à deux ont été mises dans une matrice  $M$ . Des groupes de paramètres similaires peuvent être créés avec un seuil donné. Dans notre cas, on a utilisé l'erreur de type 1  $\alpha = 0.05/n$  (avec  $n$  le nombre de paramètres à comparer).

Pour chaque facteur (i.e.  $\alpha$ ,  $\beta$  et  $\theta$ ), les estimations de chaque paramètre sont placées dans une matrice  $A$  de dimension  $P \times V$ , avec  $P$  le nombre de niveaux pour un facteur donné et  $V$  le nombre de variables (11 dans notre cas). Il y a autant de matrices  $A$  que de facteurs étudiés, c'est à dire trois.

### 2.3.1.2.2 Détecter des groupes de germplasm ou d'environnements

Le modèle de Finlay Wilkinson a fourni une matrice  $A$  pour chaque facteur étudié : l'effet du germplasm,  $\alpha$ , l'effet de l'environnement,  $\theta$  et la sensibilité des germplasm à l'interaction,  $\beta$ .

Nous avons réalisé une analyse multivariée sur chaque matrice  $A$  à l'aide d'une ACP. Ainsi, tous les paramètres estimés (tous les niveaux de chaque facteur) pour toutes les variables ont été analysés en même temps. Pour la suite de l'analyse, nous avons conservé les coefficients des variables sur les deux premiers axes de l'ACP. Cela permettait de s'affranchir du « bruit » qui était contenu dans les autres axes de l'ACP. Des *K-means*, basées sur la distance euclidienne, ont ensuite été calculées sur ces nouvelles valeurs pour créer des groupes. Le nombre de groupes retenu était celui qui maximisait la variance entre les groupes et minimisait la variance au sein des groupes. Pour cela, on a cherché la valeur la plus petite de la formule suivante :  $\frac{\Delta(Q)}{\Delta(Q+1)}$  avec  $\Delta(Q)$  la différence de variance entre  $Q$  groupes et  $Q-1$  groupes et  $\Delta(Q+1)$  la différence de variance entre  $Q+1$  groupes et  $Q$  groupes. Les fonctions R PCA et HCPC du package FactoMineR sont utilisées<sup>131</sup>.

## 2.3.2 Résultats et discussion

Les résultats sur les jeux de données équilibrés avec les méthodes AMMI, GGE et EGE ne sont pas présentés. J'ai réalisé les modèles classiques de type AMMI et GGE sur des sous jeux de données équilibrés. J'ai codé ces analyses dans R, ce qui m'a permis de mieux comprendre l'analyse des interactions  $G \times E$ . Sur les jeux de données équilibrés, les résultats étaient intéressants : certains environnements se comportaient de la même manière toutes les années, d'autres répondaient différemment. En augmentant le déséquilibre, les

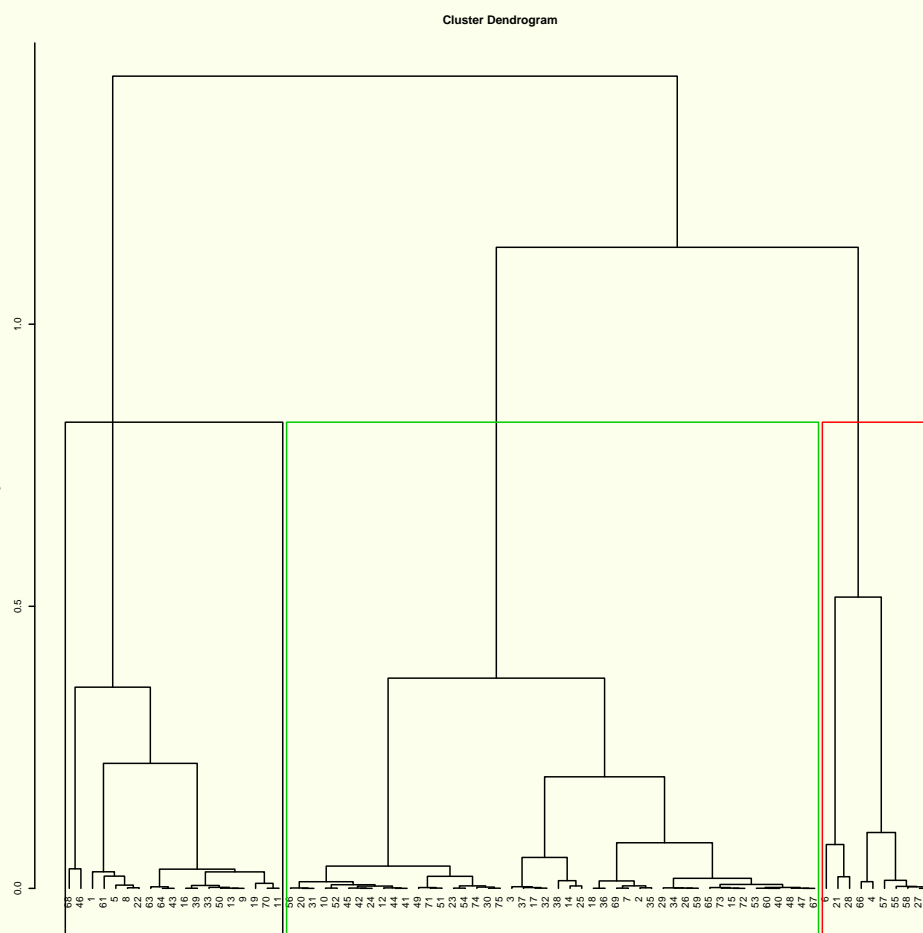
résultats changeaient et certaines valeurs aberrantes émergeaient. Les méthodes d'estimation des données manquantes pouvaient amener à considérer des interactions nulles. Supposer cela, sachant leur importance dans la sélection décentralisée, ne me convenait pas. Je me suis donc concentré sur l'analyse avec le modèle bayésien hiérarchique. L'article en annexe L présente les résultats et la discussion du modèle hiérarchique de Finlay Wilkinson. J'illustre quelques résultats dans cette partie.

Le modèle donne deux informations : les effets de chaque paramètre pour chaque variable, ceux-ci sont représentés sous forme graphique comme pour les comparaisons de moyennes, et des dendrogrammes qui synthétisent l'analyse multivariée. Des différences ont été trouvées entre les  $\alpha_i$ . De même, pour les  $\theta_j$ , des groupes significativement différents ont été détectés. Cependant, pour les  $\beta_i$ , aucune différence n'a été détectée, sauf pour le poids de l'épi où deux groupes ont été identifiés. Il est possible que le déséquilibre soit trop important pour détecter des différences.

Un exemple de résultats d'analyse multivariée est donné avec la Figure 2.15 pour les effets environnement  $\theta$ .

Concernant les environnements, il n'y a pas de tendance systématique à ce que les différentes années sur une même ferme soient regroupées. Toutefois, j'ai remarqué que les trois années (2010, 2011 et 2012) de la ferme *ROW* se trouvaient dans le même groupe, indiquant un comportement plus constant dans le temps. D'un autre côté, les quatre années d'expérimentations sur la ferme de *RAB* se sont retrouvées dans trois groupes différents. Une analyse plus détaillée en relation avec les conditions climatiques pourra être réalisée ultérieurement, l'incorporation des données 2013 nous permettra peut être de tirer des conclusions plus claires.

Aucune tendance n'a été observée dans le regroupement des germplasms selon leurs types ou leurs parents.

Figure 2.15 - Groupes créés pour  $\theta$ .

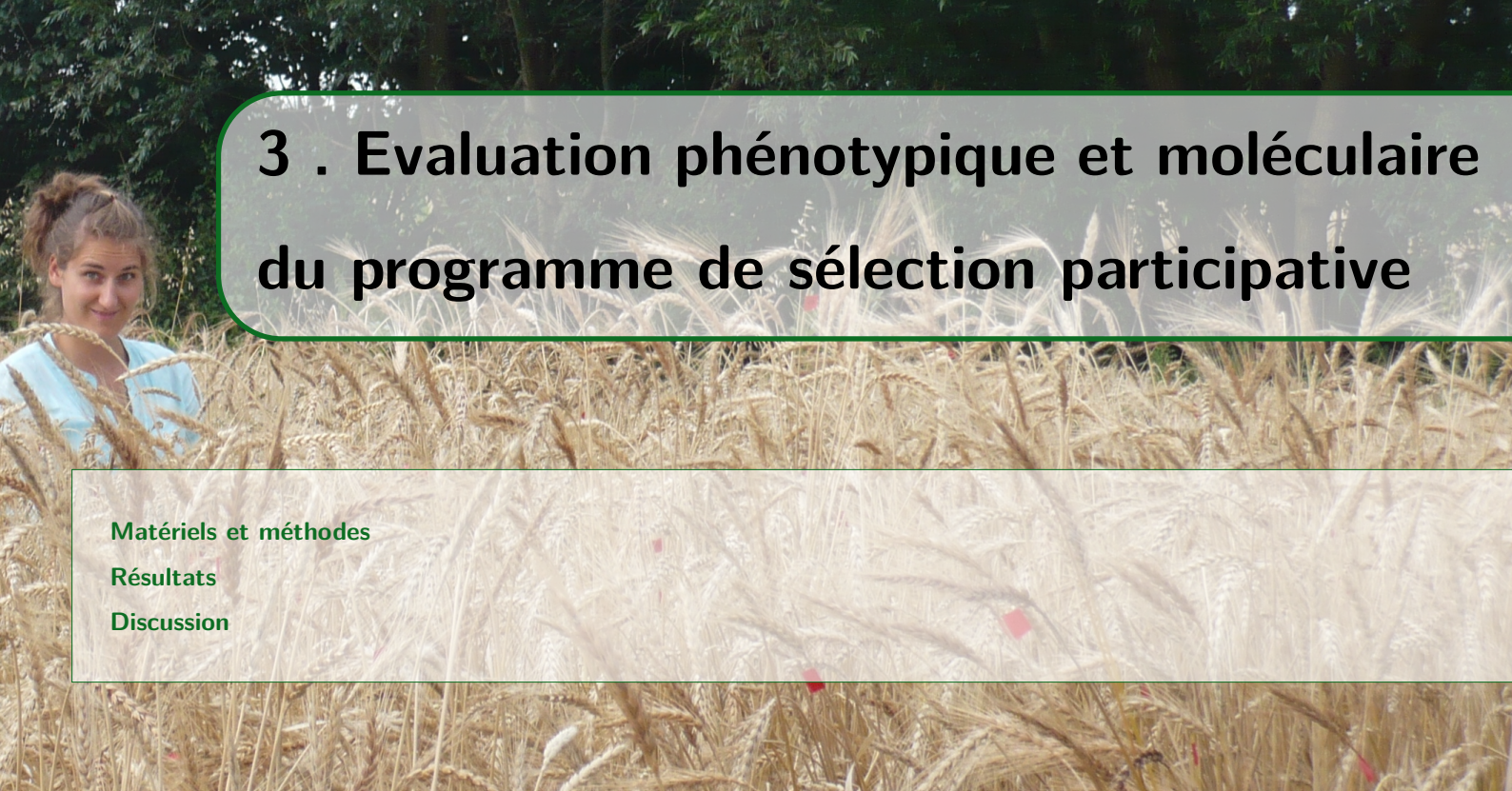
Légende :

1	ALP2011	16	FLM2011	31	JAB2009	46	JUB2012	61	PHG2009
2	ANB2011	17	FLM2012	32	JAB2010	47	LUD2010	62	PIS2010
3	ANB2012	18	FLP2009	33	JAB2011	48	LUD2011	63	RAB2009
4	BER2009	19	FLP2010	34	JFB2010	49	MAV2010	64	RAB2010
5	BER2010	20	FLP2011	35	JFB2011	50	MAV2011	65	RAB2011
6	BRE2011	21	FLP2012	36	JFB2012	51	MAV2012	66	RAB2012
7	CHD2009	22	FRC2009	37	JJG2009	52	MPH2009	67	RAL2011
8	CHD2010	23	FRC2010	38	JJG2010	53	MPH2010	68	RAL2012
9	CHD2011	24	FRC2011	39	JMC2010	54	MPH2012	69	RIH2011
10	CHD2012	25	FRC2012	40	JMC2011	55	OLM2010	70	RIH2012
11	CHP2009	26	GIM2010	41	JMC2012	56	OLR2009	71	ROW2010
12	CHP2012	27	GIM2012	42	JPB2009	57	OLR2010	72	ROW2011
13	EUK2012	28	HEC2009	43	JSG2010	58	OLR2011	73	ROW2012
14	FLM2009	29	HEC2010	44	JSG2011	59	OLR2012	74	STP2012
15	FLM2010	30	HEC2012	45	JSG2012	60	PAJ2012	75	VIC2009

Les résultats obtenus permettent aux paysans d'utiliser le réseau d'essai pour choisir de nouveaux germplasms à tester dans leurs fermes. Pour cela, trois options sont proposées aux paysans :

1. La première option est de choisir un germplasm sans *a priori*. Les germplasm les moins sensibles à l'interaction avec une valeur de  $\beta$  proche de 1 ainsi qu'une valeur extrême, positive ou négative selon les critères et les variables, pour  $\alpha$  peuvent être privilégiés.
2. La deuxième option est, à l'aide de l'analyse multivariée, de repérer une ferme qui se trouvera proche de la ferme du paysan. Une fois la ferme choisie, il est possible de choisir des populations évaluées dans cette ferme grâce aux comparaisons de moyenne (partie 2.2).
3. La troisième option est de prédire le comportement qu'aurait eu un germplasm sur un environnement en utilisant le modèle 2.11. Ce dernier est assez prédictif (annexe L). Les cinq germplasms avec des valeurs les plus élevées, ou les plus faibles selon les critères, pourraient être proposés pour chaque environnement pour chaque variable.





# 3 . Evaluation phénotypique et moléculaire du programme de sélection participative

Matériels et méthodes

Résultats

Discussion

Nous avons vu dans le chapitre 1 l'importance des ressources génétiques pour développer des agrosystèmes durables. Ces ressources peuvent être gérées de manière *ex-situ* ou *in-situ*. La gestion dynamique à la ferme est une forme de gestion *in-situ*.

L'équipe DEAP a travaillé sur le rôle des paysans et de la gestion à la ferme dans la conservation des ressources génétiques. Il a été montré que le réseau social et les pratiques des paysans ont un impact sur la diversité génétique maintenue. Le réseau d'acteur permet de maintenir de manière dynamique la diversité, voire même de l'augmenter<sup>229;231</sup>.

Le programme de sélection participative, issu de la co-construction entre chercheurs et paysans a un impact sur les pratiques des paysans<sup>i</sup>. Les paysans créent, en collaboration avec les chercheurs, de nouvelles populations qu'ils multiplient, échangent, sélectionnent, mélangent. La sélection est locale et décentralisée. Le fonctionnement en réseau permet des échanges de savoirs et savoir-faire entre les acteurs et la circulation des semences. On s'attend à ce que ce système de sélection participative maintienne et génère de la diversité génétique mais cela reste à démontrer.

Ce chapitre a deux objectifs. Le premier est d'étudier la diversité créée et maintenue par cette sélection locale décentralisée. Le deuxième est de caractériser l'effet de différents facteurs, tels que la diversité initiale, les environnements et les pratiques de sélection sur la réponse des populations et le maintien de la diversité génétique.

Pour cela, nous avons étudié un échantillon de populations qui se voulait représentatif des facteurs étudiés. L'impact de trois facteurs a été étudié au niveau moléculaire et

---

i. On revient sur ce point dans le chapitre 4

phénotypique : le type de parent utilisé dans les croisements ; la ferme où les populations ont été multipliées ; la présence, ou non, de sélection massale intra-population.

Une première partie de ce chapitre porte sur les analyses moléculaires sur des marqueurs neutres et sur des marqueurs situés dans des gènes candidats principalement associés à la précocité de floraison. Une deuxième partie porte sur les analyses phénotypiques. La discussion reprend les résultats marquants et replace le projet de sélection participative dans la gestion des ressources génétiques.

## 3.1 Matériels et méthodes

### 3.1.1 Expérimentation et acquisition des données moléculaires et phénotypiques

#### 3.1.1.1 Plan d'expérience et choix des populations

Un échantillon de 210 populations, issues de 25 croisements réalisés au départ du projet et de certains parents utilisés dans les croisements, a été étudié (Tableau 3.1). Chaque population a reçu la nomenclature suivante :  $[germplasm]_{[ferme-où-il-a-été-récolté]}[F_n(s)]$  avec  $F_n$  la génération depuis le croisement et  $(s)$  la génération depuis la sélection pour le cas particulier des populations issues de sélection massales. Dans ce chapitre, contrairement au chapitre 2, le germplasm ne change pas pour une population issue de sélection massale intra-population. Par exemple, la population  $C21\#S$  (l'événement de sélection est marqué par le « # »), sélection massale intra-population dans le germplasm  $C21$ , est associée au germplasm  $C21$ . Ceci nous a permis de travailler sur l'effet de la diversité initiale et celui de la sélection massale intra-population.

Les populations ont été choisies pour maximiser trois facteurs : le nombre de fermes où les germplasms avaient été cultivés ; le type de germplasm (Tableau 2.3 dans la partie 2.1) ; la présence, ou non, de sélection massale intra-population. Quand cela a été possible, les parents utilisés dans les croisements ont été semés. Certains germplasms partagent le même parent (Figure 3.1). Les populations issues des croisements peuvent correspondre à différentes générations (Tableau 3.2). Le type de germplasm a été choisi à l'aide des bases de données en ligne mais également à dire d'experts, c'est à dire des paysans qui les cultivent. En effet, certains germplasms, avec leur histoire particulière, pouvaient présenter plus ou moins de diversité qu'attendu. Par exemple, *Précoce-du-Japon*, une variété a priori ancienne a été considérée comme une variété moderne. Idem pour *Poncheau*.

A la suite d'une réunion avec les paysans, il a été décidé de faire l'évaluation phénotypique sur les fermes afin d'être au plus près des conditions de culture à la ferme.



L'évaluation phénotypique a été réalisée sur trois fermes participant au projet de sélection participative : *JSG*, *JFB* et *OLR*. Ces fermes sont assez différentes d'un point de vue climatique (Figure 3.2) et des pratiques agronomiques. Toutes les populations évaluées sur ces trois fermes ont aussi été semées à la station de recherche à l'INRA du Moulon (*MLN*) avec deux répétitions. Ceci pour mesurer la date d'épiaison facilement et pour prélever les échantillons d'ADN pour le laboratoire.

Pour l'évaluation phénotypique, il y avait 149 populations, issues de 25 germplasm « croisements » et de certains germplasms parents, qui ont été réparties entre les trois fermes *JFB*, *JSG* et *OLR*. Nous avons veillés à ce qu'il y ait au moins deux populations par type de germplasm et par ferme. Les populations étudiées, issues des croisements, provenaient de 17 fermes (*FRC*, *MAV*, *JSG*, *BRE*, *JFB*, *RIH*, *MPH*, *OLR*, *ROW*, *CHD*, *RAB*, *RAL*, *FLM*, *ALP*, *JMC*, *JAB* et *GIM* (Figure 2.1 dans la partie 2.1)); 29 étaient issues d'une sélection massale intra-population; 28 étaient issues d'un croisement de type  $VA \times VP$ ; 28 étaient de type  $VP \times VP$ ; 15 étaient de type  $VM \times VM$ ; 33 étaient de type  $VM \times VP$  (Tableau 3.1 & Tableau 3.2).

Les trois fermes (*JFB*, *JSG* et *OLR*), ont semé respectivement 80, 48 et 48 populations. Les expérimentations sur ces fermes ont comporté respectivement quatre, trois et trois blocs avec des témoins (faisant partie des populations) répartis dans chaque bloc. Les fermes partageaient trois témoins : *Blé-du-Lot*, *Rouge-du-Roc* et *Blé-des-Hautes-Pyrénées*. Ces témoins étaient respectivement répétés : 4, 8, et 2 fois chez *JFB*; 2, 7, et 2 fois chez *JSG* et 2, 6, et 1 fois chez *OLR*. Quelques populations étaient présentes sur plusieurs fermes comme *Atar* et *Touselle Anone* (*JFB*, *JSG* et *OLR*); *Rouge-de-Bordeaux* (*JFB* et *JSG*); *Bladette-de-Besplas*, *Magdalena*, *Redon* et *Rallet* (*JFB* et *OLR*).

Pour l'évaluation moléculaire, 65 populations, issues de 12 germplasms « croisement » et de certains germplasms parents, ont été choisies parmi les populations étudiées dans l'évaluation phénotypique. Pour chaque germplasm, les parents ont été génotypés. Les échantillons de certains parents, qui n'étaient pas disponibles sur le réseau, ont été récupérés à la banque de graines de Clermont-Ferrand (*CLF*). Les populations choisies sont issues de 13 fermes (*JSG*, *MAV*, *FRC*, *BRE*, *JFB*, *CHD*, *ROW*, *RAB*, *OLR*, *ALP*, *JMC*, *MPH* et *FLM* (Figure 2.1 dans la partie 2.1)); 11 étaient issues d'une sélection massale intra-population; 10 étaient de type  $VA \times VP$ ; 9 de type  $VP \times VP$ ; 5 de type  $VM \times VM$ ; 25 de type  $VM \times VP$  (Tableau 3.1 & Tableau 3.2).

Tableau 3.1 - Origine des populations étudiées dans l'évaluation et type de germplasm.

Tableau 3.1 – (partie 1/3)

germplasm	père			mère			descendant		
	type	nb	germplasm	type	nb	germplasm	type	nb	
<i>Touselle-Anone</i>	VP	[3-1-0];[1-1-0]	<i>Cloche</i>	VA	[0-0-0];[1-1-0]	<i>Clomone</i>	VA × VP	[4-2-0];[2-2-0]	
<i>Blé-du-Lot</i>	VP	[3-1-0];[1-1-0]	<i>Rouge-du-Roc</i>	VA	[3-1-0];[1-1-0]	<i>Rock-and-Lot</i>	VA × VP	[2-2-2];[2-2-2]	
<i>Blé-du-Lot</i>	VP	[3-1-0];[1-1-0]	<i>Rouge-de-Bordeaux</i>	VP	[2-1-0];[1-1-0]	21 × 3	VP × VP	[4-3-0];[3-3-0]	
<i>Touselle-Rouge-de-la-Drome</i>	VP	[0-0-0];[0-0-0]	<i>Marquis</i>	VA	[0-0-0];[0-0-0]	C5	VA × VP	[1-1-0];[0-0-0]	
<i>Rallet</i>	VP	[2-1-0];[1-1-0]	<i>Cloche</i>	VA	[0-0-0];[1-1-0]	C8	VA × VP	[6-3-2];[4-3-1]	
<i>Barbu-de-l-Aveyron</i>	VP	[0-0-0];[0-0-0]	<i>Rouge-de-Bordeaux</i>	VP	[2-1-0];[1-1-0]	C13	VP × VP	[4-3-1];[0-0-0]	
<i>Poncheau-selection-fermier</i>	VM	[0-0-0];[0-0-0]	<i>Alaуда</i>	VM	[1-1-0];[0-0-0]	C14	VM × VM	[7-5-2];[0-0-0]	

Tableau 3.1 – (partie 2/3)

germplasm	père			mère			descendant		
	type	nb	germplasm	type	nb	germplasm	type	nb	
<i>Blé-des-Hautes-Pyrénées</i>	VP	[3-2-0];[1-1-0]	<i>Atar</i>	VM	[3-1-0];[1-1-0]	C16	VM × VP	[7-5-2];[6-5-2]	
<i>Blé-des-Hautes-Pyrénées</i>	VP	[3-2-0];[1-1-0]	<i>Pollux</i>	VM	[0-0-0];[1-1-0]	C21	VM × VP	[13-7-6];[8-7-2]	
<i>Touselle-barbue</i>	VA	[0-0-0];[0-0-0]	<i>Automne-Rouge</i>	VP	[0-0-0];[1-1-0]	C29	VA × VP	[2-2-1];[0-0-0]	
<i>Poncheau</i>	VM	[1-1-0];[1-1-0]	<i>Véga</i>	VM	[1-0-0];[1-1-0]	C32	VM × VM	[5-3-2];[5-3-2]	
<i>Odessa-sans-barbe</i>	VP	[0-0-0];[0-0-0]	<i>Vilmorin-27</i>	VA	[1-1-0];[0-0-0]	C33	VA × VP	[5-2-3];[0-0-0]	
<i>Atar</i>	VM	[3-1-0];[1-1-0]	<i>Touselle-Anone</i>	VP	[3-1-0];[1-1-0]	C37	VM × VP	[6-5-1];[5-5-1]	
<i>Blé-des-Hautes-Pyrénées</i>	VP	[3-2-0];[1-1-0]	<i>Pollux</i>	VM	[0-0-0];[1-1-0]	C42	VM × VP	[8-7-1];[6-5-1]	
<i>Barbu-de-l-Aveyron</i>	VP	[0-0-0];[0-0-0]	<i>Atar</i>	VM	[3-1-0];[1-1-0]	C46	VM × VP	[3-3-1];[0-0-0]	
<i>Blanc-de-la-Reole</i>	VP	[1-1-0];[0-0-0]	<i>Blé-Roumain</i>	VP	[0-0-0];[0-0-0]	C49	VP × VP	[2-2-0];[0-0-0]	
<i>Touselle-Rouge-de-la-Drome</i>	VP	[0-0-0];[0-0-0]	<i>Magdaléna</i>	VA	[2-1-0];[0-0-0]	C58	VA × VP	[4-3-1];[0-0-0]	
<i>Precoc-du-Japon</i>	VM	[0-0-0];[0-0-0]	<i>Probelle</i>	VM	[0-0-0];[0-0-0]	C66	VM × VM	[3-3-0];[0-0-0]	

Tableau 3.1 – (partie 3/3)

germplasm	père			mère			descendant		
	type	nb	germplasm	type	nb	germplasm	type	nb	
<i>Touselle-Anone</i>	VP	[3-1-0];[1-1-0]	<i>Automne-Rouge</i>	VP	[0-0-0];[1-1-0]	C67	VP × VP	[3-3-0];[3-3-0]	
<i>Cloche</i>	VA	[0-0-0];[1-1-0]	<i>Rouge-de-Bordeaux-</i> <i>INRA</i>	VP	[1-1-0];[1-1-0]	C77	VA × VP	[2-2-0];[2-2-0]	
<i>Blé-de-Langogne</i>	VP	[0-0-0];[0-0-0]	<i>Roisel</i>	VM	[0-0-0];[0-0-0]	C80	VM × VP	[5-3-2];[0-0-0]	
<i>Redon</i>	VP	[2-1-0];[1-1-0]	<i>Bladette-de-Besplas</i>	VP	[2-1-0];[1-1-0]	C82	VP × VP	[4-3-0];[3-3-0]	
<i>Inversable-Blanc</i>	VP	[0-0-0];[0-0-0]	<i>Touselle-Rouge-de-</i> <i>la-Drome</i>	VP	[0-0-0];[0-0-0]	C87	VP × VP	[2-2-0];[0-0-0]	
<i>Caucasse-Rouge</i>	VP	[0-0-0];[0-0-0]	<i>Inversable-Blanc</i>	VP	[0-0-0];[0-0-0]	C89	VP × VP	[9-5-2];[0-0-0]	
<i>Caucasse-Rouge</i>	VP	[0-0-0];[0-0-0]	<i>Japhet</i>	VA	[1-1-0];[0-0-0]	C93	VA × VP	[2-2-0];[0-0-0]	

Les types de germplasm des descendants sont la concaténation du type du père et du type de la mère.

Une ligne correspond à un germplasm et ses parents.

VP : variétés de pays.

VA : variétés anciennes.

VM : variétés modernes.

nb : [A];[B] nombre de populations pour l'évaluation phénotypique ([A]) et moléculaire ([B]), avec [A] ou [B] = [a-b-c] avec a le nombre de populations évaluées (les répétitions ne sont pas prises en compte), b le nombre de ferme d'où les populations sont originaires (i.e. les fermes où elles ont été récoltées) et c le nombre de populations issues de sélection massive.

Il n'y a qu'un C'5 à cause de problèmes de semis.

Figure 3.1 - Pedigree des populations issues de croisements

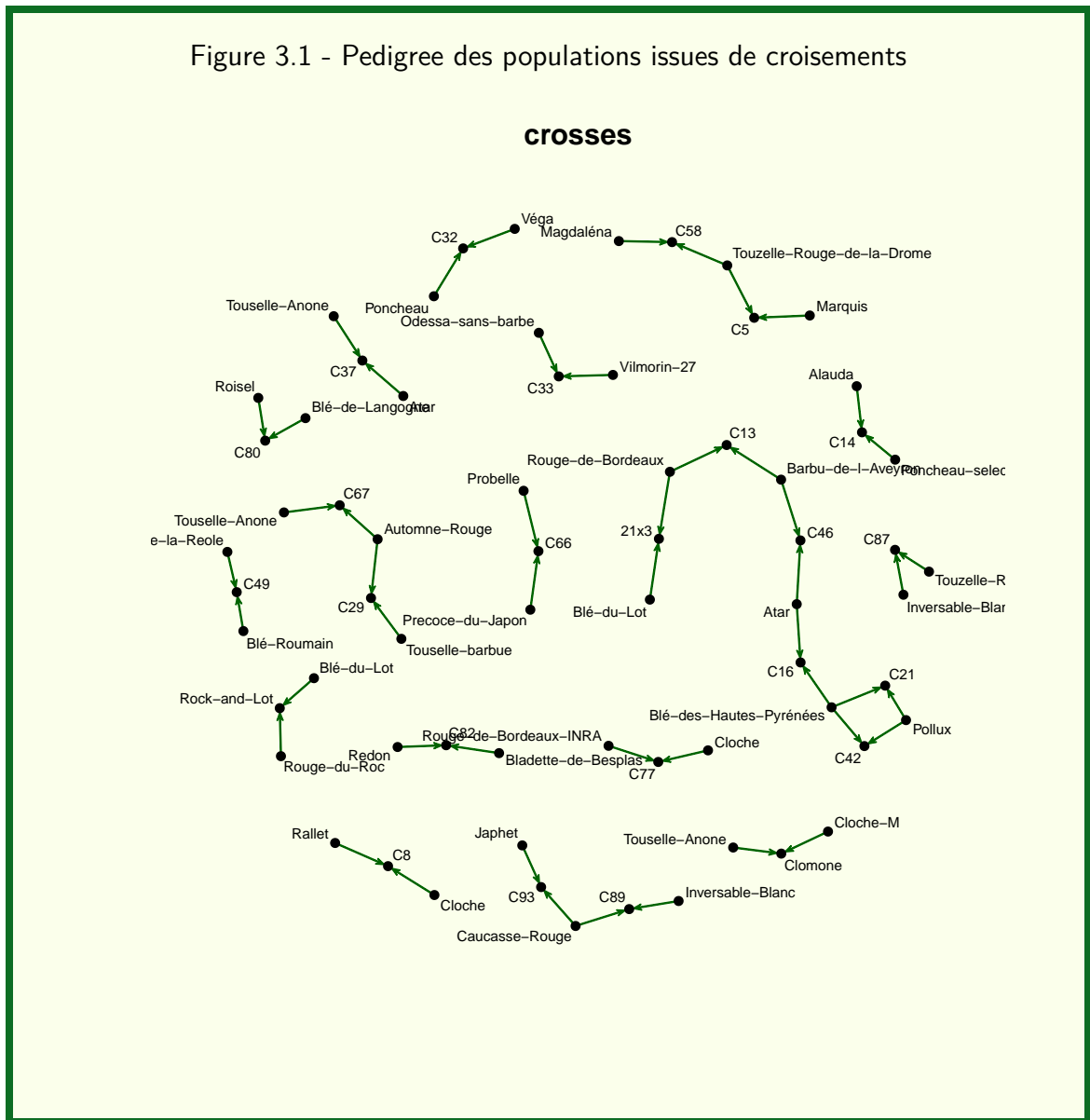
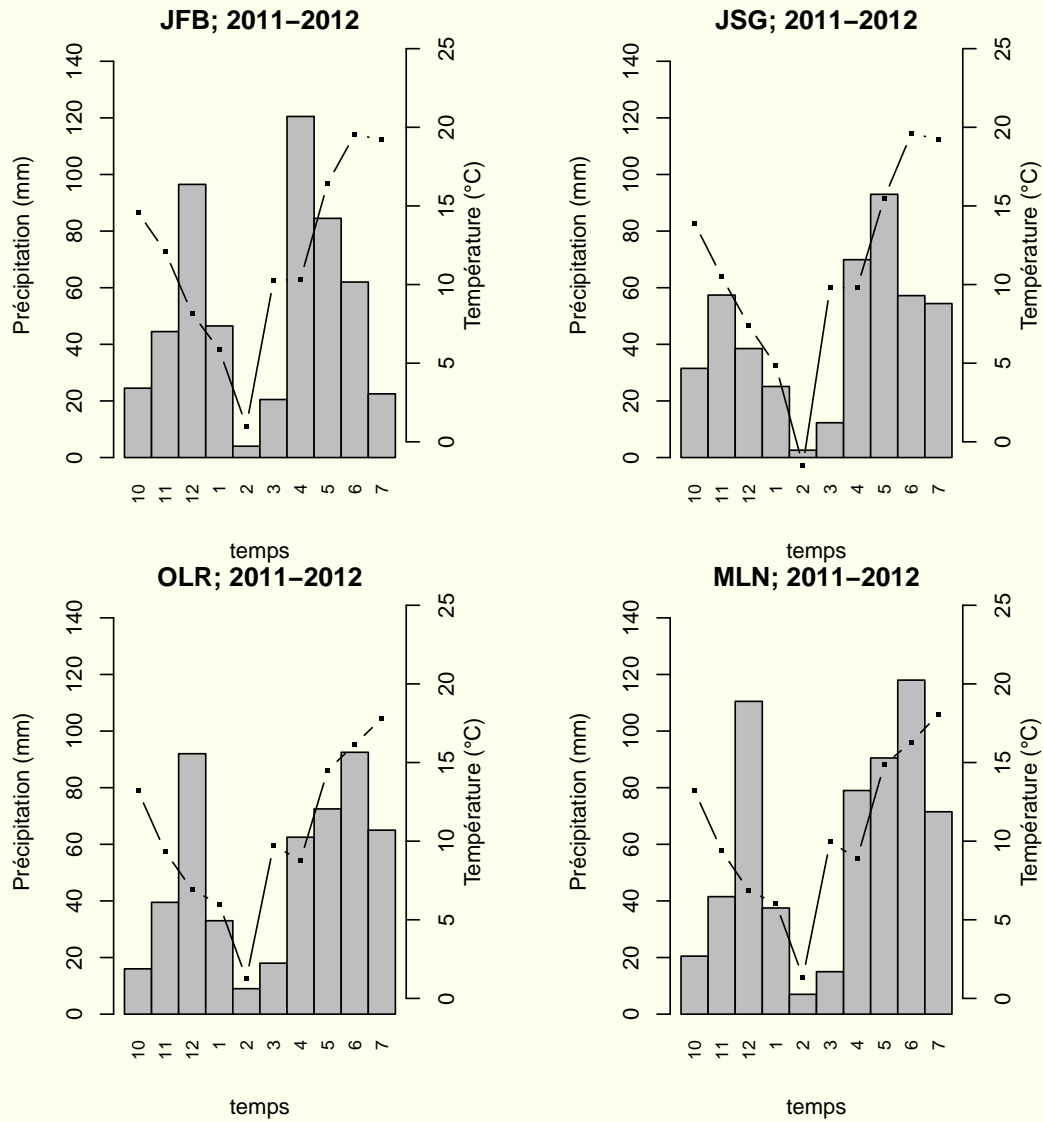


Figure 3.2 - Données météorologiques dans les fermes utilisées pour l'expérimentation pour les précipitations moyennes (barres) et les températures moyennes (lignes) par mois durant le cycle de culture du blé 2011-2012.



Les numéros des mois sont en abscisse.

Tableau 3.2 - Nombre de générations pour les populations issues des croisements et leur type de germplasm.

germplasm	type	F3	F4	F5	F6	F7	F8	?
<i>Clomone</i>	<i>VA</i> × <i>VP</i>	-	-	-	-	-	4/2	-
<i>Rock-and-Lot</i>	<i>VA</i> × <i>VP</i>	-	-	-	-	-	1/1	1/1
21x3	<i>VP</i> × <i>VP</i>	-	1/1	-	3/2	-	-	-
<i>C5</i>	<i>VA</i> × <i>VP</i>	-	-	-	1/0	-	-	-
<i>C8</i>	<i>VA</i> × <i>VP</i>	2/2	-	2/1	1/1	-	-	1/0
<i>C13</i>	<i>VP</i> × <i>VP</i>	-	1/0	-	3/0	-	-	-
<i>C14</i>	<i>VM</i> × <i>VM</i>	-	-	1/0	6/0	-	-	-
<i>C16</i>	<i>VM</i> × <i>VP</i>	-	-	2/2	2/2	-	-	3/2
<i>C21</i>	<i>VM</i> × <i>VP</i>	-	-	2/2	9/5	-	-	2/1
<i>C29</i>	<i>VA</i> × <i>VP</i>	-	-	2/0	-	-	-	-
<i>C32</i>	<i>VM</i> × <i>VM</i>	1/1	-	-	3/2	-	-	1/2
<i>C33</i>	<i>VA</i> × <i>VP</i>	-	-	1/0	2/0	-	-	2/0
<i>C37</i>	<i>VM</i> × <i>VP</i>	-	1/1	1/2	4/2	-	-	-
<i>C42</i>	<i>VM</i> × <i>VP</i>	1/1	-	4/3	2/1	-	-	1/1
<i>C46</i>	<i>VM</i> × <i>VP</i>	-	-	1/0	1/0	-	-	1/0
<i>C49</i>	<i>VP</i> × <i>VP</i>	-	-	1/0	1/0	-	-	-
<i>C58</i>	<i>VA</i> × <i>VP</i>	-	-	1/0	3/0	-	-	-
<i>C66</i>	<i>VM</i> × <i>VM</i>	1/0	1/0	-	1/0	-	-	-
<i>C67</i>	<i>VP</i> × <i>VP</i>	1/1	-	1/1	1/1	-	-	-
<i>C77</i>	<i>VA</i> × <i>VP</i>	-	-	1/1	1/1	-	-	-
<i>C80</i>	<i>VM</i> × <i>VP</i>	2/0	-	2/0	1/0	-	-	-
<i>C82</i>	<i>VP</i> × <i>VP</i>	-	-	2/2	2/1	-	-	-
<i>C87</i>	<i>VP</i> × <i>VP</i>	-	-	-	2/0	-	-	-
<i>C89</i>	<i>VP</i> × <i>VP</i>	1/0	1/0	2/0	3/0	-	-	2/0
<i>C93</i>	<i>VA</i> × <i>VP</i>	-	-	-	-	2/0	-	-
-								

Les informations concernant les données phénotypiques sont à gauche du « / » et les informations concernant les données moléculaires sont à droite.

Les types de germplasm des descendants sont la concaténation du type des parents.

*VP* : variétés de pays.

*VA* : variétés anciennes.

*VM* : variétés modernes.

*Fn* : *n*<sup>ième</sup> génération après le croisement.

? : génération inconnue.

### 3.1.1.2 Données phénotypiques

Les mesures faites sur les fermes sont les suivantes :

- Sur 25 plantes ou épis par population \* : la couleur, les barbes, la courbure, la hauteur, LLSD, la longueur de l'épi, le nombre d'épillets, la stérilité, le poids de l'épi.
- Sur le grain des 25 épis mélangés \* : protéine et poids de mille grains.

\* : ces mesures suivent le même protocole que précédemment (Tableau 2.5 dans la partie 2.1) sauf pour le taux de protéine où la prédiction a été faite sur farine pour gagner en précision.

Au Moulon, les dates au début, au milieu et à la fin de l'épiaison ont été mesurées. L'épiaison correspond à la date où la moitié de l'épi est sorti de la gaine. Le début correspond à la date où les premières plantes de la parcelle épient ; le milieu correspond à la date où 50% des plantes ont épié ; la fin correspond à la date où 95% des plantes ont épié. Les dates d'épiaisons ont été converties en degrés jours à partir des données météo de la station du Moulon.

### 3.1.1.3 Données moléculaires

Le nombre d'individus échantillonnés varie selon la diversité intra-variétale attendue. 25 individus ont été échantillonnés pour les populations issues des croisements et pour les variétés de pays (*VP*). 15 individus ont été échantillonnés pour les variétés anciennes (*VA*) et les variétés modernes (*VM*). A cause de problèmes de semis, certaines populations ont été ressemées en serre. Le nombre d'individus varie légèrement en fonction des plantes disponibles. L'ADN a été extrait à partir de 500 mg de feuilles fraîches suivi d'un protocole modifié de Dellaporta et al.<sup>61</sup> incluant la précipitation des carbohydrates<sup>166</sup>. En tout, 1497 individus ont été génotypés pour 96 Single Nucleotide Polymorphism (SNP) avec la méthode KASPAR<sup>156</sup>. Les séquences de ces marqueurs ont été développées et testées dans le cadre des thèses d'Abdul Rehman Khan<sup>148</sup> et de Stéphanie Thépot<sup>235</sup>.

Nous avons utilisé 53 marqueurs situés dans des zones non codantes et répartis sur le génome (Annexe P.1). Parmi ceux-là, 39 marqueurs polymorphes ont été sélectionnés à partir de la puce 9K iSelect assay du blé avec une diversité (*He*) de 0,5 et une bonne lisibilité de leur profil<sup>4;36;46</sup>. 14 marqueurs ont été sélectionnés à partir de la base de données Kaspar en ligne<sup>44</sup>.

Nous supposons que les populations avaient été soumises à de fortes pression de sélection dans leur environnement (abiotique, biotique, pratiques agronomiques, ...). Afin d'étudier cette hypothèse, nous nous sommes concentrés sur l'adaptation climatique et



sur des gènes candidats associés à l'épiaison. La date d'épiaison est un des caractères d'adaptation majeurs chez le blé. Elle détermine les conditions environnementales de la fécondation. La plante va caler son cycle pour que la fécondation se fasse dans les meilleures conditions<sup>123;202</sup>.

43 marqueurs étaient situés dans des gènes candidats (Annexe P.2). Neuf ont été suggérés par des collègues de l'INRA de Clermont Ferrand ; cinq ont été développés par l'INRA Le Moulon<sup>28;202;209</sup> ; trois ont été adaptés du John Innes Center<sup>19;254</sup>, un a été développé sur un polymorphisme publié<sup>225</sup>, les 25 restant ont été sélectionnés sur la puce 9K iSelect assay du blé<sup>4;36;46;238</sup>.

Six marqueurs n'ont pas fonctionné. Sur le jeu de données avec 90 SNP, nous n'avons eu que 3.6% de données manquantes. Une réassignation des données manquantes a permis d'atteindre un taux de 1.7% de données manquantes dans le jeu de données final. Ce sont surtout des individus hétérozygotes trouvés qui ont été réassignés à partir des motifs de fluorescence et de l'information a priori sur le taux d'hétérozygote attendu. Huit marqueurs étaient monomorphes dans les populations étudiées. Au final, nous avons travaillé avec 82 marqueurs : 34 dans des gènes candidats (marqueurs *CA*) et 48 dans des zones neutres (marqueurs *NE*).

### 3.1.2 Analyses statistiques

Toutes les analyses ont été réalisées avec le logiciel R<sup>197</sup>.

#### 3.1.2.1 Structure génétique

La structure a été analysée avec une Analyse Discriminante sur les Composantes Principales (DAPC). La DAPC consiste à réaliser dans un premier temps une ACP sur les données. Cela permet de générer un nouveau jeu de données avec des variables non corrélées et moins de variables que d'individus. Cela permet ensuite de réaliser une Analyse Discriminante qui crée des groupes qui maximisent la variance entre les groupes et minimisent la variance à l'intérieur des groupes. La DAPC retourne, pour chaque individu, une probabilité d'appartenance à un groupe<sup>144</sup>. Sur les 48 marqueurs neutres, une DAPC a été réalisée avec le package `adegenet`<sup>144</sup> sur les parents dans un premier temps et sur les descendants dans un deuxième temps.

Sur les descendants, pour les marqueurs *NE* et les marqueurs *CA*, les  $F_{st}$  entre paires de populations ( $pF_{st}$ ) ont été calculés (package `adegenet`). A partir de chacune des matrices de  $pF_{st}$  ainsi obtenue, nous avons construit un dendrogramme avec la méthode

UPGMA (fonction `agnes`).

Sur les marqueurs *NE*, les marqueurs *CA* et les deux types de marqueurs mélangés, des analyses de variance sur les données moléculaire (AMOVA) ont été réalisées sur un sous jeu de données correspondant aux descendants (package `pegas`). Cela représentait 4 types de germplasm, 12 germplasms et 49 populations.

Pour chaque germplasm et les parents correspondants, une ACP a été réalisée sur les marqueurs *NE* dans un premier temps et sur les marqueurs *CA* dans un deuxième temps.

Enfin, pour toutes les populations, l'hétérozygotie observée et attendue (estimateur non biaisé de Nei ( $H_{e_{NE}}$ ))<sup>178</sup> a été calculée (package `adegenet`).

### 3.1.2.2 Différenciation phénotypique

Le dispositif expérimental est déséquilibré car peu de populations sont répétées dans chaque bloc et chaque ferme. Tout d'abord, le modèle bayésien hiérarchique 2.1 (partie 2.2.1.1.1) a été utilisé en prenant toutes les données disponibles dans le projet de sélection participative. Ce modèle permet d'estimer les variances résiduelles intra-ferme et les effets des populations dans chaque ferme. Le modèle n'a pas fonctionné avec le taux de protéine (les chaînes MCMC ne convergeaient pas). Je n'ai pas eu le temps de trouver la cause de ce problème, les données avec les protéines ne sont donc pas présentées.

Dans un deuxième temps,  $Y_{ij}$ , la valeur phénotypique pour une variable donnée mesurées sur une ferme  $i$  et pour une population  $j$  a été modélisée comme :

$$Y_{ij} = \mu + \theta_i + \epsilon_{ij}$$

avec  $\theta_i$ , l'effet ferme (*OLR*, *JFB* ou *JSG*),  $\epsilon_{ij}$ , la résiduelle qui correspond aux variations entre populations. Ce modèle permet d'estimer des effets fermes et de corriger les données pour ces effets. Ce modèle a été réalisé uniquement sur un sous-jeu de données regroupant les témoins répétés dans les fermes.

Pour les mesures d'épiaisons faites au Moulon, les données sont corrigées par rapport à l'effet bloc. Ces derniers n'étaient pas significatifs avec des *p-value* supérieures à 0,7.

Les effets populations, ainsi ajustés de la variance intra et inter ferme, sont utilisés dans la suite de l'analyse. Une ACP (package `FactoMineR`) et trois modèles d'Analyse de variances ont été réalisés.

$Y_{ijk}$ , la valeur phénotypique pour une variable donnée, un type de germplasm  $i$ , un

germplasm  $j$  et une population  $k$  a été modélisée comme :

$$Y_{ijk} = \mu + \alpha_i + \beta_j(\alpha_i) + \epsilon_{ijk} \quad (3.1)$$

où  $\mu$  est la moyenne générale,  $\alpha_i$  est l'effet du type de germplasm  $i$ ,  $\beta_j$  est l'effet du germplasm  $j$  qui est hiérarchisé dans l'effet type,  $\epsilon_{ijk}$  est la résiduelle, qui correspond aux variations entre populations.

Pour étudier l'effet de la sélection massale intra-population, le modèle 3.2 a été utilisé.  $Y_{ijk}$ , la valeur phénotypique pour une variable donnée, un type de germplasm  $i$ , le statut de la sélection (sélectionné ou non)  $j$  et une population  $k$  a été modélisée comme :

$$Y_{ijk} = \mu + \alpha_i + \nu_j + \epsilon_{ijk} \quad (3.2)$$

où  $\mu$  est la moyenne générale,  $\alpha_i$  est l'effet du type de germplasm  $i$ ,  $\nu_j$  est l'effet de la sélection  $j$ ,  $\epsilon_{ijk}$  est la résiduelle qui correspond aux variations entre populations.

Pour étudier l'effet lié à la ferme, le modèle 3.3 a été utilisé.  $Y_{ijk}$ , la valeur phénotypique pour une variable donnée, un type de germplasm  $i$ , la ferme  $j$  et une population  $k$  a été modélisé comme :

$$Y_{ijk} = \mu + \alpha_i + \theta_j + \epsilon_{ijk} \quad (3.3)$$

où  $\mu$  est la moyenne générale,  $\alpha_i$  est l'effet du type de germplasm  $i$ ,  $\theta_j$  est l'effet de la ferme  $j$ ,  $\epsilon_{ijk}$  est la résiduelle qui correspond aux variations entre populations.

Des comparaisons de moyennes ont été faites avec un risque global  $\alpha = 0,1$  et un risque individuel  $\alpha/I$ , avec  $I$  le nombre de population comparées.

## 3.2 Résultats

### 3.2.1 Organisation de la diversité génétique

#### 3.2.1.1 Les parents

Nous avons tout d'abord décrit la structure des parents afin de voir d'une part si certains étaient proches génétiquement entre eux, et d'autre part si chacun apportait une contribution génétique homogène ou non.

Une Analyse Discriminante sur les Composantes Principales (DAPC)<sup>144</sup> a été réalisée sur les données de diversité neutre ( $NE$ ) pour les parents disponibles et un nombre de

groupes égal à 16 a été retenu sur la base du critère Bayésien Information Criterion (BIC) (Figure 3.3). La plupart des parents, à part *Bladette-de-Besplas*, *Rallet*, *Rouge-de-Bordeaux*, *Touselle-Anone*, appartenaient majoritairement à un groupe (Figure 3.3). *Automne-Rouge* et *Rouge-de-Clervaux* appartenaient au même groupe. Le germplasm *Rouge-de-Bordeaux* venant de chez *JFB* s'est réparti dans trois groupes. D'un autre côté, *Rouge-de-Bordeaux-INRA*, qui venait de la collection de Clermont-Ferrand, est apparu dans un seul groupe, différent des trois précédents. Ces résultats sont cohérents avec Thomas et al.<sup>231</sup> qui avaient également trouvé quatre groupes pour *Rouge-de-Bordeaux* avec le groupe de l'échantillon INRA présent en fréquence faible dans le *Rouge-de-Bordeaux* de *JFB*. La diversité faible chez *Rouge-de-Bordeaux-INRA*, confirmée avec les valeurs de  $H_e$  (Tableau 3.3) est également cohérente avec les résultats précédents. Quelques individus de *Touselle-Anone* et *Rouge-de-Bordeaux* partageaient le même groupe, ce qui avait déjà été observé par Khan<sup>148</sup>. Les individus de *Rouge-du-Roc* ont tous été assignés à un même groupe commun au *Rouge-de-Bordeaux* venant de chez *JFB*. Thomas et al.<sup>231</sup> avaient détecté dans un échantillon de *Rouge-du-Roc* trois groupes dont deux majoritaires avec 50 et 45% des individus. L'échantillon de *Rouge-du-Roc* étudié dans notre étude ayant sept générations de plus que celui étudié par Thomas et al.<sup>231</sup>, la population a pu se différencier, provoquant une perte de certains génotypes. C'est peut être aussi dû à un effet d'échantillonnage.

Le dendrogramme construit à partir des  $F_{st}$  entre paires de variétés parents ( $pF_{st}$ ) sur les marqueurs *NE* (Figure 3.4) est cohérent avec la structure révélée par la DAPC. Cependant, *Rouge-du-Roc* a été trouvé très proche de *Rouge-de-Bordeaux-INRA* et non de *Rouge-de-Bordeaux* issu de chez *JFB*, révélant la difficulté pour les  $F_{st}$  à représenter la proximité entre des entités « composites » avec des groupes communs.

Le dendrogramme à partir des  $pF_{st}$  sur les marqueurs *CA* a montré quelques profils différents mais sans changements majeurs. *Rouge-du-Roc* est toujours proche de *Rouge-de-Bordeaux-INRA* et *Automne-Rouge* et *Rouge-de-Clervaux* sont apparus très proches (Figure 3.4).

L'hétérozygotie attendue au niveau des marqueurs neutres ( $H_{e_{NE}}$ ) dans les variétés parents variait entre 0,00 et 0,28 (Tableau 3.3). Nous avons remarqué en particulier que les échantillons issus de la banque de graines de Clermont-Ferrand (*CLF*) avaient une diversité faible ( $H_{e_{NE}}(\textit{Cloche}) = 0,00$ ;  $H_{e_{NE}}(\textit{Blé-du-Lot}) = 0,01$ ;  $H_{e_{NE}}(\textit{Pollux}) = 0,04$ ). Si en général, il y avait plus de diversité génétique au sein des échantillons maintenus à la ferme, certains se sont révélés très homogènes également : pour *Automne-Rouge* et

*Poncheau* issus de *BRE*,  $H_{e_{NE}} = 0.00$  et  $0.02$  et pour *Redon* et *Rouge-de-Clairvieux* issus de *JFB*,  $H_{e_{NE}} = 0.00$ . *Rouge-du-Roc* qui est une sélection dans *Rouge-de-Bordeaux* par *JFB* avait une valeur de  $H_{e_{NE}} = 0.06$ , similaire à *Rouge-de-Bordeaux-INRA* ( $H_{e_{NE}} = 0.05$ ) mais plus faible que *Rouge-de-Bordeaux* issu de chez *JFB* ( $H_{e_{NE}} = 0.28$ ).

L'hétérozygotie observée  $H_{o_{NE}}$  était faible entre  $0.00$  et  $0.06$ .

### 3.2.1.2 Les descendants

Nous avons d'abord observé comment se regroupaient les types de parents et les types de croisement à partir d'un dendrogramme réalisé sur les  $pF_{st}$  entre les types au niveau des marqueurs *NE* (Figure 3.5). Le dendrogramme basé sur les  $pF_{st}$  a montré trois groupes. Ils se structuraient autour de *VA*, *VM* et *VP*. Tous les types avec *VP* se sont retrouvés dans un groupe. Les types *VM* et  $VM \times VM$  étaient dans un autre groupe. Le type *VA* était seul dans le dernier groupe (Figure 3.5). Le groupe contenant les populations de type  $VP$ ,  $VP \times VP$ ,  $VP \times VA$  et  $VP \times VM$  est retrouvé dans le dendrogramme construit sur la base des  $pF_{st}$  aux *CA*, mais les autres types ne sont pas apparus organisés.

D'après l'AMOVA, ce sont les différences entre germplasm qui expliquaient le plus de variabilité génétique, suivi par les populations puis les types (Tableau 3.4). Alors que le germplasm et la population sont apparus très significatifs, le type ne l'était pas du tout, indiquant pas ou peu d'effet du type de croisement initial sur la diversité. L'erreur, qui correspond à la variance intra-population, représentait autour de 50% de la variabilité. Logiquement, la variabilité révélée par les marqueurs *NE* était nettement plus importante que celle révélée par les marqueurs *CA*, mais la répartition de cette variabilité aux différents niveaux était par contre tout à fait similaire entre les deux types de marqueurs.

Pour l'ensemble des populations issues des croisements, la diversité de Nei sur les marqueurs *NE* variait entre  $0.00$  et  $0.32$ .  $H_{e_{NE}}$  au sein des populations issues de la sélection massale est apparue élevée, variant entre  $0.14$  et  $0.28$  (excepté une valeur à  $0$ ). Cependant, nous avons noté que dans trois cas sur dix, la valeur  $H_{e_{NE}}$  des sélections massales était la plus petite valeur observée au sein d'un germplasm.

L'hétérozygotie observée  $H_{o_{NE}}$  était située entre  $0.00$  et  $0.12$ . Nous avons observé une corrélation significativement négative entre le nombre de génération et l'hétérozygotie observée ( $H_o$ ). Cependant, alors qu'en génération précoce, l'hétérozygotie observée était plus faible qu'attendue sous l'hypothèse d'une reproduction en autofécondation à partir d'individus hétérozygotes. On peut penser que le maintien d'un niveau d'hétérozygotie plus élevé qu'attendu en présence d'autofécondation est dû à un faible taux de croisements

entre plantes. Aucune tendance n'a été trouvée entre le nombre de génération et la diversité  $H_e$  (Figure 3.6).

On observe une relation positive entre la distance sur le réseau de relation entre populations et la différenciation entre les populations (Figure 3.7).

Figure 3.3 - Groupes détectés par DAPC sur le jeu de données des parents avec la distribution des individus dans chaque groupe en haut et évolution du critère BIC en fonction du nombre de groupes testés pour le clustering en bas.

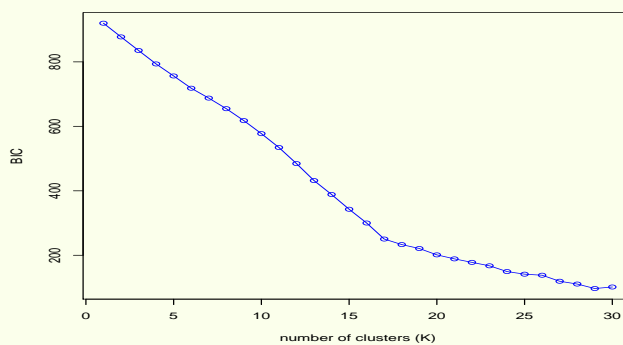
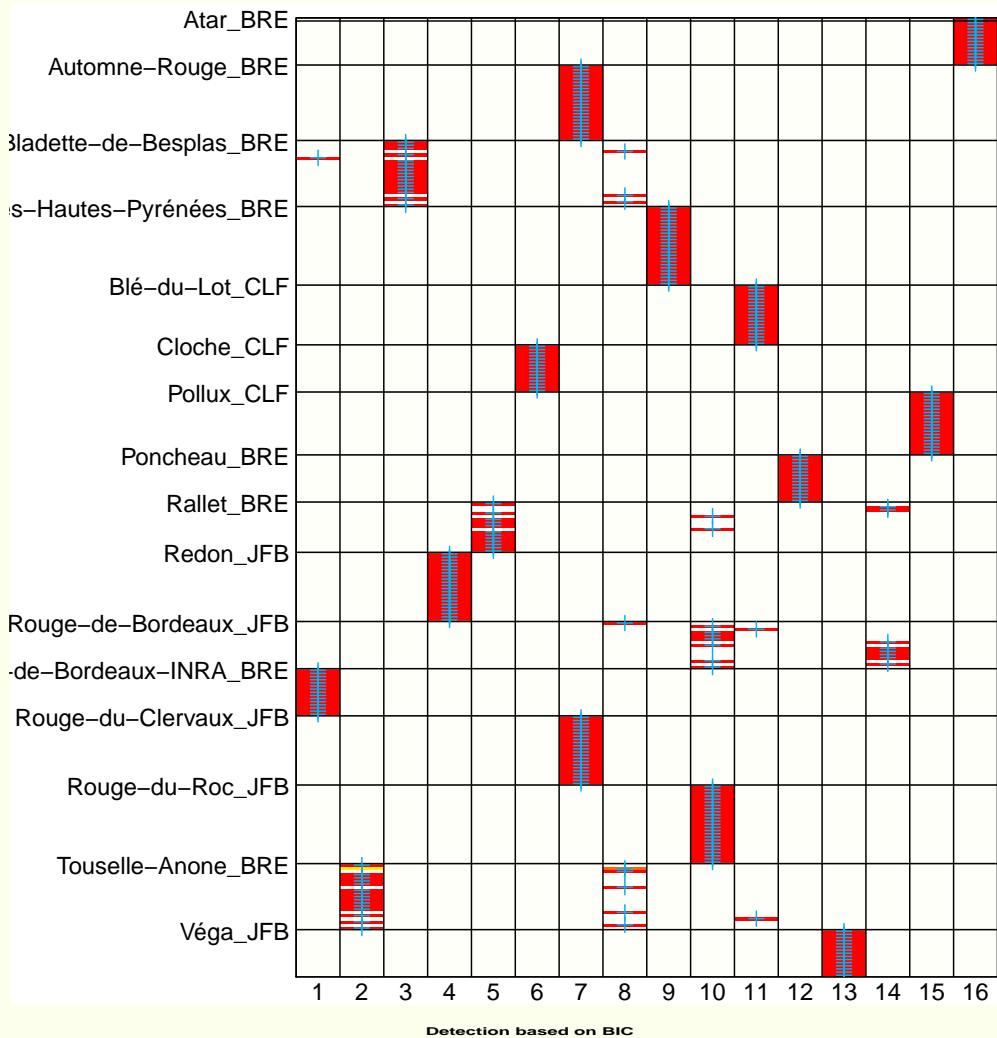


Figure 3.4 - Dendrogramme construit à partir des  $pF_{st}$  sur les marqueurs *CA* en haut et *NE* en bas pour les parents

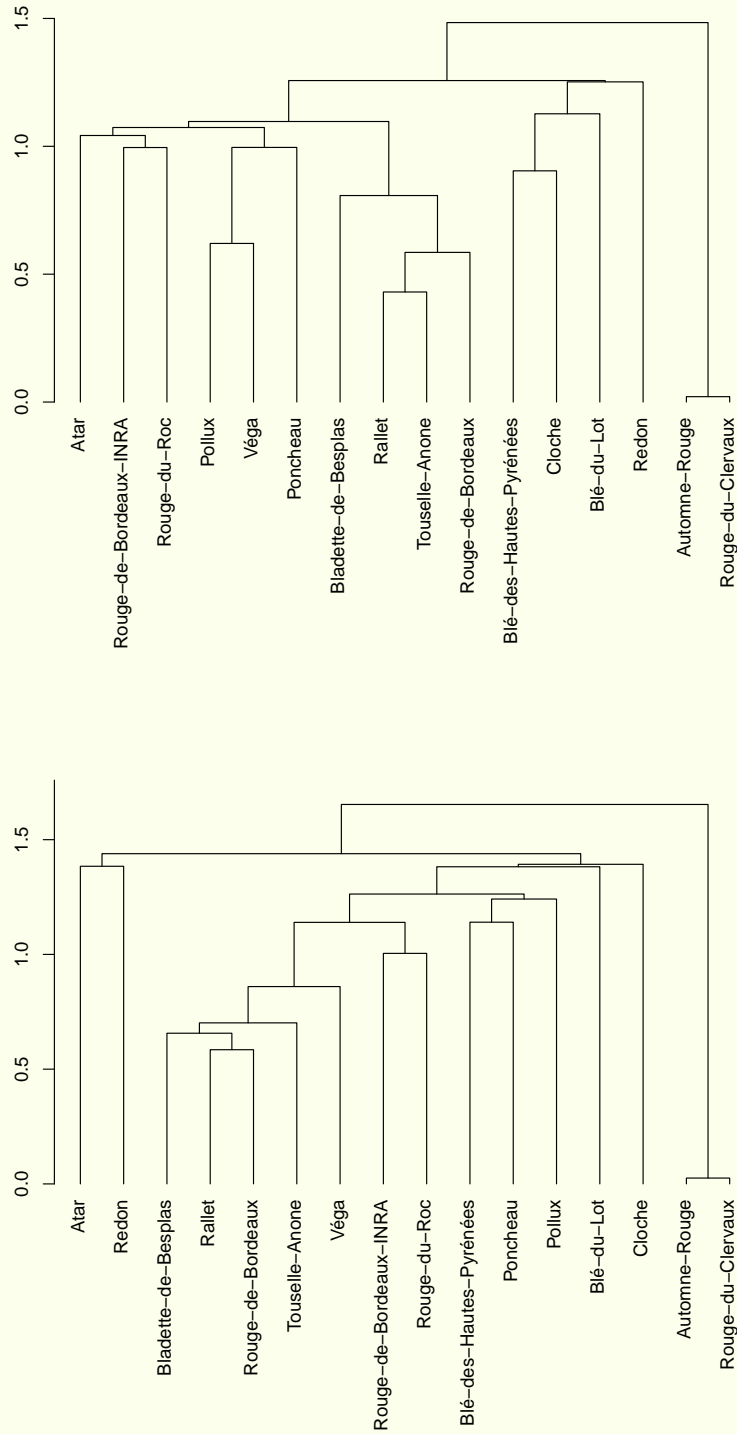




Figure 3.5 - Dendrogramme construit à partir des  $pF_{st}$  sur les marqueurs *CA* à gauche et *NE* à droite pour les types de germplasm

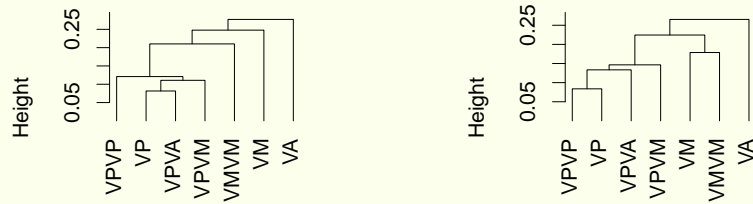
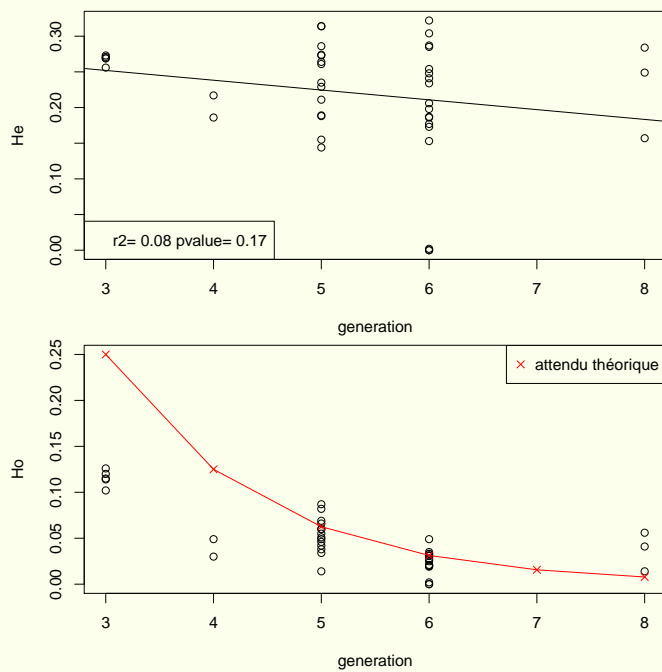
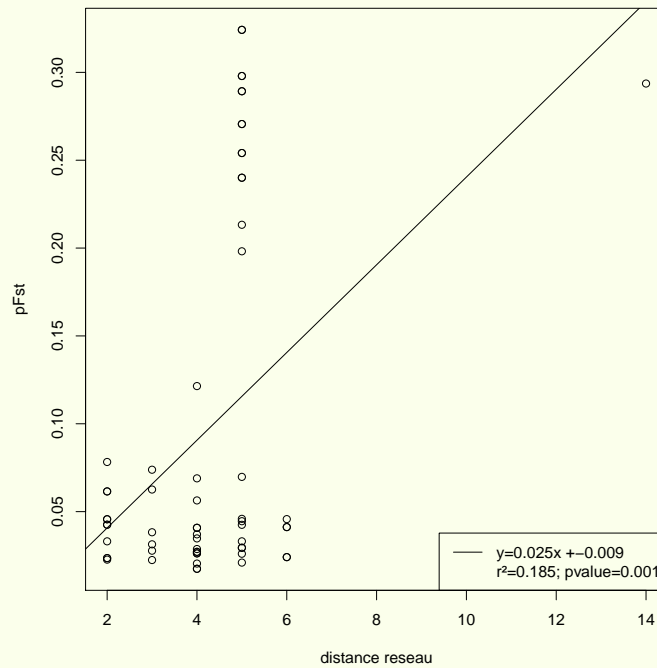


Figure 3.6 - Relation entre le nombre de générations après le croisement des populations et l'hétérozygotie attendue (en haut) et l'hétérozygotie observée (en bas).



Données sur les marqueurs *NE*.  
 $H_e$  : hétérozygotie attendue.  $H_o$  : hétérozygotie observée.

Figure 3.7 - Différenciation génétique ( $pF_{st}$ ) entre les populations en fonction de la distance sur le réseau de relation entre populations.



Données sur les marqueurs *NE*.

La distance sur le réseau correspond au nombre de multiplications qui sépare deux populations depuis leur dernier événement de diffusion commun.

La dernière reproduction, qui a été faite dans le cadre de l'évaluation, n'est pas prise en compte.

Tableau 3.3 - Indices de diversité pour chaque population sur les marqueurs *NE* et *CA*.

population	<i>nb ind</i>	<i>He<sub>NE</sub></i>	<i>Ho<sub>NE</sub></i>	<i>He<sub>CA</sub></i>
<i>Atar_BRE</i>	15	0,01	0,01	0,02
<i>Automne-Rouge_BRE</i>	24	0,00	0,00	0,02
<i>Bladette-de-Besplas_BRE</i>	21	0,19	0,02	0,09
<i>Blé-des-Hautes-Pyrénées_BRE</i>	25	0,04	0,01	0,03
<i>Blé-du-Lot_CLF</i>	19	0,01	0,01	0,05
<i>Cloche_CLF</i>	15	0,00	0,00	0,02
<i>Pollux_CLF</i>	20	0,04	0,01	0,02
<i>Poncheau_BRE</i>	15	0,02	0,00	0,02
<i>Rallet_BRE</i>	16	0,18	0,01	0,10
<i>Redon_JFB</i>	22	0,00	0,00	0,02
<i>Rouge-de-Bordeaux-INRA_BRE</i>	15	0,05	0,06	0,03
<i>Rouge-de-Bordeaux_JFB</i>	15	0,28	0,05	0,16
<i>Rouge-du-Clervaux_JFB</i>	22	0,00	0,00	0,02
<i>Rouge-du-Roc_JFB</i>	25	0,06	0,01	0,06
<i>Touselle-Anone_BRE</i>	21	0,20	0,02	0,11
<i>Véga_JFB</i>	15	0,14	0,01	0,06
<i>21 × 3_JSG_F4</i>	25	0,22	0,03	0,16
<i>21 × 3_MAV_F6</i>	7	0,00	0,00	0,04
<i>21 × 3_FRC_F6</i>	25	0,00	0,00	0,05
<i>C16#BS_JFB_F?(?)</i>	25	0,25	0,06	0,14
<i>C16_CHD_F6</i>	24	0,30	0,03	0,15
<i>C16_ROW_F5</i>	25	0,29	0,09	0,17
<i>C16_BRE_F5</i>	24	0,26	0,06	0,17
<i>C16_MAV_F6</i>	12	0,24	0,03	0,14
<i>C16#SNB-s2010_JFB_F?(?)</i>	25	0,26	0,07	0,16

population	<i>nb ind</i>	<i>He<sub>NE</sub></i>	<i>Ho<sub>NE</sub></i>	<i>He<sub>CA</sub></i>
<i>C21#b_JFB_F?(?)</i>	25	0,22	0,04	0,14
<i>C21_CHD_F6</i>	25	0,25	0,02	0,18
<i>C21_RAB_F6</i>	23	0,21	0,03	0,17
<i>C21_OLR_F6</i>	25	0,17	0,03	0,16
<i>C21_BRE_F5</i>	26	0,27	0,05	0,22
<i>C21_JSG_F5</i>	25	0,23	0,08	0,20
<i>C21_FRC_F6</i>	25	0,23	0,03	0,14
<i>C21#Sb-selec-R10_JFB_F6(2)</i>	25	0,00	0,00	0,03
<i>C32_JFB_F3</i>	25	0,26	0,12	0,16
<i>C32_ALP_F6</i>	25	0,28	0,02	0,18
<i>C32_JSG_F6</i>	26	0,32	0,04	0,22
<i>C32#S-FR_JFB_F?(?)</i>	21	0,23	0,02	0,13
<i>C32#S_JFB_F?(?)</i>	25	0,28	0,03	0,16
<i>C37#epi-long-crosse_JFB_F6(3)</i>	43	0,19	0,03	0,11
<i>C37_BRE_F5</i>	24	0,21	0,05	0,12
<i>C37_JMC_F5</i>	23	0,14	0,01	0,10
<i>C37_MAV_F6</i>	25	0,18	0,03	0,10
<i>C37_MPH_F4</i>	25	0,19	0,05	0,10
<i>C42_JFB_F3</i>	25	0,27	0,11	0,08
<i>C42_ROW_F5</i>	24	0,31	0,07	0,12
<i>C42_JMC_F6</i>	24	0,29	0,03	0,10
<i>C42_BRE_F5</i>	26	0,26	0,06	0,07
<i>C42_JSG_F5</i>	24	0,23	0,06	0,08
<i>C42#S_JFB_F?(?)</i>	24	0,26	0,08	0,05
<i>C67_JFB_F3</i>	25	0,27	0,12	0,18
<i>C67_OLR_F6</i>	24	0,19	0,03	0,17
<i>C67_JSG_F5</i>	25	0,19	0,04	0,17
<i>C77_ROW_F5</i>	25	0,31	0,07	0,15
<i>C77_OLR_F6</i>	20	0,20	0,02	0,10

population	<i>nb ind</i>	$He_{NE}$	$Ho_{NE}$	$He_{CA}$
<i>C82_FRC_F6</i>	25	0,15	0,02	0,15
<i>C82_FLM_F5</i>	24	0,15	0,03	0,17
<i>C82_OLR_F5</i>	26	0,19	0,04	0,16
<i>C8_BRE_F6</i>	25	0,25	0,05	0,08
<i>C8_RAB_F5</i>	25	0,27	0,05	0,10
<i>C8_JFB_F3</i>	25	0,27	0,13	0,08
<i>C8#S_JFB_F3(1)</i>	25	0,27	0,10	0,08
<i>Clomone_FLM_F8</i>	24	0,28	0,04	0,11
<i>Clomone_JFB_F8</i>	23	0,25	0,06	0,05
<i>Rock-and-Lot#Salvatorock</i>				
<i>_FRC_F8(7)</i>	25	0,16	0,01	0,10
<i>Rock-and-Lot#Salvatorock-S-JF</i>				
<i>_JFB_F?(?)</i>	25	0,14	0,02	0,07

*nb ind* : nombre d'individu dans la population,  $He$  : estimateur non biaisé de  $Ne_i^{178}$ ,  $Ho$  : hétérozygotie moyenne observée.

$NE$  : marqueurs neutres.  $CA$  : marqueurs dans des gènes candidats à la précocité de floraison.

Tableau 3.4 - AMOVA sur les marqueurs *NE*, les marqueurs *CA* et les deux sur le jeu de données correspondant aux descendants.

source de variation	<i>ddl</i>	<i>NE</i>			<i>CA</i>			<i>NE &amp; CA</i>		
		<i>SS</i>	%	<i>p-value</i>	<i>SS</i>	%	<i>p-value</i>	<i>SS</i>	%	<i>p-value</i>
type	3	1347,59	11,61	0,32	435,27	11,19	0,28	1779,45	11,5	0,31
germplasm	8	3135,01	27,02	9,93.10 <sup>-88</sup>	920,85	23,67	5,70.10 <sup>-74</sup>	4046,54	26,15	4,32.10 <sup>-84</sup>
population	37	1512,80	13,04	1,57.10 <sup>-38</sup>	514,51	13,22	6,51.10 <sup>-36</sup>	2024,89	13,09	7,99.10 <sup>-38</sup>
erreur	1143	5608,32	48,33		2020,05	51,92		7621,35	49,26	
total	1191	11603,73	100		3890,70	100		15472,24	100	

*NE* : marqueurs neutres. *CA* : marqueurs dans les gènes candidats. *NE & CA* : les deux types de marqueurs. *ddl* : degrés de liberté. *SS* : somme des carrés des écarts à la moyenne. % : pourcentage de variation. *p-value* : *p-value* pour chaque effet.

### 3.2.2 Organisation de la diversité intra-germplasm

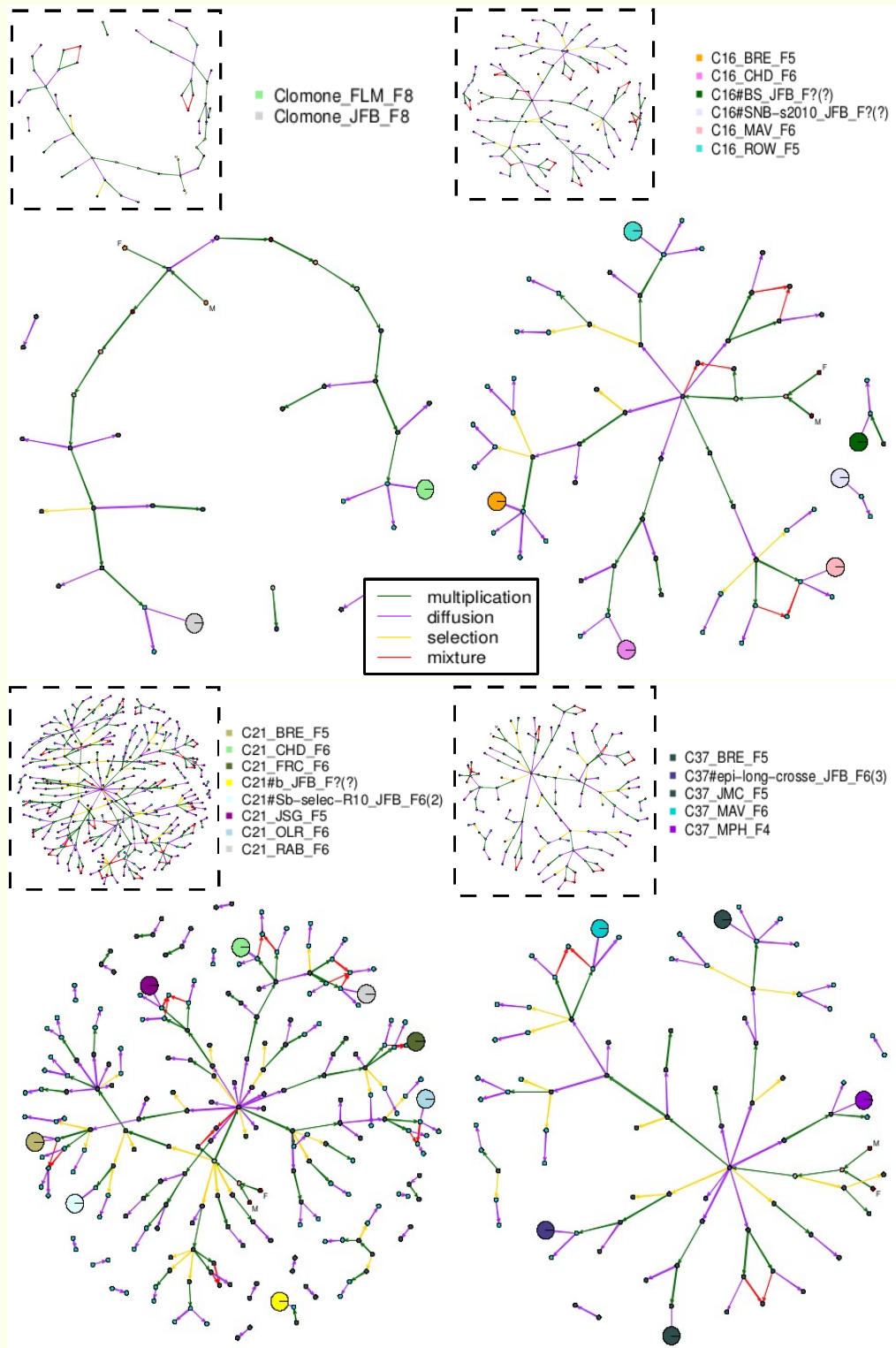
Les populations étudiées ne représentent qu'un petit échantillon des populations de chaque germplasm (Figure 3.8). Nous avons tenté de représenter les branches principales du réseau, et de couvrir une large diversité des pedigres.

Le dendrogramme basé sur les  $pF_{st}$  sur les marqueurs  $NE$  pour toutes les populations (Figure 3.9), l'ACP sur un sous-jeu de données comprenant, pour chaque germplasm, les parents et les descendants (Figure 3.10 & 3.11) et l'indice de diversité de Nei sur les marqueurs neutres ( $He_{NE}$ ) (Tableau 3.3) ont été utilisés dans la description des résultats ci dessous. L'indice de diversité pour les marqueurs  $CA$  ( $He_{CA}$ ) (Tableau 3.3) n'a été mentionné que quand ses valeurs étaient différentes au sein d'un germplasm.

#### *Clomone*

Les deux populations ont évolué dans deux fermes différentes ( $FLM$  et  $JFB$ ) (Figure 3.8). Elles sont apparues assez divergentes sur le dendrogramme (Figure 3.9). Sur l'ACP, elles ont formé deux groupes bien distincts qui englobent chacun des individus d'un des parents (*Touselle-Anone*) (Figure 3.10). L'autre parent (*Cloche*) a formé un groupe resserré encore plus différencié. Comme on l'a vu sur l'ACP, *Touselle-Anone* était très diversifiée ( $He_{NE}=0,2$ ) et certains de ses individus se rapprochaient de certains descendants (Figure 3.10). Les deux populations avaient un  $He_{NE}$  de 0,25 et 0,28, et une diversité sur les gènes candidats  $He_{CA}$  nettement plus faibles (0,05 et 0,11), signe d'une éventuelle sélection sur certains de ces gènes (Tableau 3.3) ou d'un polymorphisme réduit pour les gènes  $CA$  entre les deux parents. La différenciation plus forte sur les gènes candidats ( $pF_{st}=0,39$ ) comparée à la différence neutre ( $pF_{st}=0,29$ ) pourrait indiquer plutôt une sélection divergente à ces gènes entre les deux populations (Tableau 3.5).

Figure 3.8 - Réseau des relations entre lots de semences pour les germplasm *Clomone*, C16, C21 et C37.



Le réseau entouré de pointillé est le réseau total.  
 Une sous-partie du réseau et les relations précédant 2011 sont représentées pour plus de lisibilité.  
 Un point de couleur représente un lot de semences. Les lots de semences avec les points plus gros correspondent aux populations étudiées dans l'analyse moléculaire.  
 Une flèche représente une relation entre deux lots de semences (multiplication, diffusion, sélection, mélange).



Figure 3.9 - Dendrogramme basé sur les  $pF_{st}$  sur les marqueurs *CA* haut et *NE* en bas pour les descendants et les parents

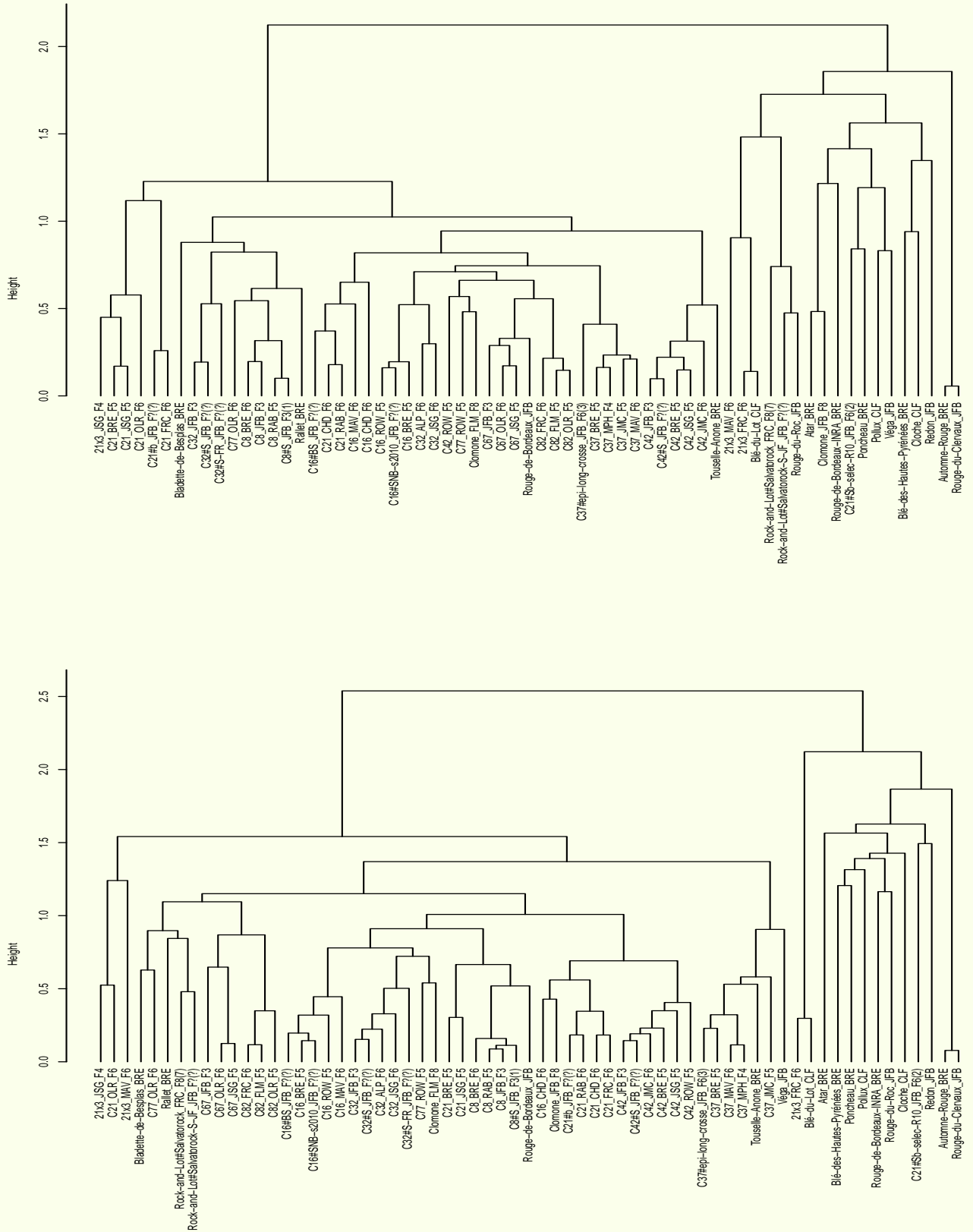
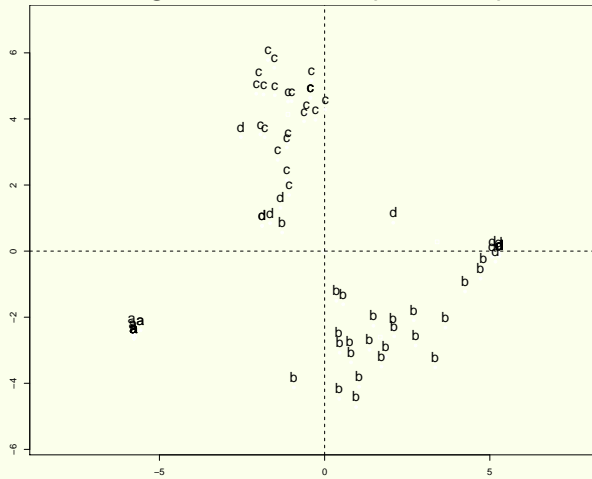
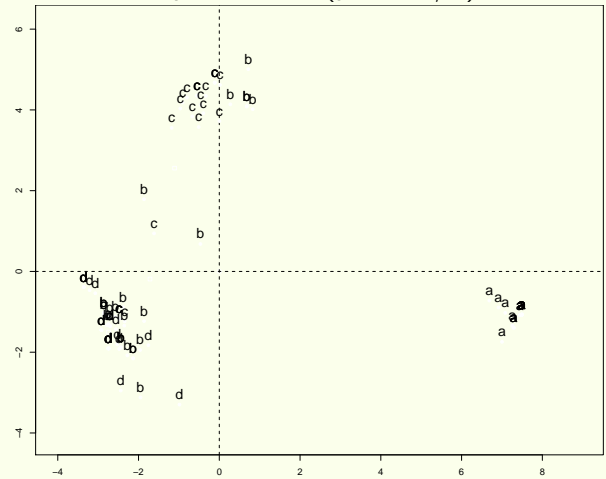


Figure 3.10 - ACP pour chaque famille sur les marqueurs *NE* (partie 1/2).



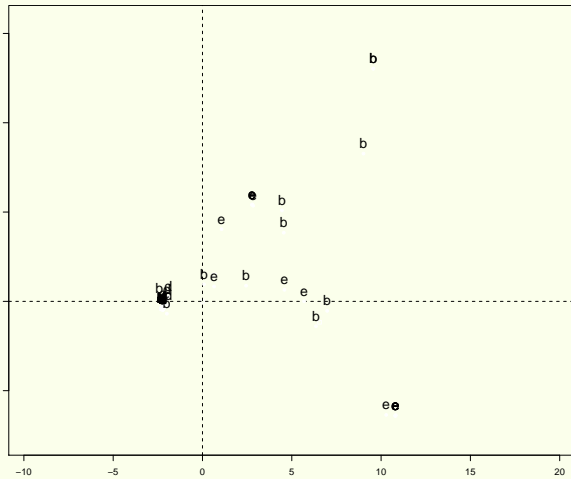
Dim1=28.2%  
Dim2=18.8%  
a. Cloche\_CLF

b. Clomone\_FLM\_F8  
c. Clomone\_JFB\_F8  
d. Touselle-Anone\_BRE



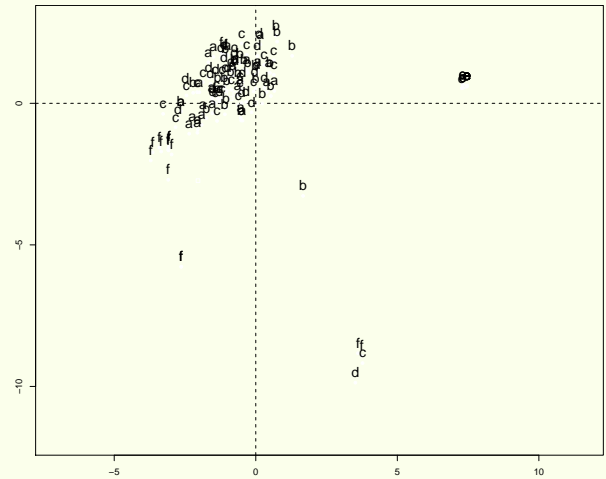
Dim1=40.9%  
Dim2=16.4%  
a. Blé-du-Lot\_CLF

b. Rock-and-Lot#Salvatorock\_FRC\_F8(7)  
c. Rock-and-Lot#Salvatorock-S-JF\_JFB\_F?(  
d. Rouge-du-Roc\_JFB



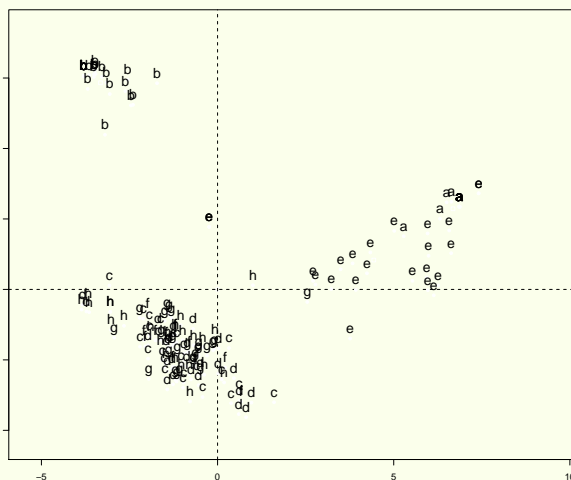
Dim1=40.8%  
Dim2=22.1%  
a. 21x3\_FRC\_F4  
b. 21x3\_JSG\_F2

c. 21x3\_MAV\_F4  
d. Blé-du-Lot\_CLF  
e. Rouge-de-Bordeaux\_JFB



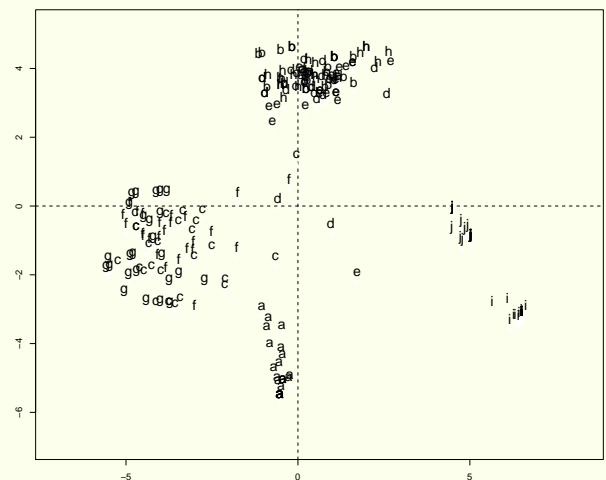
Dim1=22.8%  
Dim2=10.9%  
a. C8\_BRE\_F6  
b. C8\_JFB\_F3

c. C8\_RAB\_F5  
d. C8#S\_JFB\_F3(1)  
e. Cloche\_CLF  
f. Rallet\_BRE



Dim1=22.2%  
Dim2=17.2%  
a. Atar\_BRE  
b. Blé-des-Hautes-Pyrénées\_BRE  
c. C16\_BRE\_F5

d. C16#BS\_JFB\_F?(?)  
e. C16\_CHD\_F6  
f. C16\_MAV\_F2  
g. C16\_ROW\_F5  
h. C16#SNB-s2010\_JFB\_F?(?)



Dim1=24.6%  
Dim2=20.3%  
a. Blé-des-Hautes-Pyrénées\_BRE  
b. C21#b\_JFB\_F?(?)  
c. C21\_BRE\_F5  
d. C21\_CHD\_F6

e. C21\_FRC\_F6  
f. C21\_JSG\_F5  
g. C21\_OLR\_F3  
h. C21\_RAB\_F6  
i. C21#Sb-selec-R10\_JFB\_F6(2)  
j. Pollux\_CLF

*Rock-and-Lot*

Les deux populations partageaient le même cluster sur le dendrogramme, très éloignées de leurs parents (*Rouge-du-Roc* et *Blé-du-Lot*) (Figure 3.9). Ceci est confirmé par l'ACP où les populations se chevauchaient (Figure 3.10). Les populations sont apparues imbriquées l'une dans l'autre et très distinctes du parent *Blé-du-Lot*. La population non sélectionnée (*#Salvatorock*), présentait des individus proches du parent *Rouge-du-Roc* et d'autres plus éloignés. C'est près de ces derniers que l'on a pu trouver les individus de la population sélectionnée (*#Salvatorock-S-JF*). De façon cohérente, nous avons trouvé moins de diversité  $He_{NE}$  dans la population *#Salvatorock-S-JF*, issue de sélection, que dans la population vrac correspondante *#Salvatorock* (0,14 contre 0,16) (Tableau 3.3). Avec les marqueurs *CA*, les deux populations étaient plus proches de *Rouge-du-Roc*.

21 × 3

Sur le dendrogramme, nous avons vu que deux des populations étaient proches (*JSG*, *MAV*) et éloignées de la troisième (*FRC*), elle-même proche d'un des parents (*Blé-du-Lot*) (Figure 3.9). Les populations issues de chez *MAV* et *FRC*, très homogènes génétiquement ( $He_{NE} = 0$ ) étaient très proches de *Blé-du-Lot* (Figure 3.10). La population issue de chez *JSG*, plus diversifiée ( $He=0.22$ ), se distribue entre le groupe des individus de *Rouge-de-Bordeaux* et celui des deux autres populations et de *Blé-du-Lot*. Avec les marqueurs *CA*, les populations issues de *MAV* et de *FRC* étaient proches entre elles et de *Blé-du-Lot*. La population issue de chez *JSG* était éloignée des deux autres. De façon cohérente, nous avons trouvé beaucoup plus de diversité sur les  $He_{CA}$  chez *JSG* que chez *MAV* et *FRC* (0,16 contre 0,04 et 0,05) (Tableau 3.3).

C8

Sur le dendrogramme, ainsi que sur l'ACP, les quatre populations sont apparues particulièrement regroupées (Figure 3.10). Ces populations étaient éloignées de leurs parents mais plus proches de *Rallet* que *Cloche*. Sur l'ACP, quelques individus des descendants se rapprochaient d'individus de *Rallet*, lui-même très hétérogène ( $He_{NE}=0,18$ ). Les deux populations issues de chez *JFB* étaient les plus proches (Figure 3.9). Toutes les populations sont apparues hétérogènes génétiquement ( $He$  entre 0,25 et 0,27) (Tableau 3.3). Les populations issues de la sélection massale *#S* dans la *F2* et les populations issues du vrac correspondant avaient la même valeur élevée de  $He_{NE}$  (0,27). Les populations étaient plus proches de *Rallet* avec les marqueurs *CA*.

C16

Sur le dendrogramme, cinq populations sont retrouvées dans le même cluster, la population restante (issue de *CHD*) était relativement peu éloignée (Figure 3.9). Par contre, ces populations se sont avérées éloignées de leurs parents (*Atar* et *Ble-des-hautes-Pyrénées*) excepté pour *CHD* qui est resté proche d'*Atar*. L'ACP confirme ces observations avec cinq populations assez éloignées de leurs parents et proches entre elles et la sixième issue de la ferme *CHD* qui se distribue autour du parent *Atar* (Figure 3.10). La diversité  $He_{NE}$  des populations était élevée variant entre 0,24 et 0,30 (avec la valeur maximum pour *CHD*). Les sélections, *#BS* et *#SNB-s2010*, n'avaient pas un  $He$  plus faible que les autres (Tableau 3.3). Avec les marqueurs *CA*, cinq populations étaient ensemble et *#BS* était plus éloignée.

C21

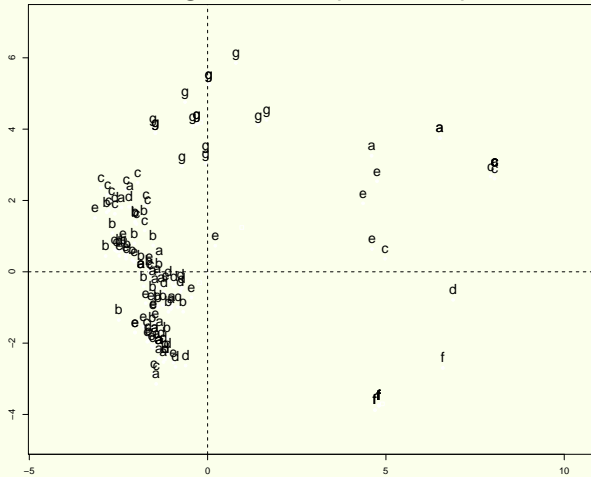
Le dendrogramme montre quatre groupes parmi le germplasm C21 (Figure 3.9). Le groupe 1 contient les populations issues de chez *BRE* et *JSG*; le groupe 2 contient *#Sb-select-R10\_JFB*; le groupe 3 contient la population issue de chez *OLR*. le groupe 4 contient les populations issues de chez *CHD*, *FRC*, *RAB* et *#b\_JFB*. Sur l'ACP, nous avons observé une structure proche avec cinq groupes : chacun des deux parents bien séparés et la sélection de *JFB*, *Sb-select-R10*, assez proche de *Pollux*, le groupe des quatre populations (*CHD*, *FRC*, *RAB* et *#b\_JFB*) puis un groupe avec les populations issues de chez *BRE*, *JSG* et *OLR* (Figure 3.10).

Toutes les populations ont présenté un  $He_{NE}$  compris entre 0,17 et 0,27, sauf *Sb-select-R10\_JFB*, qui est issue d'une sélection massale et avec une diversité nulle. L'autre sélection, *#b\_JFB* avait un  $He_{NE}$  de 0,22 (Tableau 3.3).

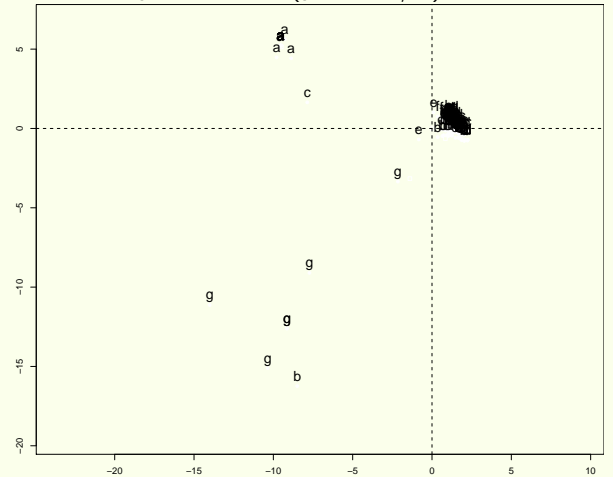
Le dendrogramme construit sur la base des  $pF_{st}$  entre marqueurs *CA* a révélé un profil assez différent avec un premier regroupe contenant les populations issues de chez *BRE*, *OLR*, *JSG*, *FRC* et *#b*, un deuxième regroupe contenant les populations issues de chez *CHD* et *RAB*. Le troisième regroupe contenait la population *#Sb-select-R10\_JFB*.

La diversité  $He_{CA}$  était la plus faible pour les sélections : 0,03 pour *#Sb-select-R10\_JFB* et 0,14 pour *#b*. Pour les autres populations, réparties sur différentes fermes,  $He_{CA}$  variait entre 0,14 et 0,22, c'est à dire sensiblement du même ordre que  $He_{NE}$  (Tableau 3.3).

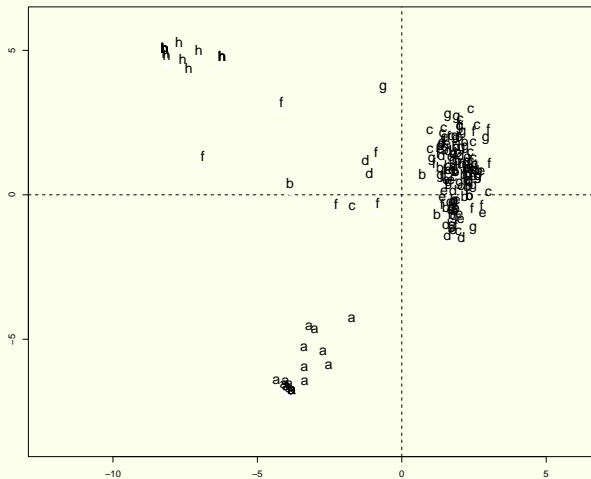
Figure 3.11 - pour chaque famille sur les marqueurs *NE* (partie 2/2).



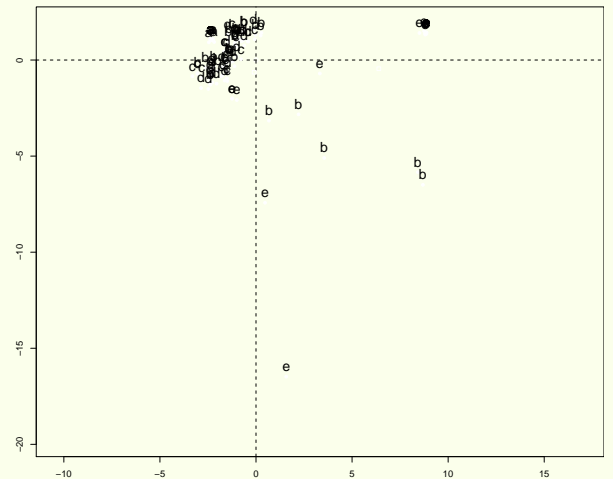
Dim1=20.6%  
Dim2=11.7%  
a. C32\_ALP\_F6  
b. C32\_JFB\_F3  
c. C32\_JSG\_F6  
d. C32#S-FR\_JFB\_F?(?)  
e. C32#S\_JFB\_F?(?)  
f. Poncheau\_BRE  
g. Véga\_JFB



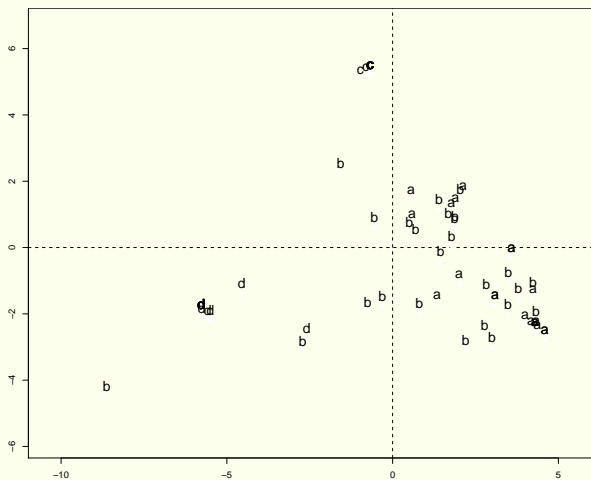
Dim1=29%  
Dim2=18%  
a. Atar\_BRE  
b. C37\_BRE\_F5  
c. C37#epi-long-crosse\_JFB\_F6(3)  
d. C37\_JMC\_F5  
e. C37\_MAV\_F6  
f. C37\_MPH\_F4  
g. Touselle-Anone\_BRE



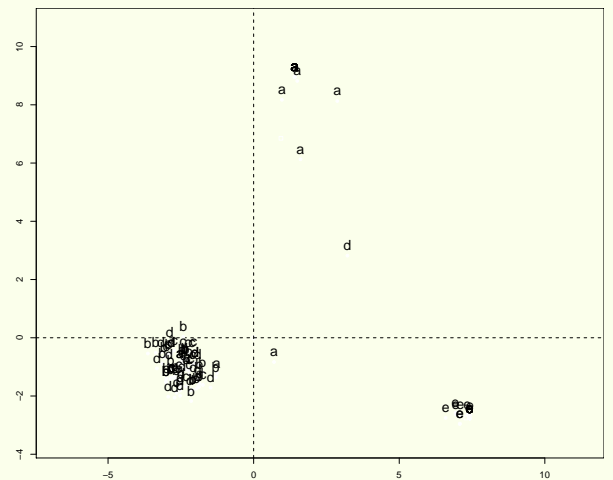
Dim1=24.4%  
Dim2=19.3%  
a. Blé-des-Hautes-Pyrénées\_BRE  
b. C42\_BRE\_F5  
c. C42\_JFB\_F3  
d. C42\_JMC\_F2  
e. C42\_JSG\_F5  
f. C42\_ROW\_F5  
g. C42#S\_JFB\_F?(?)  
h. Pollux\_CLF



Dim1=31.3%  
Dim2=10.7%  
a. Automne-Rouge\_BRE  
b. C67\_JFB\_F3  
c. C67\_JSG\_F5  
d. C67\_OLR\_F6  
e. Touselle-Anone\_BRE



Dim1=26.2%  
Dim2=18.5%  
a. C77\_OLR\_F6  
b. C77\_ROW\_F5  
c. Cloche\_CLF  
d. Rouge-de-Bordeaux-INRA\_BRE



Dim1=30.7%  
Dim2=28.6%  
a. Bladette-de-Besplas\_BRE  
b. C82\_FLM\_F5  
c. C82\_FRC\_F6  
d. C82\_OLR\_F5  
e. Redon\_JFB

C32

Sur le dendrogramme, les quatre populations étaient regroupées dans le même cluster avec les deux parents *Véga* et *Poncheau* éloignés (Figure 3.9). L'ACP confirme l'imbrication des cinq populations avec en plus une dizaine d'individus dispersés entre les deux parents, *Poncheau* et *Véga* étaient opposés (Figure 3.11).

La diversité  $He$  des descendants était élevée entre 0,26 et 0,32. La population  $\#S\_FR\_JFB$  issue d'une sélection massale avait la diversité la plus faible ( $He_{NE}=0,23$ ) alors que la population  $\#S\_JFB$  issue d'une autre sélection massale avait un  $He_{NE}$  de 0,28 (Tableau 3.3).

Avec le dendrogramme avec les marqueurs *CA*, nous avons observé deux groupes avec les populations issues de *JFB* d'un côté et les populations issues de chez *ALP* et *JSG* de l'autre.

La diversité  $He_{CA}$  tend également à être plus faible pour les populations issues de sélection. Comme pour les marqueurs neutres, ce sont les populations de *JSG* et *ALP* qui sont apparues les plus hétérogènes avec un  $He_{CA}$  de 0,18 et 0,22 (Tableau 3.3). Les populations issues de chez *ALP* et *JSG* avaient toutes les deux, comparées à la population de *JFB* leurs  $pF_{st}$  sur les marqueurs *CA* plus élevés que leurs  $pF_{st}$  avec les marqueurs *NE*, signe de sélection divergente intra-germplasm (Figure 3.12).

C37

Sur le dendrogramme construit à partir des  $pF_{st}$  neutres, les cinq populations et un des parents (*Touselle-Anone*) apparaissaient dans le même cluster (Figure 3.9). L'ACP confirme la grande proximité des cinq populations qui se superposent, à part deux individus, et sont intermédiaires aux deux parents (Figure 3.11). La diversité  $He_{NE}$  des descendants variait entre 0,18 et 0,21. On peut noter quelques individus de *Touselle-Anone*, très diversifiée ( $He_{NE}=0,2$ ), qui se rapprochent du cluster des cinq populations (Tableau 3.3).

C42

Sur le dendrogramme, les six populations appartenait au même cluster (Figure 3.9) et nous les retrouvons très imbriquées dans le premier plan de l'ACP (Figure 3.11), intermédiaires entre les deux parents qui s'opposent sur l'axe 2. La diversité  $He_{NE}$  des populations était forte, variant entre 0,23 et 0,31 (Tableau 3.3). Au sein de la population issue de chez *ROW*, certains individus se rapprochaient de *Pollux* (ACP). Comme dans les deux cas précédent, le pool de diversité des descendants est apparu relativement inter-

médiaire entre les deux parents. Avec les marqueurs *CA*, la population issue de chez *ROW* était plus éloignée des autres. Celle-ci a un  $He_{CA}$  plus élevé que les autres populations. Les  $pF_{st}$  sur les marqueurs *CA*, comparés aux marqueurs *NE*, sont plus élevés pour la population issue de chez *ROW* comparé aux populations issues de chez *JMC*, *BRE* et *JSG* (Figure 3.12).

#### C67

Dans le dendrogramme, les trois populations ont été regroupées sur le même cluster (Figure 3.9). Les parents (*Automne-Rouge* et *Touselle-Anone*) étaient assez divergents des descendants. Sur l'ACP, les descendants sont apparus plus proches d'*Automne-Rouge* sauf pour la population *C67\_JFB* (génération *F3*) plus dispersée qui se rapprochait de *Touselle-Anone* (Figure 3.11).

La diversité  $He_{NE}$  des descendants variait entre 0,19 et 0,27 (Tableau 3.3).

#### C77

Sur le dendrogramme, les deux populations issues de chez *OLR* et *ROW* apparaissent éloignées l'une de l'autre ainsi que des parents (Figure 3.9). Sur l'ACP, nous avons remarqué la grande dispersion des deux populations notamment *ROW* qui se recoupe l'une de l'autre et rejoignent le parent *Rouge-de-Bordeaux* (Figure 3.11). C'est cohérent avec la forte diversité intra-population de la descendance *ROW* ( $He_{NE} = 0,31$ ) et dans une moindre mesure *OLR* ( $He_{NE} = 0,20$ ) (Tableau 3.3).

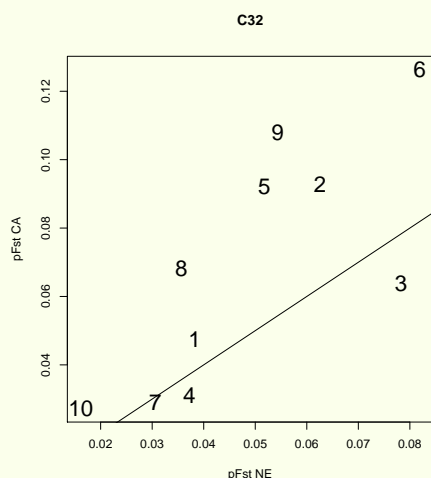
La diversité  $He_{CA}$  est plus faible chez *OLR* (0,10) que chez *ROW* (0,15).

#### C82

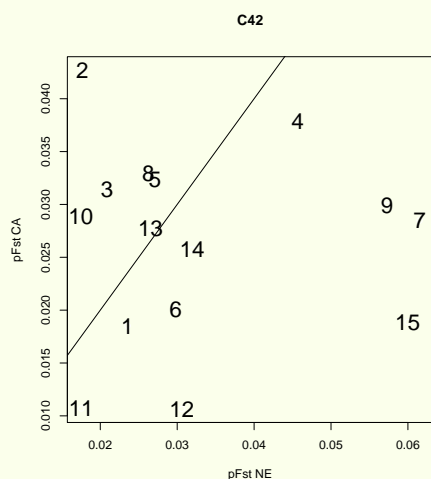
Les trois populations se sont retrouvées dans le même cluster sur le dendrogramme construit à partir des  $pF_{st}$  neutres à proximité de *Bladette-de-Besplas*, un des parents, l'autre parent, *Redon*, étant éloigné (Figure 3.9). Sur l'ACP, les trois populations se superposent avec quelques individus de *Bladette-de-besplas* à proximité (Figure 3.11). Les deux parents se trouvent sinon opposés sur l'axe 2 et distants de ces trois populations.

La diversité  $He_{NE}$  des descendants variait entre 0,15 et 0,19 (Tableau 3.3). Les trois paires de populations deux à deux présentaient des  $pF_{st}$  sur les marqueurs *CA* plus fort que les  $pF_{st}$  sur les marqueurs *NE* (Figure 3.12).

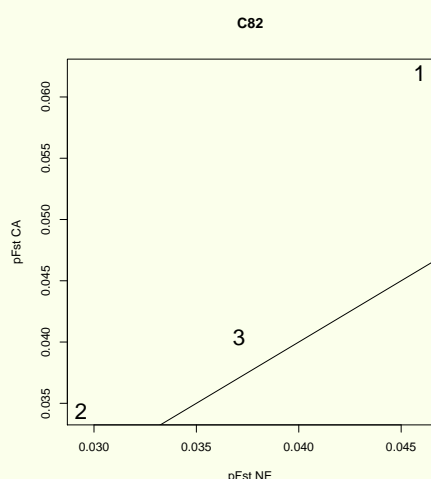
Figure 3.12 - Relation entre les  $pF_{st}$  sur les marqueurs *CA* en fonction des  $pF_{st}$  sur les marqueurs *NE* pour les germplasms *C32*, *C42* et *C82*.



1. C32\_ALP\_F6:C32\_JFB\_F3
2. C32\_JSG\_F6:C32\_JFB\_F3
3. C32\_JSG\_F6:C32\_ALP\_F6
4. C32#S-FR\_JFB\_F?(?):C32\_JFB\_F3
5. C32#S-FR\_JFB\_F?(?):C32\_ALP\_F6
6. C32#S-FR\_JFB\_F?(?):C32\_JSG\_F6
7. C32#S\_JFB\_F?(?):C32\_JFB\_F3
8. C32#S\_JFB\_F?(?):C32\_ALP\_F6
9. C32#S\_JFB\_F?(?):C32\_JSG\_F6
10. C32#S\_JFB\_F?(?):C32#S-FR\_JFB\_F?(?)



1. C42\_ROW\_F5:C42\_JFB\_F3
2. C42\_JMC\_F6:C42\_JFB\_F3
3. C42\_JMC\_F6:C42\_ROW\_F5
4. C42\_BRE\_F5:C42\_JFB\_F3
5. C42\_BRE\_F5:C42\_ROW\_F5
6. C42\_BRE\_F5:C42\_JMC\_F6
7. C42\_JSG\_F5:C42\_JFB\_F3
8. C42\_JSG\_F5:C42\_ROW\_F5
9. C42\_JSG\_F5:C42\_JMC\_F6
10. C42\_JSG\_F5:C42\_BRE\_F5
11. C42#S\_JFB\_F?(?):C42\_JFB\_F3
12. C42#S\_JFB\_F?(?):C42\_ROW\_F5
13. C42#S\_JFB\_F?(?):C42\_JMC\_F6
14. C42#S\_JFB\_F?(?):C42\_BRE\_F5
15. C42#S\_JFB\_F?(?):C42\_JSG\_F5



1. C82\_FLM\_F5:C82\_FRC\_F6
2. C82\_OLR\_F5:C82\_FRC\_F6
3. C82\_OLR\_F5:C82\_FLM\_F5

La ligne diagonale représente la droite  $pF_{stNE} = pF_{stCA}$

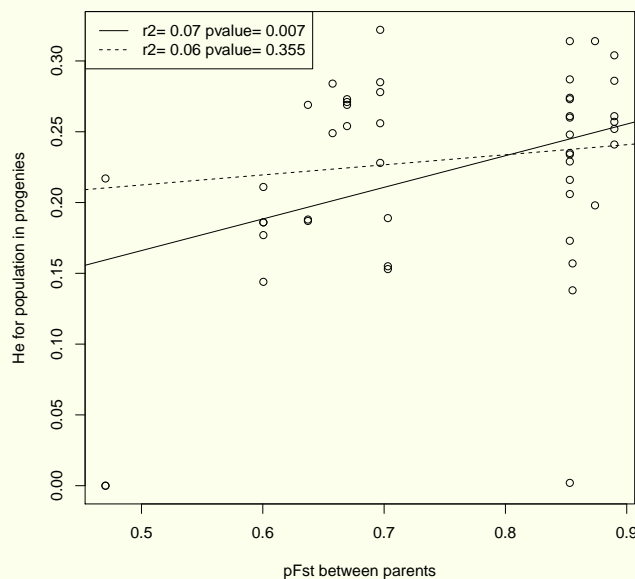
L'objectif des croisements était de générer une variabilité large et nouvelle pour la sélection. Nous disposons de deux indicateurs pour évaluer ce potentiel : la diversité eu



sein des populations issues d'un même croisement et leur différenciation.

Nous avons observé une faible relation positive entre le  $pF_{st}$  entre les parents et la diversité au sein des populations qui en dérivent ( $He_{NE}$ ) (Figure 3.13). Des croisements entre parents plus différenciés ont plus de chances d'apporter des allèles différents, ce qui augmente la diversité intra-population chez les descendants. Les populations étant issues de croisements, on s'attend à ce que les marqueurs polymorphes (et bialléliques) présentent au mieux des allèles en fréquences égales à 0,5, ce qui correspond à  $He_{max} = 0,5$ . Les valeurs inférieures observées sont dues à la présence de locus non polymorphes initialement entre les parents ou bien à de la dérive ou de la sélection sur les populations.

Figure 3.13 - Relation entre les  $pF_{st}$  entre les parents et la diversité  $He_{NE}$  des populations issues de ces parents.



Données sur les marqueurs  $NE$ .

La ligne pleine prend en compte tous les points.

La ligne pointillée ne prend pas en compte les deux points correspondants à  $He_{NE} = 0$ .

L'hétérogénéité au sein des germplasms peut être évaluée sur la base de la différenciation moyenne entre populations au sein d'un germplasm ( $pF_{st}$  sur les marqueurs  $NE$  et  $CA$ ). Globalement, le  $pF_{st}$  moyen toutes populations confondues était de 0,39 sur les marqueurs  $NE$  et de 0,33 sur les marqueurs  $CA$ . Dans tous les cas, les  $pF_{st}$  moyens sur les marqueurs  $NE$  au sein des germplasms étaient plus faibles, entre 0,02 et 0,31 (Tableau 3.5). Pour les marqueurs  $CA$ , les  $pF_{st}$  moyens étaient plus faibles dans tous les cas (entre

0,02 et 0,24) sauf pour *Clomone* où il était supérieur avec une valeur de 0,39. Ces résultats sont cohérents avec les résultats de l'AMOVA où l'effet germplasm était très fort.

Au niveau du germplasm, deux groupes peuvent être différenciés pour les  $pF_{st}$  moyen sur les marqueurs *NE*. Le premier groupe rassemblait des germplasms très diversifiés : *Rock-and-Lot*, *Clomone* et *C21* avec respectivement un  $pF_{st}$  moyen de 0,16, 0,29, et 0,31. Un deuxième groupe rassemblait les autres germplasms avec un  $pF_{st}$  moyen compris entre 0,02 et 0,10. Les valeurs les plus fortes correspondent soit aux populations les plus anciennes (*Clomone* et *Rock-and-Lot*), soit à des germplasms avec plusieurs populations sélectionnées fortement (*C21*).

Tableau 3.5 -  $pF_{st}$  moyen pour chaque germplasm issu des croisements pour les marqueurs *NE* et *CA*.

germplasm	$pF_{st}(\text{NE})$ moyen	$pF_{st}(\text{CA})$ moyen
<i>Clomone</i>	0,29	0,39
<i>Rock-and-Lot</i>	0,16	0,07
21 × 3	0,09	0,05
<i>C8</i>	0,02	0,02
<i>C16</i>	0,10	0,09
<i>C21</i>	0,31	0,24
<i>C32</i>	0,05	0,07
<i>C37</i>	0,04	0,04
<i>C42</i>	0,03	0,03
<i>C67</i>	0,03	0,02
<i>C77</i>	0,04	0,02
<i>C82</i>	0,04	0,05

### 3.2.3 Description des descendants au niveau phénotypique

#### 3.2.3.1 Les types de germplasm

D'après le modèle 3.1, l'effet type de germplasm n'est pas significatif sauf pour le poids et la courbe de l'épi. On retrouve cela avec le modèle 3.2, où le type était plus significatif pour la hauteur et la longueur de l'épi. Nous retrouvons ces quatre variables significatives avec le modèle 3.3 (Tableau 3.6). Le fait que le facteur type devienne significatif dans les

modèles 3.2 et 3.3 alors qu'il ne l'était pas dans le modèle 3.1 indique que ce sont des différences entre germplasms qui peuvent intervenir dans ces modèles.

Pour le poids de l'épi, avec le modèle 3.1 trois groupes de significativité ont été détectés. Les poids des épis du type  $VP \times VP$  sont significativement plus faibles que les types  $VM \times VM$  et  $VM \times VP$ . Le type  $VA \times VP$  était significativement plus faible que  $VM \times VM$ . Le type  $VM \times VP$  était significativement plus faible que le type  $VM \times VM$ . Le type  $VM \times VM$  a le poids des épis le plus fort et était seul dans son groupe (Figure 3.14).

Pour la courbure de l'épi, les types  $VM \times VM$  et  $VM \times VP$  avaient des épis significativement plus courbés que les épis des types  $VA \times VP$  et  $VP \times VP$  (Figure 3.14).

L'ACP montre une grande variabilité au sein de chaque type de germplasm (Figure 3.15). Le type  $VM \times VP$  est le plus diversifié.  $VP$  est le type le plus diversifié, suivi de  $VA$  et  $VM$  qui étaient aussi diversifiés. Les populations  $VM \times VP$  explorent presque la totalité de l'espace, suivies par les populations  $VP$ , les populations  $VA \times VP$  et les populations  $VM \times VM$ . Tous les types chevauchent les espaces de leurs parents sauf  $VM \times VM$ , qui se différencie.

### 3.2.3.2 Les germplasms

L'effet germplasm, hiérarchisé dans le type de germplasm, est significatif pour toutes les variables sauf pour le poids et la courbe de l'épi (Tableau 3.6).

Néanmoins, pour le poids de l'épi, trois groupes significativement différents ont été trouvés (Figure 3.14). Les groupes se chevauchent avec 14 germplasms sur 25 dans le premier groupe et 20 germplasms sur 25 dans le dernier groupe. Un germplasm se détache en tête : le  $C32$ , à l'inverse,  $C13$  est le germplasm avec le poids de l'épi le plus faible.

Alors que le premier appartient au type  $VM \times VM$ , le dernier appartient au type  $VP \times VP$ . Cependant, on distingue aussi un germplasm de type  $VA \times VP$  en cinquième position ( $C5$ ).

Pour la date de milieu d'épiaison, on distingue trois groupes significativement différents. Les germplasms  $C5$  et  $21 \times 3$  sont les plus précoces de tous les germplasms. 21 germplasms font partie du groupe le plus tardif. Pour ces derniers cas, une partie de la variabilité est captée par le type de germplasm, le germplasm à l'intérieur du type varie peu.

Les germplasms sont très diversifiés sur l'ACP (Figure 3.15). Il n'y a pas de structure

claire qui se dégage.

Certains germplasms, comme le *C21* ont une grande variabilité et explorent un espace important. D'autres, comme le *C37* ont une variabilité plus faible et explorent un espace du premier axe de l'ACP plus restreint.

### 3.2.3.3 Les sélections

Dans le modèle 3.2, l'effet lié à la version, sélectionnée ou non, n'est jamais significatif sauf pour les barbes : la sélection favorise les populations plus barbues (Tableau 3.6). L'effet version est à la limite de la signification pour la courbure de l'épi. Lors des comparaisons de moyennes, nous observons que les germplasms issus de sélection avaient des courbes plus fortes et que les épis étaient plus longs.

### 3.2.3.4 Les fermes

Dans le modèle 3.3, l'effet ferme, qui correspond à la ferme où ont été récoltées les populations, n'a jamais été significatif sur aucune variable (Tableau 3.6). Notons que ce modèle cherchait à détecter s'il y avait des effets communs entre les populations issues de germplasms différents et cultivés sur une même ferme. Il ne teste pas si les populations issues d'un même germplasm et cultivées sur des fermes différentes se différencient.

Tableau 3.6 - *p-value* des effets pour les modèles 3.1, 3.2 et 3.3 sur les données phénotypiques.

modèle	facteur	barbes	couleur	courbe	nb d'épillets	nb d'épillets stériles	hauteur	poids de l'épi	longueur de l'épi	LSD	pmg	date d'épaison début	date d'épaison milieu	date d'épaison fin
(3.1)	$\alpha_i$	0.941	0.521	<b>0.000</b>	0.693	0.544	0.129	<b>0.000</b>	0.201	0.292	0.513	0.758	0.770	0.777
(3.1)	$\beta_j$	<b>0.000</b>	<b>0.000</b>	0.307	<b>0.000</b>	<b>0.000</b>	0.066	0.688	<b>0.000</b>	<b>0.002</b>	<b>0.024</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
(3.2)	$\alpha_i$	0.718	0.057	<b>0.000</b>	0.203	0.206	<b>0.033</b>	<b>0.000</b>	<b>0.005</b>	0.060	0.296	0.344	0.318	0.341
(3.2)	$\nu_j$	<b>0.000</b>	0.123	0.225	0.871	0.221	0.569	0.248	0.052	0.711	0.646	0.118	0.304	0.989
(3.3)	$\alpha_i$	0.748	0.064	<b>0.000</b>	0.220	0.200	<b>0.032</b>	<b>0.000</b>	<b>0.008</b>	0.058	0.282	0.376	0.338	0.346
(3.3)	$\theta_j$	0.205	0.581	0.505	0.817	0.321	0.426	0.419	0.879	0.399	0.295	0.830	0.754	0.587

*p-value* associé à chaque facteur des différents modèles à 0,001 près (**en gras** : *p-value* < 5%).

$\alpha$  : effet du type de germplasm.

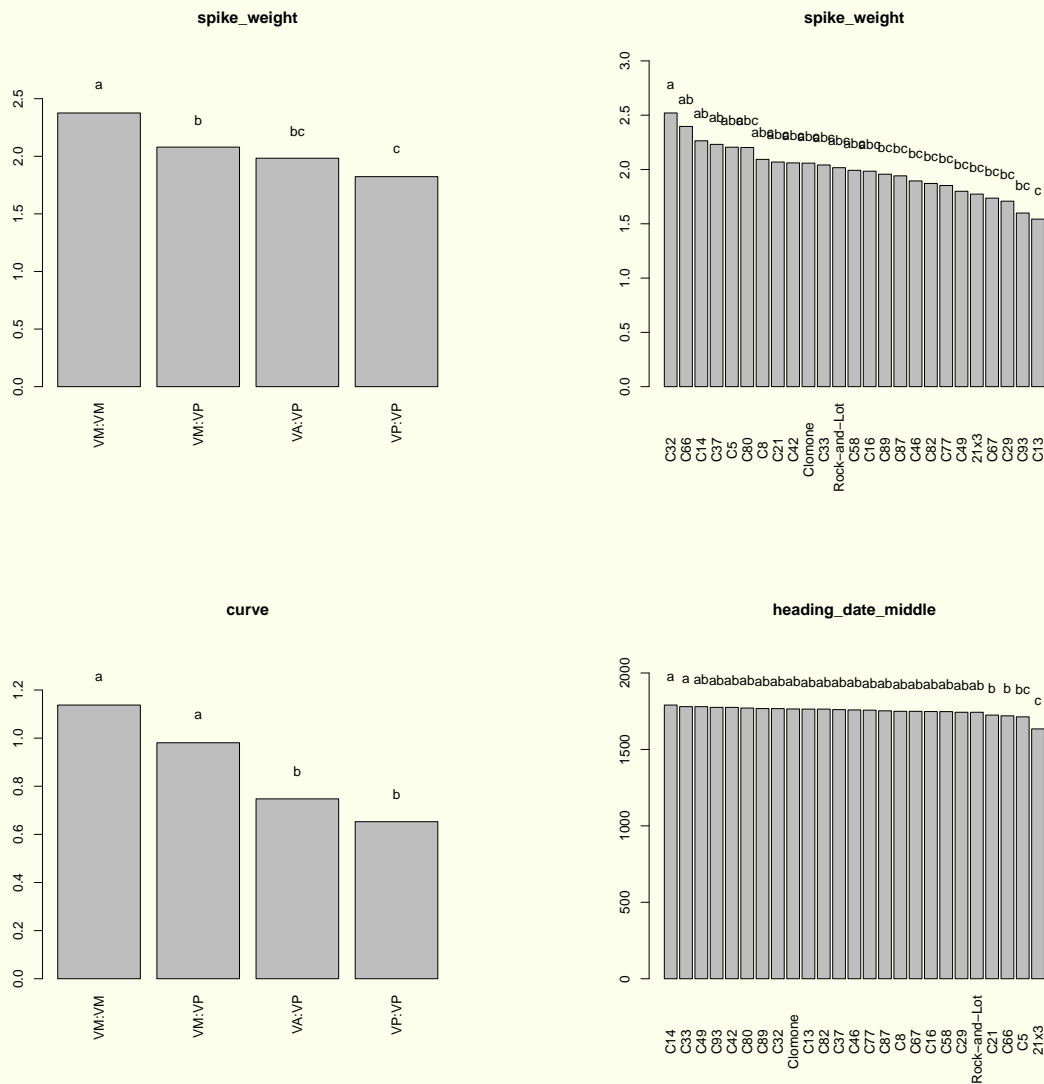
$\beta$  : effet du germplasm.

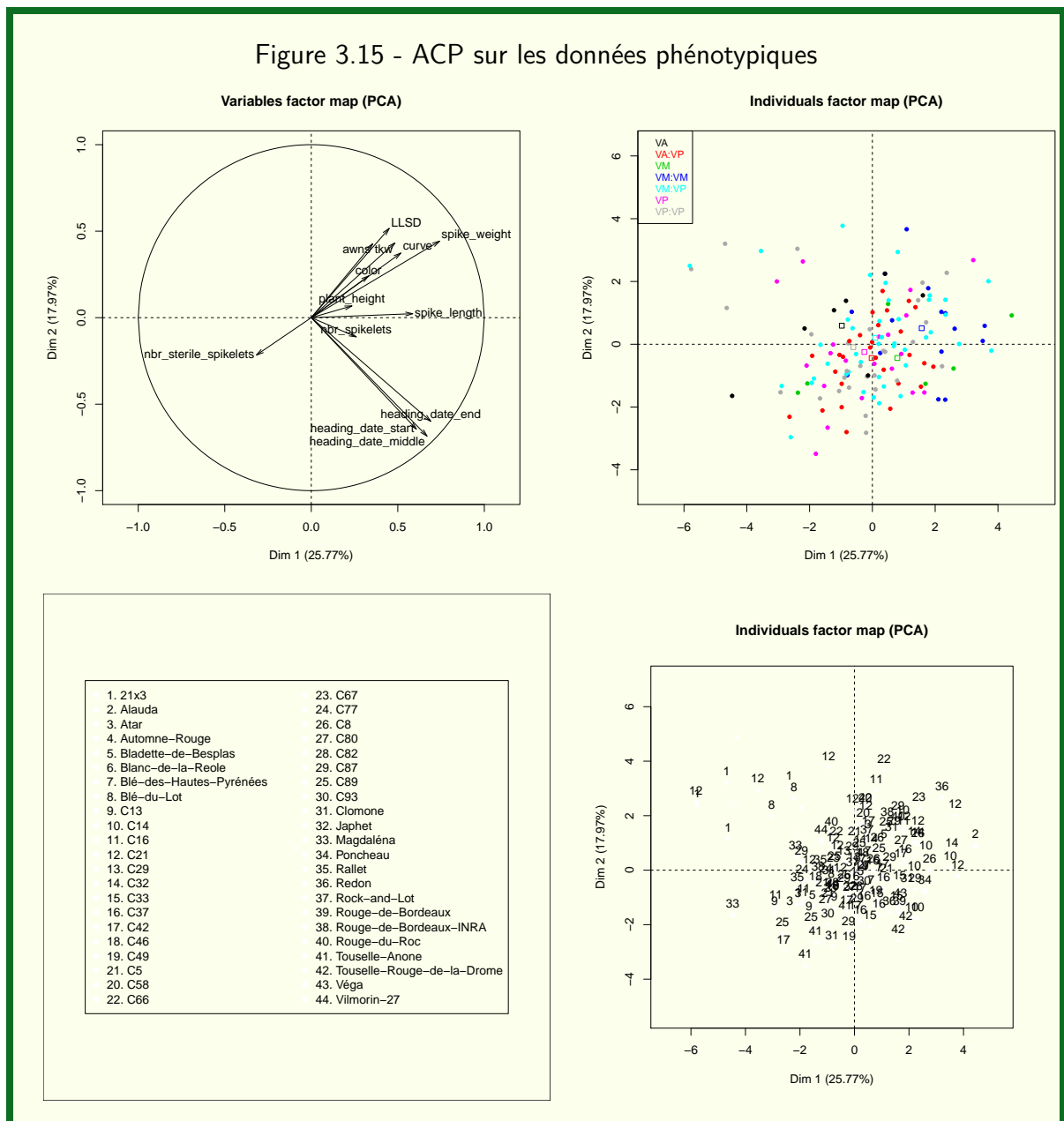
$\nu$  : effet de la version sélectionnée ou non sélectionnée.

$\theta$  : effet de la ferme.

( $\nu$ ) : nombre de groupe significativement différents estimés.

Figure 3.14 - Comparaisons de moyennes pour le poids de l'épi, la date d'épiaison et la courbure à partir du modèle 3.1





### 3.3 Discussion

Ce chapitre visait à mieux caractériser la diversité génétique qui a été créée par différents croisements et distribuée dans 25 fermes aux environnements et aux pratiques contrastées dans le cadre d'un programme de sélection participative sur le blé tendre. Ces populations ont été, pour la grande majorité, diffusées en génération  $F_3$  dans le réseau de fermes. On s'attend donc à ce que, par l'intermédiaire des différentes forces évolutives jouant sur la ségrégation, il y ait une évolution dans les fréquences alléliques des populations.

Dans un premier temps, nous discutons les résultats obtenus sur les échantillons analysés. Dans un deuxième temps, nous replaçons ces résultats dans un cadre plus global de gestion des ressources génétiques.

### 3.3.1 Évolution de la diversité dans le programme de sélection participative

#### 3.3.1.1 Organisation de la diversité génétique

##### Choix des méthodes pour l'analyse

Nous avons étudié la différenciation des populations issues de croisements. Pour ce faire, nous avons utilisé le  $F_{st}$  qui est un indicateur de la différenciation entre populations.  $F_{st} = var(inter)/(var(intra) + var(inter))$ , avec  $var$  la variance. Le  $F_{st}$  est sensible à la diversité intra-population. Si deux populations ont une diversité intra-population importante, alors on sous-estime les différences entre deux populations. Il est alors délicat de comparer des  $F_{st}$  entre populations présentant des niveaux de diversité très différents. Cela a été parfois le cas où la diversité intra-population était variable, de nulle à très forte. Pour nous affranchir de cela, nous avons également utilisé d'autres approches de type analyses multivariées. A partir des matrices des  $pF_{st}$ , nous avons regardé des dendrogrammes basés sur les distances génétiques de Reynolds, les résultats variaient peu.

##### Le choix des parents

La plupart des parents analysés se structuraient en plusieurs groupes distincts. Certains parents avaient une variabilité intra-population importante et se retrouvaient dans plusieurs groupes. Au contraire, *Automne-Rouge* et *Rouge-de-Clervaux* avaient une diversité de Nei nulle et appartenaient au même groupe. Il nous a été impossible de savoir si cette similarité était due à un problème d'étiquette, qui aurait conduit à une confusion entre le nom et les germplasms ou alors que ces deux germplasms étaient réellement très proches génétiquement.

Les différences de diversité entre les variétés parentes peuvent s'expliquer par des histoires évolutives différentes. Les parents provenant de la banque de graines de l'INRA de Clermont-Ferrand (*CLF*) se sont révélés peu diversifiés, sans doute à la suite de méthodes de conservation qui ont tendance à diminuer la diversité intra-variétale<sup>229</sup>. De la même manière, certains paysans tel que *BRE* ont des pratiques de gestion des ressources génétiques très proches de celles des banques de graines, avec élimination des hors-types et multiplication dans des micro-parcelles d'un mètre carré, ce qui peut créer un goulot d'étranglement et entraîner une baisse de la diversité. D'autres paysans, au contraire, ont des pratiques moins strictes permettant une certaine évolution<sup>64</sup>. Ces différences de pratiques sont à l'origine de la différence de diversité observée entre les deux versions de *Rouge-de-Bordeaux*. Alors que la version *INRA* est issue de *BRE* et de la banque de *CLF*, l'autre est issue du réseau de paysans<sup>231</sup>.



Les parents qui ont été utilisés au départ du projet ont été multipliés dans la ferme de *JFB* cinq années avant les croisements. Thomas<sup>229</sup> a montré que des populations issues de la banque de Clermont, et gérées au sein du RSP pouvaient gagner en diversité. Les parents utilisés lors des croisements étaient donc peut-être plus diversifiés que ceux venant de *CLF*.

Dans le projet, chaque nouveau germplasm était issu du croisement entre une à trois plantes de chacun des deux parents. Les croisements réalisés n'ont donc pas donné naissance à de vrais hybrides F1. Ainsi, les individus utilisés au sein d'un parent pour faire le croisement ne vont pas forcément apporter un seul haplotype à la descendance. Créer de nouveaux germplasms à partir d'individus hétérogènes, et issus de la même population au départ, va augmenter la diversité au sein des descendants. Par exemple, nous avons observé que *Touzelle-Anone* était très hétérogène et que ses descendants exploraient de larges gammes de diversité (c'est le cas par exemple pour *Clomone* et *C37*). Lorsque chaque parent appartenait à un seul groupe génétique détecté par la DAPC, leur diversité était a priori plus faible et donc leurs descendants présentaient moins de diversité.

Au delà de cette diversité intra-population, la distance entre les deux parents a pu permettre d'explorer plus de nouvelles combinaisons alléliques. Ceci est illustré par une relation faible toutefois entre le  $pF_{st}$  entre les parents et la diversité au sein des populations qui en dérivent ( $He_{NE}$ ) (Figure 3.13).

*JFB*, qui a été à l'initiative de ces croisements, a donc choisi des parents assez différenciés génétiquement. Il a pu caractériser a priori cette diversité grâce à l'histoire des parents et à leur observation pendant cinq années sur sa ferme dans les champs et dans le fournil. L'analyse multivariée sur les données phénotypiques ne montre néanmoins pas de clair regroupement de tel ou tel parent. Pour évaluer les populations, les paysans utilisent, au delà des critères agronomiques « classiques », des critères plus globaux qui décrivent la plante ou le peuplement tels que la brillance dans les champs ou le comportement de la farine en panification. Aucun critère n'a été mesuré concernant le comportement en panification ou la qualité nutritionnelle, alors que ces critères ont été utilisés dans le choix des parents au départ du projet. La prise en compte de ces caractères permettrait certainement de mieux discriminer les germplasms.

### Les descendants

Cette expérience n'étudie qu'une faible partie de la diversité de chaque germplasm (Figure 3.8). De même, seulement 12 germplasms ont été étudiés au niveau moléculaire

alors que sur le réseau en comptait 313 fin 2013.

Sur un jeu de données regroupant les descendants, l'AMOVA n'a pas montré d'effet lié au type de germplasm. L'origine des parents choisis ne semble pas avoir structuré la diversité des parents et leurs descendances que dans une faible mesure (Figure 3.5). Comme nous l'avons vu, la diversité spécifique présente entre et à l'intérieur des parents a peut-être eu plus d'importance que l'appartenance à certains types. Ce résultat n'était pas surprenant au vu des parents qui se structurent en germplasms et non en type (DAPC sur la Figure 3.3).

Le niveau germplasm et le niveau population expliquent une part importante de la variabilité génétique, mais le niveau individuel, qui correspond à la diversité à l'intérieur des populations, capte autour de 50% de la variabilité, ce qui est cohérent avec les valeurs de  $H_{e_{NE}}$  très fortes trouvées dans les descendances. Les populations créées sont jeunes, encore en ségrégation, et soumises à différentes pressions évolutives. En effet, à part deux populations, aucune population n'a montré une hétérozygotie nulle et les indices  $H_{e_{NE}}$  étaient importants pour beaucoup de populations. La diversité des populations ( $H_{e_{NE}}$ ) n'a pas diminué significativement avec le nombre de générations. Un fort  $H_{e_{NE}}$  signifie que la diversité créée initialement lors des croisements a été maintenue au cours des différentes générations. Les populations étaient donc encore hétérogènes et ont conservé un potentiel pour la sélection important.

La structure génétique de ces germplasms est liée à la diversité de départ entre les parents, et également à leurs histoires évolutives. Ces histoires dépendent d'au moins deux événements : les reproductions dans les fermes et la sélection massale intra-population. Les fermes présentent des conditions pédo-climatiques et des pratiques agronomiques contrastées qui sont autant de forces de sélection qui vont avoir un impact sur la diversité des populations. Nous avons observé une tendance à ce que globalement, plus la distance sur le réseau entre deux populations était grande, c'est à dire qu'elles étaient séparées par de nombreuses générations dans des environnements différents, et plus la différenciation entre les populations était grande (Figure 3.7). Afin de mieux comprendre la divergence entre les populations, il serait intéressant de lier ces événements de multiplication à des données pédo-climatiques ou des pratiques particulières.

Afin de décrire la capacité des germplasms à répondre à des sélections différentes, nous avons étudié la différenciation entre populations intra-germplasm. Ensuite nous avons séparé les germplasms en deux groupes selon le niveau de différenciation entre les

populations au sein d'un germplasm.

#### Populations peu différenciées au sein d'un germplasm

Certaines populations au sein d'un germplasm sont restées proches génétiquement, c'est le cas pour sept germplasms sur les 12 étudiés : *Rock-and-Lot*, *C32*, *C42*, *C37*, *C8*, *C67* et *C82*.

Pour les germplasms, *C8*, *C37*, *C67* et *Rock-and-Lot*, les descendants se sont plutôt rapprochés d'un des parents. Ce parent a pu apporter un fond génétique caractéristique qui structure le germplasm par rapport aux autres germplasms. Des effets d'échantillonnage dès la *F2* dus à la sélection naturelle, la sélection massale-intra-population et la dérive, ont pu conduire à faire évoluer certaines fréquences alléliques vers un des parents. Il est également possible que les croisements n'aient pas eu lieu correctement et que des autofécondations aient augmenté la fréquence d'un des deux parents dans la descendance. Enfin, il ne faut pas écarter l'hypothèse que les marqueurs n'étaient peut être pas représentatifs de la diversité. Cependant, leur nombre est élevé et ils étaient répartis sur tout le génome.

Ce n'est pas le cas pour les germplasms *C32*, *C42* et *C82* qui se sont retrouvés aussi distants de leurs deux parents. Dans ce cas, la recombinaison a bien permis de développer une diversité nouvelle et originale par rapport aux parents, ce qui était l'un des objectifs de *JFB* au début du projet.

Au delà des parents, qui confèrent un fond génétique particulier aux germplasms, les populations au sein des germplasms ont des histoires différentes qui modèlent leur structure.

Par exemple, les deux populations de *Rock-and-Lot* sont issues d'une première sélection (*#Salvatorock*) effectuée par Salvatore Ceccarelli lors de sa visite chez *JFB* en 2007 en génération *F3*. La population ainsi sélectionnée a été multipliée chez *JFB* puis diffusée chez *FRC*. En parallèle, *JFB* a réalisé une sélection dans *Rock-and-Lot#Salvatorock* pour donner *Rock-and-Lot#Salvatorock-S-JF*. On ne sait pas quand la sélection chez *JFB* a été faite. Sachant que les populations sur ces deux fermes sont très proches, il est possible que la première sélection ait appliqué un goulot d'étranglement créant ainsi un fond génétique typique. La deuxième sélection de *JFB* n'a pas pu se détacher de ce fond et est donc restée proche de la version non sélectionnée cultivées chez *FRC*.

Pour *C32*, toutes les populations étaient assez proches, mais une des sélections était

proche du vrac ( $\#S$ ) alors que l'autre sélection ( $\#S-FR$ ) était plus éloignée. Ces deux sélections ont mobilisé une diversité différente. Le peu de différences entre les populations peut être lié au faible nombre de générations sur ces fermes depuis le dernier événement de diffusion : il n'y a eu qu'une seule génération chez *ALP* et *JSG*. Néanmoins, les informations issues des marqueurs *CA* montrent une structuration avec les populations issues de *JFB* d'un côté et les populations issues de chez *ALP* et *JSG* de l'autre. Les données phénotypiques nous indiquent que le nombre de degrés jour où la date de milieu épiaison a eu lieu était de 1780 pour *JFB* contre 1762 pour *ALP* et 1752 pour *JSG*. La comparaison entre les  $pF_{st}$  sur les marqueurs *CA* et *NE* nous renseigne sur une possible sélection divergente. Ces deux derniers environnements sont plus au Nord (Saint Nazaire et Clermont-Ferrand) que *JFB* (Agen) avec des climats très différents. Ces populations, plus au Nord, ont pu évoluer vers une gamme de précocité différente. Cela paraît surprenant en seulement une seule génération de multiplication. Néanmoins, Khan<sup>148</sup> a montré que des populations pouvaient diverger sur des pas de temps très courts, de l'ordre de trois années.

Dans le germplasm *C42*, nous observons aussi des  $pF_{st}$  aux gènes candidats sensiblement plus élevés que les  $pF_{st}$  *NE* pour la population issue de chez *ROW* par rapport aux populations issues de chez *JMC*, *BRE* et *JSG*.

Pour le germplasm *C82*, la population issue de chez *OLR* a été semée deux années au printemps alors que celle de *FRC* a été semée trois années en hiver et celle de *FLM* a été semée deux années en hiver. *FLM* est situé près d'Angers et *FRC* près de Nîmes. Ce sont deux environnements très contrastés au niveau climatique. Nous avons observé que les  $pF_{st}$  sur les marqueurs neutres étaient plus faibles que les  $pF_{st}$  sur les marqueurs *CA* pour chaque paire de populations. Une sélection divergente intra-population a pu avoir lieu entre ces populations. Le semis de printemps deux années de suite chez *OLR* peut expliquer cette divergence. De même que les conditions climatiques très contrastées entre *FLM* et *FRC*.

### **Populations différenciées au sein d'un germplasm**

Nous n'avons pas observé de structuration claire en groupe pour cinq germplasms :  $21 \times 3$ , *C16*, *Clomone*, *C21* et *C77*. Alors que le premier a des populations proches de l'un ou l'autre des parents, le second a une population proche d'un parent et toutes les autres combinant la diversité des deux parents, enfin les trois derniers vont combiner de la diversité de chacun des parents.

Les deux populations de *Clomone* sont séparées par 14 multiplications : 7 chez *FLM* et 7 chez *JFB*. Ces générations successives ont augmenté la divergence par dérive et par sélection entre ces deux populations. Ainsi le  $pF_{st}$  sur les marqueurs neutres était très élevé (0,29, Tableau 3.5). De plus, le  $pF_{st}$  sur les marqueurs *CA* était plus élevé que le  $pF_{st}$  sur les marqueurs *NE* (0,39 contre 0,29), indice d'une possible sélection divergente entre ces deux populations. Néanmoins, elles restent très diversifiées avec une valeur de  $He_{NE}$  élevée. Les deux environnements sont assez contrastés avec une ferme près d'Angers et l'autre près d'Agen, ce qui peut expliquer une différenciation sur des gènes de précocité.

Pour le germplasm *C21*, deux groupes retiennent notre attention. La population *C21#Sb-select-R10\_JFB* se démarque des autres populations. La sélection a permis de garder une diversité qui n'est pas présente dans les autres versions. La sélection, sur cet exemple, est diversifiante. Néanmoins, si cette sélection a permis de conserver une partie de la diversité qui aurait été perdue sinon, la diversité au sein de cette population a été quasiment épuisée avec un  $He_{NE}$  de 0 et un  $He_{CA}$  de 0,03. Un groupe contient quatre populations issues de chez *CHD*, *FRC* et *RAB* et une sélection *C21#b* issue de chez *JFB*. Il est possible qu'elles soient issues en génération *F3* d'une même population. Nous n'avons pas eu l'information lors de la sélection du *C21#b*. Il est possible que cette sélection réalisée en génération *F2* (nous avons eu cette information après l'analyse des données) ait été diffusée dans les fermes *CHD*, *FRC* et *RAB*. Les analyses génétiques peuvent amener à affiner, voire parfois reconsidérer, certaines informations issues des entretiens avec les paysans.

Les deux populations de *C77* se sont différenciées entre *OLR* et *ROW* pour les marqueurs *NE* et les marqueurs *CA*. La diversité au sein des marqueurs *CA* est plus faible que celle au sein des marqueurs *NE*. Ces deux populations sont issues d'une même population de départ. Alors que chez *OLR*, trois multiplications, dont deux en semis de printemps ont été réalisées, chez *ROW*, deux multiplications avec semis d'hiver ont été réalisées. Ces différentes dates de semis pourraient expliquer la différenciation au niveau des gènes de précocité et la perte de diversité dans la population d'*OLR* par rapport à celle de *ROW* ( $He_{CA}$  de 0,10 dans le premier cas et de 0,15 dans l'autre). Ces semis de printemps ont été réalisés à cause de mauvaises conditions climatiques qui ont décalé le semis.

### 3.3.1.2 Analyse phénotypique

Au même titre que les analyses moléculaires, les analyses phénotypiques visaient à identifier les niveaux d'organisation structurants pour la variabilité des différents caractères. Le modèle 3.1 a la même structure que l'AMOVA avec un niveau de moins : celui des population au sein des germplasms. Cette différence vient de la méthode utilisée dans l'analyse phénotypique. En effet, nous avons corrigé les données avec le modèle bayésien dans un premier temps pour compenser le fait que toutes les populations n'étaient pas répétées. Le modèle a estimé une valeur moyenne par population. Afin d'être homogène, même si l'on avait pu ici estimer les effets populations grâce aux deux répétitions, les données d'épiaison au Moulon sont des valeurs moyennes par population.

L'effet lié à la ferme d'origine n'était jamais significatif. Toutes les populations avaient souvent circulé d'une ferme à l'autre et n'avaient pas le même nombre de génération sur la dernière ferme où elles ont été récoltées. C'est sans doute une des raisons pour lesquelles on ne trouve pas d'effet significatif, une autre raison étant que les germplasms étaient très différents les uns des autres et qu'ils ont pu répondre plus ou moins vite dans un sens ou un autre selon la variabilité disponible initialement.

Nous n'avons trouvé aucun effet de la sélection, sauf pour les barbes. Cela est peut paraître surprenant en comparaison avec les résultats du chapitre 2, où l'on trouvait des réponses à la sélection sur plus de caractères. Cependant, nous ne testons ici que l'existence d'effets communs à toutes les populations sélectionnées, celles-ci étant relativement peu nombreuses et partant de germplasms très différents. Il pourrait être intéressant de tester un modèle avec l'interaction sélection  $\times$  type de germplasm.

Il y a un effet lié aux germplasms pour la majorité des caractères étudiés. Ceci confirme qu'une très large diversité phénotypique a été créée et que ces nouvelles populations peuvent constituer une base de sélection pour les paysans.

Les croisements comportant une ou deux *VM* présentent un potentiel plus important d'amélioration de la production en grains par épi, ce qui correspond aux attendus. Cela sera à étudier sur les données du rendement, qui n'ont pas encore été analysées. Ces données manquent sur le réseau<sup>ii</sup>, elles seront donc analysées de manière plus classique en corrigeant dans un premier temps par l'effet bloc, puis par l'effet ferme avant de faire les analyses de variances et les ACP.

Par contre, en ce qui concerne la hauteur, tous les types de croisements semblent avoir

---

ii. On revient sur ce point dans le chapitre 4.

le même potentiel. Il faudra regarder au niveau de la verse comment ils se comportent (nous avons les données mais elles n'ont pas été analysées).

Notons qu'un croisement  $VP \times VA$  ( $C5$ ) présente toutefois une moyenne élevée du poids de l'épi (5<sup>ème</sup> rang), donc il existe aussi du potentiel pour ces variétés plus anciennes et de pays.

L'ACP n'a pas permis de détecter de structure particulière sauf au niveau des types de germplasms. Le type  $VM \times VP$  est le plus diversifié. Il pourrait donc être intéressant de privilégier ce type de croisement pour générer une large diversité phénotypique.

### 3.3.2 La place du programme de sélection participative dans la gestion des ressources génétiques

Le programme de sélection participative mis en place a créé de nouvelles populations qui sont gérées par les paysans en métapopulations. En effet, à partir d'une population créée par croisement, des sous-populations ont été diffusées et ont évolué au gré des pratiques des paysans et des environnements. Ces populations hétérogènes, issues d'une même population de départ, peuvent être considérées comme une métapopulation. Cette métapopulation, qui correspond à un germplasm dans notre cas, va évoluer en réponse aux forces évolutives : pression de sélection dans les fermes, sélection intra-population, dérive, migration et mutation.

Les pressions de sélection naturelle dans les fermes sont liées aux conditions pédo-climatiques mais également aux pratiques des paysans. Nous avons vu que certaines populations pouvaient se différencier au sein d'un même germplasm pour les marqueurs situés dans des gènes contrôlant la précocité ou d'autres caractères par rapport aux marqueurs neutres.

En règle générale, les populations issues de sélection massale étaient génétiquement proches des populations non sélectionnées correspondantes et n'ont pas vu leur diversité intra-population diminuer. C'est ce que l'on a observé par exemple pour la population  $C8\#S\_JFB$  qui est issue d'une sélection dans la population  $C8\_JFB$ . Parfois, la sélection a pu conduire à une population très différente des autres comme dans le cas de  $C21\#Sb-select-R10\_JFB$ , mais la diversité intra-population a alors été très réduite.

La dérive est liée aux effets d'échantillonnage. D'après Enjalbert et al.<sup>82</sup>, pour le blé tendre, une taille de 10 000 individus paraît un bon compromis pour limiter l'effet de la dérive. Dans notre projet de sélection participative, la taille des micro-parcelles était autour 5 m<sup>2</sup> (de 1m<sup>2</sup> à plus de 10m<sup>2</sup>). Avec une densité moyenne de 150 plantes au m<sup>2</sup>,

on peut arriver à environ 750 plantes, ce qui est faible. Nous observons quelques exemple où la diversité sur les marqueurs neutres a diminué avec le temps. Par exemple pour le *C67*,  $He_{NE}$  est passé de 0,27 dans la *F3* à 0,19 dans la *F5* et dans la *F6*. A l'inverse, dans d'autres populations, la diversité a été trouvée plus ou moins constante entre les générations, par exemple pour les familles *C8*, *C32* et *C42*. La dérive peut également être mise en lien avec des pratiques des paysans. Par exemple, une année, *JMC* a semé ses populations sur une seule ligne. Cela a été le cas pour la population *C37\_JMC* qui a montré une diversité  $He_{NE}$  plus faible que les autres populations du germplasm.

Si les populations sont isolées trop longtemps, leur valeur sélective peut diminuer par l'accumulation de fardeau génétique et la perte d'allèles favorables. Il est donc nécessaire qu'il y ait des migrations entre les sous-populations afin de restaurer la variabilité génétique<sup>128</sup>. Dans notre projet, la migration de pollen ou de graines lors du triage et du semis, a pu compenser la relativement faible taille des populations, et permettre qu'elles maintiennent une diversité intra-population importante. Ces migrations apportent de nouveaux allèles qui peuvent ré-introduire une certaine diversité. De plus, dès que certaines populations conviennent aux paysans, ils les multiplient dans des parcelles de plus grande taille afin de les cultiver en champ, en « pur » ou en mélange.

Nous avons trouvé une corrélation négative entre le taux d'hétérozygotie observée et la génération. C'est cohérent avec la théorie. Néanmoins, en 3<sup>ème</sup> et 4<sup>ème</sup> génération, le taux était plus faible qu'attendu et en génération plus avancée, il était plus fort. Pour ces générations précoces, l'hétérozygotie observée a pu être due à un effet d'échantillonnage. En génération plus avancée, le taux d'hétérozygotie supérieur à l'attendu étaient probablement dû à des croisements résiduels qui ont permis le maintien d'un niveau d'hétérozygotie non nul. En effet, les paysans dans le projet de sélection participative évaluent les populations côté à côté et des échanges de pollen ont pu avoir lieu entre micro-parcelles. Le blé se croise avec un taux autour de 5%<sup>81</sup> et les populations étaient très hétérogènes (cf valeur de  $He$ ). Par conséquent, il est probable que des recombinaisons efficaces aient eu lieu et que des individus différents se soient croisés. Une autre explication pourrait être l'existence d'un avantage de vigueur aux plantes les plus hétérozygotes (hétérosis) qui permettrait la maintien de l'hétérozygotie<sup>79</sup>.

De même, les paysans échangent des semences entre eux dans le cadre du projet de sélection participative, mais aussi au delà<sup>229</sup>. Lors de ces échanges, de nouveaux allèles arrivent dans la ferme et peuvent « contaminer » les populations, participant ainsi au maintien de la diversité au sein des métapopulations. Ces échanges de semences et de



pollen sont au coeur d'une gestion efficace des ressources génétiques. Ils doivent être encouragés entre les paysans. Nous avons observé une multiplication des échanges au fur et à mesure de l'avancement du projet. Les données que l'équipe de recherche apportent aux paysans leurs permettent de sélectionner de nouvelles populations sur le réseau et ainsi de faire circuler la diversité. Les méthodes développées pendant le projet de sélection participative s'insèrent donc dans une gestion dynamique des ressources génétiques en favorisant la migration des semences et la culture d'une grande diversité de populations.

La structure du réseau de fermes au sein du projet de sélection participative peut être comparée à celui de l'expérience de gestion dynamique en stations initié par l'INRA dans les années 1980<sup>80</sup>. Rhone et al.<sup>202</sup> ont montré qu'avec cette gestion, des mutations pouvaient apporter de nouveaux allèles liés à la précocité.

Notre étude ne permet pas de détecter d'éventuelles mutations, mais il est probable que certaines se fixent dans la métapopulation. En effet, en 2012, il y avait plus de 900 populations sur le réseau. Pour se donner une idée, avec des micro parcelles de 5 m<sup>2</sup> avec une densité de 150 plantes au m<sup>2</sup>, on peut arriver à environ  $900 \times 5 \times 150 = 675\,000$  plantes évaluées. Si le taux de mutation à un locus est de  $10^{-6}$ , cela fait 0,675 mutation à chaque locus sur l'ensemble des populations. Le réseau d'essais peut générer des mutations qui pourront être sélectionnées par les paysans<sup>119</sup>.

La sélection participative, en favorisant l'utilisation de variétés hétérogènes dans des environnements très contrastés, participe à la gestion des ressources génétiques. Il est largement reconnu que les pratiques des paysans, notamment dans les systèmes vivriers des pays du Sud, participent à la gestion des ressources génétiques<sup>191;230</sup>. Sur un programme de sélection participative sur le maïs,<sup>247</sup> ont montré que la diversité génétique ne diminuait pas, mais restait stable.

### 3.3.3 Perspectives

Ce chapitre est loin d'être terminé. Les résultats ont été générés à la fin de la thèse, et il m'a manqué du temps pour en faire une analyse poussée. Il faudra notamment comprendre pourquoi les effets populations sur les protéines n'ont pas pu être estimés avec le modèle bayésien. Il faudra aussi analyser les données de rendement et de verse qui ont été collectées lors de l'évaluation.

De plus, une étude bibliographique plus poussée sera nécessaire pour comparer avec d'autres études moléculaires réalisées sur des populations issues de programmes de sélection participative.

Enfin, une prise de recul sera indispensable pour passer d'un catalogue germplasm par germplam à une analyse plus globale de l'évolution de la diversité des germplasms.



## 4 . Description et analyse de la méthodologie mise en place

Organisations sociales dans lesquelles s'inscrivent les acteurs

Co-construction de la méthodologie

Innovations organisationnelles et génétiques

Evolution des pratiques des acteurs

Ce chapitre décrit et analyse la méthodologie et le mode d'organisation mis en place dans le cadre du projet de sélection participative. On peut analyser la méthodologie en considérant deux niveaux. Le niveau collectif, basé sur la co-construction d'une méthodologie de la sélection et de la gestion des ressources génétiques entre les acteurs, va créer une base de travail commune à tous. Le niveau individuel, où chaque paysan a une grande flexibilité dans ses choix. Cet équilibre entre contraintes liées au collectif et flexibilité pour l'individu permet de générer de l'innovation sur mesure. L'innovation organisationnelle engendre les innovations génétiques. L'organisation se traduit par un fonctionnement en réseau dynamique et une co-construction entre les acteurs du projet. Les innovations génétiques se traduisent à travers la création de nouvelles variétés-populations adaptées à une grande diversité de système agroécologique et de nouvelles ressources génétiques gérées de manière dynamique dans les fermes.

Tout d'abord, les différentes organisations sociales qui regroupent les acteurs du projet sont décrites. Les acteurs de chaque organisation sociale, de chaque espace social, partagent des règles, des enjeux. Ensuite, la co-construction de la méthodologie entre les acteurs est décrite et analysée. Puis, les différentes innovations générées par la méthodologie sont présentées. Enfin, j'aborderai l'évolution des pratiques dans les différentes organisations sociales.

## 4.1 Organisations sociales dans lesquelles s'inscrivent les acteurs

### 4.1.1 Présentation des différentes organisations sociales

Dans notre projet, on peut distinguer trois types d'acteurs. Le premier regroupe les paysans et le second les associations : toutes les deux correspondent à la société civile (partie 1.3.3.1). Le troisième regroupe les membres de l'équipe de recherche. Ces acteurs font partie d'un espace social qui leur est propre.

Le premier espace est composé des paysans. Par la suite, on appellera cet espace social « paysans ». Les « paysans » englobent une grande diversité d'acteurs : paysans-meuniers-boulangers, paysans-meuniers, paysans-meuniers-éleveurs, paysans. Dans tous les cas, ces acteurs se revendiquent comme « paysans » (cf partie 1.3.3.1). Les paysans sont insérés dans un territoire caractérisé par son environnement agroécologique et socio-économique. L'enjeu pour les paysans est de développer des variétés adaptées à leur demande et de se réapproprier la sélection et la gestion des semences pour gagner en autonomie.

La deuxième espace est composé par les associations du groupe blé du Réseau Semences Paysannes (RSP). Dans la suite, on appellera cet espace social le groupe « blé RSP ». Le groupe « blé RSP » rassemble les personnes qui assurent l'animation et la gestion administrative des maisons des semences paysannes<sup>1</sup>. Toutes les maisons des semences paysannes n'ont pas forcément d'animateurs (Tableau 2.1). L'enjeu pour ces associations est de développer une méthodologie qui permette de créer de nouvelles variétés de blé tendre, adaptées aux pratiques des paysans et de conserver la biodiversité cultivée. Au delà de ces méthodes, elles souhaitent développer des outils qui puissent faciliter la mise en place de cette méthodologie. Les dates de création des associations varient ; elles ont plus ou moins d'expérience dans la gestion et la sélection des semences. Chaque maison des semences a ses règles de fonctionnement. Par exemple, elles peuvent différer sur les conventions mises en place, sur l'indemnisation ou non de la préparation des semences, sur les pratiques d'échanges, etc. Nous reviendrons sur cette diversité dans la partie 5.2.1.

Le troisième espace social est composé des chercheurs. Dans notre cas, il s'agit de l'équipe DEAP et d'autres chercheurs qui participent au projet. Dans la suite on appellera cet espace social l'équipe « DEAP ». L'équipe « DEAP » englobe toutes les personnes de l'équipe de recherche DEAP : stagiaires, doctorants, techniciens et chercheurs. Au delà de l'équipe, Y. de Oliveira qui est bio-informaticien et O. David qui est statisticien participent au projet. L'enjeu pour l'équipe de recherche est de mettre en place une démarche

---

i. On appelle maison des semences paysannes les associations régionales

scientifique, faire progresser les connaissances dans le domaine de la méthodologie de la sélection et publier les résultats de ces travaux dans des revues scientifiques. De plus, les chercheurs doivent accompagner au mieux les paysans et les maisons des semences paysannes pour développer des variétés-populations, adaptées à des pratiques agricoles durables, et mettre en place une gestion dynamique de la diversité cultivée.

Le RSP, créé en 2003, a un rôle de coordination et de mise en réseau des acteurs du groupe « blé RSP ». Il a par exemple organisé dans le cadre du groupe « blé RSP » une visite chez un paysan-boulangier en Bretagne en 2004, suivie en 2005 d'une autre visite dans le Sud-Ouest. Le RSP a une action de coordination au delà du groupe blé. En 2012, par exemple, il a organisé les « Rencontres Internationales des Maisons des Semences Paysannes » sur le thème des modes d'organisation collective pour gérer la biodiversité cultivée afin de fédérer différentes initiatives régionales en France et dans d'autres pays.

#### 4.1.2 L'objet frontière « biodiversité »

Le monde des semences peut être représenté par une chaîne avec de multiples maillons. Chaque maillon représente un espace social spécifique. Faire de la recherche en isolant chacun de ces maillons est illusoire. C'est l'association des acteurs dans une co-construction qui est une véritable force pour résoudre des problèmes complexes regroupant de nombreux acteurs<sup>152</sup>.

Les « paysans » et l'équipe « DEAP » ont une culture épistémologique différente et une façon différente d'appréhender le vivant<sup>29;62</sup>. La mise en place des dispositifs expérimentaux dans les fermes illustre ces différences. D'un côté, les chercheurs, avec leur formation biométrique, souhaitent contrôler au mieux la variation environnementale pour détecter les différences dues à la génétique. De l'autre côté, les paysans ont une vision holistique de leur système et évaluent plutôt leurs populations et leurs pratiques à travers des répétitions dans le temps et la comparaison entre les fermes. Ces postures épistémologiques représentent le principal point de divergence entre ces deux types d'acteurs. Alors que les chercheurs ont une posture réductionniste, les paysans ont une posture plutôt holistique.

Ces deux organisations d'acteurs communiquent et articulent une action commune grâce à un travail collectif articulé autour de la notion de « biodiversité », analysée par certains auteurs comme un objet frontière<sup>62</sup>. L'objet frontière est un concept qui permet à des acteurs de cultures épistémiques différentes de se coordonner, de communiquer, d'articuler leurs actions autour d'un même projet<sup>220</sup>. Ici, le terme « biodiversité » recoupe les termes « semences », « variétés », « populations ». La « biodiversité » a permis de



rassembler des passionnés de la diversité des plantes et une équipe de recherche travaillant sur la gestion de la biodiversité cultivée. Cet objet frontière est changeant et évolue au cours des interactions entre paysans et chercheurs<sup>62</sup>.

Cette collaboration a pu voir le jour car, des deux côtés, des valeurs éthiques et morales étaient partagées<sup>29</sup>. Aucune charte n'a encore été mise en place dans le projet, tout le travail est basé sur la confiance entre les acteurs.

## 4.2 Co-construction de la méthodologie

Ce projet est un projet de recherche-action : tout en effectuant le programme de sélection participative pour répondre aux questions initiales, la sélection, la méthodologie et l'organisation peuvent être modifiées à chaque étape à la suite de décisions collectives.

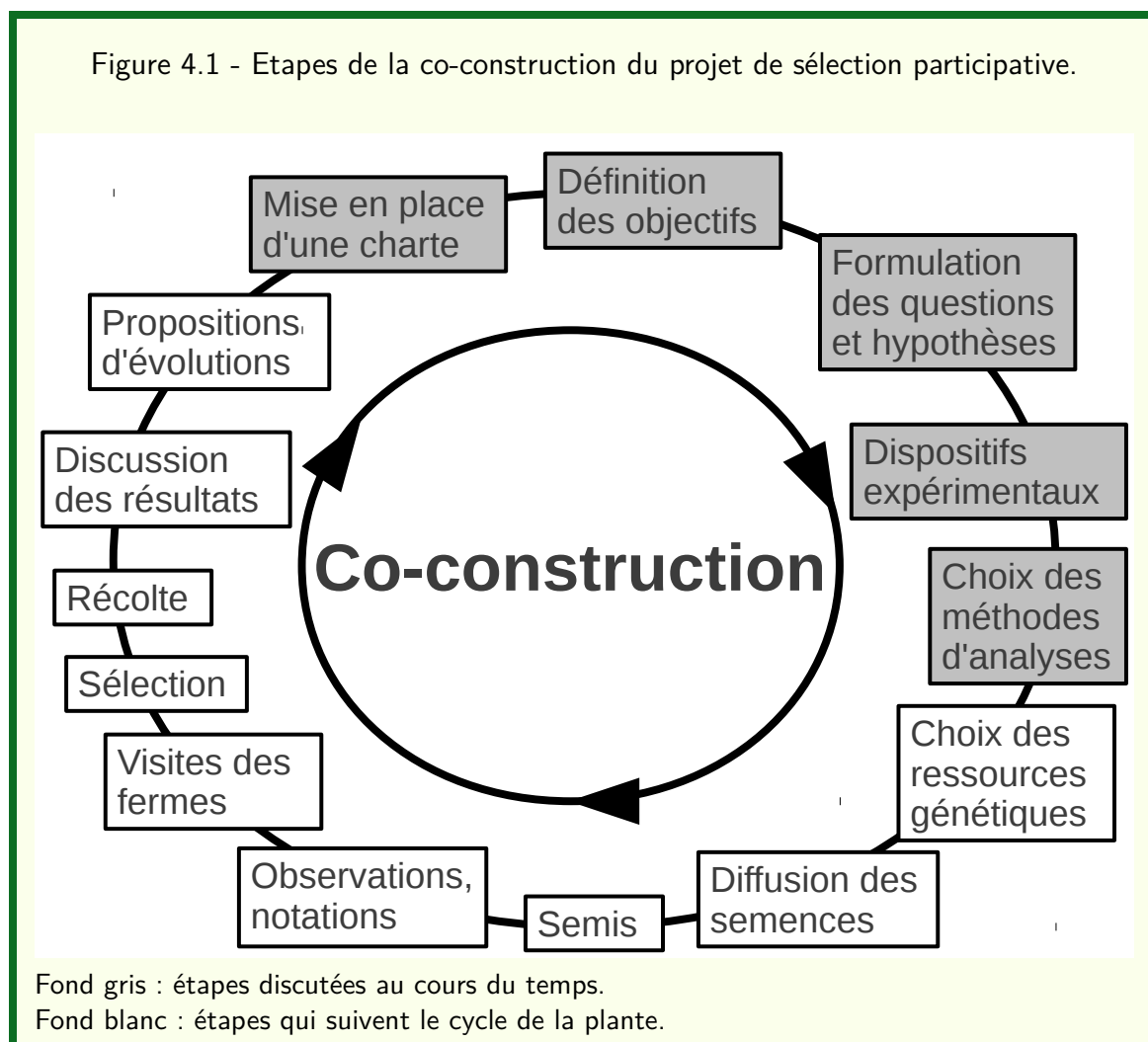
Dans notre projet, on dénombre 14 étapes qui suivent pour une partie le cycle de la plante (Figure 4.1). Chacune de ces étapes est le fruit d'interactions entre les acteurs du projet et se traduit par des décisions prises collectivement. Cette collaboration a évolué au cours du temps.

Dans cette partie, on lira la description des différentes étapes de la construction collective. Acteur dans ce projet et non observateur, j'ai été, pendant trois ans, un animateur qui a fait le lien entre les différents acteurs tout en appartenant à l'équipe « DEAP ». Cette position privilégiée m'a permis de mieux cerner le projet. Je me suis appuyé sur le travail de Bérandère Storup<sup>224</sup> et Guillemette Fruchaud<sup>108</sup>. B. Storup a travaillé dans le cadre du projet REPERE<sup>ii</sup> sur le suivi et le contexte du programme de sélection participative. En plus du rapport écrit à cette occasion, elle m'a fourni la retranscription des entretiens semi-directifs. Ces entretiens ont été réalisés avec des acteurs de l'équipe « DEAP » et des « paysans ». G. Fruchaud, qui a fait son stage de master 2 sur le projet au sein de l'équipe « DEAP », a également réalisé des entretiens avec les acteurs du projet. Son travail s'inscrit dans la continuité de celui de B. Storup, qui a co-encadré G. Fruchaud. J'ai utilisé ses entretiens semi-directifs retranscrits. Ces entretiens ont été réalisés avec des acteurs de l'équipe « DEAP », du groupe « blé RSP » et des « paysans ». J'ai aussi utilisé les comptes-rendus de cinq réunions avec les acteurs du projet qui ont eu lieu le 9 et 10 février 2010, le 7 janvier 2011, le 16 septembre 2011, le 11 septembre 2012 et le 18 septembre 2013, ainsi que les notes que j'ai pu prendre lors de réunions ou de visites chez

---

ii. Réseau d'échange et de projets sur le pilotage de la recherche et l'expertise, financé par le Ministère de l'écologie du développement durable et de l'énergie

les paysans, ma mémoire des échanges avec eux et les associations.



Tout d'abord, une première partie décrit la naissance du projet. Ensuite, une deuxième partie décrit la construction à chaque étape du programme de sélection participative. Enfin, une dernière partie analyse la co-construction.

#### 4.2.1 Naissance du projet

Pour que tout projet se développe, il faut qu'une graine germe. La graine a germé en avril 2006, quand une douzaine de membres du groupe blé du RSP sont allés en Syrie voir les essais de sélection participative avec Salvatore Ceccarelli : « *Dès que je suis rentré de ce voyage, je me suis dit que puisque d'autres l'avaient fait de l'autre côté du monde, on allait essayer de le faire et on verrait comment ça serait accueilli et repris* »<sup>iii</sup>.

A la suite de ce voyage, Jean-François Berthelot (*JFB*), paysan-boulangier près d'Agen, a décidé de faire des croisements entre des populations de blé. Son objectif était d'associer

iii. *JFB*, lors d'un entretien réalisé par B. Storup

des qualités agronomiques et nutritionnelles de différentes variétés de pays, variétés anciennes et variétés modernes créées pour l'agriculture biologique. Le croisement permet de recombinaison des caractères et d'accélérer les processus de création de diversité. Ceux-ci se mettraient en place beaucoup plus lentement dans un champ, sans intervention humaine. Les croisements permettent de trouver plus rapidement des génotypes intéressants. *JFB* a été accompagné par l'équipe « DEAP » dans la création des nouvelles populations de blé en 2005-2006. *JFB* a choisi les parents pour les croisements. L'équipe de recherche n'est pas intervenue sur cet aspect. Elle a accompagné *JFB* techniquement dans la réalisation des croisements dans sa ferme qui ont eu lieu en 2006.

Ce travail a intéressé certains paysans qui ont vu une expérience à tenter, un moyen pour créer de la variabilité et brasser les caractères pour se donner plus de chances de développer des variétés paysannes de blé tendre adaptées aux modes de cultures agroécologiques et aptes à offrir un pain de qualité. En 2008, après deux années de multiplications dans la ferme de *JFB*, 22 paysans sont volontaires pour tester ces nouvelles variétés. C'est le RSP qui a proposé à certains paysans de rentrer dans la dynamique du projet, l'équipe de recherche n'est pas intervenue dans cette étape.

Ces populations issues de croisements, réparties dans le réseau de fermes, ont été le point de départ d'un programme de sélection participative. On passe d'une collaboration entre un paysan et l'équipe de recherche à une collaboration entre paysans, associations et équipe de recherche.

#### 4.2.2 Description de l'évolution de la co-construction à chaque étape

Le type de participation dans un programme de sélection peut être décrit selon trois points de vues<sup>218</sup> : les étapes de la participation, le degré de participation et le rôle des acteurs dans la participation. Cette partie reprend chacune des 14 étapes présentées dans la figure 4.1. Le degré de participation et le rôle des acteurs y sont décrits chronologiquement.

##### Mise en place d'une charte

Au début du projet, les acteurs se connaissent déjà depuis plusieurs années et ils se font confiance. Ainsi, aucune charte n'a été mise en place. Même si ces discussions sur le sujet sont régulières dans le groupe et que la formalisation d'une charte est complexe, au bout du compte, c'est un contrat moral qui est engagé entre les acteurs. Ces discussions ont commencé dès la première réunion de 2010 de manière informelle, sans que cela soit



à l'ordre du jour.

Les débats de l'AG<sup>iv</sup> du RSP à Colmar en 2011 ont été à la base des discussions de la réunion de 2012 sur la définition du terme « variété paysanne ». Cette définition constitue le principal enjeu dans la construction du groupe RSP<sup>63</sup>. La sélection participative a pour objectif la création de variétés paysannes et se trouve impliquée aussi dans cette réflexion. Lors de la réunion de 2012, la mise en place d'une charte est à l'ordre du jour. En effet, le projet commence à prendre de l'ampleur, de nouveaux paysans s'insèrent dans le processus sans forcément avoir pris le temps de discuter avec les acteurs du projet. Les nouveaux adhérents qui participent au programme sont cooptés par les associations ou les paysans. Certains paysans ne se sont jamais rencontrés. Il est alors nécessaire de mieux cadrer le projet afin que des règles claires soient mises en place pour accueillir de nouveaux participants.

C'est le RSP et l'équipe de recherche qui ont proposé de mettre la charte à l'ordre du jour de cette réunion. Tous les paysans et les associations y étaient favorables. Tous les acteurs du projet ont pu donner leur avis sur la question. Aucune divergence n'est apparue. Parmi les premières idées qui ont émergé, on retiendra : maintenir l'agriculture paysanne, créer un système de cooptation, échanger de petites quantités de semences lors des échanges, décrire et identifier la démarche en précisant le contexte historique, social en plus des aspects liés à l'agronomie, la nutrition, la biodiversité. Cette charte n'est pas encore formalisée.

### Définition des objectifs

Les objectifs ont été fixés par les paysans et le RSP dès le début et n'ont pas évolué depuis. L'équipe de recherche s'est jointe à ces acteurs pour les accompagner dans la réalisation de ces objectifs. Il s'agit de

1. créer des variétés-populations adaptées aux conditions locales, aux pratiques des paysans,
2. développer des méthodes et des outils opérationnels pour la gestion de la biodiversité cultivée et la sélection à la ferme,
3. renforcer l'apprentissage et l'autonomie des paysans en matière de gestion et de sélection des semences.

Au delà de ces objectifs, d'autres, intermédiaires, sont apparus au cours du projet comme des étapes permettant d'atteindre ces buts.

---

iv. Assemblée Générale

Pour atteindre l'objectif 2., cinq objectifs intermédiaires sont apparus. Il sont nommés ici et sont développés dans les étapes suivantes qui leur correspondent.

L'objectif 2.1. est apparu naturellement pour tous les acteurs dès le début de projet : élaborer des fiches d'observation et de notation des parcelles.

L'objectif 2.2. a été proposé en 2009 avec la thèse de Mathieu Thomas<sup>229</sup> : développer un système de gestion de l'information pour organiser et optimiser toutes les données disponibles dans le programme. Cet objectif vient de la convergence des besoins de l'équipe « DEAP » puis du groupe « blé RSP ». Les associations, lors de rencontres informelles ou formelles, soulevaient souvent des problèmes liés à la gestion des données, à la perte de temps associée à une mauvaise organisation des données. L'équipe de recherche, devant le grand nombre des informations liées aux échanges de semences et aux observations sur les parcelles de sélection, a jugé indispensable de développer une base de données pour stocker, analyser et valoriser les données.

L'objectif 2.3. est coopté en 2010 : créer un « dossier retour » regroupant les résultats. Cette demande est venue des paysans qui souhaitaient avoir un retour sur les mesures réalisées sur leur ferme. C'est l'équipe de recherche qui a développé cet outil.

L'objectif 2.4. est apparu dès le début du projet : valoriser statistiquement les mesures au niveau de chaque ferme. En 2010, grâce à une collaboration avec O. David, statisticien à l'INRA de Jouy en Josas, de nouvelles méthodes statistiques ont été développées. Elles prennent en compte les données mesurées sur l'ensemble des fermes malgré le déséquilibre rencontré au sein des essais expérimentaux dans les fermes. Cela donne aux paysans des résultats statistiques. C'est l'équipe de recherche qui a été à l'initiative de cet objectif.

L'objectif 2.5. est apparu en 2011 : évaluer l'impact du programme de sélection participative sur l'évolution de la variabilité génétique et phénotypique. Cela permet également de donner des pistes pour une bonne gestion de ces ressources au sein du programme de sélection. Ce sont les chercheurs qui ont été à l'initiative de cette expérimentation. Cette dernière contribue à atteindre les objectifs de départ. L'équipe de recherche a mis en place cette expérience avec les paysans. Par exemple, les paysans ont souhaité que l'évaluation phénotypique des populations se fasse dans leurs fermes. Cependant, l'équipe de recherche souhaitait semer toutes les populations testées dans la station de recherche, afin de faciliter le prélèvement des feuilles pour les analyses moléculaires et le suivi de l'épiaison. Il a été convenu que toutes les populations seraient évaluées dans les fermes et semées également au Moulon, afin de faciliter les prélèvements de feuilles et les notations d'épiaisons.

Pour atteindre l'objectif 3., deux objectifs intermédiaires ont été demandés par les paysans ou par les associations qui relayaient également des demandes d'autres paysans extérieurs au projet.

L'objectif 3.1. demandé en 2010 : former les paysans à la technique des croisements.

L'objectif 3.2. demandé en 2011 : les former à des notions de génétique et de sélection.

Atteindre les objectifs 2. et 3. sont nécessaires mais pas suffisants pour atteindre l'objectif 1.

### Formulation des questions et des hypothèses

Cette partie de la méthodologie est importante car elle permet de cadrer le travail de recherche.

Pour chaque objectif, une question est posée et des hypothèses à vérifier sont posées. Les chercheurs sont à l'initiative de cette étape. Malgré tout, chaque acteur a été moteur dans la conceptualisation des questions et des hypothèses.

On peut distinguer deux situations : le projet de sélection dans les fermes (chapitre 2) et l'évaluation du programme (chapitre 3).

Pour le projet de sélection participative en « routine » dans les fermes, l'étape de formulation des questions et des hypothèses n'a jamais été clairement énoncée comme telle. Elle n'a pas fait l'objet d'un point dans l'ordre du jour d'aucune réunion. Néanmoins, l'équipe de recherche, en discutant au cours des réunions ou lors des visites dans les fermes ou de rencontres informelles, a mis en avant les questions et les hypothèses qui reflètent les besoins et les attentes des acteurs. Pour chaque objectif de la partie précédente, des questions et des hypothèses peuvent être proposées. Les hypothèses peuvent être des hypothèses à tester, des hypothèses de travail ou des propositions. Si l'on reprend les objectifs de la partie précédente, les questions ( $Q$ ) et les hypothèses ( $H$ ) sont :

objectif 2.1. -  $Q$  : Quelles fiches mettre en place qui permettent un suivi des essais et qui correspondent aux besoins des paysans?  $H$  : Les paysans connaissent leurs besoins et choisissent les critères. Cela s'est construit au cours des réunions et des échanges.

objectif 2.2. -  $Q$  : Quel outil pour gérer les données issues du programme de sélection participative?  $H$  : La spécificité des données générées dans le cadre du projet nécessite un développement particulier.

objectif 2.3. - *Q* : Comment redonner aux « paysans », les résultats générés par les analyses ? *H* : Les paysans doivent avoir accès à des données personnalisées qui correspondent à leurs besoins.

objectif 2.4. - *Q* : Comment évoluent les populations des paysans dans leurs fermes, par rapport à la sélection de l'environnement, la sélection massale ? Quelles différences y a-t-il entre deux populations sur une ferme ? Quels germplasms sont les plus hauts, versent le moins, etc ? Quelles fermes se ressemblent le plus ? Quel est le potentiel génétique des germplasms ? *H* : Les hypothèses correspondantes sont les hypothèses statistiques décrites dans le chapitre 2.

objectif 3.1. - *Q* : Comment les paysans s'approprient l'étape de création de variabilité ? Est-ce qu'ils peuvent créer leurs croisements à la ferme ? *H* : Les paysans peuvent choisir des parents qui sont croisés par l'équipe de recherche à la station de recherche. Le processus de croisement est transposable dans les fermes mais nécessite une formation technique.

objectif 3.2. - *Q* : Quels concepts peuvent aider pour faire la sélection au champ ? *H* : Les concepts de base de génétique mendélienne, génétique des populations et génétique quantitative, permettent aux paysans de mieux appréhender la sélection dans leur champ.

Pour l'évaluation du programme de sélection participative (objectif 2.5. et chapitre 3) , les questions et les hypothèses ont été largement discutées par le groupe. *Q* : Quelle contribution du programme de sélection participative à la gestion de la diversité génétique ? Quelle diversité initiale pour créer les nouvelles variétés paysannes ? *H* : La diversité des lieux et des pratiques liées à la multiplication et à la sélection permet la création et le maintien de la biodiversité cultivée. L'origine des parents choisis conditionne la qualité des variétés paysannes obtenues.

### Dispositifs expérimentaux

Pour répondre aux questions posées, des dispositifs expérimentaux doivent être mis en place. Dans ce travail, on peut différencier trois dispositifs. Le premier a été mis en place par les chercheurs à la station de recherche pour étudier la réponse à la sélection de *JFB* en génération précoce (chapitre 2). Le deuxième est celui mis en place dans les fermes pour évaluer chaque année les populations sélectionnées par les paysans (chapitre 2). Le troisième est le dispositif mis en place pour l'expérience d'évaluation (chapitre 3).

Dans un essai en réseau de stations expérimentales, l'idéal est d'avoir les mêmes popu-

lations répétées dans toutes les stations. Pour les chercheurs, s'il n'y a aucune répétition, il est impossible de générer des résultats assortis de tests statistiques. Cela a deux impacts négatifs pour les chercheurs : ils ne pourront pas donner de résultats fiables aux paysans et il leur sera impossible de publier une analyse scientifique du travail. Lors des premières réunions, les chercheurs ont insisté sur l'intérêt pour les paysans d'avoir accès à des résultats scientifiques qui pourront les aider à atteindre leurs objectifs. De même, au delà du fait que les publications sont obligatoires pour les chercheurs<sup>v</sup>, les publications permettront de faire connaître le programme et ses résultats au delà des organisations sociales « paysans », « blé RSP » et « DEAP ».

Les paysans ont été d'accord pour mettre en place des répétitions. Cela dit, ils ne souhaitent pas reproduire les pratiques de la station de recherche dans leurs champs pour quatre raisons. La première est le manque de place et de temps pour gérer les essais complètement répétés. La deuxième est l'envie d'optimiser l'espace disponible pour tester le plus de populations possible. La troisième est le besoin de flexibilité : les paysans ne veulent pas rentrer dans un cadre trop strict qui les oblige à semer des populations qui ne les intéressent pas. Pour finir, les répétitions et le contrôle des conditions expérimentales sont le symbole de l'agriculture industrielle à travers l'avènement de la biométrie (partie 1.3). Même si tous les paysans ne se rejoignent pas sur ce dernier point, il a joué un rôle dans la mise en place des dispositifs.

Les chercheurs ont réfléchi à des dispositifs intermédiaires entre le tout répétition et le non répétition. Ces dispositifs doivent permettre à la fois de répondre aux exigences des paysans et à celles des chercheurs, chacun étant conscient des contraintes des autres acteurs.

L'objectif des dispositifs est d'estimer une variance environnementale dans les fermes et de pouvoir comparer les populations sur le réseau (plus de détails dans le chapitre 2). Il a été imaginé au départ de répéter un témoin deux fois dans toutes les fermes qui seront appelées par la suite fermes satellites. Le témoin a été choisi par les paysans : *Rouge-du-Roc* pour son comportement dans de nombreuses fermes. Ensuite, dès 2010, une autre étape a été franchie avec la mise en place des fermes régionales qui accueilleront deux blocs avec dans chaque bloc quatre témoins. *Rouge-du-Roc* est conservé. Les trois autres témoins ont été proposés par l'équipe de recherche : *C21* et *C14* car ils étaient déjà très fréquents sur le réseau et *Renan* pour avoir un point de comparaison avec une

---

v. D'autant plus dans le cadre d'une thèse où le doctorant doit légitimer son travail par des publications, gages de recherche sérieuse.

variété connue en agriculture biologique. Les paysans ont été d'accord avec le choix de ces témoins. Les paysans choisissent d'être ferme régionale ou satellite. Dans les faits, on observe que les paysans les plus expérimentés ont pris le statut de fermes régionales.

Concernant l'expérience d'évaluation (chapitre 3), ce sont les paysans qui ont demandé à ce que les essais se fassent dans leurs fermes. Celles-ci sont représentatives « *des vraies conditions de la ferme* » au contraire de la station du Moulon. Un dispositif, inspiré de celui mis en place dans les fermes régionales, a été mis en place dans les trois fermes des paysans volontaires.

### Choix des méthodes d'analyse

Que ce soit pour les analyses du programme de sélection participative en « routine » (chapitre 2) ou les analyses du programme d'évaluation (chapitre 3), ce sont les chercheurs qui ont choisi les méthodes d'analyse. Cela a été fait sans préjuger si les paysans pouvaient apporter un point de vue sur ces aspects.

Le chapitre 2 présente les outils statistiques qui ont été développés dans le programme. Les statistiques permettent de prendre en compte le déséquilibre des essais grâce au réseau des fermes.

Une visite du laboratoire a eu lieu en 2009 au Moulon. Les paysans présents ont pu visiter le laboratoire. Ils ont assisté à une extraction d'ADN, à l'observation de l'ADN par des gels d'agarose, et à divers autres techniques de laboratoires. Les paysans comprennent l'intérêt des outils d'analyse de la diversité génétique mais sont mal à l'aise à cause de la toxicité de certains réactifs utilisés.

### Choix des ressources génétiques

Une fois le protocole en place, il faut choisir quelles ressources génétiques semer. Trois choix s'offrent au paysan.

1. Récupérer des variétés déjà présentes au sein du RSP. C'est le principal choix des paysans. Dans ce cas, le RSP est totalement autonome sur ses ressources génétiques.
2. Récupérer des variétés présentes dans des banques de graines. Ce choix a été le seul possible (à quelques rares variétés près) au départ pour les paysans souhaitant recultiver des variétés de pays ou des variétés anciennes. Aujourd'hui, ce choix est plus rare, même s'il reste possible. Dans ce cas, les semences viennent essentiellement du Centre de Ressources Biologiques (CRB) de l'« INRA » de Clermont-Ferrand.
3. Créer de nouvelles populations par croisement ou par mélange. Alors que les mé-

langes sont courants au sein du groupe « blé RSP », les croisements sont apparus plus tard, portés par la dynamique du projet de sélection participative (partie 4.4.1). Ce sont les paysans qui décident quelles ressources semer dans leurs essais. L'équipe de recherche n'intervient pas dans ce choix mais elle joue le rôle de facilitateur.

Tout d'abord, elle peut aider les associations à récupérer certaines ressources génétiques mais cela reste très rare, les associations sont en général autonomes sur ce point.

Ensuite, elle accompagne les paysans dans leurs croisements, comme au départ du projet avec les croisements créés par *JFB*. Dès 2009, les paysans ont demandé à suivre une formation sur les croisements. Celle-ci a été réalisée en mars 2010 au Moulon. Elle a été suivie par d'autres formations à Graines de Noé en 2011 et au Centre d'Etude et Terre d'Accueil des Blés (CETAB) et à Kernà ùn Shoma en 2013. Des fiches ont été créées pour l'occasion par l'équipe de recherche (Annexe Q).

L'opération de croisement est délicate et dépendante des conditions climatiques. L'équipe de recherche a donc proposé aux paysans qui le souhaitent de réaliser les croisements à la station de recherche dans un tunnel bâché. Cependant, ce point n'avait pas été discuté avec le groupe, et certains paysans ont regretté que cette étape clé pour la création autonome de leurs semences ne se fasse pas uniquement à la ferme et par les paysans eux-mêmes. Ce point a été mis à l'ordre du jour et discuté lors de la réunion des semis de 2012. Tous les acteurs ont été d'accord pour laisser le choix à chaque paysan de réaliser leurs croisements eux-mêmes sur leurs fermes ou par l'intermédiaire de l'équipe de recherche.

### **Diffusion des semences**

Lorsque des variétés sont récupérées par l'intermédiaire du CRB, il faut remplir un Accord de Transfert de Matériel (ATM) et les graines sont envoyées par la poste. Lorsque les variétés sont issues du RSP, il y a trois intermédiaires possibles : les paysans, les associations et l'équipe de recherche. Le rôle que doit avoir tel ou tel acteur dans la diffusion des semences n'a pas été discuté collectivement.

### **Semis**

A partir de 2010, à son initiative, l'équipe de recherche a fourni les plans des semis aux paysans. Chaque paysan propose les populations qu'il souhaite semer et les plans sont créés. Les discussions au sujet des plans de semis, entre les paysans et l'équipe de recherche, sont fréquentes et les plans sont parfois modifiés. Par exemple, certains plans ne sont pas adaptés à la parcelle, des semences manquent, de nouvelles semences arrivent dans les fermes où se font les expérimentations.

Les semis sont gérés par les paysans. Ce sont eux qui préparent la terre, préparent les lots de semences et sèment. Pour certaines fermes régionales, les associations aident les paysans. Pour l'expérience d'évaluation (chapitre 3), c'est l'équipe de recherche qui a préparé les lots semences prêts à être semés.

### Observations et notations

Les observations et les notations dans les essais sont de deux types : les notations par l'intermédiaire de fiches de suivi au cours du cycle et les notations sur les épis récoltés (partie 2.1.3).

Toutes les fiches de suivi ont été construites selon les demandes des paysans. L'équipe de recherche a participé aux discussions concernant les fiches. Lors de la réunion de février 2010, le groupe s'est mis d'accord sur trois fiches : sortie d'hiver (Annexe F), printemps (Annexe G) et maturité (Annexe H). Ces fiches ont été distribuées pour évaluer le cycle 2009-2010. Pour les semis de 2010, une fiche semis (Annexe E) a été mise en place à l'initiative de l'équipe de recherche afin de récupérer les plans de semis (indispensable pour faire ensuite les analyses statistiques) mais également des informations pour caractériser les environnements des cultures. En plus, une fiche « climat et accidents de cultures » est ajoutée à chaque fiche semis, hiver, printemps et été (Annexe D).

Au delà des critères agronomiques « classiques », comme le port au tallage, la date d'épiaison ou la hauteur, les paysans ont insisté, lors de la réunion de 2010, pour donner une note globale sur les populations ainsi que la possibilité de laisser une large place pour leurs commentaires. En effet, les critères définis ensemble ne permettent pas d'envisager la population dans toute sa globalité.

Ces fiches ont évolué au cours du temps. On peut repérer quatre changements importants mis en place en 2010 à l'initiative de l'équipe de recherche.

Certains paysans ne comprenaient pas comment remplir tel ou tel critère sur la fiche. Des notices explicatives ont donc été créées pour chaque fiche. Elles ont été améliorées d'année en année au fil des échanges avec les paysans. Une notice sur la reconnaissance des maladies en 2010 et une autre sur la reconnaissance des ravageurs du blé en 2013 ont ainsi été créées. Ces points sont largement appréciés par les paysans : « *[J'ai] l'impression que les feuilles d'observations deviennent à la fois plus simples et plus précises ; avec l'apparition des notices* »<sup>vi</sup>.

---

vi. RAB, paysan en Rhône-Alpes, lors d'un entretien réalisé par G. Fruchaud



Le deuxième changement a consisté à ajouter sur chaque fiche le nom des populations dans l'ordre du semis afin de limiter les risques d'erreurs.

Le troisième est la mise en place de cases à cocher pour limiter le risque d'erreur sur les observations demandant des notes. Par exemple, pour la verse, certains paysans attribuent un « 5 » pour une population versée et un « 1 » pour une population droite. Un autre paysan fera l'inverse. Bien que la fiche mentionne le protocole, des confusions pouvaient se produire.

Le dernier changement est à l'initiative des paysans qui souhaitaient avoir sur chaque fiche un bilan des notations précédentes. Cela a été mis en place par l'équipe de recherche. En plus de cela, des enveloppes portant l'adresse du Moulon sont données aux paysans afin de faciliter le retour des fiches qui doivent être envoyées au fur et à mesure pour maintenir une dynamique et permettre la mise à jour des informations dans la base de données <sup>vii</sup>.

Les notations quantitatives sur les plantes et les épis récoltés sont effectuées par l'équipe de recherche. Les critères retenus sont ceux d'intérêt agronomique direct pour les paysans : la hauteur, la distance entre la feuille drapeau et l'épi, le poids de grains par épi, le poids de mille grains ou le taux de protéine. Ces critères, choisis par les paysans, sont classiquement mesurés par l'équipe de recherche. Le taux de protéine et la force boulangère (qui sera mesurée dans le cadre de l'évaluation du programme) font débat au sein du groupe aux réunions de 2011 et 2012. Alors que les paysans-boulangers n'y attachent que peu d'importance, les paysans dans des filières plus longues s'y intéressent. Au final, le groupe est d'accord pour conserver ce critère.

Les autres éléments de caractérisation sont la longueur de l'épi, le nombre d'épillets, le nombre d'épillets stériles. Ces critères, moins intéressants pour les paysans, ont été proposés par l'équipe de recherche afin de mieux caractériser la diversité et la différenciation des populations.

A l'initiative de l'équipe de recherche, des notices ont été mises en place pour expliquer aux paysans la marche à suivre pour envoyer au Moulon les épis afin qu'ils soient mesurés (Annexe I). Cela a permis de cadrer cette étape très importante dans la méthodologie : l'échantillonnage.

---

vii. On revient sur la base de données plus tard dans le texte

### Visites des fermes

Les visites des fermes participant au projet sont une étape importante dans le processus de sélection. Au delà des mesures effectuées dans les fermes, c'est un moment d'échange et de partage entre les acteurs : temps pour discuter et parler des fiches, expériences en cours, projet, vie dans le réseau, vie dans le laboratoire de recherche, etc.

En 2009 et 2010, les visites ont lieu dans toutes les fermes. Mais c'est devenu impossible au niveau de la logistique. L'organisation en fermes régionales et satellites a permis de réorganiser les visites : l'équipe de recherche va faire les mesures sur les plantes uniquement dans les fermes régionales.

En 2011, à l'initiative de l'équipe de recherche, les visites pour effectuer les mesures dans les fermes régionales sont couplées aux journées portes ouvertes habituellement organisées par les associations (Figure 4.2). Fusionner ces deux visites a eu plusieurs avantages reconnus par tous lors des réunions et lors des discussions informelles :

1. L'équipe de recherche a davantage communiqué avec les associations. Les associations ont pu mieux se rendre compte de l'organisation liée au projet et s'appropriier le suivi des populations et la sélection.
2. L'équipe de recherche a pu, lors de ces rencontres, présenter le programme de sélection participative aux différentes personnes présentes : paysans, boulangers, éleveurs, consommateurs. Un jeu sur la sélection a également été créé à l'initiative de l'équipe de recherche. Dans un espace d'une trentaine de populations, chaque participant choisit cinq populations qu'il évalue positivement, et cinq populations qu'il évalue négativement. Une discussion est ensuite organisée sur les critères de choix des différents blés par les participants. Ces rencontres ont permis d'intégrer de nouveaux adhérents au sein des associations mais aussi au sein du programme de sélection.
3. Ces visites et le fait que les associations soient impliquées dans une dynamique nationale ont permis de dynamiser les adhérents des associations.

Les protocoles liés aux visites des fermes sont décrits en annexe I.

### Sélection

En accord avec les objectifs de départ selon lesquels les paysans se (ré)approprient la sélection et deviennent autonomes, les membres du groupe « blé RSP » et de l'équipe « DEAP » ne font pas de sélection, sauf lorsqu'ils y sont invités par les paysans. Sinon, ce sont uniquement les paysans qui la font. Ce point a toujours été clair dans la mise en

place du projet.

Figure 4.2 - Visites des essais dans les fermes au moment de la récolte.



De gauche à droite et de haut en bas : visite avec Bernard Ronot en 2011 à la Bretenière, visite chez François Cazeirgues en 2011 dans sa ferme, visite chez Jean-François Berthelot en 2011 dans sa ferme, visite chez Jean-Sébastien Gascuel en 2011 dans sa ferme, visite chez Raphaël Baltassat en 2012 dans sa ferme, visite chez Romain Wittrisch en 2011 dans sa ferme.

L'équipe de recherche a balisé les différents niveaux de sélection possibles : la sélection naturelle, la sélection entre les populations et la sélection à l'intérieur des populations. Cela a formalisé le vocabulaire utilisé entre les acteurs et a permis à certains paysans de découvrir ces différents niveaux qu'ils n'avaient pas forcément envisagés. De plus, des no-

tions de génétique quantitative, de génétique des populations et de génétique mendélienne sont abordées lors des réunions. Par exemple, lorsque l'on sélectionne sur le phénotype, une partie de ce qui est observé vient de l'environnement : la sélection ne sera pas forcément efficace. Pour illustrer ce point, les répétitions dans les champs sont très parlantes. Un autre exemple est la notion de dérive génétique sur la taille des échantillons à ressemer lors de sélection : plus le nombre d'épi sera grand, moins la diversité sera affectée par des variations aléatoires. Enfin, la ségrégation des caractères après un croisement est évoquée. A la demande des paysans, pour approfondir ces points, des formations ont été organisées. En 2013, deux formations ont eu lieu avec le CETAB et Kernà un Shoma. Des formations couplées aux formations sur les croisements.

C'est la sélection à l'intérieur des populations qui a produit le plus d'échanges entre les acteurs. Lors de la réunion de 2010, l'équipe de recherche a présenté les résultats obtenues suite à l'étude de la réponse aux sélections de *JFB* dans les *F2* en 2008. Cela a déclenché l'intérêt de nombreux paysans. Il y avait une demande pour savoir comment sélectionner à l'intérieur de populations ainsi qu'une difficulté de leur part à préciser leurs « critères » de sélection. Pour répondre à cela, l'équipe de recherche a proposé de mettre en place un dispositif de réponse à la sélection chez les paysans qui le souhaitent. Sur ce point, il y a eu un manque de communication entre les acteurs sur la démarche expérimentale à suivre pour étudier les sélections. En effet, il a été proposé par écrit aux paysans (Annexe I) d'envoyer à l'équipe de recherche leurs bouquets de sélection. Tous les paysans qui ont souhaité participer ont envoyé des bouquets de 25 épis ! Il n'était donc pas clairement formulé que les bouquets pouvaient contenir autant d'épis que les paysans le souhaitent. Ce point a été discuté lors de la réunion de septembre 2011 ainsi que le protocole pour étudier la réponse à la sélection. Néanmoins, peu de paysans ont suivi le protocole. Là encore, peu de paysans avaient compris l'intérêt de semer côte à côte les semences issues des bouquets de sélection et les semences issues du vrac. Au vu des premiers résultats, lors de la réunion de 2012, de nombreux paysans ont mis en place le dispositif lors des semis suivants.

### Récolte

Les récoltes sont gérées par les paysans et leurs associations. Peu d'entre eux ont accès à de petites moissonneuses. Les récoltes se font souvent avec les membres des associations qui donnent un coup de main pour l'occasion. Avant la récolte en masse de la parcelle, il y a celle des échantillons d'épis.

L'équipe de recherche a pris l'initiative de mettre au point des protocoles pour que les récoltes se passent au mieux. En effet, avec les expériences de réponse à la sélection, il y a de nombreux sacs et c'est assez compliqué pour s'y retrouver (Annexe I). Des étiquettes, les plus parlantes possible, ont été placées sur chaque sac. A chaque visite des fermes régionales, les sacs sont distribués aux paysans présents. Sinon, les sacs sont envoyés par la poste. L'équipe de recherche explique alors, avec la notice, aux paysans comment faire les récoltes. A part la demande d'un paysan qui souhaitait des sacs plus grands, il n'y a pas eu de modifications sur ce point.

Par ailleurs, dans les fermes régionales, les plantes mesurées lors des visites sont étiquetées avec des rubans adhésifs de couleur, ce qui permet de les suivre entre le champ et la salle technique. Ce sont les paysans qui doivent retrouver les plantes marquées par le ruban et les déposer dans un sac prévu à cet effet. Au début, jaune et vert, et ensuite, à la demande de quelques paysans, le rouge, plus visible, a été utilisé.

#### **Discussion des résultats**

A chaque réunion de semis, les résultats sont discutés avec les paysans. Les réunions ont lieu au Moulon ou à Paris. Tous les acteurs sont invités à s'exprimer grâce à un tour de table pour chaque point de l'ordre du jour. L'ordre du jour est proposé par l'équipe de recherche et envoyé par mail avant la réunion à tous les participants qui ont ainsi la possibilité d'ajouter des points à discuter.

Au delà de la présentation des résultats, ces réunions, qui sont un temps important d'échanges entre les paysans et entre les associations, sont, pour certains acteurs, l'unique occasion de se rencontrer dans l'année. Elles permettent de faire le point sur le projet.

Le temps alloué à la discussion des résultats, prévu dans l'ordre du jour, s'est révélé très court. Jusqu'à aujourd'hui, toutes les réunions ont présenté des résultats liés à l'évolution de la méthodologie et quelques résultats marquants sur certaines fermes. Malheureusement, il y a très peu de temps pour discuter des résultats avec chaque paysan individuellement.

Afin de proposer des résultats synthétiques et qui reprennent l'ensemble de l'information dans le temps et dans les fermes, l'équipe de recherche a créé des outils permettant d'organiser et d'analyser la masse de résultats présent sur le réseau. Au delà des fiches de suivi et des aspects statistiques qui génèrent des résultats, deux outils ont été développés : la base de données et le « dossier retour ».

### La base de données

La base de données a été développée au départ pour répondre aux besoins de l'équipe de recherche. La nécessité de classer et d'organiser les informations a émergé avec la thèse de Mathieu Thomas dans l'équipe « DEAP »<sup>229</sup> et s'est accentuée avec le programme de sélection participative. En parallèle, certains paysans et certaines associations faisaient souvent part à l'équipe de recherche de leur besoin d'un outil pour gérer leurs données. Notamment afin de répondre à une question récurrente : « Où se trouve telle variété ? ». Les acteurs de l'équipe « DEAP », sans interactions avec les autres acteurs, ont géré la mise en place de la base de données. Cela a été possible grâce à une collaboration avec Yannick de Oliveira (équipe ABI<sup>viii</sup> du Moulon). L'équipe « DEAP » et Y. de Oliveira ont discuté ensemble de la structure de la base et de l'organisation des données. Y. de Oliveira s'est occupé de la programmation.

Néanmoins, l'équipe de recherche a gardé en tête les demandes de certaines associations afin de créer un outil qui pourrait être générique. Une présentation de la base a eu lieu à la réunion de septembre 2011. Les associations ont renouvelé leur intérêt pour l'outil.

La base de données, SHiNeMaS (Seed History and Network Management System), permet de gérer plusieurs types d'informations : les lots de semences, les relations entre les lots de semences (Figure 4.3), les personnes liées aux lots de semences et les données liées aux lots de semences ou aux relations (Figure 4.4).

### Structuration de l'information

SHiNeMas a été codée en PostgreSQL. La base de données est centrée sur le lot de semences (en beige dans la figure 4.4). Un lot de semences est un ensemble de graines défini par un germplasm présent physiquement sur une ferme une année donnée dans une parcelle. C'est cette combinaison de facteurs qui assure l'unicité des lots de semences. Par exemple, le germplasm *C77* présent chez *OLR* en 2010 se nommera :

$$\underbrace{C77}_{\text{germplasm}} - \underbrace{OLR}_{\text{lieu}} - \underbrace{2010}_{\text{année}} - \underbrace{0001}_{\text{digit}}$$

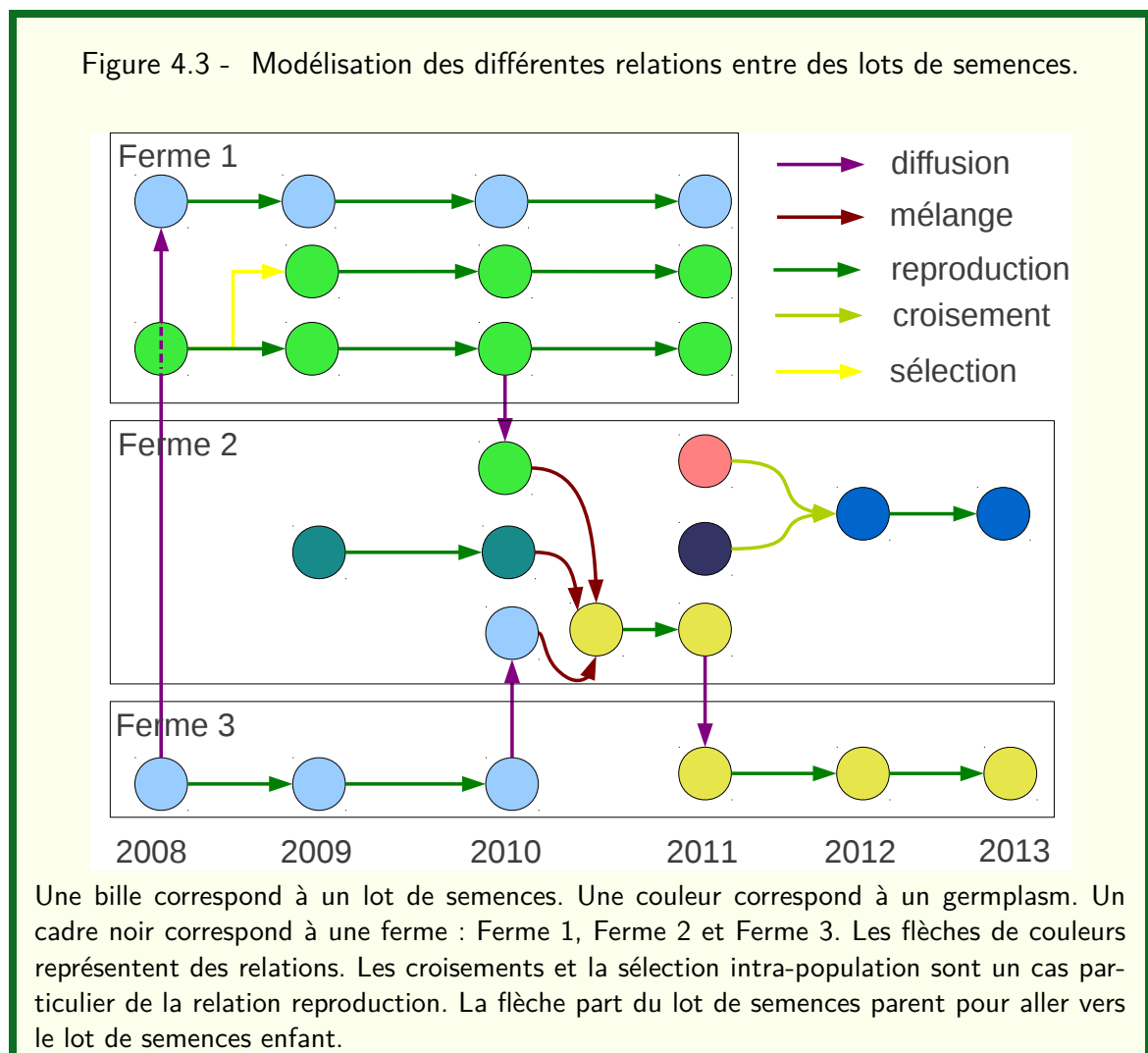
Le terme « \_0001 », que l'on nomme digit, permet de différencier des lots de graines qui auraient le même nom. Cela arrive par exemple pour les populations semées dans deux répétitions différentes. A la récolte, il y aura deux lots de semences issus de deux reproductions différentes qui seront « \_0001 » et « \_0002 ». Cette nomenclature a été mise en place sans concertation préalable et a été annoncée aux paysans lors des semis

---

viii. Atelier de Bio-Informatique

de 2011. Les paysans et les associations ont compris l'intérêt de renommer les lots de semences afin d'assurer un meilleur suivi de l'histoire de leurs populations.

Chaque relation entre lots de semences est modélisée par une relation qui relie un lot de semences parent à un lot de semences enfant. Il y a cinq relations : diffusion, mélange, reproduction, croisement et sélection intra-population (Figure 4.3 et en bleu dans la Figure 4.4). Ces relations permettent de détailler l'histoire d'un lot de semences.



La diffusion correspond à un lot de semences qui change de lieu. Cette étape change la partie *lieu* dans le nom du lot de semences.

Un mélange correspond à un mélange physique de plusieurs lots de semences. Un événement de mélange ne peut se faire que dans un même lieu, ce qui est cohérent avec la réalité, et va créer un nouveau germplasm.



La reproduction est l'étape de génération de nouvelles graines et correspond au cycle de culture : du semis à la récolte. Cette étape ajoute une génération au lot de semences et changera la partie *année* dans le nom du lot de semences.

Le croisement correspond à la création d'un nouveau germplasm, résultat d'un croisement entre deux lots de semences issus de deux germplasms différents.

La sélection correspond à une sélection massale intra-population. Le paysan choisit des épis au sein d'une population et crée un nouveau lot de semences. Les sélections issues d'un germplasm données sont notées d'un #. Par exemple, le lot de semences issus d'une sélection massale dans le germplasm *C24*, sera notée *C24#S*. Le *S* est un nom unique qui caractérise cet événement de sélection. Dans la pratique, on assigne des lettres à chaque nouvelle sélection dans un germplasm de *A* à *Z*<sup>ix</sup>. La sélection ne crée pas de nouveaux germplasms. Cependant, dans le chapitre 2, le mot « germplasm » est également utilisé pour désigner une variété créée par sélection massale. Dans l'avenir, on veillera à harmoniser le vocabulaire sur celui de la base de données.

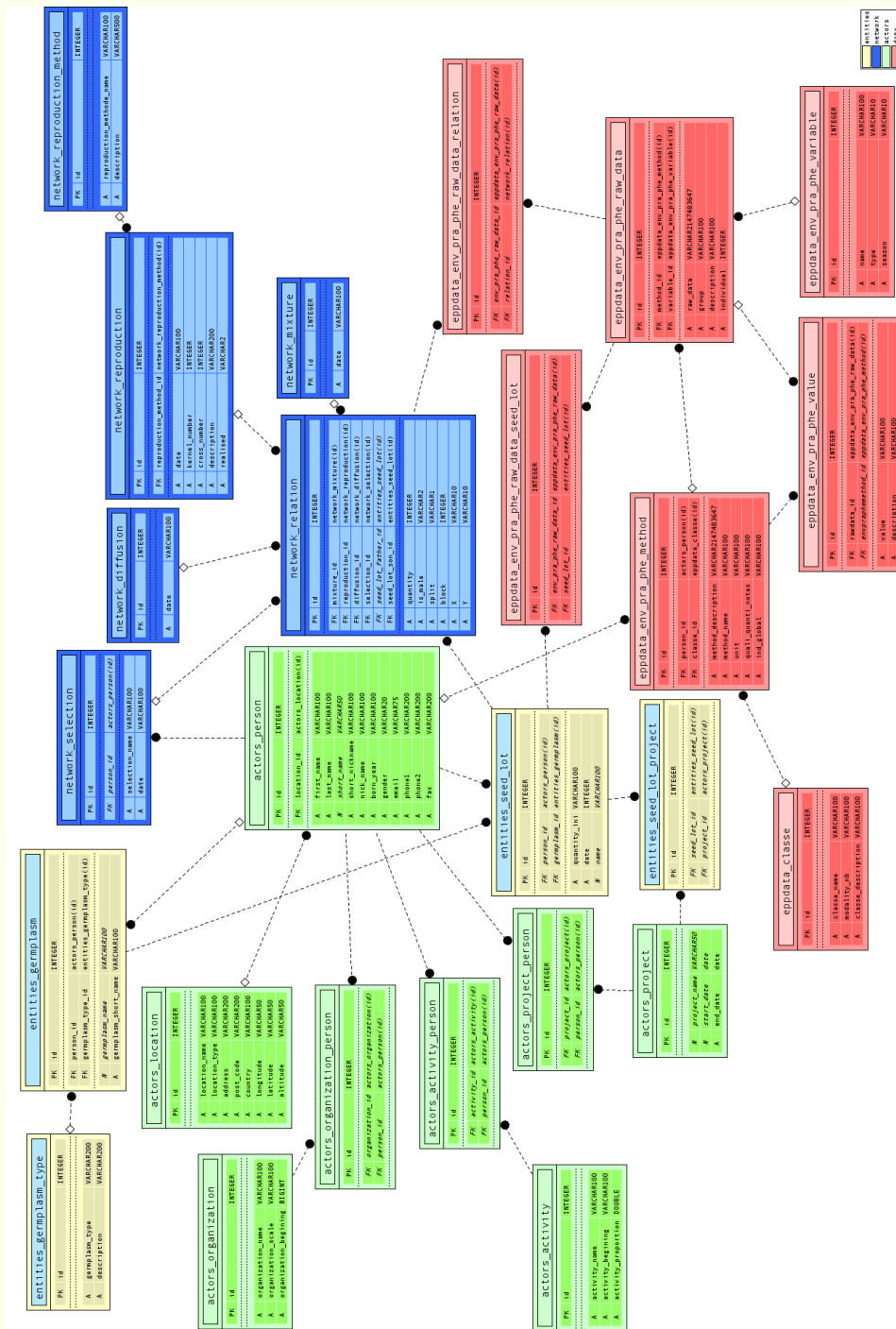
Chaque relation, lot de semences ou personne est reliée à des données (en rouge dans la Figure 4.4). Ces données peuvent être qualitatives (confiance en l'information selon laquelle une diffusion a effectivement eu lieu, statut professionnel d'une personne ... ) ou quantitatives (hauteur, poids de mille grains, données moléculaires ...). Le choix des variables est libre, ce qui permet une grande flexibilité. Chaque variable est liée à une méthode d'obtention de cette variable.

---

ix. Si il y a plus de 26 événements de sélection dans un germplasm donné, les lettres seront alors *AA*, *AB*, *AC*, etc. Bien sûr, on peut imaginer n'importe quels noms à la place des lettres.



Figure 4.4 – Relations au sein de la base de données SHiNeMaS.



Bleu : relations. Beige : personnes. Rouge : données. Vert : relations. Lot de semences. Beige : lot de semences. Rouge : données.

### L'interface graphique

Afin de faciliter l'utilisation de la base de données, une interface graphique a été développée par Y. de Oliveira. Darkawi Madi a effectué son stage de master 2 sur ce sujet. L'interface est écrite en langage `python`. La bibliothèque `Django` qui permet de requêter facilement les informations de la base a été utilisée. L'interface permet de rentrer les données dans la base et de ressortir de l'information par l'intermédiaire de requêtes.

La rentrée des données dans la base se fait par l'intermédiaire de fichiers types. Il y a six fichiers types : croisement, diffusion, mixture, reproduction, sélection et phénotype. Chacun de ces fichiers a un format standard qui permet de déclarer une relation entre deux lots de semences et d'associer des données soit aux lots de semences, soit à la relation entre deux lots de semences. Ces fichiers sont toujours couplés à un fichier méthode. Le fichier méthode contient les méthodes associées aux variables déclarées dans les fichiers types. Des messages d'erreurs sont affichés dans l'interface lorsque les données rentrées ne sont pas cohérentes (par exemple date de récolte antérieure à la date de semis, semis d'un lot de semences qui n'est pas présent dans la ferme, ...). Lorsque le fichier a été soumis, un résumé est affiché dans l'interface.

Les requêtes permettent de récupérer facilement de nombreuses informations. Par exemple, le nom des populations récoltées une année donnée dans une ferme, le nombre de sélections dans un germplasm, le nombre de générations d'une population depuis le croisement. Dès lors que les informations sont organisées dans la base, toutes les requêtes mobilisant ces informations sont possibles.

Des fonctions `R` ont été développées pour réaliser ces requêtes sur la base de données. Cela permet à la fois d'étudier le fonctionnement du réseau (nombre de multiplications, de sélections, d'échanges, etc.), et de mettre en place des fiches de suivi et des dossiers personnalisés des résultats pour chaque paysan. Les résultats sont générés par des scripts `R` et `bash` qui génèrent des scripts `LATEX` qui créent les fiches et les dossiers `.pdf`. Ces fonctions seront organisées au sein d'un package `R`. Ce package sera ensuite directement intégré dans l'interface de SHiNeMas. Actuellement, l'interface permet d'avoir accès à un nombre de requêtes limitées et ne génère pas de graphique.

### Le « dossier retour »

Les paysans ont souligné, dès le début, l'importance d'avoir des retours sur les résultats des évaluations et de leurs sélections. En 2011, l'équipe de recherche a créé un « dossier retour » inspiré par les discussions informelles avec les différents paysans. Il a été créé

indépendamment de la base de données qui était encore en construction. Présenté à la réunion de 2011, il a permis aux paysans de compléter leurs observations pour choisir les populations à semer et de baser leur sélection (inter- et intra-population) sur les résultats obtenus. Ils ont pu également choisir de nouveaux germplasms sur le réseau.

Ce dossier est structuré en deux parties (Annexe R). Il s'appuie sur les résultats obtenus par les analyses statistiques (chapitre 2).

La première partie présente les résultats obtenus dans la ferme : un bilan des notations qui ont été faites au cours du cycle et les comparaisons de moyennes sur les données quantitatives. Les résultats sur plusieurs années sont également présentés. Ces trois résultats permettent aux paysans de sélectionner entre les populations. Pour les paysans concernés, les résultats sur la réponse à la sélection et l'héritabilité réalisée sont présentés. Ces résultats leur permettent d'évaluer l'impact de leur sélection sur les critères qu'ils considèrent d'intérêt.

La deuxième partie présente les résultats sur le réseau pour que les paysans puissent identifier de nouvelles populations à tester sur leurs fermes.

Lors de la réunion de 2011, il y a eu des divergences entre les paysans sur les résultats à présenter. A la fin du tour de table, il apparaissait qu'il aurait fallu présenter toutes les mesures sous toutes les formes possibles (graphiques, tableaux, textes ...). Il a donc été décidé de créer un dossier personnalisé pour chaque paysan. Ce dossier regroupe des tableaux pour les fiches de suivis des cultures et des graphiques pour les données quantitatives suivantes : la hauteur, la verse, le poids de mille grains et le taux de protéine. En plus, un dossier exhaustif avec tous les résultats de toutes les fermes était disponible sur demande au format numérique pour les paysans qui le souhaitaient.

Les résultats concernant les expériences complémentaires au programme comme l'étude de la réponse à la sélection de *JFB* ou l'expérience d'évaluation sont également abordées avec les participants et les méthodes employées sont discutées. Les modèles statistiques sont présentés aux paysans et l'importance du réseau a notamment été mise en avant.

#### **Proposition d'évolution**

Les propositions d'évolution interviennent à deux moments : lors de la réunion des semis avec tout le groupe et lors de discussions informelles ou formelles (à travers des entretiens) entre paysans, associations et équipe de recherche. Lors des réunions, les futures orientations du projet ont été discutées et validées par consensus jusqu'à présent.

Le « dossier retour » a évolué au cours du temps. Des entretiens ont permis de l'améliorer. Certains paysans préfèrent un texte rédigé plutôt qu'un graphique. Le temps a manqué pour approfondir ce point.

Un des objectifs du programme est que les associations deviennent autonomes, il a été décidé, lors de la réunion de semis de 2012, que les associations s'approprient les outils, accompagnées par l'équipe de recherche dès la saison 2012-2013 (Annexe S).

Cette réunion-clé, a marqué le début de la décentralisation de l'organisation, jusque là gérée par l'équipe de recherche, vers les associations régionales. Tous les acteurs ont été d'accord pour effectuer cette décentralisation du processus pour plusieurs raisons. Tout d'abord, l'équipe de recherche n'a pas vocation, ni les capacités, à accompagner sur le terrain la conduite de la sélection participative au jour le jour. Ensuite, l'objectif d'autonomisation ne peut passer que par une gestion des essais par les associations régionales elles-mêmes. Enfin, décentraliser la méthodologie dans les associations permet de mieux s'adapter aux attentes des paysans.

Etant en dernière année de thèse, ma responsabilité d'organisateur et d'animateur des essais a été transférée à Sophie Pin, technicienne dans l'équipe, et aux associations. S. Pin a été chargée d'accompagner les associations dans la mise en place et le suivi des essais chez leurs adhérents. Les associations remplissaient les fichiers types que S. Pin intégrait dans la base de données afin de générer les fiches de suivi.

Cette transition a été difficile : peu de fiches ont été remplies et les sacs nous sont parvenus trop tardivement. En plus des difficultés de transfert de responsabilité, les conditions climatiques ont été très mauvaises conduisant à des retards exceptionnels à tous les stades. Il a été alors impossible de faire les mesures pour l'ensemble des échantillons et de rentrer les données dans la base pour ensuite générer le « dossier retour ». De plus, la base de données pouvant dorénavant créer des sorties standardisées, un travail de recodage des scripts R et  $\text{\LaTeX}$  était nécessaire pour créer les « dossiers retour ». Ce travail n'a pas été fait par manque de temps et les paysans n'ont pas eu de retour pour les semis de 2013.

On a pu faire le point sur les problèmes rencontrés lors de la réunion des semis de 2013. Il a été souligné la nécessité d'être très diligent lors des échanges d'informations et notamment d'envoyer les fiches sans attendre les retardataires pour ne pas pénaliser les paysans qui ont rendu leur fiche dans les délais.

L'année prochaine, 2013-2014, va poursuivre sur les mêmes bases : les associations gèrent les essais localement et S. Pin coordonne le réseau, rentre les fichiers dans la base de données, génère les fiches, crée le « dossier retour ». Cela implique une très bonne

articulation entre l'équipe de recherche et les associations. Un travail sera nécessaire pour que ces outils soient plus facilement utilisables avec la base de données.

La co-construction du projet repose sur de nombreuses étapes (Figure 4.1) le processus de sélection dans les fermes peut être divisé en six épisodes : le semis, les notations et les observations, la sélection intra-variétale, la récolte, les décisions pour le nouveau semis et les échanges de lots de semences (Figure 4.5).

Le bon déroulement de chacune de ces étapes ne peut se faire que s'il y a un animateur comme lien entre les différentes organisations sociales que sont l'équipe « DEAP » et les « paysans ». Il maintient la dynamique, relance les paysans sur les fiches et les chercheurs sur les résultats, organise les réunions, diffuse les informations aux acteurs, répond aux appels d'offre. Ce travail d'animation est la clé de réussite du projet<sup>29;224</sup>. Ce processus d'animation est primordial pour le bon fonctionnement du programme (Figure 4.6).

L'article en annexe M, publié lors du colloque DinABio 2013<sup>x</sup>, présente la méthodologie actuellement en place.

---

x. Colloque national visant à diffuser et mettre en débat les acquis récents des recherches en agriculture biologique, 13 et 14 novembre 2013 à Tours

Figure 4.5 - Les différentes étapes du processus de sélection dans les fermes

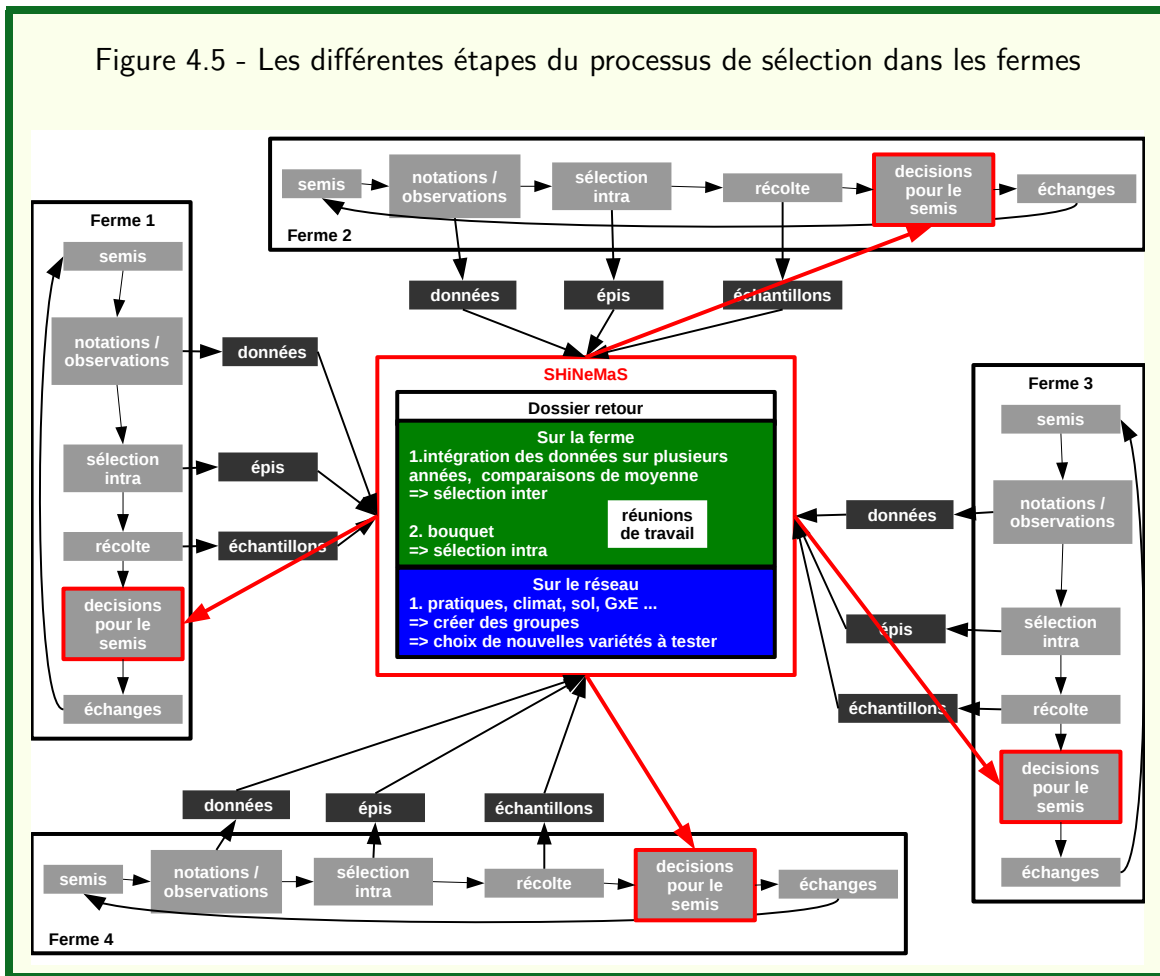
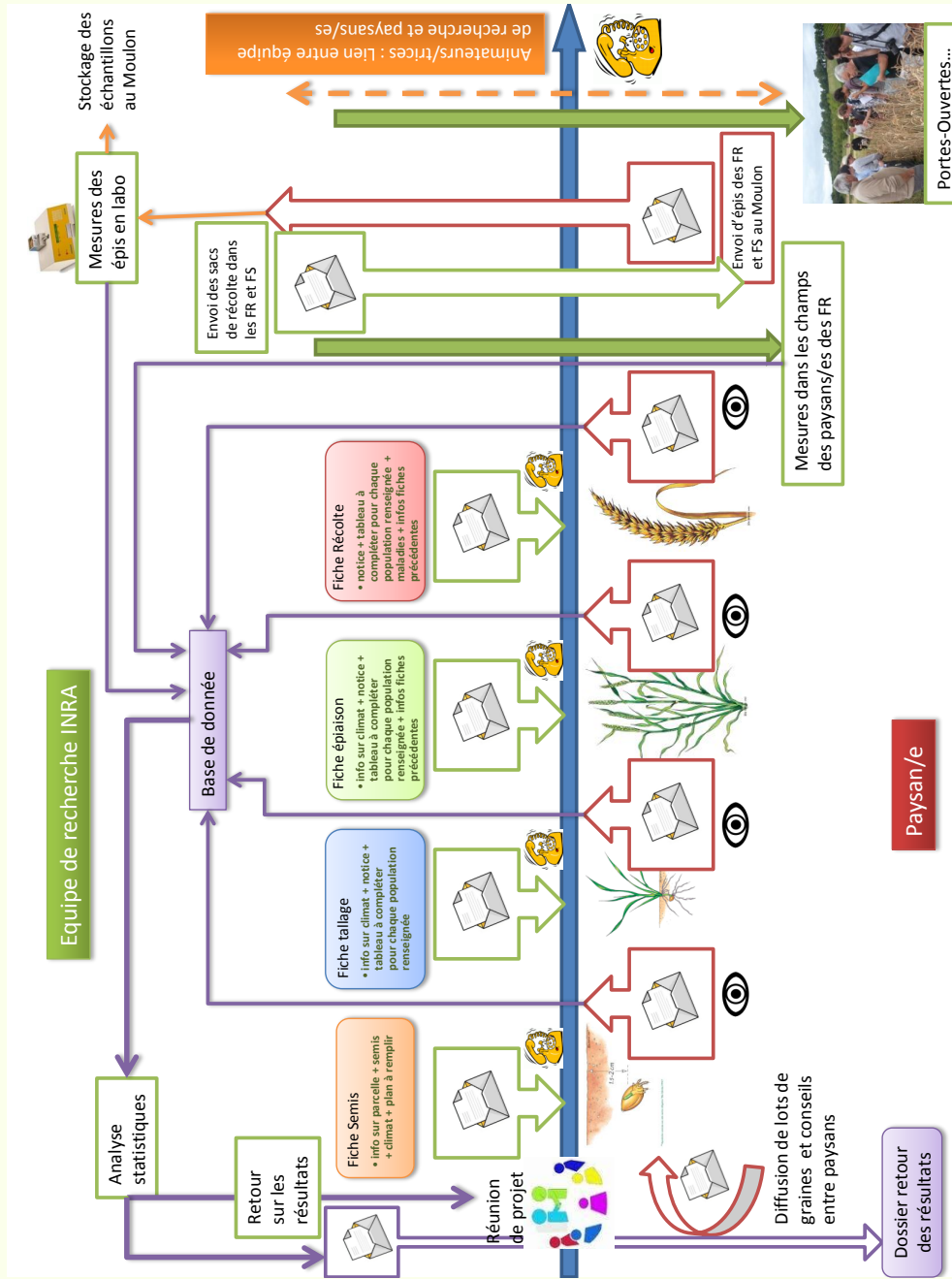
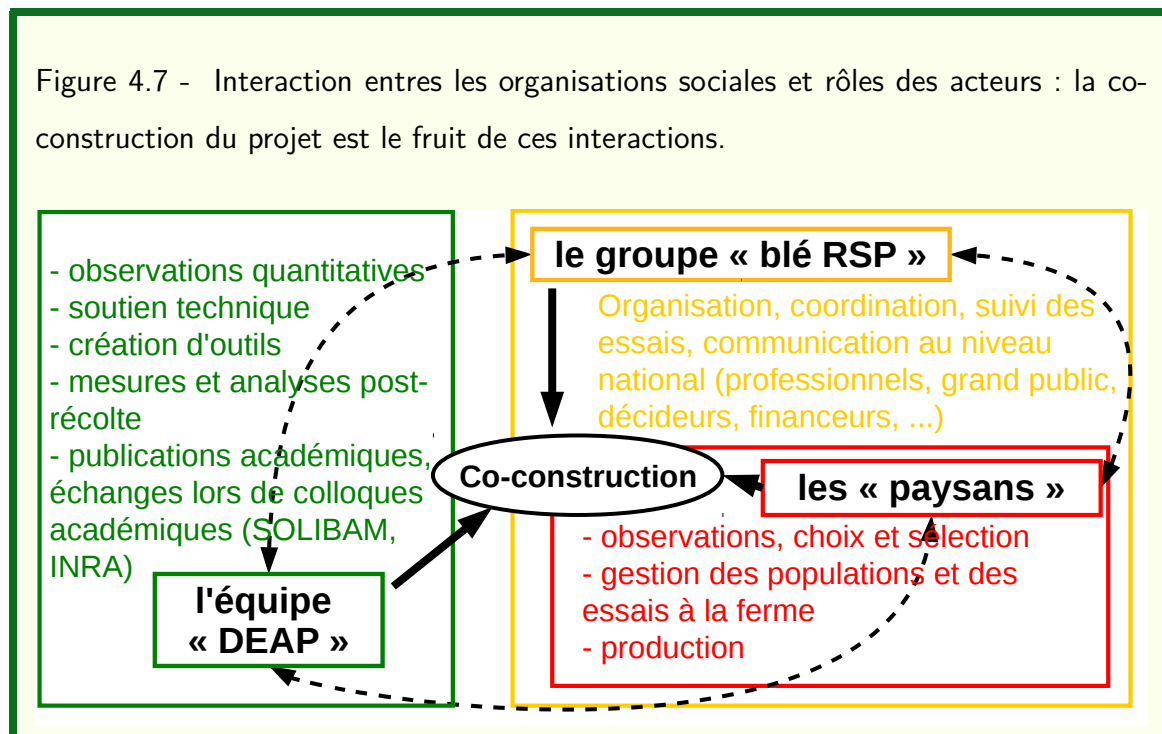


Figure 4.6 - Les étapes d'interactions entre les acteurs du projet pour animer le processus de sélection dans les fermes.



### 4.2.3 Analyse de la co-construction

La description de chaque étape nous montre qu'au delà d'une simple participation, c'est une réelle co-construction qui s'opère : chaque décision est le résultat d'échanges entre les acteurs qui contribuent à la mise en place et à l'organisation de la méthodologie<sup>224</sup>. Il n'y a pas de relations hiérarchiques entre les acteurs, cela n'empêche pas que le rôle de chacun soit identifié (Figure 4.7).



Ce travail a pris une autre dimension avec le projet européen SOLIBAM qui a permis de financer le fonctionnement de 2010 à 2014. A cela ce sont ajoutés le financement du post-doc de Julie Dawson par l'« INRA » de 2009 à 2010 et le financement de ma thèse de 2010 à 2013 par la région Ile de France (à travers le DIM ASTREA).

Toutes les actions des acteurs dans la co-construction du projet sont transparentes et chaque acteur peut proposer des évolutions. Cette approche est originale comparée à celle d'autres programmes de sélection participative où les paysans sont consultés au départ des programmes mais ne participent que rarement à toutes les étapes<sup>218</sup>.

Dans la partie qui suit, quatre thèmes sont analysés : le respect mutuel et la confiance entre les acteurs, les associations, la temporalité, la communication orale et écrite.



### Le respect mutuel et la confiance

Dès le début, cette collaboration a été basée sur la confiance et le respect entre les acteurs. Ce sont les conditions indispensables pour qu'un projet de recherche participative fonctionne<sup>169</sup>. Dans le processus de co-construction, on peut parler de décisions prises de pairs à pairs<sup>23</sup>. Cela est inédit dans l'histoire des sciences où les savoirs des chercheurs ont très souvent été estimés supérieurs à ceux des paysans<sup>2</sup>.

Certains paysans au sein du groupe ne se connaissent pas, mais pourtant des semences peuvent s'échanger entre eux<sup>108</sup>. Les paysans font confiance aux personnes qui ont coopté ces nouveaux arrivants.

### Les associations

Les premières années du programme, les associations n'étaient pas systématiquement informées. Elles suivaient le programme d'assez loin. C'est en 2011 que les premières visites dans les fermes en compagnie des associations ont eu lieu. Le degré de participation des associations est lié à l'expérience des animateurs, à la proximité géographique avec les paysans moteurs du projet, à la sensibilité des groupes à la sélection.

Par exemple, l'association ARDEAR Rhône Alpes est très dynamique : l'animateur est présent depuis longtemps et maîtrise parfaitement les enjeux liés au projet. Le CETAB, dont *JFB* est adhérent, a eu un rôle particulier dans le projet : l'association, malgré de nombreux changements d'animateur, communique souvent avec l'équipe de recherche et est très impliquée dans le suivi des essais. Enfin, Graines de Noé est représentative d'une association plus sensible à la conservation qu'à la sélection et, de fait, le programme de sélection participative peine à s'insérer dans leur dynamique, bien que l'équipe de recherche ait une relation ancienne et privilégiée avec son fondateur Bernard Ronot.

Les interactions dépendent beaucoup de l'appropriation du programme par les animateurs. Certains animateurs invitent les membres de l'équipe de recherche aux visites, à des journées sur les semences, partageront avec eux des informations sur la vie de l'association : « [...] *ce qui semble important c'est le lien entre la recherche et les paysans, et ça, il faut le garder, c'est super enrichissant* »<sup>xi</sup>.

Au contraire, d'autres animateurs travaillent de façon plus centrée sur les projets locaux et ne pensent pas à inviter l'équipe de recherche pour les visites de récolte par exemple. Avec le temps et les échanges nombreux, les associations intègrent de mieux en

---

xi. *ALH*, animateur de l'ARDEAR Rhône Alpes, lors des entretiens avec G. Fruchaud

mieux le projet dans leur dynamique. Lors de son stage de master 2, G. Fruchaud a créé un livret technique de 94 pages : Des blés en (R)évolution, qui a permis de faire un bond dans les interactions avec les associations : elles ont eu à leur disposition un document qui recensait la majorité des pratiques de recherche du programme de sélection participative. Ce genre de support a facilité la communication.

### La temporalité

La question de la temporalité est centrale dans le projet de sélection participative et ressort dans les entretiens de B. Storup et de G. Fruchaud. Ce point peut être analysé à deux niveaux.

Tout d'abord, les contraintes de temps dans une ferme ne sont pas les mêmes que celle dans un laboratoire. Le cycle des plantes et des animaux, les aspects administratifs, la répartition des outils, les réunions dans diverses associations, rythment la vie des paysans.

Le rythme des chercheurs est lié au cycle du blé, aux séminaires, aux durées des stages et des thèses, à l'écriture d'articles, et aux échéances des projets. Les paysans s'insèrent à long terme dans ce projet. Le financement de SOLIBAM dure quatre ans et demi et durant ce délai, l'équipe de recherche doit produire des rapports et des publications. Ce travail se place dans le WP6<sup>xii</sup> et doit contribuer à développer des stratégies de sélection participative. En tant que doctorant, je dois produire en trois ans un manuscrit et des publications. Les acteurs ont pris conscience de ces différences qui n'ont pas eu de réel impact sur le projet mais illustrent des visions hétérogènes au sein des acteurs. De plus, les paysans souhaitent des résultats rapidement, alors qu'analyser et publier des résultats prend du temps. Par exemple, l'analyse des sélections que *JFB* a faites en 2008 ont été analysées en 2009 et 2010 mais n'ont été publiées qu'en 2013!

Le temps de discussion entre les acteurs permet de mieux s'approprier le projet, de mieux comprendre les envies, les enjeux, les contraintes de chacun. Du temps est nécessaire pour établir un lien de confiance<sup>224</sup> et créer une organisation sociale du groupe<sup>29</sup>.

A force d'échanges et de discussions, certains paysans sont passés de la formule « *des croisements de Jean-François* » à celle « *mes populations* »<sup>xiii</sup>.

Cette question du temps revient dans chacune des réunions : « *On a une meilleure*

xii. WP pour Work Package, c'est un axe de travail du programme Européen qui compte dix WP.

xiii. J'ai observé cela lors de discussions informelles avec les paysans. On retrouve également de tels discours dans les entretiens réalisés par B. Storup et G. Fruchaud

*intuition au bout de trois ans d'observation, l'appréciation devient plus facile* »<sup>xiv</sup> ; « *Il faudrait plus de temps pour faire des tours de champs ensemble* »<sup>xv</sup>.

On s'aperçoit que les acteurs les plus impliqués et les plus motivés sont en général ceux qui ont le plus d'expérience ou qui ont passé le plus de temps à discuter avec l'équipe de recherche et les paysans moteurs.

Ce manque de temps a parfois entraîné un manque de communication qui a pu jouer sur les différences d'implications entre les acteurs : « *On n'a pas été assez explicite au début, il aurait fallu plus de temps pour expliquer. On n'était pas au même stade de réflexion non plus* »<sup>xvi</sup>. Le fait que certains paysans ne puissent pas se libérer pour assister aux réunions de semis amplifie ces différences.

### Oral et écrit

Deux modes de communication existent dans notre projet : oral et écrit.

Toutes les étapes importantes de la méthodologie (les fiches, les notices, le « dossier retour ») se font par courrier ou par mail. L'oral vient compléter certains points lors des réunions ou des rencontres informelles ainsi qu'au cours des nombreux échanges téléphoniques. Certains paysans ne lisent pas les documents et fonctionnent surtout par l'oral. Les échanges qu'ils peuvent avoir avec les autres acteurs se font dans les champs ou au téléphone. Par exemple, au moment de la récolte 2013, il a suffi d'une conversation téléphonique pour qu'un paysan nous annonce qu'il souhaitait étudier la réponse à la sélection pour 15 populations. Beaucoup d'informations sont également échangées autour de repas ou de moments festifs<sup>xvii</sup>.

Les paysans qui sont plus à l'aise à l'oral qu'à l'écrit sont moins bien intégrés dans le projet. Pour peu qu'ils ne puissent pas se rendre aux réunions, ils peuvent se retrouver totalement à l'écart. L'animation et un temps suffisant pour discuter permettent d'adapter des modes de communications appropriés à chacun.

---

xiv. MPH, paysanne dans les Landes, lors de la réunion de juillet 2011

xv. ISG, lors de la réunion de juillet 2011

xvi. JFB, lors des entretiens avec G. Fruchaud

xvii. « *Les meilleures idées émergent autour d'une bière!* », Philippe Baret à Rome en 2011 lors de la présentation de clôture du colloque SOLIBAM. Néanmoins, certains auteurs, tel Grim<sup>124</sup>, illustrent le contraire.

## 4.3 Innovations organisationnelles et génétiques

### 4.3.1 Le réseau d'échange au centre de l'innovation

Notre travail promeut deux types d'innovations concernant la méthodologie de la sélection : génétique et organisationnelle. Le réseau de fermes et d'acteurs est au centre de ces innovations.

Les innovations génétiques portent sur le nombre de nouvelles populations créées pour l'agroécologie (chapitre 2). Ces variétés en cours de sélection sont un réservoir de nouvelles combinaisons alléliques pour poursuivre de futures sélections (chapitre 3). Ces innovations génétiques sont le résultat de l'innovation organisationnelle qui apporte une nouvelle modalité pour la gestion et la valorisation des ressources génétiques.

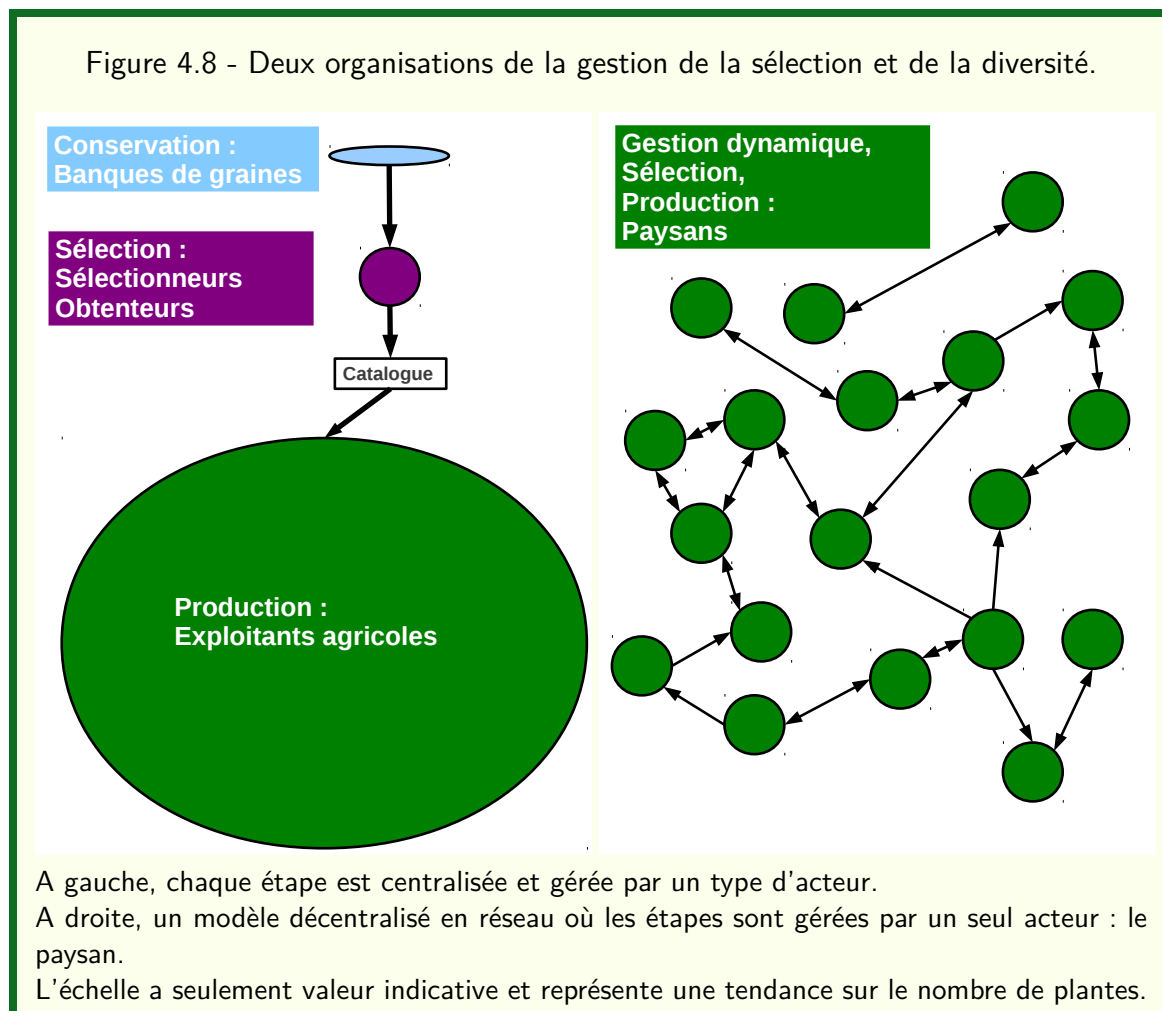
Seulement 1 à 2 % de la diversité est utilisée en sélection et les ressources génétiques ne sont pas bien caractérisées<sup>256</sup>. Dans notre projet, le réseau de fermes évalue une très large diversité qui est soumise à la sélection par les conditions environnementales et par les paysans. Le réseau relie des fermes qui font à la fois la gestion, la sélection et la production. Il n'y a aucun cadre, tel que le catalogue, pour limiter la création de nouvelles variétés-populations. Ce mode d'organisation diffère du mode d'organisation centralisé où la conservation, la sélection et la production sont séparées et gérées par des acteurs différents<sup>27</sup>. C'est le cas dans la sélection des variétés de blé tendre pour l'agriculture biologique en France<sup>206</sup>. Dans ce dernier cas, le catalogue représente un goulot d'étranglement vers la production (Figure 4.8). En effet, le catalogue contraint les variétés à être DHS et VATE (partie 1.4).

Dans un large réseau comme celui du projet de sélection participative, chaque paysan observe et évalue la diversité dans des conditions très différentes. La co-construction permet aux paysans de gagner une compétence dans l'observation des plantes, elle renforce la dimension collective. Cela peut permettre par exemple de repérer des résistances aux maladies rapidement. Bien plus rapidement que s'il faut sortir toute une partie de la diversité présente dans les banques de graines et les multiplier pour créer des lignées résistantes. En plus d'une évaluation et d'une gestion de la diversité, ce réseau permet d'augmenter la probabilité que des mutations d'intérêt se fixent, créant ainsi une variabilité nouvelle qui pourra ensuite être sélectionnée<sup>119</sup>.

Au delà d'un réseau d'acteurs et de fermes, c'est un réseau d'échanges qui est au coeur de l'innovation.

Tout d'abord, les échanges de semences. Ceux-ci permettent un maintien de la diversité

(chapitre 3) et le test par les paysans de nouvelles populations dans leur ferme (chapitre 2).



Ensuite les échanges de savoirs et de savoir-faire. La Figure 4.9 représente les échanges de savoirs et de savoir-faire entre les « paysans » et l'équipe « DEAP » : la diversité des acteurs vient enrichir le savoir collectif<sup>35;66</sup>. Les animateurs facilitent ces échanges. Il existe aussi des échanges entre le groupe « blé RSP » et l'équipe « DEAP », notamment en ce qui concerne la gestion des données, et entre le groupe « blé RSP » et les « paysans ».

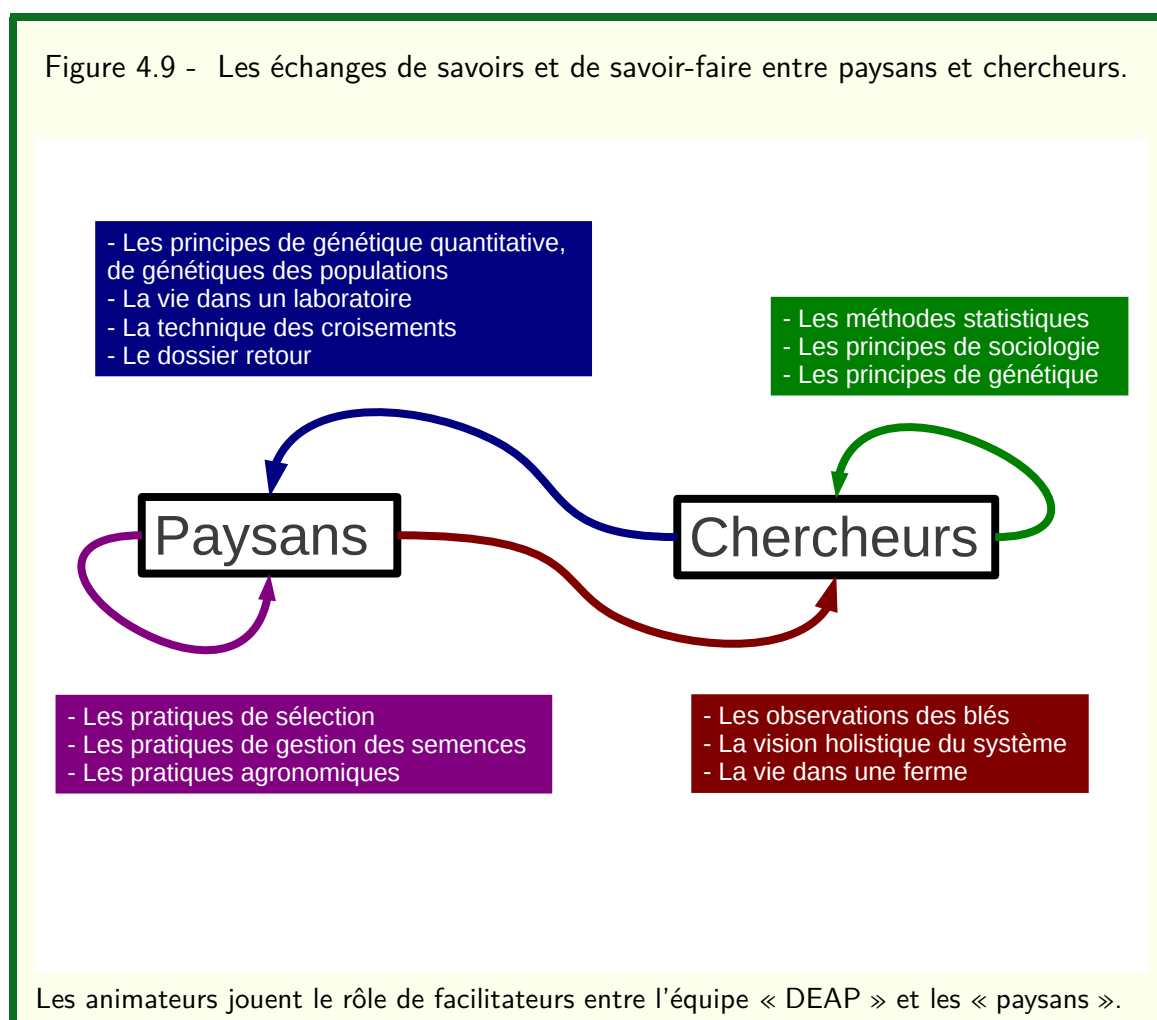
Ces échanges se font dans les champs, lors des réunions de semis, lors des formations, lors de colloques, lors de moments informels.

Pour Bruno Latour, cité par Warner<sup>252</sup>, le savoir scientifique en action devient plus puissant s'il circule de façon dynamique, s'il est connecté à divers acteurs, s'il « vascularise la société ». Les savoirs créés collectivement dans ce projet sont transférés par le RSP,

c'est à dire le réseau social d'associations et de paysans. Le groupe « blé RSP » appuie et favorise l'action des acteurs.

Ce réseau a permis de diffuser la dynamique initiée par *JFB* et *ISG* dès le début du programme. Il a permis de désenclaver des paysans isolés qui ne font pas partie de dynamiques locales. Le fait que les acteurs se sentent appartenir à un réseau augmente leur implication dans le projet.

L'organisation en réseau permet au système en entier d'être plus stable, plus résilient et donc plus durable. La diversité des structures (les maisons des semences) et la diversité des acteurs, permet de limiter les risques que le système ne s'effondre. La gouvernance de tels systèmes est une clé pour leur durabilité<sup>xviii</sup>.



Ce réseau de fermes et d'acteurs ne doit pas faire oublier les acteurs en dehors des « paysans » et du groupe « blé RSP ». Une collaboration entre les centres de ressources génétiques et les maisons des semences est très importante. Les deux types de conservation *ex-situ* et *in-situ* étant complémentaires<sup>65;80;231</sup>.

xviii. On revient sur ces aspects dans le chapitre 5

Au-delà de ces avantages du réseau valorisés à travers le projet, c'est le moyen qui est innovant : le processus de co-construction. L'innovation en tant que processus est au coeur de la recherche action<sup>35</sup>. Les outils mis en place permettent de cadrer ce processus et de valoriser la mise en réseau.

#### 4.3.2 Les innovations sur mesure

Le RSP est un réseau où la singularité de chacun peut être prise en compte<sup>62</sup>. La co-construction au niveau du collectif, c'est à dire du réseau, crée un cadre conceptuel et une façon de fonctionner commune à tous les acteurs qui préserve une flexibilité pour que chaque singularité puissent s'exprimer (Tableau 4.1). Les paysans expriment des voeux différents, ont des visions personnelles : « *Je recherche une certaine brillance* »<sup>xix</sup> ; « *C'est une sélection multifactorielle, une moyenne de facteurs* »<sup>xx</sup> ; « *Un blé, il faut que ça crache* »<sup>xxi</sup> ; « *Il me faut au moins un W de 180 pour les meuniers derrière* »<sup>xxii</sup>. Les paysans se retrouvent dans ce projet qui apporte une grande souplesse et s'impliquent selon leurs rythmes et leurs envies. La mise en réseau valorise indéniablement toutes ces initiatives et ces envies individuelles.

Un programme de sélection classique est séparé en différentes étapes de manière linéaire. Au début un cahier des charges est élaboré, puis la création de variabilité, la sélection, l'évaluation, la multiplication, l'inscription au catalogue et la diffusion sur le marché suivent<sup>67</sup>.

Dans notre cas, les étapes sont : la création de variabilité, la sélection inter-population (positive et négative), la sélection intra-population, les mélanges, l'évaluation, la multiplication, la diffusion (envoi et réception). Ces étapes s'enchaînent de manière différente pour chaque paysan. Cela illustre que chaque individu a une grande souplesse dans le projet. Par exemple, les paysans les plus expérimentés tels que *FLM* ou *JFB* sont très actifs sur de nombreuses étapes. Dès 2008, d'autres paysans rentrent dans la dynamique. Quelques-uns ne resteront dans le projet que deux ans (*JJG*), d'autres font des pauses (*CHP*). Les paysans ne font pas tous des croisements mais peuvent profiter du travail du groupe pour tester de nouvelles populations. Certains se mettent à faire des croisements quelques années après leur arrivée dans le projet (Figure 4.10).

---

xix. *JFB*, lors des entretiens avec B. Storup

xx. *OLR*, lors des entretiens avec B. Storup

xxi. *BER*, lors des entretiens avec B. Storup

xxii. *JSG*, lors de la réunion des semis en 2012.

Tableau 4.1 - Flexibilité et cadres apportés par la méthode pour les paysans et les maisons des semences

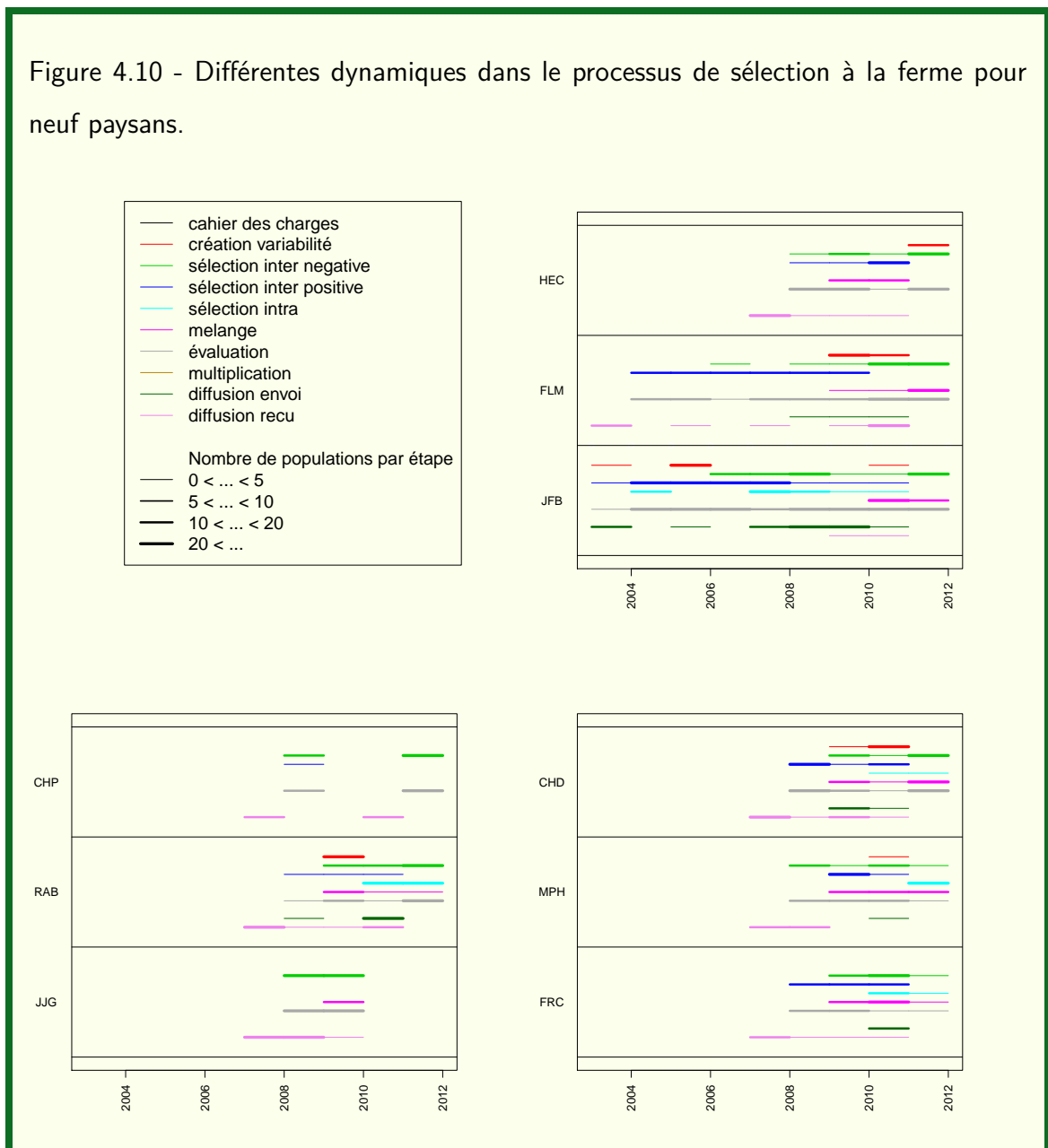
Paysans		
	Cadre	Flexibilité
Les dispositifs expérimentaux	Nécessité d'avoir au moins un témoin pour les fermes satellites et plusieurs témoins pour les fermes régionales.	Le choix des autres populations est libre. Les expériences de réponse à la sélection sont optionnelles.
Les fiches	Il faut remplir les critères choisis par le groupe.	Les commentaires sont libres.
Le dossier retour	Les informations sont en deux temps : sur la ferme et sur le réseau.	Les paysans peuvent choisir les résultats qu'ils souhaitent voir <sup>a</sup>
Maisons des semences		
	Cadre	Flexibilité
Rentrer les données dans la base de données	Les fichiers-types sont interchangeable.	Ni les variables ni les méthodes ne sont figées. Les codes sont libres et modifiables au gré des besoins.
Accéder aux informations de la base de données	L'interface fournit quelques requêtes-type	De nouvelles requêtes peuvent être mises en place dans l'interface selon les besoins.

<sup>a</sup>. Cela reste à développer mais est possible à réaliser.

Les outils ont été créés à partir de logiciels libres (`python`, `PostgreSQL`, `R`, `LATEX`). Ces logiciels sont libres d'accès et gratuits et reposent sur une large communauté très active qui se rapproche des valeurs défendues par le RSP<sup>168</sup>. Utiliser de tels logiciels permet d'être flexible et facilement accessible aux associations. Afin d'augmenter l'appropriation de ces outils, des développements sont en train de créer des interfaces « clique bouton » qui permettront à des néophytes d'utiliser ces logiciels ainsi que le développement d'un



package R qui permettra à ces outils d'être diffusés plus largement.



Le principe de cette flexibilité permet de décentraliser la méthode. La méthode est adaptable à d'autres groupes, elle pourra évoluer selon les contextes. Ce besoin de diversité s'est révélé au cœur des préoccupations des acteurs : « *Il ne faut surtout pas qu'il y ait une méthode arrêtée qui envahisse demain toute la campagne, parce que sinon il n'y a pas de diversité.* »<sup>xxiii</sup>. Cette décentralisation, en cours dès la saison 2012-2013, permettra de territorialiser la gestion des ressources génétiques et de décentraliser l'innovation. Chaque association ayant ses propres modes de fonctionnement, de nouvelles idées peuvent émerger et profiter à tout le groupe en réseau. Plus il y aura de cadres de

xxiii. JFB, lors des entretiens avec B. Storup

pensées et d'actions différents, et plus les outils et les techniques pourront s'améliorer et être partagés. La sélection participative ainsi décentralisée permet de créer une diversité de solutions à différentes échelles de temps et d'espace. La méthode mise en place, avec son cadre et sa flexibilité, favorise l'innovation des maisons des semences et des paysans.

Alors que les résultats concernant les travaux en agroécologie sont souvent critiqués car peu exportables à d'autres échelles<sup>54;246</sup>, le travail réalisé au sein de ce collectif est transposable à d'autres situations et d'autres groupes. La méthode, co-construite par le collectif, permet de valoriser les spécificités de chaque paysan et ainsi de valoriser leurs actions. C'est une action positive d'amélioration en boucle qui se met ainsi en place.

## 4.4 Evolution des pratiques des acteurs

D'après Latour<sup>152</sup> : « *toute recherche est action* ». Les résultats de la recherche ne sont pas neutres et transforment le monde dans lequel elle se situe. Le travail scientifique recompose le monde rural par son impact sur les pratiques et sur les visions du vivant qu'ont les paysans<sup>62</sup>. Nous n'étudions pas ici « le monde » mais simplement les organisations sociales qui participent au projet. Dans un souci de clarté, nous étudions, dans un premier temps, l'évolution des pratiques des « paysans » et, dans un deuxième temps, cette évolution dans le groupe « blé RSP », bien que ces deux organisations interagissent et évoluent ensemble. Enfin, je présente l'évolution de ma posture de recherche, et donc de mes pratiques de recherche, au fil de la co-construction avec les acteurs du projet.

### 4.4.1 Les « paysans »

Plus on avance dans le projet, plus le nombre de populations cultivées dans le réseau augmente pour atteindre plus de 900 pour la récolte de 2012 (Figure 4.11).

Le nombre de paysans qui font des croisements augmente également (Figure 4.12). Les croisements sont, en général, faits au Moulon sur demande des paysans mais aussi, pour quelques-uns, par les paysans eux-mêmes dans leur ferme. Réaliser des croisements où les parents sont choisis par les paysans est une démarche nouvelle, fruit de la co-construction entre les acteurs du projet. Cette pratique est inédite dans l'histoire de la sélection. Les paysans se sont appropriés le fait qu'il est possible de faire des croisements et que cela n'est pas réservé aux seuls sélectionneurs. De telles pratiques bouleversent les « paysans » et, au delà, toute la gestion des ressources génétiques et la sélection. Elles représentent un levier de création considérable.

Les sélections intra-populations évoluent également au cours du temps (Figure 4.13). De 2009 à 2010, elles sont réalisées par des paysans qui ont plus d'expérience. Dès 2011, le nombre de sélection augmente en réponse aux propositions de l'équipe de recherche d'étudier la réponse à la sélection.

La dynamique des échanges évolue avec de nouveaux paysans qui envoient des lots de semences (Figure 4.14). Ces dynamiques évoluent également en fonction des territoires. La région de départ du projet est le Sud-Ouest avec *JFB* qui concentre tous les envois en 2008, 2009 (sauf pour le témoin *Rouge-du-Roc* qui est envoyé par *FLM*) et 2010 (à l'exception de quelques populations envoyées par *OLR*). Au départ du projet, *JFB* ou l'équipe de recherche choisissent quelles populations envoyer et à qui. Dès 2011, en réponse aux retours parvenus à l'équipe de recherche, de nombreux paysans choisissent leurs populations et participent aux envois.

Les paysans qui renvoient les fiches s'investissent dans le suivi des populations (Figure 4.15). La première année, aucun retour n'a été fait, sauf pour la fiche été. Les années suivantes, il y a une tendance à ce que les retours diminuent au fur et à mesure que l'on se rapproche de la récolte. Par conséquent, certaines données comme la verse ou le poids récoltés sont souvent manquantes. Lorsque les paysans renvoient les fiches, le projet a un impact dans leurs pratiques : ils vont observer les micro-parcelles, crayon et fiche à la main. Bien sûr, certains paysans observent les blés mais ne remplissent pas les fiches ... Dans ce cas, les interactions avec les autres acteurs au sein du projet a pu avoir une influence sur leur comportement.

Il est certes impossible d'avoir accès à toute l'information<sup>27</sup> et les données sous-estiment les dynamiques qui ont lieu dans le projet. Par exemple, en 2010, un paysan a diffusé une population issue de croisement qu'il aimait bien à ses voisins. De même, en 2012, un autre paysan d'Alsace avait reçu deux populations issues des croisements du projet, alors qu'il n'en faisait pas partie (bien qu'il soit membre du RSP).

Au fur et à mesure que le projet avance, les territoires et les acteurs se l'approprient en participant de plus en plus à la vie du réseau et même, on vient de le voir, hors du réseau. Cette appropriation est d'autant plus rapide que les paysans ou les régions ont de l'expérience.

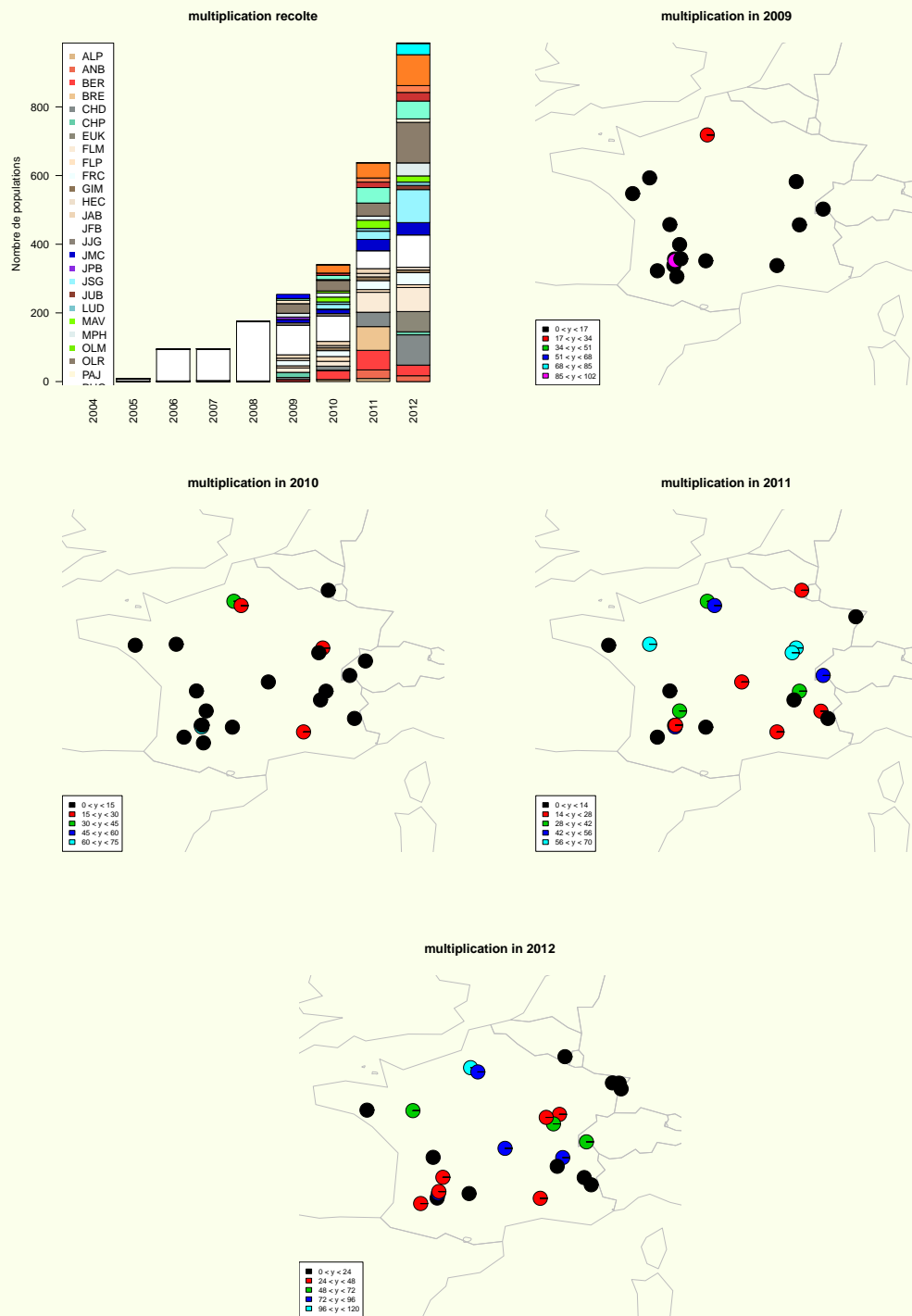
Au contact des chercheurs, de nombreux paysans sont passés d'une vision de la variété « *phare qui réussirait partout* », à une vision dynamique de mélanges issus de bouquet de sélections qui seraient adaptés localement. Les paysans ont également intégré un nouveau comportement contraignant mais librement adopté pour gérer leurs collections : « *S'ils*

*ne demandaient pas de faire un plan bien ordonné, avec des répétitions, on ne le ferait pas. [...] on est obligé de préparer les sacs de semences en fonction des petits carrés, des répétitions. Ce n'est pas ce qu'on aime le plus. C'est tout sauf poétique et marrant. »<sup>xxiv</sup>.*

---

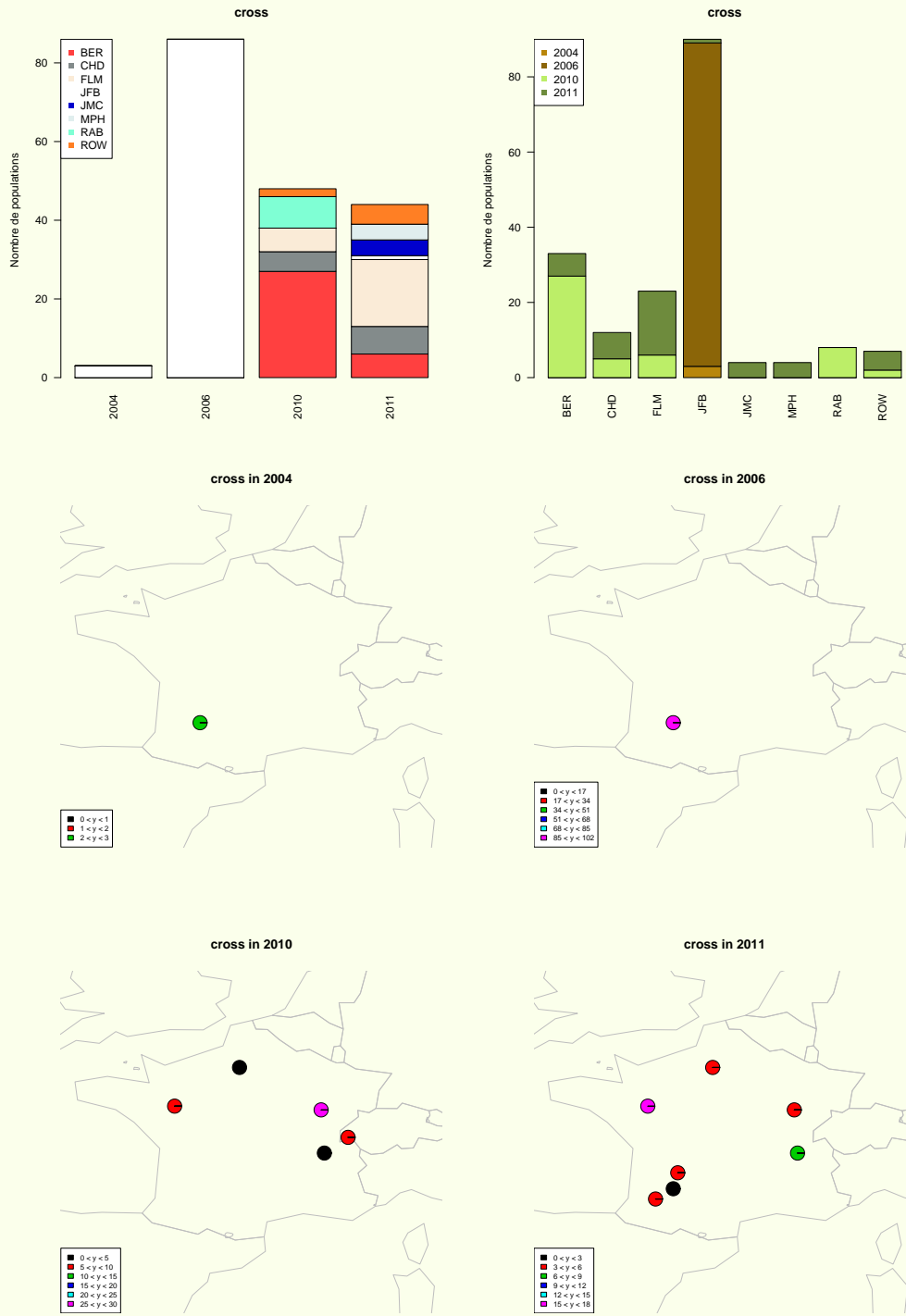
xxiv. *JFB*, lors des entretiens avec B. Storup

Figure 4.11 - Dynamique du nombre de populations cultivées dans le réseau de fermes depuis le début du projet.



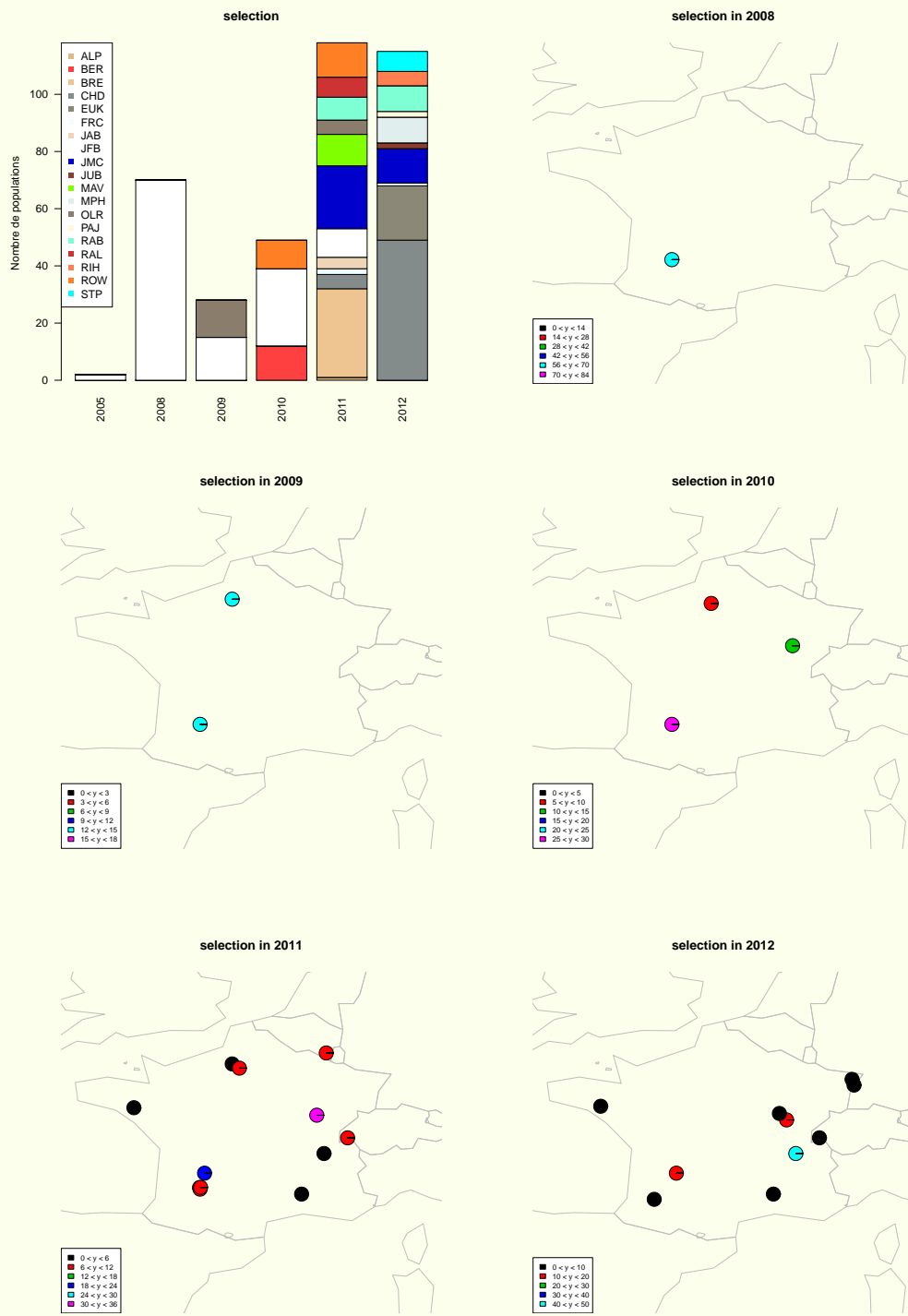
Histogramme : une couleur correspond à un paysan. Carte : une couleur correspond aux nombres de populations cultivées.

Figure 4.12 - Dynamique des croisements réalisés dans le réseau de fermes depuis le début du projet.



Si une année n'est pas représentée, c'est qu'il n'y a pas eu de croisements. Histogramme : une couleur correspond à un paysan ou une année. Carte : une couleur correspond aux nombres de croisements réalisés.

Figure 4.13 - Dynamique des sélections intra-populations dans le réseau de fermes depuis le début du projet.



Histogramme : une couleur correspond à un paysan. Carte : une couleur correspond aux nombres de populations cultivées.

Figure 4.14 - Dynamique des diffusions de populations dans le réseau de fermes depuis le début du projet.

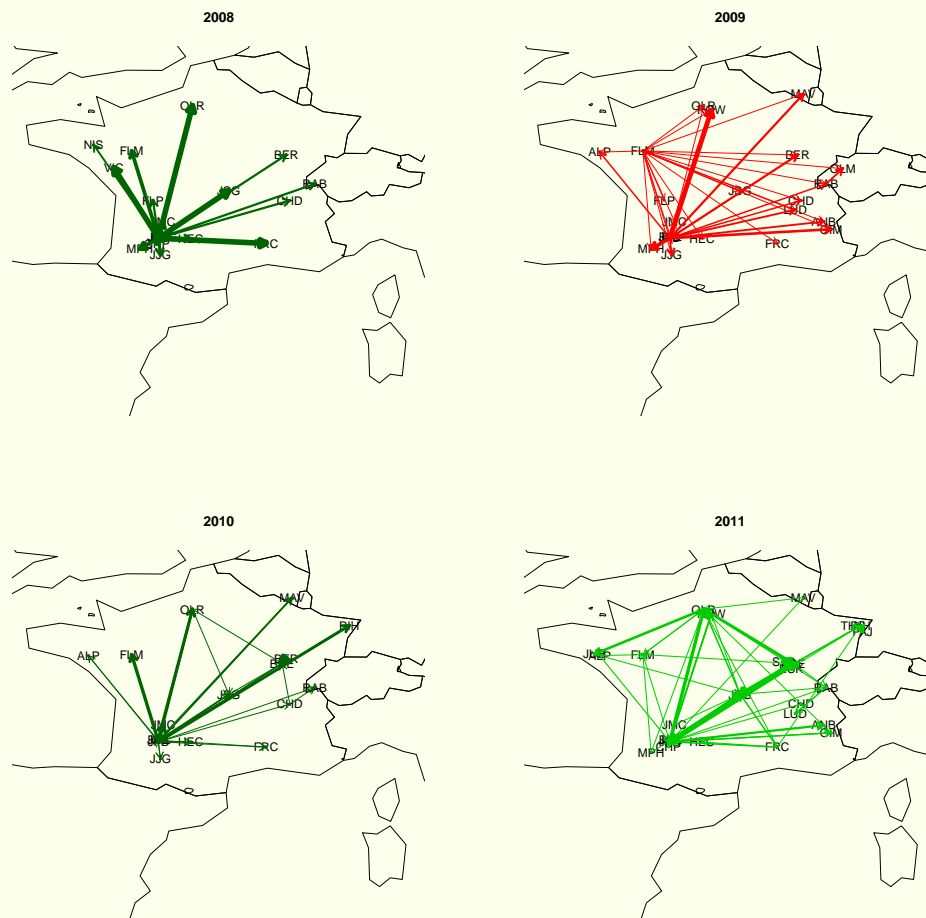
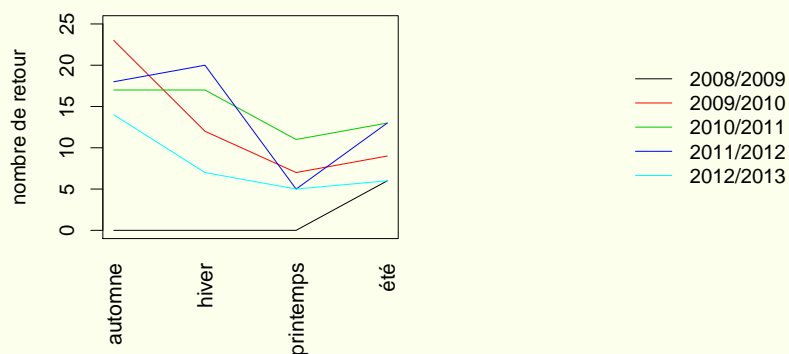


Figure 4.15 - Evolution des retours des fiches de suivi des essais pour les différentes saisons.





#### 4.4.2 Le groupe « blé RSP »

Grâce aux différentes discussions avec les paysans et l'équipe de recherche, les animateurs ont gagné en expertise et en autonomie. Cela est particulièrement vrai pour la gestion des dispositifs expérimentaux et les notions de sélection et de génétique. La plupart des animateurs ont des formations de niveau master en agronomie ou sociologie qui leur permettent de mieux appréhender ces notions.

La gestion des données a eu un impact important. Les bases de données sont des outils très utilisés dans la gestion *ex-situ*<sup>232</sup>. Cependant, de tels outils ne sont pas, à notre connaissance, utilisés dans la gestion *in-situ*. La création d'une base de données adaptée aux informations issues de sélection à la ferme et de gestion *in-situ* est une vraie nouveauté.

Depuis la réunion des semis de septembre 2011, l'assemblée générale de Colmar en décembre 2011 et la décentralisation du processus en 2012, le concept « base de données » est entré dans le groupe « blé RSP » mais aussi au delà, au « RSP » qui est l'organisation sociale qui regroupe toutes les associations et leurs membres. Il est important d'insister sur le fait que c'est le concept qui est entré au « RSP » et non pas l'outil. Ce dernier n'est pas sorti du laboratoire pour deux raisons. La première raison est technique : la première version stable n'est pas prête. La deuxième raison est déontologique : il est primordial que les « paysans », le groupe « blé RSP » et le « RSP » discutent de l'impact que pourrait avoir un tel outil dans l'organisation des groupes.

La base de données rend visibles des informations qui ne l'étaient pas. Par exemple, le nombre de populations cultivées sur le réseau ou échangées. Elle pourrait devenir un outil de surveillance. L'utilisation d'un tel outil peut mener à des fonctionnements panoptiques, tels que décrits et dénoncés dans la philosophie de Michel Foucault, pour instaurer des relations de pouvoir asymétriques. C'est à dire un fonctionnement où une personne observe les autres, sans que ces dernières se sachent observées. Un moyen d'éviter de telles dérives est de construire de manière transparente avec les acteurs une charte de bonne utilisation de l'outil. De même, un système de sécurité doit permettre de bloquer l'accès à toute personne « non autorisée ».

Les questions relatives à l'accès à l'information et au type d'informations stockées sont primordiales.

L'accès tout d'abord. Celui qui a accès à l'information peut devenir un référent clé dans le réseau. Il détient un certain pouvoir qui peut profondément modifier les relations

entre les individus d'une association. De même, des personnes malveillantes pourraient avoir accès à l'information et nuire au réseau.

Le type d'information ensuite. Certains paysans ne souhaitent pas que d'autres personnes sachent combien de populations ils cultivent, quelles sélections ils effectuent. La question du type d'information rejoint les questions d'accès. Par peur d'un accès mal géré, certains préféreraient ne rien mettre dans la base.

Au delà de ces aspects, le fait que chercheurs et paysans collaborent a constitué pour le RSP, qui soutenait cette démarche, un argument de poids dans la reconnaissance politique et juridique des pratiques de gestion de la biodiversité cultivée et de sélection des paysans dans leur ferme.

#### **4.4.3 L'équipe « DEAP » : analyse de ma posture de recherche par rapport à l'action**

Le chercheur se situe dans les sciences de l'expérimentation. C'est à dire une science où l'étude de l'effet de certain facteur se fait « tout égal par ailleurs », c'est à dire que tous les facteurs sont fixés sauf un, qui est étudié. Cette position est réductionniste, elle vise à réduire la complexité du monde en le décomposant en petites unités. Cela s'oppose aux sciences de terrain, pratiquées par les paysans, qui prend en compte la complexité des phénomènes dans les conditions réelles. Cette position est holistique, elle prend en compte la globalité et la complexité des systèmes.

La génétique des populations et la génétique quantitative se placent dans les sciences de l'expérimentation et sont réductionnistes. Elles appréhendent la réalité dans le cadre de leur champ disciplinaire, c'est à dire essentiellement sur la décomposition de la variabilité génétique et environnementale. C'est une vision probabiliste de la réalité. A titre de comparaison, ces sciences sont moins réductionnistes que la biologie moléculaire qui se concentre sur le fonctionnement d'un gène particulier.

La co-construction entre l'équipe « DEAP », les « paysans » et le groupe « blé RSP » ne nous a pas conduits à modifier notre posture réductionniste. Nous travaillons sur le blé tendre, toujours avec le même point de vue ; de ce fait, nous sommes réductionnistes. Néanmoins, la co-construction nous a conduits à modifier notre posture de recherche en passant du laboratoire au terrain.

Dans cette thèse, mon lien à l'action est très fort : j'ai animé les interactions avec les

« paysans ». Ainsi, j'ai eu un rôle particulier avec eux. J'ai commencé la thèse avec cette posture. Néanmoins, cette posture est différente de celle que j'avais lors de mon stage de fin d'étude où j'ai étudié la réponse à la sélection de *JFB* dans les *F2*. Même si les populations étaient issues d'une ferme et sélectionnées par un paysan dans un contexte donné, l'étude de la réponse à la sélection a été réalisée au Moulon : tout se passait dans le confinement du laboratoire.

Avec la thèse, ma posture a changé et j'ai migré vers le « terrain ». Je suis acteur, avec les autres membres de l'équipe de recherche, les paysans et les associations, dans le projet de sélection participative. Ma recherche et les résultats sont conditionnés, en partie, par les paysans : ce sont eux qui ont fixé les objectifs et ce sont eux qui gèrent les essais dans leurs fermes. Les productions de cette thèse sont totalement dépendantes des acteurs avec qui je co-construis le projet. Ce travail devait donc, à la fois produire des résultats tangibles dans le groupe « blé RSP » afin que les paysans continuent le projet, et valoriser scientifiquement les résultats, dans mon équipe. En d'autres termes, ma thèse s'est inscrite à deux niveaux d'action : tout d'abord, l'action avec les « paysans » et le groupe « blé RSP » en produisant des connaissances utiles pour ces acteurs ; ensuite, l'action en dehors de ces organisations sociales, en produisant des connaissances reconnues par les chercheurs en génétique quantitative, génétique des populations et méthodologie de la sélection. Cette dichotomie est propre à la recherche action<sup>35</sup>.

Ma position par rapport à l'action étant posée, il est maintenant intéressant d'analyser les articulations qui ont eu lieu entre une activité de conception et de recherche. Pour illustrer cette articulation, nous prenons comme exemple la base de données dans un premier temps, suivi des protocoles expérimentaux et des analyses statistiques dans un deuxième temps.

La base de données a été développée pour gérer la diversité des données générées par les « paysans » et le groupe « blé RSP ». La conception de cet outil est basée sur une observation fine du fonctionnement de la dynamique des lots de semences entre les « paysans » et au sein du groupe « blé RSP ». Elle a également été nourrie par des discussions avec certains animateurs des associations et certains paysans. Les interactions entre les acteurs ont permis de conceptualiser l'outil afin qu'il intègre les préoccupations des paysans. En effet, la mise en place de la base de données a été pensée pour pouvoir être utilisable par les associations et les paysans. Cet outil n'aurait jamais vu le jour si nous n'étions pas sortis de notre laboratoire. Ce sont les dynamiques de lots de semences

observées entre les « paysans » et le groupe « blé RSP » qui ont permis de développer cet outil original. La base de données m'a permis de mieux comprendre le fonctionnement du réseau d'acteurs et du réseau de semences à deux niveaux. Le premier niveau est dans la conceptualisation du réseau d'échanges. Disséquer chaque étape, penser au plus grand nombre de situations possibles, m'ont aidé à mieux comprendre le fonctionnement « théorique » du réseau. Le deuxième niveau réside dans une meilleure compréhension du réseau d'acteurs et du réseau de semences dans notre projet. Les sorties de la base de données (Figures 4.11, 4.12, 4.13, 4.14 par exemple) permettent de prendre du recul sur les événements qui se sont passés sur le réseau et ainsi de mieux analyser l'impact de notre action sur les pratiques des paysans. Prendre du recul sur ces points permet de moduler notre action avec plus de pertinence. Par exemple en interagissant plus avec les régions où les nombres d'acteurs et de populations récoltées sont plus faibles.

Les protocoles expérimentaux mis en place dans le cadre du projet de sélection participative sont un parfait exemple de co-conception qui a débouché sur la mise en place de méthodes d'analyses originales. En amélioration des plantes, une batterie de dispositifs expérimentaux et d'analyses associées à ces dispositifs existe. Nous avons vu dans la partie 4.2 comment s'est déroulée la co-conception des dispositifs expérimentaux. Les dispositifs créés sont originaux car ils ne présentent pas beaucoup de répétitions et laissent un large choix possible aux paysans pour tester les populations qu'ils souhaitent semer. De tels dispositifs n'auraient pas vu le jour dans un laboratoire de recherche. En effet, les protocoles expérimentaux, et les analyses statistiques qui leur sont associées, sont reconnus par la communauté scientifique et utilisés en routine dans les stations de recherche, pourquoi changer ? Au delà des dispositifs expérimentaux, l'originalité réside dans leurs analyses. La co-conception de ces dispositifs a transformé les pratiques liées aux analyses statistiques couramment utilisées en génétique quantitative. Grâce à une collaboration avec O. David, nous avons utilisé un modèle employé dans un autre contexte : celui des analyses d'expression de gènes. La co-conception nous a fait explorer de nouvelles voies d'expérimentation basées sur de nombreux germplasms cultivés dans un large réseau d'essais. De tels dispositifs pourront être utilisés dans un autre contexte, avec d'autres objectifs. La co-conception entre les chercheurs et les paysans a permis d'explorer des méthodes qui n'auraient jamais été explorées dans le cadre d'expérimentation standard en station de recherche. La co-conception nous a amenés à nous rapprocher d'autres disciplines. L'interdisciplinarité est au coeur de la réussite de cette co-conception. Je n'aurais jamais eu l'idée de fouiller dans les modèles statistiques utilisés dans les analyses d'expression des


gènes, si loin des champs des paysans !

Les savoirs générés ont deux domaines de validité. Le premier correspond aux « paysans » : les résultats sur le comportement des populations est vrai dans les fermes, selon les pratiques des paysans. Le deuxième correspond au groupe « blé RSP » et à l'équipe « DEAP » voire au delà, à travers les articles scientifiques et les outils développés : analyses statistiques, analyses de la diversité au niveau moléculaire et phénotypique, principes méthodologiques qui promeuvent des modes d'organisation décentralisés et participatifs, base de données, fiches de suivi, « dossier retour ». Ces résultats créés peuvent aller jusqu'aux décideurs politiques<sup>xxv</sup>.

---

xxv. On revient sur ce point dans le chapitre 5





## 5 . Transition vers la création d'un système semencier adapté à l'agroécologie

Résultats principaux de la thèse et propositions au collectif de travail

Propositions pour une transition au niveau local et régional

Propositions au niveau national pour accompagner la transition

Ce dernier chapitre reprend dans un premier temps les principaux résultats de cette thèse et fait des propositions au collectif de travail, qui est représenté par les différentes organisations sociales que sont les paysans, les associations et l'équipe de recherche. Ces premiers résultats constituent une étape dans un processus de transition vers la création d'un système semencier adapté à l'agroécologie.

Dans un deuxième temps, nous faisons des propositions pour opérer cette transition au niveau des paysans et des maisons des semences paysannes. Il s'agit respectivement du niveau local et du niveau régional.

Enfin, nous faisons des propositions au niveau national afin d'accompagner une transition plus globale qui répondrait aux évolutions locales et régionales.

## 5.1 Résultats principaux de la thèse et propositions au collectif de travail

Nous avons vu dans le chapitre 1 que l'agroécologie est reconnue dans la littérature scientifique et dans des textes internationaux comme une alternative à mettre en place pour produire de la nourriture tout en préservant les ressources naturelles<sup>91;92;133;162;240</sup>. Parmi les multiples maillons qui composent l'agroécologie, nous nous sommes intéressés à la question de l'amélioration variétale et de la gestion des ressources génétiques. La sélection décentralisée et la gestion à la ferme des ressources génétiques sont proposées comme pistes de travail. Ces deux aspects sont étroitement liés et reposent sur l'interaction entre de nombreux acteurs. Ainsi, nous proposons de rendre la sélection et la gestion des ressources génétiques participatives, issues d'un processus de construction collective entre chercheurs, associations et paysans.

Cette collaboration entre les acteurs a un impact dans les champs des paysans et aussi dans les laboratoires des chercheurs. La mise en place de cette recherche action est au coeur des principes méthodologiques de l'agroécologie<sup>222</sup>.

Un changement de paradigme est nécessaire pour prendre en compte la complexité des agrosystèmes d'un point de vue agronomique et socio-économique. Le projet de sélection participative sur le blé tendre, dans lequel ma thèse s'est insérée, s'intègre dans cette réflexion.

Notre travail a permis de mettre en place une méthodologie innovante de sélection variétale et de gestion des ressources génétiques pour l'agroécologie, basée sur la co-construction entre les acteurs ainsi que la décentralisation de la sélection. Plus précisément, les résultats de cette thèse ont été divisés en trois chapitres. Ces chapitres sont le fruit de discussions avec les membres de l'équipe DEAP et d'interactions avec des chercheurs d'autres disciplines. Alors que le coeur de métier de mon équipe d'accueil (DEAP) sont la génétique quantitative et la génétique des populations, mobilisées dans les chapitres 2 et 3, j'ai pu utiliser les statistiques bayésiennes grâce à une collaboration avec O. David, chercheur à l'unité MIA<sup>i</sup> de l'INRA de Jouy en Josas, dans le chapitre 2. Enfin, des interactions avec L. Hazard, chercheur dans l'unité AGIR à l'INRA de Toulouse et E. Demeulenaere, chercheuse dans l'équipe « Anthropologie de la conservation » au Centre National de la Recherche Scientifique (CNRS) à Paris, m'ont permis de réaliser le chapitre 4, qui renvoie à des concepts de sciences sociales. Le chapitre 4 traite également des

---

i. Mathématique et Informatique Appliquées



questions de gestion des données. La construction d'une base de données est issue d'une collaboration avec Y. de Oliveira, ingénieur dans l'équipe Atelier de Bio-informatique et d'Informatique (ABI) à l'INRA du Moulon.

Dans cette partie, je reprends les résultats principaux de chaque chapitre, analyse les limites et les contraintes que j'ai rencontrées et propose des pistes d'amélioration aux acteurs du projet.

### 5.1.1 L'analyse phénotypique

Nous pouvons scinder en deux les résultats de l'analyse phénotypique du chapitre 2. Il y a dans un premier temps les résultats méthodologiques, c'est à dire les dispositifs expérimentaux et les modèles statistiques développés pendant la thèse. Dans un deuxième temps, il y a les résultats propres au programme de sélection participative, qui caractérisent les sélections des paysans.

#### 5.1.1.1 Apports méthodologiques

Les apports méthodologiques font l'objet de deux manuscrits d'articles, présentés en annexe K et L. La collaboration avec O. David m'a permis d'explorer de nouveaux modèles d'analyses des données basés sur les statistiques bayésiennes. Ces modèles permettent de compenser le déséquilibre des données que l'on peut rencontrer dans le réseau de fermes. L'approche hiérarchique permet de profiter de l'information obtenue sur le réseau de fermes, qui regroupe de nombreux essais ayant une même structure, en faisant l'hypothèse que la variance environnementale au sein de chaque ferme est issue d'une distribution commune à toutes les variances environnementales de toutes les fermes. Un nombre important de fermes avec au moins un témoin répété est nécessaire pour que le modèle puisse fonctionner. Il serait intéressant de faire des simulations pour tester les limites et les potentialités du modèle. J'ai commencé à explorer ce point en codant des fonctions dans R qui génèrent un réseau d'essais avec une certaine proportion de fermes régionales et de fermes satellites avec plus ou moins de blocs et de témoins répétés. Une fois le réseau construit, j'ai simulé les valeurs de chaque parcelle de chaque ferme. Ces valeurs ont été tirées dans une loi normale. Une variance environnementale pour chaque ferme a été tirée dans une loi inverse gamma de paramètres  $\nu$  et  $\rho$ , censé représenter la loi de distribution commune à toutes les variances environnementales de toutes les fermes. Avec un réseau d'essai équilibré, le modèle prédisait parfaitement les vraies valeurs. Cela nous a rassurés quant à la validité du modèle. Je n'ai pas eu le temps de regarder plus en détail l'impact

du déséquilibre sur les prédictions des valeurs, mais ce travail sera poursuivi après la thèse.

Nous avons utilisé le même type de modèle bayésien hiérarchique pour étudier les interactions  $G \times E$  avec le modèle de Finlay Wilkinson. Le fait de ne pas retrouver systématiquement les différentes années d'expérimentation sur chaque ferme regroupées par environnement pose question sur les estimations de ces effets ou sur leur intérêt à les utiliser en sélection. Néanmoins, ce modèle s'est avéré en général meilleur que le modèle additif ou le modèle de Finlay Wilkinson modélisé de façon non hiérarchique. Afin de mieux explorer ce modèle, il faudrait également réaliser des simulations. J'ai réalisé des fonctions dans R qui permettent de simuler un réseau d'essais avec un certain nombre de germplasms répartis dans plus ou moins d'environnements. Une fois le réseau construit, j'ai simulé les valeurs de chaque parcelle de chaque ferme. Les différents effets ont été tirés dans des lois normales. Avec un réseau d'essais équilibré, le modèle prédisait les vraies valeurs parfaitement. Je n'ai pas eu le temps d'aller plus loin dans les analyses. Il faudra explorer la capacité prédictive du modèle pour des niveaux de déséquilibre croissants.

Au delà de ces deux modèles, il serait intéressant de tester un modèle « deux en un » qui combinerait à la fois une estimation des moyennes dans les fermes et des effets globaux génétiques, environnementaux et d'interactions. Ce modèle peut être écrit, il reste à savoir si les paramètres convergeront et si le modèle sera suffisamment prédictif. Un travail pour fusionner les deux fonctions de simulation sera nécessaire pour tester un tel modèle.

Ces modèles peuvent être pertinents pour l'expérimentation dans un large réseau d'essais, ce qui va au delà de notre projet. Ces outils statistiques pourraient permettre de repenser l'évaluation des ressources génétiques dans de larges réseaux d'expérimentation. Il pourrait également présenter un intérêt dans les projets de génomique haut débit dans lesquels plusieurs milliers de génotypes sont génotypés ou séquencés et où le facteur limitant devient le phénotypage.

#### 5.1.1.2 Les résultats dans le cadre du projet

Les principaux résultats du programme concernent l'étude des sélections faites par les paysans. Il y a deux niveaux de sélection : la sélection entre les populations et la sélection à l'intérieur des populations. Pour la sélection entre les populations, il a été difficile d'avoir un jeu de données avec des informations fiables. En effet, savoir quelles populations n'ont délibérément pas été semées, est compliqué. Néanmoins, sur le peu d'informations dont nous disposons (seulement 40 populations!), nous avons observé une tendance à ce que les populations non semées l'année suivante aient en moyenne une valeur faible pour des

caractères d'intérêt pour les paysans, tels que le taux de protéine, le poids de mille grains, ou encore la hauteur des plantes. Pour la hauteur, les paysans se séparent des populations avec les plantes plus hautes, ce qui est cohérent avec leur souhait de sélectionner des populations qui ne versent pas.

Si la sélection entre les populations est quelque chose de très répandue parmi les paysans participant au projet, la sélection massale intra-population n'est venue que plus tard.

Les premiers résultats sur la réponse à la sélection intra-population ont été obtenus avec l'expérience d'évaluation du Moulon sur les effets de la sélection de *JFB* en génération précoce. Les résultats ont été très encourageants (Annexe J). En effet, 24 populations sur 38 ont répondu significativement sur au moins un caractère, quatre populations ont répondu sur cinq caractères ou plus. Pour des caractères d'intérêt comme le poids de mille grains ou le poids de grain par épi, quand la réponse était significative, elle était toujours positive.

Les paysans ont souhaité eux aussi étudier la réponse à leur sélection. Les expériences se sont faites dans les fermes avec un protocole mis en place pour l'occasion. On peut retenir deux résultats principaux de cette expérience : la premier concerne sa mise en place avec les acteurs, le deuxième concerne les résultats sur la réponse elle même.

Pour le premier point, il y a eu au départ des problèmes de communication entre les partenaires. C'est pour cela que seulement 37 réponses à la sélection ont été analysées alors que l'on avait au départ 183 bouquets de sélection. Par conséquent, les bouquets analysés ne représentent que 20% de ce qui avait été sélectionné ; sans compter les sélections qui ont été faites mais auxquelles nous n'avons pas eu accès. L'équipe de recherche a mieux expliqué l'intérêt de cette expérimentation et, le temps passant, les paysans ont mieux compris l'intérêt de sa mise en place dans leur ferme. Les résultats sur la réponse à la sélection à la ferme devraient être plus nombreux à l'avenir.

Pour le deuxième point, les résultats sur les sélections des paysans dans leurs fermes ne confirment pas les premiers résultats des sélections de *JFB*. La réponse est significative et le plus souvent dans le même sens pour la couleur, qui devient plus foncée et le nombre d'épillets qui augmente. Alors que pour la sélection de *JFB*, lorsque la réponse était significative pour des caractères d'intérêt comme le poids de mille grains ou le poids du grain par épi, elle était toujours positive, on ne retrouve pas cela dans les fermes. Même si quelques populations répondent positivement sur ces caractères, il n'y a pas de tendance globalement positive. Ceci peut s'expliquer par des héritabilités faibles. En effet, pour le

différentiel de sélection, les épis plus lourds et avec un poids de mille grains élevé étaient bien sélectionnés. Si l'héritabilité est faible, c'est soit que la variabilité génétique a été épuisée car on est en génération plus avancée, mais les résultats du chapitre 3 semblent dire le contraire, soit que la variance environnementale est trop forte. Pour remédier à cela, on pourra proposer différentes évolutions : augmenter la taille des parcelles pour augmenter la taille des populations et sélectionner des plantes, ou des épis, selon un quadrillage assez dense comparant les plantes à leurs voisines. Ces résultats peuvent aussi être dus à des problèmes d'estimation des réponses. Alors que l'expérience avec les populations de *JFB* a été conduite à la station expérimentale avec trois blocs complets répétés, les réponses des paysans ont été étudiées dans leurs fermes. De plus, alors que l'expérience au Moulon était dédiée à l'étude de la réponse à la sélection, les expériences dans les fermes ne l'était pas. Le dispositif expérimental n'était donc pas optimal dans les fermes pour étudier la réponse à la sélection. Pour la récolte 2012-2013, 50 épis ont été mesurés pour quantifier la réponse à la sélection dans les fermes. Les données n'ont pas été analysés, mais il sera intéressant de voir si cela a un impact sur la cohérence entre les différentiels de sélection et les réponses. D'autres mécanismes ont pu jouer, tels que les interactions germplasm  $\times$  année qui viennent diminuer la corrélation parents-descendants et donc la réponse à la sélection. Le maintien d'une diversité intra-population ainsi que l'assemblage de populations sont les garanties pour les paysans d'une stabilité de leur culture au cours du temps.

### 5.1.2 L'analyse d'évaluation du programme de sélection participative

Le chapitre 3 a permis d'étudier, sur des données phénotypiques et moléculaires, l'évolution de la diversité créée par les différents croisements et dans 25 fermes aux environnements et aux pratiques contrastés. Si ce chapitre est le moins abouti de la thèse par manque de temps, il permet néanmoins d'obtenir quelques résultats intéressants pour la gestion du programme de sélection participative en ce qui concerne le choix des parents et l'impact de la sélection décentralisée.

Nous avons observé peu d'influence du type de germplasm sur la diversité génétique créée. C'est plutôt le choix de parents hétérogènes et distants génétiquement qui vont constituer un levier pour la création de germplasms avec un fort potentiel de diversification.

Au niveau phénotypique, le type a présenté un effet uniquement pour la courbe et le poids de l'épi. Il nous reste à compléter les analyses pour le rendement, la teneur en

protéine et la verse. S'il s'avère que certains types de germplasms présentent plus de potentiels pour des caractères prioritaires pour les paysans, ils pourront utiliser ces résultats pour choisir leurs croisements. Avec l'analyse multivariée, le type variétés modernes  $\times$  variétés de pays ( $VM \times VP$ ) apparaissait le plus diversifié. Il pourrait donc être intéressant de privilégier ce type de croisement pour générer une large diversité phénotypique. Cela restera à confirmer une fois que l'on aura ajouté les autres données.

Nous avons observé que le programme de sélection participative avait généré de nouveaux germplasms qui présentaient des structurations de leur diversité génétique très contrastées. Certains germplasms étaient composés de populations proches génétiquement, alors que d'autres étaient composés de populations beaucoup plus différenciées. Cette diversité est liée à la distance entre les parents et à l'histoire évolutive des différentes populations. En effet, chaque germplasm se comporte comme une métapopulation et est soumis aux différentes forces évolutives. Parmi les points remarquables, on retiendra que la sélection massale intra-population ne réduit pas forcément la diversité intra-population. De plus, elle a permis dans certains cas d'explorer une diversité originale comparée aux autres populations. Nous avons également observé des cas de sélection divergente sur les marqueurs associés aux gènes de précocité, témoins d'une réponse différenciée des sous-populations à la diversité des environnements. Malgré ces évolutions, la diversité au sein des populations est restée en général importante. Ces résultats permettent d'étendre les connaissances sur l'évolution de la diversité au sein des populations en gestion dynamique en station et à la ferme<sup>80</sup> aux populations soumises à la sélection humaine au sein d'un programme de sélection participative. Ce travail permet de documenter l'impact de la sélection participative sur la gestion des ressources génétiques et de caractériser la sélection à la ferme comme un élément de la gestion dynamique de la diversité.

### 5.1.3 Le réseau d'acteurs et la co-construction entre acteurs au coeur de l'innovation

Stassart et al.<sup>222</sup> citent trois principes socio-économiques qui sous-tendent l'agroécologie. Le premier est de créer des connaissances et des capacités collectives d'adaptation à travers des réseaux. Le deuxième est de valoriser la diversité des savoirs. Le troisième est de favoriser la possibilité de choix et d'autonomie dans les débouchés économiques. Si les deux premiers points sont en voie d'être atteints, voire le sont, par l'intermédiaire de notre méthodologie, le dernier ne l'est pas car en dehors de notre champ d'action.

Les résultats issus des chapitres 2 et 3 sont indissociables du mode de fonctionnement

entre les acteurs du projet. La co-construction des différentes étapes a amené à se poser de nouvelles questions, qui ont permis de développer de nouvelles méthodes, de nouveaux outils, mais aussi de nouvelles connaissances. Ces innovations sont le fruit de la construction collective.

Parmi les outils générés lors de cette construction entre les acteurs, on retiendra notamment les fiches de notations, le « dossier retour », la base de données spécifique, les dispositifs expérimentaux et les statistiques. Ces outils ont été pensés pour rester flexibles afin de s'adapter au mieux aux spécificités individuelles de chaque groupe local, chaque paysan, mais également pour pouvoir s'adapter à d'autres groupes dans d'autres contextes et sur d'autres espèces. Alors que le projet co-construit permet de créer un cadre conceptuel, les outils permettent une grande souplesse d'action pour chaque acteur (choix des populations à évaluer dans la ferme notamment).

Il est important de garder à l'esprit que ces outils et ces méthodes ont été construits à la suite de l'interaction entre les partenaires des différentes organisations sociales, mais aussi grâce à l'approche interdisciplinaire ou à l'interface avec d'autres disciplines que j'ai eu la chance de pouvoir pratiquer pendant ma thèse, notamment avec les statistiques et les sciences humaines et sociales. Comme nous l'avons vu dans la partie 1.1.3, cette interdisciplinarité est au coeur de l'agroécologie<sup>54</sup>.

Le réseau d'acteurs et la co-construction sont à la base de la réussite de ce projet. C'est le réseau qui permet de développer deux types d'innovations : organisationnelle et génétique. La seconde dérivant de la première. Au delà du progrès génétique, indicateur classiquement proposé par les sélectionneurs, le progrès social est au centre des enjeux de la sélection et de la gestion des ressources génétiques pour l'agroécologie<sup>69</sup>.

Une question peut être posée : jusqu'où aller dans la décentralisation ? La place du blé dans la rotation peut varier (Figure 1.2 dans la partie 1.2), ce qui aura un impact sur son comportement. Il pourrait être intéressant d'avoir deux sélections selon les types de sols ou la place du blé dans la rotation, mais cela paraît compliqué à gérer. L'idéal serait plutôt de sélectionner une diversité qui soit suffisamment large et permette ainsi de tamponner les différences environnementales rencontrées dans la ferme. Ces points restent à discuter et à étudier.

Dans le cadre du projet, nous avons identifié six paysans (*JFB*, *CHD*, *RAB*, *FRC*, *JSG*, *BER*) qui avaient utilisé certaines populations issues du programme de sélection participative dans des mélanges cultivés en plein champ et utilisés en production. Ces

mélanges pouvaient être uniquement composés de populations issues des croisements, de sélections massales intra-population, ou alors être un assemblage de populations issues de croisements et d'autres populations de type variétés anciennes, de pays, voire modernes, dans ce dernier cas il s'agit de variétés créées pour l'agriculture biologique. Cette pratique débouche sur de nouvelles questions de recherches. Beaucoup de paysans nous demandent comment mener un mélange, comment sélectionner pour un mélange. On peut imaginer les populations testées dans le cadre du programme comme un réservoir prêt à alimenter le mélange dynamique déjà cultivé dans la ferme ou alors prêt à créer un nouveau mélange. On peut imaginer des cycles de sélection récurrente sur un mélange.

En plus de ces questions sur la conduite et la sélection des mélanges, les paysans sont préoccupés par les aspects sanitaires comme la carie par exemple. De plus, les paysans aimeraient mieux connaître les déterminismes de la qualité du pain qui est souvent le produit final du travail de sélection, notamment au niveau intolérances au gluten, ces dernières pouvant être en partie une conséquence de la sélection moderne<sup>244</sup>. Notre travail ne peut pas être dissocié d'une étude sur la transformation du blé en pain. Une collaboration étroite avec des disciplines travaillant sur les analyses organoleptiques et nutritionnelles semble nécessaire dans la suite du projet<sup>174;248</sup>.

#### 5.1.4 Retours sur les objectifs de départ de la thèse et perspectives

Cette thèse avait deux objectifs :

1. le développement de méthodes et d'outils opérationnels pour la gestion de la biodiversité cultivée et la sélection à la ferme,
2. le renforcement de l'apprentissage et de l'autonomie des paysans en matière de gestion et de sélection des semences.

Ma thèse a contribué à développer de nouvelles méthodes, basées sur la co-construction entre acteurs et la décentralisation de la sélection et, comme nous l'avons vu, de nouveaux outils. Le processus de co-construction est, de mon point de vue, un des résultats les plus importants de cette thèse. C'est cette construction entre les acteurs qui est à l'origine des innovations organisationnelles et génétiques. C'est le réseau d'acteurs, l'organisation sociale, qui est au centre de notre méthodologie de sélection variétale et de gestion des ressources génétiques pour l'agroécologie.

Les résultats que nous avons générés ont permis de renforcer l'apprentissage des paysans et de les rendre plus autonomes dans leur sélection et dans la gestion de leur semence. Nous avons vu que les pratiques des paysans avaient évolué depuis le début du projet : créa-

tion de croisements, augmentation du nombre de sélection massale intra-population, augmentation du nombre de populations évaluées sur les fermes, augmentation des échanges de semences entre paysans. En plus de ces pratiques, leurs visions du vivant ont pu changer, en lien avec la vision dynamique de la diversité des chercheurs<sup>62</sup>. Ces évolutions sont autant d'illustrations de l'impact de la co-construction dans les champs. Un point remarquable est également l'évolution de notre position de recherche qui nous a amenés vers plus d'interdisciplinarité.

Tout cela est encourageant mais comment envisager la suite ? Après ce travail de recherche qui laisse encore beaucoup de questions ouvertes, il sera nécessaire de passer par une phase de développement. L'année 2012-2013 a été une première année de transition pour décentraliser la méthodologie mise en place. L'équipe de recherche a aidé les associations dans la prise en main des outils et des bonnes pratiques pour accompagner les paysans localement. Cette première phase permettra aux associations de s'autonomiser. Au même moment, l'équipe de recherche réfléchit à la manière de transmettre ce travail collectif au « monde », c'est à dire d'autres groupes d'acteurs, d'autres contextes et d'autres espèces. Dans cette optique, un travail de développement est en cours pour rendre les outils plus facilement utilisables par les associations et les paysans. Il est notamment prévu dans la suite de créer un package R qui regroupera les fonctions d'analyse et de visualisation des données de la base. De plus, l'interface sera améliorée dans le cadre d'un stage de master 2 débutant en 2014.

La suite de ce chapitre vise à proposer des pistes pour que cette décentralisation se poursuive au delà de notre collectif de travail. Des propositions sont avancées pour améliorer les projets de sélection participative au niveau local et régional dans un premier temps, puis au niveau national dans un deuxième temps.

## 5.2 Propositions pour une transition au niveau local et régional

Après le développement de l'organisation et des méthodes pour la sélection participative par le collectif de travail, la décentralisation des méthodes et le renforcement de l'organisation locale sont des étapes importantes dans la transition vers la mise en place d'un nouveau système semencier adapté à l'agroécologie.

Les maisons des semences paysannes sont organisées au niveau régional ou local et se construisent à partir du réseau local que représentent leurs adhérents : paysans, artisans



semenciers, jardiniers, amateurs soucieux de la biodiversité cultivée. Au sein de chaque maison des semences paysannes, des échanges de savoirs, de savoir-faire, de semences sont courants. Ces maisons des semences paysannes peuvent être organisées au sein d'un réseau plus large comme le Réseau Semences Paysannes (RSP).

Afin de faire des propositions qui puissent être utiles pour proposer une transmission des acquis de notre collectif aux maisons des semences paysannes en développement au sein du RSP, dans un premier temps nous caractérisons la diversité rencontrée au sein des maisons des semences paysannes. Ensuite, des propositions sont faites pour diffuser les outils et les méthodes développées et permettre leur adaptation dans ces contextes différents.

### 5.2.1 La diversité au sein des maisons des semences paysannes

Le RSP regroupe de nombreuses associations. Ces associations ont des histoires et des modes d'organisation différents. A ces histoires différentes s'ajoutent des adhérents différents qui vont apporter des idées et des orientations diverses<sup>13;76;132;155;181</sup>.

D'après Levrouw<sup>155</sup>, qui a réalisé une étude sur la diversité des maisons des semences paysannes au sein du RSP, le fonctionnement des collectifs est influencé par quatre critères principaux (Figure 5.1)

1. les acteurs, leurs motivations, leurs cultures : paysans, jardiniers, citoyens, chercheurs, artisans semenciers, etc ;
2. l'origine du collectif ;
3. le contexte et la dimension des collectifs (qui comprend également leurs moyens humains et techniques) ;
4. les types de plantes : autogames, allogames, végétatives, pérennes, annuelles, bisannuelles.

Les maisons des semences paysannes doivent accompagner les paysans dans la gestion et l'utilisation des ressources génétiques au sein de leurs systèmes de production. Elles permettent de formaliser une action collective<sup>13</sup>. Elles doivent faciliter l'accès à la diversité et à l'information, faciliter leur utilisation et enfin accompagner les paysans dans l'évaluation de leurs ressources<sup>143</sup>. Les maisons des semences paysannes sont contrôlées par leurs adhérents, notamment les paysans, ce sont eux qui gouvernent leur fonctionnement<sup>ii</sup>. Ce point est essentiel car c'est la gouvernance des semences qui est au coeur des enjeux de

---

ii. Le type de gouvernance variera selon les maisons des semences. Certaines donnent un poids plus important aux paysans plutôt qu'aux jardiniers par exemple.

l'agrobiodiversité<sup>211</sup>. Le contrôle de la gestion des semences permet *in-fine* le contrôle de la nourriture. Les maisons des semences paysannes peuvent être vues comme un modèle « territorialisé participatif » en opposition à un modèle « oligopolistique intégré »<sup>24</sup>.

Figure 5.1 - Diversité des maisons des semences paysannes au sein du RSP.

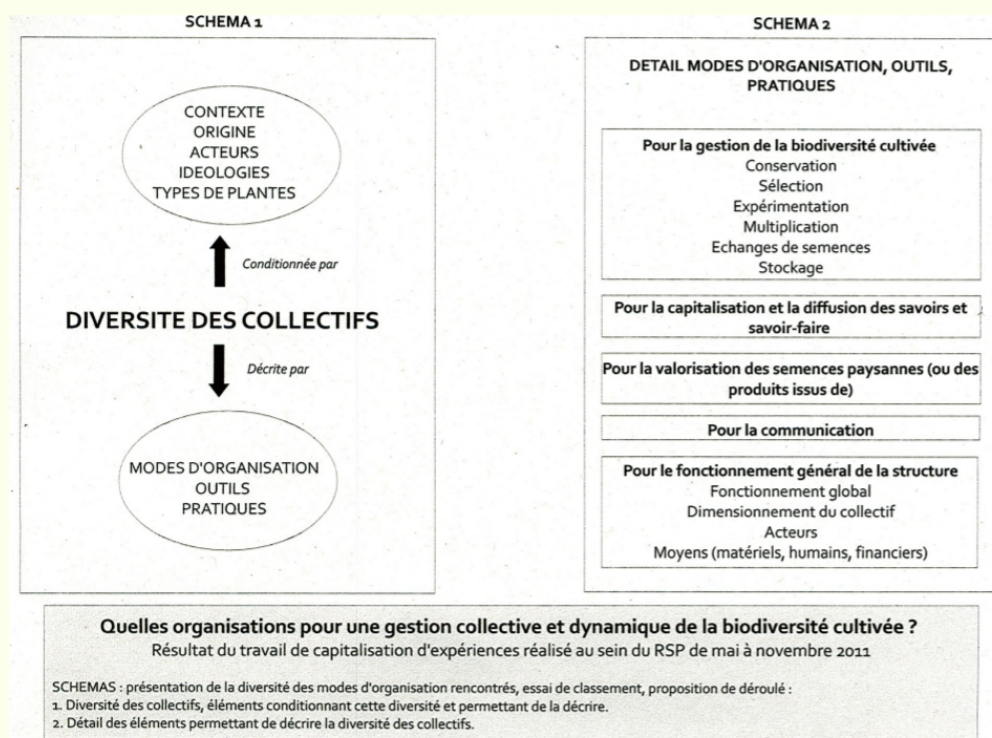


Schéma issu du travail de F. Levrouw, présenté lors de l'assemblée générale du RSP à Colmar en 2011.

Levrouw<sup>155</sup> définit les maisons des semences paysannes comme des structures qui peuvent potentiellement regrouper les activités suivantes :

1. la prospection, la recherche de variétés anciennes ou locales ;
2. la gestion dynamique des semences (conservation, sélection, expérimentation, multiplication, échanges de semences, stockage) ;
3. les échanges de savoirs et savoir-faire ;
4. la valorisation des semences paysannes et des produits qui en sont issus ;
5. la communication, la sensibilisation ;
6. l'animation du collectif (gestion des moyens humains, matériels et financiers).

Notre travail, dans le cadre du projet de sélection participative, peut venir compléter les points 1., 2., 3. et 6. Dans la partie suivante, je propose un modèle flexible de sélection participative, inspiré de notre projet, et qui pourrait s'adapter à la diversité des collectifs.

### 5.2.2 Proposition d'un modèle flexible de sélection et de gestion de la diversité pour les maisons des semences paysannes

Notre méthode n'est pas une solution en soi qui serait universelle et qu'il faudrait reproduire telle quelle dans n'importe quel contexte avec n'importe quel groupe d'acteurs. Nous pouvons distinguer trois grands enseignements dans la mise en place de la méthodologie de sélection participative sur le blé : la méthode de co-construction avec les acteurs, les outils développés, eux-mêmes issus de cette co-construction et la valorisation du large réseau d'acteurs.

Une maison des semences paysannes sans animateur ou animatrice aura de grande difficulté à mettre en place un programme de sélection participative. Nous avons vu dans notre projet le rôle central de ce dernier dans le suivi des fiches, l'acquisition des données, l'organisation des réunions, le retour des résultats aux paysans.

Par contre, la mise en place de projet de sélection participative ne nécessite pas forcément une collaboration prolongée avec les chercheurs. Ces derniers peuvent être impliqués de manière ponctuelle lors de formations ou de réunions d'échanges, mais le processus de sélection participative « au jour le jour » peut être géré par les maisons des semences paysannes. Bien sûr, cela n'empêche pas de travailler sur de nouveaux projets de recherche participative entre chercheurs, associations et paysans.

J'ai identifié trois éléments qui me semblent importants pour construire un programme de sélection participative : prendre le temps de la construction entre les acteurs ; adapter les outils pour répondre aux divers besoins ; maintenir un fonctionnement en réseau pour les maisons des semences paysannes. Ces trois points sont détaillés ci-dessous.

#### 5.2.2.1 Prendre le temps de la construction entre les acteurs

Notre projet est le témoin d'une expérience entre un groupe de paysans, d'associations et de chercheurs. Les étapes de co-construction entre les acteurs, dans le cadre de notre programme de sélection participative apportent des éléments de réflexion sur la mise en place de tels projets. Ces réflexions doivent également s'appuyer sur d'autres expériences décrites tels que des mémoires sur des stages réalisés au sein du RSP<sup>13;76;132;181</sup>.

Afin que chaque collectif puisse répondre à ses objectifs, il est primordial que des échanges entre les paysans et les animateurs (voire les chercheurs de manière plus ponctuelle) se fassent. Cela prend beaucoup de temps mais nous paraît indispensable, en effet la co-construction est au centre du processus de sélection participative. Par exemple, notre groupe n'a pas répondu au premier appel d'offre « Ecosystèmes, Agricultures, Alimenta-

tion » de la Fondation de France en mars 2013 mais au deuxième en octobre 2013. Cela afin de prendre le temps d'informer les acteurs du projet de l'appel d'offre et de construire collectivement la réponse.

S'il n'y a pas ce temps d'échanges, le groupe ne mettra pas en place lui-même ses objectifs, ses fiches de suivi et ses critères d'observation. Le groupe risque alors de reprendre les fiches ou les mesures d'autres projets, telles quelles, sans analyser si elle sont adaptées ou non. J'ai pu observer cela avec des associations qui parlent des « fiches de l'INRA ». Elles ne se sont pas approprié le processus, par manque de recul sur le projet mais aussi par manque de temps de discussion avec les paysans. De telles situations ne s'inscrivent pas dans la durée car elles ne reposent pas sur une construction au sein du collectif.

Il faut une discussion collective dans le groupe pour évaluer si les fiches déjà développées répondent à ces objectifs et les mettre à l'épreuve sur son propre terrain. Les fiches proposées peuvent être vues comme une clé donnée au groupe pour entrer dans le champ de la sélection participative, à retravailler ensuite avec l'expérience. Les paysans ne doivent pas être de simples usagers de l'innovation, comme c'est le cas dans un programme de sélection classique<sup>25</sup>. Les paysans doivent être acteurs du processus. Évidemment, leurs actions doivent se faire dans la limite de leurs possibilités matérielles. Par exemple, nous encourageons les paysans qui souhaitent évaluer des populations, même s'ils n'ont pas envie de le faire au sein d'un dispositif organisé. Certains veulent juste cultiver un mélange en surface assez importante pour ensuite sélectionner dedans. Même si cela ne rentre pas dans le cadre de la méthode, c'est à dire qu'il n'y a pas de témoin ni de suivi, ce type de démarche peut utilement être reliée à celle du groupe afin de permettre des échanges et éventuellement une évolution des pratiques.

Le travail de description réalisé dans le chapitre 4 a permis de caractériser les actions du groupe. Cette prise de recul est importante sur deux aspects.

Le premier concerne les interactions qui ont eu lieu. Cela permet de mieux cerner les points qui ont posé problème et de trouver des solutions ensemble pour les résoudre.

Le deuxième est que décrire un mode de fonctionnement peut permettre la mise en place d'une certaine « labellisation » du processus. Je mets ici des guillemets car les collectifs ne vont pas tous chercher à labelliser le processus de sélection participative. Cependant, cette idée revient souvent, comme l'idée de marque sur les « semences paysannes », dont il était question à l'assemblée générale du RSP à Colmar en 2011. Au même titre qu'une variété ne pourrait revendiquer un Certificat d'Obtention Végétale (COV) si

elle est notoirement connue, si le processus de co-construction est clairement décrit, il sera plus difficilement détourné de ses vraies valeurs. L'histoire nous montre avec Nature & Progrès que des initiatives citoyennes peuvent être reprises pour être institutionnalisées et perdre de leur contenu (partie 1.5). Sans cet effort de description, qui permettrait de se protéger, certains paysans imaginent une boule bio issue de « sélection participative » dans les rayons des hypermarchés d'ici quelques années<sup>iii</sup> !

Tous ces éléments renvoient à un point primordial : la mise en place d'une charte entre les acteurs. C'est elle qui posera les limites et les attentes du collectif de travail. Elle posera aussi les bases éthiques du collectif, point central dans ce type de projet qui regroupe une grande diversité d'acteurs animés par des valeurs communes<sup>106</sup>.

Une charte permet d'avoir un cadre clair pour de nouveaux acteurs souhaitant intégrer le projet. Cela permettrait de limiter les risques de dérive. Par exemple que des variétés créées dans le cadre du projet sortent du cadre des échanges en quantité limitée entre pairs. Certaines associations du RSP ont mis en place des conventions<sup>iv</sup>, que l'on peut difficilement assimiler à une charte, mais qui structurent indéniablement le projet.

La mise en place d'une telle charte est compliquée, prend du temps. Elle est toujours en cours dans notre groupe. De même, les adhérents du RSP sont en train d'élaborer une définition du terme « semences paysannes ». La diversité des acteurs complique cette tâche<sup>63</sup>.

#### 5.2.2.2 Adapter les outils pour répondre aux divers besoins des acteurs

Une fois les objectifs des acteurs posés. Les outils que nous avons développés peuvent aider à la mise en place du programme de sélection participative. Ces outils ont été pensés pour être flexibles et adaptables à la diversité des situations. Mais il faut veiller à ce que des outils ne figent pas les idées et la représentation que peuvent avoir les acteurs de leurs populations. Ces outils ne doivent pas faire la sélection à la place des paysans, ils apportent des éléments aux paysans pour les guider dans leurs sélection et ne doivent pas être la réponse à toutes les questions ! Un objectif du projet est que les paysans se réapproprient la sélection. Les outils créés ne sont là que pour faciliter les échanges et permettre aux paysans de gagner en expertise.

Si les critères des fiches et les notices peuvent servir de base de réflexion pour les autres

---

iii. Il en a été question lors de la réunion des semis de 2012

iv. Beaucoup d'entre elles sont liées aux contraintes juridiques, et permettent de faire les essais et les échanges de semences sous couvert d'expérimentation de recherche.

groupes travaillant sur le blé ou les céréales à paille, elles ne seront pas mobilisables pour d'autres espèces. Cependant le format des fiches, leur fréquence (une par étape du cycle de la plante), les rappels des notations faites sur chaque fiche, les notices peuvent être pris comme base de travail.

Les dispositifs expérimentaux et les analyses statistiques sont également tout à fait mobilisables sur d'autres espèces et dans d'autres groupes. Le choix des témoins doit être fait par les paysans. Il est crucial d'avoir un point de comparaison commun qui convienne à tout le groupe. Il faut également veiller à ce qu'il y ait assez d'essais avec des témoins répétés pour que les analyses statistiques puissent fonctionner. Néanmoins, une mise en réseau des maisons des semences paysannes peut limiter ce problème, j'y reviens dans la partie 5.2.2.3.

Ce type de dispositif dans lequel l'évaluation des populations est faite en même temps que la multiplication et la sélection est plutôt adapté aux espèces autogames. Pour les espèces allogames, en général, ces deux types d'activités sont séparées et chaque paysan cultive une population dans sa ferme. L'évaluation des populations se fait sur une ou très peu de plateformes<sup>3</sup>.

A cause des pollinisations croisées, les semences issues des populations évaluées ne pourront pas être ré-utilisées si l'on veut contrôler leur origine. Le processus de sélection participative est donc différent, l'espèce déterminant le mode d'organisation<sup>155</sup>. Il faudra donc réfléchir avec les groupes intéressés pour adapter les méthodes développées. Un travail est déjà en cours en collaboration avec Agro Bio Périgord sur ces aspects.

La base de données est assez souple pour prendre en compte n'importe quelle variable, qui sera associée à n'importe quelle méthode de mesure. Si la base semble bien répondre aux contraintes des projets concernant des plantes annuelles, elle n'est pas adaptée aux plantes pérennes. Une évolution de la base vers les plantes pérennes (fourragère, arbres) est en cours dans le cadre d'un partenariat avec Agro Bio Périgord et le projet Sélection Participative d'Espèces Annuelles et Ligneuses adaptées à des pratiques agroécologiques (SPEAL) SPEAL porté par l'ARDEAR de la région Centre et l'INRA d'Orléans.

Dans tous les cas, il est nécessaire de définir le cadre d'utilisation d'un tel outil. En effet, comme nous l'avons vu précédemment, une base de données peut bouleverser les relations sociales au sein des maisons des semences paysannes (partie 4.4.2). Son utilisation peut (doit ?) être mentionnée dans la charte du projet.

Enfin le « dossier retour » peut être modifié et aménagé selon les besoins des groupes. Ces retours aux paysans sont très importants : sans retour de résultats, il est très difficile de motiver les paysans à s'investir dans le projet. Sélectionner des semences dans sa ferme dans un projet de sélection participative est quelque chose de nouveau pour de nombreux paysans. Afin de les accompagner au mieux, il est nécessaire de leur apporter des résultats tous les ans. Autrement, le groupe risque de perdre sa motivation, ce qui est très compréhensible, et le projet risque de périr.

Au delà de ces outils « techniques », une question fait débat au sein des maisons des semences paysannes : le moyen de générer de la variabilité. Deux visions de la relation entre l'homme et la nature se confrontent<sup>63;224</sup>. Obellianne<sup>181</sup> illustre cette divergence au sein de la maison des semences paysannes Graine de Noé. Certains adhérents ne sont pas favorables à une sélection interventionniste, c'est à dire qui repose sur des croisements. D'un autre côté, certains adhérents sont moteur pour tester et apprendre à faire des croisements.

Un projet de sélection participative ne repose pas nécessairement sur la création de populations par croisement. En effet, il existe d'autres manières de créer de la diversité. Le Tableau 5.1 présente quelques techniques qui peuvent être utilisées notamment sur le blé tendre. Ces méthodes de création de variabilité pourront changer selon le type de reproduction de l'espèce. Notre méthodologie de sélection participative peut s'adapter à n'importe quel type de diversité, du moment qu'elle est assez large pour valoriser la sélection décentralisée et l'adaptation des populations. Les acteurs sont libres de choisir les approches qui leur conviennent le mieux d'un point de vue éthique et philosophique.

Je me suis également demandé si les marqueurs moléculaires pourraient avoir un apport dans le programme de sélection participative<sup>233</sup>. Pour l'instant, les marqueurs moléculaires ont été utilisés pour étudier la diversité génétique et sa structuration au sein du réseau (chapitre 3 et Thomas<sup>229</sup>). Est-ce que l'apport des marqueurs en sélection serait intéressant ? On pourrait l'imaginer pour sélectionner rapidement des populations avec certains gènes de résistances aux maladies. La sélection génomique pourrait aussi être envisagée, cependant peu de résultats sont disponibles sur le blé en présence d'interactions  $G \times E$ <sup>31</sup>. Il est préférable d'attendre les résultats de la sélection génomique en station avant de vouloir l'appliquer dans la sélection décentralisée dans les fermes.

Quoi qu'il en soit, cette question devrait faire l'objet de discussions avec les acteurs du groupe. Les outils moléculaires sont typiquement des outils qui ne sont pas à la portée des paysans. Une approche reposant sur la sélection génomique semble contraire au souhait des paysans de baser la sélection sur leur observation des plantes dans leur milieu. De plus, les maisons des semences paysannes ne pourront pas, a priori, être autonome sur ces aspects qui sont très techniques et coûteux.

Tableau 5.1 - Méthodes permettant de créer de la diversité chez le blé tendre.

Tableau - 5.1 (partie 1/2)

Technique	Description
Mélange	C'est le fait de mélanger des populations où les individus pourront se croiser entre eux et ainsi créer de nouveaux individus. Le processus de croisement et de brassage entre les individus est long pour le blé qui est autogame, avec un taux de croisements de 5%. Ces croisements ne peuvent pas être dirigés ni orientés. Cette technique est très facile à mettre en place dans les fermes et est une pratique courante des paysans du RSP.
Croisements bi-parentaux	Cette technique permet de créer de nouvelles populations par des croisements entre deux individus différents. Il faut castrer une plante qui deviendra femelle et féconder cette plante avec du pollen issu d'une autre plante qui sera le mâle. Les croisements sont dirigés dans le but d'associer des caractéristiques complémentaires entre des populations parents. Cette opération est délicate et prend du temps. Il faut veiller à ce qu'il n'y ait pas d'autofécondation. Les graines issues du croisement sont peu nombreuses et il faut plusieurs multiplications avant d'avoir des semences en quantité suffisante. Les croisements sont réalisables par les paysans dans leur ferme.



Tableau - 5.1 (partie 2/2)

Technique	Description
Composite Cross Population (CCP)	C'est le fait de créer une population à partir de croisements composites, c'est à dire plusieurs croisements entre plusieurs parents <sup>78</sup> . Créer une CCP, c'est à dire réaliser les croisements au départ, est très long. Ces populations sont encore plus diversifiées que les croisements entre deux parents.
Brassage avec un gène de stérilité mâle	C'est le fait d'introduire un gène de stérilité mâle dans une population afin d'augmenter le taux de recombinaisons efficaces <sup>192</sup> . Chez le blé, la stérilité est apportée par un variant présent dans la variété : <i>Probus</i> . La mise en place au départ est un peu lourde et pour conserver 100% d'allofécondation, il faut repérer les mâles-stériles, ce qui est très contraignant.
Blé synthétique	Le blé tendre est peu diversifié sur son génome D (issu de <i>Aegilops squarrosa</i> ). Pour augmenter cette diversité, des laboratoires ont créé des blés synthétique en assemblant le génome D avec le génome du blé dur AB. Cela permet de créer un nouveau blé tendre ABD <sup>74</sup> . Cette opération est technique et demande un matériel de laboratoire adapté. Il semble très difficile de faire une telle manipulation dans les fermes.
Mutagenèse	Cette technique permet d'augmenter le taux de mutation au niveau de l'ADN par l'intermédiaire d'agents extérieurs mutagènes comme les rayons UV par exemple. La mutagenèse est considéré par les acteurs du RSP comme un processus aboutissant à la création d'un OGM, ce qui renvoi à des réflexions plus larges liées à un possible impact environnemental, social, sur la santé, que l'on ne traite pas ici.

### 5.2.2.3 Maintenir un réseau de maisons des semences paysannes

Chaque maison des semences paysannes va participer à la gestion des ressources génétiques par la conservation et la sélection pour, *in-fine*, que chaque ferme produise une nourriture saine. Nous avons identifié trois acteurs principaux au sein des maisons des semences paysannes : les paysans, les jardiniers et les artisans semenciers (Figure 5.2). L'approche que nous décrivons s'ajoute à une diversité existante de moyens de gérer et valoriser la diversité<sup>51</sup>.

Chaque acteur au sein des maisons des semences paysannes peut avoir un rôle diffé-

rent selon ses envies et ses besoins. Alors que certains vont être plus concernés par les activités de la conservation, c'est à dire le maintien d'une variété présentant des critères qui permettent de la reconnaître, d'autres vont plutôt s'orienter vers la sélection et la création de nouvelles variétés adaptées à leurs besoins (Figure 5.2). Bien sûr, conservation et sélection sont très liées. La sélection à la ferme participe à la gestion dynamique des ressources génétiques (chapitre 3).

Il est recommandé qu'il n'y ait pas un lieu unique qui regroupe toutes les variétés. On peut imaginer que des variétés soient distribuées entre les acteurs des maisons des semences paysannes. Cela limite les risques de perte, mais aussi que certains acteurs s'essouffent à être responsables de trop de variétés. Il est important que les acteurs aient des moyens de stockage qui préservent la qualité des semences et que les variétés soient au moins sur deux lieux pour limiter les risques de pertes. Les variétés conservées, qui sont un puits de diversité prêt à être mobilisé en sélection, peuvent être multipliées par exemple tous les deux ou trois ans<sup>v</sup>, afin de ne pas trop limiter leur pouvoir germinatif et leur permettre de s'adapter aux environnements où elles sont multipliées sans trop surcharger leurs « mainteneurs ».

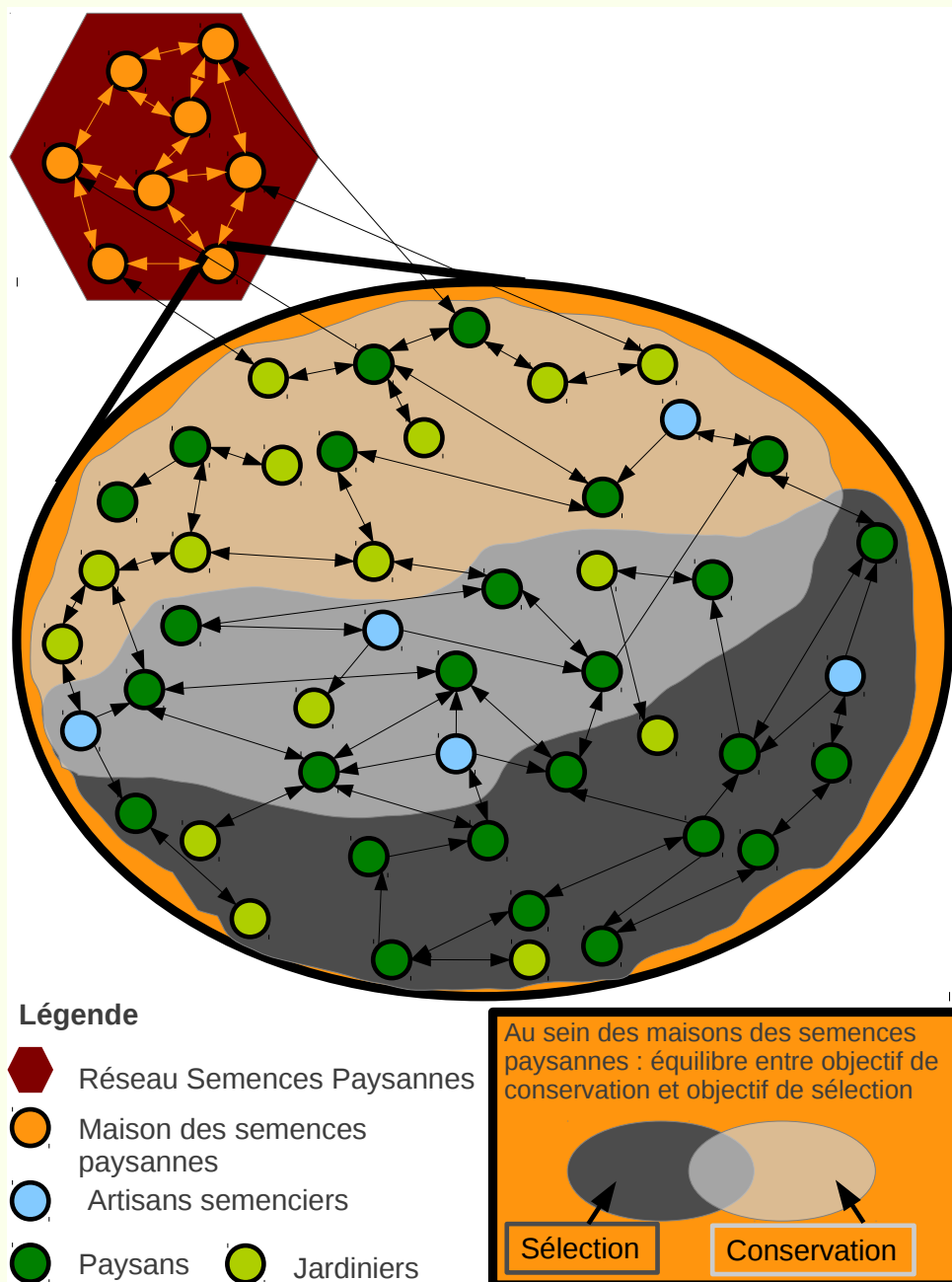
A un moment de sa sélection, un paysan peut considérer que la population a atteint un « optimum » et qu'il faille la conserver. Le moment où l'on considère qu'une variété doit être conservée n'est pas facile à évaluer ... car tout est en perpétuel mouvement.

Il faut peut être s'affranchir de cette idée de conservation mais aller plus vers une idée de gestion dynamique, c'est à dire évolutive, qui n'est pas figée dans le temps. Cela se rapproche d'une vision darwinienne de l'agriculture<sup>22</sup>. Cette vision s'oppose à la vision patrimoniale de la variété de pays qui a fédéré les paysans du RSP à sa création<sup>63</sup>. Néanmoins, il semble important de conserver des populations avec des caractéristiques reconnues sur leur comportement agronomique ou nutritionnel par exemple. On peut imaginer que les jardiniers se chargent de la conservation quand les paysans se concentrent sur la sélection. Dans certaines maisons des semences paysannes comme Kerna ùn Shoma, les jardiniers jouent un rôle dans la conservation<sup>13</sup>. Bien sûr ces frontières ne sont pas étanches. Le processus est dynamique et ne s'arrête jamais.

---

v. Ce sont des pratiques courantes chez les paysans du RSP

Figure 5.2 - Les maisons des semences paysannes en réseau, actrices de la gestion dynamique des ressources génétiques : conservation et sélection.



Chaque acteur va avoir des objectifs différents : conservation ou sélection, qui participent au maintien de la diversité génétique. Certains acteurs peuvent avoir ces deux objectifs. La répartition des acteurs entre conservation et sélection est hypothétique mais se veut représentative d'une diversité de pratiques.

Les échanges de semences, savoirs et savoir-faire sont représentés par les flèches. Celles-ci peuvent être à sens unique ou à double sens au sein d'une maison des semences paysannes et aussi au delà.

Il est important que chaque maison des semences paysannes puisse partager et s'inspirer des travaux des autres maison des semences paysannes à travers un réseau plus large, il peut s'agir du RSP. Le réseau comporte une grande diversité d'organisations sociales. Cela le rend plus stable et résilient face à d'éventuelles perturbations environnementales, liées aux stress biotiques et abiotiques, et aussi politiques, si une maisons des semences paysannes se dissout et que les variétés sont perdues par exemple <sup>vi</sup>.

Sans une mise en réseau efficace, le développement de programme de sélection participative sera beaucoup moins performant. C'est la communication qui permet de huiler les rouages de ce réseau. Bien sûr les paysans communiquent entre eux en dehors des maisons des semences paysannes, mais ces dernières se veulent le catalyseur des échanges et de l'organisation de la gestion et de la sélection des semences. La mise en réseau des maisons des semences paysannes permet de favoriser et d'encourager les échanges d'informations, d'outils, de résultats, de savoirs, de savoir-faire et de semences.

Concernant les informations, il peut s'agir d'appel d'offre, de formations, de rapports, d'articles spécialisés, etc. Ces échanges d'information se font par mails (il existe des listes de diffusions au sein du RSP) ou lors de rencontres. Par exemple, lors de réponse à des appels d'offre, il est souhaitable que chaque maison des semences paysannes soit au courant et dise, de manière transparente, si elle souhaite répondre ou non. Ceci afin de ne pas surcharger les réponses et risquer de se faire de la concurrence.

Les outils regroupent la base de données, les fiches, le « dossier retour », les dispositifs expérimentaux et les statistiques. Beaucoup d'associations ont déjà mis en place des fiches de suivi et ont un certain recul sur leur fonctionnement <sup>13;76;181</sup>. Ce sont autant d'exemples mobilisables pour d'autres groupes. Ces outils permettent de générer des résultats.

Les résultats peuvent être décuplés si les maisons des semences paysannes les mettent en commun. Dans le cadre de notre méthode, si les témoins ne sont pas les mêmes entre les maison des semences paysannes, une partie des analyses statistiques (la comparaison des moyennes intra-ferme) est toujours possible. On peut faire l'hypothèse que les variances des témoins sont issues d'une même loi. Cependant, il y aura beaucoup plus de potentiel pour les analyses si les témoins sont communs.

Les échanges de semences peuvent se faire à la suite de la réunion des semis sur la base des « dossiers retours » qui sont donnés aux paysans. Des échanges peuvent également avoir lieu lors de rencontres des semis ou de manière informelle en dehors des regrou-

---

vi. On ne parle pas de risque de guerre, mais dans un tel cas, il vaut mieux ne pas avoir toutes les ressources au même endroit ...

pement organisés. Ces échanges de semences sont souvent liés à des échanges de savoirs et de savoir-faire<sup>66</sup>. Si un intermédiaire se glisse entre les paysans, les savoirs associés à ces semences risquent de circuler moins bien ou plus du tout. Cela pose la question suivante : Mise en réseau, oui, mais jusqu'à quel point ? Est-ce que les acteurs doivent tous se connaître pour partager leurs observations, leurs résultats, leurs méthodes ? Est-ce que la taille du réseau doit rester « humaine » pour que la gestion des semences ait un sens ? Les réseaux d'échanges au sein du RSP reposent sur la confiance entre les acteurs et l'échange de petites quantités de semences<sup>23</sup>. Est-ce que la mise en réseau plus large changerait les relations entre les acteurs ? Autant de questions difficiles dont il convient de discuter avec les acteurs.

Il n'est pas possible que les paysans fassent la sélection sur toutes les espèces présentes dans leurs fermes. Par contre il est envisageable que les acteurs du réseau puissent se partager la sélection et la conservation pour différentes espèces.

Nous n'avons pas mentionné les espèces potagères. La sélection pour les espèces potagères est plus compliquée. En effet, la plupart des espèces sont allogames, ce qui demande une gestion très rigoureuse pour éviter qu'il n'y ait de croisement entre les variétés. De plus, les conditions de productions de semences sont plus contraignantes pour les légumes de type racine ou feuille. Ainsi, pour un maraîcher cultivant de nombreuses espèces, il est difficile de concevoir qu'il puisse gérer en parallèle la sélection pour chacune d'elles.

Enfin, parmi les missions des maisons des semences paysannes, il y a les aspects liés à la communication et la sensibilisation<sup>155</sup>. Ce point est très important car il permettra de fédérer de nouveaux acteurs. Plus il y aura d'acteurs dans le réseau, plus il y aura de diversité. Diversité d'idées, de savoirs, de savoir-faire, de semences ... C'est cette diversité qui est source d'innovation. Et c'est le réseau qui permettra à cette innovation de diffuser, de s'adapter, d'évoluer. C'est un cercle vertueux.

Sur les questions liées au financement, il pourrait être envisageable que les consommateurs soutiennent ce type de projet en cotisant aux maisons des semences paysannes. De même, les paysans qui adhèrent aux maisons des semences paysannes pourraient reverser une partie de la vente de leur produit. Dans ce cas, les questions de « labellisation » peuvent être mises en avant pour valoriser une hausse des prix. A titre d'exemple, nous pourrions compter 5 centimes par kilo de farine. Il serait possible de dégager 25 000 euros (l'équivalent d'un animateur à temps partiel) avec 500 tonnes de farine. Dans un réseau

de 25 paysans, cela revient à 20 tonnes par paysan et par an. En comptant un rendement (minimal) de 2,5 tonnes de farine par hectare, cela correspondrait à 8 hectares.

L'organisation au sein des maisons des semences paysannes regroupées en réseau se heurte à plusieurs contraintes qui l'empêche de pouvoir se développer davantage. La dernière partie de ce chapitre fait des propositions pour que ce mode d'organisation puisse se développer plus largement. Ces propositions visent à accompagner la transition vers un système semencier adapté à l'agroécologie.

### 5.3 Propositions au niveau national pour accompagner la transition

Le modèle de sélection et de gestion des ressources génétiques pour l'agroécologie que nous proposons fonctionne à contre courant du système actuel. Il contourne les normes du système en place. C'est parce que nous nous sommes affranchi des contraintes du système, telles que la nécessité que les variétés soient homogènes génétiquement ou la sélection en station, que le collectif d'acteurs a pu innover au niveau organisationnel mais aussi génétique.

La meilleure connaissance du fonctionnement du programme de sélection participative, et les propositions qui en découlent pour les maisons des semences paysannes, sont autant d'éléments qui permettent de participer à la transition vers un autre système semencier adapté à l'agroécologie et, au final, vers des systèmes alimentaires durables.

Cependant, une telle transition n'est possible que si le système en place n'est pas bloqué par un ensemble de verrous socio-économiques<sup>17</sup>. Or, dans le cas du système semencier français, où un paradigme et un système institutionnalisé dominant, de tels verrouillages peuvent apparaître (Vanloqueren et Baret<sup>245</sup> illustrent ce point avec les variétés de blé multi-résistantes). Une analyse fine de ces verrouillages permettrait de proposer des politiques adaptées pour faire évoluer le système vers une différenciation des systèmes semenciers selon les acteurs et les marchés. En effet, il nous semble qu'il peut y avoir co-existence entre un système avec sélection centralisée, inscription au catalogue, protection Certificat d'Obtention Végétale (COV), commercialisation pour les grands marchés et un système avec sélection participative, variétés populations évolutives non inscrites, échanges entre pairs de quantité limitée.

Dans cette dernière, nous avons identifié trois points principaux qui permettraient à notre modèle de se développer : former les animateurs, créer un espace juridique pour

les semences paysannes et allouer des moyens financiers pour les maisons des semences paysannes. Le développement de l'agroécologie nécessite des politiques adaptées<sup>60</sup>. Ces points sont autant de propositions pour accompagner cette transition qui, pour être réalisée, devra être portée au niveau politique.

### 5.3.1 Former les animateurs

La dynamique au sein du réseau est très liée aux animateurs. Ils doivent être sur tous les fronts, mobiliser de nombreux concepts des sciences agronomiques aux sciences juridiques, organiser les rencontres et être disponibles pour les paysans et les chercheurs. Cette polyvalence en fait un nouveau métier clé dans la sélection et la gestion des ressources génétiques pour l'agroécologie. Leur comportement n'est pas neutre et il va influencer le fonctionnement du réseau. Il est donc primordial qu'ils soient formés pour réaliser cette tâche ardue.

La formation peut être de deux ordres : la formation sur les principes de la sélection participative et la formation à partir des échanges d'expériences de chacun des animateurs.

La sélection participative est interdisciplinaire. Le but d'une formation sur ses méthodes n'est pas que les animateurs deviennent des experts, mais qu'ils acquièrent des outils et des connaissances suffisantes qui leur permettent d'interagir avec d'autres acteurs tels que des paysans, des chercheurs, d'autres animateurs. Ces formations peuvent porter, d'un point de vue théorique et pratique, sur les méthodes de co-construction, la juridiction des semences (droit de propriété intellectuelle, accès aux ressources génétiques, commercialisation), la gestion des données (d'un point de vue technique et aussi déontologique), les analyses statistiques (théorie et utilisation de logiciels), les bases de génétique quantitatives et des populations, les principes de l'agroécologie, les concepts de sociologie des organisations, la gestion et l'organisation de rapports bibliographiques.

Au delà de ces aspects, les échanges d'expériences entre animateurs peuvent être très utiles. Cela permet de prendre le temps de se poser<sup>vii</sup> et de réfléchir à ses pratiques et celles des autres. Cela représente aussi un temps où les animateurs peuvent échanger sur les problèmes qu'ils ont rencontrés, comment ils les ont résolus, etc.

Ce type de formation pourrait être suivie dès l'arrivée de nouveaux animateurs. En effet, les mouvements d'animateurs sont très fréquents dans les associations et commencer avec de bonnes bases ne peut qu'être un plus. Ces nombreux mouvements sont représentatifs de la difficulté de ce métier, à l'interface entre les acteurs des maisons des semences

---

vii. Chose qui n'arrive jamais, au vu des échanges que j'ai eus avec eux pendant ma thèse.

paysannes et des champs disciplinaires très différents.

On pourrait aussi imaginer mettre en place un master « sélection participative et agroécologie » ainsi que des journées d'échanges avec les citoyens en général, afin de rapprocher d'un côté paysans et citoyens et de l'autre chercheurs et citoyens.

### 5.3.2 Créer un espace juridique pour les variétés issues du programme de sélection participative

Nous avons vu dans la partie 1.4 les trois aspects liés à la juridiction des semences : la propriété intellectuelle, l'accès aux ressources et la commercialisation. Ces aspects juridiques renvoient à la gouvernance des semences, qui est au coeur des enjeux de l'agro-biodiversité<sup>211</sup>.

Actuellement, il n'existe aucun cadre juridique qui prenne en compte ces nouvelles variétés paysannes, issues de la co-construction au sein d'un collectif d'acteurs. Pour plus de clarté dans la suite du texte, nous les appelons « variétés paysannes ».

Pour que ces « variétés paysannes » soient reconnues, il est nécessaire de créer un espace juridique qui permette de définir un droit de propriété intellectuelle (ou un système équivalent), de rendre ces variétés accessibles librement pour de futures sélections et de créer un cadre de commercialisation.

Pour chacun de ces points, nous faisons des propositions issues de la littérature, des résultats de notre projet, des notes prises lors de conférences de G. Kastler (délégué général du RSP), des discussions avec A-C. Möy, (ancienne animatrice de la veille juridique du RSP), et enfin des discussions issues de la réunion des semis de 2012.

Bien que cela valorise les « variétés paysannes », on ne mentionnera pas la protection et la commercialisation des produits issus de ces semences. Cela recoupe d'autres domaines juridiques liés aux appellations d'origines contrôlées (AOC), aux indications géographiques protégées (IGP) et aux marques<sup>51</sup>.

#### Les droits de propriété intellectuelle

Le système du COV n'est pas applicable dans le cas des « variétés paysannes »<sup>172</sup>. En effet, le COV repose sur les caractères de distinction, d'homogénéité et de stabilité. Ces critères ne correspondent pas aux « variétés paysannes » qui sont des populations hétérogènes.

Le droit de propriété intellectuelle peut être envisagé comme un moyen de se protéger contre le bio-piratage, plutôt qu'un moyen de rémunérer le travail de sélection<sup>172</sup>. Le critère



de nouveauté, également nécessaire pour le dépôt d'un COV, peut aider à se protéger de la bio-piraterie. En effet, aucun droit de propriété intellectuelle n'est possible dès que la variété est « *notoirement connue* » d'après l'Union Internationale pour la Protection des Obtentions Végétales (UPOV).

Une approche pour déterminer la nouveauté serait de décrire l'histoire des populations. La démarche collective de l'obtention d'une variété doit être identifiée avec le contexte historique et social, en plus de critères classiques liés aux performances agronomiques, nutritionnelles et en rapport avec la biodiversité et les services éco-systémiques. De même que les informations concernant le mode de création (croisement, mélange), les lieux où elle a été cultivée et combien de temps, etc. Ces données pourraient être recensées dans un passeport.

Il est tentant de faire le parallèle entre les « semences paysannes » et les logiciels libres<sup>172;211</sup>. Il pourrait être envisageable de transposer la licence publique générale (General Public Licence (GPL) en anglais) aux semences. A la différence du domaine public, il y a moyen de contrôler et poser les conditions que l'on souhaite sur le devenir des variétés et des ressources génétiques. Cependant, à la différence d'un logiciel, la variété est vivante et donc n'est pas stable. De plus, les paysans ne veulent pas forcément que leurs variétés soient accessibles à tout le monde (à la différence du système GPL où tout le monde peut avoir accès sans restrictions)<sup>172</sup>.

#### **L'accès aux ressources**

Les savoirs paysans et leurs actions à travers la co-construction du programme sont une clé pour générer l'innovation. De fait les droits des paysans doivent être reconnus et affirmés par les législations nationales. C'est ce que demande le Traité International sur les Ressources Phytogénétiques pour l'Alimentation et l'Agriculture (TIRPAA) (partie 1.4), celui-ci a été ratifié par la France mais n'a pas été transféré dans le droit français. Il est primordial que les articles du TIRPAA soient retranscrits dans le droit français afin de créer un espace légal pour échanger, utiliser et produire ses semences.

Cet accès aux ressources pourrait être compliqué par la contamination des « variétés paysannes » par des gènes brevetés<sup>172</sup>. Il est donc également souhaitable que la législation puisse protéger les paysans contre ce type de menace.

#### **La commercialisation**

Des évolutions sur la commercialisation qui permettraient de « *définir un statut réglementaire clair et adapté pour les variétés « populations » ou issues de la sélection*

*participative* » sont encouragées par le commissariat général à la stratégie et à la prospective<sup>75</sup>. Mais actuellement, le travail de sélection participative se fait sous le couvert de l'expérimentation pour la recherche et les variétés qui en sont issues ne peuvent pas être commercialisées. Cela représente un frein pour que ce modèle puisse se développer. En effet, l'absence de commercialisation signifie l'absence d'échanges, qui sont à la base de notre méthodologie, que cela soit pour le développement des variétés (chapitre 2) ou la gestion de la biodiversité cultivée (chapitre 3 et Thomas<sup>229</sup>).

Pour être commercialisées, les variétés doivent être inscrites au catalogue et donc DHS et VATE. Au même titre que pour le COV, les « variétés paysannes » ne peuvent pas être DHS. De plus, il y a peu de chances qu'elles surpassent les témoins dans les réseaux d'évaluation du CTPS. En effet, ces « variétés paysannes » ont été développées pour être adaptées localement à des environnements spécifiques. Les évaluer en dehors de leur environnement n'a pas de sens, sauf si ce dernier est proche de l'environnement de sélection. Dans la suite du projet, nous souhaiterions évaluer quelques « variétés paysannes » issues du programme de sélection participative dans différentes fermes. Cela apportera des éléments sur l'évaluation de telles variétés, qui devra très certainement prendre en compte les interactions  $G \times E$ .

La contrainte de l'homogénéité pourrait être remplacée par des variances et des moyennes phénotypiques en fonction de quelques témoins. On pourrait imaginer considérer qu'une population n'est plus la même si elle s'écarte de deux écarts-types de la variance décrite sur un catalogue<sup>56</sup>.

Le catalogue de conservation n'est pas non plus satisfaisant. Il est trop restrictif avec des quantités échangeables trop petites, un niveau d'hétérogénéité trop faible (taux de hors type de 10%). Ensuite, l'impossibilité pour la variété d'être diffusée hors de sa zone géographique est une contrainte.

Pour le contrôle qualité des semences, on peut imaginer un système de contrôle participatif, comme cela est fait avec Nature & Progrès. Ce type de contrôle se base sur la participation active des acteurs concernés. Il est construit sur la confiance et un réseau d'échanges de connaissances (d'après l'IFOAM) .

### 5.3.3 Allouer des moyens financiers pour les maisons des semences paysannes

Mettre en place des formations et créer un espace juridique adapté sont deux éléments indispensables pour développer notre modèle. Mais pour que celui-ci puisse s'installer et soit consolidé dans la durée, il faut des moyens financiers.

Nous avons vu que les maisons des semences paysannes pourraient s'autofinancer si les consommateurs et les paysans cotisaient à un fond dédié. Cependant, afin de permettre une réelle transition, il faut de plus grands moyens. Le financement de ce modèle revient, selon moi, à l'Union Européenne par l'intermédiaire de la Politique Agricole Commune (PAC). En effet, notre modèle prône le développement de variétés populations et la gestion des ressources génétiques pour l'agroécologie. Cette dernière confère des avantages considérables pour la société, que cela soit dans la production de nourriture saine, dans le maintien des services écosystémiques et la réduction d'usage d'intrant (partie 1.1.3).

On pourrait aussi imaginer de donner une mission officielle d'animation et d'accompagnement de la sélection participative aux instituts tel que l'ITAB ou les chambres d'agriculture.

\*\*\*

Le projet de sélection participative constitue une niche d'innovation<sup>viii</sup>. Développer cette niche permettra la construction et la consolidation de systèmes diversifiés de sélection variétale et de gestion des ressources génétiques pour l'agroécologie. Cela passe entre autres par la formation des animateurs, la création d'un espace juridique pour les semences paysannes et l'allocation de budget pour les maisons des semences paysannes.

De plus, il sera important d'inciter le régime sociotechnique standard en place, dans notre cas le système semencier français, à ouvrir des opportunités pour que ces approches nouvelles puissent se développer. Elles dépasseront alors leur statut de niche et pourront compléter le modèle standard, voire s'hybrider avec lui et ainsi le faire évoluer.

---

viii. La réflexion qui suit est largement inspirée de Meynard et al.<sup>165</sup> qui travaillaient sur les freins et leviers à la diversification des cultures.





## Conclusion

Cette thèse s'est insérée dans une recherche action qui visait à répondre à deux objectifs. Le premier était de développer des méthodes et des outils opérationnels pour la gestion de la biodiversité cultivée et la sélection à la ferme du blé tendre dans des agrosystèmes basés sur les principes de l'agroécologie. Le deuxième était de renforcer l'apprentissage et l'autonomie des paysans en matière de gestion et de sélection des semences.

Ces objectifs ont été en partie atteints avec la création de méthodes et d'outils : fiches de suivi de la culture, base de données, méthodes statistiques, dossier de retour des résultats pour chaque paysan, d'un livret technique. De plus, les paysans ont renforcé leur apprentissage et leur autonomie car ils étaient acteurs du programme et non simples utilisateurs de variétés, comme en témoigne leur implication croissante à faire de la sélection massale et à créer de nouveaux croisements.

Une analyse moléculaire et phénotypique a montré que le programme de sélection participative s'insère dans la gestion *in-situ* des ressources génétiques par la création de nouvelles variétés-populations et par leur gestion dans un large réseau de fermes.

Ce projet a permis deux innovations : organisationnelle et génétique. L'innovation organisationnelle correspond au mode de fonctionnement du projet, basé sur la décentralisation de la sélection, la co-construction entre les acteurs et leur mise en réseau. Les innovations génétiques découlent de ce mode d'organisation : une large diversité a été créée, est maintenue et évaluée dans les fermes et de nouvelles variétés populations sont utilisées par des paysans dans leurs fermes.

Atteindre de tels résultats a été possible grâce à la co-construction entre les acteurs du

projet : paysans, associations et équipe de recherche. Cette construction est au coeur de la sélection participative. La rencontre entre savoirs locaux et connaissances scientifiques a été source d'innovation.

Le changement de paradigme nécessaire en agroécologie se retrouve dans l'évolution de notre posture de recherche. Ce travail repose sur une collaboration interdisciplinaire associant statisticiens, généticiens quantitativistes et des populations, bio-informaticiens et chercheurs en science humaine et sociale. L'interdisciplinarité et la co-construction ont donné un sens à l'action de notre recherche.

La méthodologie que nous avons développée est évolutive, flexible et adaptable. Elle met les paysans aux coeur de la sélection et de la gestion des ressources génétiques. Ce projet participe à la transition vers un nouveau système semencier adapté à la diversité des systèmes agroécologiques. Cette transition ne pourra se faire à une plus large échelle qu'à travers une évolution de la législation actuellement en place et l'identification d'un mode de financement pérenne. Les semences sont le premier maillon de la production alimentaire. Mettre les paysans au coeur de la création variétale et de la gestion des ressources génétiques est une étape nécessaire pour atteindre la souveraineté alimentaire.





## Bibliographie

- [1] Agence-Bio, 2013. Une agriculture sans produits chimiques de synthèse ni OGM. <http://www.agencebio.org/une-agriculture-sans-produits-chimiques-de-synthese-ni-ogm>, 5 octobre 2013.
- [2] Agrawal A., 1995. Dismantling the divide between indigenous and scientific knowledge. *Development and Change* 26(3), 413–439.
- [3] AgroBioPérigord, Bio d'Aquitaine, 2013. Expérimentations en variétés populations. Rapport.
- [4] Akhunov E.D., Akhunova A.R., Anderson O.D., Anderson J.A., Blake N., Clegg M.T., Coleman-Derr D., Conley E.J., Crossman C.C., Deal K.R., Dubcovsky J., Gill B.S., Gu Y.Q., Hadam J., Heo H., Huo N., Lazo G.R., Luo M.C., Ma Y.Q., Matthews D.E., McGuire P.E., Morrell P.L., Qualset C.O., Renfro J., Tabanao D., Talbert L.E., Tian C., Toleno D.M., Warburton M.L., You F.M., Zhang W., Dvorak J., 2010. Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. *Bmc Genomics* 11.
- [5] Almekinders C., Centeno J., Torrez R., Olivera L., Suarez J., Carrasco J., 2006. Experiencias y aprendizajes del desarrollo de variedades de frijol de manera participativa en el norte de nicaragua. *Agronomía Mesoamericana* 17(3), 327.
- [6] Almekinders C., Elings A., 2001. Collaboration of farmers and breeders : Participatory crop improvement in perspective. *Euphytica* 122(3), 425–438.
- [7] Altieri M., 1989. Agroecology : A new research and development paradigm for world agriculture. *Agriculture, Ecosystems & Environment* 27(1-4), 37–46.
- [8] Altieri M.A., 2002. Agroecology : the science of natural resource management for poor farmers in marginal environments. *Agriculture, Ecosystems & Environment* 93(1–3), 1–24.
- [9] Altieri M.A., Merrick L.C., 1987. In situ conservation of crop genetic resources through maintenance of traditional farming systems. *Economic Botany* 41, 86–96.

- [10] AMAP, 2013. Annuaire national des AMAP - associations pour le maintien d'une agriculture paysanne. <http://www.reseau-amap.org/>, 28 septembre 2013.
- [11] Annicchiarico P., 2002. Genotype x environment interactions. Challenges and opportunities for plant breeding and cultivar recommendations. FAO, Rome, Italie.
- [12] Annicchiarico P., Chiapparino E., Perenzin M., 2010. Response of common wheat varieties to organic and conventional production systems across italian locations, and implications for selection. *Field Crops Research* 116(3), 230–238.
- [13] Arenou D., 2013. Création d'une maison des semences en alsace, pour une gestion dynamique locale de la biodiversité cultivée : fonctionnement et articulation entre les acteurs en valorisant les conservatoires de blés anciens. Rapport, Ecole Supérieure d'Agriculture, Angers.
- [14] ARVALIS, 2013. Stratégies de fertilisation du blé tendre : fractionner l'azote en trois apports. <http://www.arvalis-infos.fr:80/view-12293-arvarticle.html>, 5 octobre 2013.
- [15] Atlin G.N., Cooper M., Bjornstad A., . A comparison of formal and participatory breeding approaches using selection theory. *Euphytica* 122(3).
- [16] Balvanera P., Pfisterer A.B., Buchmann N., He J.S., Nakashizuka T., Raffaelli D., Schmid B., 2006. Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters* 9(10), 1146–1156.
- [17] Baret P., Stassart P., Vanloqueren G., Van Damme J., 2013. Dépasser les verrouillages de régimes socio-techniques des systèmes alimentaires pour construire une transition agroécologique : Thème 2 Alimentation, agriculture, élevage. Université catholique de louvain, Namur.
- [18] Baron C., 2003. La gouvernance : débats autour d'un concept polysémique. *Droit et Société* (54), 329–351.
- [19] Beales J., Turner A., GriYths S., Snape J.W., Laurie D.A., 2007. A pseudo-response regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (*triticum aestivum* l.). *Theoretical and Applied Genetics* 115(5), 721–733.
- [20] Bengtsson J., Ahnström J., Weibull A.C., 2005. The effects of organic agriculture on biodiversity and abundance : a meta-analysis. *Journal of Applied Ecology* 42(2), 261–269.
- [21] Bommelear O., Devaux J., 2011. Coûts des principales pollutions agricoles de l'eau. Rapport, Ministère de l'Ecologie du Développement durable des Transports et du Logement.
- [22] Bonneuil C., 2009. Savoirs, pouvoirs et imaginaires de la gestion du vivant de darwin à aujourd'hui. In : *Sciences en campagne : regards croisés, passés et à venir : colloque de Cerisy*, B. Hervieu et B. Hubert. L'Aube Edition., pp. 77–98.
- [23] Bonneuil C., Demeulenaere E., 2007. Vers une génétique de pair à pair ? L'émergence de la sélection participative. In : *Des sciences citoyennes ? La question de l'amateur dans les sciences naturalistes*, F Charvolin, A Micoud, L K Nyhart. L'Aube Edition., pp. 122–147.



- [24] Bonneuil C., Demeulenaere E., Thomas F., Joly P., Allaire G., Goldringer I., 2006. Innover autrement ? La recherche face à l'avènement d'un nouveau régime de production et de régulation des savoirs en génétique végétale. Dossier de l'environnement de l'INRA (30), 29–51.
- [25] Bonneuil C., Fenzi M., 2011. Des ressources génétiques à la biodiversité cultivée. Revue d'anthropologie des connaissances Vol. 5, n° 2(2), 206–233.
- [26] Bonneuil C., Goffaux R., Bonnin I., Montalent P., Hamon C., Balfourier F., Goldringer I., 2012. A new integrative indicator to assess crop genetic diversity. Ecological Indicators 23, 280–289.
- [27] Bonneuil C., Thomas F., 2009. Gènes, pouvoirs et profits : Recherche publique et régimes de production des savoirs de Mendel aux OGM. Quae.
- [28] Bonnin I., Rousset M., Madur D., Sourdille P., Dupuits L., Brunel D., Goldringer I., 2008. FT genome a and d polymorphisms are associated with the variation of earliness components in hexaploid wheat. Theoretical and Applied Genetics 116(3), 383–394.
- [29] Brac de la Perrière R.A., de Kochko P., Neubauer C., Storup B., 2011. Visions paysannes de la recherche dans le contexte de la sélection participative : comment co-construire et mutualiser les connaissances sur les plantes ? Pour l'Emergence d'une Université du Vivant, Donzy le National, France.
- [30] Brunel S., Action contre la faim, 2000. Geopolitique de la faim. Presses universitaires de France, Paris.
- [31] Burgueño J., de los Campos G., Weigel K., Crossa J., 2012. Genomic prediction of breeding values when modeling genotype  $\times$  environment interaction using pedigree and dense molecular markers. Crop Science 52(2), 707.
- [32] Buttel F., 2003. Envisioning the future development of farming in USA : agroecology between extinction and multifunctionality ? New Directions in Agroecology Research and Education ,Wisconsin, University.
- [33] Bänziger M., Cooper M., 2001. Breeding for low-input conditions and consequences for participatory plant breeding : Examples from tropical maize and wheat. Euphytica 122(3), 503–519.
- [34] Camacho Villa T.C.C., Maxted N., Scholten M., Ford-Lloyd B., 2005. Defining and identifying crop landraces. Plant Genetic Resources 3(03), 373–384.
- [35] Casabianca F., Albaladejo C., 1997. Des multiples légitimités de la recherche-action. Etudes et Recherches sur les systèmes Agraires et le Développement (30), 11–25.
- [36] Cavanagh C.R., Chao S., Wang S., Huang B.E., Stephen S., Kiani S., Forrest K., Saintenac C., Brown-Guedira G.L., Akhunova A., See D., Bai G., Pumphrey M., Tomar L., Wong D., Kong S., Reynolds M., da Silva M.L., Bockelman H., Talbert L., Anderson J.A., Dreisigacker S., Baenziger S., Carter A., Korzun V., Morrell P.L., Dubcovsky J., Morell M.K., Sorrells M.E., Hayden M.J., Akhunov E., 2013. Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. Proceedings of the National Academy of Sciences of the United States of America 110(20), 8057–8062.

- [37] Ceccarelli S., 1989. Wide adaptation : How wide ? *Euphytica* 40(3), 197–205.
- [38] Ceccarelli S., 2012. *Plant Breeding with farmers : a technical manual*. ICARDA, PO Box 5466, Aleppo, Syria.
- [39] Ceccarelli S., Grando S., 2007. Decentralized-participatory plant breeding : an example of demand driven research. *Euphytica* 155(3), 349–360.
- [40] Ceccarelli S., Grando S., 2009. *Participatory Plant Breeding in Cereals : Handbook of Plant Breeding*, vol. 3. Springer New York.
- [41] Ceccarelli S., Grando S., Bailey E., Amri A., El-Felah M., Nassif F., Rezgui S., Yahyaoui A., 2001. Farmer participation in barley breeding in syria, morocco and tunisia. *Euphytica* 122(3), 521–536.
- [42] Ceccarelli S., Grando S., Singh M., Michael M., Shikho A., Al Issa M., Al Saleh A., Kaleonjy G., Al Ghanem S., Al Hasan A., Dalla H., Basha S., Basha T., 2003. A methodological study on participatory barley breeding. II. Response to selection. *Euphytica* 133(2), 185–200.
- [43] Ceccarelli S., Grando S., Tutwiler R., Baha J., Martini A., Salahieh H., Goodchild A., Michael M., 2000. A methodological study on participatory barley breeding i. selection phase. *Euphytica* 111(2), 91–104.
- [44] CerealsDB, 2013. Wheat SNP database. [http://www.cerealsdb.uk.net/CerealsDB/SNPs/Documents/D0C\\_snps.php](http://www.cerealsdb.uk.net/CerealsDB/SNPs/Documents/D0C_snps.php), 1 octobre 2013.
- [45] Chable V., Conseil M., Serpolay E., Le Lagadec F., 2008. Organic varieties for cauliflowers and cabbages in Brittany : from genetic resources to participatory plant breeding. *Euphytica* 164, 521–529.
- [46] Chao S., Dubcovsky J., Dvorak J., Luo M.C., Baenziger S.P., Matnyazov R., Clark D.R., Talbert L.E., Anderson J.A., Dreisigacker S., Glover K., Chen J., Campbell K., Bruckner P.L., Rudd J.C., Haley S., Carver B.F., Perry S., Sorrells M.E., Akhunov E.D., 2010. Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (*triticum aestivum* l.). *Bmc Genomics* 11.
- [47] Chateil C., Goldringer I., Tarallo L., Kerbirou C., Le Viol I., Ponge J.F., Salmon S., Gachet S., Porcher E., 2013. Crop genetic diversity benefits farmland biodiversity in cultivated fields. *Agriculture, Ecosystems & Environment* 171, 25–32.
- [48] Cochard N., 2006. Société civile. *Alternatives Economiques Poche* (22).
- [49] Coffman W., Smith M.E., 1991. Role of public, industry, and international research center breeding programs in developing germplasm for sustainable agriculture. In : *Plant breeding and sustainable agriculture, Considerations for Objectives and Methods*, no., D.A. Sleeper, T.C. Barker, and E.J. Bramel Cox, no. 18 :19. Madison, WI : CSSA special publication Edition.
- [50] Confédération Paysanne, 2012. *Droits de propriété industrielle et droits des agriculteurs dans le cadre de l’UPOV et des réglementations européennes et française sur le COV, le brevet et la commercialisation des semences*. Rapport.

- [51] Couix N., Lauvie A., Charrier F., Hazard L., 2013. Des ressources génétiques mobilisées dans une diversité des formes de valorisation : entre tensions et dynamiques de développement. *Innovations Agronomiques* (29), 99–112.
- [52] CTPS, 2013. Protocole d'Expérimentation céréale à paille, essais de valeur agronomique technologique et environnementale. Rapport.
- [53] Dagnelie P., 1975. *Théorie et méthodes statistiques*, vol. 2. presse agronomique de Gembloux.
- [54] Dalgaard T., Hutchings N.J., Porter J.R., 2003. Agroecology, scaling and interdisciplinarity. *Agriculture, Ecosystems & Environment* 100(1), 39–51.
- [55] David J., Charrier A., 2009. La domestication des plantes, un processus ancien tellement moderne ! Présentaiton à l'Académie d'Agriculture.
- [56] Dawson J., Goldringer I., 2012. Breeding for genetically diverse populations : variety mixture and evolutionary populations. In : *Organic Crop Breeding*, Edith T. Lammerts van Bueren and James R. Myers. Wiley-backwell Edition., pp. 77–98.
- [57] Dawson J., Murphy K., Jones S., 2008. Decentralized selection and participatory approaches in plant breeding for low-input systems. *Euphytica* 160(2), 143–154.
- [58] Dawson J.C., Serpolay E., Giuliano S., Schermann N., Galic N., Berthelot J.F., Chesneau V., Ferté H., Mercier F., Osman A., Pino S., Goldringer I., 2013. Phenotypic diversity and evolution of farmer varieties of bread wheat on organic farms in europe. *Genetic Resources and Crop Evolution* 60(1), 145–163.
- [59] Dawson J.C., Serpolay E., Giuliano S., Schermann N., Galic N., Chable V., Goldringer I., 2012. Multi-trait evolution of farmer varieties of bread wheat after cultivation in contrasting organic farming systems in europe. *Genetica* 140(1-3), 1–17.
- [60] de Molina M.G., 2013. Agroecology and politics. how to get sustainability? about the necessity for a political agroecology. *Agroecology and Sustainable Food Systems* 37(1), 45–59.
- [61] Dellaporta S.L., Wood J., Hicks J.B., 1983. A plant DNA miniprep: Version II. *Plant Molecular Biology Reporter* 1(4), 19–21.
- [62] Demeulenaere E., . *Agriculteurs et chercheurs dans la gestion à la ferme des ressources génétiques : dynamiques d'apprentissage autour de la biodiversité. Sciences en campagne, Regards croisés, passés et à venir* .
- [63] Demeulenaere E., Bonneuil C., 2010. Cultiver la biodiversité. semences et identité paysanne. *Les mondes agricoles en politique. De la fin des paysans au retour de la question agricole* , 73–92.
- [64] Demeulenaere E., Bonneuil C., 2011. Des semences en partage : construction sociale et identitaire d'un collectif "paysan" autour de pratiques semencières alternatives. *Techniques & Culture* (57), 202–221.
- [65] Demeulenaere E., Bonneuil C., Balfourier F., Basson A., Berthelot J., Chesneau V., Ferté H., Galic N., Kastler G., Koenig J., Mercier F., Payerment J., Pommart A., Ronot B., Rousselle Y., Supiot N., Zaharia H., Goldringer I., 2008. Étude des complémentarités entre gestion dynamique à la ferme et gestion statique en collection. *Les Actes du BRG* , 117–138.

- [66] Demeulenaere E., Goulet F., 2012. Du singulier au collectif. agriculteurs et objets de la nature dans les réseaux d'agricultures «alternatives ». *Terrains & travaux* (20), 121–138.
- [67] Desclaux D., 2006. Sélection participative : spécificités et enjeux pour des agricultures paysannes et durables. *Courrier de l'environnement de l'INRA* (30), 119–127.
- [68] Desclaux D., Ceccarelli S., Navazio J., Coley M., Trouche G., Aguirre S., Weltzien E., Lançon J., 2012. Centralized or decentralized breeding : the potentials of participatory approaches for low-input and organic agriculture. In : *Organic Crop Breeding*, Edith T. Lammerts van Bueren and James R. Myers. Wiley-backwell Edition.
- [69] Desclaux D., Nolot J.M., Chiffolleau Y., 2009. La sélection participative pour élaborer des variétés de blé dur pour l'agriculture biologique. *Innovations Agronomiques* (7), 65–78.
- [70] Desclaux D., Nolot J.M., Chiffolleau Y., Leclerc C., Gozé E., 2008. Changes in the concept of genotype x environment interactions to fit agriculture diversification and decentralized participatory plant breeding : pluridisciplinary point of view. *Euphytica* 163, 533–546.
- [71] Diamond L., 1997. Civil society and the development of democracy. Working paper.
- [72] Dicke M., Baldwin I.T., 2010. The evolutionary context for herbivore-induced plant volatiles : beyond the 'cry for help'. *Trends in Plant Science* 15(3), 167–175.
- [73] Doré T., Meynard J.M., 2006. Itinéraire technique, système de culture : de la compréhension du fonctionnement du champ cultivé à l'évolution des pratiques agricoles. introduction. In : *L'agronomie aujourd'hui*, Doré T., Le Bail M., Martin P., Ney B., Roger-Estrade J., c/o Inra, RD 10, 78026 Versailles Cedex. quae Edition., p. 367.
- [74] Doussinault G., Pavoine M.T., Jaudeau B., Jahier J., 2001. Évolution de la variabilité génétique chez le blé. *Dossier de l'environnement de l'INRA* (21), 91–103.
- [75] Ducos G., Pujol J.L., Teillant A., 2013. Pour un secteur des semences diversifié et innovant. La note d'analyse du Commissariat général à la stratégie et à la prospective (5).
- [76] Duvacher P., 2010. Inventaire des variétés de pays circulant sur le Rhône-Alpes, et mise en perspectives des forces et faiblesses de l'organisation collective autour du programme semences. Rapport, Université Paris 1 - Panthéon Sorbonne.
- [77] Döring T., Pautasso M., Wolfe M., Finckh M.R., 2012. Pest and disease management in organic farming : Implications and inspirations for plant breeding. In : *Organic Crop Breeding*, Edith T. Lammerts van Bueren and James R. Myers. Wiley-backwell Edition.
- [78] Döring T.F., Knapp S., Kovacs G., Murphy K., Wolfe M.S., 2011. Evolutionary plant breeding in Cereals—Into a new era. *Sustainability* 3(12), 1944–1971.
- [79] Enjalbert J., David J.L., 2000. Inferring recent outcrossing rates using multilocus individual heterozygosity : Application to evolving wheat populations. *Genetics* 156(4), 1973–1982.

- [80] Enjalbert J., Dawson J., Paillard S., Rhoné B., Rousselle Y., Thomas M., Goldringer I., 2011. Dynamic management of crop diversity : From an experimental approach to on-farm conservation. *Comptes Rendus Biologies* 334(5-6), 458–468.
- [81] Enjalbert J., Goldringer I., David J., Brabant P., 1998. The relevance of outcrossing for the dynamic management of genetic resources in predominantly selfing *Triticum aestivum* L. (bread wheat). *Genet. Sel. Evol.* 30(Suppl. 1), 197–211.
- [82] Enjalbert J., Goldringer I., Paillard S., Brabant P., 1999. Molecular markers to study genetic drift and selection in wheat populations. *J. Exp. Bot.* 50(332), 283–290.
- [83] Even M.A., 2009. L'IAASTD, une expertise internationale qui marque un changement de paradigme pour l'agriculture et le développement. Analyse, centre d'étude et de prospective du Ministère de l'agriculture de l'agroalimentaire et de la forêt (6).
- [84] FADEAR, Confédération Paysanne, 2013. Charte de l'agriculture paysanne. [www.agriculturepaysanne.org](http://www.agriculturepaysanne.org), 28 octobre 2013.
- [85] Falconer D.S., 1952. The problem of environment and selection. *American Naturalist* (86), 293–298.
- [86] Falconer D.S., Mackay T.F.C., 1996. *Introduction to Quantitative Genetics*. Longman.
- [87] FAO, 1996. Déclaration de Rome sur la sécurité alimentaire mondiale.
- [88] FAO, 1996. The state of the world's plant genetic resources for food and agriculture. Rapport, Rome, Italie.
- [89] FAO, 1999. Agricultural biodiversity, multifunctional character of agriculture and land conference. In : Background papers 1. FAO, Maastricht, Pays-Bas.
- [90] FAO, 2007. International conference on organic agriculture and food security. Rapport, Rome, Italie.
- [91] FAO, 2009. Déclaration du sommet mondial sur la sécurité alimentaire.
- [92] FAO, 2010. Biodiversity for food and agriculture contributing to food security and sustainability in a changing world. Rapport, Rome, Italie.
- [93] FAO, 2010. Report of the twenty-second session of the committee on agriculture. Rapport, FAO, Rome, Italie.
- [94] FAO, 2010. The second report on the state of the world's plant genetic resources for food and agriculture. Rapport, Rome, Italie.
- [95] FAO, 2013. Statistical yearbook 2013. Rapport, Rome, Italie.
- [96] FAO, FIDA, PMA, 2012. L'état de l'insécurité alimentaire dans le monde L'état de l'insécurité alimentaire dans le monde 2012. la croissance économique est nécessaire mais elle n'est pas suffisante pour accélérer la réduction de la faim et de la malnutrition. Rapport, FAO, Rome, Italie.
- [97] Finckh M., Gacek E.S., Goyeau H., Lannou C., Merz U., Mundt C., Munk L., Nadziak J., Newton A., de Vallavieille-Pope C., 2000. Cereal variety and species mixtures in practice, with emphasis on disease resistance. *Agronomie* 20(7), 813–837.

- [98] Finckh M.R., 2008. Integration of breeding and technology into diversification strategies for disease control in modern agriculture. *European Journal of Plant Pathology* 121, 399–409.
- [99] Finckh M.R., Mundt C.C., 1992. Plant competition and disease in genetically diverse wheat populations. *Oecologia* 91(1), 82–92.
- [100] Finlay K., Wilkinson G., 1963. The analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research* 14(6), 742–754.
- [101] Fisher R.A., 1918. The correlation between relatives on the supposition of mendelian inheritance. *Philosophical Transactions of the Royal Society of Edinburgh* (52), 399–433.
- [102] FNAB, 2011. Rapide historique des règles françaises de l'agriculture biologique. [http://www.fnab.org/index.php?option=com\\_content&view=article&id=16:rapide-historique-des-regles-francaises-de-lagriculture-biologique&catid=8:reglementation&Itemid=17](http://www.fnab.org/index.php?option=com_content&view=article&id=16:rapide-historique-des-regles-francaises-de-lagriculture-biologique&catid=8:reglementation&Itemid=17), 27 septembre 2013.
- [103] Folke C., Carpenter S., Walker B., Scheffer M., Elmqvist T., Gunderson L., Holling C., 2004. Regime shifts, resilience, and biodiversity in ecosystem management. *Annual Review of Ecology, Evolution, and Systematics* 35(1), 557–581.
- [104] Fontaine L., Fourrié L., Garnier J.F., Mangin M., Colomb B., Carof M., Aveline A., Prieur L., Quirin T., Chareyron B., Maurice R., Glachant C., Gouraud J.P., 2012. Connaître, caractériser et évaluer les rotations en systèmes de grandes cultures biologiques. In : Résultats de l'appel à projet CASDAR 2007. Paris, pp. 1–14.
- [105] FranceAgriMer, 2012. Variétés et rendement des céréales biologiques. Récolte 2011. Rapport.
- [106] Francis C., Lieblein G., Gliessman S., Breland T.A., Creamer N., Harwood R., Salomonsson L., Helenius J., Rickerl D., Salvador R., Wiedenhoef M., Simmons S., Allen P., Altieri M., Flora C., Poincelot R., 2003. Agroecology : The ecology of food systems. *Journal of Sustainable Agriculture* 22(3), 99–118.
- [107] Frison E., Cherfas J., Hodgkin T., 2011. Agricultural biodiversity is essential for a sustainable improvement in food and nutrition security. *Sustainability* 3(1), 238–253.
- [108] Fruchaud G., 2012. L'organisation du projet SOLIBAM-croisement du Roc 2011-2012. Rapport, Ecole Supérieur d'Agriculture, Angers.
- [109] Fu D., Szűcs P., Yan L., Helguera M., Skinner J.S., Zitzewitz J.v., Hayes P.M., Dubcovsky J., 2005. Large deletions within the first intron in VRN-1 are associated with spring growth habit in barley and wheat. *Molecular Genetics and Genomics* 274(4), 442–443.
- [110] Fufa F., Grando S., Kafawin O., Shakhathreh Y., Ceccarelli S., 2010. Efficiency of farmers' selection in a participatory barley breeding programme in Jordan. *Plant Breeding* 129, 156 – 161.
- [111] Gallais A., 1990. Théorie de la sélection en amélioration des plantes. Sciences Agronomiques. Masson Edition.

- [112] Gate P., 2009. Formation physiologique des céréales ARVALIS. Rapport, La minière, Guyancourt.
- [113] Gauch H., 2006. Statistical analysis of yield trials by AMMI and GGE. *Crop Sci* 46(4), 1488–1500.
- [114] Gauch H., Piepho H., Annicchiarico P., 2008. Statistical analysis of yield trials by AMMI and GGE : further considerations. *Crop Sci* 48(3), 866–889.
- [115] Ghaouti L., Vogt-Kaute W., Link W., 2008. Development of locally-adapted faba bean cultivars for organic conditions in germany through a participatory breeding approach. *Euphytica* 162, 257–268.
- [116] Gianinazzi S., Gollotte A., Binet M.N., Tuinen D.v., Redecker D., Wipf D., 2010. Agroecology : the key role of arbuscular mycorrhizas in ecosystem services. *Mycorrhiza* 20(8), 519–530.
- [117] Goffaux R., Goldringer I., Bonneuil C., Montalent P., Bonnin I., 2011. Quels indicateurs pour suivre la diversité génétique des plantes cultivées ? le cas du blé tendre en france depuis un siècle. Rapport, FRB.
- [118] Goldringer I., 1993. Optimisation d'un schéma de sélection récurrente à cycle court appliqué au blé tendre d'hiver (*Triticum aestivum* L.). Thèse de doctorat, Institut National Agronomique, Paris-Grignon.
- [119] Goldringer I., Bonnin I., David J., 2013. Rôle de la sélection dans la dynamique de la diversité des blés cultivés : de la domestication à la sélection moderne au 20ème siècle. *Innovations Agronomiques* (29), 1–15.
- [120] Goldringer I., Brabant P., Gallais A., 1997. Estimation of additive and epistatic genetic variances for agronomic traits in a population of doubled-haploid lines of wheat. *Heredity* 79(1), 60–71.
- [121] Goldringer I., Brabant P., Kempton R.A., 1994. Adjustment for competition between genotypes in single-row-plot trials of winter wheat (*triticum aestivum*). *Plant Breeding* 112(4), 294–300.
- [122] Goldringer I., Enjalbert J., David J., Paillard S., Pham J., Brabant P., 2001. Dynamique management of genetic resources : a 13-year experiment on wheat. *Broadening the Genetic Base of Crop Production* , 245–260.
- [123] Goldringer I., Prouin C., Rousset M., Galic N., Bonnin I., 2006. Rapid differentiation of experimental populations of wheat for heading time in response to local climatic conditions. *Annals of Botany* 98(4), 805–817.
- [124] Grim T., 2008. A possible role of social activity to explain differences in publication output among ecologists. *Oikos* 117(4), 484–487.
- [125] Hajjar R., Jarvis D.I., Gemmill-Herren B., 2008. The utility of crop genetic diversity in maintaining ecosystem services. *Agriculture, Ecosystems and Environment* 123, 261–270.
- [126] Hammer K., Arrowsmith N., Gladis T., 2003. Agrobiodiversity with emphasis on plant genetic resources. *Naturwissenschaften* 90(6), 241–250.

- [127] Hanski I., Gilpin M., 1991. Metapopulation dynamics : brief history and conceptual domain. *Biological Journal of the Linnean Society* 42(1-2), 3–16.
- [128] Henry J.P., Pontis C., David J.L., Gouyon P.H., 1991. An experiment on dynamic conservation of genetic resources with metapopulations. *Species conservation : A population biological approach* , 185–198.
- [129] Higgins J.A., Bailey P.C., Laurie D.A., 2010. Comparative genomics of flowering time pathways using brachypodium distachyon as a model for the temperate grasses. *PLoS ONE* 5(4), e10065.
- [130] Hoad S., Bertholdsson N., Neuhoﬀ D., Köpke U., 2012. Approches to breed for improved weed suppression in organically grown cereals. In : *Organic Crop Breeding*, Edith T. Lammerts van Bueren and James R. Myers. Wiley-backwell Edition.
- [131] Husson F., Josse J., Pagès J., 2010. Principal component methods - hierarchical clustering - partitional clustering : why would we need to choose for visualizing data? Rapport, Agrocampus Ouest.
- [132] Hyacinthe A., 2009. Des relations nouvelles ? semences paysannes, projets agricoles et débat environnemental. Rapport, Université Lumière Lyon 2. Faculté d'Anthropologie et de Sociologie, Lyon.
- [133] IAASTD, 2008. Agriculture at a crossroads. Évaluation internationale des connaissances, des sciences et des technologies agricoles pour le développement. Résumé analytique du rapport de synthèse. Rapport.
- [134] IFOAM, 2013. Principles of organic agriculture. Rapport.
- [135] INC, 2011. A quel label bio se fier? [http://www.60millions-mag.com/guides\\_d\\_achat/conseils\\_d\\_experts/a\\_quel\\_label\\_bio\\_se\\_fier](http://www.60millions-mag.com/guides_d_achat/conseils_d_experts/a_quel_label_bio_se_fier), 24 septembre 2013.
- [136] Incroyables-comestibles, 2013. Incroyables comestibles france. <http://www.incredible-edible.info/>, 28 septembre 2013.
- [137] INRA, 2011. Lancement du projet AMAIZING. <https://colloque4.inra.fr/lancementamaizing/>, 24 septembre 2013.
- [138] INRA, 2013. Projet Breed Wheat : lettre d'information numéro 1. <http://www.breedwheat.fr/fr/news.php?optimurl=1ere-lettre-d%27information-Breedwheat>, 1 octobre 2013.
- [139] ITAB, 2011. Rotations en grandes cultures biologiques sans élevage 8 fermes-types, 11 rotations repères agronomiques, économiques, techniques et environnementaux. Rapport.
- [140] ITAB, 2012. Désherber mécaniquement les grandes cultures. projet « optimiser et promouvoir le désherbage mécanique ». CASDAR 2009/2011. Rapport, ITAB.
- [141] ITAB, ARVALIS, 2013. Référentiel blé AB, les résultats du réseau ITAB. Rapport.
- [142] Jackson L., Pascual U., Hodgkin T., 2007. Utilizing and conserving agrobiodiversity in agricultural landscapes. *Agriculture, Ecosystems & Environment* 121(3), 196–210.



- [143] Jarvis D., Hodgkin T., Sthapit B.R., Fadda C., Lopez-Noriega I., 2011. An heuristic framework for identifying multiple ways of supporting the conservation and use of traditional crop varieties within the agricultural production system. *Critical Reviews in Plant Science* 30, 125–176.
- [144] Jombart T., Devillard S., Balloux F., 2010. Discriminant analysis of principal components : a new method for the analysis of genetically structured populations. *BMC Genetics* 11(1), 94.
- [145] Joshi K., Musa A., Johansen C., Gyawali S., Harris D., Witcombe J., 2007. Highly client-oriented breeding, using local preferences and selection, produces widely adapted rice varieties. *Field Crops Research* 100(1), 107–116.
- [146] Joshi K., Witcombe J., 2003. The impact of participatory plant breeding (PPB) on landrace diversity : A case study for high-altitude rice in Nepal. *Euphytica* 134, 117–125.
- [147] Josse J., Denis J.B., 2012. Inferring biadditive model with the bayesian paradigm. Rapport, INRA, Jouy en Josas.
- [148] Khan A.R., 2013. Short term response of European wheat populations to contrasted agro-climatic conditions : a genetic analysis and first step towards development of epigenetic markers in earliness gene VRN-A1. Thèse de doctorat, Université Paris 11.
- [149] Kiiza B., Kitembo L., Mwangi R.O.M., 2012. Participatory plant breeding and selection impact on adoption of improved sweetpotato varieties in Uganda. *Journal of Agricultural Science and Technology* 2, 673–681.
- [150] Kiær L.P., Skovgaard I.M., Østergård H., 2009. Grain yield increase in cereal variety mixtures : A meta-analysis of field trials. *Field Crops Research* 114(3), 361–373.
- [151] Lammerts Van Bueren E., Myers J., 2012. Organic crop breeding : Integrating organic agricultural approaches and traditional and modern plant breeding methods. In : *Organic Crop Breeding*, Edith T. Lammerts van Bueren and James R. Myers. Wiley-backwell Edition., pp. 3–13.
- [152] Latour B., 1997. Toute recherche est action ! Etudes et recherches sur les systèmes agraires et le développement (30), 197–208.
- [153] Le Boulc'h V., David J., Brabant P., De Vallavieille-Pope C., 1994. Dynamic conservation of variability : responses of wheat populations to different selective forces including powdery mildew. *Genetics Selection and Evolution* 26, 221–240.
- [154] Le Foll S., 2013. La loi d'avenir pour l'agriculture, l'alimentation et la forêt. Ministère de l'agriculture, de l'agroalimentaire et la forêt.
- [155] Levrouw F., 2013. Les Maisons des Semences Paysannes en France. Étude des organisations collectives de gestion de la biodiversité cultivée au sein du Réseau Semences Paysannes. Version de travail, Réseau Semences Paysannes, Aiguillon.
- [156] LGC Genomics, 2013. KASP genotyping reagents. <http://www.lgccgenomics.com/genotyping/kasp-genotyping-reagents/>, 1 octobre 2013.

- [157] Li J., Lammerts van Bueren E.T., Huang K., Qin L., Song Y., 2013. The potential of participatory hybrid breeding. *International Journal of Agricultural Sustainability* 11(3), 234–251.
- [158] Lobell D.B., Schlenker W., Costa-Roberts J., 2011. Climate trends and global crop production since 1980. *Science* 333(6042), 616–620.
- [159] Masson F., 2013. La démarche VATE. Présentation à Agrocampus Ouest - Centre d'Angers, GEVES.
- [160] Maxted N., Guarino L., Myer L., Chiwona E.A., 2002. Towards a methodology for on-farm conservation of plant genetic resources. *Genetic Resources and Crop Evolution* 49(1), 31–46.
- [161] McDonald B.A., Linde C., 2002. The population genetics of plant pathogens and breeding strategies for durable resistance. *Euphytica* 124(2), 163–180.
- [162] MEA, 2005. *Ecosystems and Human Well-being : General Synthesis*. Millennium Ecosystem Assessment, Washington DC, USA, Island Press Edition.
- [163] Messmer M., Hildermann I., Thorup-Kristensen K., Rengel Z., 2012. Nutrient management in organic farming and consequences for direct and indirect selection strategies. In : *Organic Crop Breeding*, Edith T. Lammerts van Bueren, James R. Myers. Wiley-backwell Edition., pp. 15–38.
- [164] Meynard J.M., Cresson C., 2011. Le conseil scientifique de l'Agriculture biologique identifie 8 priorités de recherche-développement. *Notes et études socio-économiques* (35), 27–40.
- [165] Meynard J.M., Messéan A., Charlier A., Charrier F., Fares M., Le Bail M., Magrini M., Savini I., 2013. Freins et leviers à la diversification des cultures. Étude au niveau des exploitations agricoles et des filières. Rapport, INRA.
- [166] Michaels S., John M., Amasino R., 1994. Removal of polysaccharides from plant DNA by ethanol precipitation. *Biotechniques* 17(2), 274–276.
- [167] Mille B., Fraj M., Monod H., de Vallavieille-Pope C., 2006. Assessing four-way mixtures of winter wheat cultivars from the performances of their two-way and individual components. *European Journal of Plant Pathology* 114(2), 163–173.
- [168] Millot G., Bonneuil C., 2005. Savoirs libres et production de biens communs en réseau. Compte-rendu de la rencontre entre acteurs des semences paysannes et acteurs de projets coopératifs numériques libres, organisée le 3 décembre 2005 par la Fondation Sciences Citoyennes, le Réseau Semences Paysannes et l'Association pour la Promotion et la Recherche en Informatique Libre.
- [169] Millot G., Neubauer C., Storup B., 2013. La recherche participative comme mode de production de savoirs. un état des lieux des pratiques en France.
- [170] Moreira P.M.R.M., Pêgo S.E., Patto C.V., Hallauer A.R., 2008. Comparison of selection methods on pigarro, a portuguese improved maize population with fasciation expression. *Euphytica* 163, 481–499.
- [171] Morris M., Bellon M., 2004. Participatory plant breeding research : Opportunities and challenges for the international crop improvement system. *Euphytica* 136, 21–35.

- [172] Moÿ A., 2010. Etude sur le devenir juridique des variétés issues de sélection participative et paysanne. Rapport, programme DIVERBA.
- [173] Murphy K., Campbell K., Lyon S., Jones S., 2007. Evidence of varietal adaptation to organic farming systems. *Field Crops Research* 102(3), 172–177.
- [174] Murphy K.M., Hoagland L.A., Reeves P.G., Baik B.K., Jones S.S., 2009. Nutritional and quality characteristics expressed in 31 perennial wheat breeding lines. *Renewable Agriculture and Food Systems* 24(04), 285–292.
- [175] Murphy K.M., Lammer D., Lyon S.R., Carter B., Jones S.S., 2005. Breeding for organic and low-input farming systems : An evolutionary-participatory breeding method for inbred cereal grains. *Renewable Agriculture and Food Systems* 20(1), 48–55.
- [176] Nabugoomu F., Kempton R., Talbot M., 1999. Analysis of series of trials where varieties differ in sensitivity to locations. *Journal of agricultural, biological and environmental Statistics* 4(3), 310–325.
- [177] Negri V., 2003. Landraces in central italy : where and why they are conserved and perspectives for their on-farm conservation. *Genetic Resources and Crop Evolution* 50(8), 871–885.
- [178] Nei M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89(3), 583–590.
- [179] Newton A., Akar T., Baresel J., Bebeli P., Bettencourt E., Bladenopoulos K., Czembor J., Fasoula D., Katsiotis A., Koutis K., Koutsika-Sotiriou M., Kovacs G., Larsson H., Carvalho M.P.d., Rubiales D., Russell J., Santos T.D., Patto M.V., 2010. Cereal landraces for sustainable agriculture. a review. *Agronomy for Sustainable Development* 30(2), 237–269.
- [180] Ninkovic V., Åhman I.M., 2009. Aphid acceptance of hordeum genotypes is affected by plant volatile exposure and is correlated with aphid growth. *Euphytica* 169(2), 177–185.
- [181] Obellianne S., 2012. Diagnostic et perspectives pour le projet de conservation et de sélection de blés anciens à graines de noé. Rapport, Ecole Supérieur d’Agriculture, Angers.
- [182] Olesen J.E., Askegaard M., Rasmussen I.A., 2009. Winter cereal yields as affected by animal manure and green manure in organic arable farming. *European Journal of Agronomy* 30(2), 119–128.
- [183] Olivieri I., Michalakis Y., Gouyon P.H., 1995. Metapopulation genetics and the evolution of dispersal. *The American Naturalist* 146(2), 202.
- [184] ONU, 2012. Drought in U.S. pushing up world food prices. <http://www.un.org/en/development/desa/news/policy/wesp.html>, 15 septembre 2013.
- [185] Osman A., Chable V., 2009. Inventory of initiatives on seeds of landraces in europe. *Journal of Agriculture and Environment for International Development* 103, 95–130.

- [186] Ostergard H., Finckh R., Fontaine L., Goldringer I., Hoad S.P., Kristensen K., Lammerts van Bueren E., Mascher F., Munk L., Wolfe M., 2009. Time for a shift in crop production : embracing complexity through diversity at all levels. *Journal of the Science of Food and Agriculture* 89(9), 1439–1445.
- [187] Paillard S., Goldringer I., Enjalbert J., Doussinault G., de Vallavieille-Pope C., Brabant P., 2000. Evolution of resistance against powdery mildew in winter wheat populations conducted under dynamic management. I. Is specific seedling resistance selected ? *TAG Theoretical and Applied Genetics* 101(3), 449–456.
- [188] Paillard S., Goldringer I., Enjalbert J., Trottet M., David J., de Vallavieille-Pope C., Brabant P., 2000. Evolution of resistance against powdery mildew in winter wheat populations conducted under dynamic management. II. Adult plant resistance. *TAG Theoretical and Applied Genetics* 101(3), 457–462.
- [189] Papaïx J., Goyeau H., Du Cheyron P., Monod H., Lannou C., 2011. Influence of cultivated landscape composition on variety resistance : an assessment based on wheat leaf rust epidemics. *New Phytologist* 191(4), 1095–1107.
- [190] Papy F., Goldringer I., 2011. Cultiver la biodiversité. *Courrier de l'environnement de l'INRA* (60), 55–62.
- [191] Pautasso M., Aistara G., Barnaud A., Caillon S., Clouvel P., Coomes O.T., Deletre M., Demeulenaere E., De Santis P., Doering T., Eloy L., Emperaire L., Garine E., Goldringer I., Jarvis D., Joly H.I., Leclerc C., Louafi S., Martin P., Massol F., McGuire S., McKey D., Padoch C., Soler C., Thomas M., Tramontini S., 2012. Seed exchange networks for agrobiodiversity conservation. a review. *Agronomy for Sustainable Development* 33(1), 151–175.
- [192] Phillips S., Wolfe M., 2005. Evolutionary plant breeding for low input systems. *The Journal of Agricultural Science* 143(04), 245–254.
- [193] Pimentel D., Hepperly P., Hanson J., Douds D., Seidel R., 2005. Environmental, energetic and economic comparisons of organic and conventional farming systems. *BioScience* 55(7), 573–582.
- [194] Pinheiro de Carvalho M.A.A., Bebeli P.J., Bettencourt E., Costa G., Dias S., Dos Santos T.M.M., Slaski J.J., 2013. Cereal landraces genetic resources in worldwide GeneBanks. a review. *Agronomy for Sustainable Development* 33(1), 177–203.
- [195] Power A.G., 2010. Ecosystem services and agriculture : tradeoffs and synergies. *Philosophical Transactions of the Royal Society B-Biological Sciences* 365(1554), 2959–2971.
- [196] Purugganan M.D., Fuller D.Q., 2009. The nature of selection during plant domestication. *Nature* 457(7231), 843–848.
- [197] R Core Team, 2013. *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- [198] Raquin A.L., Brabant P., Rhoné B., Balfourier F., Leroy P., Goldringer I., 2008. Soft selective sweep near a gene that increases plant height in wheat. *Molecular ecology* 17(3), 741–756.

- [199] Reif J.C., Zhang P., Dreisigacker S., Warburton M.L., Ginkel M.v., Bohn D.H.M., Melchinger A.E., 2005. Wheat genetic diversity trends during domestication and breeding. *Theoretical and Applied Genetics* 110, 859–864.
- [200] Reusch T.B.H., Ehlers A., Hämmerli A., Worm B., 2005. Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *Proceedings of the National Academy of Sciences of the United States of America* 102(8), 2826–2831.
- [201] Rhone B., 2008. Étude de mécanismes génétiques impliqués dans l'adaptation climatique de populations expérimentales de blé tendre. Thèse de doctorat, Université Paris 11.
- [202] Rhone B., Remoue C., Galic N., Goldringer I., Bonnin I., 2008. Insight into the genetic bases of climatic adaptation in experimentally evolving wheat populations. *Molecular Ecology* 17(3), 930–943.
- [203] Robert C., 2001. *The Bayesian Choice*. Springer Texts in Statistics. Springer, second Edition.
- [204] Roger-Estrade J., 2012. Caractérisation du champ cultivé. Rapport, AgroParisTech.
- [205] Roger-Estrade J., Richard G., Caneill J., 2006. Effets des systèmes de culture sur l'évolution des états du milieu cultivé. introduction. In : *L'agronomie aujourd'hui*, Doré T., Le Bail M., Martin P., Ney B., Roger-Estrade J., c/o Inra, RD 10, 78026 Versailles Cedex. quae Edition., p. 367.
- [206] Rolland B., Le Campion A., Oury F.X., 2012. Pourquoi sélectionner de nouvelles variétés de blé tendre adaptées à l'agriculture biologique ? *Courrier de l'environnement de l'INRA* (62), 71–85.
- [207] Roussel V., Koenig J., Beckert M., Balfourier F., 2004. Molecular diversity in french bread wheat accessions related to temporal trends and breeding programmes. *TAG Theoretical and Applied Genetics* 108(5), 920–930.
- [208] Roussel V., Leisova L., Exbrayat F., Stehno Z., 2005. SSR allelic diversity changes in 480 european bread wheat varieties released from 1840 to 2000. *Theor Appl Genet* 111, 162–170.
- [209] Rousset M., Bonnin I., Remoué C., Falque M., Rhoné B., Veyrieras J.B., Madur D., Murigneux A., Balfourier F., Gouis J.L., Santoni S., Goldringer I., 2011. Deciphering the genetics of flowering time by an association study on candidate genes in bread wheat (*triticum aestivum* l.). *Theoretical and Applied Genetics* 123(6), 907–926.
- [210] Réseau Semences Paysannes, 2008. Voyage autour des blés paysans. Témoignages Edition.
- [211] Santilli J., 2011. *Agrobiodiversity and the law*. Routledge Edition.
- [212] Sebillotte M., Papy F., 2010. Michel sebillotte, agronome : penser l'action. *Natures Sciences Sociétés* Vol. 18(4), 446–451.
- [213] Serpolay E., Schermann N., Dawson J., van Bueren E.T.L., Goldringer I., Chable V., 2011. Phenotypic changes in different spinach varieties grown and selected under organic conditions. *Sustainability* 3(12), 1616–1636.

- [214] Sherman J.D., Yan L., Talbert L., Dubcovsky J., 2004. A PCR marker for growth habit in common wheat based on allelic variation at the gene. *Crop Science* 44(5), 1832.
- [215] Slow Food, 2010. Slow Food France. <http://www.slowfood.fr/le-memento-slow-food>, 29 septembre 2013.
- [216] Smith M.E., Castillo G. F., Gomez F., 2001. Participatory plant breeding with maize in Mexico and Honduras. *Euphytica* 122(3), 551–565.
- [217] Smith P., Martino D., Cai Z., Gwary D., Janzen H., Kumar P., McCarl B., Ogle S., O'Mara F., Rice C., Scholes B., Sirotenko O., 2007. Agriculture. In : *Climate Change 2007 : Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer, Cambridge, United Kingdom and New York, NY, USA. Cambridge University Press Edition., pp. 497–540.
- [218] Sperling L., Ashby J.A., Smith M.E., Weltzien E., McGuire S., 2001. A framework for analyzing participatory plant breeding approaches and results. *Euphytica* 122, 439–450.
- [219] Sperling L., Loevinsohn M., Ntabomvura B., 1993. Rethinking the farmer's role in plant breeding : local bean experts and on-station selection in Rwanda. *Experimental Agriculture* 29(4), 509–519.
- [220] Star S.L., Griesemer J.R., 1989. Institutional Ecology, Translations and Boundary Objects : Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science* 19(3), 387–420.
- [221] Stassart P., 2010. The potential of agroecology to contribute to ecological transition in agriculture. Proceedings of the international conference. Rapport, Arlon Campus, Belgique.
- [222] Stassart P., Baret P., Grégoire J., Hance T., Mormont M., Reheul D., Stilmant D., Vanloqueren G., Visser M., 2012. L'agroécologie : trajectoire et potentiel pour une transition vers des systèmes alimentaires durables. In : *Agroécologie entre pratiques et sciences sociales*, Van Dam D., Nizet J., Streith M., Stassart P.M., 26 Boulevard du Docteur Petitjean, 21079 Dijon Cedex. Educagri Edition., p. 309.
- [223] Stassart P.M., Jamar D., 2008. Steak up to the horns! the conventionalization of organic stock farming : knowledge lock-in in the agrifood chain. *GeoJournal* 73(1), 31–44.
- [224] Storup B., de Kochko P., 2011. Suivi d'une expérience de sélection participative en cours l'exemple du projet européen SOLIBAM. Rapport, Association pour une Fondation Science Citoyenne.
- [225] Su Z., Hao C., Wang L., Dong Y., Zhang X., 2011. Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 122(1), 211–223.
- [226] Suarez A., Gitay H., Watson R., Jon Dokken D., 2002. Les changements climatiques et la biodiversité. document technique V du GIEC.

- [227] Suneson C.A., 1956. An evolutionary plant breeding method. *Agronomy Journal* 48, 188–191.
- [228] Thapa D., Sharma R., Mudwari A., Ortiz-Ferrara G., Sharma S., Basnet R., Witcombe J., Virk D., Joshi K., 2009. Identifying superior wheat cultivars in participatory research on resource poor farms. *Field Crops Research* 112(2–3), 124–130.
- [229] Thomas M., 2011. Gestion dynamique à la ferme de l’agrobiodiversité : relation entre la structure des populations de blé tendre et les pratiques humaines. Thèse de doctorat, Université Paris Diderot.
- [230] Thomas M., Dawson J.C., Goldringer I., Bonneuil C., 2011. Seed exchanges, a key to analyze crop diversity dynamics in farmer-led on-farm conservation. *Genetic Resources and Crop Evolution* 58(3), 321–338.
- [231] Thomas M., Demeulenaere E., Dawson J.C., Khan A.R., Galic N., Jouanne-Pin S., Remoue C., Bonneuil C., Goldringer I., 2012. On-farm dynamic management of genetic diversity : the impact of seed diffusions and seed saving practices on a population-variety of bread wheat. *Evolutionary Applications* 5(8), 779–795.
- [232] Thormann I., Gaisberger H., Mattei F., Snook L., Arnaud E., 2012. Digitization and online availability of original collecting mission data to improve data quality and enhance the conservation and use of plant genetic resources. *Genetic Resources and Crop Evolution* 59(5), 635–644.
- [233] Thro A., Spillane C., 2000. Biotechnology-assisted participatory plant breeding : Complement or contradiction ? Working Document 4, CGIAR.
- [234] Thrupp L.A., 2000. Linking agricultural biodiversity and food security : the valuable role of agrobiodiversity for sustainable agriculture. *International Affairs* 76(2), 283–297.
- [235] Thépot S., 2014. Utilisation d’une population multi-parentale et hautement recombinante de blé tendre pour l’étude de l’architecture génétique de la précocité de floraison. Thèse de doctorat, Université Paris 11.
- [236] Tomich T.P., Brodt S., Ferris H., Galt R., Horwath W.R., Kebreab E., Leveau J.H.J., Liptzin D., Lubell M., Merel P., Micheltore R., Rosenstock T., Scow K., Six J., Williams N., Yang L., 2011. Agroecology : A review from a global-change perspective. *Annual Review of Environment and Resources* 36, 193–222.
- [237] Tooker J.F., Frank S.D., 2012. Genotypically diverse cultivar mixtures for insect pest management and increased crop yields. *Journal of Applied Ecology* 49(5), 974–985.
- [238] Triticeae Toolbox, 2013. The Triticeae Toolbox (T3). <http://malt.pw.usda.gov/t3/sandbox/wheat/about.php>, 1 octobre 2013.
- [239] Trochard R., Bouthier A., Morvan T., Grall J., 2012. Effet azote à court, moyen et long terme des produits résiduels organiques issus d’élevages. In : Résultats de l’appel à projet CASDAR 2007. Paris, pp. 15–30.
- [240] UNCTAD (United Nation Conference On Trade And Development), 2013. Wake up before it is too late : Make agriculture truly sustainable now for food security in a changing climate. Trade and environment review. Rapport.

- [241] UNISDR (United Nations Office for Disaster Risk Reduction), 2012. Economic losses from disasters set new record in 2012. Rapport.
- [242] UPOV, 2002. Introduction générale à l'examen de la distinction, de l'homogénéité et de la stabilité et à l'harmonisation des descriptions des obtentions végétales. Rapport.
- [243] UPOV, 2013. Mission de l'UPOV. <http://www.upov.int/overview/fr/>, 15 septembre 2013.
- [244] Van den Broeck H.C., de Jong H.C., Salentijn E.M.J., Dekking L., Bosch D., Hamer R.J., Gilissen L.J.W.J., van der Meer I.M., Smulders M.J.M., 2010. Presence of celiac disease epitopes in modern and old hexaploid wheat varieties : wheat breeding may have contributed to increased prevalence of celiac disease. TAG. Theoretical and applied genetics. *Theoretische und angewandte Genetik* 121(8), 1527–1539.
- [245] Vanloqueren G., Baret P.V., 2008. Why are ecological, low-input, multi-resistant wheat cultivars slow to develop commercially ? a belgian agricultural 'lock-in' case study. *Ecological Economics* 66, 436–446.
- [246] Vanloqueren G., Baret P.V., 2009. How agricultural research systems shape a technological regime that develops genetic engineering but locks out agroecological innovations. *Research Policy* 38(6), 971–983.
- [247] Vaz Patta M.C., Moreira P.M., Almeida N., Satovic Z., Pego S., 2008. Genetic diversity evolution through participatory maize breeding in portugal. *Euphytica* 161, 283–291.
- [248] Vindras C., Sinoir N., 2012. Guide pratique : Comment évaluer la qualité gustative d'un produit ? Rapport, ITAB, Paris, France.
- [249] Virk D., Chakraborty M., Ghosh J., Prasad S., Witcombe J., 2005. Increasing the client orientation of maize breeding using farmer participation in eastern india. *Experimental Agriculture* 41(04), 413–426.
- [250] Virk D., Witcombe J., 2007. Evaluating cultivars in unbalanced on-farm participatory trials. *Field Crops Research* 106(2), 105–115.
- [251] Vom Brocke K., Trouche G., Weltzien E., Barro-Kondombo C.P., Gozé E., Chantereau J., 2010. Participatory variety development for sorghum in burkina faso : Farmers' selection and farmers' criteria. *Field Crops Research* 119(1), 183–194.
- [252] Warner K.D., 2008. Agroecology as participatory science emerging alternatives to technology transfer extension practice. *Science Technology & Human Values* 33(6), 754–777.
- [253] Wezel A., Bellon S., Dore T., Francis C., Vallod D., David C., 2009. Agroecology as a science, a movement and a practice. a review. *Agronomy for Sustainable Development* 29(4), 503–515.
- [254] Wilhelm E.P., Turner A.S., Laurie D.A., 2009. Photoperiod insensitive ppd-a1a mutations in tetraploid wheat (*triticum durum* desf.). *Theoretical and Applied Genetics* 118(2), 285–294.
- [255] Wolfe M., 2000. Crop strength through diversity. *Nature* 406(6797), 681–682.



- [256] Wolfe M., Baresel J., Desclaux D., Goldringer I., Hoad S., Kovacs G., Löschenberger F., Miedaner T., Østergård H., Lammerts van Bueren E., 2008. Developments in breeding cereals for organic agriculture. *Euphytica* 163(3), 323–346.
- [257] Wood D., Lenne J.M., 1997. The conservation of agrobiodiversity on-farm : questioning the emerging paradigm. *Biodiversity and Conservation* 6(1), 109–129.
- [258] Wricke G., 1962. Über eine methode zur erfassung der ökologischen streubreite in feldversuchen. *Zeitschrift Fur Pflanzenzuchtung-Journal of Plant Breeding* 47(1), 92–&.
- [259] Yan W., Kang M., Ma B., Woods S., Cornelius P., 2007. GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Science* 47, 643–355.
- [260] Zhu Y., Chen H., Fan J., Wang Y., Li Y., Chen J., Fan J., Yang S., Hu L., Leung H., Mew T., Teng P., Wang Z., Mundt C., 2000. Genetic diversity and disease control in rice. *Nature* 406(6797), 718–722.





# Table des matières

Avant-propos	9
Remerciements	11
Résumé	17
Liste des abréviations	19
Sommaire	24
1 Système semencier et souveraineté alimentaire	25
1.1 Souveraineté alimentaire, agroécologie et semences	26
1.1.1 Du concept de sécurité alimentaire à celui de souveraineté alimentaire	26
1.1.2 Etat des lieux et menaces pour la sécurité alimentaire dans le monde	27
1.1.3 Les alternatives : agroécologie et participation des acteurs	29
1.1.3.1 L'agriculture doit changer de paradigme	29
1.1.3.2 Les principes de l'agroécologie	30
1.1.3.3 La participation des acteurs	37
1.1.4 Semences et ressources génétiques au sein de l'agrobiodiversité	38
1.2 Stratégies de gestion des ressources génétiques et de sélection variétale pour l'agroécologie	39
1.2.1 La gestion des ressources génétiques	39
1.2.1.1 La gestion <i>ex-situ</i>	40

1.2.1.2	La gestion <i>in-situ</i>	42
1.2.2	La sélection décentralisée	44
1.2.3	La structure génétique des variétés	51
1.2.3.1	La variabilité pour s'adapter	52
1.2.3.2	La variabilité pour conférer stabilité et résilience	53
1.3	Participation des acteurs pour la sélection et la gestion des ressources génétiques	55
1.3.1	Les paysans sélectionneurs et gestionnaires des ressources génétiques depuis le néolithique	55
1.3.2	L'institutionnalisation de la sélection et de la gestion des ressources génétiques et son impact sur la diversité cultivée	56
1.3.3	La réappropriation de la sélection et de la gestion des ressources génétiques par la société civile en France	60
1.3.3.1	L'émergence de la société civile sur la question des semences	60
1.3.3.2	La sélection décentralisée et participative	61
1.3.3.3	La gestion des ressources génétiques par les paysans	63
1.4	Gouvernance des semences et des ressources génétiques	68
1.4.1	Le droit des semenciers et le système de propriété intellectuelle	69
1.4.1.1	Le système des brevets	69
1.4.1.2	Le système <i>sui generis</i>	72
1.4.1.3	Différences et interactions entre le système <i>sui generis</i> et le système des brevets	74
1.4.2	Le droit des paysans	74
1.4.3	L'accès aux ressources génétiques et le partage des bénéfices	75
1.4.3.1	Convention sur la Diversité Biologique	75
1.4.3.2	Traité International sur les Ressources Phytogénétiques pour l'Alimentation et l'Agriculture	76
1.4.4	La commercialisation des semences en Europe et en France	78
1.4.4.1	Le catalogue officiel européen	78
1.4.4.2	Le catalogue de conservation	79

1.4.4.3	Le catalogue officiel français	79
1.5	Contexte et objectifs du projet de sélection décentralisée et participative sur le blé tendre en France	80
1.5.1	Le contexte français	80
1.5.2	Objectifs du travail de thèse	85
2	Analyse phénotypique du programme de sélection participative	87
2.1	Dispositifs expérimentaux dans les fermes	88
2.1.1	Fermes participant au projet	88
2.1.2	Dispositifs mis en place dans les fermes	92
2.1.3	Variables mesurées	97
2.2	Analyse phénotypique dans les fermes	100
2.2.1	Méthodes	100
2.2.1.1	Comparaison de moyennes dans les fermes	100
2.2.1.2	Sélection massale intra-population	104
2.2.2	Résultats	109
2.2.2.1	Modèle Bayésien hiérarchique	109
2.2.2.2	Comportement des populations dans les fermes	114
2.2.2.3	Sélection inter-population	116
2.2.2.4	Sélection massale intra-population	118
2.2.3	Discussion	128
2.2.3.1	Méthode statistique utilisée et perspectives pour les programmes de sélection participative	128
2.2.3.2	Résultats du programme de sélection participative	130
2.3	Analyse des interactions $G \times E$	134
2.3.1	Méthodes	134
2.3.1.1	Travail préliminaire sur les données	134
2.3.1.2	L'analyse de l'interaction avec le modèle bayésien hiérarchique	138
2.3.2	Résultats et discussion	140
3	Evaluation phénotypique et moléculaire du programme de sélection participative	145
3.1	Matériels et méthodes	146

3.1.1	Expérimentation et acquisition des données moléculaires et phénotypiques	146
3.1.1.1	Plan d'expérience et choix des populations	146
3.1.1.2	Données phénotypiques	154
3.1.1.3	Données moléculaires	154
3.1.2	Analyses statistiques	155
3.1.2.1	Structure génétique	155
3.1.2.2	Différenciation phénotypique	156
3.2	Résultats	157
3.2.1	Organisation de la diversité génétique	157
3.2.1.1	Les parents	157
3.2.1.2	Les descendants	159
3.2.2	Organisation de la diversité intra-germplasm	169
3.2.3	Description des descendants au niveau phénotypique	180
3.2.3.1	Les types de germplasm	180
3.2.3.2	Les germplasms	181
3.2.3.3	Les sélections	182
3.2.3.4	Les fermes	182
3.3	Discussion	185
3.3.1	Évolution de la diversité dans le programme de sélection participative	186
3.3.1.1	Organisation de la diversité génétique	186
3.3.1.2	Analyse phénotypique	192
3.3.2	La place du programme de sélection participative dans la gestion des ressources génétiques	193
3.3.3	Perspectives	195
4	Description et analyse de la méthodologie mise en place	197
4.1	Organisations sociales dans lesquelles s'inscrivent les acteurs	198
4.1.1	Présentation des différentes organisations sociales	198
4.1.2	L'objet frontière « biodiversité »	199
4.2	Co-construction de la méthodologie	200
4.2.1	Naissance du projet	201

4.2.2	Description de l'évolution de la co-construction à chaque étape	202
4.2.3	Analyse de la co-construction	226
4.3	Innovations organisationnelles et génétiques	230
4.3.1	Le réseau d'échange au centre de l'innovation	230
4.3.2	Les innovations sur mesure	233
4.4	Evolution des pratiques des acteurs	236
4.4.1	Les « paysans »	236
4.4.2	Le groupe « blé RSP »	243
4.4.3	L'équipe « DEAP » : analyse de ma posture de recherche par rapport à l'action	244
5	Transition vers la création d'un système semencier adapté à l'agroécologie	249
5.1	Résultats principaux de la thèse et propositions au collectif de travail	250
5.1.1	L'analyse phénotypique	251
5.1.1.1	Apports méthodologiques	251
5.1.1.2	Les résultats dans le cadre du projet	252
5.1.2	L'analyse d'évaluation du programme de sélection participative	254
5.1.3	Le réseau d'acteurs et la co-construction entre acteurs au coeur de l'innovation	255
5.1.4	Retours sur les objectifs de départ de la thèse et perspectives	257
5.2	Propositions pour une transition au niveau local et régional	258
5.2.1	La diversité au sein des maisons des semences paysannes	259
5.2.2	Proposition d'un modèle flexible de sélection et de gestion de la diversité pour les maisons des semences paysannes	261
5.2.2.1	Prendre le temps de la construction entre les acteurs	261
5.2.2.2	Adapter les outils pour répondre aux divers besoins des acteurs	263
5.2.2.3	Maintenir un réseau de maisons des semences paysannes	267
5.3	Propositions au niveau national pour accompagner la transition	272
5.3.1	Former les animateurs	273

5.3.2 Créer un espace juridique pour les variétés issues du programme de sélection participative	274
5.3.3 Allouer des moyens financiers pour les maisons des semences paysannes	276
<b>Conclusion</b>	<b>279</b>
<b>Bibliographie</b>	<b>281</b>
<b>Table des matières</b>	<b>301</b>









UNIVERSITÉ PARIS-SUD  
ÉCOLE DOCTORALE : SCIENCES DU VÉGÉTAL

Laboratoire de génétique végétale du Moulon

**DISCIPLINE : génétique**

**THÈSE DE DOCTORAT**

Soutenue le 15 janvier 2014

par

**Pierre RIVIÈRE**

---

**Méthodologie de la sélection décentralisée et  
participative : un exemple sur le blé tendre.**

---

**Annexes**





# Table des Annexes

A	Données météorologiques dans les fermes	5
B	Exemples de plan de semis dans les fermes	11
	B.1 Fermes régionales	11
	B.2 Fermes satellites	12
C	Parents des populations issues de croisements cultivées dans le projet	13
D	Fiche climat	25
E	Fiches de semis	27
F	Fiches de sorties d'hiver	39
G	Fiches de printemps	47
H	Fiches d'été	63
I	Protocoles	71
	I.1 Notice pour les récoltes des épis des essais et renvoi des sacs des fermes au Moulon	71
	I.2 Visite des fermes Régionales et mesures phénotypiques au champ	74
	I.3 Mesures dans la salle technique	77
J	Réponse à la sélection massale paysanne en génération précoce dans des populations de blé tendre issues de croisements	81
K	Modélisation hiérarchique de la variance pour des essais déséquilibrés	119
L	Modélisation hiérarchique des interactions génotype × environnement	165
M	Mise en place d'une méthodologie de sélection participative sur le blé tendre en France	217
N	Les statistiques bayésiennes	233
	N.1 Modélisation bayésienne	233
	N.2 Calcul de la distribution a posteriori	234

O	Décomposition des covariances entre individus dans une population consanguine	237
	O.1 Le modèle génétique en population panmictique	237
	O.2 Le modèle génétique en population consanguine : effet du régime de reproduction	239
	O.3 Décomposition de la covariance entre deux individus d'une population consanguine	242
P	Marqueurs utilisés dans l'analyse moléculaire	247
	P.1 Marqueurs dans des zones neutres	247
	P.2 Marqueurs situés dans des gènes candidats	249
Q	Fiche sur les croisements	251
R	Le dossier retour	261
S	Répartitions des tâches entre les paysans, les associations et le Moulon dès 2012	273

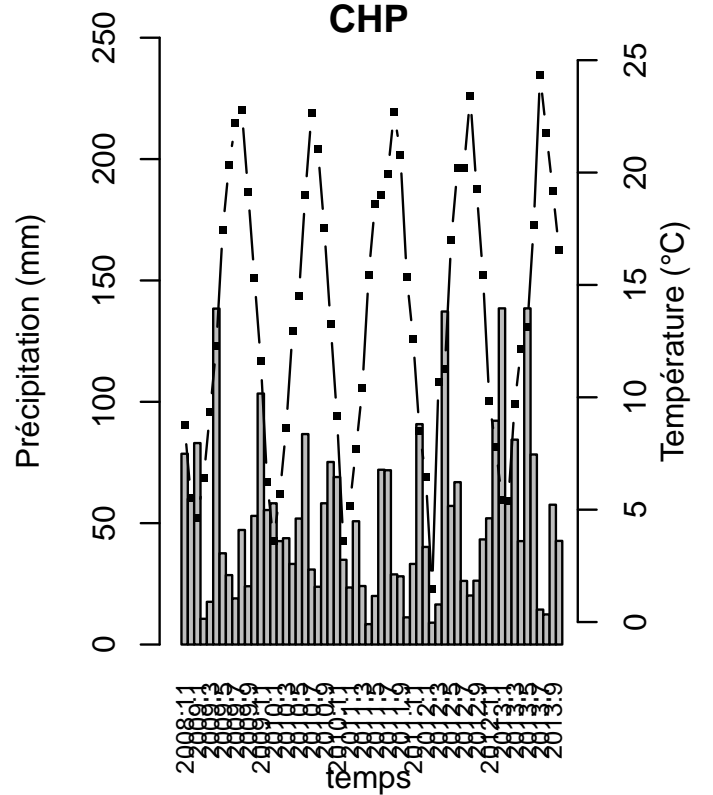
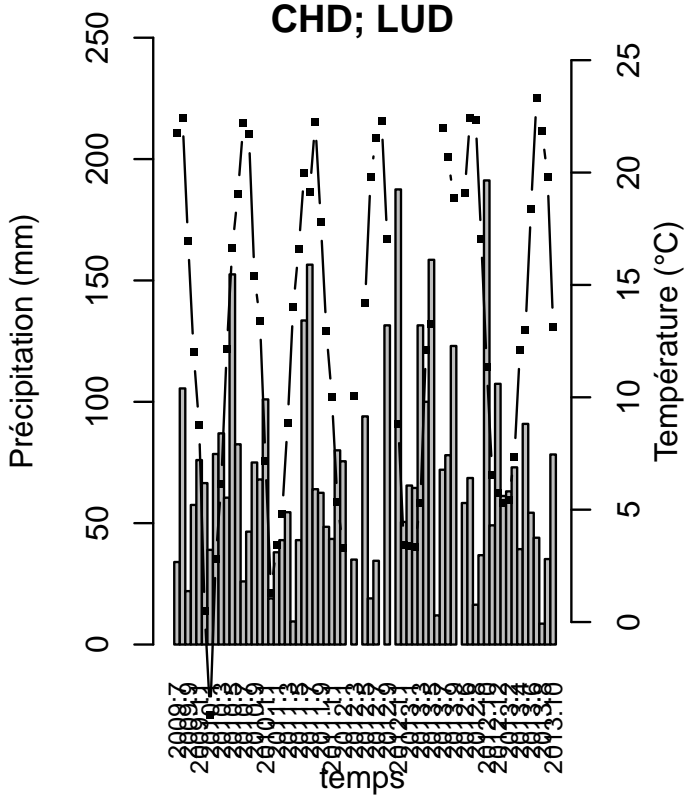
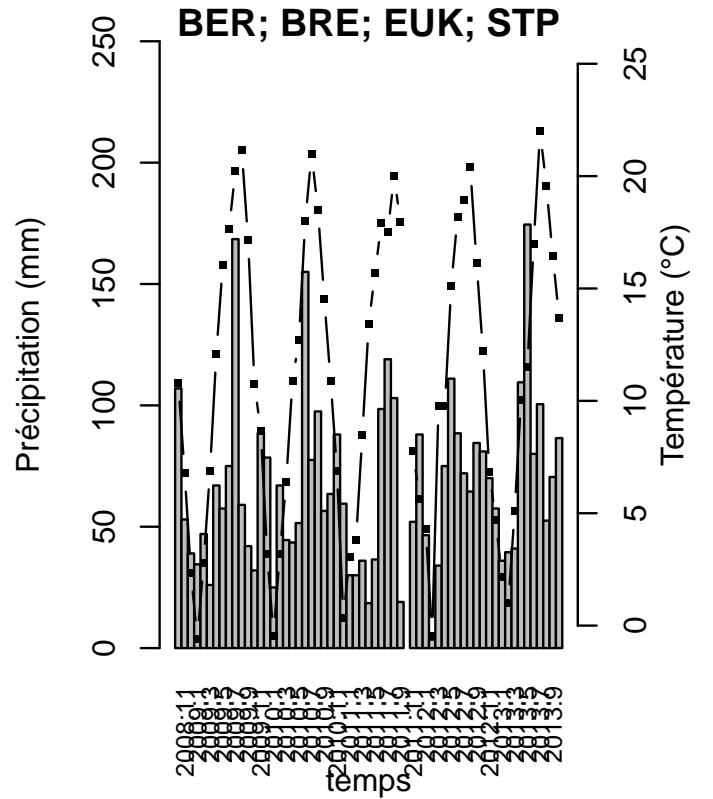
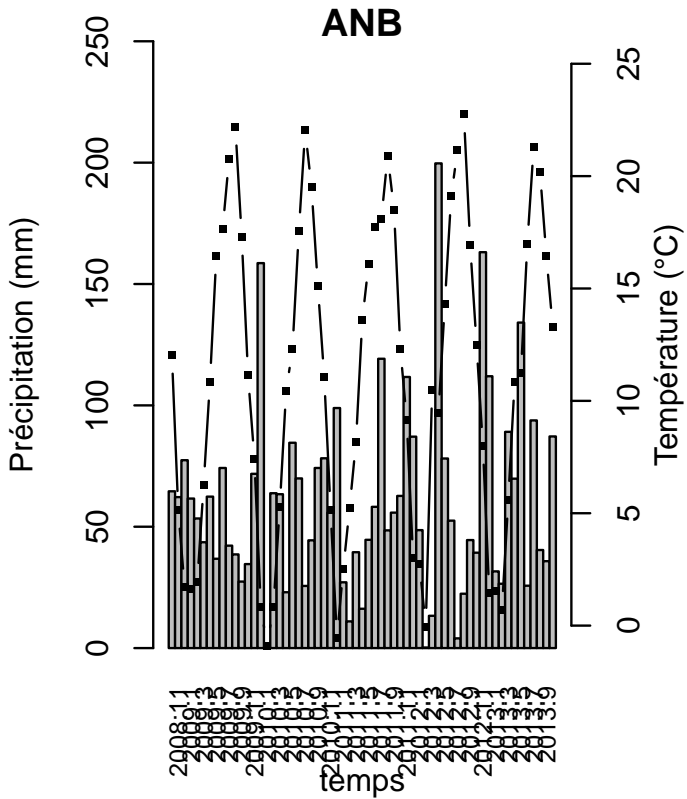
# Annexe A : Données météorologiques dans les fermes

Les données météorologiques ont été récupérées par l'intermédiaire des stations INRA les plus proches des fermes. Même si ces données ne sont pas les plus représentatives de la météo dans les fermes, elles donnent une idée de la diversité des conditions climatiques rencontrée dans le réseau de fermes.

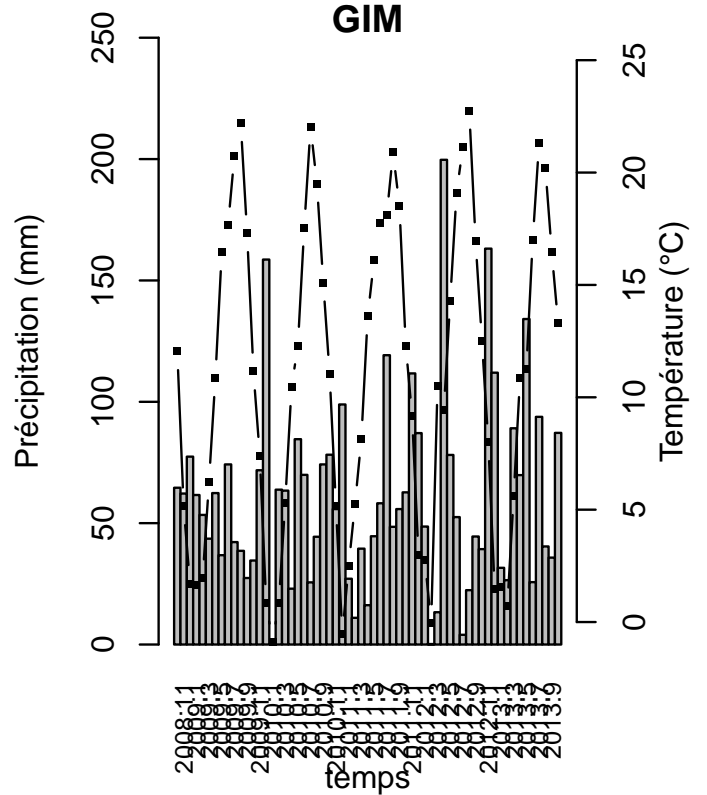
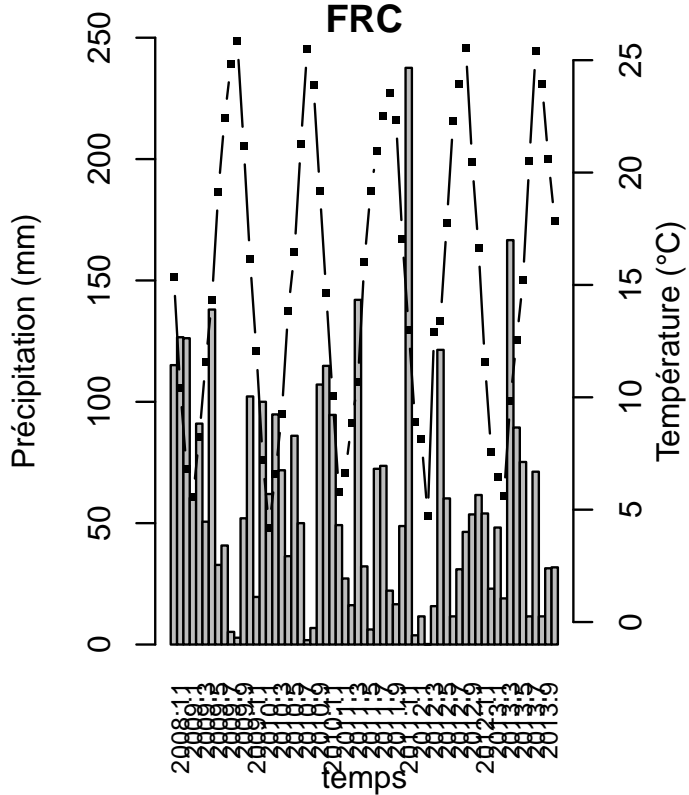
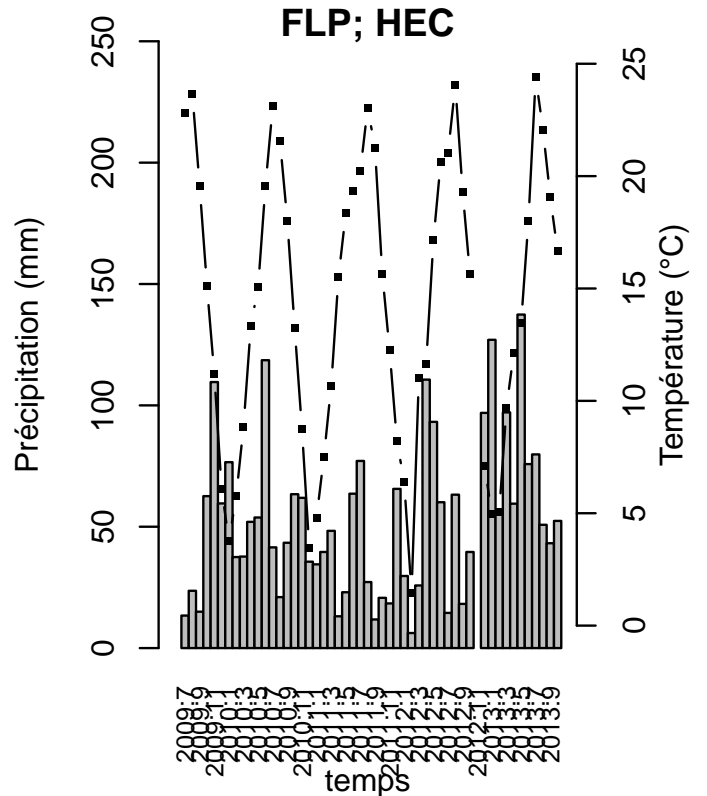
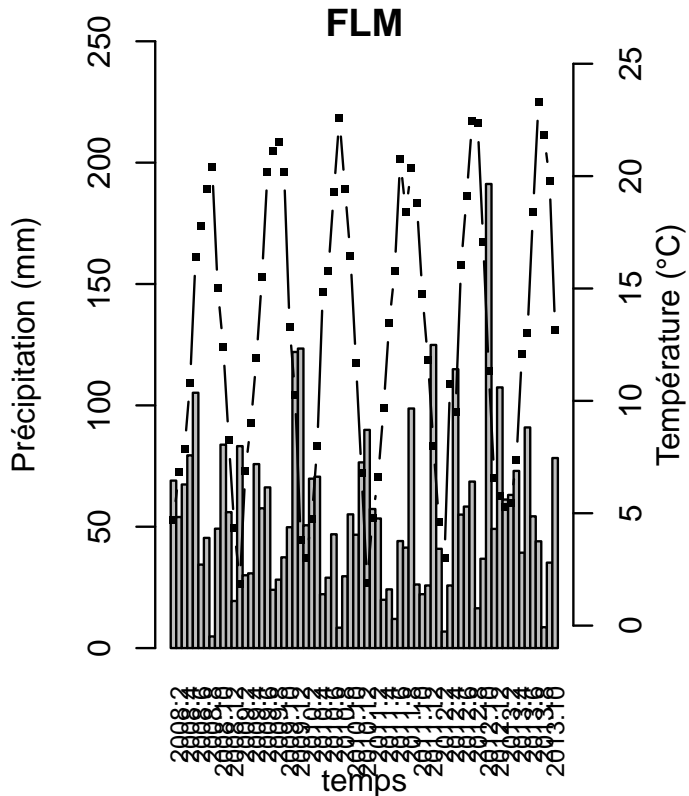
Tableau A.1 - Stations météo les plus proches des fermes

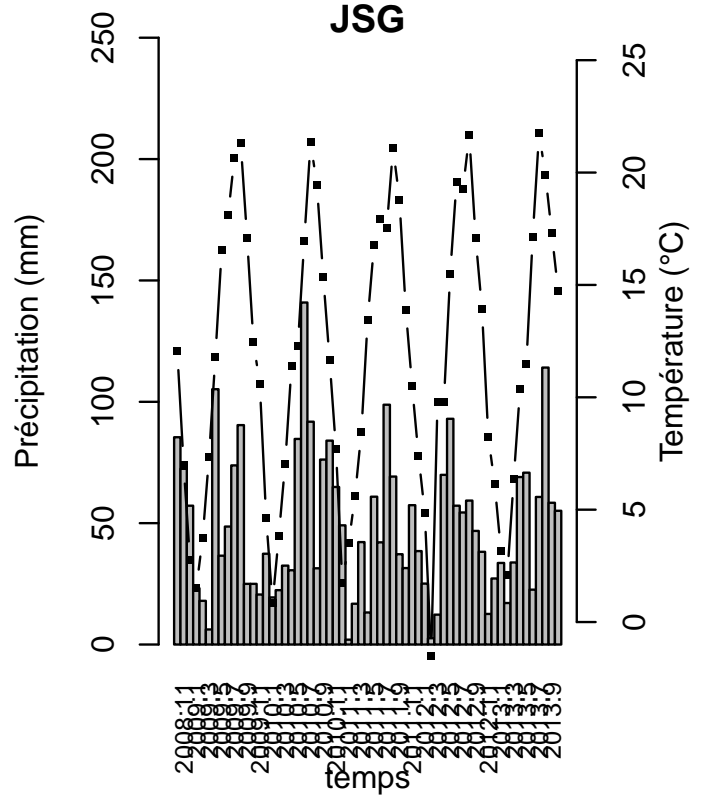
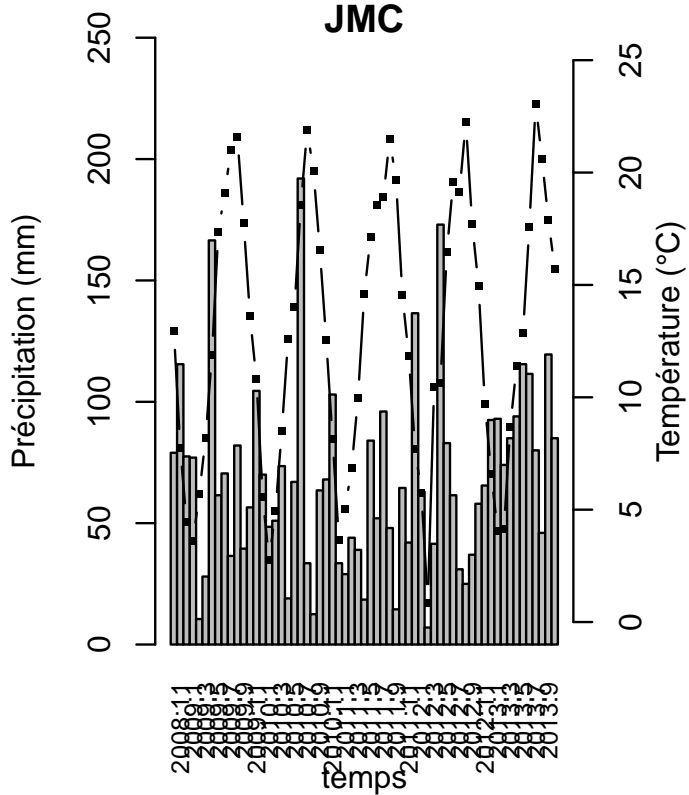
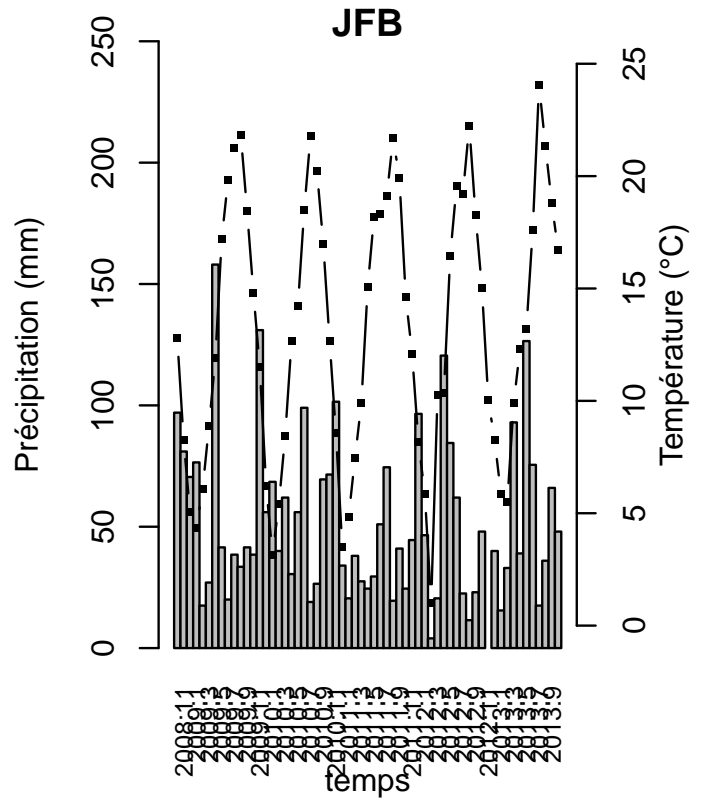
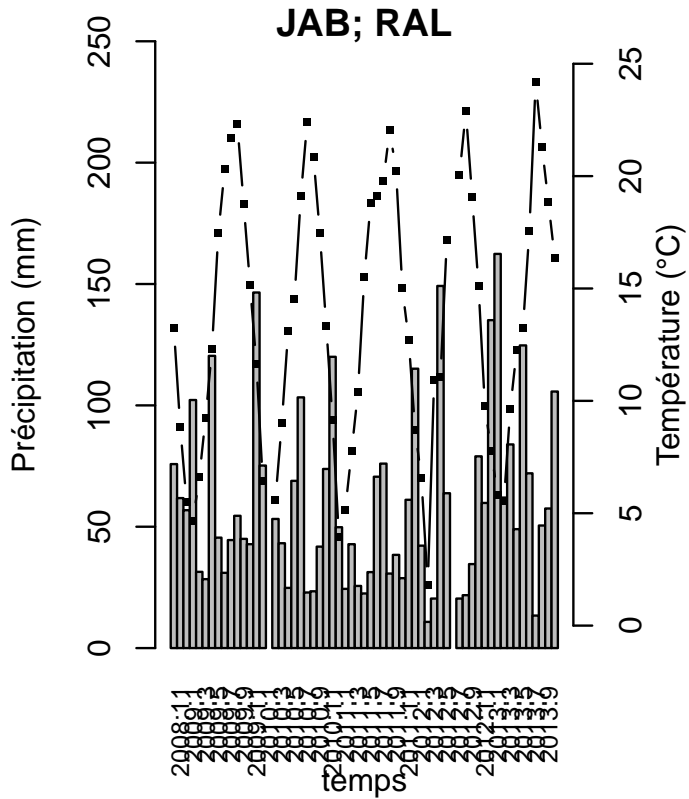
fermes	station météo la plus proche	distance ferme-station météo (km)
ANB ; GIM	Embrun	50 ; 43
BER ; BRE ; EUK ; STP	Bretenières	51 ; 0 ; 12 ; 53
CHD ; LUD	Chatte	56 ; 24
CHP	Estillac	29
FLM	Beaucouze	8
FLP ; HEC	Montauban	58 ; 45
FRC	Bellegarde	16
JAB ; RAL	Marmande	23 ; 29
JFB	Bourran	14
JMC	Creysse	15
JSG	Clermont-Ferrand	9
JUB	Naizin	63
MPH	Mont de Marsan	22
OLR	Thiverval-Grignon	61
RIH ; THG	Entzheim	4 ; 69
RAB	Thonon -les-Bains	35
ROW	Champhol	49

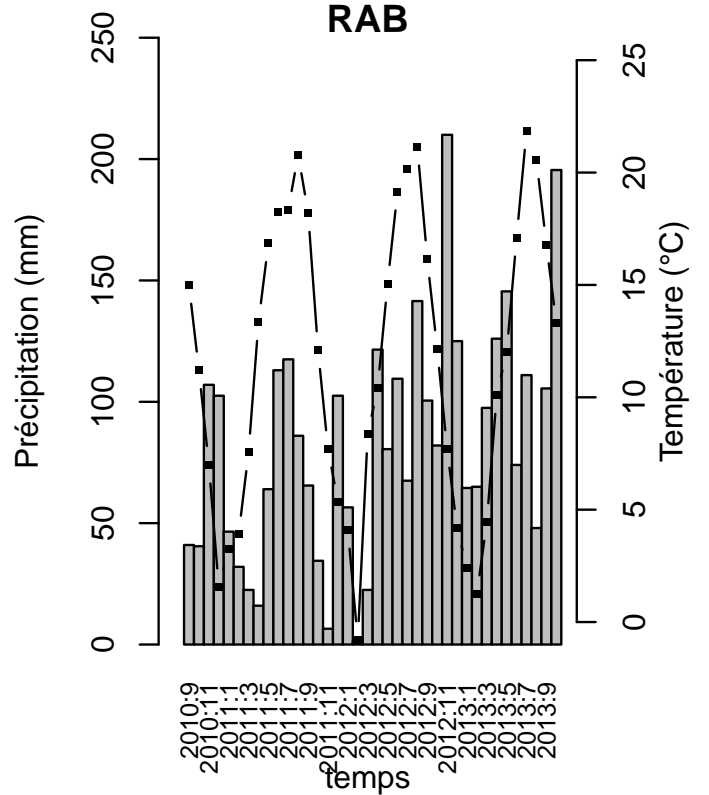
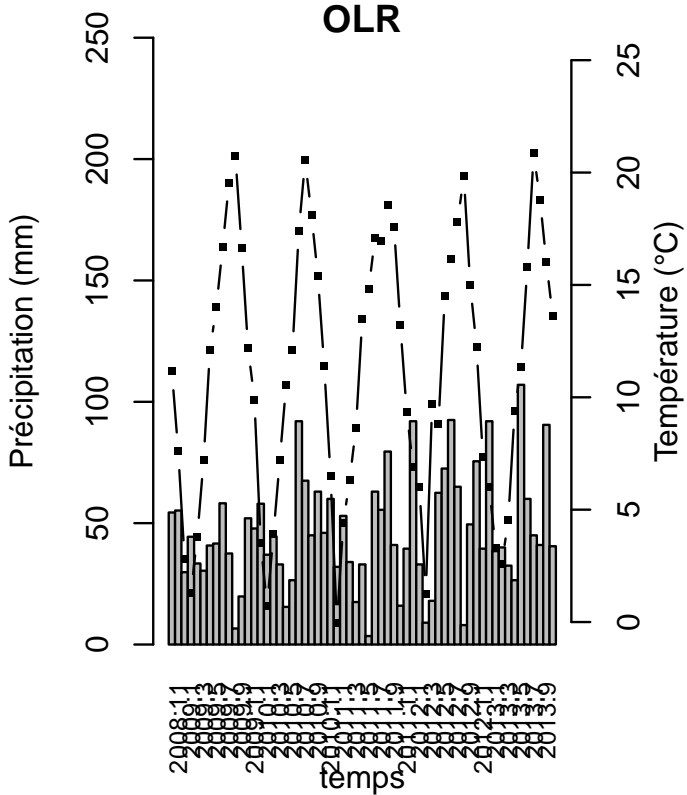
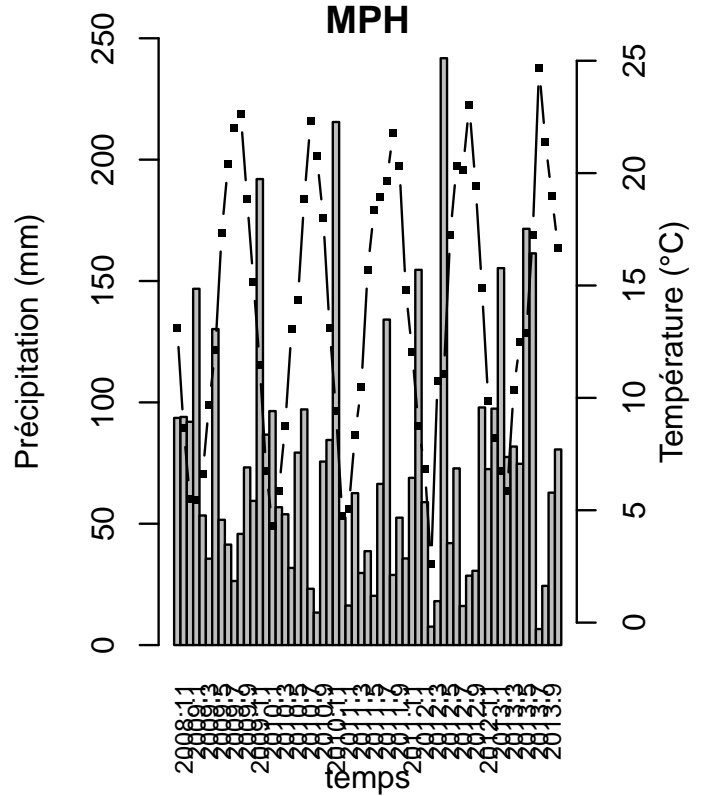
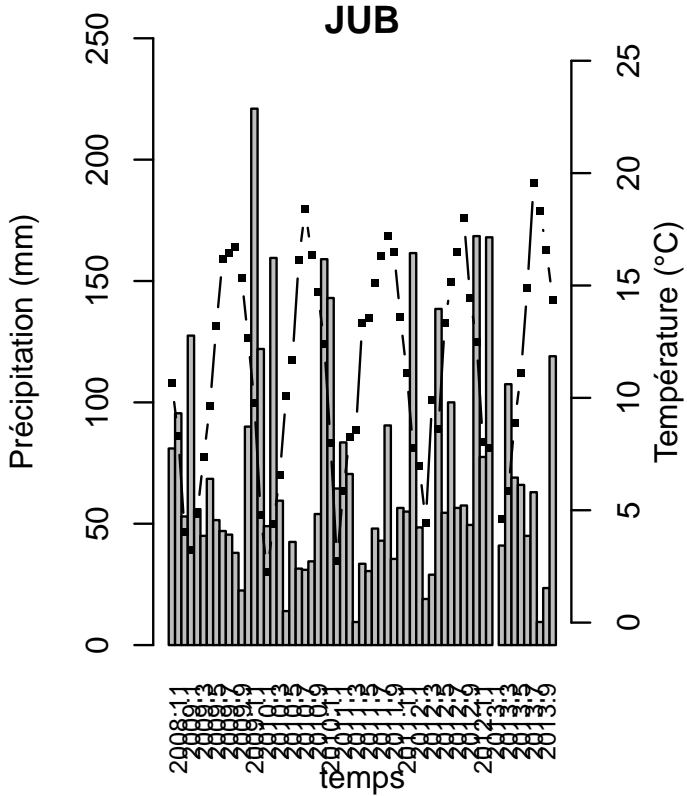
Uniquement les fermes à moins de 70 km des stations météo ont été retenues.

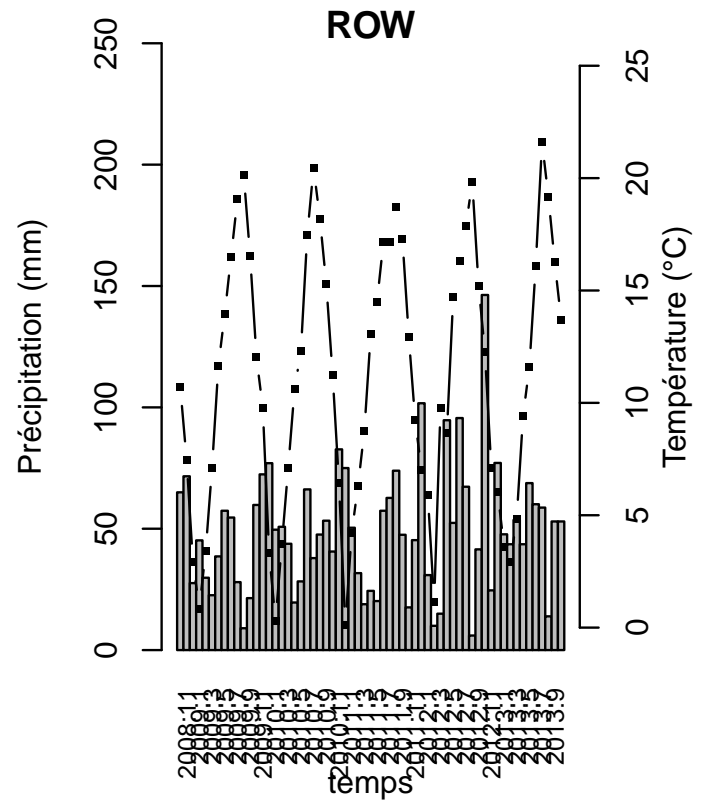
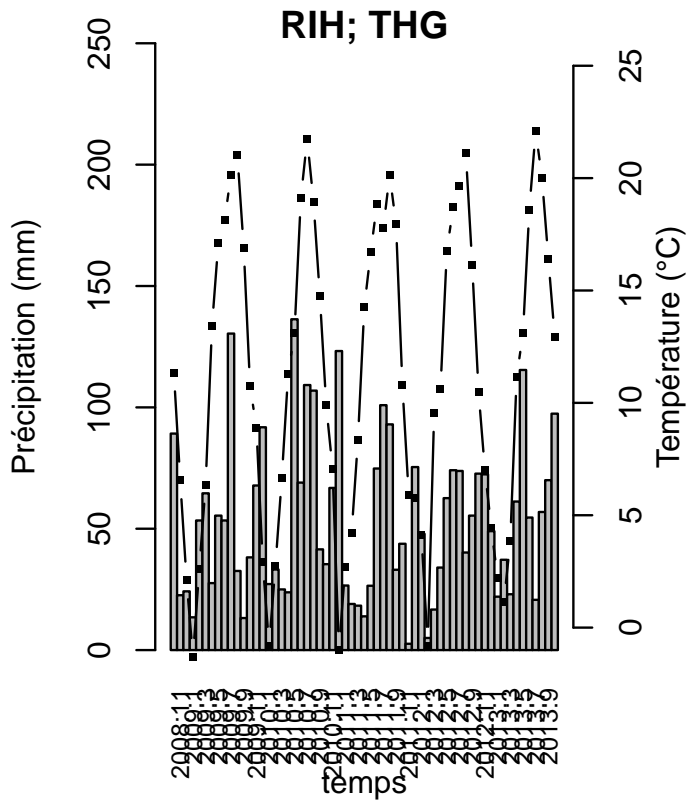












# Annexe B : Exemples de plan de semis dans les fermes

## B.1 Fermes régionales

<b>A.</b>	Rouge-du-Roc rep 1	C7	C20	Renan rep 1
	melange des croisements	C21 rep 1	C5	C48
	X	X	C14 rep 1	X
	Renan rep 2	C81	rietl	C16
	C64	melange de Loex	Rouge-du-Roc rep 2	C34
	C21#Sb	C14 rep 2	F2 (C162, C163, C164, C165, C166)	C21 rep 2
<b>B.</b>	C21#Sb	C21 rep 2	C90#C	C90
	C14 rep 2	C60	C60#D	Renan rep 2
	F2 (C173, C174, C175)	F2 (C176, C177, C178)	Rouge-du-Roc rep 2	X
	C(4+18)	C(4+18)#C	Rouge-du-Roc rep 1	C86
	C14 rep 1	C14#H	Rouge-du-Roc#I	C80
	C14#I	C1	C1#C	Renan rep 1
	C8	C21 rep 1	C21#L	C37
<b>C.</b>	C14 rep 1	C24	C50	
	C84	C45	Renan rep 1	
	M(Atrium, Renan)	Rouge-du-Roc rep 1	C6	
	C21 rep 1	C64	M(JSJG-2010)	
	C58#S	Renan rep 2	C56	
	C5	C60	C14 rep 2	
	C21 rep 2	C42	C32	
	C46	Rouge-du-Roc rep 2	X	

Figure B.1 - Exemple de trois dispositifs expérimentaux dans les fermes régionales (A., B. et C.). Les témoins sont sur fond noir.

## B.2 Fermes satellites

A.	Rouge-du-Roc rep 1	C32
	C42	C42#D
	C76	C89
	C21	Touzelle Anone X Atar
	C34#BR	C21#S
	C63	Rouge-du-Roc rep 2

B.	Rouge-du-Roc rep 1	C33#S-Isa
	C14	C77
	C58	C21
	C93	Rouge-du-Roc rep 2

C.	Rouge-du-Roc rep 1	X
	C60#C	60#S-JF
	C37	C37#F
	C16#F	C37#G
	C16	C16#E
	C50	C50#C
	C50#D	C83#D
	C83	C83#C
	21x3	Rouge-du-Roc#H
	Rouge-du-Roc#G	Rouge-du-Roc rep 2

D.	Rouge-du-Roc rep 1	C60#C
	C37	60#S-JF
	C16	C37#F + C37#G
	C50	C16#F + C16#E
	C83	C50#C + C50#D
	21x3	C83#D + C83#C
	Rouge-du-Roc#G	Rouge-du-Roc rep 2

Figure B.2 - Exemple de quatre dispositifs expérimentaux dans les fermes satellites (A., B., C. et D.). Les témoins sont sur fond noir.

# Annexe C : Parents des populations issues de croisements cultivées dans le projet

Tableau C.1 - Parents des populations issues de croisements, date de réalisation et paysan à l'initiative du croisement. NA : Donnée manquante. NOT : croisement qui a échoué. X : le nom du croisement n'est pas donné par souhait du paysan qui l'a réalisé VP : Variété de Pays, VA : Variété Ancienne, VM : Variété Moderne créée pour l'agriculture biologique, C : population issue de croisement.

croisement nom	père type nom	mère		nom du paysan	année du croisement
		type	nom		
Clomone	VPxVA Touselle-Anone	VP	Cloche	JFB	2004
Clorichelle	VPxVP Richelle	VP	Rouge-Clourauz	JFB	2004
Rock & Lot	VPxVA Blé-du-Lot	VP	Rouge-du-Roc	JFB	2004
X	VPxVP Blé-du-Lot	VP	Rouge-de-Bordeaux	JFB	2006
X	NAxVM Inconnu1	NA	Poncheau	JFB	2006

croisement		père		mère		nom du paysan		année du croisement	
nom	type	nom	type	nom	type	nom du paysan	type	année du croisement	type
X	NAxVP	Inconnu2	NA	Blé-du-Mesnil	VP	JFB	VP	2006	VP
X	NAxVP	Inconnu3	NA	Rouge-de-Bordeaux	VP	JFB	VP	2006	VP
X	NAxVM	Inconnu4	NA	PGR-28s	VM	JFB	VM	2006	VM
X	VPxVP	Richelle	VP	Bladette-velue	VP	JFB	VP	2006	VP
X	VMxVA	Rousselin	VM	Gros-bleu	VA	JFB	VA	2006	VA
X	VPxVM	Touselle-Anone	VP	Atar	VM	JFB	VM	2006	VM
X	VPxVA	Touzelle-Rouge-de-la-Drome	VP	Marquis	VA	JFB	VA	2006	VA
X	NAxVP	Inconnu5	NA	Bladette-de-Besplas	VP	JFB	VP	2006	VP
X	NAxVP	Meunier-d-Apt	NA	Blé-du-Mesnil	VP	JFB	VP	2006	VP
X	VPxVA	Rallet	VP	Cloche	VA	JFB	VA	2006	VA
X	VPxVP	Inversible	VP	Petanielle-Noire-de-Nice	VP	JFB	VP	2006	VP
X	VPxVP	Rouge-d-Alsace	VP	Richelle-hative-à-gros-grain	VP	JFB	VP	2006	VP
X	VPxVP	Barbu-de-l-Aveyron	VP	Rouge-de-Bordeaux	VP	JFB	VP	2006	VP
X	VAxVP	Jade-de-Tourneba	VA	Blé-de-Haute-Loire	VP	JFB	VP	2006	VP
X	NAxVP	Inconnu7	NA	Touzelle-Rouge-de-la-Drome	VP	JFB	VP	2006	VP
X	VPxVM	Blé-des-Hautes-Pyrénées	VP	Pollux	VM	JFB	VM	2006	VM
X	VPxVP	Caucasse-Rouge	VP	Inversible-Blanc	VP	JFB	VP	2006	VP
X	VPxVM	Moulon-Rouge	VP	Rendor	VM	JFB	VM	2006	VM
X	VPxVM	Blanc-de-la-Reole	VP	Lutèce	VM	JFB	VM	2006	VM



croisement		père		mère		nom du paysan		année du croisement	
nom	type	nom	type	nom	type	nom du paysan	type	année du croisement	type
X	VMxVM	Véga	VM	Poncheau	VM	JFB	VM	2006	VM
X	NAxVA	Inconnu8	NA	Magdaléna	NA	JFB	VA	2006	VA
X	VAxVP	Aubaine-Blanche	VA	Blé-du-Lot	VA	JFB	VP	2006	VP
X	VPxVM	Rouge-de-Mongolie	VP	Capitole	VP	JFB	VM	2006	VM
X	VAxVP	Touselle-barbue	VA	Automne-Rouge	VA	JFB	VP	2006	VP
X	VMxVM	Poncheau-selection-fermier	VM	Alauda	VM	JFB	VM	2006	VM
X	VPxVA	Rouge-de-Bordeaux	VP	Marquis	VP	JFB	VA	2006	VA
X	VPxVM	Blé-des-Hautes-Pyrénées	VP	Atar	VP	JFB	VM	2006	VM
X	VPxVM	Blé-du-Lot	VP	Véga	VP	JFB	VM	2006	VM
X	NAxVA	Inconnu6	NA	Gros-bleu	NA	JFB	VA	2006	VA
X	VPxVA	Rouge-de-Mongolie	VP	Japhet	VP	JFB	VA	2006	VA
X	VMxVA	Rousselin	VM	Gros-bleu	VM	JFB	VA	2006	VA
X	NAxVP	Inconnu15	NA	Blé-des-Hautes-Pyrénées	NA	JFB	VP	2006	VP
X	NAxVP	Inconnu16	NA	Richelle-hative-à-gros-grain	NA	JFB	VP	2006	VP
X	VMxVM	Poncheau	VM	Véga	VM	JFB	VM	2006	VM
X	VPxVA	Odesa-sans-barbe	VP	Vilmorin-27	VP	JFB	VA	2006	VA
X	VAxVM	Rock-and-Lot	VA	Véga	VA	JFB	VM	2006	VM
X	VMxVP	Probelle	VM	Rouge-du-Roussillon	VM	JFB	VP	2006	VP
X	VMxVP	Atar	VM	Touselle-Anone	VM	JFB	VP	2006	VP

croisement		père		mère		nom du paysan		année du croisement	
nom	type	nom	type	nom	type	nom du paysan	type	année du croisement	type
X	NAxVA	Inconnu9	NA	Touselle-barbue	VA	JFB	VA	2006	VA
X	VPxVM	Blé-des-Hautes-Pyrénées	VP	Pollux	VM	JFB	VM	2006	VM
X	NAxVP	Inconnu11	NA	Bladette-de-Besplas	VP	JFB	VP	2006	VP
X	VMxVP	Colmar	VM	Touselle-Anone	VP	JFB	VP	2006	VP
X	VPxVP	Blanc-de-la-Reole	VP	Blé-Roumain	VP	JFB	VP	2006	VP
X	NAxVA	Inconnu12	NA	Marquis	VA	JFB	VA	2006	VA
X	NAxVP	Inconnu13	NA	Inversable-Blanc	VP	JFB	VP	2006	VP
X	VPxVM	Coutarel	VP	Lutèce	VM	JFB	VM	2006	VM
X	NAxVP	Inconnu14	NA	Rouge-de-Bordeaux-INRA	VP	JFB	VP	2006	VP
X	VPxVA	Rallet	VP	Cloche	VA	JFB	VA	2006	VA
X	VMxVM	Atar	VM	Poncheau	VM	JFB	VM	2006	VM
X	NAxVA	Cousco	NA	Vilmorin-27	VA	JFB	VA	2006	VA
X	VPxVA	Touzelle-Rouge-de-la-Drome	VP	Magdaléna	VA	JFB	VA	2006	VA
X	VAxVP	Jade-de-Tourneba	VA	Richelle-Blanche-de-Waple	VP	JFB	VP	2006	VP
X	VAxNA	Vilmorin-27	VA	Meunier-d-Apt	NA	JFB	NA	2006	NA
X	NAxVP	Vieux-Ferret	NA	Rouge-du-Clervaux	VP	JFB	VP	2006	VP
X	NAxVP	Inconnu17	NA	Blé-des-Hautes-Pyrénées	VP	JFB	VP	2006	VP
X	VPxVP	Caucasse-Rouge	VP	Inversable-Blanc	VP	JFB	VP	2006	VP
X	VPxVM	Blé-de-l-Aveyron-barbu	VP	Pollux	VM	JFB	VM	2006	VM

croisement		père		mère		nom du paysan		année du croisement	
nom	type	nom	type	nom	type	nom du paysan	type	année du croisement	type
X	NAxVP	Inconnu25	NA	Rallet	VP	JFB	VP	2006	VP
X	VPxVM	Precoce-du-Japon	VP	Probelle	VM	JFB	VM	2006	VM
X	VPxVP	Touselle-Anone	VP	Automne-Rouge-Barbu	VP	JFB	VP	2006	VP
X	NAxVP	Inconnu18	NA	Blé-des-Hautes-Pyrénées	VP	JFB	VP	2006	VP
X	VPxNA	Rouge-de-Bordeaux	VP	Naimara	NA	JFB	NA	2006	NA
X	NAxVP	Inconnu19	NA	Blé-des-Hautes-Pyrénées	VP	JFB	VP	2006	VP
X	NAxVP	Vieux-Ferret	NA	Blé-de-l-Aveyron-barbu	VP	JFB	VP	2006	VP
X	NAxVP	Inconnu20	NA	Inversable-Blanc	VP	JFB	VP	2006	VP
X	VAxVP	Cloche	VA	Rouge-de-Bordeaux-INRA	VP	JFB	VP	2006	VP
X	NAxVA	Inconnu21	NA	Touselle-barbue	VA	JFB	VA	2006	VA
X	VPxVM	Rouge-de-Clairaux-ou-Barbu-de-l'Argnon	VP	Atar	VM	JFB	VM	2006	VM
X	VPxVM	Blé-de-Longogne	VP	Roisel	VM	JFB	VM	2006	VM
X	NAxVA	Inconnu22	NA	Cloche	VA	JFB	VA	2006	VA
X	VPxVP	Blé-des-Hautes-Pyrénées	VP	Blé-du-Mesnil	VP	JFB	VP	2006	VP
X	VPxVP	Rouge-de-Bordeaux	VP	Rouge-de-Mongolie	VP	JFB	VP	2006	VP
X	NAxVM	Inconnu10	NA	Atar	VM	JFB	VM	2006	VM
X	VPxVM	Barbu-de-l-Aveyron	VP	Atar	VM	JFB	VM	2006	VM
X	VPxVP	Redon	VP	Bladette-de-Besplas	VP	JFB	VP	2006	VP
X	NAxVP	Inconnu23	NA	Inversable-Blanc	VP	JFB	VP	2006	VP

croisement		père		mère		nom du paysan		année du croisement	
nom	type	nom	type	nom	type	nom du paysan	type	année du croisement	type
X	VAxVP	Japhet	VA	Blé-des-Hautes-Pyrénées	VP	JFB	VP	2006	VP
X	VMxVP	Atar	VM	Rouge-de-Bordeaux	VP	JFB	VP	2006	VP
X	VPxVA	Blé-Roumain	VP	Aurelie-Gaby	VA	JFB	VA	2006	VA
X	VPxVP	Inversible-Blanc	VP	Touzelle-Rouge-de-la-Drome	VP	JFB	VP	2006	VP
X	NAxVP	Inconnu24	NA	Rouge-de-Bordeaux-INRA	VP	JFB	VP	2006	VP
X	NAxVP	Inconnu26	NA	Automne-Rouge	VP	JFB	VP	2006	VP
X	VPxVA	Caucasse-Rouge	VP	Japhet	VA	JFB	VA	2006	VA
X	VAxVM	Aragon	VA	Bondinave	VM	JFB	VM	2006	VM
A17	VMxVP	Alauda	VM	Blanc-de-Lorraine	VP	BER	VP	2010	VP
NOT	VMxVP	Alauda	VM	Blanc-de-Lorraine	VP	BER	VP	2010	VP
A19	VMxVP	Atar	VM	Blanc-de-Lorraine	VP	BER	VP	2010	VP
A7	VMxVP	Atar	VM	Blé-de-la-Saone	VP	BER	VP	2010	VP
A2	VPxVM	Blé-de-la-Saone	VP	Alauda	VM	BER	VM	2010	VM
A6	VPxVP	Blé-de-la-Saone	VP	Caucasse-Rouge	VP	BER	VP	2010	VP
NOT	VPxVM	Blé-de-la-Saone	VP	Atar	VM	BER	VM	2010	VM
A4	VPxVP	Blé-de-la-Saone	VP	Blé-des-Vosges	VP	BER	VP	2010	VP
A3	VPxVP	Blé-de-la-Saone	VP	Goldendrop	VP	BER	VP	2010	VP
A10	VPxVP	Blé-de-la-Saone	VP	Haut-Brionnais-Mutique	VP	BER	VP	2010	VP

croisement		père		mère		nom du paysan		année du croisement
nom	type	nom	type	nom	type	nom du paysan	année du croisement	
A16	VPxVP	Blé-du-Jura	VP	Blanc-du-Morvan	VP	BER	2010	
A15	VPxVP	Blé-du-Jura	VP	Rouge-d-Altkirch	VP	BER	2010	
A14	VPxVA	Blé-du-Jura	VP	Rouge-du-Roc	VA	BER	2010	
A20	VPxVP	Haut-Brionnais-Mutique	VP	Blanc-de-Lorraine	VP	BER	2010	
NOT	NAxVP	Huabey	NA	Blé-de-la-Saone	VP	BER	2010	
A5	NAxVP	Huabey	NA	Blé-de-la-Saone	VP	BER	2010	
A21	NAxVP	Huabey	NA	Blé-de-la-Saone	VP	BER	2010	
A22	VPxVM	Mottin	VP	Alauda	VM	BER	2010	
A23	VPxVP	Mottin	VP	Barbu-du-Maconnais	VP	BER	2010	
A18	VPxVP	Mottin	VP	Blanc-de-Lorraine	VP	BER	2010	
A1	VPxNA	Mottin	VP	Blé-Allemand	NA	BER	2010	
A8	VPxVP	Mottin	VP	Blé-de-la-Saone	VP	BER	2010	
A11	VPxNA	Rouge-d-Altkirch	VP	Blé-Allemand	NA	BER	2010	
A12	VPxNA	Rouge-d-Alsace	VP	Blé-Allemand	NA	BER	2010	
NOT	VPxNA	Rouge-d-Alsace	VP	Blé-Allemand	NA	BER	2010	
A9	VAxNA	Rouge-du-Roc	VA	Blé-Allemand	NA	BER	2010	
A13	VAxVP	Rouge-du-Roc	VA	Blé-du-Jura	VP	BER	2010	
C162	VPxNA	Mottin	VP	Barbot-du-Forez	NA	CHD	2010	
C163	NAxNA	Romanello	NA	Gua	NA	CHD	2010	

croisement		père		mère		nom du paysan		année du croisement	
nom	type	nom	type	nom	type	nom du paysan	année du croisement	nom du paysan	année du croisement
C164	NAxVP	Rosso-delle-Langhe	NA	Rouge-d-Alsace	VP	CHD	2010	CHD	2010
C165	VPxNA	Saissette-de-Provence	VP	Rouge-de-Marchissy	NA	CHD	2010	CHD	2010
C166	NAxNA	Tuelle	NA	Alentejano	NA	CHD	2010	CHD	2010
C167	VPxNA	Barbu-du-Maconnais	VP	Redon-Sixt-sur-Aff	NA	FLM	2010	FLM	2010
C168	NAxVM	Sabando	NA	Alauda	VM	FLM	2010	FLM	2010
C169	NAxVP	Sabando	NA	Barbu-du-Maconnais	VP	FLM	2010	FLM	2010
C170	NAxNA	Sabando	NA	Oulianowska	NA	FLM	2010	FLM	2010
C171	NAxNA	Sabando	NA	Redon-Sixt-sur-Aff	NA	FLM	2010	FLM	2010
C172	VPxVM	Saissette-de-Provence	VP	Alauda	VM	FLM	2010	FLM	2010
C173	VMxVP	Alauda	VM	Blé-de-la-Saone	VP	RAB	2010	RAB	2010
C174	VPxNA	Blé-de-la-Saone	VP	Inconnu27	NA	RAB	2010	RAB	2010
NOT	VPxVP	Blé-de-la-Saone	VP	Mottet-Blanc	VP	RAB	2010	RAB	2010
C175	VPxVP	Blé-de-la-Saone	VP	Mottet-Blanc	VP	RAB	2010	RAB	2010
NOT	VPxVP	Blé-de-la-Saone	VP	Rouge-de-Bordeaux	VP	RAB	2010	RAB	2010
C176	VPxVP	Blé-de-la-Saone	VP	Rouge-de-Bordeaux	VP	RAB	2010	RAB	2010
C177	VPxVP	Mottet-Rouge	VP	Blé-du-Jura	VP	RAB	2010	RAB	2010
C178	VPxNA	Mottet-Rouge	VP	Inconnu28	NA	RAB	2010	RAB	2010
C179	NAxNA	2	NA	1	NA	ROW	2010	ROW	2010

croisement		père		mère		nom du paysan		année du croisement	
nom	type	nom	type	nom	type	nom du paysan	type	année du croisement	type
C180	NAxNA	3	NA	4	NA	ROW	NA	2010	ROW
NOT	VPxVP	Blé-Autrichien	VP	Automne-Rouge	VP	BER	VP	2011	BER
NOT	VPxVP	Blé-Autrichien	VP	Petit-rouge-du-Morvan	VP	BER	VP	2011	BER
C100	VPxVP	Blé-de-la-Saone	VP	Automne-Rouge	VP	BER	VP	2011	BER
C101	VPxVP	Petit-rouge-du-Morvan	VP	Blé-de-la-Saone	VP	BER	VP	2011	BER
C102	VAxVP	Rouge-du-Roc	VA	Rouge-de-Bordeaux	VP	BER	VP	2011	BER
NOT	NAxNA	Poulard-d-Europe-centrale	NA	Blé-dur-de-Calabre	NA	CHD	NA	2011	CHD
NOT	NAxVP	Senatore-Capelli	NA	Poulard-des-Alpes	VP	CHD	VP	2011	CHD
NOT	VPxNA	Rouge-d-Alsace	VP	Rieti	NA	CHD	NA	2011	CHD
C103	NAxVP	Cabazon-de-Valdegoni	NA	Haut-Brionnais	VP	CHD	VP	2011	CHD
C104	VPxNA	Haut-Brionnais	VP	Marat-Barbu	NA	CHD	NA	2011	CHD
C105	VPxVP	Touselle-rouge-des-Hautes-Alpes	VP	Mottet-Blanc	VP	CHD	VP	2011	CHD
C106	NAxNA	Rojo-de-Humanes	NA	Petit-Rouge-de-pays	NA	CHD	NA	2011	CHD
NOT	NAxVP	Gigante-Lampino-de-Najera	NA	Nonette-de-Lausanne	VP	FLM	VP	2011	FLM
C107	VPxVP	Saint-Priest	VP	Bladette-de-Provence	VP	FLM	VP	2011	FLM
C108	NAxVP	Blanco-de-Corella	NA	Nonette-de-Lausanne	VP	FLM	VP	2011	FLM
C108	VPxNA	Nonette-de-Lausanne	VP	Blanco-de-Corella	NA	FLM	NA	2011	FLM
C109	NAxVM	Soandres-Laracha	NA	Alauda	VM	FLM	VM	2011	FLM

croisement		père		mère		nom du paysan	année du croisement
nom	type	nom	type	nom	type		
C110	NAxVM	Involcable	NA	Alauda	VM	FLM	2011
C114	NAxNA	Gigante-Lampino-de-Najera	NA	Blanco-de-Corella	NA	FLM	2011
C114	NAxNA	Blanco-de-Corella	NA	Gigante-Lampino-de-Najera	NA	FLM	2011
C116	NAxNA	Rojo-de-Sabando	NA	Hernani	NA	FLM	2011
C117	VPxVP	Petit-rouge-du-Morvan	VP	Mottin	VP	FLM	2011
C118	NAxVP	Jejar-de-Valencia	NA	Nonette-de-Lausanne	VP	FLM	2011
C119	VPxNA	Nonette-de-Lausanne	VP	Poulard-d-Italie-B198	NA	FLM	2011
C120	NAxNA	Royo-de-Pamplona	NA	Rojo-de-Sabando	NA	FLM	2011
C116	NAxNA	Hernani	NA	Rojo-de-Sabando	NA	FLM	2011
C122	VPxVP	Petit-rouge-du-Morvan	VP	Saint-Priest	VP	FLM	2011
C123	NAxVP	Royo-de-Pamplona	NA	Saint-Priest	VP	FLM	2011
C126	VPxNA	Nonette-de-Lausanne	VP	Turgidum-di-maliani	NA	FLM	2011
C127	NAxNA	Inconnu30	NA	Inconnu29	NA	JMC	2011
C128	NAxNA	Inconnu32	NA	Inconnu31	NA	JMC	2011
C129	NAxNA	Inconnu34	NA	Inconnu33	NA	JMC	2011
C130	NAxNA	Inconnu36	NA	Inconnu35	NA	JMC	2011
C131	VAxVP	Rouge-du-Roc	VA	Barbu-du-Maconnais	VP	BER	2011
C135	VAxVP	Rouge-du-Roc	VA	Blé-de-Langogne	VP	BER	2011



croisement		père		mère		nom du	année du
nom	type	nom	type	nom	type	paysan	croisement
C132	VPxVP	Blé-de-Langogne	VP	Barbu-du-Maconnais	VP	MPH	2011
C133	NAxVP	Murat	NA	Barbu-du-Maconnais	VP	MPH	2011
C134	NAxVP	Murat	NA	Blé-de-Langogne	VP	MPH	2011
C136	NAxVA	Murat	NA	Rouge-du-Roc	VA	MPH	2011
C137	NAxVA	Ridelle-et-BB	NA	Bon-Fermier	VA	ROW	2011
C138	VMxC	Kampman	VM	Clomone	C	ROW	2011
C139	NAxNA	Hermes	NA	Heinpe	NA	ROW	2011
C140	NAxVP	Succes	NA	Miracle	VP	ROW	2011
C140	VPxNA	Miracle	VP	Succes	NA	ROW	2011



# **Annexe D : Fiche climat**

La fiche climat accompagne chaque fiche semis (Annexe E), hiver (Annexe F), printemps (Annexe G) et été (annexe H). La fiche climat présentée était associée à la fiche hiver de 2010-2011.

## Informations générales sur le climat : hiver 2010

**Pluies (par rapport aux normales saisonnières) :**

faibles                       normales                       excès

**Températures (par rapport aux normales saisonnières) :**

froides                       normales                       chaudes

**Accidents climatiques (grêle, gel, orage, inondation, neige, etc.) :**

**Accidents sur la culture (sanglier, etc.) :**

**Informations complémentaires :**

## **Annexe E : Fiches de semis**

Il n'y avait pas de fiche semis entre 2008 et 2010. La fiche présentée est celle de 2010-2011, elle n'a pas évoluée depuis 2010. Il n'y avait pas de notice avec cette fiche. Une notice a été créée en 2013 par Sophie Pin (équipe DEAP) et Florian Géry (stagiaire BTS en 2013 dans l'équipe DEAP). Un fiche sur les ravageurs a également été créée en 2013 par Florian Géry.

## Fiche de renseignements sur la parcelle à retourner

Nom :

Prénom :

### Informations générales sur la parcelle 2010/2011

L'idéal est que toutes les populations aient le même itinéraire technique. Cependant, si des informations sont différentes pour certains croisements, merci de l'indiquer.

Nom de la parcelle :

Taille de la parcelle :

Taille des micro-parcelles :

Topographie :  Pente si oui :  < 3%  de 3 à 10%  > 10% PlateauType de sol :  limons  argiles  argilo-calcaire  sableux  autre : ... ..Sol :  séchant  intermédiaire  hydromorpheBattance :  oui  nonDrainage :  oui  non

Informations complémentaires :

---

## Informations générales sur le semis 2010/2011

Si les informations sont différentes pour certains croisements, merci de l'indiquer.

**Labour avant semis :**       oui       non

**Date de semis :**

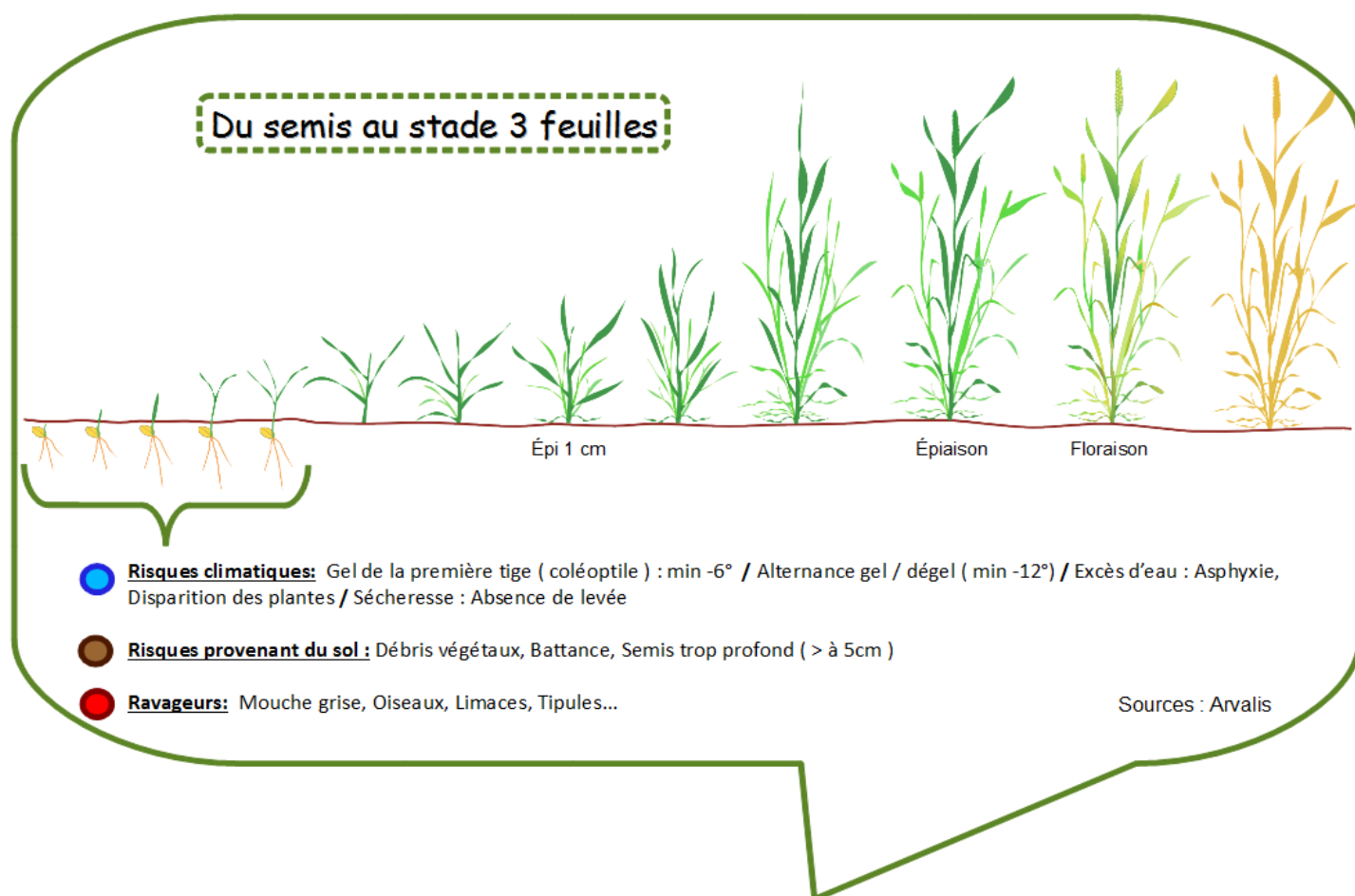
**Méthode de semis :**  volée       semoir

**Nombre de grains au m<sup>2</sup> :**

**Levée :**       lente       rapide      et       régulière       irrégulière

**Précédent cultural :**

**Informations complémentaires :**



**Notice de la fiche**  
**« Automne » 2014**  
**à retourner à votre animateur**

**La fiche Automne permet de recenser toutes les informations sur la parcelle ainsi que sur les populations grâce au plan de semis.**

L'idéal est que toutes les populations aient le même itinéraire technique. Cependant, si des informations sont différentes pour certaines populations, merci de l'indiquer.

**La fiche climat permet de recenser quelques informations sur les conditions climatiques.**



**Plan des micro-parcelles 2013/2014**

Si vous avez modifié le plan de semis des micro-parcelles, merci de bien vouloir nous l'indiquer.



---

## **Reconnaissance des Ravageurs :**

Cette fiche montre quelques ravageurs que l'on peut observer sur le blé lors de son cycle. Cette fiche n'est pas du tout détaillée, elle est simplifiée au maximum ! Elle permet une identification rapide des ravageurs à partir de photos. Quelques idées de lutttes biologiques y sont proposées.

La notation des ravageurs peut se faire dans la partie « commentaire » de la fiche « printemps » et de la fiche « été ». Il ne vous est pas demandé de passer tous les jours dans la parcelle pour faire des observations. Cependant, il est intéressant de voir quel type de ravageur est présent sur les cultures pour une année donnée ainsi que les résistances.

Exemple de commentaires que l'on peut faire sur la fiche : *forte pression de pucerons des épis*

Les informations qui suivent sont tirées du site d'ARVALIS-Institut du végétale *Diagnostic des accidents du blé* et du site Gerbeaud ( <http://www.gerbeaud.com> ).

- I. Les ravageurs du semis jusqu'à la montaison**
  - 1. **La mouche grise**
  - 2. **Les tipules**
  - 3. **Les pucerons**
  - 4. **Les cicadelles**
  - 5. **Les taupins**
  
- II. Les ravageurs des épis**
  - 1. **Les larves tordeuses ( Cnephasia )**
  - 2. **Les Cécidomyies**
  - 3. **Les pucerons des épis**

## I. Les ravageurs du semis jusqu'à la montaison

La levée jusqu'au stade 3 feuilles est une période critique dans le cycle du blé car la jeune pousse est très vulnérable.

Par la suite, la pression des ravageurs diminue.

### 1. La mouche grise

#### Description :

Cette mouche est présente en particulier dans les régions céréalières en Europe et dans toute la France. Ses œufs sont blancs, d'environ 1 mm de diamètre et sont déposés dans la terre au mois de septembre. La larve est blanche nacréée puis devient blanche crémeuse ; elle mesure entre 1mm et 8 mm. Au stade adulte, elle a un corps gris et mesure 6 à 7 mm.

#### Dégâts :

Les larves perforent la tige souterraine et dévorent le bourgeon terminal des jeunes pieds à partir de février jusqu'à mi-avril où elles sortent de la tige et se nymphosent. Les attaques peuvent entraîner la perte de talles voir de la plante.



**Larve de mouche grise** (crédit arvalis)



**Mouche grise** (crédit INRA)

#### Les moyens de luttés :

Éviter les variétés qui sont sensibles au froid, choisir les dates de semis en fonction de l'activité des mouches et semer pas trop profond à densité élevée, éviter les précédents blé (plus de 2 ans), betterave, oignon, pomme de terre...

### 2. Les tipules ( appelés plus couramment « Cousins » )

#### Description :

Les œufs sont noirs, étroits et rigides et mesurent environ 1 mm. La larve de 3 à 4 cm est grise terreuse. Une fois adulte, la tipule fait entre 15 et 25 mm. Il est de couleur brun clair avec un corps et des ailes allongés et de très longues pattes. Son vol est lent et hésitant.

#### Dégâts :

Les larves de tipules sectionnent la partie souterraine de la plante ou la tige. Les plantes meurent par plaques dans la parcelle. On les retrouve souvent à la suite d'un retournement de prairie en milieu humide. La larve monte la nuit. Les attaques débutent à l'automne et se poursuivent au printemps.



**Larve de tipule** (crédit INRA)



**Tipule adulte** (crédit INRA)

Les moyens de luttés :

Travailler le sol en inter-culture, perturber le parasite grâce au travail du sol au cours de l'année (Binage...), éviter les précédents luzerne et céréales. Labourer avant le mois d'août pour détruire le plus de nymphes. Retourner la prairie avant la ponte début septembre afin d'éviter les jeunes larves.

**3. Les pucerons ( *Rhopalosiphum padi* )**

Description :

Puceron mesurant entre 1,5 à 2,3 mm, de couleur vert foncé avec une zone brune rougeâtre à l'extrémité postérieure. On les remarque à l'automne sur les feuilles par beau temps.

Dégâts :

Ce puceron est vecteur de la jaunisse nanifiante des céréales (virus). Les symptômes s'observent à la sortie de l'hiver où les extrémités des feuilles sont décolorées (rougissement pour le blé). Les dégâts peuvent être importants lors d'automne doux et de semis précoces (avant le 15 octobre).



**Les colonies de pucerons** (crédit INRA)

Moyens de lutte :

Éviter les repousses de blés avant le semis et favoriser le développement des auxiliaires comme les coccinelles (haies...). Favoriser les cultures intermédiaires et limiter l'enherbement ; broyer les

#### 4. Les cicadelles ( *Pasmotettix alienus* )

##### Description :

L'adulte est un insecte sauteur de 3,6 à 4,4 mm de couleur brune claire avec des taches foncées et des ailes disposées en toit. La larve ressemble à l'adulte. Elle est présente particulièrement dans le centre de la France. On remarque la présence de l'adulte en octobre.

##### Dégâts :

Cet insecte est vecteur du virus de la maladie des pieds chétifs qui prélève directement la sève. On observe les symptômes à la reprise de la végétation : les plantes atteintes restent naines et ont parfois un fort tallage. Les feuilles sont décolorées et courtes. À partir de février, on observe la mort des plantes atteintes si il y a eu une forte attaque ou alors les plantes ne montent pas ou produisent des épis partiellement stériles.



**Cicadelle** (crédit aramel.free.fr)

##### Moyens de lutte :

Éviter les semis très précoces, les repousses de céréales et de graminées sauvages.

#### 5. Les taupins

##### Description :

Il existe plus de 50 genres de taupins. On les observe à l'aide de pièges.

Les adultes ont un corps très allongé de 6 à 12 mm recouvert d'une fine pubescence grise blanchâtre ; leur tête est large et presque complètement dissimulée par le thorax. La larve mesure entre 17 et 20 mm et possède une tête aplatie avec de courtes mandibules ; son corps jaune pâle est extrêmement dur et résistant.

Les risques d'attaques sont élevés sur des sols riches en matières organiques, après des cultures pérennes et particulièrement les prairies.

##### Dégâts :

La larve est très polyphage et se développe aux dépens des racines de nombreuses cultures ; Sur le blé, la larve pénètre à la base de la plantule et détruit le bourgeon qui entraîne un jaunissement de la feuille centrale et la mort de la plante. L'adulte hiberne dans le sol puis il mange les feuilles des plantes au printemps. Les attaques ont lieu de la sortie hiver jusqu'à l'automne.





**Larve de Taupin** (crédit INRA)



**Taupin adulte** (crédit INRA)

#### Moyens de lutte :

Il existe des pièges à taupins et différents moyens de luttés biologiques :

La rotation des cultures ainsi qu'un apport de chaux sont de bons moyens préventifs. Le purin de fougère semble être un bon répulsif aussi.

Le piégeage est aussi une bonne solution : découpez des pommes de terre en rondelles, plantez-les dans le sol, et relevez vos "pièges" tous les 2 jours. Cette méthode vous permet d'une part de vérifier si votre sol est infesté, et d'autre part de ramasser les larves afin de les détruire

Les auxiliaires à privilégier : Oiseaux, Taupes, Carabes, Musaraignes

On constate de fortes pressions après les précédents prairie, jachère, luzerne (éviter une culture d'automne après retournement d'une prairie).

Un travail du sol régulier pendant l'inter-culture détruit les individus et leur nourriture.

La lutte est difficile car le cycle complet de l'insecte se déroule sur plusieurs années. La lutte doit donc se faire durant plusieurs années.

## **II. Les ravageurs des épis :**

### **1. Les larves tordeuses ( Cnephasia )**

#### Description :

La larve mesure environ 15mm, de couleur orangée avec une tête brune. À la sortie de l'hiver, la jeune larve est transportée par le vent. À l'âge adulte, le papillon mesure entre 16 et 18 mm et est de couleur grise. Il vole à partir de juillet.

#### Dégâts :

La larve pénètre dans la feuille puis attaque les épis. Les épis peuvent être blancs (la tige est sectionnée sous l'épi), mutilés (la larve consomme le contenu des épillets) ou atrophiés (la larve a sectionnée une partie de la tige et la sève est plus ou moins interrompue).



**La larve dévorant un épi** (crédit INRA)



**L'adulte** (crédit INRA)

Moyens de lutte :

Surveiller les cultures de la fin de montaison au gonflement du grain.

Il existe plusieurs moyens de luttés biologiques :

- Le piégeage des adultes mâles grâce à des phéromones qui attirent les papillons mâles au fond d'un piège ou les désorientes
- Les prédateurs naturels : Trichogrammes, certains hyménoptères...
- La culture du maïs diminue les risques pour l'année suivante.

**2. Les Cécidomyies des fleurs de blés**

• **Cécidomyie jaune du blé :**

Description :

La larve est un asticot de 3 mm. L'adulte est une mouche de 2 à 3 mm. Les larves passent l'hiver dans un cocon. La nymphose a lieu début juin et les adultes pondent à l'intérieur des fleurs du blé.

Dégâts :

Les attaques réduisent la vigueur du blé, ce qui engendre une diminution du poids de mille grains et de la capacité germinative, la teneur en gluten le sucre et l'amidon.



**La larve dans l'épillet** (crédit INRA)



**Adulte** (crédit INRA)

- **Cécidomyie orangée du blé**

**Description :**

La larve mesure entre 2 et 3mm de longueur, elle est jaune-orangée. L'adulte a les mêmes caractéristiques mais possède de longues pattes marron clair.

Son cycle de développement dure un an. Il y a donc une génération par an. La larve passe l'hiver dans un cocon enfoui dans le sol puis elle se nymphose au printemps. Les adultes apparaissent durant le mois de juin, s'accouplent et les femelles pondent après le coucher du soleil à l'intérieur des fleurs. Les jeunes larves éclosent 5 à 10 jours après et consomment les grains durant 3 / 4 semaines.

**Dégâts :**

La grosseur et la qualité du blé diminuent car les larves s'alimentent aux dépens des grains.



**La larve sur un grain** (crédit INRA)



**Adulte** (crédit INRA)

**Moyens de lutte :**

Surveiller les cultures du début de l'épiaison jusqu'à la fin floraison au moment des pontes. Les vols ont lieu en soirées par temps lourd et en absence de vent

Le choix de la date de semis, l'incinération des chaumes et l'accroissement de la vigueur de la plante par le travail du sol peuvent être des moyens de lutttes efficaces.

1. **Les Pucerons des épis**

**Description :**

L'adulte mesure de 2 à 3mm, il est allongé et de couleur variable (vert, jaune , rose, brun) ; ils apparaissent dans un premier temps sur les feuilles puis certains vont sur l'épis où ils se développent fortement.

**Dégâts :**

Les pucerons se nourrissent en piquant les épis pour prélever de la sève ; le poids des grains est alors réduit, voir les nombre de grains en cas de fortes attaques.





**Pucerons des épis (crédit INRA)**

Moyens de lutte :

Favoriser des inter-cultures propres avant l'implantation.



# **Annexe F : Fiches de sorties d'hiver**

Il n'y avait pas de notice pour les fiches pour les saisons 2008-2009 et 2009-2010. Une notice est créée dès 2010. Le nom des populations a été ajouté (à la main) pour chaque paysan pour la saison 2009-2010.

Dès 2010, le nom des populations cultivées chez chaque paysan est automatisé.

En 2011, la notice évolue pour prendre en compte une remarque d'un paysan qui ne comprenait pas le port au tallage.

En 2013, la notice est améliorée par Sophie Pin (équipe DEAP) et Florian Géry (stagiaire BTS en 2013 dans l'équipe DEAP).



Tableau de renseignement des parcelles sortie d'hiver 2011!!! La méthode pour remplir le tableau est dans la notice!!!

Croisement	Date d'observation	Globale	Reprise	Densité de peuplement (plantes/m <sup>2</sup> )	Port au tallage	Attitude des feuilles	Commentaires
ALP				<input checked="" type="checkbox"/> 0 à 100 <input type="checkbox"/> 100 à 200 <input type="checkbox"/> 200 à 300 <input type="checkbox"/> 300 à 400 <input type="checkbox"/> > 400	<input type="checkbox"/> dressé <input checked="" type="checkbox"/> demi-dressé <input type="checkbox"/> intermédiaire <input type="checkbox"/> demi-étalé <input type="checkbox"/> étalé	<input type="checkbox"/> horizontale <input type="checkbox"/> de biais <input checked="" type="checkbox"/> verticale	
Exemple	03 04 11	4	3	<input type="checkbox"/> 0 à 100 <input type="checkbox"/> 100 à 200 <input type="checkbox"/> 200 à 300 <input type="checkbox"/> 300 à 400 <input type="checkbox"/> > 400	<input type="checkbox"/> dressé <input type="checkbox"/> demi-dressé <input type="checkbox"/> intermédiaire <input type="checkbox"/> demi-étalé <input type="checkbox"/> étalé	<input type="checkbox"/> horizontale <input type="checkbox"/> de biais <input type="checkbox"/> verticale	Belle population, mais densité faible
219				<input type="checkbox"/> 0 à 100 <input type="checkbox"/> 100 à 200 <input type="checkbox"/> 200 à 300 <input type="checkbox"/> 300 à 400 <input type="checkbox"/> > 400	<input type="checkbox"/> dressé <input type="checkbox"/> demi-dressé <input type="checkbox"/> intermédiaire <input type="checkbox"/> demi-étalé <input type="checkbox"/> étalé	<input type="checkbox"/> horizontale <input type="checkbox"/> de biais <input type="checkbox"/> verticale	
132				<input type="checkbox"/> 0 à 100 <input type="checkbox"/> 100 à 200 <input type="checkbox"/> 200 à 300 <input type="checkbox"/> 300 à 400 <input type="checkbox"/> > 400	<input type="checkbox"/> dressé <input type="checkbox"/> demi-dressé <input type="checkbox"/> intermédiaire <input type="checkbox"/> demi-étalé <input type="checkbox"/> étalé	<input type="checkbox"/> horizontale <input type="checkbox"/> de biais <input type="checkbox"/> verticale	

Nom :

Prénom :

## Fiche de renseignements sortie d'hiver 2011 à retourner

### Notice de la fiche

**La notice permet de savoir quelle note attribuer à quel caractère.** Pour chaque micro-parcelle de croisement de votre essai, pouvez-vous indiquer les caractères suivants : **note globale, note de reprise, note de densité peuplement, note de port au tallage, note sur l'attitude des feuilles, commentaires.**

- **date de notation** : date à laquelle vous effectuez les mesures

- **globale** : la note globale va de 1=moche à 5= magnifique ; cette note sera difficile à comparer entre les paysans mais donne une indication tout de même sur le comportement global de la variété.

- **reprise** : de 1: les plantes n'ont pas bougé depuis la période de froid à 5 : les plantes ont doublé de taille depuis la période de froid

- **densité de peuplement**

1 : de 0 à 100 plantes/m<sup>2</sup>

2 : de 100 à 200 plantes /m<sup>2</sup>

3 : de 200 à 300 plantes /m<sup>2</sup>

4 : de 300 à 400 plantes /m<sup>2</sup>

5 : supérieur à 400 plantes/m<sup>2</sup>

Pour les personnes qui ont semé à la main, une estimation sera parfaite!

- **port au tallage** :

**Indiquer le type de talles majoritaire dans la population pour chaque classe :**

1: dressé

2: demi-dressé

3: intermédiaire

4: demi-étalé

5: étalé

- **attitude des feuilles**

**Indiquer le type majoritaire dans la population pour chaque classe :**

1: horizontale

2: de biais

3: verticale

- **commentaires**

Libre à vous ...

!!! Un exemple est mis dans la première ligne du tableau à remplir !!!

## Fiche de renseignements sortie d'hiver 2012 à retourner

### Notice de la fiche

**La notice permet de savoir quelle note attribuer à quel caractère.** Pour chaque micro-parcelle de croisement de votre essai, pouvez vous indiquer les caractères suivants : **note globale, note de reprise, note de densité peuplement, note de port au tallage, note sur l'attitude des feuilles, commentaires.**

**Date de notation :** date à laquelle vous effectuez les mesures

**Globale :** la note globale va de 1=moche à 5= magnifique ; cette note sera difficile à comparer entre les paysans mais donne une indication tout de même sur le comportement global de la variété.

**Reprise :** de 1: les plantes n'ont pas bougé depuis la période de froid à 5 : les plantes ont doublé de taille depuis la période de froid

#### Densité de peuplement

de 0 à 100 plantes/m<sup>2</sup>

de 100 à 200 plantes /m<sup>2</sup>

de 200 à 300 plantes /m<sup>2</sup>

de 300 à 400 plantes /m<sup>2</sup>

supérieur à 400 plantes/m<sup>2</sup>

Pour les personnes qui ont semé à la main, une estimation sera parfaite!

#### Port au tallage :

**Indiquer le type de talles majoritaire dans la population pour chaque classe :**

dressé

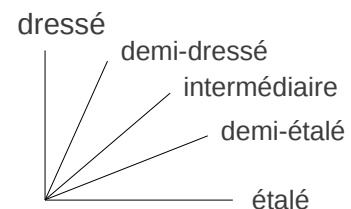
demi-dressé

intermédiaire

demi-étalé

étalé

Les différents ports s'observent comme suit:



#### Attitude des feuilles (assez redondant avec le port au tallage)

**Indiquer le type majoritaire dans la population pour chaque classe :**

horizontale

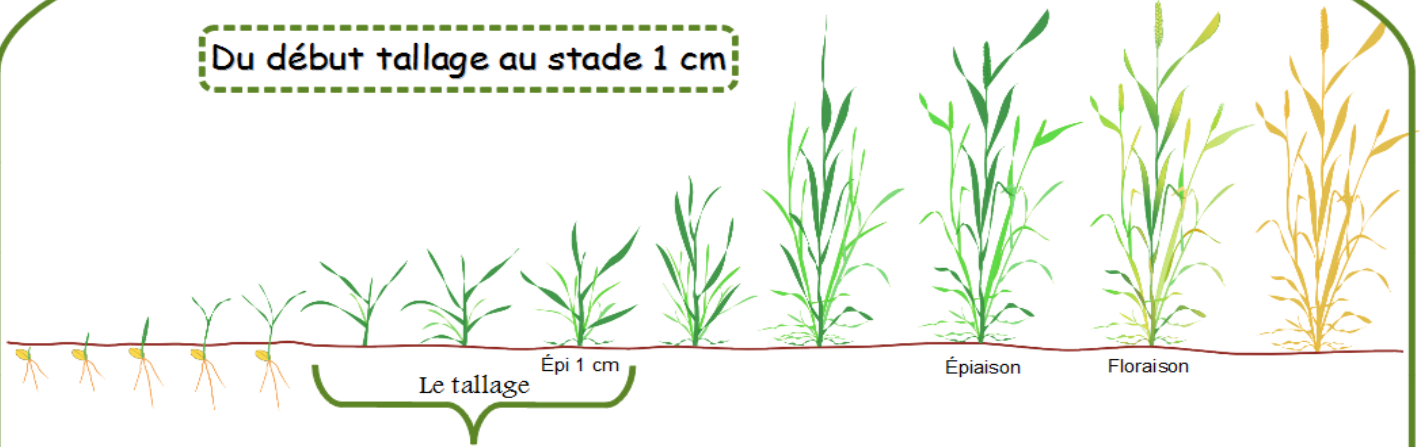
de biais

verticale

**Commentaires :** Libre à vous ... une petite remarque sur la résistance au gel ...

!!! Un exemple est mis dans la première ligne du tableau à remplir !!!

Du début tallage au stade 1 cm



**Risques climatiques:** Alternance gel / dégel ( min -12°) / Excès d'eau : Ralentissement du tallage / Excès en azote : Biomasse trop importante / Carence en azote : Ralentissement de l'émission de talles

**Risques provenant du sol :** Concurrence des adventices

**Maladies et ravageurs:** Pourriture hivernale / Pucerons, Citadelles, Taupins ...

Sources : Arvalis

### Notice de la fiche

### « sortie d'hiver » 2014

### à retourner à votre animateur

La notice permet de savoir quelle note attribuer à quel caractère.

Pour chaque micro-parcelle de croisement de votre essai, pouvez-vous indiquer les caractères suivants :

- **Date de notation :** date à laquelle vous effectuez les mesures
- **Globale :** la note globale va de 1 = moche à 5 = magnifique ; cette note sera difficile à comparer entre les paysans mais donne une indication tout de même sur le comportement global de la variété.
- **Reprise :** de 1= les plantes n'ont pas bougé depuis la période de froid à 5 = les plantes ont doublé de taille depuis la période de froid
- **Densité de peuplement**
  - o De 0 à 100 plantes/m<sup>2</sup>
  - o De 100 à 200 plantes /m<sup>2</sup>
  - o De 200 à 300 plantes /m<sup>2</sup>
  - o De 300 à 400 plantes /m<sup>2</sup>
  - o Supérieur à 400 plantes/m<sup>2</sup>

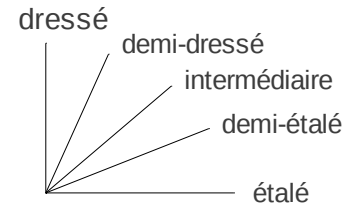
Pour les personnes qui ont semés à la main, une estimation sera parfaite!

- **Port au tallage :**

**Indiquer le type de talles majoritaire dans la population pour chaque classe :**

Les différents port s'observe comme suit :

- dressé
- demi-dressé
- intermédiaire
- demi-étalé
- étalé



Port dressé



Port intermédiaire



Port étalé

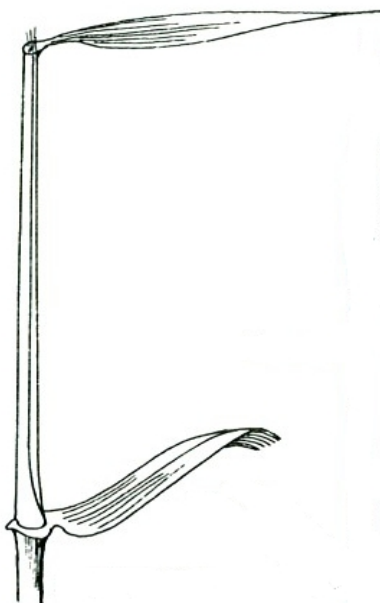
Source : Les cultures céréalières, Jacques Boyeldieu

- **Attitude des feuilles (assez redondant avec le port au tallage)**

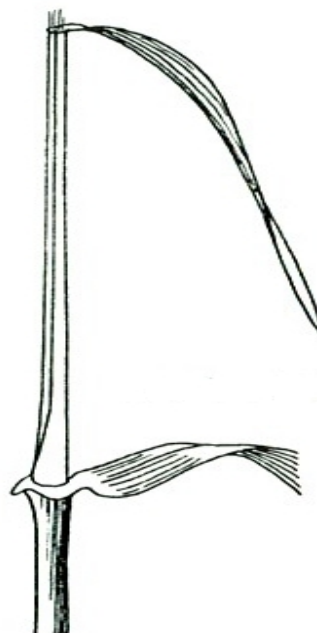
**Indiquer le type majoritaire dans la population pour chaque classe :**

Les différents ports s'observent comme suit :

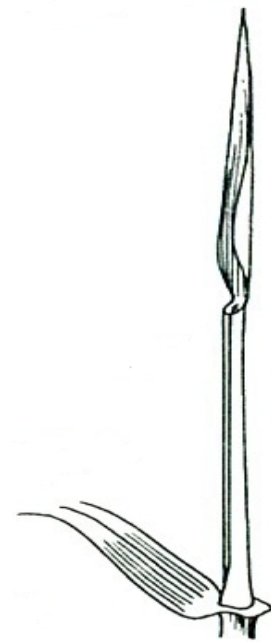
- horizontale
- de biais
- verticale



Horizontale



De biais



Verticale

Source : Les cultures céréalières, J. Boyeldieu

- **Commentaires :**

Libre à vous ... une petite remarque sur la résistance au gel ...

!!! Un exemple est mis dans la première ligne du tableau à remplir !!!





# Annexe G : Fiches de printemps

Il n'y avait pas de notice pour les fiches pour les saisons 2008-2009 et 2009-2010. Le nom des populations a été ajouté (à la main) pour chaque paysan pour la saison 2009-2010.

Dès 2010, le nom des populations cultivées chez chaque paysan est automatisé. La fiche maladie est créée, elle n'évoluera pas.

En 2012, un note sur les maladies est demandée sur la fiche. Auparavant, elles étaient attendues dans la case « Commentaires ; maladies ; épiaison ». La notation de l'épiaison évolue suite à des remarques de paysans qui trouvaient cette notation trop compliquée. La notice évolue également.

En 2013, la notice est améliorée par Sophie Pin (équipe DEAP) et Florian Géry (stagiaire BTS en 2013 dans l'équipe DEAP).

**Croisements du Roc 2010**

Sortie d'épiaison / Échelle de notation précocité : 1=le plus précoce 5= le plus tardif,  
Échelle de notation globale : 1= pas beau 5 = magnifique

**Nom :** \_\_\_\_\_

Autres précisions :

Roc = le témoin rouge de roc, avec deux réplifications

Date d'Observation	Numéro	note globale de précocité	Note globale de la parcelle	Vos observations : notations spécifiques (aspects remarquables du feuillage, biomasse, maladies, verse, enherbage etc)
	Roc Rep 1			
	Roc Rep 2			

Croisement	Date	Globale	Tallage	Vigueur	Couleur	adventices (nombre et type)	Commentaires ; ma- ladies ; épiaison
Exemple	25 05 11	5	<input type="checkbox"/> faible <input type="checkbox"/> moyen <input checked="" type="checkbox"/> fort	1	<input type="checkbox"/> Vert clair <input checked="" type="checkbox"/> Vert <input type="checkbox"/> Vert foncé		
	Date mesure épiaison	3 jours ou plus avant aujourd'hui	avant- hier	aujourd hui	demain après de- main	3 jours ou plus après aujourd'hui	Bla bla bla
	15/05/11	0	0	10	0	70	
	18/05/11	30	0	0	50	20	
	21/05/11	80	0	0	20	0	
Rappel hiver : date : 10/04/11, globale= 2, reprise=3, densité= 300 à 400, Port au tallage : intermédiaire, Attitude des feuilles : horizontale							
Commentaires hiver : belle population mais levée hétérogène							
Croisement	Date	Globale	Tallage	Vigueur	Couleur	adventices (nombre et type)	Commentaires ; ma- ladies ; épiaison
ALP RdR1			<input type="checkbox"/> faible <input type="checkbox"/> moyen <input type="checkbox"/> fort		<input type="checkbox"/> Vert clair <input type="checkbox"/> Vert <input type="checkbox"/> Vert foncé		
	Date mesure épiaison	3 jours ou plus avant aujourd'hui	avant- hier	aujourd hui	demain après de- main	3 jours ou plus après aujourd'hui	
Rappel hiver : date : globale= , reprise= , densité= , Port au tallage : , Attitude des feuilles :							
Commentaires hiver :							

Nom :

Prénom :

## Fiche de renseignements printemps 2011 à retourner

### Notice de la fiche

**La notice permet de savoir quelle note attribuer à quel caractère.** Pour chaque croisement de votre essai, pouvez vous indiquer les caractères suivant : **date de notation, globale, tallage, vigueur, couleur, présence d'adventices, épiaison, commentaires et maladies.**

**date de notation :** date à laquelle vous effectuez les mesures

**globale :** la note globale va de 1=moche à 5= magnifique ; cette note sera difficile à comparer entre les paysans mais donne une indication tout de même sur le comportement global de la variété.

**Tallage :** Indiquer si le tallage est faible, moyen ou fort.

**Vigueur :** note de 1 à 5

**Couleur :** Indiquer la tendance générale de la population : vert clair, vert ou vert foncé.

**Adventices (nombre et type) :** indiquer s'il y a des adventices et si oui, leur nom et leur nombre. Une « fiche adventices » est jointe.

**-Commentaires, maladies, épiaison**

Libre à vous ... vous pouvez indiquer s'il y a des maladies à l'aide de la « fiche maladie » jointe.

En bas du tableau sont reportées les notations qui ont été faites en hiver, pour mémoire.

**Dans le tableau, il y a une partie spéciale dédiée à l'épiaison**

**Epiaison : ATTENTION, LE PLUS COMPLIQUE A NOTER !**

On considère qu'un épi est épié quand la moitié de l'épi est sortie de la gaine (figure 1)



Dans la partie **épiaison** du tableau il faut :

- noter la date d'observation
- pour chaque croisement indiquer si l'épiaison:
  - **a lieu aujourd'hui**
  - **a eu lieu** il y a un jour, deux jour, ...
  - **va avoir lieu** dans un jour, deux jours, ...

**Cette fiche peut être remplie plusieurs fois, avec plusieurs passages** (3 passages paraissent l'idéal : le fait d'aller plusieurs fois sur la parcelle permet d'être plus précis et d'éviter de fausses estimations. Par exemple le 5 mai, je dis que RdR va épié dans 3 jours. Et finalement, quand je retourne au champ le 8 mai, je vois qu'il va épié dans 2 jours ...). **Dans le tableau, on a mis trois lignes qui correspondent à trois passages. Mais un passage sera déjà très bien !**

Figure 1:épi épié  
(source ITAB)

Exemple: Aujourd'hui nous sommes le 15 mai, j'observe que 20% des épis de la population ont été épiés, 10% ont été épiés hier, le reste n'a pas encore été épié et ne sera pas épié avant 3 jours.

Ce qui donne dans la fiche :

Croisement	Date	3 jours ou plus avant aujourd'hui	avant-hier	hier	aujourd'hui	demain	après demain	3 jours ou plus après aujourd'hui
Exemple	15/05/11	0	0	10	20	0	0	70
	18/05/11	30	0	0	50	0	20	0
	21/05/11	80	0	0	0	20	0	0

**!!! ATTENTION : la somme de la ligne doit être égale à 100% !!!**

nb : pour les mesures du 18/05/11, on retrouve les 30% dans la case « 3 jours ou plus avant aujourd'hui », ces 30% avaient été observés le 15/05/11 dans la case « hier » et « aujourd'hui ». Ici c'est un cas idéal, mais parfois, on se trompe, d'où l'intérêt de passer plusieurs fois si possible.

Un exemple est mis dans la première ligne du tableau à remplir.

JUB Tableau de renseignement des parcelles printemps 2012!!! La méthode pour remplir le tableau est dans la notice!!!

Croisement	Date	Globale	Tallage	Vigueur	Couleur		adventices (nombre et type)	Commentaires généraux
Exemple	10 05 12	3	<input type="checkbox"/> faible <input type="checkbox"/> moyen <input checked="" type="checkbox"/> fort	3	<input type="checkbox"/> Vert clair <input checked="" type="checkbox"/> Vert <input type="checkbox"/> Vert foncé			Bla bla bla
<b>Date</b>	<b>0%</b>	<b>0-30%</b>	<b>30-60%</b> <b>60% 100%</b>	<b>100%</b>	<b>sain</b>	<b>malade</b>	<b>très malade</b>	<b>Commentaires ; maladies ; épiaison</b>
10 05 12	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	X	<input type="checkbox"/>	<input type="checkbox"/>	
18 05 12	<input type="checkbox"/>	<input type="checkbox"/>	X	<input type="checkbox"/>	<input type="checkbox"/>	X	<input type="checkbox"/>	
25 05 12	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	X	<input type="checkbox"/>	X	<input type="checkbox"/>	
Rappel hiver : date : 10/04/11 globale= 2 reprise=3 densité= 300 à 400 Port au tallage : intermédiaire Attitude des feuilles : horizontale								
Commentaires hiver : belle population mais levée hétérogène								
Croisement	Date	Globale	Tallage	Vigueur	Couleur		adventices (nombre et type)	Commentaires généraux
Rouge-du-Roc rep 1			<input type="checkbox"/> faible <input type="checkbox"/> moyen <input type="checkbox"/> fort		<input type="checkbox"/> Vert clair <input type="checkbox"/> Vert <input type="checkbox"/> Vert foncé			
<b>Date</b>	<b>0%</b>	<b>0-30%</b>	<b>30-60%</b> <b>60% 100%</b>	<b>100%</b>	<b>sain</b>	<b>malade</b>	<b>très malade</b>	<b>Commentaires ; maladies ; épiaison</b>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Rappel hiver : date : globale= , reprise= , densité= , Port au tallage : , Attitude des feuilles :								
Commentaires hiver :								

## Fiche de renseignements printemps 2012 à retourner

### Notice de la fiche

La notice permet de savoir quelle note attribuer à quel caractère. Pour chaque population de votre essai, pouvez vous indiquer les caractères suivant : **date de notation, globale, tallage, vigueur, couleur, présence d'adventices, épiaison, maladies, commentaires**. Les notations faites en hiver sont indiquées sur la fiche pour mémoire.

**Date de notation** : date à laquelle vous effectuez les mesures

**Globale** : la note globale va de 1=moche à 5= magnifique ; cette note sera difficile à comparer entre les paysans mais donne une indication tout de même sur le comportement global de la variété.

**Tallage** : Indiquer si le tallage est faible, moyen ou fort.

**Vigueur** : note de 1 à 5 (1=pas de vigueur à 5=très bonne vigueur)

**Couleur** : Indiquer la tendance générale de la population : vert clair, vert ou vert foncé.

**Adventices (nombre et type)** : indiquer s'il y a des adventices et si oui, leur nom et leur nombre.

**Commentaires, maladies, épiaison** : Libre à vous ...

**Épiaison et maladies** : **ATTENTION, LE PLUS COMPLIQUÉ ET LONG A NOTER !**

On considère qu'un épi est épié quand la moitié de l'épi est sortie de la gaine (*Figure 1*)



Dans la partie **épiaison** et **maladie** du tableau il faut noter la date d'observation

1) Dans la partie épiaison :

Pour chaque population indiquer si l'épiaison a lieu pour:

- **0%** des plantes de la population
- entre **0 et 30%** des plantes de la population
- entre **30 et 60%** des plantes de la population
- entre **60 et 100%** des plantes de la population
- **100%** des plantes de la population

2) Dans la partie maladie

Pour chaque population indiquer, par rapport aux maladies si la population est :

- **saine**
- **malade**
- **très malade**

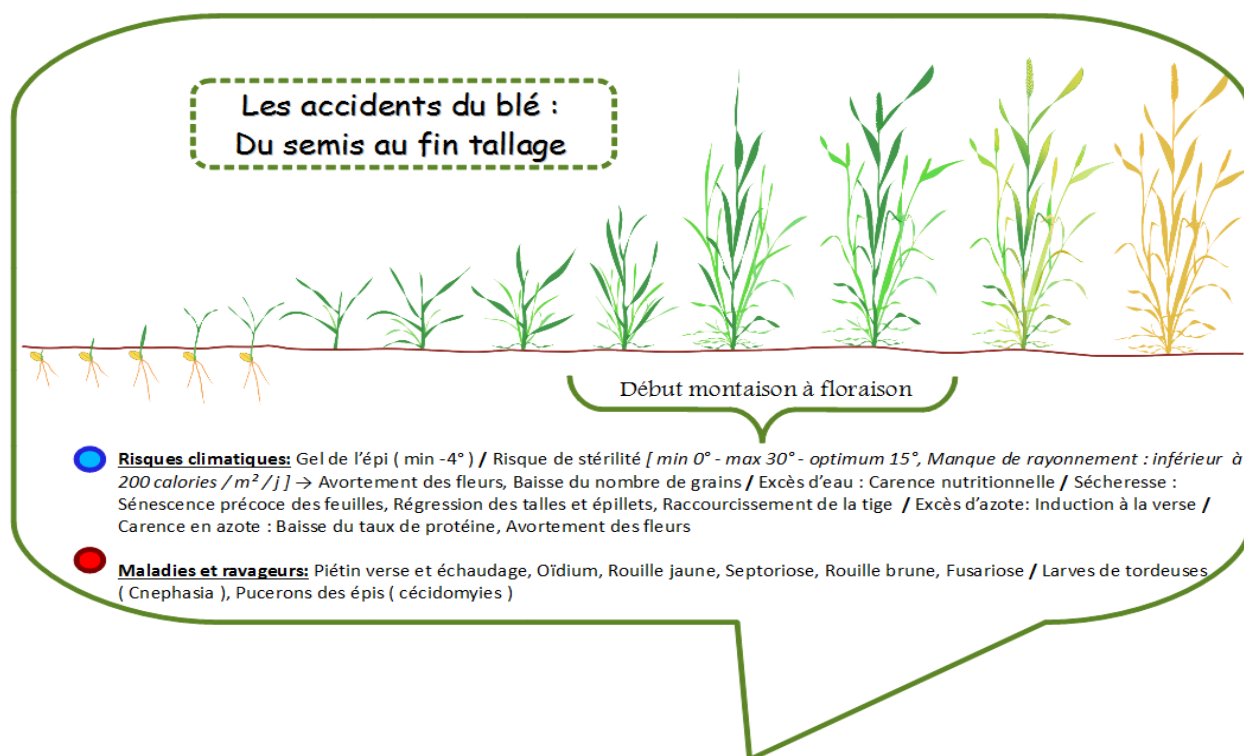
*Figure 1:épi épié*  
(source ITAB)

On pourra se rapporter à la « Fiche reconnaissance des maladies » pour ajouter dans les commentaires des observations.

Cette fiche peut être remplie plusieurs fois, avec plusieurs passages (3 passages paraissent l'idéal, mais un passage serait déjà bien: le fait d'aller plusieurs fois sur la parcelle permet d'être plus précis). Dans le tableau, on a mis trois lignes qui correspondent à trois passages. Mais un passage sera déjà très bien !

Si vous faites des photos, nous sommes preneurs !

!!! Un exemple est mis dans la première ligne du tableau à remplir !!!



**Notice de la fiche**  
**« printemps » 2013**  
**à retourner à votre animateur**

**La notice permet de savoir quelle note attribuer à quel caractère.**

Les notations faites en hiver sont indiquées sur la fiche pour mémoire.

Pour chaque population de votre essai, pouvez-vous indiquer les caractères suivant :

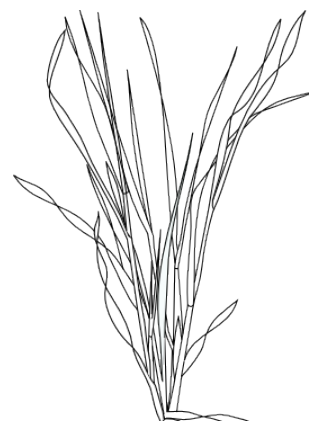
- **Date de notation :** date à laquelle vous effectuez les mesures
- **Globale :** la note globale va de 1=moche à 5= magnifique ; cette note sera difficile à comparer entre les paysans mais donne une indication tout de même sur le comportement global de la variété.
- **Tallage :** Indiquer si le tallage est faible, moyen ou fort.



Faible



Moyen



Fort



- **Vigueur** : note de 1 à 5 (1=pas de vigueur à 5=très bonne vigueur)
- **Couleur** : Indiquer la tendance générale de la population : vert clair, vert ou vert foncé.
- **Adventices (nombre et type)** : indiquez, s'il y a des adventices et si oui, leur nom et leur nombre.
- **Commentaires, maladies, épiaison** : Libre à vous ...
- **Épiaison et maladies** :



**ATTENTION, LE PLUS COMPLIQUE ET LONG A NOTER !**

On considère qu'un épi est épié quand la moitié de l'épi est sortie de la gaine (Dessin ci-dessous).

Dans la partie **épiaison** et **maladie** du tableau il faut noter la date d'observation



Epi épié

Source : ITAB

1) Dans la partie épiaison :

Pour chaque population indiquer si l'épiaison a lieu pour:

- **0%** des plantes de la population
- entre **0 et 30%** des plantes de la population
- entre **30 et 60%** des plantes de la population
- entre **60 et 100%** des plantes de la population
- **100%** des plantes de la population

2) Dans la partie maladie :

Pour chaque population indiquer, par rapport aux maladies si la population est :

- **saine**
- **malade**
- **très malade**

On pourra se rapporter à la « Fiche de reconnaissance des maladies » pour ajouter dans les commentaires des observations.

Date	Tallage		Vigneux	Couleur			adventice (nombre)	Commentaires ; maladies ; épiaison
	faible <input type="checkbox"/>	moyen <input type="checkbox"/>		fort <input checked="" type="checkbox"/>	Vert clair <input type="checkbox"/>	Vert <input checked="" type="checkbox"/>		
Exemple			3					bla bla
Date	30-60%	60-100%	100%					
10.05.12	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
18.05.12	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>					
25.05.12	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>					
Rappel livrer : de 300 à 400 Port au tallage : intermédiaire Attitude des feuilles : horizontale								
Commentaires hi								

**Date** du ou des passages (dans l'exemple 3 passages)

### Partie épiaison

Pour une date donnée, je coche le pourcentage d'épiaison

### Partie maladie

Pour une date donnée, je coche si la micro parcelle est : saine, malade ou très malade

Cette fiche peut être remplie plusieurs fois, avec plusieurs passages (3 passages paraissent l'idéal, mais un passage serait déjà bien: le fait d'aller plusieurs fois sur la parcelle permet d'être plus précis). Dans le tableau, on a mis trois lignes qui correspondent à trois passages. Mais un passage sera déjà très bien !

Si vous faites des photos, nous sommes preneurs !

!!! Un exemple est mis dans la première ligne du tableau à remplir !!!

---

## Fiche reconnaissance des maladies

Cette fiche montre les différentes maladies que l'on peut observer sur la blé lors de son cycle. Cette fiche n'est pas du tout détaillée, elle est simplifiée au maximum! Elle permet une identification rapide des maladies à partir de photos.

La notation des maladies peut se faire par un commentaire. On ne demande pas de passer tout les jours dans la parcelle pour faire des observations et voir comment évolue telle ou telle maladie. Cependant, il est intéressant de voir quelle est la pression de maladie pour une année donnée.

Exemple de commentaires que l'on peut faire sur la fiche : beaucoup de rouille jaune et de fusariose, sinon rien de spécial.

Les informations qui suivent sont tirées de «Maladies des blés et des orges» de l'Institut Technique des Céréales et des Fourrages (ITCF) et d'une présentation de Maxime Trottet «Les maladies du blé» Groupe technique céréales à paille Le Rheu - 7 juin 2007.

Il y a une fiche par type de maladie.

### **I. Les maladies du pied et des épis : les fusarioses**

### **II. Les maladies des feuilles et des épis**

#### **II. 1. La septoriose**

#### **II. 2. la rouille jaune, brune et noire**

#### **II. 3. l'oïdium des céréales**

## I. Les maladies du pied et des épis : les fusarioses



source : Maxime Trottet



source : Maxime Trottet

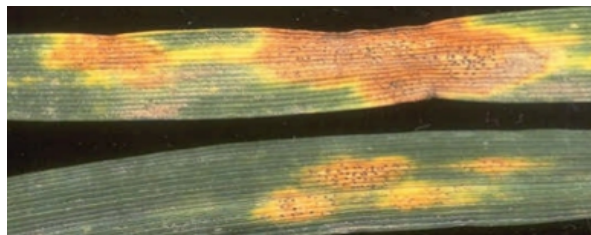
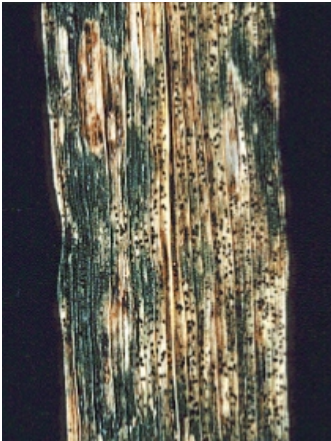
La fusariose est facilement reconnaissable sur les épis qui sont échaudé avant la maturité. Les épillets rougissent au début de la maladie.

## II. Les maladies des feuilles et des épis

### II. 1. La septoriose

les symptômes sont le plus visibles sur le feuillage au printemps.

Il y a deux types de septoriose: *tritici* qui est facilement reconnaissable grâce aux points noirs sur la feuille ; les feuilles sont jaunes et brunâtre.



Symptômes de la septoriose *tritici*. Source : Maxime Trottet

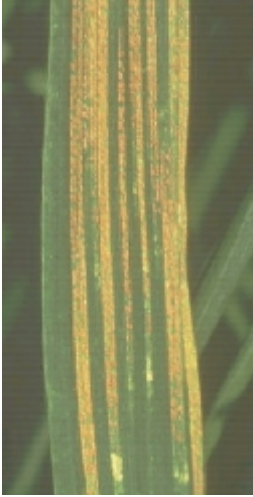
Et la septoriose *nodurum* qui montre sur les feuilles le même profil mais sans les points noirs.



Symptômes de la septoriose *nodurum*. Source : Maxime Trottet

## II. 2. la rouille jaune, brune et noire

**La rouille jaune** se reconnaît grâce aux taches jaunes en relief qui suivent les nervures de la feuille. De petites pustules sont présentes et suivent les nervures. Elles sont alignées en stries. Un foyer de tache jaune est repérable dans le champ.



Symptômes de la rouille jaune. Source : Maxime Trottet

**La rouille brune** est reconnaissable grâce aux pustules brunes qui sont sur les feuilles. Elles ne suivent pas les nervures.



Symptômes de la rouille brune. Source : Maxime Trottet

**La rouille noire** est rare (présente surtout dans le Sud). Les pustules sont de couleur assez foncée ou noire qui forment des stries essentiellement localisées sur les tiges.

### II. 3. l'oïdium des céréales

Les symptômes apparaissent sur les feuilles, les tiges et les épis.



Symptômes de l'oïdium des céréales. Source : Maxime Trottet





# Annexe H : Fiches d'été

Il n'y avait pas de notice pour les fiches pour les saisons 2008-2009 et 2009-2010. Le nom des populations a été ajouté (à la main) pour chaque paysan pour la saison 2009-2010.

Dès 2010, le nom des populations cultivées chez chaque paysan est automatisé.

En 2011, le piétin verse est enlevé, une hauteur moyenne et le poids de la récolte sont ajoutés.

En 2013, la notice est améliorée par Sophie Pin (équipe DEAP) et Florian Géry (stagiaire BTS en 2013 dans l'équipe DEAP).

**Croisements du Roc 2010**

Maturité / Échelle de notation biomasse: 1-5 (1=moins productif 5=plus productif)

Globale 1=pas beau 5= magnifique

Nom : \_\_\_\_\_

Autres précisions :

Roc = le témoin rouge de roc, avec deux réplifications

Date d'Observation	Numéro	note globale de biomasse	Note globale de la parcelle	Vos observations : notations spécifiques (aspects remarquables des épis, plantes, couleur, productivité, verse, enherbage etc)
	Roc Rep 1			
	Roc Rep 2			

Remarques éventuelles :

Croisement	Date	Globale	Biomasse	Crosses			Couleur		hétéro	Barbes		Verse	Commentaires
Exemple	25 05 11	5	5	pas croisé	50	blanc	50	5	barbus	50	50	blablaba	
				intermediaire	15	intermediaire	20		intermediaire	5	<input type="checkbox"/> piétin		
				croisé	35	rouge	30		nonbarbus	45	verse		
Rappel hiver : date : 10/04/11, globale= 2, reprise=3, densité= 300 à 400, Port au tallage : intermédiaire, Attitude des feuilles : horizontale													
Commentaires hiver : belle population mais levée hétérogène													
Rappel printemps : date : 10/04/11, globale= 2, reprise=3, densité= 300 à 400, Port au tallage : intermédiaire, Attitude des feuilles : horizontale													
Commentaires printemps : belle population mais levée hétérogène													

Croisement	Date	Globale	Biomasse	Crosses			Couleur		hétéro	Barbes		Verse	Commentaires
ALP RdR1				pas croisé		blanc			barbus				
				intermediaire		intermediaire			intermediaire		<input type="checkbox"/> piétin		
				croisé		rouge			nonbarbus		verse		
Rappel hiver : date : globale= , reprise= , densité= , Port au tallage : , Attitude des feuilles :													
Commentaires hiver :													
Rappel printemps : date : globale= , tallage= , vigueur= , Couleur : , Adventices :													
Commentaires printemps :													

Croisement	Date	Globale	Biomasse	Crosses			Couleur		hétéro	Barbes		Verse	Commentaires
6				pas croisé		blanc			barbus				
				intermediaire		intermediaire			intermediaire		<input type="checkbox"/> piétin		
				croisé		rouge			nonbarbus		verse		
Rappel hiver : date : globale= , reprise= , densité= , Port au tallage : , Attitude des feuilles :													
Commentaires hiver :													
Rappel printemps : date : globale= , tallage= , vigueur= , Couleur : , Adventices :													
Commentaires printemps :													

Nom :

Prénom :

## Fiche de renseignements été 2011 à retourner

### Notice de la fiche

**La notice permet de savoir quelle note attribuer à quel caractère.** Pour chaque croisement de votre essai, pouvez-vous indiquer les caractères suivant : **date de notation, globale, tallage, vigueur, couleur, présence d'adventices, épiaison, commentaires et maladies.**

**date de notation** : date à laquelle vous effectuez les mesures

**globale** : la note globale va de 1=moche à 5= magnifique ; cette note sera difficile à comparer entre les paysans mais donne une indication tout de même sur le comportement global de la variété.

**biomasse** : la note va de 1=moche à 5= magnifique ; cette note sera difficile à comparer entre les paysans mais donne une indication tout de même sur le comportement global de la variété.

**crosses** : indiquer le pourcentage d'épis dans la population qui sont crossés, intermédiaires ou pas crossés.

**couleur des épis**: indiquer le pourcentage d'épis dans la population qui sont blancs, intermédiaires ou rouges.

**hétérogénéité de la population (vue globale)** : indiquer si la populations est, dans son ensemble, très homogène (1) ou très hétérogène (5).

**barbes** : indiquer le pourcentage d'épis dans la population qui sont barbus, intermédiaires ou non barbus.

**verse** : indiquer le pourcentage des plantes dans la populations qui sont versées. Précisez si la verse est due au piétin verse en cochant la case. La verse peut être du au piétin verse : pour cela il suffit de voir si la base de la tige n'est pas noire.

Piétin verse : taches régulières.



piétin verse, tache régulière source : ITCF



forte attaque de piétin verse entraînant un échaudage complet (feuilles et épi) de la talle correspondant. Source : ITCF

**commentaires** :

Libre à vous ...

En bas du tableau sont reportées les notations qui ont été faites en hiver et printemps, pour mémoire.

Croisement	Date	Globale	Biomasse	Crosses		Couleur		hétéro	Barbes		Verse	Hauteur (m)
				pas croisé	intermédiaire	blanc	intermédiaire		barbus	intermédiaire		
Exemple	10 07 12	3	4	50	15	50	20	5	50	5	<input type="checkbox"/> à plat	1.1 m
				30	30	30	30		45	<input type="checkbox"/> intermédiaire		
										<input checked="" type="checkbox"/> presque droit		

**Date de la récolte :**

**Poids du grain récolté sur la micro-parcelle : 2,3 Kg**

Commentaires pour cette pop : Bla bla bla

Rappel hiver : date : 10/04/11 globale= 2 reprise=3 densité= 300 à 400 Port au tallage : intermédiaire Attitude des feuilles : horizontale

Commentaires hiver : belle population mais levée hétérogène

Rappel printemps : date :15/05/11 globale= 4 tallage= moyen vigueur= 5 Couleur : vert foncé Adventices : <10 pied de gaillet

Commentaires printemps : belle pop, bonne reprise mais mauvaise concurrence avec le gaillet

Croisement	Date	Globale	Biomasse	Crosses		Couleur		hétéro	Barbes		Verse	Hauteur (m)
				pas croisé	intermédiaire	blanc	intermédiaire		barbus	intermédiaire		
Rouge-du- Roc rep 1				pas croisé		blanc			barbus		<input type="checkbox"/> à plat	
				intermédiaire		intermédiaire			intermédiaire		<input type="checkbox"/> couché	
				croisé		rouge			nonbarbus		<input type="checkbox"/> intermédiaire	

**Date de la récolte :**

**Poids du grain récolté sur la micro-parcelle :**

Commentaires pour cette pop :

Rappel hiver : date : Attention, plan hypothétique car pas retour fiche automne globale= , reprise= , densité= , Port au tallage : , Attitude des feuilles :

Commentaires hiver :

Rappel printemps : date : NA globale= NA, tallage= NA, vigueur= NA, Couleur : NA, Adventices : NA

Commentaires printemps : NA

## Fiche de renseignements été 2012 à retourner

### Notice de la fiche

**La notice permet de savoir quelle note attribuer à quel caractère.** Pour chaque population de votre essai, pouvez-vous indiquer les caractères suivant : **date de notation, globale, biomasse, crosses, couleur des épis, hétérogénéité de la population, barbes, verse, hauteur, date de récolte, poids du grain récolté sur la micro-parcelle, commentaires.** Les notations faites en hiver et au printemps sont indiquées sur la fiche pour mémoire.

**Date de notation :** date à laquelle vous effectuez les mesures

**Globale :** la note globale va de 1=moche à 5= magnifique ; cette note sera difficile à comparer entre les paysans mais donne une indication tout de même sur le comportement global de la variété.

**Biomasse :** la note va de 1=moche à 5= magnifique ; cette note sera difficile à comparer entre les paysans mais donne une indication tout de même sur le comportement global de la variété.

**Crosses :** indiquer le pourcentage d'épis dans la population qui sont crossés, intermédiaires ou pas crossés. La crosse est la courbure de l'épi.

**Couleur des épis:** indiquer le pourcentage d'épis dans la population qui sont blancs, intermédiaires ou rouges.

**Hétérogénéité de la population (vue globale) :** indiquer si la populations est, dans son ensemble, très homogène (1) ou très hétérogène (5).

**Barbes :** indiquer le pourcentage d'épis épiés dans la population qui sont barbus, intermédiaires ou non barbus.

**Verse :** indiquer une catégorie de verse pour la populations :

à plat

couché

intermédiaire

presque droit

droit

**Hauteur :** indiquer la moyenne de la population en mètre.

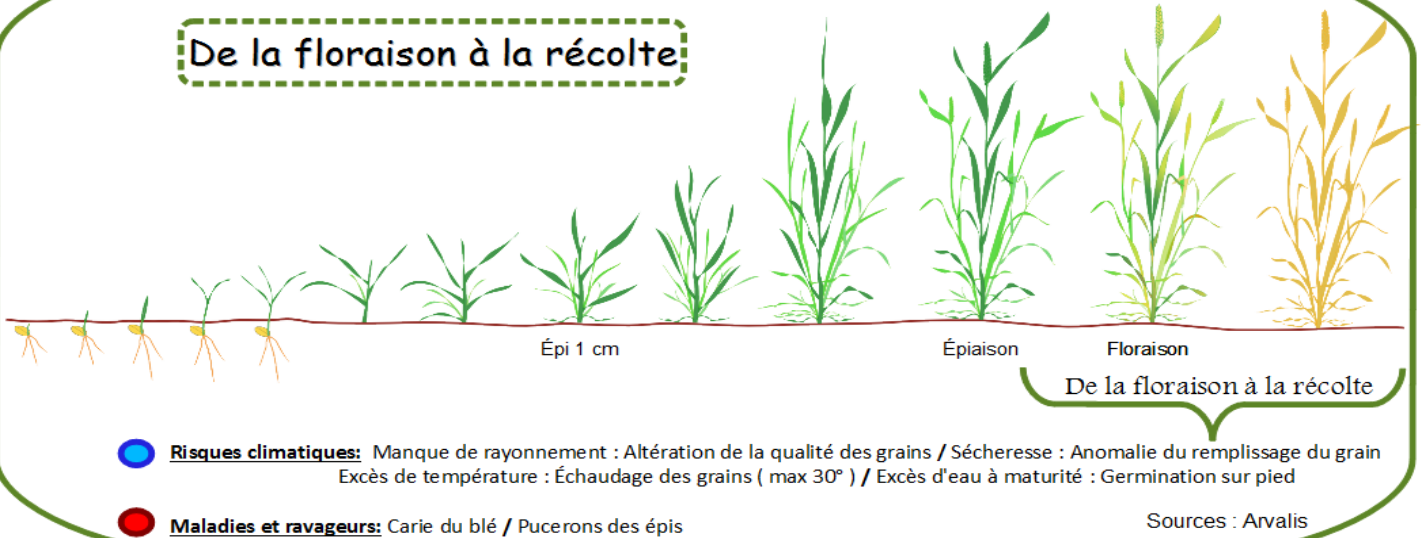
**Date de récolte :** indiquer la date de la récolte

**Poids du grain récolté sur la micro-parcelle :** indiquer le poids en Kg de grains récolté sur la micro-parcelle (cela nous permet de faire une estimation de rendement).

**Commentaires :** Libre à vous ...

!!! Un exemple est mis dans la première ligne du tableau à remplir !!!

## De la floraison à la récolte



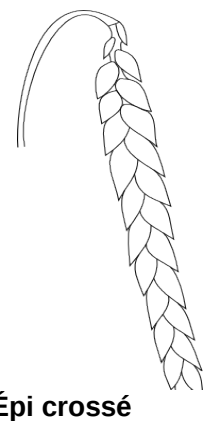
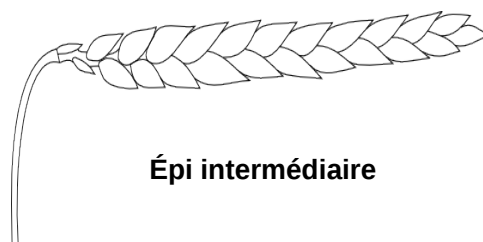
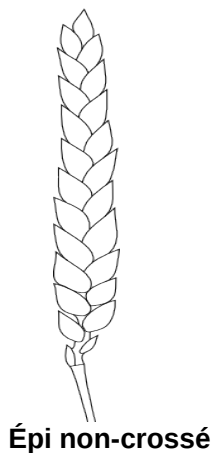
## Notice de la fiche

« été » 2013

### à retourner à votre animateur

**La notice permet de savoir quelle note attribuer à quel caractère.** Pour chaque population de votre essai. Les notations faites en hiver et au printemps sont indiquées sur la fiche pour mémoire. Pour chaque population de votre essai, pouvez-vous indiquer les caractères suivants :

- **Date de notation :** date à laquelle vous effectuez les mesures
- **Globale :** la note globale va de 1 = moche à 5 = magnifique ; cette note sera difficile à comparer entre les paysans mais donne une indication tout de même sur le comportement global de la variété.
- **Biomasse :** la note va de 1 = moche à 5 = magnifique ; cette note sera difficile à comparer entre les paysans mais donne une indication tout de même sur le comportement global de la variété.
- **Crosses :** indiquer le pourcentage d'épis dans la population qui sont crossés, intermédiaires ou non-crossés. La crosse est la courbure de l'épi :



70 • **Couleur des épis:** indiquer le pourcentage d'épis dans la population qui sont blancs, intermédiaires ou rouges. **Fiches d'été**

- **Barbes :** indiquer le pourcentage d'épis épiés dans la population qui sont barbus, intermédiaires ou non barbus.
- **Hétérogénéité de la population (vue globale) :** La note va de 1 = très homogène (les plantes se ressemblent entre elles) à 5 = très hétérogène (les plantes sont très différentes).
- **Verse :** indiquer une catégorie de verse pour la population :
  - à plat
  - couché
  - intermédiaire
  - presque droit
  - droit
- **Hauteur :** indiquer la moyenne de la population en mètre.
- **Date de récolte :** indiquer la date de la récolte
- **Poids du grain récolté sur la micro-parcelle :** indiquer le poids en Kg de grains récolté sur la micro-parcelle (cela nous permet de faire une estimation de rendement).
- **Commentaires :** Libre à vous ...

!!! Un exemple est mis dans la première ligne du tableau à remplir !!!



# Annexe I : Protocoles

## I.1 Notice pour les récoltes des épis des essais et renvoi des sacs des fermes au Moulon

**!!! Merci de nous renvoyer les sacs même si vous ne les avez pas utilisés !!!**

### **I. Récolte des échantillons : sac « 25 épis mesurés »**

Nous avons mesuré 25 plantes dans chaque parcelle, chaque plante est marquée avec un scotch vert et jaune et une laine rouge. Pour chaque micro-parcelle, **nous envoyer ces 25 épis marqués.**

**Le sac a utiliser est « sac 25 épis mesurés ». Attention à ce que les épis ne soient pas trop abimés.**

### **II. Récolte des échantillons : sac « 25 épis au hasard »**

Pour chaque parcelle, **nous envoyer 25 épis de 25 plantes différentes AU HASARD : petit, gros, moche, barbu, non barbu ...** ces épis au hasard doivent être représentatifs de la population. **Le sac a utiliser est « sac 25 épis au hasard ».** Cela nous permet d'avoir du stock au Moulon.

**Si vous voulez nous mettre du grain à la place, pas de souci (au contraire, c'est plus simple pour nous!).**

### **III. Récolte de vos bouquets de sélection si vous le souhaitez : sac « bouquet sélections (optionnel) » (Autant d'épis que vous le souhaitez!).**

Si vous le souhaitez, vous pouvez **sélectionner à l'intérieur** des variétés des épis qui vous plaisent. **Vous pouvez sélectionner autant d'épis que vous le souhaitez.** Vous pouvez nous envoyer ces épis sélectionnés pour qu'on les mesure. On vous les renvoie dès que possible pour le semis.

**On vous laisse mettre sur le sac dans quelle variété vous avez fait votre sélection.** Si vous nous donnez un mélange de bouquet de plusieurs variétés, merci de nous le noter.

**N'oubliez pas de noter dans votre petit carnet fétiche le pourquoi de vos sélections,** cela sera intéressant pour vous pour voir si votre sélection a répondu comme vous le souhaitiez.

Dans l'idée l'étude de vos sélections se décomposera comme suit :

**Cette année 2011-2012 :** On va comparer les épis que vous avez sélectionné avec les autres épis de la variété qui n'ont pas été sélectionné (c'est le différentiel de sélection).

**L'année prochaine 2012-2013 :** Pour chaque variété, semer les graines issues des épis sélectionnés et les épis non sélectionnés côte à côte. On mesurera des échantillons de chaque essai. On arrangera le plan en fonction.

N'hésitez pas à m'appeler si vous voulez plus de précision ! Ce sera avec plaisir ! **Merci de nous renvoyer les sacs même si vous ne les avez pas utilisés.**

### **IV. Envoi des sacs au Moulon**

**Avec vos sacs, mettez la facture pour être remboursé.** Devant la complexité administrative que sont les remboursements ... on a mis des sous de côté, on vous remboursera directement par chèque !

Merci d'envoyer le tout à :

Pierre Rivière  
UMR de Génétique Végétale  
Ferme du Moulon  
91190 Gif-sur-Yvette, France

## Récolte 2012 des essais : fermes satellites

**!!! Merci de nous renvoyer les sacs même si vous ne les avez pas utilisés !!!**

### I. Récolte des échantillons : sac « 50 épis au hasard »

Pour chaque parcelle, **nous envoyer 50 épis de 50 plantes différentes AU HASARD : petit, gros, moche, barbu, non barbu ...** ces épis au hasard doivent être représentatifs de la population. Cela nous permet d'avoir du stock au Moulon.

### II. Récolte de vos bouquets de sélection si vous le souhaitez : sac « bouquet sélections (optionnel) » (Autant d'épis que vous le souhaitez!).

Si vous le souhaitez, vous pouvez **sélectionner à l'intérieur** des variétés des épis qui vous plaisent. **Vous pouvez sélectionner autant d'épis que vous le souhaitez.** Vous pouvez nous envoyer ces épis sélectionnés pour qu'on les mesure. On vous les renvoie dès que possible pour le semis.

**On vous laisse mettre sur le sac dans quelle variété vous avez fait votre sélection.** Si vous nous donnez un mélange de bouquet de plusieurs variétés, merci de nous le noter.

**N'oubliez pas de noter dans votre petit carnet fétiche le pourquoi de vos sélections,** cela sera intéressant pour vous pour voir si votre sélection a répondu comme vous le souhaitiez.

Dans l'idée l'étude de vos sélections se décomposera comme suit :

**Cette année 2011-2012 :** On va comparer les épis que vous avez sélectionné avec les autres épis de la variété qui n'ont pas été sélectionné (c'est le différentiel de sélection).

**L'année prochaine 2012-2013 :** Pour chaque variété, semer les graines des épis sélectionnés et les épis non sélectionnés côte à côte. On mesurera des échantillons de chaque essai. On arrangera le plan en fonction.

N'hésitez pas à m'appeler si vous voulez plus de précision ! Ce sera avec plaisir ! **Merci de nous renvoyer les sacs même si vous ne les avez pas utilisés.**

### III. Envoi des sacs au Moulon

**Avec vos sacs, mettez la facture pour être remboursé.** Devant la complexité administrative que sont les remboursements ... on a mis des sous de côté, on vous remboursera directement par chèque !

Merci d'envoyer le tout à :

Pierre Rivière  
UMR de Génétique Végétale  
Ferme du Moulon  
91190 Gif-sur-Yvette, France

## **I.2 Visite des fermes Régionales et mesures phénotypiques au champ**

Le protocole n'a pas changé d'une saison à l'autre.

## SOLIBAM : Visite des fermes régionales cet été

### *Matériels:*

#### *mesures*

- mètres x 2 (demander la perche aux paysans)
- pads + recharges x 2
- appareil photo
- pc portable pour mettre les photos et les données au fur et à mesure
- Grosses enveloppes récolte pour chaque ferme de la région (avec sacs pour les mesures, le vrac et les éventuelles sélection + notices + fiches été)
- paire de ciseaux x 2
- feutre indélébile x 3
- pelote de laine rouge pour marquer les épis mesurés x 15
- rouleau scotch d'électricien pour numéroter les plantes x 15
- cahier + crayon à papier x 3

#### *vêtements*

- chapeau, crème solaire, bouteille
- parapluie, k-way, bottes
- tente + ...

#### *animation*

- livret de présentation du projet x 100
- poster du projet
- fiche avec les premiers résultats (1 personnalisée par FR visitée)
- pelote de laine rouge et verte x 3
- bambou pour délimiter la zone où on fait le jeu et à mettre en face des micro parcelles au cas où il n'y ait pas de piquet-étiquette x 40
- drapeau pour mettre au sommet des bambous x 5
- carte de visite x 100

Les objectifs de ces visites sont:

1. mesurer les populations qui sont cultivées
2. communiquer sur le projet, rencontrer les paysans et les citoyens du coin
3. faire discuter les personnes autour des populations pour qu'ils échangent leurs point de vue (jeu de la sélection)
4. distribuer les sacs de récolte et fiches été aux FS si paysans présents.

### **1. mesurer les populations qui sont cultivées**

**1.a** Les caractères mesurés aux champs sur 25 individus sont:

- la distance sol-sommet de l'épi
- la distance sol-base de l'épi
- la distance sol-base de la feuille drapeau
  
- les barbes (Absence - 0, 0.5, 1, 1.5, 2 - Forte présence)
- la couleur (Clair (blanc) - 0, 0.5, 1, 1.5, 2 - Foncé (rouge ou noir))
- la courbe (Droit - 0, 0.5, 1, 1.5, 2 - Très courbé)

**1.b** Pour chaque population, on prend 2 photos : 1 de l'étiquette, 1 de l'ensemble de la population

**1.c** si à maturité, les 25 épis mesurés (+ scotch numéroté) sont récoltés et mis dans le sac adéquat sinon, on marque les épis avec la laine rouge, le paysan, les récoltera et les enverra au Moulon.

**1.d** si à maturité, récolter 25 épis en vrac que l'on met dans le sac adéquat sinon, le paysan les récoltera et les enverra au Moulon.

## **2. animation des visites**

On présente ce que l'on fait sur SOLIBAM dans une présentation de 15 minutes (+ poster + fiche avec premiers résultats). On distribue des livrets. On donne des cartes avec nos coordonnées.

## **3. faire discuter les personnes autour des populations pour qu'ils échangent leurs point de vue.**

L'idée est de faire prendre conscience à tout le monde que l'on peut sélectionner, voir des différences dans les blés, en apprécier certains plus que d'autre.

Nous mettons ça sous forme d'un jeu.

Règle:

### temps 1 : sélection inter-population

- chaque personne dispose les rubans selon ses préférences: vert: j'aime bien , rouge: je n'aime pas. Il y a moins de rubans qu'il n'y a de parcelles; Par exemple, pour 30 parcelles, les personnes auront 5 verts et 5 rouges: ça oblige à faire des choix!!!

### temps 2 : sélection intra-population

- on se focalise sur les 3 parcelles ayant eu le plus de rubans, chaque personne marque 1 épis qu'il aime par parcelle (marquage avec des laines de couleur vive enroulées autour de l'épi). BIEN PRENDRE LE TEMPS DE DISCUTER DES CHOIX DE CHACUN.
- NB : On peut prendre une couleur de laine pour les paysans et une pour le reste du public et confronter leur regards

### *Notations pour nous*

- prendre des photos de tout ça!

- noter combien il y a de rubans par parcelles
- Sur les populations qui font débat, ou pas, noter les remarques de chacun dans le cahier!

## **4. distribuer les sacs de récolte et fiches été au FS .**

Si les paysans des FS sont présents, on leur distribue le pack avec les sacs de récolte + la fiche été + la notice.

Si ils ne sont pas là on les envoie par la poste.

### **I.3 Mesures dans la salle technique**

Les protocoles ont varié d'une année à l'autre selon les dispositifs mis en place. La différence réside dans le nombre d'épis mesurés et les variables mesurées.

## Fermes Satellites

### « Sac Sélections (optionnel) »

- ⇒ 25 épis au hasard dans le bouquet
  - couleur
  - barbes
  - courbe
  - poids de l'épi
  - nombre d'épis (qui peut être différents de 25)
  - tout battre ensemble
  - nombre de grains
  - poids du grain
- ⇒ le reste, si il y a
  - tout battre ensemble



On met tous les grains dans le même sac pour PAYSAN

### « sac 50 épis au hasard »

- ⇒ S'il y a « Sac Sélections (optionnel) », 25 épis au hasard dans le bouquet
  - couleur
  - barbes
  - courbe
  - poids de l'épi
  - nombre de plantes (qui peut être différents de 25)
  - tout battre ensemble
  - mesurer nombre de grains et poids du grain
- ⇒ Si il N'y a PAS « Sac Sélections (optionnel) », sur 25 épis
  - poids de l'épi
  - nombre d'épis (qui peut être différents de 25)
  - tout battre ensemble
  - nombre de grains total
  - poids du grain total
- ⇒ S'il y a plus de 25 épis dans le sac
  - tout battre, pour le stock, pas de mesure



On met tous les grains dans le même sac pour stocker



## Fermes Régionales

### « Sac Sélections (optionnel) »

⇒ 25 épis au hasard dans le bouquet

- couleur
- barbes
- courbe
- longueur de l'épi
- nombre d'épillets
- nombre d'épillets stériles
- nombre d'épillets manquants
- poids de l'épi
- nombre d'épis (qui peut être différents de 25)
- tout battre ensemble
- nombre de grains
- poids du grain

⇒ le reste, si il y a

- tout battre ensemble

**On met tous les grains dans le même sac pour PAYSAN**

### « sac 25 épis mesurés »

- longueur de l'épi
- nombre d'épillets
- nombre d'épillets stériles
- nombre d'épillets manquants
- poids de l'épi
- nombre d'épis (qui peut être différents de 25)
- tout battre ensemble
- nombre de grains
- poids du grain



NE PAS oublier les « ● »

### « sac 25 épis au hasard »

- tout battre ensemble, pour le stock, pas de mesure



# **Annexe J : Réponse à la sélection massive paysanne en génération précoce dans des populations de blé tendre issues de croisements**

Cet article a été publié dans *Renewable Agriculture and Food Systems* en 2013.

1 **Response to farmer mass selection in early generation progeny of**  
2 **bread wheat landrace crosses.**

3 Pierre RIVIERE<sup>1\*</sup> - Isabelle GOLDRINGER<sup>1</sup> - Jean-François BERTHELLOT<sup>2</sup> - Nathalie GALIC<sup>1</sup> - Sophie PIN<sup>1</sup>

4 - Patrick De KOCHKO<sup>2</sup> - Julie C. DAWSON<sup>1,3</sup>

5 <sup>1</sup>UMR Génétique Végétale, INRA – Univ Paris-Sud - CNRS, Ferme du Moulon, F-91190 Gif-sur-Yvette,

6 France

7 <sup>2</sup>Réseau Semences Paysannes, 3, avenue de la gare, F-47190 Aiguillon

8 <sup>3</sup>Current address: Cornell University, Department of Plant Breeding and Genetics, 417 Bradfield Hall, Ithaca,

9 NY 14853, USA

10 \*corresponding author: [pierre.riviere@moulon.inra.fr](mailto:pierre.riviere@moulon.inra.fr)

11

12 **Abstract**

13 A Participatory plant breeding (PPB) program involving the French farmers' association  
14 "Réseau Semences Paysannes" and the French National Agricultural Research Institute  
15 (INRA) at Le Moulon was initiated in 2005. In the process of designing the breeding  
16 scheme, we evaluated the impact of farmer selection at an early stage (F<sub>2</sub>) on bread wheat  
17 cross progeny populations. The objectives were to characterize the effect of farmer  
18 selection, to evaluate the impact of farmer selection on intra-varietal diversity, to provide  
19 farmers with relevant information that they can use to improve their selection practices.  
20 Early selection was found efficient for some traits and for some of the 35 F<sub>2</sub>-derived F<sub>3</sub>  
21 families. For traits of interest such as thousand kernel weight or grain weight per spike,  
22 when the response was significant, it was always positive. For most of the traits studied,  
23 the among-family genetic variance increased after selection while the average within-  
24 family genetic variance decreased. This study provides the first quantitative results for this  
25 PPB program and information that will help optimize it in the future.

26

27 **Key words**

---

1

1 bread wheat, participatory plant breeding, population varieties, biodiversity management,

2 organic breeding

## 1 Introduction

2 Organic farming often leads to specific environmental conditions, which are more stressful  
3 for the plants since the use of chemical inputs is not permitted (1). While modern varieties  
4 fit the needs of conventional farming in industrialized countries (homogeneity, adaptation  
5 to mechanization and high input management), they often are not adapted to agronomic  
6 practices that decrease the use of inputs and fossil energy. In addition, dimensions of  
7 integrated systems such as the need for biomass for animal feed, competitive ability with  
8 weeds, efficient uptake and utilization of nitrogen and end-use quality for traditional foods  
9 are often not taken into account in conventional breeding programs (2-5). This often leads  
10 farmers in North Africa, Latin America or Asia to cultivate landraces or historic varieties  
11 instead of modern cultivars (6-7), since landraces may be adapted to heterogeneous  
12 environments and specific objectives (8).

13  
14 To succeed in developing new varieties adapted to these kinds of heterogeneous  
15 environments, participatory plant breeding (PPB) has been implemented in several cases  
16 (9,13). PPB aims at developing varieties adapted to farmers' needs in contrasted low input  
17 environments (such as organic management in Europe), while maintaining genetic  
18 diversity. It is based on (i) accounting for genotype x environment x management  
19 interactions through decentralized selection, (ii) collaboration between researchers,  
20 breeders, farmers and other stakeholders; and (iii) the development and use of appropriate  
21 genetic diversity for breeding.

22  
23 Efficient breeding in stressful environmental conditions will require that environmental  
24 complexity is taken into account (14). Decentralized selection in many sites is needed in  
25 order to conduct direct selection in the target environment (9), which has proved more  
26 efficient than indirect selection from favorable to stressful environments (2). In addition,

1

1 participation of farmers is required to benefit from their experience and expertise in varietal  
2 evaluation in their particular environment (15), and to implement selection in a way that will  
3 respond to their specific needs. This participation also empowers farmers and leads to  
4 more autonomy with respect to varietal choices and the promotion of farmers' rights  
5 (3,16). This approach has already proved to be efficient in developing countries  
6 (7,12,15,17–20) but has started only recently in Europe, with new programs showing  
7 promising results (13,21–25).

8

9 In France, the demand of organic farmers for adapted varieties first resulted in the  
10 cultivation of landraces and historic varieties (26). In 2003, a group of organic farmers that  
11 want to conserve agricultural biodiversity and enhance their seed autonomy founded the  
12 association Réseau Semences Paysannes (RSP, the Farmers' seed network). The RSP is  
13 a network of farmers' associations that conserve, multiply and exchange landraces, old  
14 varieties and other farmers' varieties (27). On-farm management has been shown to be an  
15 effective method to conserve agricultural biodiversity, complementary to *in situ*  
16 conservation on research stations and *ex situ* conservation in gene banks (28,29). On-  
17 farm management is a key activity for the conservation of genetic resources, as underlined  
18 by the 2004 International Treaty on Plant Genetic Resources for Food and Agriculture  
19 under the control of the FAO (ITPGRFA). Nevertheless, an *in situ* conservation system  
20 solely based on landraces and old varieties may not fit all farmers' needs. Farmers from  
21 the RSP have become interested in the development of new varieties that both conserve  
22 crop biodiversity and are adapted to the current organic farming practices (27). Because in  
23 the context of industrialized agriculture and an institutionalized seed supply system, the  
24 classical inter-generational transmission of knowledge has disappeared, PPB development  
25 presents unique challenges both at the technical (genetic, agronomic, analytical) and at  
26 the organizational level (26,30).

2

1

2 Several strategies can be applied to start a new breeding scheme, such as crossing,  
3 mixing or selecting in the available landraces or locally adapted populations (31). Here, we  
4 chose to create new populations through manual crosses, thus generating a broad range  
5 of new allelic combinations. The farmer-baker who initiated the project, Jean-François  
6 Berthelot (JFB), chose the parents according to their baking and milling quality, their  
7 history and geographical area of cultivation and their agronomic behavior based on the  
8 knowledge he has acquired by growing them on his farm for the past five years, and  
9 through collaborations with other farmers from the RSP. The goal of making crosses was  
10 to combine the bread-making quality and agronomic resilience of landraces with historic  
11 varieties (1<sup>st</sup> half of the 20<sup>th</sup> century) and a few more recent varieties, which are more  
12 resistant to lodging. Before the populations derived from these crosses were distributed to  
13 a large network of farms mainly in France, an experiment was performed to assess the  
14 impact of farmer's mass selection in an early generation.

15

16 From conversations with farmers, it is clear that while they are looking for certain  
17 characteristics in their varieties, they are also looking to maintain more phenotypic  
18 diversity within varieties than is normally present in modern varieties. They often mention  
19 that one of the benefits of more heterogeneous population-varieties is their increased  
20 stability over years, due to the within-variety heterogeneity that buffers environmental  
21 fluctuations. Within-field varietal diversity has been shown to increase the functionality,  
22 resilience and stability of agricultural ecosystems (32,33). Genetically diverse varieties  
23 may combine quantitative and qualitative resistance, thus providing more durable disease  
24 resistance (34-36). Phenotypic variability within varieties has also been found to be  
25 associated with an increase in associated biodiversity (37,38). Within-field varietal diversity  
26 may not only contribute to yield stability but also to the stability of quality as shown in



1

1 wheat that varietal mixtures increase uptake efficiency of nitrogen (39). Finally, the on-farm  
2 management of such diverse varieties contributes to the *in situ* conservation of genetic  
3 resources for plant breeding (8,13,28,40-42).

4 While for an autogamous plant like wheat, most of the selection is often done among  
5 varieties that are nearly pure lines, at a more advanced generation, and not within  
6 segregating populations (6,9,10,43), this study assessed the diversity created in the  
7 program and the response to farmers' mass selection in early generations after crossing  
8 ( $F_2$ ) in terms of trait means and genetic variance within and among families. Collaboration  
9 with farmers (participation) at all stages of the research study was critical to reach these  
10 objectives.

11

## 12 **Materials and Methods**

### 13 *Context and experimental design*

14 On the initiative of JFB, a farmer active in the RSP, a participatory plant-breeding project  
15 was started in 2005 with researchers from INRA Le Moulon. The project was extensively  
16 discussed with all farmers, RSP coordinators and researchers. Ninety crosses were made  
17 on JFB's farm between different historic wheat varieties, landraces and modern varieties  
18 created for organic agriculture. These landraces, historic varieties and modern varieties  
19 had been cultivated on his farm for at least five years. Most of these populations came  
20 from the national seed banks at INRA Clermont Ferrand, France; and from Switzerland  
21 (varieties created by Peter Kuntz) and Germany (varieties created by Bertold Heyden).

22

23 The first ( $F_1$ ) and second ( $F_2$ ) generations of progeny of the 90 different families (one  
24 family is derived from each cross, numbered 1-90) have been grown on his farm in 2006-  
25 2007 and 2007-2008. Selections were made of individual spikes in a sub-set of the  $F_2$   
26 families. Seed of these spikes was bulked for each family and 35 of these families were

2

1 evaluated in the F3 generation with their corresponding unselected bulk at INRA Le  
2 Moulon (Gif sur Yvette, France) in 2008-2009. Three families had two different selected  
3 versions, so the total number of selected populations was 38. When two selections were  
4 made within a family, this was indicated by a letter following the number (14a and 14b, 34a  
5 and 34b, 42a and 42b). There were three complete block replicates of the 35 families, with  
6 paired rows of each version (selected or bulk) for each family. Rows were 1.20m long with  
7 20 seeds sown per plot. Paired versions within families were randomized in each block but  
8 the selected and the bulk families of each pair were always grown side by side. This  
9 maximized the power to detect differences between the two versions. In addition, Renan, a  
10 pure line cultivar frequently used in organic agriculture in France was used as a check  
11 variety and as a point of comparison for the farmers. Renan was included twice in each  
12 replicate of the experiment.

13  
14 *Measurements*  
15 Qualitative observations and quantitative measures were taken on the main tiller of each of  
16 five plants for each version of each family in each replicate (i.e. a total of 15 plants for  
17 each version of each family). The traits measured and their abbreviations are listed in  
18 Table 1. These traits were chosen by farmers and also based on phenotypic descriptors  
19 used for variety registration. After field measurements were made on each plant, the spike  
20 was cut and individually bagged. At the technical facilities at INRA Le Moulon,  
21 measurements were taken on all spikes collected. Grain from each replication of each  
22 version within families was analyzed for technology traits at INRA Clermont Ferrand using  
23 near infrared spectroscopy (Foss NIRSystem 6500), using whole grain. Although this  
24 prediction is less precise than on whole grains flour, it was chosen because it is not  
25 destructive and this allowed us to replant the seeds. The correlation between estimated  
26 and true value using this method in the calibration sample was (Table 1) : protein (0.86),

1

1 hardness (0.77), TW (0.80), MTI (0.7) and W (0.75). This was assessed on modern  
2 cultivars (G. Branlard, pers. comm.).

3

#### 4 *Statistical analysis*

5 First, we tested for an overall version effect using an ANOVA model with all effects fixed:

$$6 Y_{ijkl} = \mu + family_i + rep_j + version_k + \varepsilon_{ijkl} \quad (1)$$

7 Where  $Y_{ijkl}$ , is the phenotypic value measured for plant  $l$  of version  $k$  of family  $i$  in

8 replication  $j$ ,  $\mu$  is the general mean,  $family_i$  is the effect of family  $i$ ,  $rep_j$  is the effect of

9 repetition  $j$ ,  $version_k$  is the effect of the version ( $k$ =selected or non-selected) and  $\varepsilon_{ijkl}$  is the

10 random error term.

11 Model (2) was used to test for specific selection effect dependent on the family. This was

12 similar to Model 1 but the version effect (selected or non-selected) was nested within

13 family ( $version(family)_{ik}$ ).

14

15 The statistical analysis was implemented in SAS v 9.2 proc GLM (44). Two sided tests

16 between the least-square (LS) means for selected and non-selected versions within each

17 family were made with Tukey's multiple comparison procedure in SAS. This was done

18 using the SLICE function on the LS Means for version(family), which tests for an effect of

19 version within each family. To visualize the response to selection for multiple traits and

20 families, the log(pvalue) of the two sided test described above were recorded in a matrix of

21 traits x families (18x38). We used a log transformation in order to weight the significant

22 changes. When the mean of the selected version were higher than the bulk, then

23 log(pvalue) were multiply by -1 (so that the values were positive), else by 1 (so that the

24 values remained negative). Ward's clustering procedure was used on the data in this

25 matrix to group each version of the different families by the similarities in their responses

26 to selection. This clustering and the resulting dendrogram and heatmap visualization was

2

1 done with the heatmap function in R (45).

2

3 Repeatability is a measure of the proportion of phenotypic variation due to genetic causes.

4 It was estimated using the genetic and residual variance estimated for each version

5 (selected or bulk) from the ANOVA mixed model:

$$6 Y_{ijk} = \mu + family_i + rep_j + R_{ijk} \quad (3)$$

7 Where  $Y_{ijk}$  is the phenotypic value measured for plant  $k$  of family  $i$  in repetition  $j$ ,  $\mu$  is the

8 general mean,  $family_i$  is the family  $i$  (random effect),  $rep_j$  is the repetition  $j$  (fixed effect) and

9  $R_{ijk}$  is the error. Mixed models were used to estimate variance components using the

10 function VarCorr in the R package lme4 (45). Thus the repeatability was estimated as:

$$11 r = \frac{\text{var}(G)}{\text{var}(G) + \frac{\text{var}(R)}{3}}$$

12 Where  $\text{var}(G)$  was the estimated among-family genetic variance, and  $\text{var}(R)$  the estimated

13 residual variance. As the families are progenies of F2 spikes, the error  $R_{ijk}$  of the model

14 includes both the within-family genetic variance plus the environmental variance.

15 We estimated the average within-family genetic variance as the difference between the

16 error variance ( $\text{var}(R)$  from model (3)) and the residual variance of the check variety

17 Renan. As the check is a pure line, i.e. genetically homogeneous, the variation observed

18 was only due to the environment. More information on the calculation of within-family

19 variance is provided in supplemental material. For the data where there was no individual

20 data but only plot mean values, only the replication effect was used in the model.

21 Correlations among traits before and after selection were also calculated, methods and

22 results are given in the supplementary information.

23

## 24 **Results**

### 25 *Trait variation*

1

1 PH was rather high in both the selected and bulk versions with a mean of 134.5 and 132.7  
2 cm respectively. This was also the case for LLSD with means of 28.6 and 27.7 cm,  
3 respectively, and for SL with means of 12.8 and 12.7 cm, respectively (Table 2). Such high  
4 values are typical of landraces and historic varieties. In general, the range of variation of  
5 the trait values was quite large indicating the high level of diversity generated in these  
6 populations.

7 In the model (1) and (2), the family effect was significant for all traits (Table 3.a. and 3.b.) ,  
8 while the effect of replication was significant only for protein, W and MTI which might be  
9 due to the fact that the grain composition traits are highly dependent on the environment  
10 (46,47).

11

### 12 *Response to selection for trait means*

13 Seven traits changed significantly after selection with a common trend over all families  
14 among which TKW and protein showed highly significant changes (respectively positive  
15 and negative). Six other traits changed significantly but the direction of the response  
16 depended on the family (Figure 1). For 24 families out of 38, significant differences  
17 between bulk and selected versions were found for at least one trait.

18

19 The overall version effect was significant at  $p < 0.01$  for TKW and protein (Table 3.a.), and  
20 significant at  $p < 0.05$  for PH, earliness, GW\_Spike, TW, W and MTI. This indicated a  
21 unidirectional response to selection for these traits: although the response was not always  
22 significant for the different families, we found a trend towards an increase for TKW,  
23 GW\_Spike and PH and towards a decrease for protein (Figure 1). The effect of the  
24 version within family was significant for seven traits out of 16: PH, LLSD, SL, SpTot, TKW,  
25 sterility and earliness (Table 3.b.) indicating that for LLSD, SL, SpTot and sterility the  
26 response was dependent on the family.

2

1

2 While there was a significant global version effect for PH (Table 3.a), the magnitude of the  
3 selection response was specific to each family (Table 3.b). PH significantly increased from  
4 120 cm to 150 cm for family 64; from 118cm to 140cm for family 80; from 70 cm to 110 cm  
5 for family 34a, and from 100 cm to 115 cm for family 34b, whereas PH decreased  
6 significantly from 130 cm to 125 cm for family 24 (Figure 1.b.).

7

8 As can be seen in Figures 1.a., 1.c. & 2, selection for GW\_Spike, KN\_Spikelet, TKW,  
9 earliness and TW always increased trait values when response was significant. Moreover,  
10 TKW increased for 26 families over the 38, of which 4 cases were significant (Figure 1.c.).  
11 In contrast, Hardness, W and MTI selection always decreased trait values when the  
12 response was significant. Finally, PH, LLSD, sterility, awns, SpTot, color, density, protein,  
13 and SL changed in both directions. It is interesting to point out that protein tended to  
14 decrease after selection for 30 families over the 38 selected, but only two families  
15 decreased significantly. For one family (family 60), protein increased significantly after  
16 selection. There were 14 families that did not respond to selection for any of the traits  
17 measured, 11 families that responded to selection for only one trait, six families that  
18 responded to selection for two traits, six families for three to eight traits and one family for  
19 12 traits.

20

21 For family 34, the two selections ("a" and "b"), led to different responses: eight significant  
22 changes for "b" and two significant changes for "a" that were in the same direction as for  
23 "b" (for PH and LLSD). For family 42, the two selections ("a" and "b"), led to similar  
24 response patterns. For family 14, neither of the selections led to many changes from the  
25 bulk version. The dendrogram shows that variables linked to technological properties  
26 (Protein, MTI, W) had similar response patterns (Figure 2). LLSD and PH also responded

1

1 similarly to each other as did TKW and TW.

2

3 *Response to selection at the variance level*

4 Repeatability ranged from 0.13 for GW\_Spike within the bulk versions to 0.81 for PH within  
5 the bulk versions, with higher values for morphological traits such as PH, LLSD, SL,  
6 earliness and SpTot and for grain composition traits (Table 4). This indicated good control  
7 of environmental variation and the ability of the experimental design to discriminate among  
8 families both before and after selection. Note that the lowest value was for the trait that is  
9 the closest to grain yield (GW\_Spike).

10 GW\_spike, TKW, protein and MTI had greater repeatability in the selected version  
11 compared to the bulk (increases from +32.2% to +139.4%). Repeatability for SL, TW, W,  
12 PH, KN\_Spike, SpTot, earliness and density did not change much in the selected version  
13 (changes from -4.8% to +10.5%). Repeatability for hardness, sterility, KN\_Spikelet and  
14 LLSD was lower in the selected version (decreases from -9.2% to -28.3%) (Table 4).  
15 Selection increased the among-family genetic variance for GW\_Spike, TKW and MTI  
16 (increases from +25.4% to +66.0%) while it was reduced in a more limited proportion for  
17 LLSD, KN\_Spikelet, earliness and hardness (decreases from -11.4% to -22.7%). Little  
18 change was observed for the other traits (Table 4). Selection increased the average within-  
19 family genetic variance for sterility (+32.7%) while a marked decrease was observed for  
20 GW\_Spike, TKW, protein, TW and MTI (between -44.7% and -13.3%) (Table 4).

21

## 22 **Discussion**

23 In this study, we analyzed F2 derived F3 families from 35 crosses among a wide range of  
24 landraces, historic and more recent varieties and selected by an organic farmer. First we  
25 comment on the diversity created by the crosses. Then, we discuss the response to

2

1 farmer's selection within early generation families (mean and variance). We provide  
2 specific examples of directional selection, correlations among selected traits, and the  
3 influence of parental varieties on the response to selection in the supplementary  
4 information. Finally, we discuss how these results can be integrated in the ongoing PPB  
5 program.

6

7 *1. Creation of diversity for selection*

8 In conventional selection programs, breeders usually seek to decrease the within-family  
9 genetic variance in order to obtain uniform lines. However, in this project, one of the goals  
10 was to maintain within-family variance and thus maintain the genetic potential for  
11 continuing on-farm selection. Overall, after selection, this variation remained high. This is  
12 positive as it will allow the farmers to continue selecting within populations.

13

14 Both before and after selection, differences among families were highly significant for all  
15 traits (Table 3a & 3b). This was consistent with the objective of the crosses to generate a  
16 large range of diversity in order to increase the chance of developing populations that  
17 might adapt to contrasting environmental conditions. Traits where uniformity is of  
18 importance for standardized production may also not be as critical in situations where  
19 production and value-added processing occurs on-farm or for a small artisanal market.  
20 For the specific example of baking force, W, see supplementary information.

21

22 Because this experiment was done in a common garden with all populations grown in a  
23 single environment, the observed phenotypic diversity is likely due to the genetic diversity  
24 found in our panel of populations. The populations are in the F3 generation derived from  
25 crosses among very diverse parents. Assessing the relative contribution of among and  
26 within-family variability is of importance to develop appropriate selection procedure in PPB.



1

1

2 In general, genetic among-family variation was high for most traits (Table 4), leading to  
3 high repeatability, except for some characteristics of the spike such as KN\_spike and  
4 GW\_spike. This is consistent with classical findings in quantitative genetics for such  
5 complex traits. When selecting, selection among families can be the first step  
6 (9,19,23,43). With diverse and distinguishable families, farmers can find populations that  
7 better suit their specific environments and practices.

8 The within-family genetic variance also was quite high for many traits (Table 4). This is  
9 expected as we are dealing with segregating populations. After selection among families,  
10 selection within families can be the second step in PPB. Within-family diversity can be  
11 used in an evolutionary plant breeding approach (48) where populations are mainly  
12 submitted to natural selection. To increase selection efficiency, farmers may apply mass  
13 selection to guide the evolution of the population towards phenotypes of interest for them  
14 (7,10,25,49,50). Researchers and farmers conducting selection have to be careful to  
15 monitor competition among individuals within the population (for light and uptake of  
16 nutrients in the soil) when using either evolutionary breeding or mass selection, as  
17 competition ability might be negatively correlated with some traits desired by farmers  
18 (quality and yield) (2,31).

19

## 20 *II. Assessing the impact of farmers' mass selection*

21 A significant positive response was found for morphological and phenological traits such  
22 as PH, LLSL, earliness or SL. Within the RSP, farmers are often looking for plants that are  
23 tall but resistant to lodging because they have noticed a positive relationship between PH  
24 and the maintenance of grain filling under stress which impacts yield and grain quality.  
25 From farmers' observations, taller plants also are more competitive with weeds, and  
26 farmers can use the extra straw for livestock or for soil fertility management. The measure

2

1 of LLSD was proposed by the farmers, because they found that a greater distance  
2 between the spikes and the foliage may prevent leaf diseases from jumping to spikes.  
3 They also observed an improvement in grain filling and grain maturation with longer LLSD  
4 if leaves die because of disease or abiotic stress, as the stem resources can be used to  
5 continue grain filling. This is corroborated in the scientific literature (51).

6  
7 Traits such as TKW or GW\_Spike, which are more difficult to assess visually in the field,  
8 also increased after selection in families where there were significant differences. Family  
9 80 was very heterogeneous and responded to 12 traits out of 18 measured. This can be  
10 related to the parents used (see supplementary information). Ceccarelli et al. (43) also  
11 showed the ability of farmers to select superior populations when working with early  
12 generations.

13 In discussions among farmers, they said that they may have particular selection criteria but  
14 always adapt to the specific populations they observe and take a more holistic approach to  
15 selection. As already noticed by Ghaouti et al. (23), farmer selection is integrative, it does  
16 not favor individual traits but instead overall plant vigor, field productivity and quality. In our  
17 study, traits such PH, SL, SpTot, density, and color are indicators of such vigor for the  
18 farmer, leading to homogenization and uniform selection. However, if there are plants that  
19 look interesting but do not fit the “type”, the farmers will select them anyway, leading to  
20 more heterogeneous samples.

21 Selection in the F<sub>2</sub> may be early for an efficient response to selection as segregation is not  
22 complete (i.e. (1/2)<sup>2</sup> heterozygotes expected), and the covariance between F<sub>2</sub> plants and  
23 F<sub>3</sub> plants from the same family is expected to be moderate. Moreover, as there is a high  
24 level of heterozygosity, differences are difficult to assess. Most of the selection in PPB  
25 program described in the literature are carried out within more advanced generations, for  
26 example F<sub>3</sub>, F<sub>4</sub>, F<sub>5</sub> in barley and bean (6,17,43); F<sub>5</sub> for rice (10); and F<sub>3</sub> and F<sub>4</sub> in sorghum

1

1 (12). In this study, the farmer's objective was to save time and subject the populations to  
2 the conditions of the target environment as soon as possible. An original aspect of this  
3 experiment is that all crosses were performed on-farm and the early generations were also  
4 cultivated on the same farm by the farmer who initiated the project. In most reported cases  
5 of participatory plant breeding, crosses were made on the research station and the early  
6 generations were also grown at the research station (9,12,17,43). Here, the farmer  
7 applied mass selection without being influenced by the researchers. This approach differs  
8 from other programs where researchers train farmers (49) or apply selection before the  
9 farmers (12,50). The objective here was to assess specifically the effect of the farmer's  
10 selection, which is based on his unique knowledge of his farming system. The efficiency of  
11 farmers' selection has been demonstrated in other cases on rice, barley or quinoa but it  
12 was most often a screening among families at a later generation (6,10,19).

13

14 The farmer whose selections were studied is not representative of all farmers in the RSP.  
15 He has extensive experience growing and observing population-varieties and may be  
16 described as an "expert farmer". The objective of the study was to characterize the  
17 response to mass selection by an expert farmer in order to assess the potential of and  
18 limits to this approach for motivated farmers involved in PPB programs. These results  
19 show that mass selection within families can be effective for some traits even at early  
20 generations, and that genetic diversity is maintained within families for future selection.  
21 These results will contribute to improve farmers' understanding of the impact of their  
22 selection and thus may affect their future selection.

23

## 24 **Conclusion**

25 The objective of the farmer in this study was to select improved populations, based on  
26 farmers' criteria, while maintaining the potential for future selection within populations.

2

1 Based on our results this goal seems to have been achieved. This study has helped lay  
2 out the basis for the implementation of a PPB program by creating new populations with  
3 broad diversity which can then be distributed for selection according to farmers' criteria.  
4 The understanding and the analysis of the results were possible because of the interaction  
5 between farmers and the research team. This collaboration is the basis of the program and  
6 has lead to a better knowledge of farmer variety management and its impact on genetic  
7 diversity.

8

9 Selection on-farm is new for farmers in France (30). More time and exchanges of  
10 knowledge are needed for farmers to regain the knowledge and skills of selection (52).  
11 These newly created families were sent to farmers all over France and are now managed  
12 by around 25 farmers under different environments and practices. These farmers are  
13 collaborating in the PPB program which is the basis of the methodology. Several farmers  
14 have started mass selection within the populations. The next step will be to study the  
15 evolution of these newly created families under farmers' selection and management  
16 practices as well as evolutionary pressure in diverse environments in terms of molecular  
17 and phenotypic diversity and on-farm agronomic and quality traits.

18

## 19 **Acknowledgements**

20 We thank Jérôme Enjalbert, Mathieu Thomas for useful discussion. We thank Gerard  
21 Branlard from INRA Clermont Ferrand for his time when doing NIRS analysis and his help  
22 understanding quality measurements. P. Rivière was funded by a grant from DIM ASTREA  
23 (Ile de France region). J. Dawson was funded by an INRA postdoctoral research fellowship  
24 and during 5 months by a postdoctoral fellowship from the European Community's  
25 Seventh Framework Programme (FP7/ 2007–2013) under the Grant Agreement n245058-  
26 Solibam (Strategies for Organic and Low-input Integrated Breeding and Management).

1

---

1 The research leading to these results has received funding from the above mentioned FP7  
2 Solibam programme and from the Specific Targeted Research Project of the European  
3 Commission 6th Framework Program Priority 8.1 SSP: Opportunities for farm seed  
4 conservation, breeding and production Proposal/Contract no.: SSP-CT-2006-04434.

5

6

**References:**

1. Agence Bio. La réglementation en agriculture biologique: principes et spécificités 2010. [Internet]. <http://www.agencebio.org>. [cited 2 August 2010]. Available on: [http://www.agencebio.org/upload/pagesEdito/fichiers/synthese\\_reglementation0110.pdf](http://www.agencebio.org/upload/pagesEdito/fichiers/synthese_reglementation0110.pdf)
2. Wolfe M, Baresel J, Desclaux D, Goldringer I, Hoad S, Kovacs G, et al. 2008. Developments in breeding cereals for organic agriculture. *Euphytica*;163:323-346.
3. Morris ML, Bellon MR. 2004. Participatory Plant Breeding Research: Opportunities and challenges for the international crop improvement system. *Euphytica*. 136:21-35.
4. Alvarez N, Garine E, Khasah C, Dounias E, Hossaert-McKey M, McKey D. 2005. Farmers' practices, metapopulation dynamics, and conservation of agricultural biodiversity on-farm: a case study of sorghum among the Duupa in sub-Saharan Cameroon. *Biological Conservation*. 121:533-543.
5. Dawson J, Murphy K, Jones S. 2008. Decentralized selection and participatory approaches in plant breeding for low-input systems. *Euphytica* 160:143-154.
6. Ceccarelli S, Grando S. 2007. Decentralized-participatory plant breeding: an example of demand driven research. *Euphytica* 155:349-360.
7. Smith ME, Castillo G. F, Gomez F. 2001. Participatory plant breeding with maize in Mexico and Honduras. *Euphytica* 122:551-565.
8. Jarvis DI, Hodgkin T, Sthapit BR, Fadda C, Lopez-Noriega I. 2011. An Heuristic Framework for Identifying Multiple Ways of Supporting the Conservation and Use of Traditional Crop Varieties within the Agricultural Production System. *Critical Reviews in Plant Sciences* 30:125-176.
9. Ceccarelli S, Grando S, Bailey E, Amri A, El-Felah M, Nassif F, et al. 2001. Farmer participation in barley breeding in Syria, Morocco and Tunisia *Euphytica* 122:521-536.
10. Joshi KD, Witcombe JR. 2003. The impact of participatory plant breeding (PPB) on landrace diversity: A case study for high-altitude rice in Nepal. *Euphytica* 134:117-125.
11. Almekinders CJM, Elings. 2001. A. Collaboration of farmers and breeders: Participatory crop improvement in perspective. *Euphytica* 122:425-438.
12. Vom Brocke K, Trouche G, Weltzien E, Barro-Kondombo CP, Gozé E, Chantereau J. 2010. Participatory variety development for sorghum in Burkina Faso: Farmers' selection and farmers' criteria. *Field Crops Research* 119:183-194.
13. Patto MCV, Moreira PM, Almeida N, Satovic Z, Pego S. 2008. Genetic diversity evolution through participatory maize breeding in Portugal. *Euphytica* 161:283-291.
14. Desclaux D, Nolot JM, Chiffolleau Y, Leclerc C, Gozé E. 2008. Changes in the concept of genotype X environment interactions to fit agriculture diversification and decentralized participatory plant breeding: pluridisciplinary point of view. *Euphytica* 163:533-546.
15. Ceccarelli S, Grando S. 2009. Participatory Plant Breeding in Cereals. *Cereals*.

Springer New York. p. 1-20.

16. Halewood M, Deupmann P, Sthapit BR, Vernoy R, Ceccarelli S. 2009. Participatory Plant Breeding to Promote Farmer's Rights. Bioversity International.
17. Almekinders C, Centeno JM, Torrez RH, Olivera LSM, Suarez JMG, Carrasco JG. 2006. Experiencias y aprendizajes del desarrollo de variedades de frijol de manera participativa en el norte de Nicaragua. *Agronomía Mesoamericana*. 17:327-346.
18. Sperling L, Loevinsohn M, Ntabomvura B. 1993. Rethinking the farmer's role in plant breeding: local bean experts and on-station selection in Rwanda. *Experimental Agriculture*. 29:509-519.
19. McElhinny E, Peralta E, Mazón N, Danial DL, Thiele G, Lindhout P. 2007. Aspects of participatory plant breeding for quinoa in marginal areas of Ecuador. *Euphytica*. 153:373-384.
20. Bachmann L, Cruzada E, Wright S. 2009. Food Security and Farmer Empowerment. *Magsasaka at Siyentista Tungo sa Pag-unlad ng Agrikultura*.
21. Osman AM, Almekinders CJM, Struik PC, Bueren ETL van. 2008. Can conventional breeding programmes provide onion varieties that are suitable for organic farming in the Netherlands? *Euphytica*. 163:511-522.
22. Chable V, Conseil M, Serpolay E, Le Lagadec F. 2008. Organic varieties for cauliflowers and cabbages in Brittany: from genetic resources to participatory plant breeding. *Euphytica* 164:521-529.
23. Ghaouti L, Vogt-Kaute W, Link W. 2008. Development of locally-adapted faba bean cultivars for organic conditions in Germany through a participatory breeding approach. *Euphytica* 162:257-268.
24. Mendes Moreira PMR, Pêgo SE, Vaz Patto C, Hallauer AR. 2008. Comparison of selection methods on 'Pigarro', a Portuguese improved maize population with fasciation expression. *Euphytica* 163:481-499.
25. Dawson J, Rivière P, Berthelot JF, Mercier F, De Kochko P, Galic N, et al. 2011. Collaborative Plant Breeding for Organic Agricultural Systems in Developed Countries. *Sustainability* 3:1206-1223.
26. Thomas M, Dawson JC, Goldringer I, Bonneuil C. 2011. Seed exchanges, a key to analyze crop diversity dynamics in farmer-led on-farm conservation. *Genetic resources and Crop Evolution* 58:321-338.
27. Réseau Semences Paysannes. 2008. Voyage autour des blés paysans. *Témoignages*.
28. Enjalbert J, Dawson JC, Paillard S, Rhoné B, Rousselle Y, Thomas M, et al. 2011. Dynamic management of crop diversity: From an experimental approach to on-farm conservation. *Comptes Rendus Biologies*. 334:458-468.
29. Thomas M, Demeulenaere E, Dawson JC, Khan AR, Galic N, Jouanne-Pin S, et al. 2012. On-farm dynamic management of genetic diversity: the impact of seed diffusions and seed saving practices on a population-variety of bread wheat.

Evolutionary Applications 5:779-795.

30. Bonneuil C, Demeulenaere E. 2007. Vers une génétique de pair à pair? L'émergence de la sélection participative. Des sciences citoyennes? La question de l'amateur dans les sciences naturalistes. de l'aude. Des sciences citoyennes? La question de l'amateur dans les sciences naturalistes p. 122-147.
31. Goldringer I, Enjalbert J, David J, Paillard S, Pham JL, Brabant P. 2001. Dynamic management of genetic resources: a 13 year experiment on wheat. In: Cooper HD, Spillane C, Hodgkin T, éditeurs. Broadening the Genetic Base of Crop Production. Rome, Italy: IPGRI/FAO p. 245-260.
32. Newton AC, Akar T, Baresel JP, Bebeli PJ, Bettencourt E, Bladenopoulos KV, et al. 2010. Cereal landraces for sustainable agriculture. A review. Agronomy for Sustainable Development 30:237-269.
33. Hajjar R, Jarvis DI, Gemmill-Herren B. 2007. The utility of crop genetic diversity in maintaining ecosystem services. Agriculture, Ecosystems and Environment 123:261–270 .
34. Wolfe MS. 2000. Crop strength through diversity. Nature 406:681-682.
35. Finckh MR, Gacek ES, Goyeau H, Lannou C, Merz U, Mundt CC, et al. 2000. Cereal variety and species mixtures in practice, with emphasis on disease resistance. Agronomie 20:813-37.
36. Zhu Y, Chen H, Fan J, Wang Y, Li Y, Chen J, et al. 2000. Genetic diversity and disease control in rice. Nature 406:718-722.
37. Chateil C, Goldringer I, Tarallo L, Kerbiriou C, Le Viol I, Ponge J-F, et al. 2013. Crop genetic diversity benefits farmland biodiversity in cultivated fields. Agriculture, Ecosystems & Environment 171:25-32.
38. Tooker JF, Frank SD. 2012. Genotypically diverse cultivar mixtures for insect pest management and increased crop yields. Journal of Applied Ecology 49:974-985.
39. Mille B, Fraj M, Monod H, De Vallavieille-Pope C. 2006. Assessing Four-Way Mixtures of Winter Wheat Cultivars from the Performances of their Two-Way and Individual Components. European Journal of Plant Pathology 114:163-173.
40. Louette D, 2000. Smale M. Farmers' seed selection practices and traditional maize varieties in Cuzalapa, Mexico. Euphytica 113:25-41.
41. Lançon J, Pichaut JP, Djaboutou M, Lewicki-Dhainaut S, Viot C, Lacape JM. 2008. Use of molecular markers in participatory plant breeding: assessing the genetic variability in cotton populations bred by farmers. Annals of Applied Biology 152:113-119.
42. Pujol B, David P, McKey D. 2005. Microevolution in agricultural environments: how a traditional Amerindian farming practice favours heterozygosity in cassava (*Manihot esculenta* Crantz, Euphorbiaceae). Ecology Letters 8:138-47.
43. Ceccarelli S, Grando S, Tutwiler R, Baha J, Martini AM, Salahieh H, et al. 2000. A methodological study on participatory barley breeding I. Selection phase. Euphytica



111:91-104.

44. SAS. SAS Institute Incorporated. 2008. SAS/STAT User's Guide. Cary, NC: SAS Institute, Inc.
45. R Development Core Team. 2010. R: A Language and Environment for Statistical Computing. [Internet]. Vienna, Austria: R Foundation for Statistical Computing. Available on: <http://www.R-project.org>
46. Bassett M, Allan RE, Rubenthaler GL. 1989. Genotype × environment interactions on soft white winter wheat quality. *Agronomy Journal* 81:955-960.
47. Herndl M, White JW, Graeff S, Claupein W. 2008. The impact of vernalization requirement, photoperiod sensitivity and earliness perse on grain protein content of bread wheat (*Triticum aestivum* L.). *Euphytica* 163:309-320.
48. Murphy K, Lammer D, Lyon S, Carter B, Jones SS. 2005. Breeding for organic and low-input farming systems: an evolutionary-participatory breeding method for inbred cereal grains. *Renewable Agriculture and Food Systems* 20:48-55.
49. Gyawali S, Sunwar S, Subedi M, Tripathi M, Joshi KD, Witcombe JR. 2007. Collaborative breeding with farmers can be effective. 2007. *Field Crops Research* 101:88-95.
50. Virk DS, Chakraborty M, Ghosh J, Prasad SC, Witcombe JR. 2005. Increasing the client orientation of maize breeding using farmer participation in Eastern India. *Experimental Agriculture* 41:413-426.
51. Blum A. 1998. Improving wheat grain filling under stress by stem reserve mobilisation. *Euphytica* 100:77-83.
52. Storup B, De Kochko P. 2011. Suivi d'une expérience de sélection participative en cours l'exemple du projet européen SOLIBAM. [Internet]. <http://portail.documentation.developpement-durable.gouv.fr/> [cited 26 July 2013]. Available on: [http://portail.documentation.developpement-durable.gouv.fr/dri/document.xsp?id=Temis-0077262&n=22&q=%28%2B%28question\\_word%3Aapproche+question%3A|approche%29%29&order=descendant&sort=titre\\_field&](http://portail.documentation.developpement-durable.gouv.fr/dri/document.xsp?id=Temis-0077262&n=22&q=%28%2B%28question_word%3Aapproche+question%3A|approche%29%29&order=descendant&sort=titre_field&)

1

1 Figure 1. Evolution between selected (grey) and bulk (black) versions for four traits: (a.) :  
2 grain weight per spike (GW\_Spike) in grams, (b.) : plant height (PH) in cm, (c.) : thousand  
3 kernel weight (TKW) in grams and (d.) : protein concentration (Protein) in %. Stars  
4 represent significant differences between the means: \*:  $0.05 < p\text{-value} < 0.01$ ; \*\*:  $0.01 <$   
5  $p\text{-value} < 0.001$ ; \*\*\*:  $p\text{-value} < 0.001$ . tem= control (Renan).

6

7 Figure 2. Change in the phenotypic mean between bulk and selected versions of F3  
8 families for several traits. If the mean decreases: - :  $0.05 > p\text{-value} > 0.01$ ; -- :  $0.01 > p\text{-}$   
9  $\text{value} > 0.001$ ; --- :  $p\text{-value} < 0.001$ . If the mean increases: + :  $0.05 > p\text{-value} > 0.01$ ; ++ :  
10  $0.01 > p\text{-value} > 0.001$ ; +++ :  $p\text{-value} < 0.001$ . Earliness = days to flowering, so an increase  
11 means that a family flowers later. Color that increases means darker spike. Awns that  
12 increases means more awns.

2

1

1 Table 1. Traits measured. In bold are the traits analysed.

2 Field measurements:

Traits	Measures methods
<b>earliness</b>	number of degree-days when the main tiller spike is 50% emerged, recorded for each plant by daily observations in the field from 24/05/2009 to 19/06/2009, and converted to degree-days using meteorological data collected on the research station
<b>Awns</b>	on a scale from 0 (awnless) to 2 (fully awned)
<b>Color</b>	on a scale from 0 (white) to 2 (dark red)
<b>PH</b>	Plant height in cm, without the awns, at the top of the highest spikelet, even if sterile
SB	height in cm of the bottom of the lowest spikelet, even if it was sterile
LI	height in cm of the insertion of the stem into the flag-leaf in mm
<b>LLSD</b>	Last-Leaf-to-Spike-Distance = SB-LI

3

4 Spike measurements:

Traits	Measures methods
<b>SL</b>	Length of the spike in cm
<b>SpTot</b>	Total number of spikelets, including those that may be sterile or missing
SpMi	Missing spikelets that had fallen off in the field or the bag, used for correction of other measurements
SpSt	Sterile spikelets at the base or summit of the spike that had zero kernels
SpW	Weight of the spike before threshing measured to the nearest 0.01 g, after the stem was cut to the base of spike
KN	Number of kernels per spike counted after threshing for each spike individually
<b>KN_Spikelet</b>	$\text{Kernels per spikelet} = \text{KN} / (\text{SpTot} - \text{SpSt} - \text{SpMi})$
<b>KN_Spike</b>	$\text{KN} + \text{KN\_Spikelet} * \text{SpMi}$
GW	Grain weight measured for each spike to the nearest 0.01 g
<b>GW_Spike</b>	$\text{KN\_Spikelet} * \text{KN} / \text{GW}$
<b>TKW</b>	Thousand kernel weight = $\text{GW\_Spike} / \text{KN\_Spike} * 1000$
<b>density</b>	$\text{SpTot} / \text{SL}$ (spikelets/cm)
<b>sterility</b>	$\text{SpSt} / \text{KN\_Spike}$

2

1  
1

2 If the end of the spike was broken, the existing spike length was used for calculations of  
3 spike density, and the SL itself was treated as missing data for the analysis. If only the last  
4 spikelet was missing, 5mm was added to the spike length measurement.

5

6 Grain composition measurements:

Traits	Measures methods
<b>Protein</b>	Protein content (%)
<b>W</b>	Dough strength in $10^{-4}$ J. It is the force needed to break the dough.
<b>Hardness</b>	Hardness
<b>MTI</b>	Mixing Time Integral, associated with W
<b>TW</b>	Test weight, which is similar to mass density

7

8

9

2

1

1 Table 2. a. Summary for each trait for the data for selected version.

2

	PH (cm)	LLSD (cm)	SL (cm)	SpTot	KN_Spikelet	KN_Spike	GW_Spike (g)	TKW (g)	density	sterility	earliness	Protein (%)	hardness	TW	W
Min.	55.0	3.00	7.7	14	1.05	17.0	0.37	21.53	0.00	0.00	1155	7.6	25.0	76.6	127
1st Qu.	122.7	23.0	11.6	21	2.22	46.0	2.22	45.43	1.64	0.00	1266	9.0	51.3	79.5	184
Median	136.5	29.2	12.8	22	2.46	52.8	2.61	49.57	1.77	0.05	1302	9.5	57.0	80.7	216
Mean	134.5	28.6	12.8	23	2.47	52.9	2.62	49.51	1.80	0.06	1298	9.5	57.4	80.7	212
3rd Qu.	148.2	34.7	14.0	24	2.72	60.0	3.00	53.36	1.94	0.09	1329	10.0	65.0	81.7	241
Max.	172.4	51.6	18.8	29	3.72	85.6	5.10	91.02	2.74	0.97	1548	11.5	91.0	85.1	302

4 Table 2. b. Summary for each trait for the data for bulk version.

5

	PH (cm)	LLSD (cm)	SL (cm)	SpTot	KN_Spikelet	KN_Spike	GW_Spike (g)	TKW (g)	density	sterility	earliness	Protein (%)	hardness	TW	W
Min.	76.0	15.0	8.2	17	1.17	21.0	0.26	10.88	1.32	0.00	1155	7.9	31.0	76.3	118
1st Qu.	122.4	22.3	11.5	21	2.18	44.7	2.12	44.15	1.65	0.00	1246	9.3	51.8	78.7	190
Median	134.8	28.5	12.5	22	2.45	52.0	2.49	48.13	1.78	0.05	1285	9.7	59.0	80.3	223
Mean	132.7	27.7	12.7	23	2.45	52.7	2.53	47.98	1.81	0.05	1291	9.8	59.5	80.2	221
3rd Qu.	146.2	34.0	14.0	24	2.71	59.9	2.94	52.59	1.94	0.09	1329	10.3	66.0	81.5	253
Max.	173.3	49.8	18.3	30	4.23	97.4	4.23	64.21	2.81	0.35	1481	12.9	93.0	83.8	328

2

1

1 Table 3.a. Results of the ANOVA (model 1) with  $df(\text{family})=37$ ,  $df(\text{version})=1$ ,  $df(\text{rep})=2$ . F values are presented. Stars indicate the

2 significance of the F test; \*:  $0.05 > p\text{-value} > 0.01$ ; \*\*:  $0.01 > p\text{-value} > 0.001$ ; \*\*\* :  $p\text{-value} < 0.001$ .

	PH (cm)	LLSD (cm)	SL (cm)	SpTot	KN_Spikelet	KN_Spike	GW_Spike (g)	TKW (g)	density	sterility	earliness	Protein (%)	hardness	TW	W	MTI
family	32.37***	21.51***	7.39***	20.64***	4.68***	5.76***	3.44***	7.86***	5.3***	3.41***	15.12***	2.36***	3.95***	4.56***	3.88***	3.15***
rep	0.71	2.81	1.29	1.43	1.44	2.13	2.03	1.13	0.02	0.78	3.9*	10.03***	1.74	0.79	6.52**	12.41***
version	4.15*	2.1	0.66	0.24	0.19	0.19	5.1*	14.39***	1.48	0.19	3.2	11.38***	2.55	5.98*	4.44*	6.47*

Table 3.b. Results of the

5 ANOVA (model 2) with  $df(\text{family})=37$ ,  $df(\text{version in family})=38$ ,  $df(\text{rep})=2$ . F values are presented. Stars indicate the significance of the F

6 test; \*:  $0.05 > p\text{-value} > 0.01$ ; \*\*:  $0.01 > p\text{-value} > 0.001$ ; \*\*\* :  $p\text{-value} < 0.001$ .

7

	PH (cm)	LLSD (cm)	SL (cm)	SpTot	KN_Spikelet	KN_Spike	GW_Spike (g)	TKW (g)	density	sterility	earliness	Protein (%)	hardness	TW	W	MTI
family	34.54***	22.45***	4.10***	21.14***	4.66***	5.83***	3.48***	7.95***	5.65***	3.48***	15.20***	2.36***	3.78***	4.77***	3.78***	3.05***
rep	0.76	2.93	2.39	2.96	1.28	1.29	2.15	2.05	1.15	3.9*	0.28	10.04***	1.66	0.83	6.35**	12.02***
version (family)	2.83***	2.15***	1.47*	1.74*	0.79	1.3	1.38	1.66**	1.4	1.2	1.47*	1.28	0.83	1.37	0.96	0.99

2

1

1 Table 4. Differences in repeatability, genetic variance among F<sub>3</sub> and genetic variance within F<sub>3</sub> for each version: bulk and selected for2 each trait;  $\sigma = \sqrt{\text{variance}}$  .

3

4

5

6

	PH (cm)	LLSD (cm)	SL (cm)	SpTot	KN_Spikelet	KN_Spike	GW_Spike (g)	TKW (g)
repeatability bulk	0.81	0.77	0.43	0.68	0.29	0.37	0.13	0.36
repeatability S	0.79	0.67	0.45	0.70	0.26	0.35	0.32	0.51
% between repeatability S and bulk	-2.32%	-13.18%	+4.61%	+3.62%	-10.63%	-4.84%	+139.39%	+42.8%
$\sigma$ among-family bulk	14.75	6.55	0.81	1.46	0.15	4.88	0.14	2.61
$\sigma$ among-family S	13.62	5.24	0.83	1.43	0.13	4.47	0.23	3.28
% between $\sigma$ among-family S and bulk	-7.47%	-19.96%	+2.14%	-1.48%	-11.6%	-8.38%	+65.96%	+25.36%
$\sigma$ within-family bulk	11.06	5.50	1.35	0.77	0.00	0.00	0.37	4.90
$\sigma$ within-family S	10.82	5.70	1.31	0.43	0.00	0.00	0.32	4.25
% between $\sigma$ within-family bulk and S	-2.18%	+3.6%	-2.75%	-44.16%	0%	0%	-14.22%	-13.32%

	density	sterility	earliness	Protein (%)	hardness	SW	W	MTI
repeatability bulk	0.36	0.31	0.59	0.36	0.62	0.66	0.54	0.34
repeatability S	0.34	0.22	0.54	0.48	0.50	0.68	0.60	0.60
% between r S and r bulk	-5.21%	-28.35%	-8.74%	+32.2%	-19.26%	+3.93%	+10.53%	+74.28%
$\sigma$ among-family bulk	0.09	0.02	33.10	0.34	7.12	1.14	23.54	10.43
$\sigma$ among-family S	0.09	0.02	30.27	0.36	5.57	1.11	21.74	15.56
% between $\sigma$ among-family S and bulk	-1.12%	-0.52%	-8.55%	+6.64%	-21.77%	-1.99%	-7.64%	+49.3%
$\sigma$ within-family bulk	0.15	0.05	24.59	0.51	7.80	0.84	18.32	22.20
$\sigma$ within-family S	0.16	0.06	26.08	0.28	7.75	0.64	0.00	18.86
% between $\sigma$ within-family bulk and S	+6.09%	+32.69%	+6.03%	-44.73%	-0.57%	-23.64%	0%	-15.03%

2

1  
1 **Supplementary methods and results**

2  
3 *Correlations among traits:*

4 Correlations among traits were calculated with the software R and the lm procedure (1) on  
5 the selected version on one hand and bulk versions on the other hand for all the family.  
6 Checks were not included in this calculation. In order to obtain correlations between all  
7 traits, they have been calculated on the mean over the 15 plants.

8 Differences in slope for each correlation have been calculated as follows:

9 
$$t_{test} = \frac{b_1 - b_2}{\sqrt{s_c^2 \left[ \frac{1}{\sum (x_{1l} - \bar{x}_1)^2} + \frac{1}{\sum (x_{2l} - \bar{x}_2)^2} \right]}}$$

10 With  $s_c^2$ , the common residual variance (SCE1r + SCE2r)/(n1+n2-4);  $b_1$  the coefficient of the  
11 regression for the bulk version;  $b_2$  the coefficient of the regression for the selected version  
12  $x_{1l}$ , the value of individual  $l$  for variable  $x_1$ .  $\bar{x}_1$  is the mean of variable  $x_1$ . The same for  
13 variable 2. We concluded that there were no significant differences between the slopes  
14 when the t-test value was in the confidence interval of 5%.

15  
16 The correlations within the selected versions and within the bulk version were similar  
17 except for TKW and KN\_Spike, GW\_Spike and protein, W and MTI (Supplementary Table  
18 1).

19 PH and LLSD were highly correlated (r between 0.83 and 0.90). In addition, there were 2  
20 groups of correlated traits: (1) GW\_Spike, KN\_Spikelet, KN\_Spike, SpTot, SL and (2)  
21 protein, W and MTI (Supplementary Table 1).

22  
23 *Estimation of within-family genetic variation:*

24 We used the check, Renan, which is a pure line, i.e. genetically homogeneous, to assess  
25 the within-plot environmental variation. As the check was placed randomly twice in each



1  
1 replication (6 plots in total) we assumed it captured a significant part of the environmental  
2 variation present in the field.  
3 We used the following model:  
4  $C_{klm} = \mu + plot_k + rep_l + E_{klm}$  (4)  
5 Where  $C_{klm}$  is the phenotypic value measured on the check varieties  $m$  in plot  $k$  in  
6 repetition  $l$ ,  $\mu$  is the general mean,  $plot_k$  is the plot where the check was cultivated (one row  
7 has several plants),  $rep_l$  is the replication and  $E_{klm}$  is the experimental error, corresponding  
8 to within-row environmental variation. As those two variances were estimated  
9 independently, the within-family genetic variance was estimated as the difference between  
10  $var(R)$  from model (3) and  $var(E)$  from model (4). For the three cases where the within-  
11 family genetic variance was estimated to be slightly negative, it was set to zero.  
12 The estimation of the within-family average environmental and genetic variances is based  
13 on several assumptions. Our control is a modern variety that we assume represents a  
14 single genotype but in fact it may not be 100% genetically homogeneous. As there were  
15 only 6 lines of the check in the trial, this may not be enough to catch the environmental  
16 variation. On the other hand, we have seen that the repetition effect is negligible for almost  
17 all traits (Table 3). Yet, the differences among rows may be smaller than the differences  
18 within rows (micro environmental variation). This may also explain why we occasionally  
19 found negative values for the average within-family genetic variance (Table 4). One limit of  
20 the estimation of diversity is that it has been done from the phenotypic data. A molecular  
21 approach would be more precise and give us complementary information on the genetic  
22 diversity of these families. Genetic diversity is currently being assessed with molecular  
23 markers on a set of families issued from farmers' selection at subsequent generations.  
24 Yet, as farmers select on the observed phenotypic variation, genetic variation estimated  
25 from phenotypic measures is relevant to help design their selection practices.

1  
1

## 2 **Supplementary discussion**

### 3 *Response to selection and drift*

4 We did not test in this experiment that trait evolution was due specifically to selection and  
5 not to genetic drift which might arise due to the limited number of spikes (from 25 to 50) in  
6 the selected samples. Drift would also affect the bulk version which is derived from seed  
7 sample of the same size. However, when changes were in the same direction for several  
8 families (general version effect) or when they corresponded to a selection objective of the  
9 farmer, it is more likely to be due to the effects of selection than of random drift.

10

### 11 *Uniform response to selection*

12 Selection tended to increase PH, except for one family in which plant height decreased  
13 significantly (Figure 1.b). Both among- and within-family genetic variances slightly  
14 decreased. Stabilizing plant height around an optimum of 100cm to limit lodging was an  
15 important objective of JFB in crossing landraces with more recent varieties. In practice, the  
16 selection also took other factors into account. For example, family 64 was taller than  
17 optimal and still increased for PH after selection because, despite its height, it was very  
18 resistant to lodging due to its strong stem. Tall plants with strong stems that do not lodge  
19 are interesting for farmers because of better competition with weeds or volunteers from  
20 previous crops. This is corroborated in the scientific literature (2). In other cases such as  
21 family 34, the farmer wanted to discard the tall plants but we found that selection  
22 increased PH. During discussions with the farmers, it appears that they naturally select  
23 spikes that are more visible, i.e. above the average height of the family. While they are  
24 aware of this, they initially did not pay enough attention to that point when selecting plants  
25 within families that have very heterogeneous heights. The selection was carried out in

2

1  
1 2008, since then, farmers now focus more on this. If the experiment would carried out  
2 again with selections from the current year, the changes in PH might be more dependent  
3 on the initial average height of the population rather than always showing increases when  
4 significant (except for family 24).

5  
6 *Differential response to selection*  
7 For some traits, an increase in among-family variance corresponded to differential  
8 responsiveness to selection in the same direction. For instance, the among-family genetic  
9 variance of TKW and of GW\_Spike increased with selection and the average within-family  
10 variance decreased, while there was also a positive overall response to selection.  
11 Increasing TKW and GW\_Spike was in fact an objective of JFB's selection and this  
12 selection was generally effective across families, possibly due to a high genetic variability  
13 available for these traits within families. When selection was not efficient within some  
14 families this could be due to a lack of diversity within some families or to a large  
15 environmental effect compared to the genotypic effect (GW\_Spike and TKW have among  
16 the lowest repeatability).

17  
18 *Multi-trait dimension of the response to selection*  
19 Correlations between traits may either contribute to response to selection, or negative  
20 correlations may reduce the response to selection of traits that are negatively correlated.  
21 Using indirect selection can be useful in the case of traits that are difficult to assess in the  
22 field (i.e. traits that are not visible such as protein or TKW). These traits can be evaluated  
23 over time by farmers engaged in milling and baking, but are difficult to select when  
24 choosing individual plants in the field. In this study, most traits linked to grain composition  
25 and technological transformation were only slightly correlated to morphological traits and

1  
1 were negatively correlated with production traits. We take the case of protein and TKW as  
2 examples (Supplementary Table 1). Protein is important since these populations are often  
3 developed with the objective of improving bread-making quality. There was a tendency for  
4 protein to decrease after selection for 30 families out of 38 (2 significantly), and a tendency  
5 for TKW to increase after selection for 26 families over 35 (4 significantly). Moreover, the  
6 three families that changed most in the unfavorable direction for protein content (90, 80  
7 and 53) were among the six families that best responded for TKW and GW\_Spike (Figure  
8 1.a & 1.c.). This is consistent with results from the literature showing that yield is  
9 negatively correlated to protein content (3).

10  
11 This negative correlation is not a major concern for JFB and many of the farmers in the  
12 Réseau Semences Paysannes since his strategy is to cultivate a certain number of  
13 varieties as a mixture, where the quality of the best varieties will compensate for lower  
14 quality varieties. As a baker, he knows how to deal with varieties that have lower values of  
15 technical quality measures compared with those needed for industrial baking, and  
16 emphasizes the need for organoleptic and nutritional quality. So, first, there is a selection  
17 of populations that fit his agronomic practices and environment. Then the bread-making  
18 quality will be evaluated. This general approach is common to many conventional and  
19 participatory breeding programs: choose high-quality parents, make crosses and evaluate  
20 the progeny for agronomic traits, then evaluate selected populations for bread-making  
21 quality (4).

22  
23 Finally, it is interesting to note that, after selection, some correlations significantly changed.  
24 The correlation between TKW and KN\_spikelet, and between GW\_Spike and protein,  
25 respectively changed from -0.14 to +0.30 and from -0.40 to 0.04 indicating that selection

1  
1 broke the undesirable negative correlation between traits. Some families responded to  
2 selection for one of the traits without diminishing the other (Family 60 for example  
3 significantly increased in protein and remained stable for GW\_Spike and TKW).

4

#### 5 *Influence of parental varieties on the selection response*

6 The choice of parents is very important in achieving breeding goals (5). In this program,  
7 the parents were mainly chosen according to their agronomic performance and quality in  
8 on-farm breadmaking. When selection acts on qualitative traits such as spike color it is  
9 easy to observe the contributions of each parent. For example, in family 64, a cross  
10 between Vieux Ferrette and Rouge de Clervaux, the red spike color of Rouge de Clervaux  
11 was selected. Parental contributions are more subtle for quantitative traits. The progeny of  
12 the cross between Blé de Langogne (male) and Roisel (female) (family 80), showed high  
13 responsiveness (12 traits out of 18 measured). The two parents used are very different. Blé  
14 de Langogne is a landrace, with no awns, a large stem, tall spike, low spikelet density, tall  
15 plants, very heterogeneous, while the characteristics of Roisel are: awns, dense spikes  
16 (high number of spikelets per spike), and rather tall plants (because of Russian  
17 genealogy). Roisel is a modern cultivar (released in 1967) with higher values of baking  
18 technical properties (higher W).

19 Family 80 was highly heterogeneous compared to the other families. Roisel was  
20 “dominant” according to JFB, meaning that most of its attributes are seen in the family (for  
21 visible traits). The selection significantly decreased the density (Blé de Langogne type)  
22 and significantly increased the awns (Roisel type). Selection also significantly decreased  
23 W, which is related to the selection of individuals closer to the Blé de Langogne type. The  
24 changes observed may be due to the large differences between the parents which led to a  
25 heterogeneous family compared to the other (i.e. individuals different for awns, color,

2

1  
1 spike). The selection favored individuals sharing attributes of both parents (for example  
2 awns from Roisel and low spike density from Blé de Langogne), which was the objective of  
3 this selection.

4  
5 *Range of variation for baking force and artisanal bread-making*  
6 W is related to the energy the dough can withstand during mixing. We found a large  
7 variation for W with values rather low compared to conventional varieties (6). In a study  
8 conducted under similar conditions (on a research station with conventional management),  
9 Branlard et al. (6) estimated a W mean of 236 with a maximum at 563 within a panel of  
10 130 modern cultivars. In our study, the mean is 221 in the bulk families before selection  
11 and 212 in the selected versions, with a maximum of 328. While these values and  
12 variation would be problematic for industrialized baking, it is not the case of artisanal baker  
13 in the project who can adapt their baking technique to the variety. These techniques  
14 involve the use of natural levain (sourdough), less intensive mixing and longer  
15 fermentation and baking times.

16 The values of W we observed may be related to the parents used to create those families,  
17 especially those from landraces and historic varieties. These varieties were originally  
18 baked with artisanal practices and therefore they have not been selected for a high W,  
19 which is only required in industrial baking (7). To better assess baking force, the  
20 populations should be evaluated on several farms, since W is correlated to protein ( $r=0.61$ ,  
21 Supplementary Table 1) and protein content is very dependent on the environment  
22 (primarily differences in nitrogen level) and the interaction between genotype and  
23 environment (8,3). The relevance of measuring W for artisanal bakers is a major point of  
24 discussion among participants in the PPB program but it is a common standard in  
25 conventional breeding programs.

1

1

## 2 Supplementary References:

1. R Development Core Team. R: A Language and Environment for Statistical Computing. 2010. [Internet]. Vienna, Austria: R Foundation for Statistical Computing. Available on: <http://www.R-project.org>
2. Hoad S, Bertholdsson NO, Neuhoﬀ D, Köpke U. Approches to Breed for Improved Weed Suppression in Organically Grown Cereals. 2012. Organic Crop Breeding. Wiley-Backwell. Edith T. Lammerts van Bueren and James R. Myers.
3. Herndl M, White JW, Graeff S, Claupein W. The impact of vernalization requirement, photoperiod sensitivity and earliness perse on grain protein content of bread wheat (*Triticum aestivum* L.). 2008. *Euphytica* 163:309-320.
4. Murphy K, Lammer D, Lyon S, Carter B, Jones SS. Breeding for organic and low-input farming systems: an evolutionary-participatory breeding method for inbred cereal grains. 2005. *Renewable Agriculture and Food Systems* 20:48-55.
5. Wolfe M, Baresel J, Desclaux D, Goldringer I, Hoad S, Kovacs G, et al. 2008. Developments in breeding cereals for organic agriculture. *Euphytica*;163:323-346.
6. Branlard G, Oury FX, Martre P, Aussenac T, Regnard J-B. Étude et prédiction des fractions protéiques de réserve et indices de sélection pour la qualité. 2008. Synthèse des programmes de recherche FSOV, Actes de la rencontre scientifique p. 68-73.
7. Réseau Semences Paysannes. Voyage autour des blés paysans. 2008. Témoignages.
8. Bassett M, Allan RE, Rubenthaler GL. Genotype × environment interactions on soft white winter wheat quality. 1989. *Agronomy Journal* 81:955-960.

**Réponse à la sélection massive paysanne en génération précoce dans des populations de blé tendre issues de croisements**

1  
 1 Supplementary Table 1. Correlation between traits (version selected are above gray cells  
 2 in bold, version bulk are under the grey cell), r is represented. When the slopes were  
 3 different, the r value are in a black box.

	PH	LLSD	SL	SpTot	KN_Spikelet	KN_Spike	GW_Spike	TKW	density	sterility	earliness	Protein	Hardness	TW	W	Mti
PH		<b>0.83</b>	<b>0.04</b>	<b>0.1</b>	<b>-0.38</b>	<b>-0.18</b>	<b>-0.14</b>	<b>-0.01</b>	<b>-0.01</b>	<b>-0.07</b>	<b>0.35</b>	<b>-0.32</b>	<b>-0.18</b>	<b>0.21</b>	<b>-0.27</b>	<b>-0.40</b>
LLSD	0.90		<b>-0.02</b>	<b>-0.06</b>	<b>-0.42</b>	<b>-0.29</b>	<b>-0.27</b>	<b>-0.04</b>	<b>-0.09</b>	<b>-0.09</b>	<b>0.06</b>	<b>-0.27</b>	<b>-0.23</b>	<b>0.17</b>	<b>-0.25</b>	<b>-0.39</b>
SL	0.03	-0.02		<b>0.68</b>	<b>0.36</b>	<b>0.67</b>	<b>0.51</b>	<b>-0.13</b>	<b>-0.54</b>	<b>0.03</b>	<b>0.10</b>	<b>-0.20</b>	<b>0.07</b>	<b>0.17</b>	<b>-0.10</b>	<b>-0.05</b>
SpTot	0.08	-0.02	0.66		<b>0.20</b>	<b>0.70</b>	<b>0.34</b>	<b>-0.42</b>	<b>0.24</b>	<b>0.34</b>	<b>0.22</b>	<b>-0.41</b>	<b>0.16</b>	<b>0.39</b>	<b>-0.12</b>	<b>-0.07</b>
KN_Spikelet	-0.43	-0.39	0.41	0.20		<b>0.80</b>	<b>0.70</b>	<b>-0.02</b>	<b>-0.23</b>	<b>-0.05</b>	<b>0.07</b>	<b>0.03</b>	<b>0.26</b>	<b>-0.10</b>	<b>0.03</b>	<b>0.17</b>
KN_Spike	-0.32	-0.31	0.68	0.65	0.85		<b>0.77</b>	<b>-0.18</b>	<b>-0.08</b>	<b>-0.06</b>	<b>0.09</b>	<b>-0.16</b>	<b>0.26</b>	<b>0.10</b>	<b>-0.02</b>	<b>0.09</b>
GW_Spike	-0.23	-0.28	0.60	0.46	0.67	0.80		<b>0.48</b>	<b>-0.26</b>	<b>-0.37</b>	<b>-0.02</b>	<b>0.04</b>	<b>0.07</b>	<b>0.06</b>	<b>0.03</b>	<b>0.16</b>
TKW	0.16	0.06	-0.23	-0.39	<b>-0.41</b>	-0.48	0.15		<b>-0.27</b>	<b>-0.48</b>	<b>-0.17</b>	<b>0.30</b>	<b>-0.23</b>	<b>-0.08</b>	<b>0.11</b>	<b>0.16</b>
density	0.05	0.01	-0.52	0.28	-0.30	-0.14	-0.25	-0.14		<b>0.34</b>	<b>0.08</b>	<b>-0.2</b>	<b>0.1</b>	<b>0.2</b>	<b>0.04</b>	<b>0.04</b>
sterility	0.32	0.14	0.02	0.45	-0.19	-0.07	-0.22	-0.20	0.50		<b>0.32</b>	<b>-0.41</b>	<b>0.06</b>	<b>0.33</b>	<b>-0.16</b>	<b>-0.09</b>
earliness	0.43	0.20	0.05	0.22	-0.07	-0.06	-0.23	-0.26	0.19	0.63		<b>-0.20</b>	<b>-0.14</b>	<b>0.12</b>	<b>-0.33</b>	<b>-0.29</b>
Protein	-0.24	-0.21	-0.31	-0.41	-0.11	-0.28	<b>-0.40</b>	-0.13	-0.05	-0.21	-0.10		<b>0.08</b>	<b>-0.51</b>	<b>0.61</b>	<b>0.63</b>
Hardness	-0.34	-0.35	0.24	0.26	0.38	0.43	0.33	-0.23	-0.03	-0.04	-0.14	0.07		<b>0.07</b>	<b>0.51</b>	<b>0.35</b>
TW	0.22	0.16	0.19	0.31	0.06	0.14	0.11	-0.04	0.06	0.41	0.20	-0.35	0.09		<b>-0.01</b>	<b>0.01</b>
W	-0.21	-0.18	-0.23	-0.26	0.11	-0.05	-0.20	-0.22	-0.02	-0.16	-0.08	0.65	0.57	0.08		<b>0.85</b>
Mti	-0.22	-0.18	-0.28	-0.27	0.03	-0.14	-0.24	-0.13	0.02	-0.09	-0.12	0.59	0.42	0.27	<b>0.89</b>	

5

6



# **Annexe K : Modélisation hiérarchique de la variance pour des essais déséquilibrés : application à des expérimentations flexibles de sélection décentralisée et participative**

Nous souhaitons publier dans la revue Field Crop Research, qui traite de l'amélioration des plantes. Cet article est presque prêt à être soumis. Il sera un peu remanié pour être moins technique sur les aspects statistiques. Les aspects plus techniques et des figures seront placés dans les données supplémentaires.

1 Hierarchical modeling of variance for flexible experiments  
2 in decentralized participatory plant breeding

3 Pierre Rivière<sup>1</sup> Julie C. Dawson<sup>2</sup> Isabelle Goldringer<sup>1</sup>

4 Olivier David<sup>3</sup>

5 27 novembre 2013

6 <sup>1</sup> INRA Le Moulon, UMR 0320 de Génétique Végétale, Ferme du Moulon F-91190 Gif sur  
7 Yvette, France

8 <sup>2</sup> Department of Horticulture, University of Wisconsin-Madison, 1575 Linden Dr. Madison,  
9 WI 53706, USA

10 <sup>3</sup> INRA, UR 341 de Mathématiques et Informatique Appliquées, Domaine de Vilvert  
11 F-78352 Jouy en Josas, France

12 **Abstract**

13 Decentralized participatory plant breeding (PPB) is a method for creating new va-  
14 rieties adapted to agroecology based systems. Selection is decentralized in the target  
15 environments and relies on collaboration between farmers, NGOs and researchers. In  
16 the wheat PPB project carried out in France, the decentralization is extreme as each  
17 farmer participating in the project has its own experiment on farm. This leads to unba-  
18 lanced designs with few residual degrees of freedom that require appropriate statistical  
19 analyses. In this paper we investigate a hierarchical Bayesian modeling of trial residual  
20 variances for a large network of 48 environments (an environment being a farm in a

1 year). We found that the hierarchical approach brought robust results leading to reliable  
2 mean comparisons on farms. It allows farmers to compare population-varieties on their  
3 farm and to carry out selection. This was possible because the farms shared a common  
4 experimental design and the number of environments in the network was large. For  
5 each variable studied, repeatability on farm was high enough to carry out selection (i.e.  
6 differentiation of populations was possible). Our approach benefits farmers who cannot  
7 set up replicated trials (because of space, time and material available) on their farm  
8 but wish to participate in the programme. The only constraint being the replication of  
9 a control population, farmers can choose which populations to test : this allows for the  
10 selection of a wide range of diversity over the network of farms. By developing such an  
11 approach, we give new tools to deal with on farm decentralized breeding that can be  
12 implemented for a large range of species. R scripts are available under request.

13 **Keywords** : decentralized participatory plant breeding ; unbalanced data ; hierar-  
14 chical model ; MCMC ; bread wheat ; bayesian statistics.

# 1 Introduction

Agroecology has been proposed as a strategy for sustainable agriculture (De Schutter, 2010). New varieties must be created to be part of an agroecological system which "addresses the production by a biodiverse agroecosystem able to sponsor its own functioning" as well as ensure participation of farmers to answer their needs (Altieri, 2002). These new varieties should have better resilience and stability to cope with rapid and unpredictable environment variations (biotic abiotic as well as societal) (Dawson et al., 2008; Shiferaw et al., 2013).

Decentralized participatory plant breeding (PPB) has been proposed to create such varieties (Ceccarelli et al., 2001).

In decentralized selection, selection is done directly in the target environments (i.e. on farm). This approach takes advantage of genotype  $\times$  environment interactions that are important as the environment is no longer controlled by inputs (Dawson et al., 2008). This method has already been shown to be very effective (Ceccarelli et al., 2000a; Smith et al., 2001; Virk et al., 2005).

In participatory selection, farmers do the selection in early generations in collaboration with researchers (Witcombe et al., 1996). This increases the efficiency of the breeding programme as farmers bring their own knowledge of their environment and of the required traits in this environment for their varieties (Ceccarelli and Grando, 2007). In addition, farmers are involved in the decision process leading to empowerment and new governance of the seed system (Bonneuil and Demeulenaere, 2007; Morris and Bellon, 2004).

PPB programmes have been set up worldwide on different species (Ceccarelli et al., 2000a; Smith et al., 2001; Virk et al., 2005), including in western countries (Mendes Moreira et al., 2008; Chable et al., 2008; Dawson et al., 2011; Rivière et al., 2013).

A wheat PPB programme started in France in 2005. It is a collaboration between the DEAP (Diversity Evolution and Adaptation of Population) team at INRA Le Moulon and the Réseau Semences Paysannes (RSP, the Farmers' seed network). Initially, 90 populations were created by crossing parents chosen by Jean François Berthelot (JFB),

1 a farmer-baker.

2 These 90 populations were spread in 2008 in a network of farms in France belonging  
3 to the RSP. This PPB programme had three objectives : (i) develop population-varieties  
4 adapted to farmers' practices and needs (organic management, bread quality ...) in a  
5 participatory breeding programme, (ii) develop strategies for preserving genetic diver-  
6 sity through on-farm dynamic management and plant breeding and (iii) learn and im-  
7 prove farmers' breeding techniques and diffuse it broadly. This last point is original as  
8 it has not been the priority of farmers in other projects ([Almekinders and Elings, 2001](#)).

9

10 In this article, we focus on the first objective that consists in developing population-  
11 varieties. Such varieties are genetically diverse and are expected to be more adapted  
12 to stochastic environments ([Newton et al., 2010](#); [Hajjar et al., 2008](#); [Wolfe et al., 2008](#);  
13 [Ceccarelli et al., 2010](#)).

14 As farmers participating in the project had organic systems spanning a large range  
15 of conditions, selection was decentralized in order to develop population-varieties that  
16 would suit their requirements. According to [Ceccarelli \(2012\)](#), the choice of an expe-  
17 rimental design is dictated by the amount of seed and the space available to do the  
18 experiment. In the wheat PPB project, the main problem was the space available on-  
19 farm, as these breeding activities had to find a place in the activities of the farm. This  
20 led to original designs with few replicated populations which made standard estimates  
21 of residual variances and population comparisons unstable (due to few residual degrees  
22 of freedom).

23 Several statistical methods for analysing data with a small number of replications  
24 have been proposed. In plant breeding, unreplicated trials are often analysed using  
25 spatial or linear modeling ([Kempton and Fox, 1997](#); [Ceccarelli et al., 2000b](#); [Singh  
26 et al., 2003](#); [Virk and Witcombe, 2007](#)).

27 However these methods usually require that some control varieties are sufficiently  
28 replicated which was not the case in the wheat PPB programme. Indeed, our designs  
29 were closer to micro-array experiments which also have small numbers of replicates.  
30 The models used to analyse their data involve additional assumptions to stabilise esti-

1 mates of gene variances. In some models, gene variances are assumed to be the same  
2 (Kerr, 2003) or to take a given number of values (Delmar et al., 2005). In other mo-  
3 dels, a separate variance is allowed for each gene but gene variances are assumed to  
4 come from a common distribution. Such hierarchical models take advantage of the pa-  
5 rallel structure of the data, that share the same organization (Baldi and Long, 2001).  
6 Bayesian methods are well adapted for such models as they assume that parameters  
7 are random variables (Smyth, 2004; Lewin et al., 2006). Hierarchical models allow ro-  
8 bust estimation of residual variances, however they have not been applied to variety  
9 trials. While bayesian statistics have been used in agriculture experiments (Besag and  
10 Higdon, 1999; Theobald et al., 2002), they have generally been applied to balanced data.

11

12 In this article, we develop a hierarchical Bayesian model to perform mean compa-  
13 risons on farm with the highly unbalanced trials of the wheat PPB programme. First,  
14 a hierarchical model for residual variances is developed and fitted to the data using a  
15 Bayesian approach. Then, the developed model is compared to a non hierarchical model  
16 and the between-year variation of residual variances is quantified. Finally the influence  
17 of the number of environments on the results of the analysis is studied.

## 1      **2    Materials and methods**

2            In this study a germplasm was a biological entity which gathered plants that be-  
3            longed to a landrace, an old varietie, a mixture, a farmers' mass selection or a derived  
4            cross population. A population was a given germplasm on one farm in one year.

### 5      **2.1    Experimental design and data set**

#### 6      **2.1.1    Experimental design**

7            All the farmers in the project grew wheat under organic farming methods. They  
8            applied the same farming practices on the experimental trial and on their other wheat  
9            fields. The farm network was divided into seven geographic regions. In each region there  
10           were two types of farms based on their experimental design : one or two regional farms  
11           and several satellite farms. Farmers chose to be a regional farm or a satellite farm mainly  
12           based on their experience. This could change from one year to another. Regional farms  
13           had between 20 and 32 plots and were visited each summer during "summer field visits".  
14           The visits gathered the research team, the farmers, the bakers, the citizens. Satellite  
15           farms had between 5 and 20 plots. For 2011-2012, there were nine regional farms and  
16           17 satellite farms (among which 14 sent phenotypic observation back) (Figure 1).



Figure 1 – Farms participating in the programme for 2011-2012 : regional farms (red), satellite farms (blue), associations from RSP (yellow) and INRA Le Moulon (green).

1 Experimental designs were set up in collaboration between farmers and the research  
 2 team (Dawson et al., 2011). It was a compromise between the number of populations  
 3 and the number of replications (Storup and de Kochko, 2011).

4 A farmer's variety, *Rouge-du-Roc*, was chosen as control by farmers. On satellite  
 5 farms, Rouge du Roc was replicated twice. Regional farms had two blocks and four  
 6 controls present in each block : Rouge du Roc, *Renan* (a pure line widely used in orga-  
 7 nic agriculture in France); *C14* and *C21* (populations derived from crosses). Farmers  
 8 chose the other populations to be sown that were not replicated (Figure 2). In 2009-  
 9 2010, there were only satellite farms. The organisation with regional farms and satellite  
 10 farms started in 2010-2011.



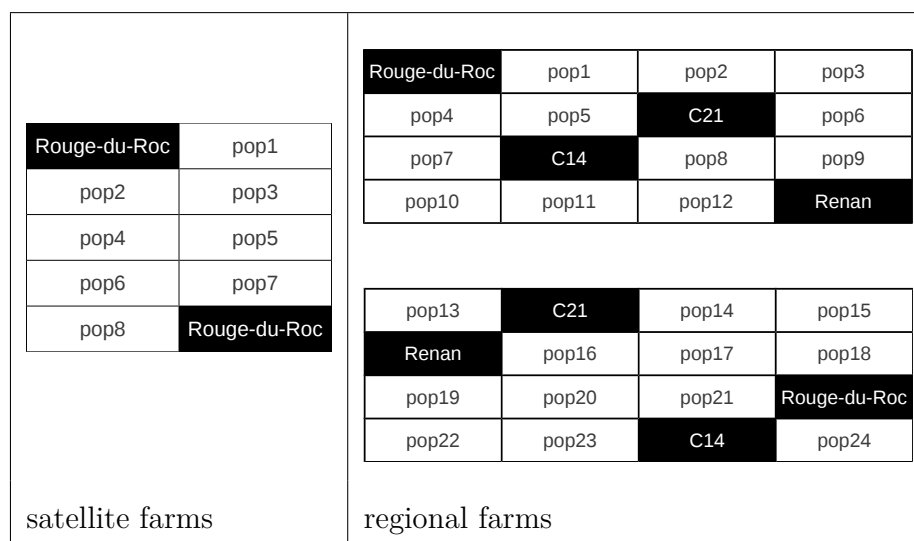


Figure 2 – Experimental design of satellite and regional farms. Controls replicated are in black boxes. pop stands for "tested population".

1           The residual had only three degrees of freedom ( $df$ ) for regional farms and one  $df$   
 2           for satellite farms, leading to a poor estimation of the residual variance and to a lack  
 3           of power for comparing populations. A minimum of 30 residual  $df$  is recommended to  
 4           have a good estimate of the environmental variation (Ceccarelli, 2012).

5           Overall, there were 254 different germplasms tested in 75 environments. However  
 6           only 48 environments followed the experimental design presented in Figure 2 and were  
 7           used in the analysis, which gathered 232 germplasms (Figure 3). 95% of germplasm  $\times$   
 8           environment combinations were missing, thus the data were severely non orthogonal.

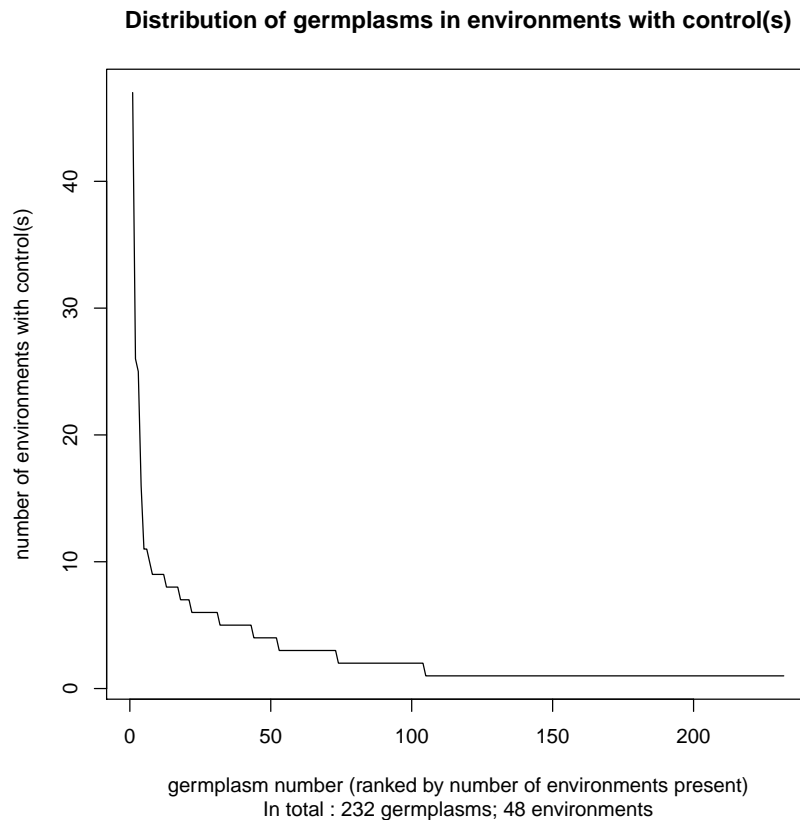


Figure 3 – Number of environments (farm×year with control replicated) where each germplasm was tested. Most of the germplasms were only tested in one or a few environments.

1           A constraint was the lack of facilities available for sowing and harvesting leading to  
 2 smaller sized trials. The majority of the farmers involved in the project did not have  
 3 machines to sow and harvest. The farmers manage their experiments from sowing to  
 4 harvest. Plot size was between five and ten m<sup>2</sup> and the experiment was set up in a wheat  
 5 field on farm so that all the agronomic practices of the field were the same on the trial.  
 6 This allowed for selection according to farmers' practices. In order to have as much  
 7 information as possible, we used the data from three seasons : 2009-2010, 2010-2011  
 8 and 2011-2012. We considered as an environment a combination of a farm and a year. A  
 9 farm was denoted by three letters. Environments then had the form FRC2011 for farm  
 10 FRC and season 2010-2011, 2010 being the sowing date and 2011 the harvesting date.  
 11 In the following, we will only use the harvesting year to denote seasons, for example  
 12 2011 will correspond to 2010-2011.

1 The proportion of regional farms and satellite farms varied during the project (Table  
2 1).

year	type of farms	number of farms in total	number of farms with control(s)
2010	RF	2	2
	SF	18	13
2011	RF	10	10
	SF	8	7
2012	RF	9	9
	SF	13	7

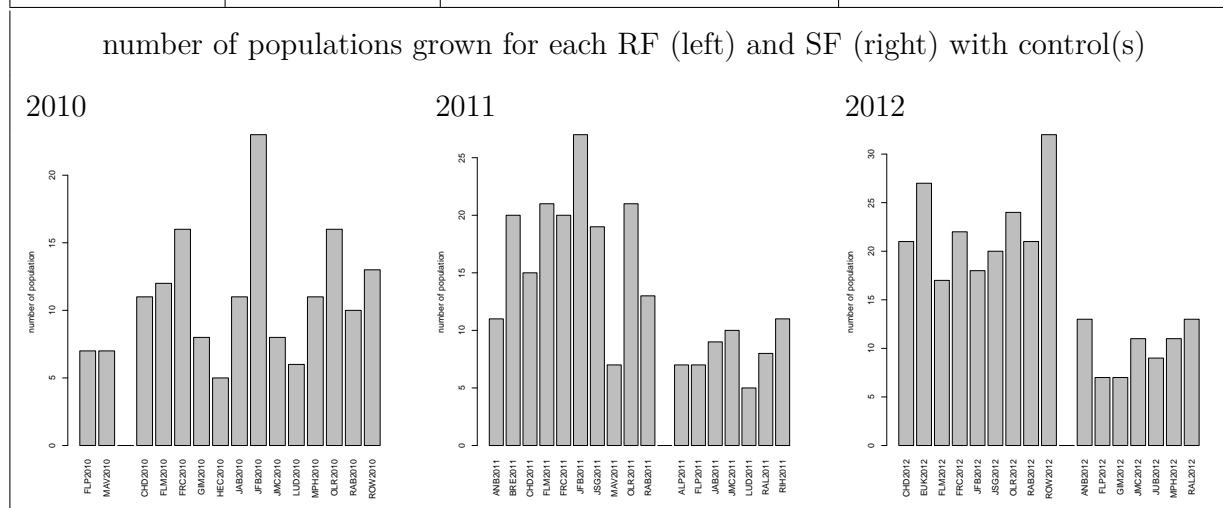


Table 1 – Number of regional farms (RF) and satellite farms (SF) during seasons 2009-2010, 2010-2011, 2011-2012 that sent back measures and spikes for analysis, number of farms that followed the experimental design with the controls and that were used in the analysis and number of populations grown for each farm.

### 3 2.1.2 Data collected

4 Two types of variables were measured : individual ones (25 measures on 25 plants  
5 within each population) and global ones (an average measure from a bulk of 25 plants  
6 in each plot) (Table 2). The variables measured were chosen after discussion between  
7 farmers and researchers. Some variables were based on phenotypic descriptors used for

- 1 variety registration. Fewer variables were recorded on satellite farms. Observations were  
 2 measured either on plants on farm or on spikes at the research station (Table 2).

variable	RF	SF	env	method
color <sup>i,1</sup>	X	X	33	color of the spikes from 0 (white) to 2 (red). Possible values were 0, 0.5, 1, 1.5 and 2.
awns <sup>i,1</sup>	X	X	34	presence of awns on the spikes from 0 (no awns) to 2 (full of awns). Possible values were 0, 0.5, 1, 1.5 and 2.
curve <sup>i,1</sup>	X	X	33	curvature of the spike at maturity from 0 (no curve) to 2 (spike up side down). Possible values were 0, 0.5, 1, 1.5 and 2.
plant height <sup>g,i,1</sup>	X		36	height of the plant from the ground to the top of the spike in mm.
last leaf spike distance (LLSD) <sup>i,1</sup>	X		31	distance from the flag leaf to the bottom of the spike in mm.
spike length <sup>i,1</sup>	X		28	length of the spike in mm.
nbr spikelets <sup>i,2</sup>	X		21	number of spikelets.
sterility <sup>i,2</sup>	X		35	sum of sterile spikelets per spike.
spike weight <sup>i,2</sup>	X	X	44	weight of the spike in g.
protein <sup>g,2</sup>	X	X	25	% of protein content in the grain measured with NIRS technology at INRA Clermont Ferrand France on grain using near infrared spectroscopy (FOSS NIRSystem 6500).
tkw <sup>g,2</sup>	X	X	45	thousand kernels weight in g.

Table 2 – Variables measured and method associated. RF : regional farms ; SF : satellite farms ; env : number of environments for a given variable ; <sup>g</sup> : global measures ; <sup>i</sup> : individual measures ; <sup>1</sup> : measure on farm ; <sup>2</sup> : measure at the research station on samples of spikes sent by farmers.

## 2.2 Hierarchical Bayesian model

All the analyses were carried out using the software R (?) and JAGS (Plummer, 2003). R scripts are available under request.

Because of the low number of residual degrees of freedom for each farm, we used a hierarchical Bayesian model in order to assess mean differences on farm. This model allowed for a residual variance for each environment. We took advantage of the parallel (i.e. similar) structure of the trials on each environment of the network to assume that trial residual variances came from an inverse gamma distribution. Such a distribution has a support bounded by 0 (consistent with variance definition) and may have various shapes including asymmetric distributions. From an agronomical point of view, the assumption that trial variances were heterogeneous was consistent with organic farming : there were as many environments as farmers leading to a high heterogeneity. Environment was here considered in a broad sense : practices (sowing date, sowing density, tilling, etc.), pedo climatic conditions, biotic and abiotic stress, ... (Desclaux et al., 2008). Moreover, the inverse gamma distribution had conjugate properties that facilitated MCMC convergence (see section 2.3). This model was therefore a good compromise between agronomy and statistics. The residual variance of a trial was estimated using all the information available rather than using the data from that particular trial only.

### 2.2.1 Likelihood

We restricted ourselves to analysing plot means. The phenotypic value  $Y_{ijk}$  for variable  $Y$ , population  $i$ , environment  $j$  and block  $k$  was modelled as :

$$Y_{ijk} = \mu_{ij} + \beta_{jk} + \varepsilon_{ijk}; \quad \varepsilon_{ijk} \sim \mathcal{N}(0, \sigma_j^2), \quad (1)$$

where  $\mu_{ij}$  was the mean of population  $i$  in environment  $j$ ;  $\beta_{jk}$  was the effect of block  $k$  in environment  $j$  satisfying the constraint  $\beta_{j2} = -\beta_{j1}$ ;  $\varepsilon_{ijk}$  was the residual error;  $\mathcal{N}(0, \sigma_j^2)$  denoted normal distribution centred on 0 with variance  $\sigma_j^2$ .

The residual variance estimated from the controls was assumed to be representative of the residual variance of the other populations. Blocks were included in the model

1 only if the trial had blocks.

2 Standardised residuals were plotted to check model assumptions.

### 3 2.2.2 Prior distributions

4 We used a hierarchical prior for the residual variances. First these parameters were  
5 assumed to follow an inverse gamma distribution :

$$\tau_j \sim \text{Gamma}(\nu, \rho),$$

6 where  $\tau_j$  was the precision for environment  $j$  defined by  $\tau_j = \sigma_j^{-2}$ . Then, we placed  
7 vague prior distributions on the hyperparameters  $\nu$  and  $\rho$  :

$$\nu \sim \text{Uniform}(\nu_{min}, \nu_{max}); \quad \rho \sim \text{Gamma}(10^{-6}, 10^{-6}).$$

8 We chose  $\nu_{min} = 2$ . Then the distribution of  $\sigma_j^2$  given  $\nu$  and  $\rho$  had a finite mean  
9 and a finite variance (Robert, 2001). We chose  $\nu_{max} = 10$ . The distribution used for  $\rho$   
10 was vague and allowed for a large range of possibilities.

11 The parameters  $\mu_{ij}$  and  $\beta_{j1}$  were assumed to follow vague prior distributions :

$$\mu_{ij} \sim \mathcal{N}(\mu_{.j}, 10^6); \quad \beta_{j1} \sim \mathcal{N}(0, 10^6).$$

12 As the prior distribution of  $\mu_{ij}$  was vague, the prior mean of  $\mu_{ij}$  had little effect on  
13 the posterior distribution of  $\mu_{ij}$  and was thus unimportant (Carlin and Louis, 2008). It  
14 was fixed to the empirical trial mean  $\mu_{.j}$ .

## 15 2.3 Posterior distributions

16 Bayesian inference is based on the posterior distribution of model parameters. This  
17 distribution could not be calculated explicitly for the hierarchical model presented  
18 in the previous section but could be estimated using Markov Chain and Monte Carlo  
19 (MCMC) methods. These methods simulate values of model parameters according to a  
20 Markov chain that converges to the posterior distribution of model parameters (Robert,  
21 2001).

1 MCMC methods were implemented using the package `rjags` of R that performed  
 2 Gibbs sampling (Robert, 2001). Two MCMC chains were run independently to test for  
 3 convergence using the Gelman-Rubin test. This test was based on the variance within  
 4 and between the chains (Gelman and Rubin, 1992). A burn-in and lots of iterations  
 5 were needed in the MCMC procedure. In our case, the burn-in had 1000 iterations, then  
 6 100 000 iterations were done with a thinning interval of 10 to reduce autocorrelations  
 7 between samples, so that 10 000 samples were available for inference for each chain. For  
 8 all the parameters convergence diagnostics did not detect any lack of convergence.

## 9 2.4 Mean comparisons

10 The mean of each population was compared to the mean of each other population.  
 11 Let  $H_0$  and  $H_1$  denote the hypotheses :

$$H_0 : \mu_{ij} \geq \mu_{i'j}, \quad H_1 : \mu_{ij} < \mu_{i'j}.$$

12 The difference  $\mu_{ij} - \mu_{i'j}$  between the means of population  $i$  and population  $i'$  in  
 13 environment  $j$  was considered as significant if either  $H_0$  or  $H_1$  had a high posterior  
 14 probability, that is if  $Pr\{H_0|y\} > 1 - \alpha$  or  $Pr\{H_1|y\} > 1 - \alpha$ , where  $\alpha$  was some specified  
 15 threshold. The difference was considered as not significant otherwise. The posterior  
 16 probability of a hypothesis was estimated by the proportion of MCMC simulations for  
 17 which this hypothesis was satisfied. The threshold  $\alpha$  was fixed to  $\alpha = 0.1/I$ , where  $I$   
 18 was the number of populations in a given environment. For an environment with 20  
 19 populations,  $\alpha = 0.005$ . For an environment with 10 populations,  $\alpha = 0.01$ . Value of  
 20  $\alpha$  was used to take account of multiple testing. It was used instead of a Bonferroni  
 21 correction as the latter was usually considered to be too strict. Westfall (1997) used a  
 22 Bonferroni correction directly on the prior. Our approach allows to change the threshold  
 23 without running the model again.

24 Populations that were not significantly different were allocated to the same group.  
 25 In order to study the allocation of populations to groups, we introduced two ratios :

- 26 – ratio 1 : number of groups / number of populations.

1            If number of groups = number of populations, then this ratio equals 1 and there  
 2            were as many groups as populations : all the populations were in different groups.  
 3            If the number of groups is 1, then this ratio is 1 / number of populations, and all  
 4            the populations were in the same group.

5            – ratio 2 : mean number of populations per group / number of populations.

6            If mean number of populations per group = number of populations, then this  
 7            ratio equals 1 and all the populations were in the same group. If mean number of  
 8            populations per group is 1, then this ratio equals 1 / number of populations, and  
 9            all the populations were in different groups.

10           These ratios provided information on the discrimination of populations.

11           We also compared block means when the trials had blocks. We used the same method  
 12           as for population comparisons, but with  $H_0 : \beta_{j1} \geq \beta_{j2}$  and  $\alpha = 0.05$ .

## 13           2.5    Comparison of hierarchical and non-hierarchical mo- 14           dels

15           In order to study the benefit of model 1, a similar analysis was run with a non  
 16           hierarchical model (model 2). In this model, the prior distribution of precision  $\tau_j$  was  
 17           vague :

$$\tau_j \sim \text{Gamma}(10^{-6}, 10^{-6}).$$

18           The analysis with this model was close to a frequentist analysis of variance (Box  
 19           and Tiao, 1973).

20           We compared the values of  $\sigma_j$  and the numbers of groups of both models. To do  
 21           so, we applied the same method as for population comparisons with  $H_0 : \sigma_{j2} \geq \sigma_{j1}$   
 22           where  $\sigma_{j2}$  was  $\sigma_j$  for model 2 and  $\sigma_{j1}$  was  $\sigma_j$  for model 1.

## 23           2.6    Between-year variation of residual variances

24           We ran model 1 on four data sets : 2010, 2011, 2012 and 2010-2011-2012 (i.e. all the  
 25           data). We studied the evolution of the estimates of  $\nu$ ,  $\rho$  and  $\sigma_j$  and the evolution of



1 the common distribution of the  $\sigma_j$ . We compared the values of  $\nu$  and  $\rho$  using the same  
 2 method as for population comparisons.

### 3 **2.7 Influence of the number of environments**

4 The data must include several environments to estimate the hyperparameters  $\nu$  and  
 5  $\rho$  of model 1. Thus we studied how many environments were required in order to have  
 6 a reliable analysis. We considered two variables, tkw and spike length.

7 For a given number of environments, it was not possible to consider all possible  
 8 environment combinations (i.e.  $\binom{1}{n} + \binom{2}{n} + \dots + \binom{n-1}{n} + \binom{n}{n}$  combinations!). There-  
 9 fore we generated combinations randomly for certain numbers of environments. More  
 10 combinations were generated for the lowest numbers of environments in order to be  
 11 representative of the variation (Table 3). The number of environments sampled was  
 12 larger than or equal to five so that enough information was available to run the model.

tkw		spike length	
Number of environments	Number of samples	Number of environments	Number of samples
5	20000	5	20000
10	12000	10	12000
20	8000	15	8000
30	4000	20	4000
40	2000	25	2000
45	1	28	1

Table 3 – Number of samples used for each number of environments.

13 For a given number of environments, we considered the distributions of the posterior  
 14 means of  $\nu$  and  $\rho$  for each sample. Then, we plotted the common distributions of residual  
 15 standard deviations obtained with values of  $\nu$  and  $\rho$  corresponding to different quartiles  
 16 of their respective distribution : minimum, 25%, median, 75% and maximum. Outliers  
 17 were not used as minimum and maximum values.

## 1      3      Results

### 2      3.1    Model assessment

3            For all the variables, the estimates of residual variances were consistent with the  
 4            assumption that they followed an inverse gamma distribution. For example, for plant  
 5            height, the majority of  $\sigma_j$  was between 50 and 80 mm, three were between 80 mm and  
 6            90 mm and two were between 130 mm and 150 mm (Figure 4). A similar pattern was  
 7            found for tkw :  $\sigma_j$  was between 2g and 4g with two extreme environments at 4.8g and  
 8            6.3g. For all variables, standardised residuals were between -2 and 2, suggesting that  
 9            the model fitted the data rather well (Figure 5).

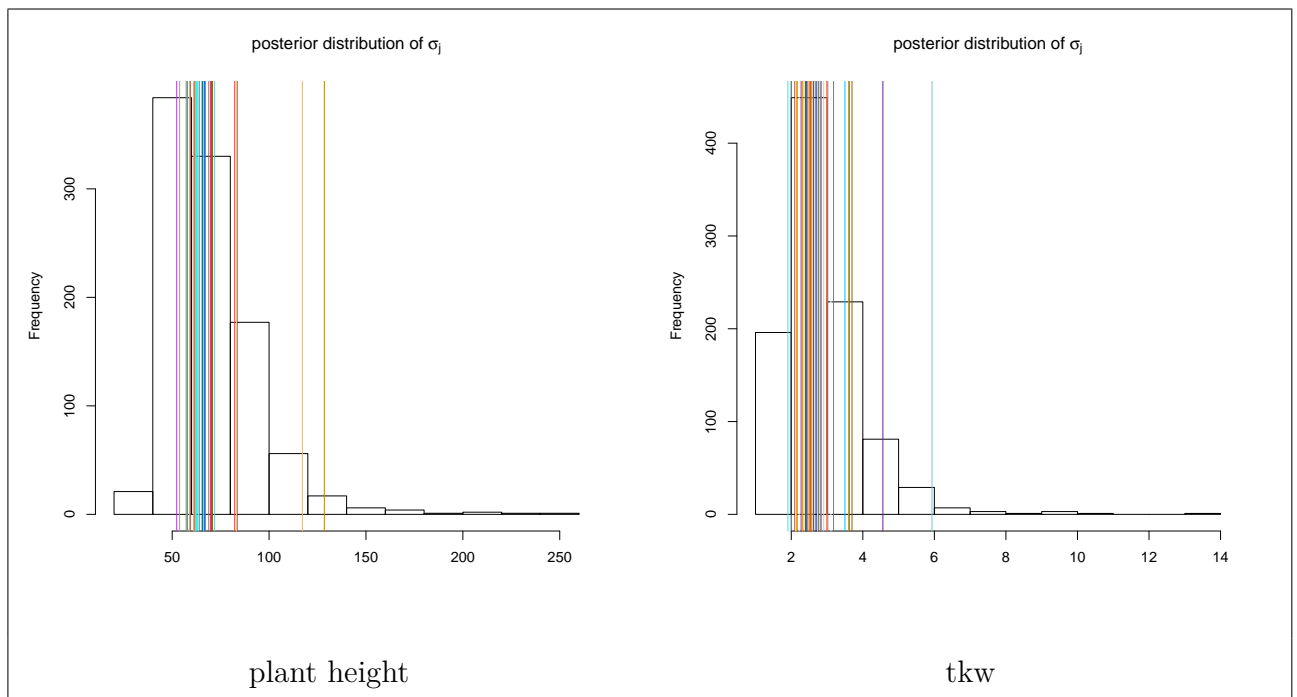


Figure 4 – Histogram of the common distribution of the  $\sigma_j$  based on the estimated values of  $\nu$  and  $\rho$ . Vertical bars are the observed values for each environment  $j$ .

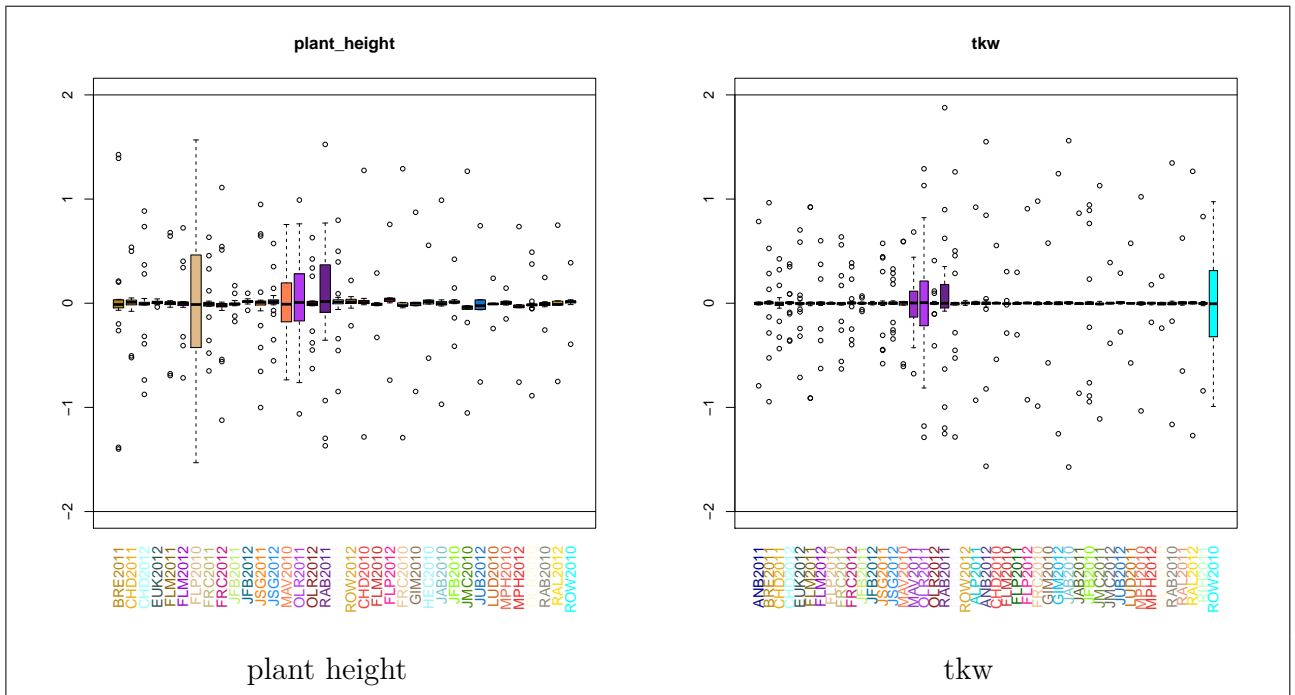


Figure 5 – Boxplots of standardised residuals for each environment for plant height and tkw.

## 1      3.2 Mean comparisons

### 2      3.2.1 Block effects

3            Block effect were significant in 11% to 16% of environments for plant height, LLSD  
 4            and spike length. For the other variables, this proportion dropped between 0% and  
 5            10%. The estimate of  $\sigma_j$  tended to be higher when block effects were significant.

### 6      3.2.2 Genotype effects

7            The confidence intervals of  $\sigma_j$  were rather large for all the environments (Figure 6).  
 8            Although regional farms were supposed to better control environmental variation, there  
 9            was no clear link between the type of farms (regional or satellite) and the number of  
 10            groups or the estimation of  $\sigma_j$ . This could be explained by the fact that many factors  
 11            varied between farms, in particular farms had different populations.

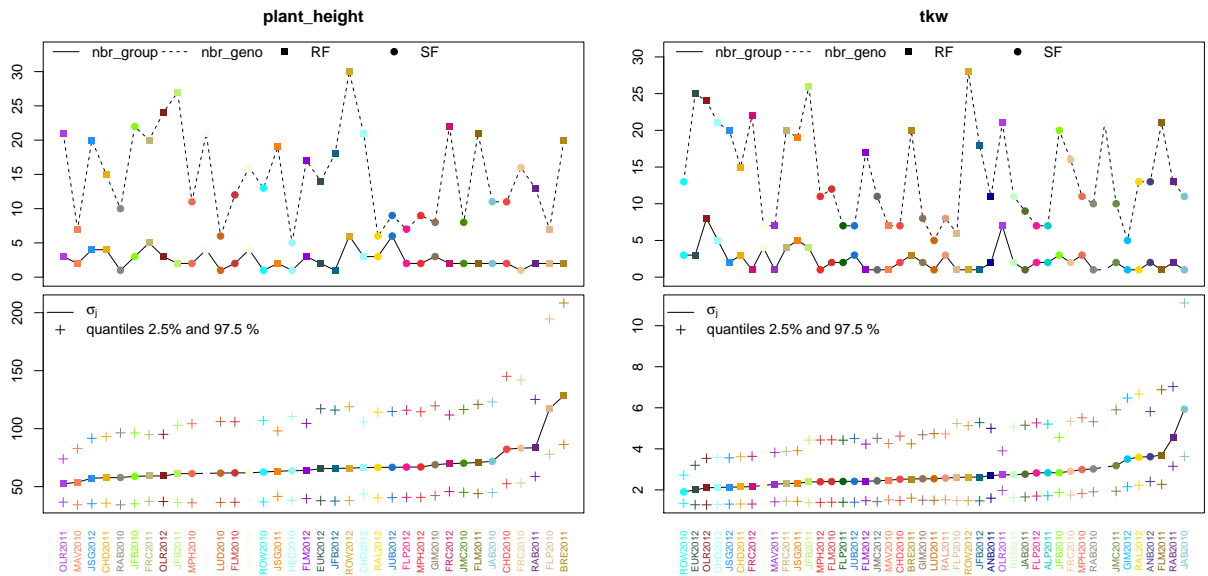


Figure 6 – Number of populations, number of groups detected and estimated  $\sigma_j$  (with confidence interval) for each environment for plant height (left) and tkw (right). RF : regional farms ; SF : satellite farms.

1 Table 4 summarises the information for each variable for ratio 1 (number of groups  
 2 / number of populations) and ratio 2 (mean number of populations per group / number  
 3 of populations). There was a negative correlation between the value of  $\sigma_j$  and ratio 1  
 4 for all the variables. This was expected as increasing  $\sigma_j$  reduced the precision of the  
 5 estimate of  $\mu_{ij}$  and then the number of groups.

6 In 32.2% of environments, there was only one group (ratio 1 in Table 4). However the  
 7 proportion of environments with a single group varied between variables. There were no  
 8 environments with one group for LLSD while there were many environments with one  
 9 group for spike weight. In 67.8% of the cases, some populations could be distinguished.

10 There was never one population in each group overall the environments and the  
 11 variables (ratio 2 in Table 4). In 55.7% of environments, ratio 2 was between  $\frac{2}{3}$  and 1  
 12 meaning that the groups overlapped each other.

	ratio 1 ( $r1$ )			ratio 2 ( $r2$ )			
	$r1 = \frac{1}{np}$	$\frac{1}{3} < r1 \leq \frac{1}{3}$	$\frac{1}{3} < r1 \leq \frac{2}{3}$	$\frac{2}{3} < r1 \leq 1$	$r2 = \frac{1}{np}$	$\frac{1}{3} < r2 \leq \frac{1}{3}$	$\frac{2}{3} < r2 \leq 1$
color	3	17	12	1	0	1	27
awns	0	6	28	0	0	12	21
curve	15	17	1	0	0	0	7
nbr sterile spikelets	21	14	0	0	0	0	2
plant height	6	28	2	0	0	2	17
LSD	0	19	12	0	0	2	25
spike length	10	16	2	0	0	0	9
nbr spikelets	5	12	5	0	0	2	8
spike weight	29	13	2	0	0	0	5
tkw	19	23	3	0	0	0	15
protein	10	13	2	0	0	0	7
% of the total	32.2	48.6	18.9	0.3	0	5.2	39.1
							55.7

Table 4 – Distribution (number of environments) of ratio 1 and ratio 2. ratio 1 = number of groups / number of populations ( $np$ ) and ratio 2 = mean number of populations per group / number of populations. ratio 1 =  $\frac{1}{np}$  means that all the populations were in the same group, ratio 1 = 1 means that all the populations were in different groups. ratio 2 =  $\frac{1}{pg}$  means that all the populations were in different groups, ratio 2 = 1 means that all the populations were in the same group.

### 1 3.3 Comparison of hierarchical and non-hierarchical models 2

3 Residual standard deviations  $\sigma_j$  were less variable between environments for the  
4 hierarchical model than for the non hierarchical model (Figure 7). This occurred for  
5 all the variables. The information given by the hierarchical model shrank the residual  
6 variances. This was consistent with the theory (Robert, 2001; Smyth, 2004).

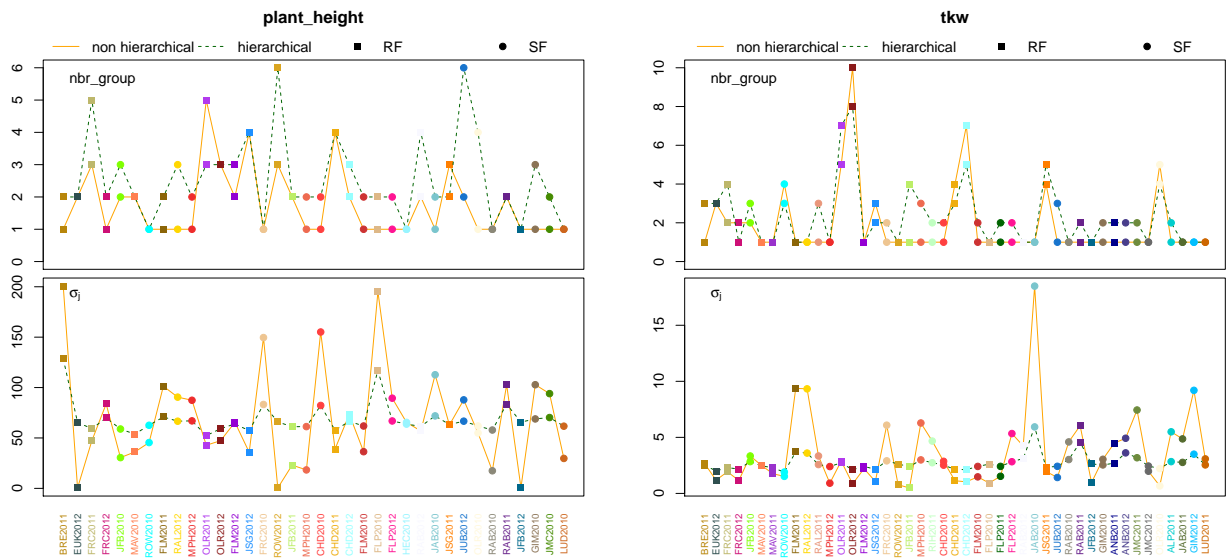


Figure 7 – Comparison of residual standard deviation ( $\sigma_j$ ) and number of groups for plant height and tkw between non hierarchical model 2 and hierarchical model 1. RF : regional farms; SF : satellite farms.

difference in the number of groups	no	↑	↓
color	8	20	5
awns	7	17	10
curve	18	13	2
nbr sterile spikelets	19	12	4
plant height	13	21	2
LLSD	3	26	2
spike length	15	12	1
nbr spikelets	8	10	4
spike weight	28	12	4
tkw	17	21	7
protein	10	14	1
% of the total	36.1	52.6	11.3

Table 5 – Occurrence of changes in number of groups for each variable. ↑ : number of groups is greater for model 1 compared to model 2, ↓ : number of groups is smaller for model 1 compared to model 2.

1            There were more groups with the hierarchical model (52.6%) compared to the non  
2            hierarchical model (11.3%) (Figure 7 and Table 5). In none cases did the  $\sigma_j$  show a  
3            significant change in model 2 compared to model 1.

#### 4            **3.4 Between-year variation of residual variances**

5            The effects of years on  $\nu$  and  $\rho$  were not significant. However the size of year effects  
6            varied between variables. For example, year effects on  $\nu$  and  $\rho$  were larger for spike  
7            length than for tkw (Figure 8). Likewise, the effects of years on the estimates of residual  
8            standard deviations were larger for spike length than for tkw (Figure 9). For each  
9            variable, there was at least a one-year data set that led to a rather flat posterior  
10            distribution for  $\nu$ .

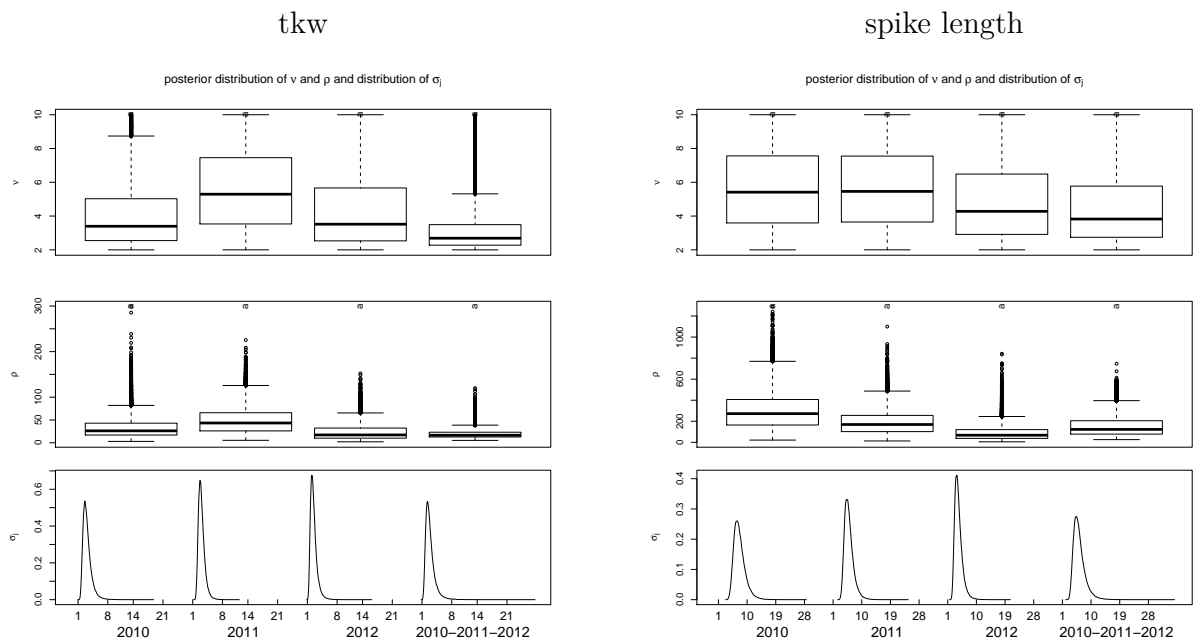


Figure 8 – Posterior distribution of  $\nu$  and  $\rho$  and common distribution of residual standard deviations ( $\sigma_j$ ) based on the mean value of  $\nu$  and  $\rho$  for tkw (left) and spike length (right) for 2010, 2011, 2012, and 2010-2011-2012 (i.e. all the data).



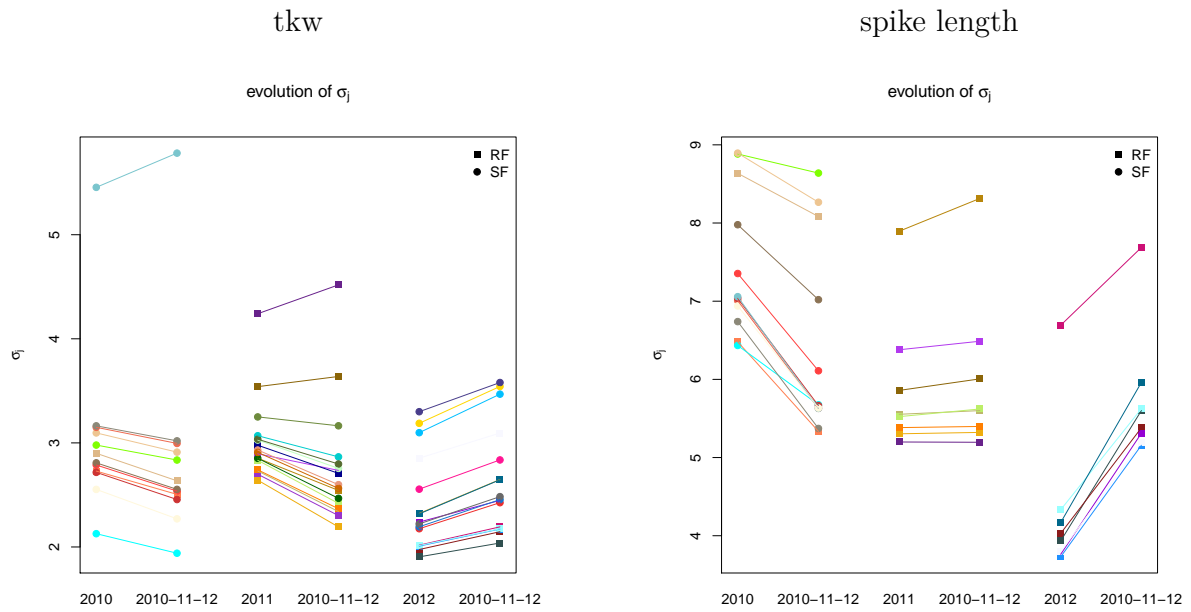


Figure 9 – Estimates of residual standard deviations ( $\sigma_j$ ) for tkw (left) and spike length (right) for 2010, 2011, 2012, and 2010-2011-2012 (i.e. all the data).

### 1      3.5 Influence of the number of environments

2            For both variables tkw and spike length, the variance and the mean of the estimates  
 3 of  $\nu$  and  $\rho$  decreased with the number of environments. The shape of the distributions  
 4 of the common distribution of residual standard deviations became increasingly precise  
 5 as the variance of  $\nu$  and  $\rho$  decreased (Figure 10 & 11).

6            The larger the number of environments, the smaller the variability in the common  
 7 distribution of  $\sigma_j$ .

8            For both variables, the median of the estimates of  $\nu$  with 5 environments was very  
 9 close to the prior median ( $(\nu_{max} + \nu_{min})/2 = 6$ ), and the estimates of  $\rho$  had extreme  
 10 values with 5 environments.

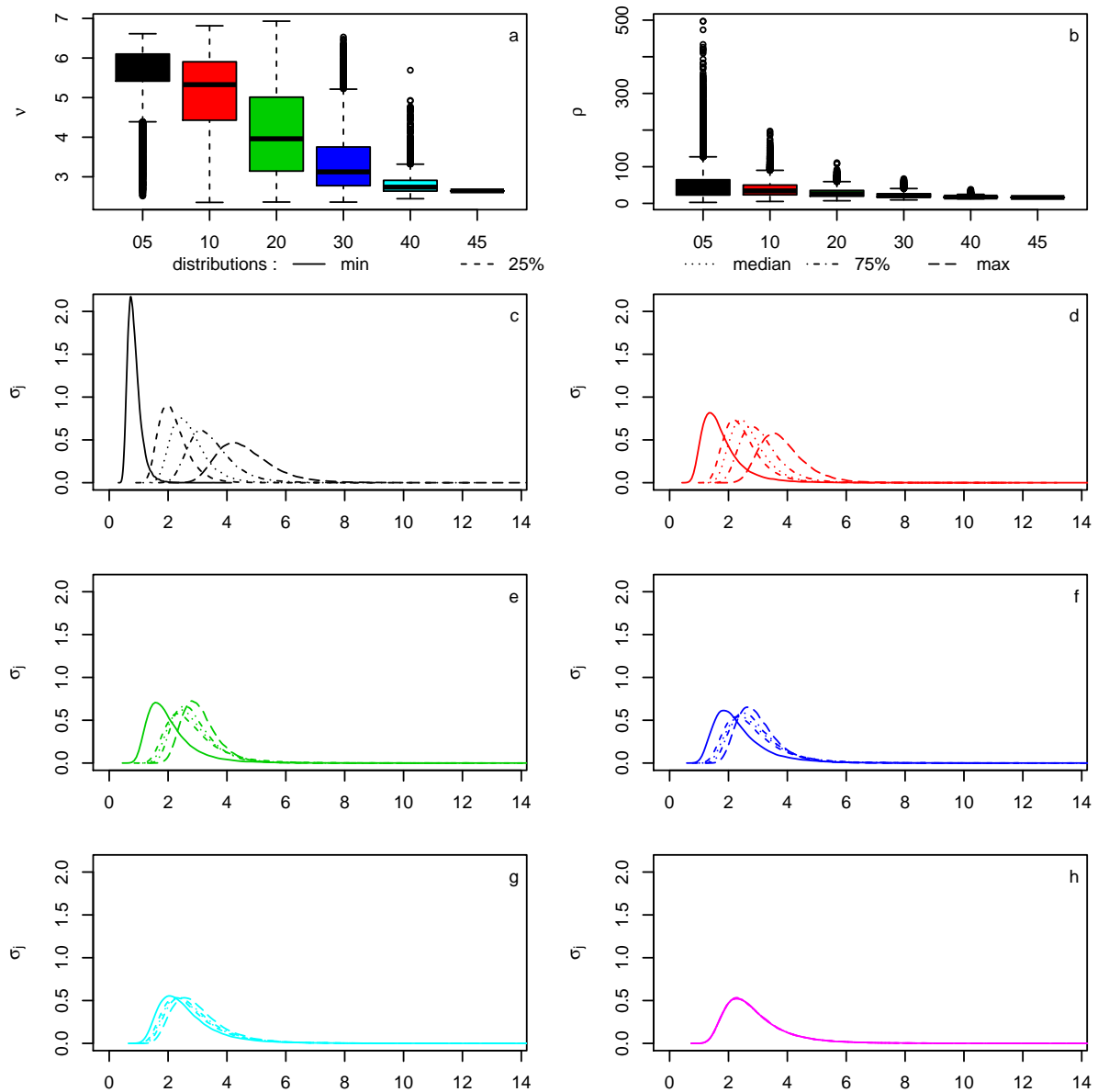


Figure 10 – For tkw, evolution of the estimation of  $\nu$  (a) and  $\rho$  (b) with the number of samples taken for the analysis (5, 10, 20, 30, 40 and 45) and respective common distribution of residual standard deviation  $\sigma_j$  for the different quartiles of  $\nu$  and  $\rho$  (c, d, e, f, g, h, i).

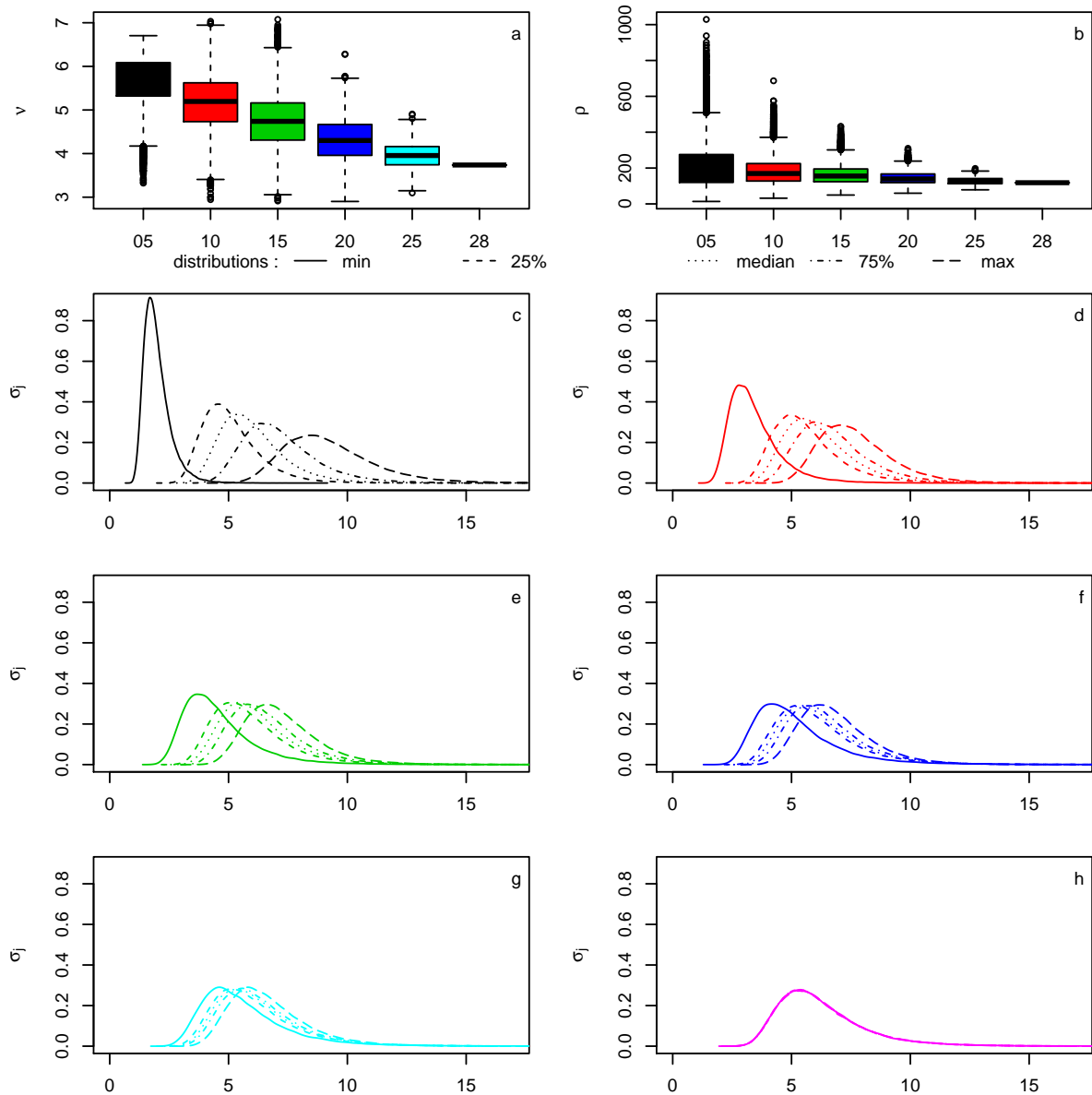


Figure 11 – For spike length, evolution of the estimation of  $\nu$  (a) and  $\rho$  (b) with the number of samples taken for the analysis (5, 10, 20, 30, 40 and 45) and respective common distribution of residual standard deviation  $\sigma_j$  for the different quartiles of  $\nu$  and  $\rho$  (c, d, e, f, g, h, i).

## 1      4    Discussion

### 2      4.1   Hierarchical model

3            The hierarchical model we used was inspired from statistical analyses of micro-array  
4            data (Baldi and Long, 2001). Micro-array data have mainly been analysed using em-  
5            pirical Bayes methods (Smyth, 2004). These methods are computationally efficient but  
6            neglect the uncertainty on hyperparameters. They are appropriate for microarray expe-  
7            riments as the latter involve very large numbers of genes (Lönnstedt and Britton, 2005;  
8            Nott et al., 2007). Yet in the wheat PPB experimental design, the number of environ-  
9            ments was not very large and it was more rigorous to take account of the uncertainty  
10           on the hyperparameters  $\nu$  and  $\rho$ . This is why prior distributions were developed for  
11           these parameters and a fully Bayesian approach was implemented.

12           The estimates of residual variances were consistent with the assumption that re-  
13           sidual variances came from an inverse gamma distribution. This shows the variability  
14           that exists in organic farms. Other results confirm that variability on farm can be  
15           heterogeneous (Hu et al., 2013).

16           Some environments showed extreme residual variances (Figure 4 and 6). For example,  
17           there was one extreme environment for tkw (JAB2010) and there were four extreme  
18           environments for plant height (CHD2010, FRC2010, FLP2010, BRE2011). These envi-  
19           ronments were subject to a high heterogeneity. For example the strong heterogeneity of  
20           trial BRE2011 could be explained by mineralisation differences in the soil that were not  
21           controlled by the blocks of the trial. This may have happened in other trials. A correct  
22           orientation of blocks is needed for efficient statistical analysis with or without spatial  
23           adjustment (Leiser et al., 2012). In the following of the PPB programme, the principles  
24           to settle experimental designs should be further explained and discussed with farmers.

25           The developed model could be extended in various ways. First it could be interesting  
26           to take account of within-plot measurements for individual variables. As we are dealing  
27           with heterogeneous populations, the variation within a plot may be high. We tried such  
28           a model but the analysis was not very satisfactory as the posterior distribution of  $\nu$   
29           was rather flat for most variables. Moreover, it took 1.5 more computational times.

1 However taking individual data into account would allow to estimate a between-plot  
2 and a within plot-variance.

3 Then it would be interesting to shrink and thus stabilise estimates of genotype  
4 effects, in particular for unreplicated genotypes This may also improve the estimates  
5 of residual variances on farm. Finally it would be interesting to develop non-gaussian  
6 models for discrete variables such as color or awns.

## 7 4.2 Population comparisons

8 As decentralised selection is carried out in the target environments, the effect of a  
9 genotype in an environment includes the genetic part of the  $G \times E$  interaction. The  
10 genotype effect is therefore a perfect estimation of the potential of the genotype in a  
11 given farm, which is the main objective of PPB (Ceccarelli and Grando, 2007).  $G \times E$   
12 interactions are a key point that justified selection in contrasting environments (Dawson  
13 et al., 2008). Moreover, heritability within each farm is expected to be higher compared  
14 to heritability at the network level, since the  $G \times E$  interaction variance should increase  
15 the denominator in the latter case.

16 Selection for a good average performance across environments is not necessarily de-  
17 sired. Farmers are not looking for broad adaptation but for on-farm specific adaptation.  
18 Such an approach is put forward by authors like Gauch et al. (2008) that study GGE  
19 biplots (mixing  $G$  and  $G \times E$  effects) to find the "which won where" genotypes.

20 We ran the model on the mean of the populations, but the within-population va-  
21 riability would allow for complementary mass selection (Rivière et al., 2013). When  
22 mass selection within populations is applied, the derived progeny can be evaluated the  
23 following year using the same approach.

24 Comparing populations to the control also provides information on stability over  
25 years. This information complements the farmers expertise acquired through cultiva-  
26 ting and taking notes on the development of the crop. Correlation between variables  
27 (between protein and tkw for example) are also of interest to some farmers.

28  
29 Some authors used a mixed classical/Bayesian approach to perform mean compa-

1        risons. They estimated variance with a Bayesian model and then use the estimated  
2        variance in a classical t-test (Baldi and Long, 2001; Smyth, 2004).

3            An alternative way is to conduct fully Bayesian comparisons (Nashimoto and Wright,  
4        2008) as we did. This approach is more rigorous as it takes account of the uncertainty  
5        of all the parameters or unequal sample size.

6            The final objective of the model is to identify groups of populations that are signifi-  
7        cantly different. This assignment is based on the threshold  $\alpha$  that depends on agronomic  
8        objectives. This threshold was set to  $\alpha = 0.1/I$  (with  $I$  the number of populations on  
9        the farm). It corresponded to a "soft" Bonferroni correction, the Bonferroni correction  
10       being very conservative. The number of groups created differed between farms. When  
11       there were few groups, the populations were similar or could not be distinguished, and  
12       therefore the repeatability was low. Conversely, when there were many groups, popu-  
13       lations could be distinguished and the repeatability was higher. A high repeatability  
14       is an important issue in breeding as it allows for selection among populations, which is  
15       the main step of selection in the breeding process (Ceccarelli and Grando, 2007).

16            Over all the environments and the variables, some populations could be distingui-  
17        shed from other populations in 67.8% of the cases. When no differences among popu-  
18        lations means were detected, selection was not possible. As one objective of this PPB  
19        programme is that farmers (re)learn selection, the threshold could be adjusted to allow  
20        the detection of at least two groups instead of having farmers choose at random. The  
21        initial value could be set to  $\alpha = 0.1/I$  (with  $I$  the number of populations on the farm)  
22        and if only one group is obtained, then this value could be adjusted to allow the detec-  
23        tion of two groups. In this cases, the farmers should be informed of the lower degree of  
24        confidence that there are significant differences among populations.

### 25        **4.3 Comparison of hierarchical and non-hierarchical mo-** 26        **dels**

27            In order to study the merit of the hierarchical model, the analysis based on this  
28        model was compared to a more standard analysis that was based on a non-hierarchical  
29        model and that was close to a split-plot analysis.

1           In this analysis, the estimated residual variance of an environment did not take  
2 account of the data of the other environments and so was very dependent on the ob-  
3 servations collected on the controls in this environment. For example, for plant height,  
4 the estimated residual standard deviation for trial FLP2010 was very high (more than  
5 200 mm). This trial had *Rouge-du-Roc* repeated twice. The difference between both  
6 repetitions was 400mm, which is very large. With the hierarchical model, the estimated  
7 residual standard deviation was still important (2<sup>nd</sup> higher) but had a more reasonable  
8 value (around 130 mm). The hierarchical model shrank variance estimates and limited  
9 the occurrence of false positives resulting from a poor estimation of residual variances.  
10 The hierarchical model seems therefore more adapted to our data.

11           In 52.6% of the cases, the hierarchical model 1 increased the number of groups, thus  
12 allowing the detection of more differences among populations than the non hierarchical  
13 model 2. Only in 11.3% of the cases was the hierarchical model 1 less powerful to  
14 detect population differences and decreased the number of groups. Thus, the hierarchical  
15 modelling of the variance has a strong impact in the creation of groups and therefore in  
16 the information given to farmers to carry out selection among populations. Overall, the  
17 hierarchical modelling increased or maintained the differences between populations.

#### 18    **4.4 Between-year variation of residual variances**

19           Residual variances did not show a large variation between years. However this va-  
20 riation varied between variables and was low in particular for the variables awns and  
21 color. This is coherent as those traits were mainly affected by genetic variations and  
22 were therefore little affected by environmental factors.

23           A different gamma distribution for a given year may be due to a low number of envi-  
24 ronments for that year. Indeed, we observed that the variables with fewer environments  
25 were more affected by years. The influence of years may therefore be due to statistical  
26 rather than agronomic reasons.

27           A hierarchical model with a separate residual-variance distribution for each year  
28 was not appropriate for the wheat PPB programme. The number of trials per year was  
29 below 20 and was too low (see Section 4.5). Moreover, it is not appropriate to carry

1 out selection only by having information on one year. It is better to have an average  
2 estimation of the value on several years (mixing within and between year variation).  
3 The farmers are aware of this and always want to evaluate a germplasm at least two  
4 or three years before applying selection. For each variable, there was at least a year for  
5 which the posterior distribution of  $\nu$  was rather flat or for which the MCMC algorithm  
6 did not behave well. In such cases, the analysis was not reliable.

## 7 4.5 Influence of the number of environments

8 Estimation of residual variances using the hierarchical model depended on the num-  
9 ber of environments used in the analysis. The mean and the dispersion of the estimates  
10 of hyperparameters  $\nu$  and  $\rho$  both decreased with the number of environments used  
11 (Figures 10 and 11).

12 When we took five environments, the estimates of  $\nu$  and  $\rho$  had a large variance and  
13 the prior influenced the posterior. The median of the estimates of  $\nu$  was really close to  
14 the prior median. As the number of environments increased, the prior was less influential  
15 leading to better estimates of hyperparameters. As the common distribution of variances  
16 played the role of a prior in the analysis, its shape influenced the estimates of these  
17 variances. On the whole, at least 20 environments seemed required to implement the  
18 hierarchical model with confidence in the wheat PPB programme. Consistently, spike  
19 length observed in fewer environments showed a higher variance in the inverse gamma  
20 distribution than *tkw*.

## 21 4.6 Within-trial experimental design

22 Our experimental design is reminiscent of *mother baby trials*. Such trials involve a  
23 *mother trial* that is replicated and several *baby trials* that are not replicated (Bellon  
24 et al., 2002; Witcombe et al., 2005a). Those trials are very flexible for the farmers (for  
25 details and example see (Witcombe et al., 2005a; Bellon et al., 2002; Gyawali et al.,  
26 2007; Ghaouti et al., 2008; Vom Brocke et al., 2010)).

27 For some *mother baby trials*, the complete statistical analysis is performed on the  
28 *mother trial* where replications are done while the *baby trial* often leads to only descrip-



1 tive or qualitative analyses (Bellon et al., 2002; Gyawali et al., 2007; Virk et al., 2005;  
2 Witcombe et al., 2005b). The analyses of *baby trials* can be completed by interviews of  
3 farmers(Witcombe et al., 2005a).

4 Some PPB projects are divided into several stages leading to specific experimental  
5 designs. This was used by ICARDA (Ceccarelli et al., 2001; Ceccarelli and Grando, 2007;  
6 Ceccarelli, 2012). Stage 1 involves many genotypes. The design is based on unreplicated  
7 trials with systematic checks replicated in the trials to quantify environmental variation.  
8 For example a typical trial can involve between 11 and 17 controls for between 134 and  
9 160 plots over 9 locations (Ceccarelli et al., 2003). Stages 2 and 3 have fewer entries.  
10 The design is based on incomplete blocks with two replicates divided into rows and  
11 columns. One advantage of incomplete block designs (IBD) is that each replicate is an  
12 independent unit that can be allocated to a different field (Ceccarelli, 2012). Stage 4  
13 has 2 or 4 entries, it is the last stage. The trials can be designed as alpha-lattices with  
14 two replications or as randomized complete blocks with farmers as replicates, or as  
15 standard replicated trials (Ceccarelli and Grando, 2007). Such designs allow statistical  
16 analyses (Singh et al., 2003; Ceccarelli et al., 2000b; Fufa et al., 2010).

17 In the wheat PPB programme, we can make a parallel between on the one hand  
18 regional farms and *mother trials* and on the other hand satellite farms and *baby trials*  
19 (Dawson et al., 2011). However these trials show many differences. Compared to *mother*  
20 *trials*, regional farms in the wheat PPB programme have the same plot size (Gyawali  
21 et al., 2007; Ceccarelli et al., 2000b) but they are less balanced than *mother trials*.  
22 Satellite farms have small plots (between 5 and 10 m<sup>2</sup>, as for regional farms) while *baby*  
23 *trials* usually have larger plots between 50 and 1000 m<sup>2</sup> (Gyawali et al., 2007; Cecca-  
24 relli and Grando, 2007). Moreover, satellite farms have more genotypes than *baby trials*  
25 (between one and four genotypes (Ceccarelli and Grando, 2007) or one or two genotypes  
26 and one control (Virk and Witcombe, 2007)). In most of the *baby trials* described in the  
27 literature, the control is not replicated which does not allow to compare the genotypes  
28 tested on farm as there are no residual degrees of freedom.

29  
30 If a control is replicated several times in the trial, it is possible to assess the within-

1 trial variation in order to assess genotypic differences. This within-trial variation is a  
2 key point in the analysis of unreplicated IBD (Singh et al., 2003). In our case satellite  
3 farms have few replications for the control but the statistical model can cope with this  
4 difficulty as long as the network involves many farms.

5 In the wheat PPB project, block effects were sometimes significant. In such cases,  
6 the use of blocks led to a better estimation of the genotypic effects. In addition, as the  
7 field used to carry out the experiment in a farm changed every year, block effects that  
8 were not significant in a farm a given year could become significant in this farm another  
9 year. Thus it seems more cautious to keep block designs for regional farms. Besides it  
10 would be interesting to optimize trial designs, in particular the number of controls, the  
11 number of replications of controls and the proportion regional and satellite farms on  
12 the network.

13 In addition, it is possible to include farms that are intermediate between regional  
14 farms and satellite farms in the programme, that is farms with or without a block  
15 structure and two or more controls replicated twice.

16  
17 At the beginning of the wheat PPB programme, the farmers did not always see the  
18 interest to have controls. After discussions based on the results during meetings, an  
19 increasing number of farmers wanted to have controls on their farms. The methodology  
20 is still under construction and the results sent back to farmers become more and more  
21 accurate and complete. This points out the time which is needed to implement a PPB  
22 programme (Storup and de Kochko, 2011) and to create a social organisation for the  
23 group (Ali Brac de la Perrière et al., 2011). Time and discussions among the actors lead  
24 to a better appropriation of the project by farmers and researchers too.

25  
26 The genetic diversity used in the wheat PPB project is large and allows selection.  
27 Even if the number of germplasm may not be representative of the diversity present  
28 (Bonneuil et al., 2012), these germplasm came from a broad genetic basis. Crosses led  
29 to numerous new genotypes available for selection within and between populations.  
30 Moreover, the landraces and old varieties used by the farmers in the RSP have been

1 shown to be genetically diverse (Thomas et al., 2012). Such a broad diversity was also  
2 used in other PPB programmes (Ceccarelli et al., 2001).

3 The total number of genotypes tested has been fairly variable in PPB projects  
4 (Almekinders and Elings, 2001). For example, Virk and Witcombe (2007) tested 38  
5 genotypes in 49 to 197 farms over 6 years while Ceccarelli et al. (2001) tested 208  
6 genotypes in 9 farms in Syria. In addition to those trials, nurseries with more than  
7 500 segregating populations were also used (Ceccarelli et al., 2001). In the wheat PPB  
8 programme, 254 germplasms were tested over four years in around 25 farms per year :  
9 this allowed an efficient screening of the available diversity in numerous environments.  
10 Among these germplasms, only 232 were tested in an experimental design with controls,  
11 which should be noted since there were no technical assistance provided to farmers to  
12 conduct their trials.

13 The farmers chose the germplasms they wanted to experiment. The only constraint  
14 was to follow the design of a regional or a satellite farm and in particular to use one  
15 or several controls. It was not necessary for farmers to have the same controls. As long  
16 as one (or several) population(s) was (were) replicated, the residual variance of a trial  
17 could be estimated from the trial data and from the common distribution of residual  
18 variances.

19 When *baby trials* are used as replication, the farmers depend on the other farmers  
20 for sowing. Witcombe et al. (2005a) pointed this dependence : trials with replicated  
21 designs can be far away, while unreplicated trials have to be close to each other. Even  
22 if some farms were in the same pedo-climatic environment, farming practices such as  
23 rotations, sowing density and soil fertility management would differ between farmers  
24 leading to different selective pressures and therefore different selected populations. In  
25 the wheat PPB programme, farmers were independent of the others and did not have  
26 to take account of the choice of the other farmers. This provided much freedom and  
27 flexibility in the selection for the farmers.

## 1      5      Conclusion

2            In the wheat PPB programme, farmers take part in the whole breeding activities  
3 and decisions for selections as they select in early generations directly on their farm  
4 and according to their objectives. The developed methodology has three main advan-  
5 tages : (i) farmers do not have to handle balanced (or slightly unbalanced) replicated  
6 designs on their farm ; (ii) farmers are not dependent on the other farmers to choose the  
7 germplasms they want to test on their farm, thereby increasing the diversity selected  
8 in the network and providing more flexibility to farmers to make their selection ; (iii)  
9 population effects can be compared on farm so that farmers can select among their  
10 populations.

11           We replace the space available on farm by the space available overall the network.  
12 The developed statistical method allows a more effective use of data from collaborative  
13 plant breeding projects and can increase the participation of farmers in such projects.  
14 Our methodology could be implemented on other crops as long as the parallel structure  
15 of the trials is kept. This is an important point as the collaboration in the network is  
16 one of the farmers' motivations. Beyond their own results, they contribute to the results  
17 for the other farmers. They are not alone and work as a group (Storup and de Kochko,  
18 2011).

19           Moreover, as some controls and some populations are tested in several farms of  
20 the network, it would be interesting to estimate genotype, farm and  $G \times E$  effects.  
21 This would bring more information to farmers to choose candidate genotypes to test  
22 on their farm. Simulation studies would also be of interest to assess which levels of  
23 disequilibrium of the network lead to a reliable analysis.

24           Participatory research raises new questions and leads to the development of in-  
25 novative methods in plant breeding. The wheat PPB programme and the developed  
26 methodology are based on collective and collaborative work and action between far-  
27 mers, associations of farmers and researchers (Ali Brac de la Perrière et al., 2011). PPB  
28 is an interdisciplinary field which aims at gathering different disciplines and addresses  
29 global and multi-dimensional challenges : social (Demeulenaere et al., 2008), regulatory  
30 (Santilli, 2011), agronomic (Papy and Goldringer, 2011), genetics (Dawson et al., 2011),

- 1 ecology (Chateil et al., 2013) and statistics in order to bring societal action (Kastenhofer
- 2 et al., 2011).

## 1 Acknowledgments

2 We thank Nathalie Galic and Sophie Pin for their great help in measuring all the  
3 populations all around France. We thank all the farmers participating in the project.  
4 We thank Gerard Branlard from INRA Clermont Ferrand for his time when doing NIRS  
5 analysis. P. Rivière was funded by a grant from DIM ASTREA (Ile de France region).  
6 The research leading to these results has received funding from the European Commu-  
7 nity's Seventh Framework Programme (FP7/ 2007–2013) under the Grant Agreement  
8 n245058-Solibam (Strategies for Organic and Low-input Integrated Breeding and Ma-  
9 nagement).

---

## 1      **References**

- 2      Ali Brac de la Perrière R., de Kochko P., Neubauer C., Storup B., 2011. Visions  
3      paysannes de la recherche dans le contexte de la sélection participative : comment  
4      co-construire et mutualiser les connaissances sur les plantes ? Pour l'Emergence d'une  
5      Université du Vivant.
- 6      Almekinders C., Elings A., 2001. Collaboration of farmers and breeders : Participatory  
7      crop improvement in perspective. *Euphytica* 122(3), 425–438.
- 8      Altieri M.A., 2002. Agroecology : the science of natural resource management for poor  
9      farmers in marginal environments. *Agriculture, Ecosystems & Environment* 93(1–3),  
10     1–24.
- 11     Baldi P., Long A.D., 2001. A bayesian framework for the analysis of microarray expres-  
12     sion data : regularized t -test and statistical inferences of gene changes. *Bioinformatics*  
13     17(6), 509–519.
- 14     Bellon M.R., Reeves J., Center I.M., Improvement W., Innovation C.S.P.o.P.R., for  
15     Technology Development G.A., Institutional, Institute I.R.R., 2002. Quantitative  
16     analysis of data from participatory methods in plant breeding. CIMMYT.
- 17     Besag J., Higdon D., 1999. Bayesian analysis of agricultural field experiments. *Journal*  
18     *of the Royal Statistical Society : Series B (Statistical Methodology)* 61(4), 691–746.
- 19     Bonneuil C., Demeulenaere E., 2007. Vers une génétique de pair à pair ? l'émergence  
20     de la sélection participative. In : *Des sciences citoyennes ? La question de l'amateur*  
21     *dans les sciences naturalistes*, F Charvolin, A Micoud, L K Nyhart. L'Aude edn., pp.  
22     122–147.
- 23     Bonneuil C., Goffaux R., Bonnin I., Montalent P., Hamon C., Balfourier F., Goldringer  
24     I., 2012. A new integrative indicator to assess crop genetic diversity. *Ecological*  
25     *Indicators* 23, 280–289.
- 26     Box G., Tiao G., 1973. *Bayesian Inference in Statistical Analysis*. Addison-Wesley,  
27     Reading.

- 1 Carlin B.P., Louis T., 2008. Bayesian Methods for Data Analysis. Texts in Statistical  
2 Science. Chapman & Hall/CRC, third edn.
- 3 Ceccarelli S., 2012. Plant Breeding with farmers : a technical manual.
- 4 Ceccarelli S., Grando S., 2007. Decentralized-participatory plant breeding : an example  
5 of demand driven research. *Euphytica* 155(3), 349–360.
- 6 Ceccarelli S., Grando S., Bailey E., Amri A., El-Felah M., Nassif F., Rezgui S., Yahyaoui  
7 A., 2001. Farmer participation in barley breeding in syria, morocco and tunisia.  
8 *Euphytica* 122(3), 521–536.
- 9 Ceccarelli S., Grando S., Maatougui M., Michael M., Slash M., Haghparast R., Rah-  
10 manian M., Taheri A., Al-Yassin A., Benbelkacem A., Labdi M., Mimoun H., Nachit  
11 M., 2010. Plant breeding and climate changes. *The Journal of Agricultural Science*  
12 148(06), 627–637.
- 13 Ceccarelli S., Grando S., Singh M., Michael M., Shikho A., Al Issa M., Al Saleh A.,  
14 Kaleonjy G., Al Ghanem S., Al Hasan A., Dalla H., Basha S., Basha T., 2003.  
15 A methodological study on participatory barley breeding II. response to selection.  
16 *Euphytica* 133(2), 185–200.
- 17 Ceccarelli S., Grando S., Tutwiler R., Baha J., Martini A., Salahieh H., Goodchild  
18 A., Michael M., 2000a. A methodological study on participatory barley breeding i.  
19 selection phase. *Euphytica* 111(2), 91–104.
- 20 Ceccarelli S., Grando S., Tutwiler R., Baha J., Martini A., Salahieh H., Goodchild  
21 A., Michael M., 2000b. A methodological study on participatory barley breeding i.  
22 selection phase. *Euphytica* 111(2), 91–104.
- 23 Chable V., Conseil M., Serpolay E., Le Lagadec F., 2008. Organic varieties for cau-  
24 liflowers and cabbages in brittany : from genetic resources to participatory plant  
25 breeding. *Euphytica* 164, 521–529.



- 1 Chateil C., Goldringer I., Tarallo L., Kerbiriou C., Le Viol I., Ponge J.F., Salmon S.,  
2 Gachet S., Porcher E., 2013. Crop genetic diversity benefits farmland biodiversity in  
3 cultivated fields. *Agriculture, Ecosystems & Environment* 171, 25–32.
- 4 Dawson J., Murphy K., Jones S., 2008. Decentralized selection and participatory ap-  
5 proaches in plant breeding for low-input systems. *Euphytica* 160(2), 143–154.
- 6 Dawson J., Rivière P., Berthelot J., Mercier F., de Kochko P., Galic N., Pin S., Serpolay  
7 E., Thomas M., Giuliano S., Goldringer I., 2011. Collaborative plant breeding for  
8 organic agricultural systems in developed countries. *Sustainability* (3), 1206–1223.
- 9 De Schutter O., 2010. Rapport du rapporteur spécial sur le droit à l'alimentation. Tech.  
10 rep., Nations Unies.
- 11 Delmar P., Robin S., Tronik-Le Roux D., Daudin J.J., 2005. Mixture model on the  
12 variance for the differential analysis of gene expression data. *Journal of the Royal*  
13 *Statistical Society : Series C (Applied Statistics)* 54(1), 31–50.
- 14 Demeulenaere E., Bonneuil C., Balfourier F., Basson A., Berthelot J., Chesneau V.,  
15 Ferté H., Galic N., Kastler G., Koenig J., Mercier F., Payement J., Pommart A.,  
16 Ronot B., Rousselle Y., Supiot N., Zaharia H., Goldringer I., 2008. Étude des com-  
17 plémentarités entre gestion dynamique à la ferme et gestion statique en collection.  
18 *Les Actes du BRG* , 117–138.
- 19 Desclaux D., Nolot J.M., Chiffolleau Y., Leclerc C., Gozé E., 2008. Changes in the  
20 concept of genotype x environment interactions to fit agriculture diversification and  
21 decentralized participatory plant breeding : pluridisciplinary point of view. *Euphytica*  
22 163, 533–546.
- 23 Fufa F., Grando S., Kafawin O., Shakhathreh Y., Ceccarelli S., 2010. Efficiency of  
24 farmers' selection in a participatory barley breeding programme in Jordan. *Plant*  
25 *Breeding* 129, 156 – 161.
- 26 Gauch H., Piepho H., Annicchiarico P., 2008. Statistical analysis of yield trials by  
27 AMMI and GGE : further considerations. *Crop Sci* 48(3), 866–889.

- 1 Gelman A., Rubin D., 1992. Inference from iterative simulation using multiple sequences  
2 (with discussion). *Statistical Science* (7), 457–511.
- 3 Ghaouti L., Vogt-Kaute W., Link W., 2008. Development of locally-adapted faba bean  
4 cultivars for organic conditions in germany through a participatory breeding ap-  
5 proach. *Euphytica* 162, 257–268.
- 6 Gyawali S., Sunwar S., Subedi M., Tripathi M., Joshi K., Witcombe J., 2007. Collabo-  
7 rative breeding with farmers can be effective. *Field Crops Research* 101(1), 88–95.
- 8 Hajjar R., Jarvis D.I., Gemmill-Herren B., 2008. The utility of crop genetic diversity  
9 in maintaining ecosystem services. *Agriculture, Ecosystems and Environment* 123,  
10 261–270.
- 11 Hu X., Yan S., Shen K., 2013. Heterogeneity of error variance and its influence on  
12 genotype comparison in multi-location trials. *Field Crops Research* 149, 322–328.
- 13 Kastenhofer K., Bechtold U., Wilfing H., 2011. Sustaining sustainability science : The  
14 role of established inter-disciplines. *Ecological Economics* 70(4), 835–843.
- 15 Kempton R.A., Fox P., 1997. *Statistical methods for plant variety evaluation*. Chapman  
16 & hall ltd edn.
- 17 Kerr M., 2003. Linear models for microarray data analysis : Hidden similarities and  
18 differences. *Journal of Computational Biology* 10(6), 891–901.
- 19 Leiser W.L., Rattunde H.F., Piepho H.P., Parzies H.K., 2012. Getting the most out  
20 of sorghum low-input field trials in west africa using spatial adjustment. *Journal of*  
21 *Agronomy and Crop Science* 198(5), 349–359.
- 22 Lewin A., Richardson S., Marshall C., Glazier A., Aitman T., 2006. Bayesian modeling  
23 of differential gene expression. *Biometrics* 62(1), 1–9. PMID : 16542224.
- 24 Lönnstedt I., Britton T., 2005. Hierarchical bayes models for cDNA microarray gene  
25 expression. *Biostatistics* 6(2), 279–291.

- 1 Mendes Moreira P., Pêgo S., Vaz Patto C., Hallauer A., 2008. Comparison of selec-  
2 tion methods on ‘Pigarro’, a portuguese improved maize population with fasciation  
3 expression. *Euphytica* 163, 481 :499.
- 4 Morris M., Bellon M., 2004. Participatory plant breeding research : Opportunities and  
5 challenges for the international crop improvement system. *Euphytica* 136, 21–35.
- 6 Nashimoto K., Wright F., 2008. Bayesian multiple comparisons of simply ordered means  
7 using priors with a point mass. *Computational Statistics & Data Analysis* 52(12),  
8 5143–5153.
- 9 Newton A., Akar T., Baresel J., Bebeli P., Bettencourt E., Bladenopoulos K., Czembor  
10 J., Fasoula D., Katsiotis A., Koutis K., Koutsika-Sotiriou M., Kovacs G., Larsson H.,  
11 Carvalho M.P.d., Rubiales D., Russell J., Santos T.D., Patto M.V., 2010. Cereal lan-  
12 draces for sustainable agriculture. a review. *Agronomy for Sustainable Development*  
13 30(2), 237–269.
- 14 Nott D.J., Yu Z., Chan E., Cotsapas C., Cowley M.J., Pulvers J., Williams R., Little  
15 P., 2007. Hierarchical bayes variable selection and microarray experiments. *Journal*  
16 *of Multivariate Analysis* 98(4), 852–872.
- 17 Papy F., Goldringer I., 2011. Cultiver la biodiversité. *Courrier de l’environnement de*  
18 *l’INRA* (60), 55–62.
- 19 Plummer M., 2003. JAGS : A program for analysis of bayesian graphical models using  
20 gibbs sampling. In : *Proceedings*. Vienna, Austria.
- 21 Rivière P., Goldringer I., Berthelot J.F., Galic N., Pin S., De Kochko P., Dawson J.C.,  
22 2013. Response to farmer mass selection in early generation progeny of bread wheat  
23 landrace crosses. *Renewable Agriculture and Food Systems FirstView*, 1–12.
- 24 Robert C., 2001. *The Bayesian Choice*. Springer Texts in Statistics. Springer, second  
25 edn.
- 26 Santilli J., 2011. *Agrobiodiversity and the law*. Routledge.

- 1 Shiferaw B., Smale M., Braun H.J., Duveiller E., Reynolds M., Muricho G., 2013.  
2 Crops that feed the world 10. past successes and future challenges to the role played  
3 by wheat in global food security. *Food Security* , 1–27.
- 4 Singh M., Malhotra R.S., Ceccarelli S., Sarker A., Grando S., Erskine W., 2003. Spatial  
5 variability models to improve dryland field trials. *Experimental Agriculture* 39, 151–  
6 160.
- 7 Smith M.E., Castillo G. F., Gomez F., 2001. Participatory plant breeding with maize  
8 in mexico and honduras. *Euphytica* 122(3), 551–565.
- 9 Smyth G., 2004. Linear models and empirical bayes methods for assessing differential  
10 expression in microarray experiments. In : *Statistical Applications in Genetics and*  
11 *Molecular Biology*, Michael P.H. Stumpf, Imperial College London, vol. 3.
- 12 Storup B., de Kochko P., 2011. Suivi d’une expérience de sélection participative en  
13 cours l’exemple du projet européen SOLIBAM. Tech. rep., Association pour une  
14 Fondation Science Citoyenne.
- 15 Theobald C.M., Talbot M., Nabugoomu F., 2002. A bayesian approach to regional and  
16 local-area prediction from crop variety trials. *Journal of Agricultural, Biological, and*  
17 *Environmental Statistics* 7(3), 403–419.
- 18 Thomas M., Demeulenaere E., Dawson J.C., Khan A.R., Galic N., Jouanne-Pin S., Re-  
19 moue C., Bonneuil C., Goldringer I., 2012. On-farm dynamic management of genetic  
20 diversity : the impact of seed diffusions and seed saving practices on a population-  
21 variety of bread wheat. *Evolutionary Applications* 5(8), 779–795.
- 22 Virk D., Chakraborty M., Ghosh J., Prasad S., Witcombe J., 2005. Increasing the  
23 client orientation of maize breeding using farmer participation in eastern india. *Ex-  
24 perimental Agriculture* 41(04), 413–426.
- 25 Virk D., Witcombe J., 2007. Evaluating cultivars in unbalanced on-farm participatory  
26 trials. *Field Crops Research* 106(2), 105–115.

- 1 Vom Brocke K., Trouche G., Weltzien E., Barro-Kondombo C.P., Gozé E., Chantereau  
2 J., 2010. Participatory variety development for sorghum in burkina faso : Farmers'  
3 selection and farmers' criteria. *Field Crops Research* 119(1), 183–194.
- 4 Westfall P., 1997. A bayesian perspective on the bonferroni adjustment. *Biometrika*  
5 84(2), 419–427.
- 6 Witcombe J.R., Gyawali S., Sunwar, S., Sthapit B.R., Joshi K.D., 2005a. Participatory  
7 plant breeding is better described as highly client-oriented plant breeding. II. optional  
8 farmer collaboration in the segregating generations. *Expl. Agric.* 42, 79–90.
- 9 Witcombe J.R., Joshi A., Joshi K.D., Sthapit B.R., 1996. Farmer participatory crop  
10 improvement. i. varietal selection and breeding methods and their impact on biodi-  
11 versity. *Expl. Agric.* 32, 445–460.
- 12 Witcombe J.R., Joshi K.D., Gyawali S., Musa A.M., Johansen C., Virk D.S., Sthapit  
13 B.R., 2005b. Participatory plant breeding is better described as highly client-oriented  
14 plant breeding. i. four indicators of client-orientation in plant breeding. *Experimental*  
15 *Agriculture* 41(3), 299–319. WOS :000231655300002.
- 16 Wolfe M., Baresel J., Desclaux D., Goldringer I., Hoad S., Kovacs G., Löschenberger  
17 F., Miedaner T., Østergård H., Lammerts van Bueren E., 2008. Developments in  
18 breeding cereals for organic agriculture. *Euphytica* 163(3), 323–346.



# Annexe L : Modélisation hiérarchique des interactions génotype × environnement pour des expérimentations flexibles de sélection décentralisée et participative

Cet article n'est qu'une première version qui sera remaniée et approfondie. Comme pour l'article sur les comparaisons de moyennes (Annexe K), les parties plus techniques et quelques figures seront dans les données supplémentaires.

Parmi les points à approfondir, je testerai plusieurs approches de validation croisée. Cela me permettra de tester si le modèle peut prédire le comportement de germplasms, d'environnements, voire d'années. Cela permettra également de pouvoir comparer avec les autres validation croisées de la littérature. Ce point est encore imprécis dans la discussion.

De plus, j'estimerai les effets populations sur chaque environnement avec le modèle décrit dans l'annexe K. Pour l'instant, les données sont corrigées pour l'effet bloc dans les fermes régionales et ne sont pas corrigées pour les fermes satellites.

Les « ? » dans les citations correspondent à l'article sur les comparaisons de moyennes (Annexe K), qui n'est pas encore référencé !

1 Hierarchical modeling of genotype × environment  
2 interaction for flexible experiments in decentralized  
3 participatory plant breeding

4 Pierre Rivière<sup>1</sup> Julie C. Dawson<sup>2</sup> Isabelle Goldringer<sup>1</sup>

5 Olivier David<sup>3</sup>

6 27 novembre 2013

7 <sup>1</sup> INRA Le Moulon, UMR 0320 de Génétique Végétale, Ferme du Moulon F-91190 Gif sur  
8 Yvette, France

9 <sup>2</sup> Department of Horticulture, University of Wisconsin-Madison, 1575 Linden Dr. Madison,  
10 WI 53706, USA

11 <sup>3</sup> INRA, UR 341 de Mathématiques et Informatique Appliquées, Domaine de Vilvert  
12 F-78352 Jouy en Josas, France

13 **Abstract**

14 Decentralized participatory plant breeding (PPB) is aimed at creating varieties  
15 adapted to agroecology based systems. In PPB, selection is decentralized in the target  
16 environments and relies on collaboration between farmers, NGOs and researchers. The  
17 method takes into account genotype × environment ( $G \times E$ ) interactions to select for  
18 specific adaptation. In a wheat PPB project carried out in France, each farmer parti-  
19 cipating in the project chooses the germplasm to be evaluated on his farm. Although  
20 there was only one germplasm present on all farms, some germplasm were evaluated on



1 several farms, leading to a dataset severely non orthogonal with 95% of combinaisons  
2  $G \times E$  missing. In order to get as much as information from the dataset, we studied  
3 the  $G \times E$  interaction with a hierarchical bayesian modeling of the Finlay Wilkinson  
4 (HFW) model. This model is more parcimonious to deal with the disequilibrium as it  
5 estimates a global sensivity to interaction for each germplasm. We found that the hie-  
6 rarchical approach gave robust results for the estimation of genotypic, environmental  
7 and sensitivity effects despite the high disequilibrium of the data set. Cross validation  
8 studies showed that our model was predictive : correlation between observed and pre-  
9 dicted values ranged from 0.42 to 0.86 according to the variables studied. Our approach  
10 allowed for a wide screen of the genetic diversity present on the network of farms as 117  
11 germplasms out of 254 were grown on at least two environments. This approach may  
12 contribute to the overall methodology of PPB : farmers can choose new germplasms to  
13 test as candidates on their farms based on these results. We provide new tools to run  
14 on-farm decentralized breeding that can be implemented for a large range of species. R  
15 scripts are available under request.

16 **Keywords** : decentralized participatory plant breeding ; unbalanced data ;  $G \times E$   
17 interaction ; hierarchial model ; MCMC ; bread wheat ; bayesian statistics.

# 1 Introduction

Varietal selection for agroecology, low input and Organic Agriculture (OA) based system is a major concern to reach sustainable agriculture (Wolfe et al., 2008). In OA, the use of chemical input (nitrogen, phytochemicals) is not allowed (Agence-Bio, 2011) and irrigation may not be always possible. Therefore the cultivation environments are not optimized by inputs and the varieties grow in very diverse conditions (Dawson et al., 2008). These environments are dependent on pedo-climatic conditions as well as on the practices of the farmers : sowing date, sowing density, tillage, etc. (Desclaux et al., 2008).

In order to develop varieties adapted to such very diverse environments two strategies can be used : (i) centralized and indirect selection or (ii) decentralised and direct selection. The key difference between both approaches is the way to take the genotype-environment ( $G \times E$ ) interactions into account. The  $G \times E$  interaction is considered by plant breeder as the main factor limiting efficiency of breeding programme (i.e. response to selection) (Ceccarelli et al., 2001). In centralised and indirect selection, breeding families are evaluated and selected in a few stations supposed to represent the target environments. This is efficient if there is high additive genetic correlation between the two environments and the narrow sense heritability is high in the selection environment (Falconer and Mackay, 1996).

In decentralized selection, evaluation and selection are conducted directly in target environments (i.e. on farm). This approach takes into account  $G \times E$  interactions that are important in OA (Murphy et al., 2007; Dawson et al., 2008). The selection maximize the positive use of  $G \times E$  interactions to select for specific adaptation. This method was already shown to be effective (Ceccarelli et al., 2000; Smith et al., 2001; Virk et al., 2005; Murphy et al., 2007) or close to direct selection (Annicchiarico et al., 2010). As theoretically shown by Falconer (1952), direct selection should often be more efficiency than indirect selection.

In participatory selection, farmers do the selection based on interactions with the researchers in early generation (Witcombe et al., 1996). This increases the efficiency of the breeding program as farmers bring their own knowledges related to their environ-

1       ments and of the required traits they wish for their varieties (Ceccarelli and Grando,  
2       2007). Beyond plant breeding, farmers take place into the decision process leading to  
3       empowerment and new governance of the seed system (Bonneuil and Demeulenaere,  
4       2007; Morris and Bellon, 2004).

5       The wheat PPB programme started in France in 2005, as a collaboration between  
6       INRA Le Moulon and the Réseau Semences Paysannes (RSP, the Farmers' seed net-  
7       work). This PPB programme has three objectives : (i) develop population-varieties  
8       adapted to farmer's practices and needs (organic management, bread quality ...) in a  
9       participatory approach, (ii) develop strategies for preserving genetic diversity through  
10      on-farm dynamic management and plant breeding and (iii) learn and improve farmers'  
11      breeding techniques and diffuse it broadly.

12      In this context, a large number of populations are evaluated over a large network  
13      of farms in the RSP (Dawson et al., 2011; Rivière et al., 2013; ?). In this article, we  
14      focused on the first objective (i.e. develop population-varieties) and tried to improve  
15      population assessment and comparison on each farm by using the information obtained  
16      all over the network.

17      Each farmer in the program chooses the population to be evaluated on his farm.  
18      This leads to lots of unbalanced data as very few populations are present on all the  
19      farms. However, some populations are present on several farms which allow for a link  
20      between the farms.

21      Several statistical models have been proposed to take  $G \times E$  into account (Pie-  
22      pho, 1998; Annicchiarico, 2002) : Additive Main Effects and Multiplicative Interaction  
23      (AMMI) (Gauch et al., 2008), GGE (Yan et al., 2007), joint regression (Finlay and Wil-  
24      kinson, 1963), Shukla stability variance (Shukla, 1972), ecovalence of Wricke (Wricke,  
25      1962). These methods have proved to be robust for balanced data. Nabugoomu et al.  
26      (1999) and Theobald et al. (2002) explored frequentist and bayesian approaches respec-  
27      tively to deal with slightly unbalanced trials. But no studies analysed very unbalanced  
28      data such as those obtained in the wheat PPB project.

29      It is impossible to assess interaction term due to a germplasm in an environment  
30      if they do not occur together. Moreover, variance heterogeneity among germplasm is

1 not orthogonal to variance heterogeneity among environments leading to bias in the  
2 estimation of genotype and environment effects. This can be overcome if a joint esti-  
3 mation is performed, as it can be done with hierarchical bayesian model (Edwards and  
4 Jannink, 2006).

5 In order to get as much information as possible from our data set, we used the  
6 model of Finlay Wilkinson. The Finlay Wilkinson model estimates germplasm effects,  
7 environmental effects and the sensitivity of a germplasm to the interaction (Finlay and  
8 Wilkinson, 1963). This sensitivity can be seen as an average behavior of a germplasm in  
9 a set of environments. As the data set is severely non orthogonal and we have many data,  
10 we decided to solve the Finlay Wilkinson model with a hierarchical bayesian approach.

11 There is a large literature on hierarchical bayesian models. For example, they have  
12 been used for modelling  $G \times E$  interactions (Theobald et al., 2002; Cotes et al., 2006;  
13 Edwards and Jannink, 2006) and micro array studies (Lönstedt and Britton, 2005;  
14 Nott et al., 2007). Micro-array experiments also have a small numbers of replicates  
15 (Smyth, 2004; Lewin et al., 2006). Thus the model used to analyse their data involve  
16 additional assumptions to stabilise estimates of gene variances. In hierarchical models,  
17 a separate variance is allowed for each gene but gene variances are assumed to come  
18 from a common distribution (Baldi and Long, 2001). Hierarchical models allow robust  
19 estimation of residual variances of genes. We transposed this method to our variety trials  
20 by assuming that the disequilibrium of the data can be compensated by modelisation  
21 of common distributions where parameters of the Finlay Wilkinson model are drawn.

22 In this study, we investigated the model of Finlay Wilkinson with a hierarchical  
23 bayesian approach to assess germplasm and environmental effects overall the network  
24 and study the sensitivity of germplasms to interaction. First, the model of Finlay Wil-  
25 kinson has been developed and fitted to the data using a bayesian approach. Then, we  
26 assessed the interest of the approach by testing (i) the disequilibrium in the estimation  
27 of the parameters, (ii) the influence of an additive and a non hierarchical model and  
28 (iii) the prediction of the model by a cross validation study.

## 2 Materials and methods

In the article a germplasm was a biological entity which could be a landrace, an old varietie, a mixture, a farmers' mass selection or a derived cross population. A population was a given germplasm in one farm in one year.

### 2.1 Experimental design and data set

#### 2.1.1 Experimental design

Data of the four years 2008-2009, 2009-2010, 2010-2011 and 2011-2012 were used in the analysis. There were between 15 and 22 farms each year. In the following the combinaison of a farm in a given year was considered as an environment. A farm is identified by three letters. Thus, environments refered to these three letters followed by the year of harvest. For example JFB2009 stands for the experiment on JFB's farm on the year 2008-2009.

Environments were divided into two types of farm based on their experimental design : regional and satellite farms (Dawson et al., 2011). Experimental design was the result of a co-construction between the reseach team and the farmers. All the farms shared one control : the population *Rouge-du-Roc*. Regional farms grew between 20 and 32 plots, satellite farms grew between 5 and 20 plots. Regional farms had a two block structure design with four controls in each block (one being *Rouge-du-Roc*). The farmers chose the other populations to grow. *Rouge-du-Roc* was also replicated twice on satellite farms and the farmers chose the other populations to grow. More details on the experimental design can be found in ?.

Overall, there were 254 germplasms grown on 75 different environments. In a  $I \times J$  matrix, with  $I$  the number of germplasms and  $J$  the number of environments, 95% of the combinaisons of germplasm  $\times$  environment were not present.

In order to assess  $G \times E$  interaction, we worked on 117 germplasms that were present on at least 2 environments, i.e. 46% of the whole data set (Figure 1).

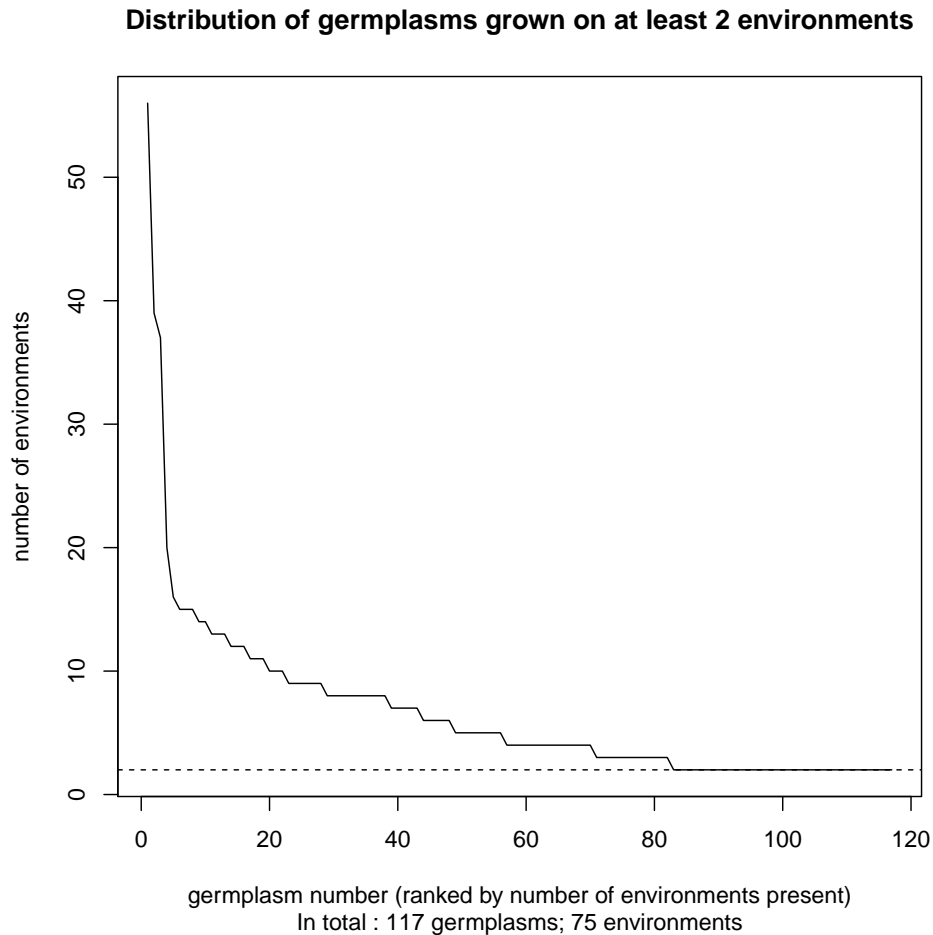


Figure 1 – Number of environment (farm × year) with germplasm that were present on at least two environments in the network of farms. The dash line represents two environments.

### 1      2.1.2 Data collected

2            Two types of variables have been measured : at the individual level (25 measures on  
3            25 plants within the population) and at the global level : based on the average measure  
4            on sample of the 25 plants. The traits measured have been chosen based on discussion  
5            between farmers and researchers. They were also based on phenotypic descriptors used  
6            for variety registration. As it was not possible to do the measures on all the farms in  
7            France, measures were done differently for regional and satellite farms. Some measures  
8            were done on farm and at the research station on the sample of spikes sent by the  
9            farmers (Table 1).

variable	RF	SF	method
color <sup>i,1</sup>	X	X	color of the spikes from 0 (white) to 2 (red). Possible values were 0, 0.5, 1, 1.5 and 2.
awns <sup>i,1</sup>	X	X	presence of awns on the spikes from 0 (no awns) to 2 (full of awns). Possible values were 0, 0.5, 1, 1.5 and 2.
curve <sup>i,1</sup>	X	X	curvature of the spike at maturity from 0 (no curve) to 2 (spike up side down). Possible values were 0, 0.5, 1, 1.5 and 2.
plant height <sup>g,i,1</sup>	X		height of the plant from the ground to the top of the spike in mm.
last leaf spike distance (LLSD) <sup>i,1</sup>	X		distance from the flag leaf to the bottom of the spike in mm.
spike length <sup>i,1</sup>	X		length of the spike in mm.
nbr spikelets <sup>i,2</sup>	X		number of spikelets.
sterility <sup>i,2</sup>	X		sum of sterile spikelets per spike.
spike weight <sup>i,2</sup>	X	X	weight of the spike in g.
protein <sup>g,2</sup>	X	X	% of protein content in the grain measured with NIRS technology at INRA Clermont Ferrand France on grain using near infrared spectroscopy (FOSS NIRSystem 6500).
tkw <sup>g,2</sup>	X	X	thousand kernels weight in g.

Table 1 – Variables measured and method associated. RF : regional farms; SF : satellite farms; <sup>g</sup> : global measures; <sup>i</sup> : individual measures; <sup>1</sup> : measure on farm; <sup>2</sup> : measure at the research station on samples of spikes sent by farmers.

- 1 The final variables used were either the means over the 25 individuals of the samples
- 2 in each population or the global measure recorded at the population level. Differences in
- 3 the number of environments for each variable were due to missing data. The data were

1 adjusted for block effect on regional farms to correct for intra environmental variation.

2 On satellite farms no correction was applied as there was no block structure.

3 In order to study the disequilibrium in data distribution, the data set was divided

4 into several subset (Table 2). The name of the subset takes the following nomenclature :

$$\underbrace{DATA-j}_{(1)} - \underbrace{I \times J}_{(2)} - \underbrace{\%NA}_{(3)}$$

5 with (1) the name of the subset,  $DATA-j$  was a data set where all germplasms were

6 present on at least  $j$  environments. For each data set  $j \in [2, J]$  with  $J$  the maximum

7 depending on the variable.  $J$  was chosen so that enough information was available to

8 run the model; (2) the dimension of the model matrix containing  $I$  rows (the number

9 of germplasms) and  $J$  columns (the number of environments) and (3) the propor-

10 tion of missing data in the subset in percentage ( $NA$  : Non Available). For example

11  $DATA-2_{106 \times 53}_{90}$  was a subset of 106 germplasms and 53 environments where all

12 germplasms were present on at least 2 environments, it had 90% of missing data.



color		awns		curve		nbr sterile spikelets		plant height		LLSD							
$j$	$I \times J$	$\%NA$	$j$	$I \times J$	$\%NA$	$j$	$I \times J$	$\%NA$	$j$	$I \times J$	$\%NA$						
2	106 × 53	90	2	112 × 67	90.8	2	106 × 52	89.9	2	103 × 60	89.4	2	102 × 60	89.5	2	101 × 52	88.8
3	76 × 52	87.3	3	79 × 66	88	3	76 × 51	87.2	3	77 × 60	87	3	76 × 60	87	3	72 × 52	85.8
4	53 × 52	84.4	4	67 × 66	86.6	4	51 × 51	83.9	4	65 × 60	85.5	4	66 × 60	85.8	4	59 × 52	83.9
5	45 × 52	82.9	5	51 × 66	84.3	5	45 × 51	82.7	5	50 × 60	83.2	5	50 × 60	83.4	5	43 × 52	80.8
6	34 × 52	80.5	6	44 × 66	83.1	6	34 × 51	80.3	6	40 × 60	81.1	6	39 × 60	81.1	6	37 × 52	79.3
7	23 × 52	76.8	7	40 × 66	82.3	7	23 × 51	76.6	7	35 × 60	79.8	7	36 × 60	80.3	7	34 × 52	78.5
8	19 × 52	74.7	8	31 × 66	80.2	8	18 × 51	73.9	8	31 × 60	78.7	8	29 × 60	78.4	8	27 × 52	76.4
9	14 × 51	70.6	9	24 × 66	78	9	14 × 50	70.3	9	21 × 60	74.9	9	20 × 60	74.7	9	16 × 51	70.1
10	11 × 50	66.7	10	19 × 65	75.4	10	11 × 49	66.4	10	16 × 59	71.3	10	16 × 59	71.6	10	12 × 51	66
11	9 × 50	63.8	11	15 × 64	72.5	11	9 × 49	63.5	11	14 × 59	69.6	11	11 × 59	66.4	11	10 × 51	63.1
13	6 × 50	56.7	12	14 × 64	71.8	13	6 × 49	56.5	12	12 × 59	67.7	12	10 × 58	64.3	12	9 × 50	60.7
			13	9 × 64	66.5				13	8 × 58	61	13	9 × 58	62.6			
			14	8 × 63	64.3				14	6 × 57	54.4						

spike length		nbr spikelets		spike weight		tkw		protein						
$j$	$I \times J$	$\%NA$	$j$	$I \times J$	$\%NA$	$j$	$I \times J$	$\%NA$	$j$	$I \times J$	$\%NA$			
2	100 × 48	88.1	2	78 × 43	87.3	2	106 × 56	89.9	2	108 × 71	90.6	2	84 × 32	86.9
3	70 × 48	84.8	3	61 × 43	85	3	80 × 55	87.5	3	79 × 71	88.2	3	48 × 32	81.8
4	58 × 48	82.9	4	43 × 43	81.7	4	54 × 55	84.2	4	68 × 71	87	4	35 × 32	78.5
5	40 × 48	79	5	36 × 43	79.9	5	48 × 55	83.1	5	52 × 71	84.7	5	21 × 32	72.5
6	36 × 48	77.8	6	32 × 43	78.9	6	36 × 55	80.5	6	46 × 71	83.6	6	14 × 31	65.4
7	33 × 48	77	7	20 × 43	74.5	7	27 × 55	77.6	7	41 × 71	82.7	7	9 × 31	57
8	27 × 48	75.1	8	14 × 43	70.6	8	20 × 55	74.3	8	32 × 71	80.6	8	6 × 31	46.8
9	14 × 47	66.7	9	10 × 42	65.5	9	16 × 55	71.5	9	26 × 71	78.7			
10	12 × 47	64.4				10	11 × 54	65.3	10	19 × 70	75.2			
11	10 × 47	61.5							11	17 × 70	73.9			
12	8 × 46	56.8							12	13 × 70	70.8			
									13	11 × 70	68.6			
									14	8 × 69	63.2			

Table 2 – Characteristics of each subset for each variable.  $j$  refers to the subset  $DATA-j$  which is a data set where all germplasms were present on at least  $j$  environments;  $I \times J$  is the dimension of the model matrix containing  $I$  rows (the number of germplasms) and  $J$  columns (the number of environments) and  $\%$  is the proportion of missing data in the subset in percentage. For example  $DATA-2_{106 \times 53_{90}}$  was a subset of 106 germplasms and 53 environments where all germplasms were present on at least 2 environments, it had 90% of missing data.

## 1      2.2    The Hierarchical Finlay Wilkinson (HFW) model

### 2      2.2.1   Likelihood

3             $Y_{ij}$ , the phenotypic value for a given variable of germplasm  $i$  and environment  $j$ ,  
4 was written as :

$$Y_{ij} = \alpha_i + \theta_j + \eta_i\theta_j + \varepsilon_{ij}; \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_e^2)$$

5            where,  $\alpha_i$  was the effect of germplasm  $i$ ;  $\theta_j$  was the effect of environment  $j$ ;  $\eta_i\theta_j$   
6 was the interaction between germplasm  $i$  and environment  $j$ ;  $\varepsilon_{ij}$  was the residual (i.e.  
7 the interaction not explained by  $\eta_i\theta_j$ ).

8

9            This model can also be written as :

$$\begin{aligned} Y_{ij} &= \alpha_i + \theta_j + \eta_i\theta_j + \varepsilon_{ij} && ; \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_e) \\ Y_{ij} &= \alpha_i + (1 + \eta_i)\theta_j + \varepsilon_{ij} && ; \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_e) \\ Y_{ij} &= \alpha_i + \beta_i\theta_j + \varepsilon_{ij} && ; \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_e) \quad (1) \end{aligned}$$

10            Where  $\beta_i = (1 + \eta_i)$  was the sensitivity of germplasm  $i$  to environments. We used  
11 the last model (model 1), which is known as the Finlay Wilkinson model (Finlay and  
12 Wilkinson, 1963). This model assess the sensitivity coefficient  $\beta$  for each germplasm.  $\beta$   
13 varies around 1 and  $\beta$  closer to 1 indicates more stable germplasm.

14            Considering the high non orthogonality and the amount of data, we decided to  
15 solve this model with a hierarchical Bayesian approach. In the following, the model 1  
16 of Finlay Wilkinson resolved with a hierarchical bayesian approach was noted as the  
17 HFW model.

### 18      2.2.2   Prior distributions

Each parameter of the model were taken from the following distributions :

$$\alpha_i \sim \mathcal{N}(\mu, \sigma_v^2); \quad \beta_i \sim \mathcal{N}(1, \sigma_s^2); \quad \theta_j \sim \mathcal{N}(0, \sigma_l^2); \quad \sigma_e^{-2} \sim \mathcal{Gamma}(10^{-6}, 10^{-6})$$

1 A vague prior distribution was used for  $\sigma_e$ .

Then, we placed vague priors on hyperparameters  $\mu$  (which was the general mean of  $\alpha_i$ ),  $\sigma_v^2$ ,  $\sigma_s^2$  and  $\sigma_l^2$  :

$$\mu \sim \mathcal{N}(\nu, \nu^2); \quad \sigma_v \sim \text{Uniforme}(0, \nu); \quad \sigma_s \sim \text{Uniforme}(0, 1); \quad \sigma_l \sim \text{Uniforme}(0, \nu),$$

2 where  $\nu$  was the arithmetic mean of the data :  $\nu = \sum_{ij} Y_{ij}/n$  where  $n$  was the  
 3 number of observations. The distribution of variance were borned to 0 as variance are  
 4 positive by definition.

## 5 2.3 Posterior distributions

6 Bayesian inference is based on the posterior distribution of model parameters. The  
 7 distributions were estimated using Markov Chain and Monte Carlo (MCMC) methods.  
 8 These methods simulate values for the parameter of the model according to a Markov  
 9 Chain that converges to the posterior distribution of model parameters (Robert, 2001).

10 MCMC methods were implemented using R, rjags and JAGS packages (Plummer,  
 11 2003) that performed Gibbs sampling (Robert, 2001). R scripts are available upon re-  
 12 quest.

13 Two MCMC chains were run independently to test for convergence. A Gelman  
 14 test was used to test for convergence of the two chains. This test was based on the  
 15 variance within and between the chains (Gelman and Rubin, 1992). A burn-in and lots  
 16 of iterations were needed in the MCMC procedure. In our case, the burn-in had 1000  
 17 iterations, then 100 000 iterations were done with a thinning interval of 10 to reduce  
 18 autocorrelations between samples, so that 10 000 samples were available for inference  
 19 for each chain. For all the parameters convergence diagnostics did not detect any lack  
 20 of convergence. Initial values for each chain were taken randomly except for  $\mu$ ;  $\sigma_v$  and  
 21  $\sigma_l$  that were the median of their respective posterior distributions from the additive  
 22 model (i.e. model 1 with  $\forall i; \beta_i = 1$ ).

23 For each factor ( $\alpha$ ,  $\beta$  or  $\theta$ ), the estimation of each parameter was put in a  $G$  matrix  
 24 of dimension  $P \times V$ , with  $P$  the number of parameters estimated and  $V$  the number

1 of variable (11 in our case). There were as many  $G$  matrix as factor studied. As all the  
 2 parameters did not have a value for each variable (all germplasms or all environments  
 3 did not have always a value for each variable), the  $G$  matrix had some missing data.

4 On each  $G$  matrix, we performed a PCA on the  $G$  matrix followed by K-means  
 5 procedure to create groups of similar parameters (Husson et al., 2010). The Kmeans  
 6 were done on the two first axes of a PCA. The first axes represent the main information  
 7 when the last represent mainly noise (Husson et al., 2010; Gauch et al., 2008). The  
 8 number of groups was choosen in order to maximise the variance between the groups.  
 9 We used the R function PCA and HCPC from package FactoMineR.

## 10 2.4 Parameter comparisons

11 Let  $par_i$  the  $i^{\text{th}}$  parameter of a factor which could be  $\alpha$ ,  $\beta$  or  $\theta$ . For a given factor,  
 12 each pair of parameters were compared. Let  $H_0$  and  $H_1$  denote the hypotheses :

$$H_0 : "par_i \geq par_{i'}", H_1 : "par_i < par_{i'}".$$

13 The difference  $par_i - par_{i'}$  between the means of parameter  $i$  and paramter  $i'$  was  
 14 considered as significant if either  $H_0$  or  $H_1$  had a high posterior probability, that is  
 15 if  $Pr\{H_0|y\} > 1 - \alpha$  or  $Pr\{H_1|y\} > 1 - \alpha$ , where  $\alpha$  was some specified threshold.  
 16 The difference was considered as not significant otherwise. The posterior probability  
 17 of a hypothesis was estimated by the proportion of MCMC simulations for which this  
 18 hypothesis was satisfied. We can assimilate this probability to a *pvalue* (Westfall, 1997;  
 19 Nashimoto and Wright, 2008).

20 Groups of parameters could be created with a given threshold. The threshold used  
 21 for the type I error was  $\alpha = 0.05/n$  (with  $n$  the number of parameter to compare).

## 1      2.5    Comparison of the HFW model to the hierarchical 2      additive model and to the non hierarchical Finlay Wilkin- 3      son model

We compared HFW model (model 1) to the hierarchical additive model (model 2) which was HFW model with  $\forall i; \beta_i = 1$  and to a non hierarchical Finlay Wilkinson model (model 3), in which the prior distributions used were non informative :

$$\alpha_i \sim \mathcal{N}(\mu, 10^6); \quad \beta_i \sim \mathcal{N}(1, 1); \quad \theta_j \sim \mathcal{N}(0, \nu^2); \quad \sigma_{ij}^{-2} \sim \mathcal{Gamma}(10^{-6}, 10^{-6})$$

4      To test for model differentiation, we used the DIC criterion, which is close to the AIC  
5      criterion (Spiegelhalter et al., 2002). We looked at *diffDIC*, the differences between  
6       $DIC_{hier}$  which was the value of DIC for model 1 and  $DIC_{model}$ , *model* being the additive  
7      model 2 or the non hierarchical model 3. If  $diffDIC < 0$  then, the hierarchical model  
8      fitted better the data than the other model (Plummer, 2008).

## 9      2.6    Impact of the increasing unbalanced data in the esti- 10     mation of the parameters

11     We run HFW model on all the subsets presented in Table 2 to see how the model  
12     responded to different degree of unbalanced data.

## 13     2.7    Cross validation

14     To test if the HFW model was predictive, we ran cross validation studies on all the  
15     subsets generated (Table 2). We run replicated cross validation which means that all  
16     the individuals of the data set were in the training population at least once (Daetwyler  
17     et al., 2013).

18     Let  $DATA-j$  be a subset generated (see section 2.1.2 and Table 2) where all germ-  
19     plasms were present on at least  $j$  environments. Here  $j \in [3, J]$  with  $J$  the maximum  
20     depending on the variable (see Table 2).

1 From each *DATA-j*, we deleted each  $Y_{ij}$  value from germplasm  $i$  and environment  
2  $j$  to create  $D_k$ . There were as many  $D_k$  as observation  $Y_{ij}$  in *DATA-j*.

3 HFW model was run on each  $D_k$ . We calculated  $\hat{Y}_{ij}$  the estimated value of  $Y_{ij}$   
4 knowing the value of  $\alpha_i$ ,  $\beta_i$  and  $\theta_j$  :  $Distribution(\hat{Y}_{ij}) = MCMC(\alpha_i) + MCMC(\beta_i) \times$   
5  $MCMC(\theta_j)$ . In total, considering all the variable, there were more than 35000  $D_k$  ran.  
6 We managed to proceed such calculation in a reasonable time because of the use of  
7 HTcondor (<http://research.cs.wisc.edu/htcondor/index.html>). R functions and  
8 bash scripts have been written to perform the parallelization and the calculation.

9 Based on  $Distribution(\hat{Y}_{ij})$ , we calculated the probability  $Pr\{\hat{Y}_{ij} = Y_{ij}\}$  (see section  
10 2.4). This probability was negatively correlated to  $|\hat{Y}_{ij} - Y_{ij}|$ .

11 Correlation between the observed value and the estimated value were done on the  
12 biggest data set, *DATA-3*.

13 An ANOVA to assess effects of parameters that could best explain the variation of  
14 the accuracy in the estimation was performed according to the following model :

$$15 \quad Pr\{\hat{Y}_{ij} = Y_{ij}\} = DATA-j + \alpha_i + \theta_j$$

16 With  $Pr\{\hat{Y}_{ij} = Y_{ij}\}$ , the probability that the estimation of the model  $\hat{Y}_{ij}$  equals the  
17 real value  $Y_{ij}$ , *DATA-j* the data set with all germplasms on at least  $j$  environments,  
18  $\alpha_i$  the effect of the germplasm,  $\theta_j$  the effect of the environment.

## 3 Results

### 3.1 Parameter estimations

HFW model was run on the entire data set. For all the variables, the parameters, as well as all  $\epsilon_{ij}$ , were taken in normal distributions. For each variable and each parameter, groups significantly different were created, except for  $\beta$  where only one variable, spike weight, had two groups, all the others variables had only one group (Table 3). When summarizing the informations for a given factor with the  $G$  matrix, groups can be created (Figure 2, 3, 4).

We looked at germplasm mass selected by farmers and their corresponding non selected germplasm. Such cases appeared 20 times. For 13 of them, the distance between the germplasms were below 0.5, meaning that the germplasms were close (Figure 2). For five of them, they were in the two different group found with HCPC. Finally,  $C21$  and its corresponding selected germplasm were divided in two groups with  $C21$  (n°23) and  $C21\#S$  (n°24) in one group and with  $C21\#Sb$  (n°25) and  $C21\#Snb$  (n°26) in another group. The same pattern was observed with  $C37\#epi-long-crosse$  (n°4) and  $C37\#S-epis-crosses-S-JF$  (n°45) in one group and  $C37$  (n°43) in another.

As for germplasm effect, sensitivities ( $\beta_i$ ) did not show any trend among germplasm.

For the environments, there was no trends according to the year. The farms were not necessary in the same group, showing the importance of the year in their behaviour. For example  $ROW$ , for all years, was in the same group (2010, 2011 and 2012) whereas  $RAB$  had 2011 in one group, 2012 in another group and 2009 and 2010 in another group.

1 : alpha[144epi-C-JFB]	39 : alpha[C33]	77 : alpha[C7]
2 : alpha[144epi-P-JFB]	40 : alpha[C33#S-Isa]	78 : alpha[C74]
3 : alpha[21x3]	41 : alpha[C34]	79 : alpha[C76]
4 : alpha[Aréno]	42 : alpha[C34#BR]	80 : alpha[C77]
5 : alpha[Barbu-du-Maconnais]	43 : alpha[C37]	81 : alpha[C78]
6 : alpha[Bianc-des-Flandres]	44 : alpha[C37#epi-long-crosse]	82 : alpha[C8]
7 : alpha[Blé-Autrichien]	45 : alpha[C37#S-epis-crosses-S-JF]	83 : alpha[C80]
8 : alpha[C1]	46 : alpha[C38]	84 : alpha[C80#S-Isa-sJF]
9 : alpha[C12]	47 : alpha[C41]	85 : alpha[C80#S-JF]
10 : alpha[C13]	48 : alpha[C(4+18)]	86 : alpha[C81]
11 : alpha[C13#S-JF]	49 : alpha[C42]	87 : alpha[C82]
12 : alpha[C14]	50 : alpha[C42#SELECTION]	88 : alpha[C83]
13 : alpha[C14#Sa]	51 : alpha[C43]	89 : alpha[C83#SELECTION]
14 : alpha[C14#Sb]	52 : alpha[C44]	90 : alpha[C84]
15 : alpha[C15]	53 : alpha[C45]	91 : alpha[C85]
16 : alpha[C16]	54 : alpha[C46]	92 : alpha[C86]
17 : alpha[C16#SELECTION]	55 : alpha[C47]	93 : alpha[C87]
18 : alpha[C17]	56 : alpha[C48]	94 : alpha[C88]
19 : alpha[C19]	57 : alpha[C49]	95 : alpha[C89]
20 : alpha[C1#S]	58 : alpha[C4#nb]	96 : alpha[C89#MP]
21 : alpha[C2]	59 : alpha[C4#S]	97 : alpha[C90]
22 : alpha[C20]	60 : alpha[C4#S-ficelles-rouges]	98 : alpha[C92]
23 : alpha[C21]	61 : alpha[C5]	99 : alpha[C92#S]
24 : alpha[C21#S]	62 : alpha[C50]	100 : alpha[C93]
25 : alpha[C21#Sb]	63 : alpha[C52]	101 : alpha[Clomone]
26 : alpha[C21#Snb]	64 : alpha[C53]	102 : alpha[Ile-de-France]
27 : alpha[C23]	65 : alpha[C56]	103 : alpha[M(Atrium, Renan)]
28 : alpha[C24]	66 : alpha[C58]	104 : alpha[Mélange-croisements]
29 : alpha[C24#S]	67 : alpha[C58#S]	105 : alpha[Mélange-de-Loex]
30 : alpha[C25]	68 : alpha[C6]	106 : alpha[Mottin]
31 : alpha[C26]	69 : alpha[C60]	107 : alpha[Redon-Sixt-sur-Aff]
32 : alpha[C26#SELECTION]	70 : alpha[C60#S-JF]	108 : alpha[Renan]
33 : alpha[C28]	71 : alpha[C63]	109 : alpha[Rock-and-Lot]
34 : alpha[C29]	72 : alpha[C64]	110 : alpha[Rock-and-Lot#RLGK]
35 : alpha[C29#S]	73 : alpha[C65]	111 : alpha[Rock-and-Lot#Salvatorock]
36 : alpha[C31]	74 : alpha[C66]	112 : alpha[Rouge-du-Roc]
37 : alpha[C32]	75 : alpha[C67]	113 : alpha[Touzelle-Bergerie]
38 : alpha[C32#S]	76 : alpha[C69]	



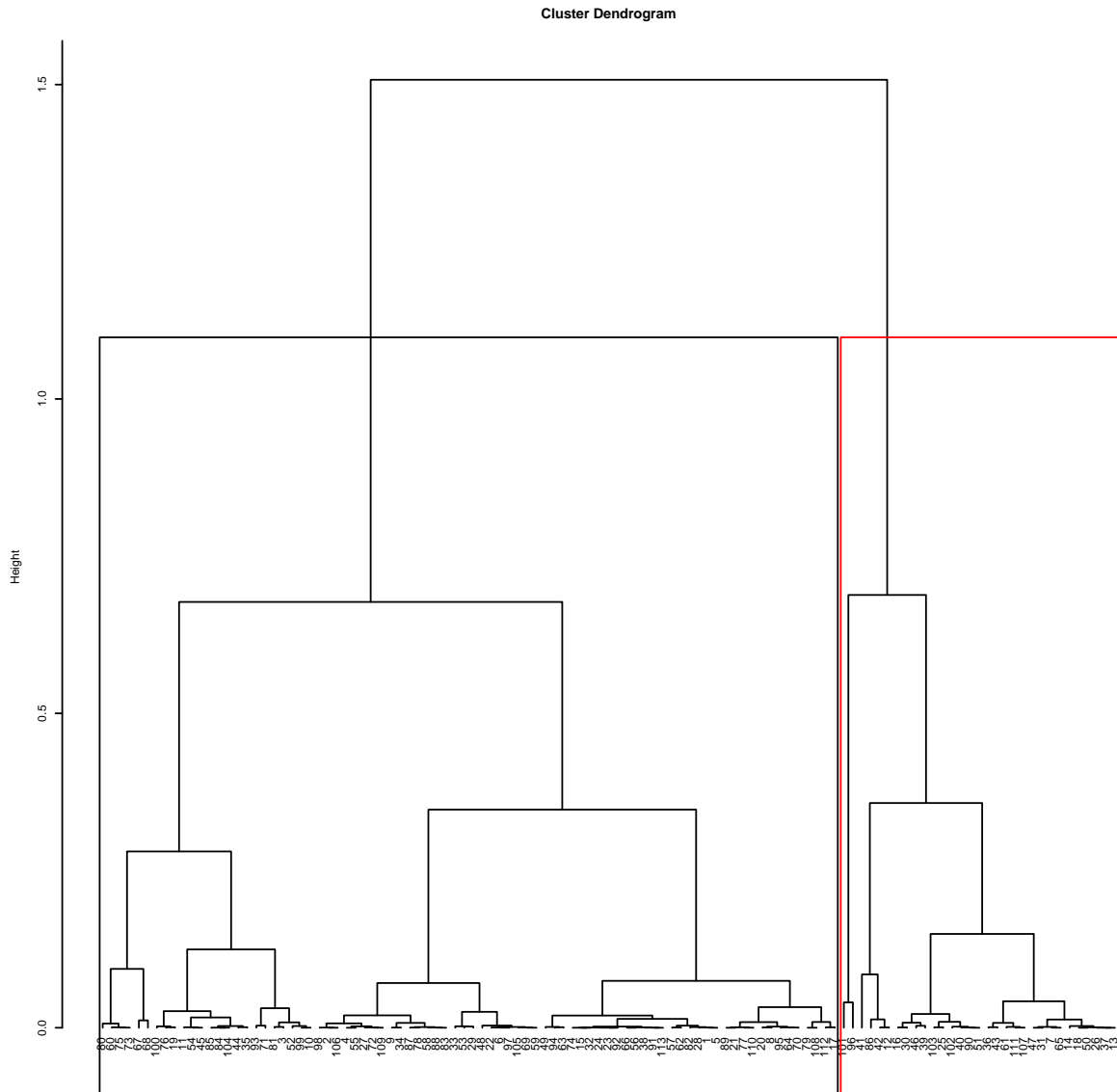


Figure 2 – Groups created for  $\alpha$ .

1 : beta[144epi-C-JFB]	39 : beta[C33]	77 : beta[C7]
2 : beta[144epi-P-JFB]	40 : beta[C33#S-Isa]	78 : beta[C74]
3 : beta[21x3]	41 : beta[C34]	79 : beta[C76]
4 : beta[Aréno]	42 : beta[C34#BR]	80 : beta[C77]
5 : beta[Barbu-du-Maconnais]	43 : beta[C37]	81 : beta[C78]
6 : beta[Blanc-des-Flandres]	44 : beta[C37#epi-long-crosse]	82 : beta[C8]
7 : beta[Blé-Autrichien]	45 : beta[C37#S-epis-crosses-S-JF]	83 : beta[C80]
8 : beta[C1]	46 : beta[C38]	84 : beta[C80#S-Isa-sJF]
9 : beta[C12]	47 : beta[C41]	85 : beta[C80#S-JF]
10 : beta[C13]	48 : beta[C(4+18)]	86 : beta[C81]
11 : beta[C13#S-JF]	49 : beta[C42]	87 : beta[C82]
12 : beta[C14]	50 : beta[C42#SELECTION]	88 : beta[C83]
13 : beta[C14#Sa]	51 : beta[C43]	89 : beta[C83#SELECTION]
14 : beta[C14#Sb]	52 : beta[C44]	90 : beta[C84]
15 : beta[C15]	53 : beta[C45]	91 : beta[C85]
16 : beta[C16]	54 : beta[C46]	92 : beta[C86]
17 : beta[C16#SELECTION]	55 : beta[C47]	93 : beta[C87]
18 : beta[C17]	56 : beta[C48]	94 : beta[C88]
19 : beta[C19]	57 : beta[C49]	95 : beta[C89]
20 : beta[C1#S]	58 : beta[C4#nb]	96 : beta[C89#MP]
21 : beta[C2]	59 : beta[C4#S]	97 : beta[C90]
22 : beta[C20]	60 : beta[C4#S-ficelles-rouges]	98 : beta[C92]
23 : beta[C21]	61 : beta[C5]	99 : beta[C92#S]
24 : beta[C21#S]	62 : beta[C50]	100 : beta[C93]
25 : beta[C21#Sb]	63 : beta[C52]	101 : beta[Clomone]
26 : beta[C21#Snb]	64 : beta[C53]	102 : beta[Ile-de-France]
27 : beta[C23]	65 : beta[C56]	103 : beta[M(Atrium, Renan)]
28 : beta[C24]	66 : beta[C58]	104 : beta[Mélange-croisements]
29 : beta[C24#S]	67 : beta[C58#S]	105 : beta[Mélange-de-Loex]
30 : beta[C25]	68 : beta[C6]	106 : beta[Mottin]
31 : beta[C26]	69 : beta[C60]	107 : beta[Redon-Sixt-sur-Aff]
32 : beta[C26#SELECTION]	70 : beta[C60#S-JF]	108 : beta[Renan]
33 : beta[C28]	71 : beta[C63]	109 : beta[Rock-and-Lot]
34 : beta[C29]	72 : beta[C64]	110 : beta[Rock-and-Lot#RLGK]
35 : beta[C29#S]	73 : beta[C65]	111 : beta[Rock-and-Lot#Salvatorock]
36 : beta[C31]	74 : beta[C66]	112 : beta[Rouge-du-Roc]
37 : beta[C32]	75 : beta[C67]	113 : beta[Touzelle-Bergerie]
38 : beta[C32#S]	76 : beta[C69]	

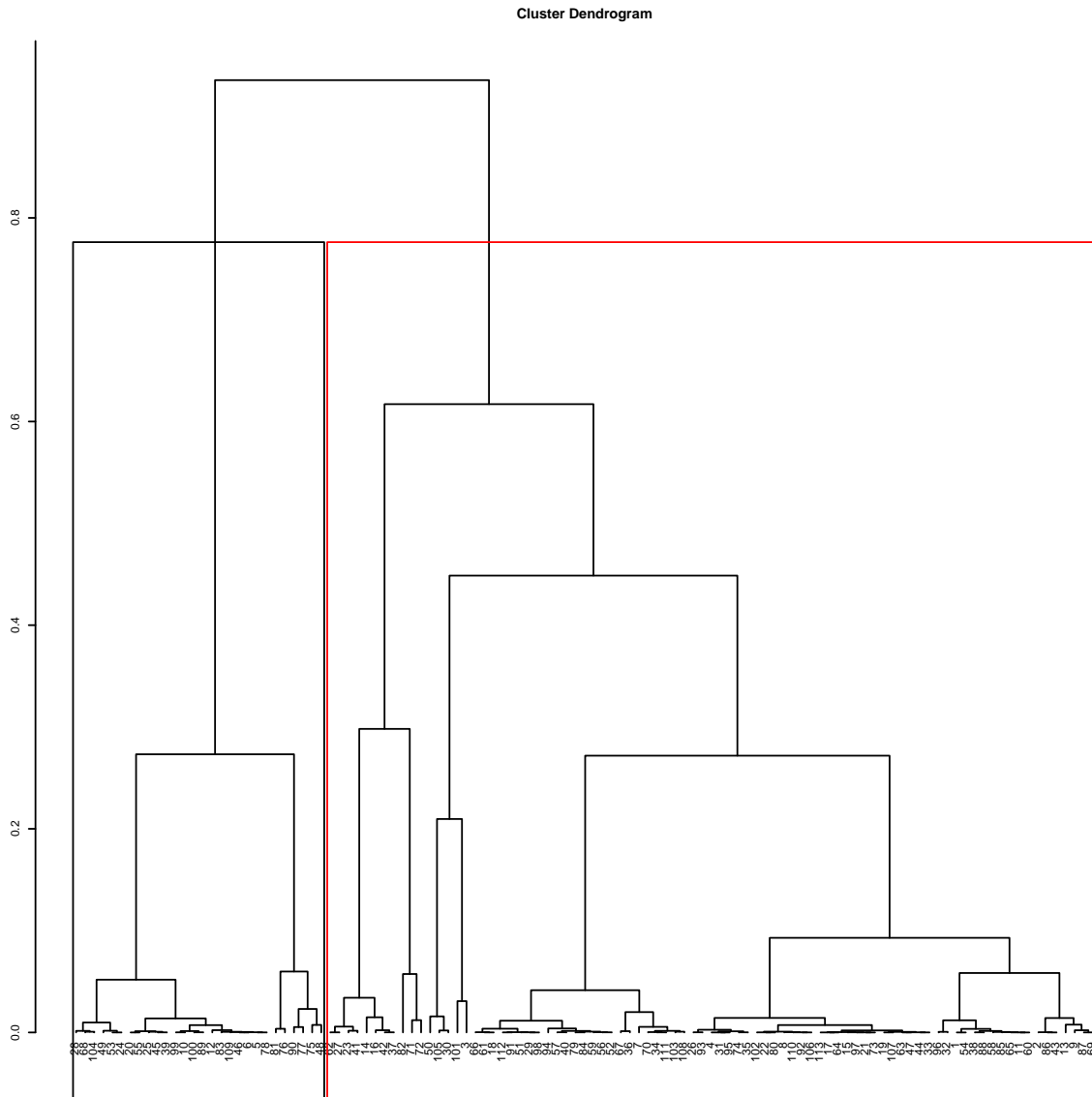


Figure 3 – Groups created for  $\beta$ .

1 : theta[ALP2011]	26 : theta[GIM2010]	51 : theta[MAV2012]
2 : theta[ANB2011]	27 : theta[GIM2012]	52 : theta[MPH2009]
3 : theta[ANB2012]	28 : theta[HEC2009]	53 : theta[MPH2010]
4 : theta[BER2009]	29 : theta[HEC2010]	54 : theta[MPH2012]
5 : theta[BER2010]	30 : theta[HEC2012]	55 : theta[OLM2010]
6 : theta[BRE2011]	31 : theta[JAB2009]	56 : theta[OLR2009]
7 : theta[CHD2009]	32 : theta[JAB2010]	57 : theta[OLR2010]
8 : theta[CHD2010]	33 : theta[JAB2011]	58 : theta[OLR2011]
9 : theta[CHD2011]	34 : theta[JFB2010]	59 : theta[OLR2012]
10 : theta[CHD2012]	35 : theta[JFB2011]	60 : theta[PAJ2012]
11 : theta[CHP2009]	36 : theta[JFB2012]	61 : theta[PHG2009]
12 : theta[CHP2012]	37 : theta[JJG2009]	62 : theta[PIS2010]
13 : theta[EUK2012]	38 : theta[JJG2010]	63 : theta[RAB2009]
14 : theta[FLM2009]	39 : theta[JMC2010]	64 : theta[RAB2010]
15 : theta[FLM2010]	40 : theta[JMC2011]	65 : theta[RAB2011]
16 : theta[FLM2011]	41 : theta[JMC2012]	66 : theta[RAB2012]
17 : theta[FLM2012]	42 : theta[JPB2009]	67 : theta[RAL2011]
18 : theta[FLP2009]	43 : theta[JSG2010]	68 : theta[RAL2012]
19 : theta[FLP2010]	44 : theta[JSG2011]	69 : theta[RIH2011]
20 : theta[FLP2011]	45 : theta[JSG2012]	70 : theta[RIH2012]
21 : theta[FLP2012]	46 : theta[JUB2012]	71 : theta[ROW2010]
22 : theta[FRC2009]	47 : theta[LUD2010]	72 : theta[ROW2011]
23 : theta[FRC2010]	48 : theta[LUD2011]	73 : theta[ROW2012]
24 : theta[FRC2011]	49 : theta[MAV2010]	74 : theta[STP2012]
25 : theta[FRC2012]	50 : theta[MAV2011]	75 : theta[VIC2009]

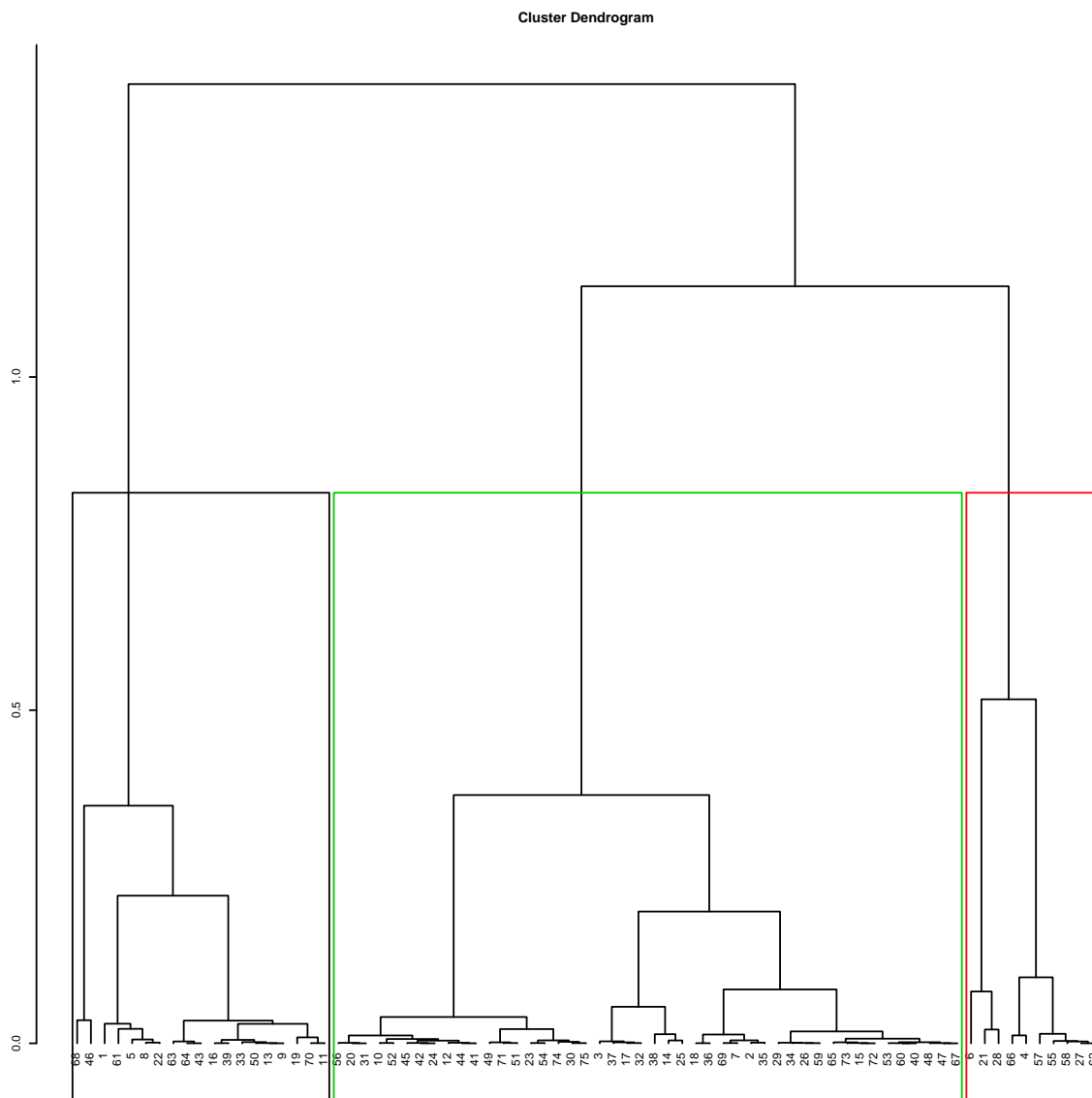


Figure 4 – Groups created for  $\theta$ .

1     **3.2 Comparison of the HFW model to the hierarchical**  
 2     **additive model and to the non hierarchical Finlay Wilkin-**  
 3     **son model**

4     **3.2.1 Hierarchical additive model versus HFW model**

5     For  $\epsilon_{ij}$ , the distribution did not change between the two models.

1            There were little changes in the distribution of mean and standard deviation for the  
2 estimations of  $\alpha$  and  $\theta$  between the hierarchical additive and the HFW model (Figure  
3 5 & 6). Yet these few changes led to differences in the number of groups detected for  
4  $\alpha$  and  $\theta$  (Table 3). There were no trends, the number of groups increased or decreased  
5 depending on variables.

6            Except for awns and nb of spikelets, the hierarchical model was more efficient based  
7 on DIC criterion (Table 4).

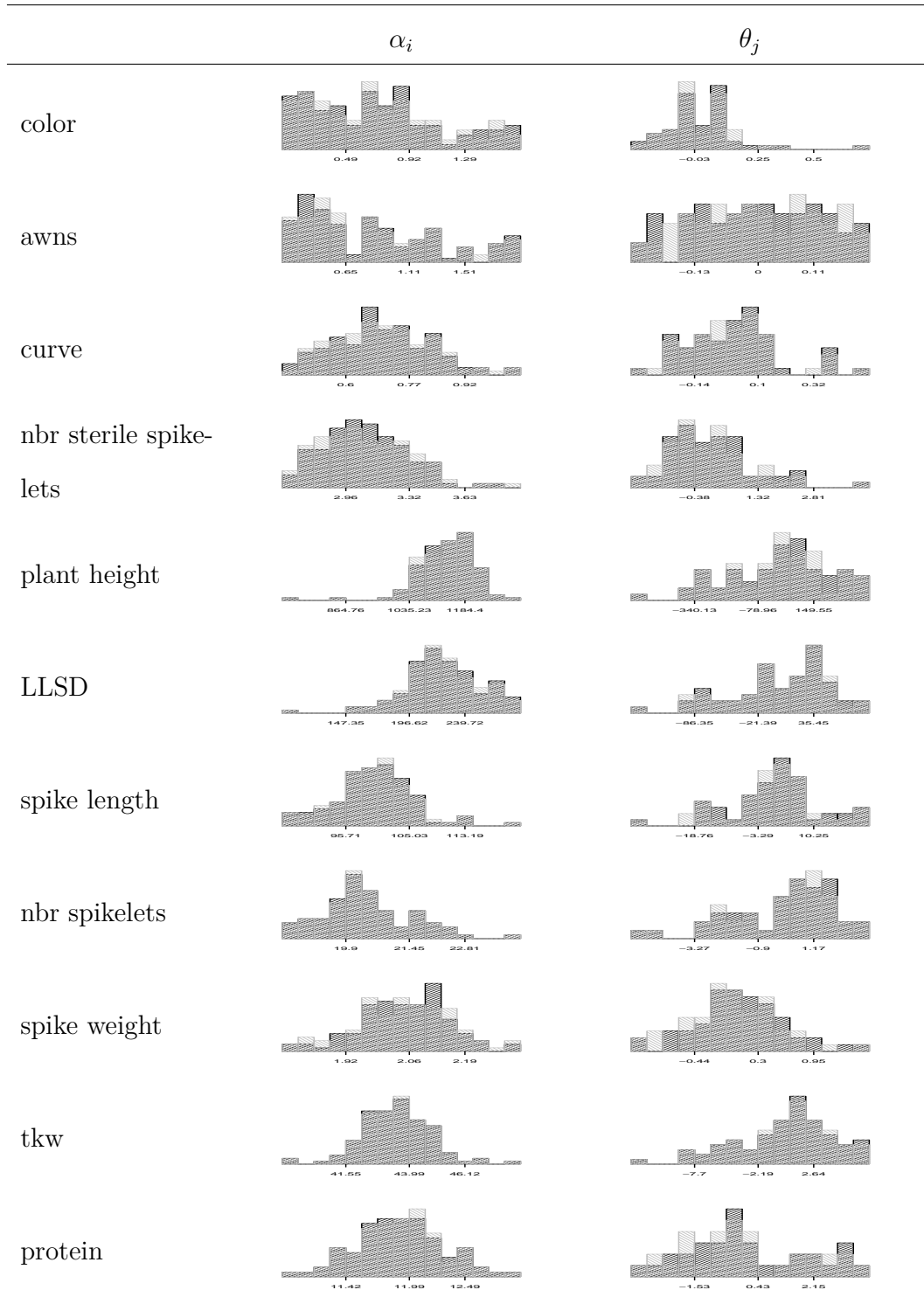


Figure 5 – Distributions of the mean for  $\alpha$  and  $\theta$  for each variable. black : HFW model, grey : hierarchical additive model.

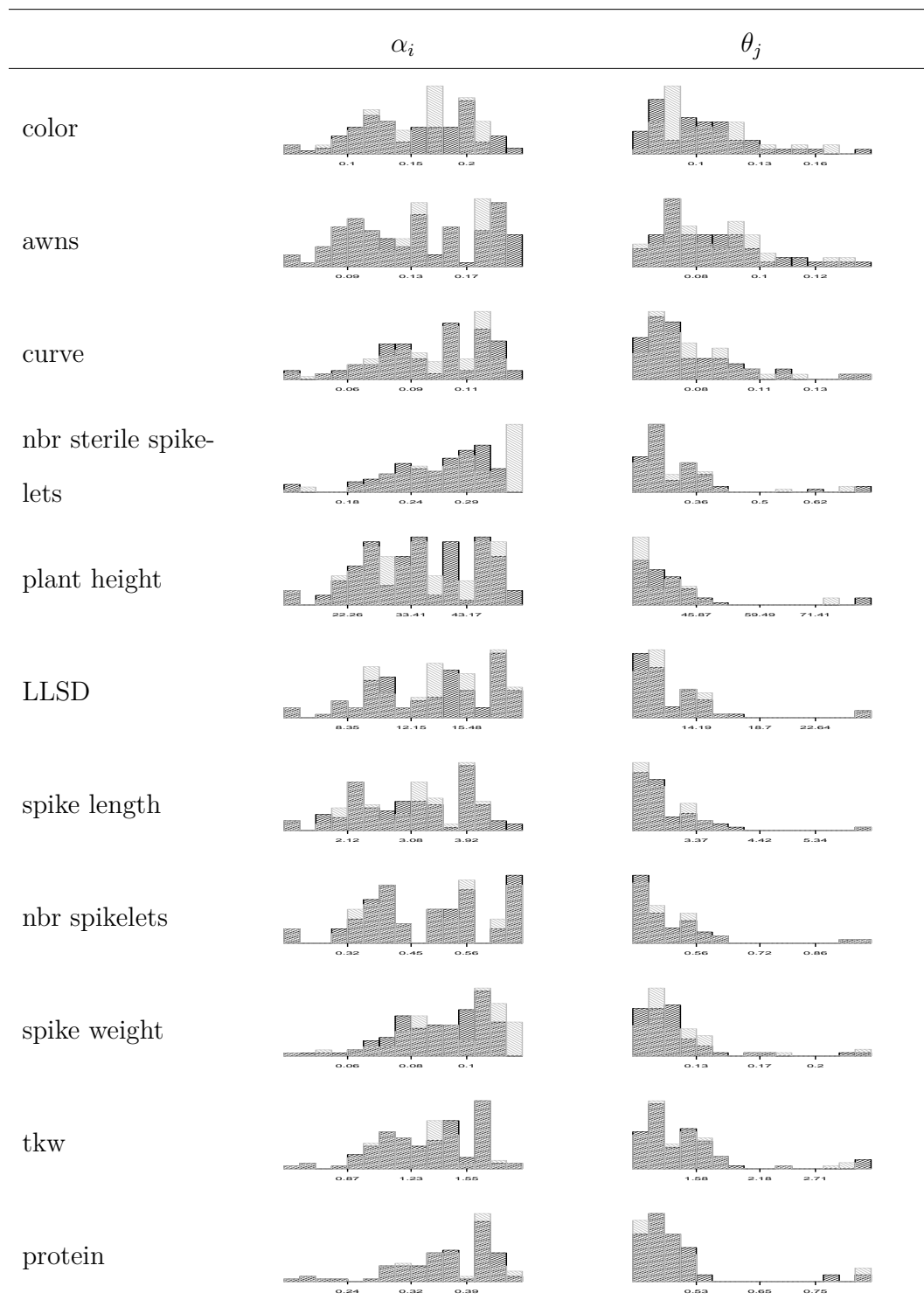


Figure 6 – Distributions of the standard deviation for  $\alpha$  and  $\theta$  for each variable. black : HFW model, grey : hierarchical additive model.



### 3.2.2 Non hierarchical FW model versus HFW model

For  $\epsilon_{ij}$ , the distribution did not change between the two models.

For all factors, the variance estimated by the hierarchical model was smaller compared to the non hierarchical model (Figure 8). The estimation of  $\alpha_i$  means, as well as the estimation of  $\theta_j$  means, did not change a lot in the hierarchical model compared to the non hierarchical model. On the other hand, the estimation of  $\beta_i$  means changed a lot with values closer to 1 in the hierarchical model compared to the non hierarchical model (Figure 7). The estimation of  $\beta_i$  with the non hierarchical model had some extremely high value when few germplasm were present, which was not the case with the hierarchical model (Figure 9). For example, for tkw, when the germplasms were present in two environments,  $\beta_i$  could have values of -1 and 2 for the non hierarchical model whereas it was 0.95 and 1.05 with the hierarchical model. This was less the case for  $\alpha_i$  and  $\theta_j$  (Figure 9). The information given by the hierarchical model shrank the estimation of the parameters leading to higher accuracy in parameter estimations. This was coherent with theory (Robert, 2001; Smyth, 2004). This led to a change in the assignation of groups for the parameters (Table 3).

Based on the DIC criterion, the hierarchical model was better than the non hierarchical model for only five variables out of 11 (color, spike length, nbr spikelets, tkw and protein). For curve and LLSD, the two model were similar (Table 4).

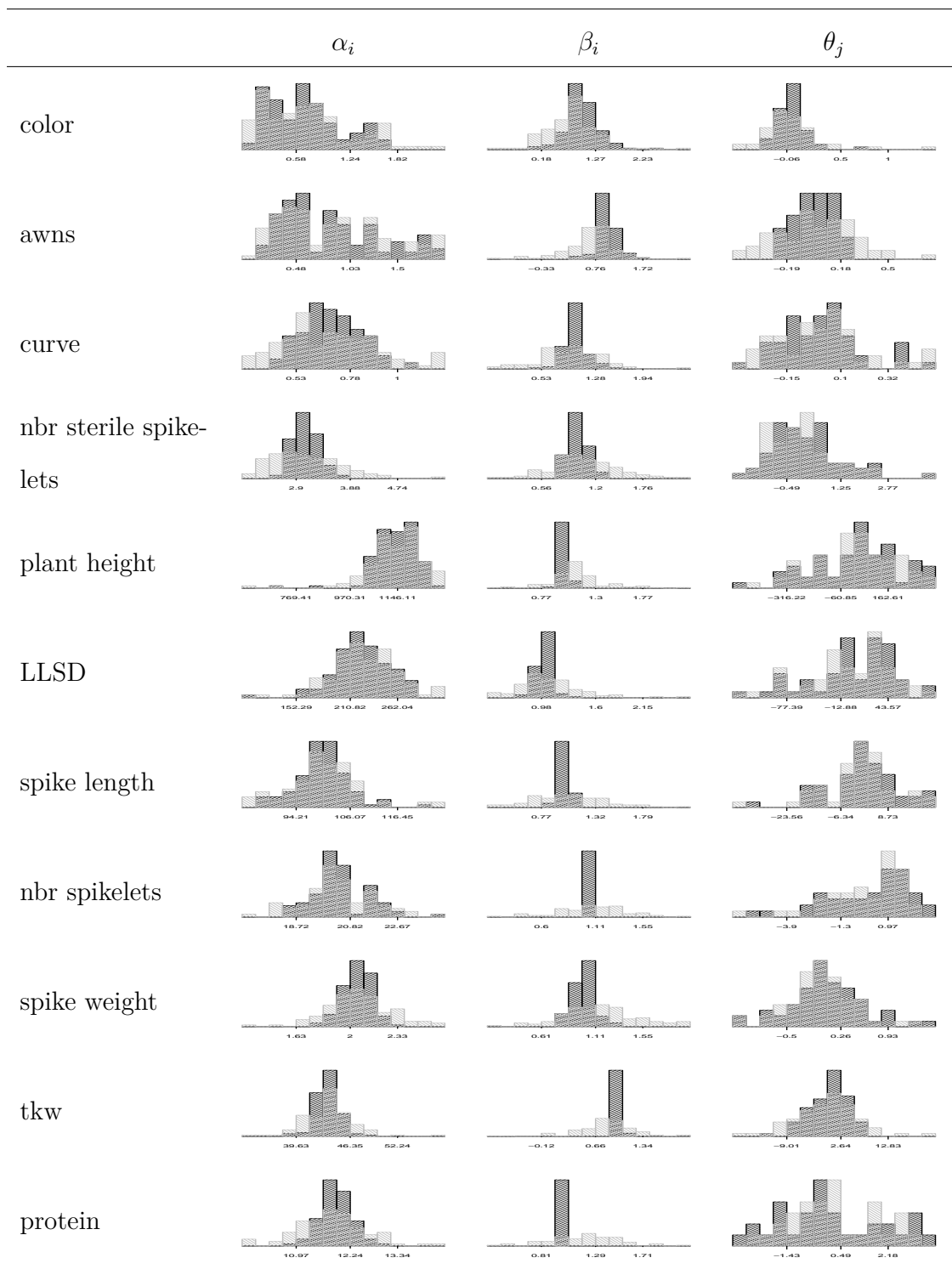


Figure 7 – Distributions of the mean for  $\alpha$ ,  $\beta$  and  $\theta$  for each variable. black : HFW model, grey : non hierarchical FW model.

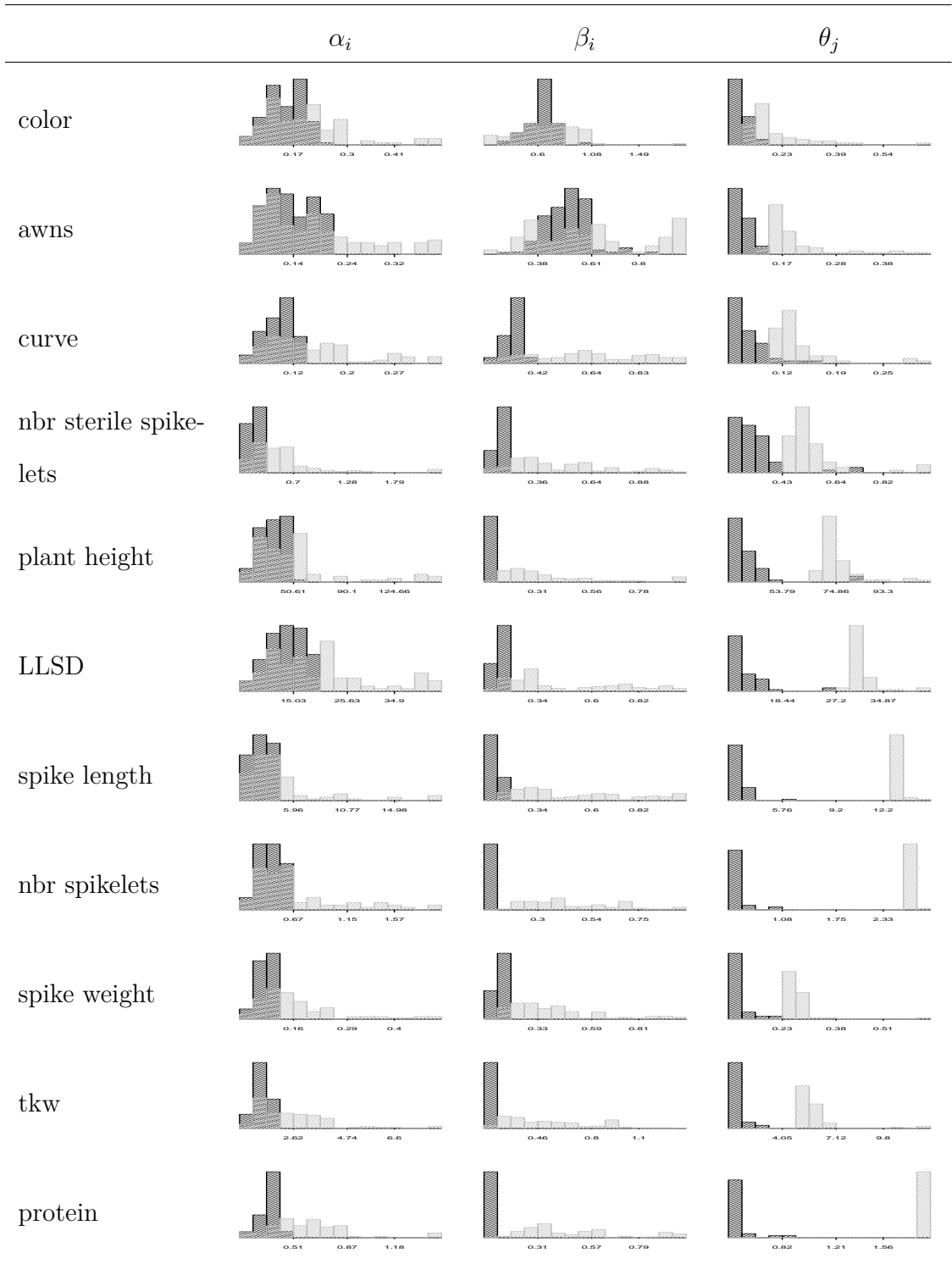


Figure 8 – Distributions of the the standard deviation for  $\alpha$ ,  $\beta$  and  $\theta$  for each variable. black : HFW model, grey : non hierarchical FW model.

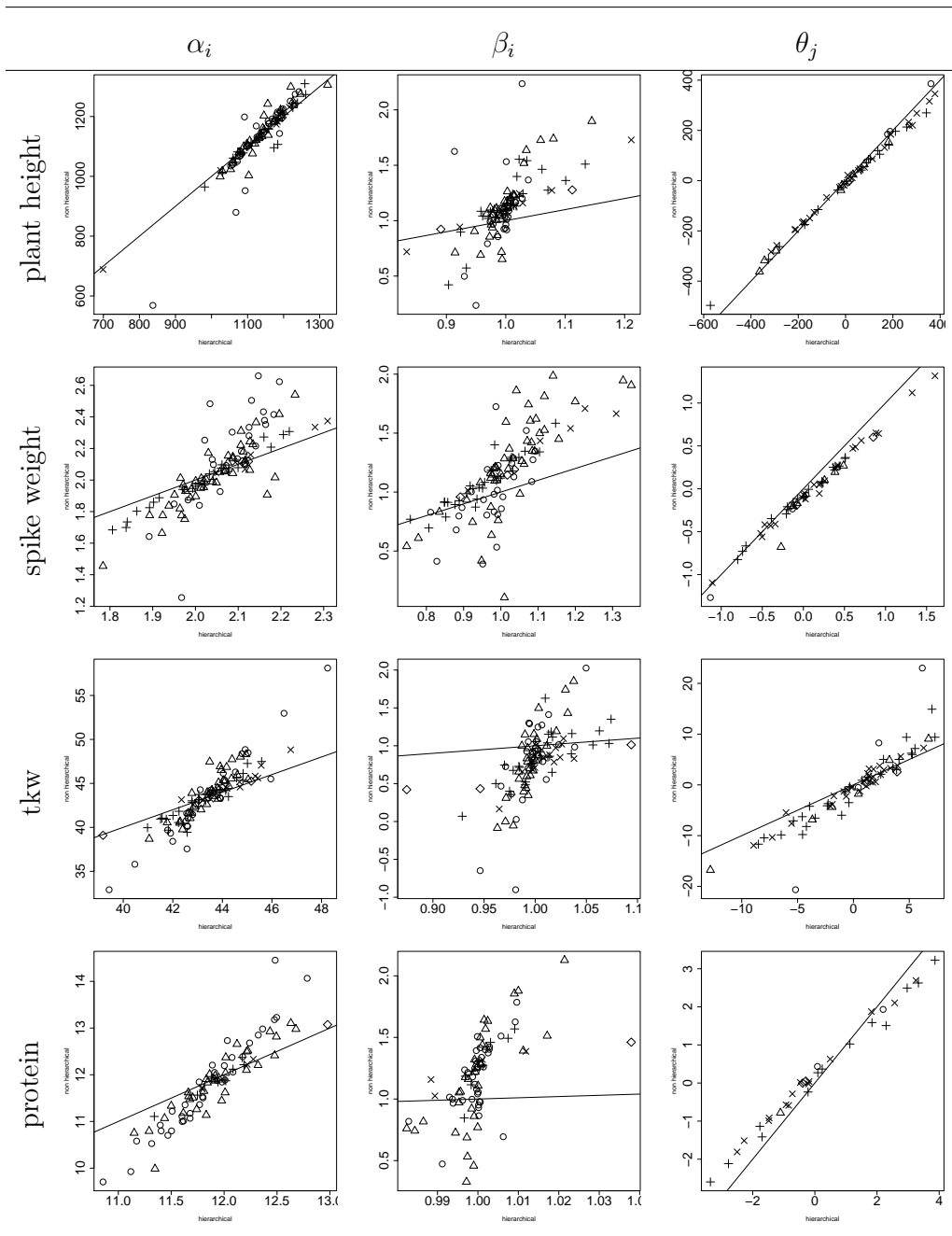


Figure 9 – Parameter estimations for  $\alpha$ ,  $\beta$  and  $\theta$ , with the non hierarchical FW model ( $y$ ) and the HFW model ( $x$ ) considering the entire data set ( $DATA-2$ ). Example for plant height, spike weight, tkw and protein. With  $\text{nbr\_germ}$ , the number of germplasm on the network, it comes :  $\circ$  :  $\text{nbr\_germ} = 2$ ;  $\triangle$  :  $2 < \text{nbr\_germ} \leq 5$ ;  $+$  :  $5 < \text{nbr\_germ} \leq 10$ ;  $\times$  :  $10 < \text{nbr\_germ} \leq 20$ ;  $\diamond$  :  $20 \leq \text{nbr\_germ}$ ;  $-$  :  $y = x$ .

variable	$\alpha$			$\beta$			$\theta$					
	nbr	gp h	gp ad	gp nh	nbr	gp h	gp ad	gp nh	nbr	gp h	gp ad	gp nh
color	106	17	15	5	106	1	NA	1	53	5	5	6
awns	112	19	21	8	112	1	NA	2	67	4	4	1
curve	106	4	6	2	106	1	NA	2	52	10	8	6
nbr sterile spikelets	103	2	4	1	103	1	NA	1	60	12	13	11
plant height	102	10	11	3	102	1	NA	1	60	18	23	20
LLSD	101	11	12	4	101	1	NA	1	52	18	19	20
spike length	100	10	11	5	100	1	NA	1	48	13	14	11
nbr spikelets	78	11	12	2	78	1	NA	1	43	12	11	12
spike weight	106	4	6	2	106	2	NA	1	56	14	14	4
tkw	108	4	5	2	108	1	NA	1	71	10	12	4
protein	84	2	3	5	84	1	NA	1	32	11	11	11

Table 3 – Number of significant groups for  $\alpha$ ,  $\beta$  and  $\theta$ , considering the whole data set. nbr : number of parameter ; gp : number of significant groups respectively for hierarchical (h) additive (ad) and non hierarchical (nh) model. NA : Non Available.

variable	<i>diffDIC</i> with additive	<i>diffDIC</i> with non hierarchical
color	-13.21	-0.41
awns	1.83	8.44
curve	-9.39	0.16
nbr sterile spikelets	-3.84	1.74
plant height	-30.26	2.17
LLSD	-4.55	0.00
spike length	-5.98	-1.89
nbr spikelets	2.04	-7.94
spike weight	-13.56	0.52
tkw	-6.69	-2.97
protein	-2.93	-2.01

Table 4 – Value of the DIC differences between the hierarchical model and the additive model (*diffDIC* with additive) and between the hierarchical model and the non hierarchical model (*diffDIC* with non hierarchical).

### 1      3.3    Impact of the increasing unbalanced data in the esti- 2      mation of the parameters

3            Whatever subsets and variables analysed, the estimation for all  $\alpha_i$ ,  $\beta_i$  and  $\theta_j$  never  
4            significantly changed except for one value of  $\theta_j$  for curve.

5            For all variables, the more the disequilibrium of the subset decreased (and the  
6            more the subset was small), the more the variance of  $\sigma_v$ ,  $\sigma_l$ ,  $\sigma_s$  and  $\sigma_e$  distributions  
7            increased. For example for plant height, there were a change in the distribution of  
8             $\sigma_v$ ,  $\sigma_l$ ,  $\sigma_s$  and  $\sigma_e$  respectively for  $DATA-11\_11 \times 59\_66.4$ ,  $DATA-10\_16 \times 59\_71.15$ ,  
9             $DATA-14\_6 \times 57\_54.4$  and  $DATA-8\_29 \times 60\_78.4$  (Figure 10).

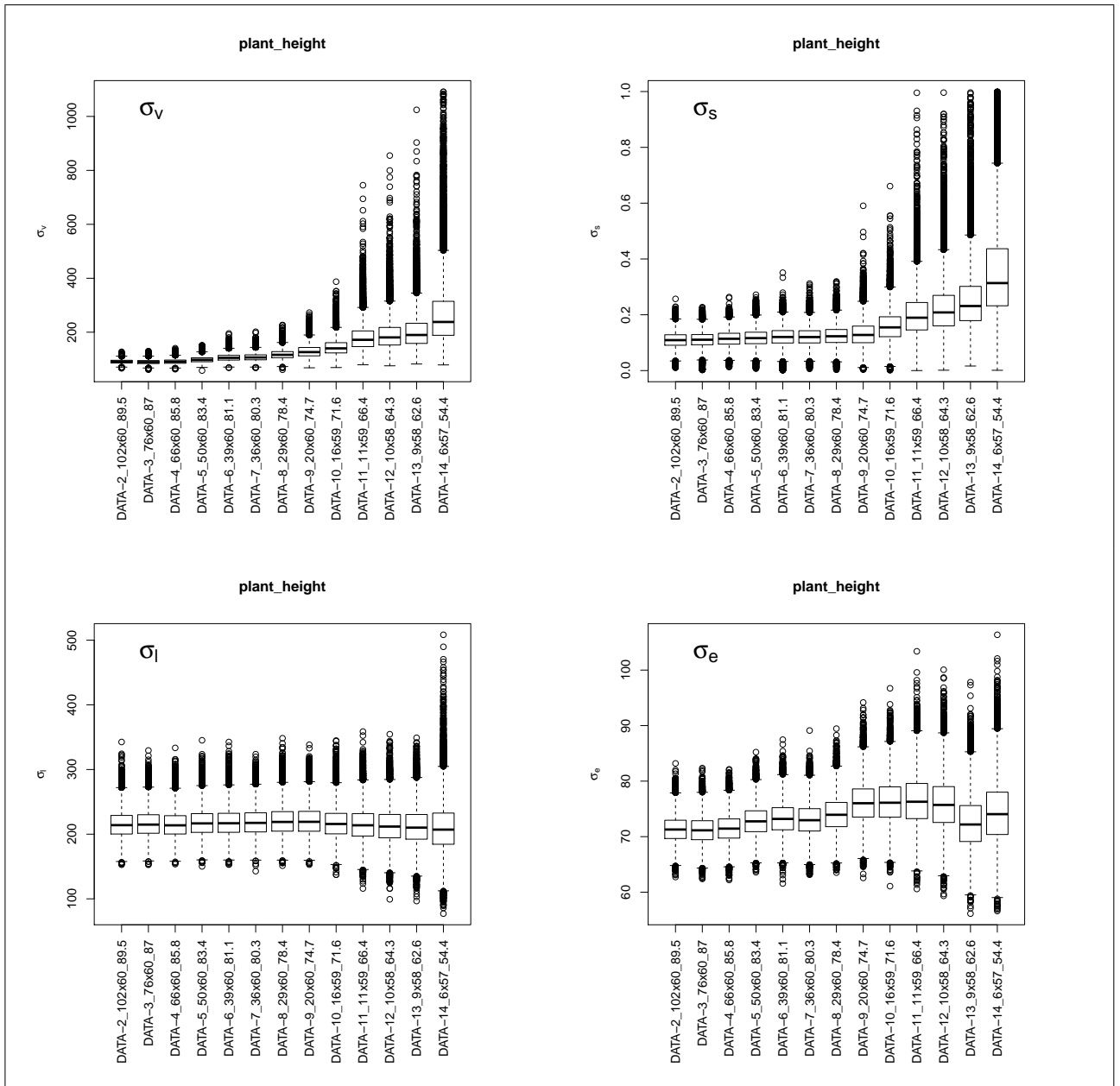


Figure 10 – Distributions of  $\sigma_v$ ,  $\sigma_s$ ,  $\sigma_l$  and  $\sigma_e$  for each subset for plant height. The name of the data refers to the disequilibrium and are explained in Table 2.

### 1 3.4 Cross validation

2 On the whole data set (i.e. *DATA-3*), for all variables except for curve variable, the  
 3 correlation between observed ( $Y_{ij}$ ) and estimated ( $\hat{Y}_{ij}$ ) values ranged from 0.49 for tkw  
 4 to 0.86 for plant height content (Figure 11). The correlation for curve was smaller than  
 5 the others with a value of 0.42.

- 1 Overall variables, between 56.8% and 67.5% of the  $Y_{ij}$  were well estimated with a  
 2 confidence of 95%. Considering the cross validation study overall the subset generated  
 3 (i.e. all the  $DATA-j$  with  $j \in [3, J]$ ), the variation of the fit was not due to the  
 4 disequilibrium of the subset but to the germplasms and the environments (Table 5).

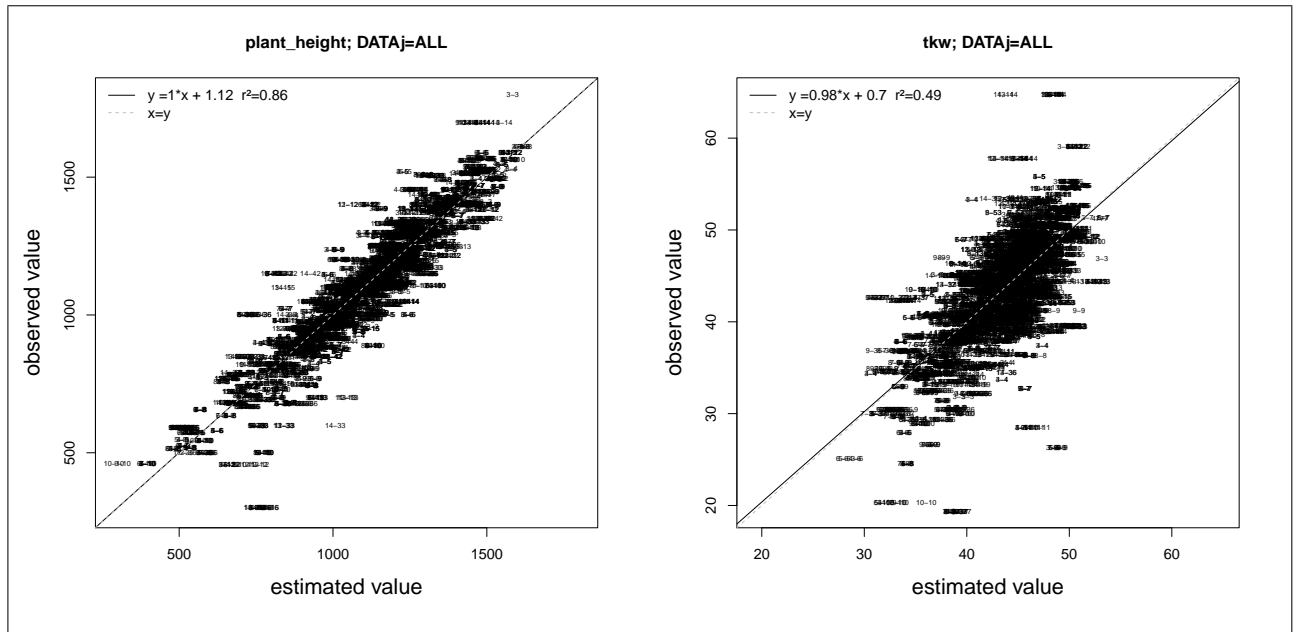


Figure 11 – Estimated values ( $\hat{Y}_{ij}$ ) with the HFW model compared to observed values ( $Y_{ij}$ ) for plant height (left) and tkw (right).



variable	$r^2$ between observed and predicted value	$\% Pr\{\hat{Y}_{ij} = Y_{ij}\} > 0.05$	$DATA_j$	$geno$	$env$
color	0.69	61.6	0.925	0	0
awns	0.77	67.6	0.102	0	0
curve	0.42	59.8	1	0	0
nbr sterile spikelets	0.68	59.4	0.169	0	0
plant height	0.86	67.5	0	0	0
LLSD	0.77	59.4	0.915	0	0
spike length	0.7	61.4	0.184	0	0
nbr spikelets	0.76	61.8	0.123	0	0
spike weight	0.76	58.1	0.048	0	0
tkw	0.49	58.2	0.531	0	0
protein	0.8	56.8	0.419	0	0

Table 5 – Summary of the cross validation for each variable.  $r^2$  : correlation between observed and predicted value for the entire data set (i.e.  $DATA - 3$ );  $\% Pr\{\hat{Y}_{ij} = Y_{ij}\} > 0.05$  : percentage of observed values that were well estimated with a confidence of 95% for the entire data set ;  $DATA_j$ ,  $DATA_jDk$ ,  $geno$  and  $env$  : p-value of the ANOVA.

## 4 Discussion

### 4.1 Modelisation of $G \times E$ interaction for unbalanced dataset

Edwards and Jannink (2006) and Cotes et al. (2006) proposed a hierarchical bayesian model to study the interaction. They did not deal with high disequilibrium data and modelled individual variance of interaction for all germplasm  $\times$  environment combinations. This leads to a better estimation of the parameters. This kind of modelisation is also found with the AMMI and GGE models, which are very popular to study the interaction. AMMI analysis allows for accurate study of the  $G \times E$  structure and GGE analysis considers the structure of genotype effect (G) together with interaction (GE) (Yan et al., 2007; Gauch et al., 2008). Those models are resolved with frequentist (Annicchiarico, 1997; Gauch et al., 2008; Yan et al., 2007) or bayesian approaches (Viele and Srinivasan, 2000; Crossa et al., 2011).

With 95% of missing combinaison, with such models, we might obtain a sufficient accuracy only on certain  $G \times E$  combinaison but could not estimate general pattern. With the HFW model, we use all the information available to assess sensitivity to interaction for each germplasm as well as general germplasm and environment effects. Sensitivity is definitely not the best parameter to study in details the interaction. Nevertheless, this model is more parcimonious to deal with the disequilibrium as it estimates a global sensitivity to interaction for each germplasm.

Restricted maximum likelihood (REML) has been used with very unbalanced PPB trials but only assessed general effects assuming that the interaction were negligible (Virk and Witcombe, 2007). The bayesian approach can overcome some issues that are met with REML such as negative variance estimates, the difficulty of having confidence intervals, the non normality of the data or the heterogeneity of within environment error variances (Theobald et al., 2002; Cotes et al., 2006; Edwards and Jannink, 2006). While bayesian approaches have been reported to be efficient for unbalanced data analysis (Theobald et al., 2002; Edwards and Jannink, 2006), it has never been tested on high unbalanced dataset such as we handle in the wheat PPB project. With the HFW model,

1 we assumed that the extreme disequilibrium of the data can be compensated by the  
2 huge information available in the data set. The common controls and the germplasm  
3 present in higher frequency on the network allow for the estimation of the different  
4 effects.

5  
6 We decided to estimate the variance of the factor distributions. This is coherent from  
7 a agronomic point of view as the germplasm used on the network came from a large  
8 diversity as well as the farms (?). From a statistical point of view, estimate the variance  
9 of the distribution leads to a better estimation of the parameters. The estimation of  
10 the variance decreases the unexplained variance leading to a shrinkage in parameters  
11 estimation.

12 [Cotes et al. \(2006\)](#) used a bayesian approach to assess the sensitivity of the germ-  
13 plasms in a Multi-Environment Trials (MET). The objective was to add *a priori* in-  
14 formation on the germplasms. The prior information came from germplasms already  
15 evaluated on other MET. A similar approach was used by [Foucteau and Denis \(2001\)](#)  
16 which was found to grealty improve the results. Here, we used a vague prior in the mo-  
17 del as we did not have any information *a priori*. The project is recent and evaluation  
18 studies will bring extra information on the behaviour of the germplasm that might be  
19 implemeted in prior distributions.

20  
21 Another way to improve the model could be to add covariates based on the pedigree  
22 of the germplasm. Such information could increase the prediction efficiency of our model  
23 ([Burgueño et al., 2012](#)). However, the pedigree of the germplasm was rather complex to  
24 assess for heterogenous population derived from crosses. Moreover, the differentiation  
25 of population on farms induce bias in the construction of a pedigree matrix. This  
26 differentiation is due to local adaptation linked to the heritability of the trait. Another  
27 way to assess pedigree would be to use molecular markers ([Burgueño et al., 2012](#)).  
28 But this rises new questions as the populations are heterogeneous and differentiated.  
29 Multidimensional analysis followed by Kmeans or DAPC analysis may be useful to find  
30 similarites between germplasm to create a pseudo kinship matrix ([Dawson et al., 2012](#);

1 [Jombart et al., 2010](#); [Husson et al., 2010](#)).

## 2 **4.2 Comparision of the HFW model to the hierarchical** 3 **additive model and to the non hierarchical Finlay Wilkin-** 4 **son model**

5 The number of significant groups detected changed between the hierarchical additive  
6 and the HFW model showing the importance of choosing one model or another. The DIC  
7 criterion gave us clue than the HFW model seems better for all variables except awns  
8 and nbr spikelets. The modelisation of interaction brought more information than the  
9 additive model. This was not obvious at first sight considering the high disequilibrium  
10 of the data. The few  $G \times E$  combinaisons available were sufficient to better explain the  
11 data.

12 The HFW model led to a shrinkage in the variance for all the parameters compared  
13 to the non hierarchical FW model. The information brought by the data constrained  
14 the possible value of parameters leading to higher accuracy in the prediction. On  
15 the other hand, the germplasm and environment mean effects did not change drastically  
16 between the HFW model and the non hierarchical FW model. No changes in the means  
17 but changes in the variances explained that for all variables (except protein), the HFW  
18 model allow to detect more groups for  $\alpha$  than the non hierarchical FW model. This  
19 was also observed for  $\theta$  for six variables, two did not changed and three showed less  
20 groups with the HFW model. The number of environments being less than the number  
21 of germplasm, there were less parameters available to estimate the  $\theta$  ditribution, which  
22 could lead to less differences with the non hierarchical model.

23 For  $\beta$ , only one group was created for both the HFW model (except for spike weight)  
24 and the non hierarchical FW model. There were differences in the estimation of the  $\beta$   
25 between the two models but it was not enough to change the number of groups. This  
26 could be explained by the distribution of the  $\beta_i$  which were tight around 1 for both  
27 models. The low number of  $G \times E$  combinaisons could explain those distributions.  
28 There were not enough interaction terms to get enough informations. Nevertheless,

1 according to the DIC criterion, modelisation of  $\beta$  increase the accuracy of the model  
2 for five variables out of 11, for two variables it was similar (Table 4).

3 The shrinkage of the parameter estimation due to the hierarchical approach is well  
4 illustrated by  $\beta$  (for details on shrinkage in hierarchical model, see discussion in [Edwards](#)  
5 [and Jannink \(2006\)](#)). The estimation of  $\beta$  change with the hierarchical approach for  
6 germplasm which were not present on many environments (Figure 9). This was expected  
7 as the estimation of the slope (i.e.  $\beta$ ) would be less accurate with less germplasm. The  
8 hierarchical approach allowed for reasonable values of  $\beta_i$  that did not reach extrem  
9 values as it was the case in the non hierarchical model. This was illustrated with tkw  
10 (Figure 9).

### 11 4.3 Impact of the increasing unbalanced data in the esti- 12 mation of the parameters

13 For all factors  $\alpha$ ,  $\beta$  and  $\theta$ ; the estimation of the parameters did not significantly  
14 change whatever subset was used. For example  $\alpha_1$  in *DATA-2* and  $\alpha_1$  in *DATA-7* were  
15 equal. No significant changes in the estimation of parameters was consistent with no  
16 significant changes in the distribution of the variances of the respective parameter  
17 distributions  $\sigma_v$ ,  $\sigma_s$ ,  $\sigma_l$  (except for plant height). Still, impact of the disequilibrium  
18 could be observed on the distribution of the variance. The more the disequilibrium  
19 of the subset decreased (i.e. the more the subset was small), the more the variance  
20 distributions increased. A lack of accuracy in the distribution can be due to statistical  
21 reasons : the number of parameters was not big enough or the data available was not  
22 big enough. Therefore, in order to get more precise distributions, it seems better to  
23 have big unbalanced data set instead of small balanced data set.

### 24 4.4 Prediction of the hierarchical model

25 [Burgueño et al. \(2012\)](#) pointed out that two predictions are relevant in plant breeding  
26 : (i) predicting the performance of untested genotypes and (ii) predicting the  
27 performance of genotypes that were not evaluated on some farms.

1 Here we are interested in the second objective. To evaluate if our model was predi-  
2 cive, we ran cross validation studies. We run a complete cross validation studies as all  
3 the individuals were predicted by all the other individuals. The partitioning of training  
4 and testing populations affect the accuracy of the prediction With our approach, the  
5 training population can not be bigger. Therefore, we avoid any correlation between  
6 individual in the training or the testing population due to sampling method that could  
7 induce biais in the results (Daetwyler et al., 2013).

8 The prediction ( $r^2$ ) between predicted and observed values ranged between 0.42 for  
9 curve (a semi quantitative variable) to 0.86 for plant height. These correlations were of  
10 a similar level to the one found with genomic selection ( $r^2$  between 0.58 and 0.9 (Zhao  
11 et al., 2012)) but is better than genomic selection dealing with  $G \times E$  interactions (for  
12 example,  $r^2$  ranged from 0.15 to 0.47 (Ly et al., 2013) or from 0.3 to 0.65 (Burgueño  
13 et al., 2012)).

14  $r^2$  of protein content was 0.80. Protein content is known to be dependant on environ-  
15 mental factor (primarily differences in nitrogen level) and interaction with genotypes  
16 (Blum, 1998). Nitrogen levels are very heterogeneous in OA as it is mainly based on  
17 mineralisation from manure or crop residues (Messmer et al., 2012). Such level of pre-  
18 diction is therefore surprising. tkw, which is also a very integrative factor, show a  $r^2$  of  
19 0.49 which is coherent with other study (Zhao et al., 2012).

20 We could use the current data as a training population as it is done in genomic  
21 selection (Ly et al., 2013; Burgueño et al., 2012). The training population brings the  
22 value for  $\alpha$  and  $\beta$  for each germplasm and  $\theta$  for each environment. It is then possible  
23 to predict the behaviour of a germplasm on a given environment. We get encouraging  
24 results as between 56.8% and 67.5% of the  $Y_{ij}$  were well estimated with a confidence  
25 of 95% overall the variables measured. This is dependant on the training population  
26 (Daetwyler et al., 2013). The more there will be data, the more accurate would be the  
27 prediction. As the wheat PPB project continues and an increasing number of farmers is  
28 participating in the project, the data set will increase in size and therefore the training  
29 population will be more robust for prediction.

## 1      4.5    Use of the results in the wheat PPB project metho- 2      dology

3            In the wheat PPB programm, the evaluation units consist in heterogenous popula-  
4            tion. Thus, these populations may vary in their genotype compositions according to the  
5            generation since the crossing, the environment and the farmers' practices. In this analy-  
6            sis, such variations were deliberately considered as part of the response of a population  
7            to a given environment. Therefore the  $G \times E$  interaction assessed was overestimated.  
8            Differentiation of a germplasm on a given environment increased the interaction and  
9            thus a large  $G \times E$  might indicate a strong differentiation and local adaptation. Yet,  
10           no trends, based of germplasm family or selection within germplasm, were found in the  
11           behaviour of germplasm sensitivity

12           Using a model with a year effect would be more efficient and more informative  
13           (Piepho, 1998; Theobald et al., 2002). We choose to work on environments that were a  
14           combination of a farm and a year. This is of interest for farmers to have a summarised  
15           information. A germplasm with a low sensitivity to interaction is potentially stable  
16           over locations and years. Moreover, considering the disequilibrium of the data, we gave  
17           priority to a strong data set with many environments rather than a more unbalanced  
18           dataset by separating years and farms. Based on the  $G$  matrix, the farms were not  
19           necessary in the same group, showing the importance of the year in their behaviour.

20           Except for regional farms, where correction for block effects was applied, the model  
21           did not take into account the variability within site. Yet, the prediction of the model  
22           might be increased by taking into account the within site variability as suggested by  
23           some authors (Cotes et al., 2006; Edwards and Jannink, 2006), however others argue it  
24           is not of importance (Bhatt et al., 1984).

25           This work is part of the methodology we set up in the wheat PPB program. The  
26           project is based on collaboration between farmers, NGOs and researchers to reach  
27           the objectives (Dawson et al., 2011). This collaboration led to original experimental  
28           designs on farm. The farmers were free to grow the population they want based on  
29           their selection : this is part of the flexibility of the methodology (?). The advantage  
30           is to test for a wide range of genetic diversity in a wide range of environments. There

1 were 75 environments resulting from a combination of years and farms. While this  
2 was rather limited compared to CIMMYT's MET which included between 500 and  
3 800 germplasm on 12 MET since 1945 and 1986 (Braun et al., 1996); the number of  
4 germplasm evaluated/selected was important compared to other studies. For example  
5 the number of germplasm, location and year were respectively (NA stands for Non  
6 Available) : 10, 5 and NA ; 10, 5 and 2 ; 9, 20 and NA (Cotes et al., 2006) ; 600, 4 and  
7 NA (Burgueño et al., 2012) ; 16, 16 and 1 (Gauch, 1997) ; 11, 11 and 3 ; 18, 31 and 3 ;  
8 9, 6 and 3 ; 10, 7 and 3 (Annicchiarico, 1997) ; 7, 10 and 2 ; 13, 10 and 1 ; 11, 10 and  
9 1 (Annicchiarico et al., 2010) ; 85, 5 and 7 (Edwards and Jannink, 2006) ; 24, 47 and  
10 3 (Annicchiarico et al., 2005). The method is based on extreme decentralisation of the  
11 selection and  $G \times E$  analysis is a step toward selection on farm.

12 One aim of the project was to bring farmers information on the germplasms that  
13 were evaluated/selected on the network. This information will allow farmers to select  
14 new germplasms to test as new candidates on their farms. This study provides statistical  
15 tools to deal with this inedit experimental design in order to bring results for the crea-  
16 tion of population varieties adapted to specific pedo-climatic conditions and farmers'  
17 practices. The objectives of the analysis are the same than in a MET analysis : (i) esti-  
18 mate and predict traits of interest for breeding, (ii) study stability of germplasms over  
19 several environments, (iii) select new germplasms to be sown in new locations (Cotes  
20 et al., 2006). Usually, results of MET analysis can be used with two philosophies : find  
21 stable germplasm that performs well in average on many locations or detect special lo-  
22 cal adaptation to certain environments (Annicchiarico et al., 2005; Gauch, 2006). Here,  
23 the two approches were used.

24 Based on the results, there were three options for the farmers to choose new germ-  
25 plasms as candidates on their farms.

- 26 1. The first option is to choose a germplasm without *a priori*. Germplasms that were  
27 less sensible to interaction ( $\beta$  close to 1) and had an extreme effect (high or low  
28  $\alpha$ ) for traits of interest can be choosen.
- 29 2. The second option is to find farms that share the same group than the farmers'  
30 own farm. This could be done by taking into account all the effects for all the



1 variables ( $G$  matrix). Only three groups of environments were detected with the  
2 method used. Nevertheless we can adjust the groups with a "personal" criterion  
3 based on the dendrogram. Once a farm is chosen, the farmers can look at  
4 germplasm grown on the farm and choose new populations. This can be done  
5 based on mean comparisons (?). In addition, results in percentage of the common  
6 control *Rouge-du-Roc*, may add extra information for the farmers.

- 7 3. The third option is to predict the behaviour that a germplasm would have in  
8 a given environment based on the HFW model. For each variable, the five best  
9 germplasms could be proposed to each farmer.

10 Following these three options, farmers can evaluate new germplasm and incorporate  
11 it in the selection process on their farm.

## 12 5 Conclusion

13 The hierarchical model seems a good candidate to estimate the parameters of the  
14 Finlay Wilkinson model despite the high inbalance of our data set. Germplasm, envi-  
15 ronment and sensitivity estimated effects allow farmers to use the information based  
16 on the network in the selection process on farm.

17 The use of molecular markers may improve the reliability of the results by asses-  
18 sing pedigree matrix and genomic prediction. Nevertheless, molecular markers may be  
19 difficult to implement on farm.

20 Simulation studies would also be of interest to assess the maximum disequilibrium  
21 on the network that allows for confident parameter estimations.

22 Participatory research addresses new questions to research and develop innovative  
23 methods in order to bring societal action (Kastenhofer et al., 2011). All the methodology  
24 is based on collective and collaborative work and action between farmers, associations of  
25 farmers and researchers (Ali Brac de la Perrière et al., 2011). New statistical methods  
26 can contribute to make the best use of such complex multi-environment data in the  
27 selection process.

## 1      **Aknowledgment**

2            We thank Nathalie Galic and Sophie Pin for their great help in measuring all the  
3      populations all around France. We thank all the farmers participating in the project.  
4      We thank Franch Gauthier for his technical help coding script for HTcondor and his  
5      efficient debug-condor hotline. We thank Gerard Branlard from INRA Clermont Fer-  
6      rand for his time when doing NIRS analysis. P. Rivière was funded by a grant from  
7      DIM ASTREA (Ile de France region). The research leading to these results has recei-  
8      ved funding from the European Community's Seventh Framework Programme (FP7/  
9      2007–2013) under the Grant Agreement n245058-Solibam (Strategies for Organic and  
10     Low-input Integrated Breeding and Management).

## References

- 1           **References**
- 2           Agence-Bio, 2011. L'agriculture bio en france. Tech. rep.
- 3           Ali Brac de la Perrière R., de Kochko P., Neubauer C., Storup B., 2011. Visions  
4           paysannes de la recherche dans le contexte de la sélection participative : comment  
5           co-construire et mutualiser les connaissances sur les plantes ? Pour l'Emergence d'une  
6           Université du Vivant.
- 7           Annicchiarico P., 1997. Additive main effects and multiplicative interaction (AMMI)  
8           analysis of genotype-location interaction in variety trials repeated over years. Theo-  
9           retical and Applied Genetics 94(8), 1072–1077.
- 10          Annicchiarico P., 2002. Genotype x environment interactions. Challenges and oppor-  
11          tunities for plant breeding and cultivar recommendations. FAO. FAO, Rome, Italy.
- 12          Annicchiarico P., Bellah F., Chiari T., 2005. Defining subregions and estimating benefits  
13          for a specific-adaptation strategy by breeding programs. Crop Science 45(5), 1741.
- 14          Annicchiarico P., Chiapparino E., Perenzin M., 2010. Response of common wheat  
15          varieties to organic and conventional production systems across italian locations,  
16          and implications for selection. Field Crops Research 116(3), 230–238.
- 17          Baldi P., Long A.D., 2001. A bayesian framework for the analysis of microarray expres-  
18          sion data : regularized t -test and statistical inferences of gene changes. Bioinformatics  
19          17(6), 509–519.
- 20          Bhatt G., Ellison F., Marshall D., 1984. A case for unreplicated plots for multi-site  
21          yield testing in wheat. Australian Journal of Agricultural Research 35(2), 107–114.
- 22          Blum A., 1998. Improving wheat grain filling under stress by stem reserve mobilisation.  
23          Euphytica 100(1-3), 77–83.
- 24          Bonneuil C., Demeulenaere E., 2007. Vers une génétique de pair à pair ? l'émergence  
25          de la sélection participative. In : Des sciences citoyennes ? La question de l'amateur

- 1 dans les sciences naturalistes, F Charvolin, A Micoud, L K Nyhart. L'Aude edn., pp.  
2 122–147.
- 3 Braun H.J., Rajaram S., vanGinkel M., 1996. CIMMYT's approach to breeding for  
4 wide adaptation. *Euphytica* 92(1-2), 175–183. WOS :A1996VT61200023.
- 5 Burgueño J., de los Campos G., Weigel K., Crossa J., 2012. Genomic prediction of  
6 breeding values when modeling genotype × environment interaction using pedigree  
7 and dense molecular markers. *Crop Science* 52(2), 707.
- 8 Ceccarelli S., Grando S., 2007. Decentralized-participatory plant breeding : an example  
9 of demand driven research. *Euphytica* 155(3), 349–360.
- 10 Ceccarelli S., Grando S., Bailey E., Amri A., El-Felah M., Nassif F., Rezgui S., Yahyaoui  
11 A., 2001. Farmer participation in barley breeding in syria, morocco and tunisia.  
12 *Euphytica* 122(3), 521–536.
- 13 Ceccarelli S., Grando S., Tutwiler R., Baha J., Martini A., Salahieh H., Goodchild  
14 A., Michael M., 2000. A methodological study on participatory barley breeding i.  
15 selection phase. *Euphytica* 111(2), 91–104.
- 16 Cotes J.M., Crossa J., Sanches A., Cornelius P.L., 2006. A bayesian approach for  
17 assessing the stability of genotypes. *Crop Science* 46(6), 2654.
- 18 Crossa J., Perez-Elizalde S., Jarquin D., Cotes J.M., Viele K., Liu G., Cornelius P.L.,  
19 2011. Bayesian estimation of the additive main effects and multiplicative interaction  
20 model. *Crop Science* 51(4), 1458.
- 21 Daetwyler H.D., Calus M.P.L., Pong-Wong R., de los Campos G., Hickey J.M., 2013.  
22 Genomic prediction in animals and plants : Simulation of data, validation, reporting,  
23 and benchmarking. *Genetics* 193(2), 347–+. WOS :000314821300003.
- 24 Dawson J., Murphy K., Jones S., 2008. Decentralized selection and participatory ap-  
25 proaches in plant breeding for low-input systems. *Euphytica* 160(2), 143–154.

- 1 Dawson J., Rivière P., Berthelot J., Mercier F., de Kochko P., Galic N., Pin S., Serpolay  
2 E., Thomas M., Giuliano S., Goldringer I., 2011. Collaborative plant breeding for  
3 organic agricultural systems in developed countries. *Sustainability* (3), 1206–1223.
- 4 Dawson J.C., Serpolay E., Giuliano S., Schermann N., Galic N., Chable V., Goldringer  
5 I., 2012. Multi-trait evolution of farmer varieties of bread wheat after cultivation in  
6 contrasting organic farming systems in europe. *Genetica* 140(1-3), 1–17.
- 7 Desclaux D., Nolot J.M., Chiffolleau Y., Leclerc C., Gozé E., 2008. Changes in the  
8 concept of genotype x environment interactions to fit agriculture diversification and  
9 decentralized participatory plant breeding : pluridisciplinary point of view. *Euphytica*  
10 163, 533–546.
- 11 Edwards J.W., Jannink J.L., 2006. Bayesian modeling of heterogeneous error and  
12 genotype  $\times$  environment interaction variances. *Crop Science* 46(2), 820.
- 13 Falconer D.S., 1952. The problem of environment and selection. *American Naturalist*  
14 (86), 293–298.
- 15 Falconer D.S., Mackay T.F.C., 1996. *Introduction to Quantitative Genetics*. Longman.
- 16 Finlay K., Wilkinson G., 1963. The analysis of adaptation in a plant-breeding pro-  
17 gramme. *Australian Journal of Agricultural Research* 14(6), 742–754.
- 18 Foucteau V., Denis J.B., 2001. Statistical analysis of successive series of experiments  
19 in plant breeding : a bayesian approach. In : Gallais A., Dillmann C., Goldringer  
20 I. (eds.), *Quantitative Genetics and Breeding Methods : The Way Ahead*, Inst Natl  
21 Recherche Agronomique, Paris, vol. 96. pp. 49–56. WOS :000167541400005.
- 22 Gauch H., 2006. Statistical analysis of yield trials by AMMI and GGE. *Crop Sci* 46(4),  
23 1488–1500.
- 24 Gauch H., Piepho H., Annicchiarico P., 2008. Statistical analysis of yield trials by  
25 AMMI and GGE : further considerations. *Crop Sci* 48(3), 866–889.
- 26 Gauch H.G.Z., 1997. Identifying mega-environments and targeting genotypes. *Crop*  
27 *Science* 37(2), 311.

- 1 Gelman A., Rubin D., 1992. Inference from iterative simulation using multiple sequences  
2 (with discussion). *Statistical Science* (7), 457–511.
- 3 Husson F., Josse J., Pagès J., 2010. Principal component methods - hierarchical clus-  
4 tering - partitional clustering : why would we need to choose for visualizing data ?
- 5 Jombart T., Devillard S., Balloux F., 2010. Discriminant analysis of principal compo-  
6 nents : a new method for the analysis of genetically structured populations. *BMC*  
7 *Genetics* 11(1), 94. PMID : 20950446.
- 8 Kastenhofer K., Bechtold U., Wilfing H., 2011. Sustaining sustainability science : The  
9 role of established inter-disciplines. *Ecological Economics* 70(4), 835–843.
- 10 Lewin A., Richardson S., Marshall C., Glazier A., Aitman T., 2006. Bayesian modeling  
11 of differential gene expression. *Biometrics* 62(1), 1–9. PMID : 16542224.
- 12 Ly D., Hamblin M., Rabbi I., Melaku G., Bakare M., Gauch H.G., Okechukwu R., Dixon  
13 A.G., Kulakow P., Jannink J.L., 2013. Relatedness and genotype × environment  
14 interaction affect prediction accuracies in genomic selection : A study in cassava.  
15 *Crop Science* 53(4), 1312.
- 16 Lönnstedt I., Britton T., 2005. Hierarchical bayes models for cDNA microarray gene  
17 expression. *Biostatistics* 6(2), 279–291.
- 18 Messmer M., Hildermann I., Thorup-Kristensen K., Rengel Z., 2012. Nutrient manage-  
19 ment in organic farming and consequences for direct and indirect selection strategies.  
20 In : *Organic Crop Breeding*, Edith T. Lammerts van Bueren, James R. Myers. Wiley-  
21 backwell edn., pp. 15–38.
- 22 Morris M., Bellon M., 2004. Participatory plant breeding research : Opportunities and  
23 challenges for the international crop improvement system. *Euphytica* 136, 21–35.
- 24 Murphy K., Campbell K., Lyon S., Jones S., 2007. Evidence of varietal adaptation to  
25 organic farming systems. *Field Crops Research* 102(3), 172–177.

- 1 Nabugoomu F., Kempton R., Talbot M., 1999. Analysis of series of trials where varieties  
2 differ in sensitivity to locations. *Journal of agricultural, biological and environnmen-*  
3 *tal Statistics* 4(3), 310–325.
- 4 Nashimoto K., Wright F., 2008. Bayesian multiple comparisons of simply ordered means  
5 using priors with a point mass. *Computational Statistics & Data Analysis* 52(12),  
6 5143–5153.
- 7 Nott D.J., Yu Z., Chan E., Cotsapas C., Cowley M.J., Pulvers J., Williams R., Little  
8 P., 2007. Hierarchical bayes variable selection and microarray experiments. *Journal*  
9 *of Multivariate Analysis* 98(4), 852–872.
- 10 Piepho H.P., 1998. Methods for comparing the yield stability of cropping systems - a re-  
11 view. *Journal of Agronomy and Crop Science-Zeitschrift Fur Acker Und Pflanzenbau*  
12 180(4), 193–213. WOS :000074884500001.
- 13 Plummer M., 2003. JAGS : A program for analysis of bayesian graphical models using  
14 gibbs sampling. In : *Proceedings. Vienna, Austria.*
- 15 Plummer M., 2008. Penalized loss functions for bayesian model comparison. *Biostatistics*  
16 9(3), 523–539. PMID : 18209015.
- 17 Rivière P., Goldringer I., Berthelot J.F., Galic N., Pin S., De Kochko P., Dawson J.C.,  
18 2013. Response to farmer mass selection in early generation progeny of bread wheat  
19 landrace crosses. *Renewable Agriculture and Food Systems FirstView*, 1–12.
- 20 Robert C., 2001. *The Bayesian Choice. Springer Texts in Statistics.* Springer, second  
21 edn.
- 22 Shukla G.K., 1972. Some statistical aspects of partitioning genotype-environmental  
23 components of variability. *Heredity* 29(2), 237–245.
- 24 Smith M.E., Castillo G. F., Gomez F., 2001. Participatory plant breeding with maize  
25 in mexico and honduras. *Euphytica* 122(3), 551–565.

- 1 Smyth G., 2004. Linear models and empirical bayes methods for assessing differential  
2 expression in microarray experiments. In : Statistical Applications in Genetics and  
3 Molecular Biology, Michael P.H. Stumpf, Imperial College London, vol. 3.
- 4 Spiegelhalter D.J., Best N.G., Carlin B.P., Van Der Linde A., 2002. Bayesian measures  
5 of model complexity and fit. *Journal of the Royal Statistical Society : Series B*  
6 (Statistical Methodology) 64(4), 583–639.
- 7 Theobald C.M., Talbot M., Nabugoomu F., 2002. A bayesian approach to regional and  
8 local-area prediction from crop variety trials. *Journal of Agricultural, Biological, and*  
9 *Environmental Statistics* 7(3), 403–419.
- 10 Viele K., Srinivasan C., 2000. Parsimonious estimation of multiplicative interaction in  
11 analysis of variance using kullback-leibler information. *Journal of Statistical Planning*  
12 *and Inference* 84(1-2), 201–219. WOS :000085006300013.
- 13 Virk D., Chakraborty M., Ghosh J., Prasad S., Witcombe J., 2005. Increasing the  
14 client orientation of maize breeding using farmer participation in eastern india. *Ex-*  
15 *perimental Agriculture* 41(04), 413–426.
- 16 Virk D., Witcombe J., 2007. Evaluating cultivars in unbalanced on-farm participatory  
17 trials. *Field Crops Research* 106(2), 105–115.
- 18 Westfall P., 1997. A bayesian perspective on the bonferroni adjustment. *Biometrika*  
19 84(2), 419–427.
- 20 Witcombe J.R., Joshi A., Joshi K.D., Sthapit B.R., 1996. Farmer participatory crop  
21 improvement. i. varietal selection and breeding methods and their impact on biodi-  
22 versity. *Expl. Agric.* 32, 445–460.
- 23 Wolfe M., Baresel J., Desclaux D., Goldringer I., Hoad S., Kovacs G., Löschenberger  
24 F., Miedaner T., Østergård H., Lammerts van Bueren E., 2008. Developments in  
25 breeding cereals for organic agriculture. *Euphytica* 163(3), 323–346.



- 1       Wricke G., 1962. Über eine methode zur erfassung der ökologischen streubreite in  
2       feldversuchen. Zeitschrift Fur Pflanzenzuchtung-Journal of Plant Breeding 47(1),  
3       92-&. WOS :A19624046C00005.
- 4       Yan W., Kang M., Ma B., Woods S., Cornelius P., 2007. GGE biplot vs. AMMI analysis  
5       of genotype-by-environment data. Crop Science 47, 643–355.
- 6       Zhao Y., Gowda M., Liu W., Wuerschum T., Maurer H.P., Longin F.H., Ranc N., Reif  
7       J.C., 2012. Accuracy of genomic selection in european maize elite breeding popula-  
8       tions. Theoretical and Applied Genetics 124(4), 769–776. WOS :000300245100012.



# **Annexe M : Mise en place d'une méthodologie de sélection participative sur le blé tendre en France**

Cet article a été publié dans la revue Innovations Agronomiques lors du colloque DinABio en 2013. DinABio est un colloque national visant à diffuser et mettre en débat les acquis récents des recherches en agriculture biologique.

## Mise en place d'une méthodologie de sélection participative sur le blé tendre en France.

Rivière P.<sup>1</sup>, Pin S.<sup>1</sup>, Galic N.<sup>1</sup>, de Oliveira Y.<sup>1</sup>, David O.<sup>2</sup>, Dawson J.<sup>3</sup>, Wanner A.<sup>4</sup>, Heckmann R.<sup>4</sup>, Obbellianne S.<sup>4</sup>, Ronot B.<sup>4</sup>, Parizot S.<sup>4</sup>, Hyacinthe A.<sup>4</sup>, Dalmasso C.<sup>4</sup>, Baltassat R.<sup>4</sup>, Bochède A.<sup>4</sup>, Mailhe G.<sup>4</sup>, Cazeirgue F.<sup>4</sup>, Gascuel J-S.<sup>4</sup>, Gasnier R.<sup>4</sup>, Berthelot J-F.<sup>4</sup>, Baboulène J.<sup>4</sup>, Poilly C.<sup>4</sup>, Lavoyer R.<sup>4</sup>, Hernandez M-P.<sup>4</sup>, Coulbeaut J.-M.<sup>4</sup>, Peloux F.<sup>4</sup>, Mouton A.<sup>4</sup>, Mercier F.<sup>4</sup>, Ranke O.<sup>4</sup>, Wittrish R.<sup>4</sup>, de Kochko P.<sup>4</sup>, Goldringer I.<sup>1</sup>

<sup>1</sup> INRA, Le Moulon ferme du Moulon F-91190 Gif sur Yvette

<sup>2</sup> INRA, Unité Mathématiques et Informatique Appliquées UR 341 Domaine de Vilvert F-78352 Jouy en Josas

<sup>3</sup> University of Wisconsin-Madison, 393 Plant Sciences / Moore Hall, 1575 Linden Dr. Madison, WI 53706, USA

<sup>4</sup> Réseau Semences Paysannes 3, avenue de la Gare, F-47190 Aiguillon

**Correspondance:** pierre.riviere@moulon.inra.fr

### Résumé

En Agriculture Biologique, les environnements des cultures et les pratiques agronomiques sont très diversifiés. Cette diversité peut être prise en compte par une sélection décentralisée associant le savoir-faire des agriculteurs et les connaissances scientifiques des chercheurs. Un travail co-construit entre associations et paysans du Réseau Semences Paysannes et l'équipe de recherche DEAP de l'INRA du Moulon a permis d'élaborer une approche innovante de sélection décentralisée et participative à la ferme.

Cette approche permet de (i) créer de nouvelles variétés populations de blé tendre adaptées localement (innovation génétique) (ii) mettre en place un mode d'organisation basé sur la co-construction entre acteurs et la décentralisation (innovation organisationnelle) et (iii) développer des dispositifs expérimentaux, créer des outils statistiques et de gestion de données qui favorisent ces innovations génétique et sociétale.

**Mots clés :** Méthodologie de la sélection, blé tendre, biodiversité cultivée, co-construction

### Abstract: Set up of a methodology for participatory plant breeding in bread wheat in France.

In Organic Agriculture, cultivation environments and agronomic practices are very diverse. This diversity can be handled with decentralized selection based on the knowledge of farmers and scientists. A collaborative work between associations from Réseau Semences Paysannes and the DEAP team from INRA du Moulon set up an innovative breeding approach on farm based on decentralization and participation of farmers.

This approach makes it possible to (i) create new population varieties of bread wheat locally adapted (genetic innovation) (ii) set up an organizational scheme based on decentralization and co construction between actors (societal innovation) and (iii) develop experimental designs, create statistical and data management tools which stimulate these genetic and societal innovations.

**Keywords:** breeding methodology, bread wheat, cultivated biodiversity, co-construction

## Introduction

L'Agriculture Biologique (AB) a des attentes fortes en termes de recherche sur les concepts et les méthodes de sélection et de gestion de la diversité.

En AB, l'utilisation d'intrants chimiques et de produits phytosanitaires issus de la chimie de synthèse est interdite. Les apports en azote proviennent essentiellement de la minéralisation des résidus, du fumier et des légumineuses qui peuvent être insérées dans les rotations. La minéralisation est irrégulière au cours du temps. Ces contraintes rendent les environnements de culture moins contrôlés et donc plus stressants pour les plantes. Actuellement, peu de variétés ont été sélectionnées pour l'AB. Même si récemment, la création variétale conventionnelle a été utilisée pour développer des lignées pures adaptées à l'AB (Rolland *et al.*, 2012). Nous proposons une autre approche méthode qui associe deux principes approches pour créer de nouvelles variétés adaptées aux pratiques de l'AB : développer des variétés avec une grande diversité inter- et intra-variétale et décentraliser la sélection en faisant participer les acteurs au processus de sélection.

Les variétés-populations sont des variétés hétérogènes qui présentent une grande diversité entre les individus. En France, jusqu'au 19<sup>ème</sup> siècle, de nombreuses variétés-populations étaient cultivées. Ces variétés étaient adaptées localement aux conditions pédo-climatiques et aux pratiques agronomiques des paysans. Les paysans semaient et sélectionnaient leurs variétés d'année en année. Dès la fin du 19<sup>ème</sup> siècle, le métier de sélectionneur s'est développé et progressivement les variétés utilisées ont été celles des sélectionneurs. Au 20<sup>ème</sup> siècle, les variétés-populations ont été remplacées par des variétés homogènes (lignées, hybrides, clones). Ce processus s'accompagne d'une réduction de la biodiversité cultivée rencontrée dans les paysages comme cela a été montré sur le blé tendre (Bonneuil *et al.*, 2012). Une telle diminution augmente la vulnérabilité des agro-systèmes (Tooker et Franck, 2012). Un paysage diversifié au niveau des espèces cultivées, des variétés cultivées (diversité inter-variétale, notamment sur les gènes de résistances) et de la diversité intra-variétale est source de services écosystémiques qui augmentent les externalités positives ainsi que la stabilité et la résilience des agro-systèmes (MEA, 2005 ; Papaïx *et al.*, 2011 ; Chateil *et al.*, 2013). La biodiversité domestique<sup>1</sup> peut donc être considérée comme un des leviers nécessaires pour répondre aux évolutions du climat, des pathogènes et de la société. Elle est gérée soit dans des banques de graines (gestion statique) soit en culture dans les stations de recherche et les champs des paysans (gestion dynamique) (Thomas *et al.*, 2011).

Les variétés hétérogènes génétiquement présentent certains avantages agronomiques. Elles sont plus stables et résilientes au cours du temps (Wolfe *et al.*, 2008), permettent de maintenir une diversité des résistances aux pathogènes (Finckh *et al.*, 2000) et contribuent à diversifier les communautés de l'agro-système, améliorant les services écologiques (Tooker et Franck, 2012 ; Chateil *et al.*, 2013).

Une population hétérogène évolue sous la pression des forces évolutives. La sélection naturelle (influencée par les conditions pédo-climatiques mais également par les pratiques agronomiques des paysans) et la sélection de l'Homme permettent aux populations d'évoluer et de s'adapter localement. Ces approches sont valorisées dans la sélection décentralisée.

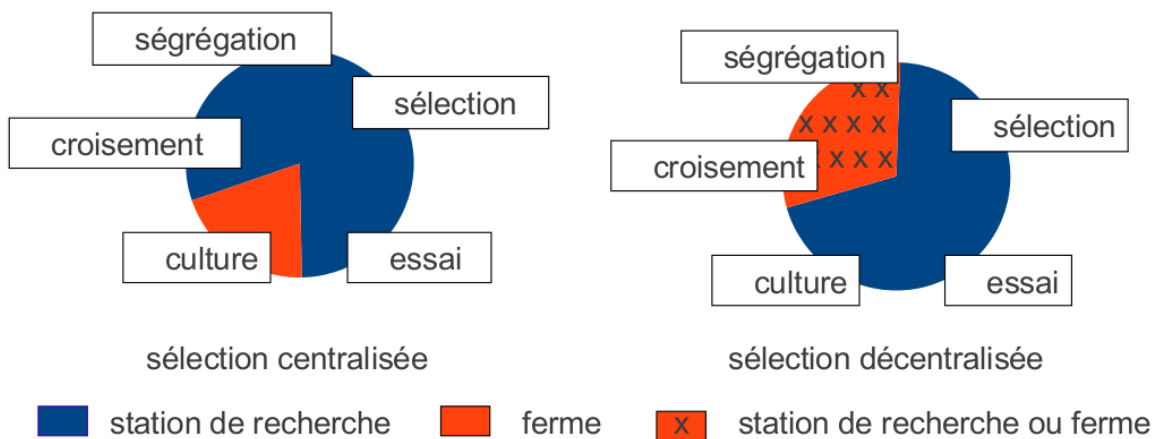
Dans le processus de sélection décentralisée, la sélection s'effectue directement dans les environnements « cibles » (Wolfe *et al.*, 2008). Les interactions entre le génotype et l'environnement sont ainsi valorisées. L'environnement est ici pris au sens large : stress biotique, abiotique, contexte pédo-climatique, pratiques agronomiques (rotations, itinéraires de cultures, etc.), et environnement socio-économique (type de filière par exemple) (Desclaux *et al.*, 2008). Au contraire, la sélection centralisée (Figure 1) sélectionne les génotypes dans des stations de recherche le plus souvent en conditions non limitantes et contrôlées (apports d'azote, irrigation, produits phytosanitaires). La sélection centralisée est dans de nombreux cas moins efficace que la sélection décentralisée pour

---

<sup>1</sup> Diversité entre les espèces et au sein des espèces animales et végétales domestiques et sélectionnés par l'Homme

développer des variétés adaptées à des conditions limitantes et variées (Falconer, 1952 ; Murphy *et al.*, 2007; Wolfe *et al.*, 2008), bien que certaines études montrent le contraire (Annicchiarico *et al.*, 2010), notamment lorsque les environnements cibles se rapprochent des environnements de sélection ou que la variabilité génétique évaluée est restreinte (variétés issues de sélection conventionnelle) (Wolfe *et al.*, 2008).

En plus d'être décentralisée, la sélection peut être participative (Ceccarelli et Grando, 2007) : les paysans, les associations et l'équipe de recherche mettent en commun les observations de terrain et les analyses scientifiques. L'utilisateur final de la variété étant le paysan, l'associé au processus de sélection permet de valoriser les interactions GxE et ainsi d'augmenter les chances d'adoption des variétés. Le paysan va sélectionner des variétés correspondant à ses pratiques culturelles, ses sols, sa filière, ses débouchés commerciaux. Au delà de la sélection, cette collaboration permet d'autonomiser les paysans et de leur redonner une place dans la gouvernance des semences (Bonneuil et Demeulenaere, 2007 ; Morris et Bellon, 2004). La sélection décentralisée et participative est utilisée dans les pays du Sud (Ceccarelli *et al.*, 2000 ; Smith *et al.*, 2001 ; Virk *et al.*, 2005) et plus récemment dans les pays du Nord (Mendes Moreira *et al.*, 2008 ; Chable *et al.*, 2008 ; Dawson *et al.*, 2011; Rivière *et al.*, 2013).



**Figure 1** : Représentation des différences entre sélection centralisée (à gauche) et décentralisée à la ferme (à droite) ; d'après Ceccarelli et Grando (2007).

En France, un programme de sélection participative (SP) a démarré en 2005 entre le Réseau Semences Paysannes (RSP) et l'équipe DEAP (Diversité Evolution et Adaptation des Populations) de l'INRA du Moulon. Le RSP regroupe 75 associations qui ont pour objectif l'utilisation et le maintien de la biodiversité cultivée et la mise en place de structures collectives de gestion des semences (Maison de la semence) (Réseau Semences Paysannes, 2008).

L'objectif de cette collaboration entre le RSP et DEAP est de mettre au point une méthodologie de la sélection basée sur la décentralisation et la co-construction entre paysans, associations et chercheurs. Cette méthodologie vise à (i) créer des variétés-populations adaptées aux conditions locales et aux pratiques des paysans, (ii) développer des méthodes et des outils opérationnels pour la gestion / sélection de la biodiversité cultivée à la ferme et (iii) renforcer l'apprentissage et l'autonomie des agriculteurs en matière de gestion/sélection des semences.

Au départ du projet en 2005-2006, Jean François Berthelot (JFB), paysan-boulangier près d'Agen, a réalisé sur sa ferme avec l'aide de l'équipe de recherche 90 croisements pour créer de nouvelles populations. Le RSP souhaitait également démontrer qu'il était possible aux paysans de créer leurs variétés. Ces nouvelles populations hétérogènes ont été multipliées sur sa ferme entre 2006 et 2008.

En 2008, ces populations ont été distribuées au sein du RSP. Entre 20 et 25 paysans, répartis dans toute la France, participent au projet de SP depuis 2008.

Cet article décrit la méthodologie mise en place dans ce processus de sélection.

## 1. Co-construction du programme entre paysans, associations et chercheurs.

La méthodologie présentée est le résultat de six années de collaboration entre paysans, associations du RSP et l'équipe de recherche. Ce projet est un projet de recherche action : Tout en effectuant le programme de recherche pour répondre à la question initiale, la sélection, la méthodologie et l'organisation peuvent être modifiées, ajustées à chaque étape à la suite de décisions collectives.

La diversité utilisée dans le programme vient en grande partie des populations issues des croisements de JFB en 2005. JFB a choisi les parents qui sont des variétés de pays, des variétés anciennes et quelques variétés modernes de Suisse et d'Allemagne créées pour l'AB. L'objectif de JFB était d'associer les qualités agronomiques (résistance à la verse, concurrence vis à vis des adventices, etc.) aux qualités gustatives (évaluées par analyse sensorielle) et nutritionnelles (teneur et qualité des protéines, contenu en minéraux, micro-nutriments et vitamines, contenu en acide phytique et teneur en fibres évaluées sur certaines variétés et informations issues de Roussel *et al.* (2005) pour les autres) du pain. Des mélanges de variétés-populations sont également utilisés comme source de diversité. Au cours du temps, certains paysans ont proposé de nouveaux croisements entre des variétés qu'ils avaient identifiées et des demandes de formation à la réalisation de croisements ont été faites. Plusieurs formations ont donc été réalisées par l'équipe de recherche à destination des paysans. Cela a permis de créer 120 nouvelles populations depuis 2006. Des formations sur les principes de la génétique et la sélection ont également été réalisées à la demande des paysans, de même qu'un livret technique (version beta) sur la SP<sup>2</sup>.

Chaque année, avant les semis, une réunion est organisée avec tous les acteurs du projet. Au début du projet, les mesures à réaliser pour évaluer les nouvelles populations créées sont choisies. Des fiches de suivi ont été mises en place pour chaque étape du cycle du blé : semis, tallage, épiaison et maturité. Au delà des critères agronomiques « classiques » (port au tallage, date d'épiaison, présence de maladies, etc.), les paysans ont souhaité donner une note globale (subjective et holistique) et un commentaire. Ils ont également demandé que figure sur les fiches un rappel des dernières notations prises. Ces demandes ont été intégrées dans la création des fiches de suivi qui évoluent d'année en année.

Un exemple de la « fiche été » est présenté en Figure 2.

Croisement	Date	Globale	Biomasse	Crosses		Couleur		hétéro	Barbes		Verse	Hauteur (m)
C21 <sub>2</sub> Sb				pas croisé	blanc			barbus	<input type="checkbox"/> à plat			
				intermédiaire	intermédiaire			intermédiaire	<input type="checkbox"/> couché			
				croisé	rouge			nonbarbus	<input type="checkbox"/> intermédiaire			
									<input type="checkbox"/> presque droit			
									<input type="checkbox"/> droit			
Date de la récolte :												
Poids du grain récolté sur la micro-parcelle :												
Commentaires pour cette pop :												
Rappel hiver : date : 20/03/2012 globale= 4, reprise= 5, densité= , Port au tallage : demi-étalé, Attitude des feuilles : de biais												
Commentaires hiver : une des plus développée												
Rappel printemps : date : 15/05/2012 globale= 4, tallage= moyen, vigueur= 3, Couleur : vert, Adventices : gaillet (beaucoup) ; vulpin (un peu) ; coquelicot (un peu)												
Commentaires printemps : joli, plus court que le mélange												

Figure 2 : Exemple de la « fiche été » à remplir au moment de la maturité.

<sup>2</sup> Fruchaud, G., Rivière, P., Goldringer, I., Des blés en (R)Evolution : création de variétés de blé tendre de qualité adaptées à l'Agriculture Biologique et conservation de la biodiversité cultivée dans un processus de sélection participative. 2011. Document technique.

En plus de ces fiches qui rapportent des données qualitatives, des mesures quantitatives sont réalisées à la récolte par l'équipe de recherche (couleurs, courbure de l'épi, barbes, hauteur, distance entre la base de l'épi et la feuille drapeau, poids de l'épi, nombre d'épillets, poids de mille grains et taux de protéines).

Les fermes sont divisées en deux types selon le dispositif expérimental qu'elles accueillent : les fermes régionales et les fermes satellites (Figure 3). La variété-population Rouge du Roc, une sélection paysanne dans Rouge de Bordeaux, est utilisée comme témoin répété dans toutes les fermes. Le dispositif expérimental des fermes régionales comporte deux blocs avec 4 populations ou variétés présentes dans chaque bloc : C14 et C21 issus de croisements de 2006, Renan qui est largement utilisé en AB et Rouge du Roc. Dans les fermes satellites, Rouge du Roc est répétée deux fois. Les fermes régionales accueillent entre 20 et 35 populations. Elles sont des lieux de rencontre l'été avec les paysans de la région, les boulangers, les consommateurs. Les fermes satellites accueillent entre 5 et 10 populations. Ce dispositif a été mis en place à partir de 2010. Il est le résultat d'un compromis entre une maximisation du nombre de génotypes testés dans le terrain alloué et la possibilité de calculer des variances environnementales sur chaque ferme pour pouvoir différencier les populations.

Rouge-du-Roc	pop1	pop2	pop3	
pop4	pop5	C21	pop6	
pop7	C14	pop8	pop9	
pop10	pop11	pop12	Renan	
pop13	C21	pop14	pop15	
Renan	pop16	pop17	pop18	
pop19	pop20	pop21	Rouge-du-Roc	
pop22	pop23	C14	pop24	

Rouge-du-Roc	pop1
pop2	pop3
pop4	pop5
pop6	pop7
pop8	Rouge-du-Roc

**Figure 3** : Dispositif sur les fermes régionales (gauche) et les fermes satellites (droite). Les témoins sont sur fond noir.

## 2. Outils d'analyse et de gestion des données mis en place

Pour appuyer notre méthodologie, deux types d'outils ont été mis en place : statistiques et gestion des données.

### 2.1 Statistiques

Les dispositifs expérimentaux sur les fermes, qui résultent de compromis entre possibilité d'analyses statistiques et surface et travail pour le paysan, sont très déséquilibrés : dans une matrice regroupant toutes les combinaisons génotype-ferme possibles, il y a plus de 95% de données manquantes. De nouvelles méthodes statistiques ont été développées pour analyser ces données. Les objectifs de ces analyses sont (i) de réaliser des comparaisons de moyennes sur les fermes et (ii) d'estimer les effets génétiques et fermes sur le réseau ainsi que la sensibilité des génotypes à l'interaction GxE.

Savoir si deux génotypes sont significativement différents l'un de l'autre permet de guider la sélection entre génotypes. Dans notre cas, une analyse de variance classique avec estimation des effets génétiques est impossible car les populations sont très peu répétées et il n'y a pas assez de degrés de liberté à la résiduelle.

Pour nous affranchir de cette contrainte, nous avons profité de la masse de données dont nous disposons pour utiliser un modèle Bayésien hiérarchique (Robert, 2001). Ce modèle fait l'hypothèse que les variances environnementales de chaque ferme sont issues d'une loi commune. Cette loi est une



loi Gamma avec une large variance qui permet de couvrir la diversité de variabilité des différents milieux. Le modèle estime la loi des variances d'où sont issues leurs valeurs. C'est possible car de nombreuses fermes ont la même structure d'essai. In fine, ce modèle permet de comparer les populations deux à deux sur chaque ferme.

Il est également intéressant pour le paysan d'avoir des informations sur le comportement des populations sur d'autres fermes. Il faut être capable de regrouper les fermes selon le comportement des populations qui y sont évaluées et savoir la part des interactions entre les populations et leurs environnements.

Nous utilisons le modèle de Finlay Wilkinson (Nabugoomu *et al.*, 1999). Ce modèle permet d'estimer des effets génotypes, fermes et une sensibilité à l'interaction pour chaque génotype. Un tel modèle ne va pas dans les détails de l'interaction comme le ferait un modèle de type AMMI (Gauch *et al.*, 2008). Cependant, il est plus parcimonieux et donc moins sensible au déséquilibre des données et permet de retirer beaucoup d'information.

Comme précédemment, nous utilisons un modèle Bayésien hiérarchique pour estimer les différents paramètres (Robert, 2001). Nous faisons l'hypothèse que les effets des génotypes, les effets des lieux et les sensibilités des génotypes sont aléatoires. L'analyse est possible car nous avons de nombreuses données dans l'espace et dans le temps. Un tel modèle permet de classer les fermes en groupes de ressemblance et d'informer sur la sensibilité des variétés à l'interaction. De telles informations sont utiles pour les paysans qui souhaitent tester de nouvelles populations. Ils peuvent tester des populations issues de fermes qui se comportent de façon proche de leur ferme, ou utiliser des populations plus stables (sensibilité à l'interaction faible).

## 2.2 Gestion des données

Une base de données a été développée pour stocker les informations relatives au projet de SP : SHiNeMas (Seed History and Network Management System). Le principe de l'organisation des données est que chaque lot de semences est relié à un autre lot de semences par une relation : reproduction, diffusion, sélection, mélange. Ces relations permettent de détailler l'historique d'un lot de semences. Chaque relation, lot de semences ou personne est relié à des données. Ces données peuvent être qualitatives (confiance en l'information qu'une diffusion a effectivement eu lieu, statut professionnel d'une personne, ...) ou quantitatives (hauteur, poids de mille grains, données moléculaires, ...). Le choix des variables est libre, cela permet une grande flexibilité. Chaque variable est liée à une méthode d'obtention de cette variable. Des fonctions R ont été développées pour réaliser des requêtes sur la base de données. Cela permet à la fois d'étudier le fonctionnement du réseau (nombre de multiplications, de sélections, d'échanges, etc.), et de mettre en place des fiches de suivi et des dossiers personnalisés des résultats pour chaque paysan.

Tous ces outils sont open source et gratuits. Ce point est primordial pour la transmission et la diffusion de ce travail dans les associations. Un package R regroupant toutes ces fonctionnalités (analyses statistiques et interface avec ShiNeMas) ainsi qu'une interface de saisie des données via des fichiers types et de requête sont en cours de développement.

## 3. Différentes étapes entre le semis et la récolte

Dans cette partie, le terme « animateur » est utilisé pour désigner une personne de l'équipe de recherche ou d'une association. L'équipe de recherche a coordonné le projet de 2005 à 2012. Pour l'année 2012-2013, les associations animent le projet, aidées par l'équipe de recherche. Ce travail d'animation est la clé de réussite du projet (Ali Brac de la Perrière *et al.*, 2011 ; Storp et De Kochko, 2011).

Les différentes étapes du programme de SP sur une année sont les suivantes : semis, notations et observations, sélection intra-variétale, récolte, décisions pour le semis et échanges (Figure 4).

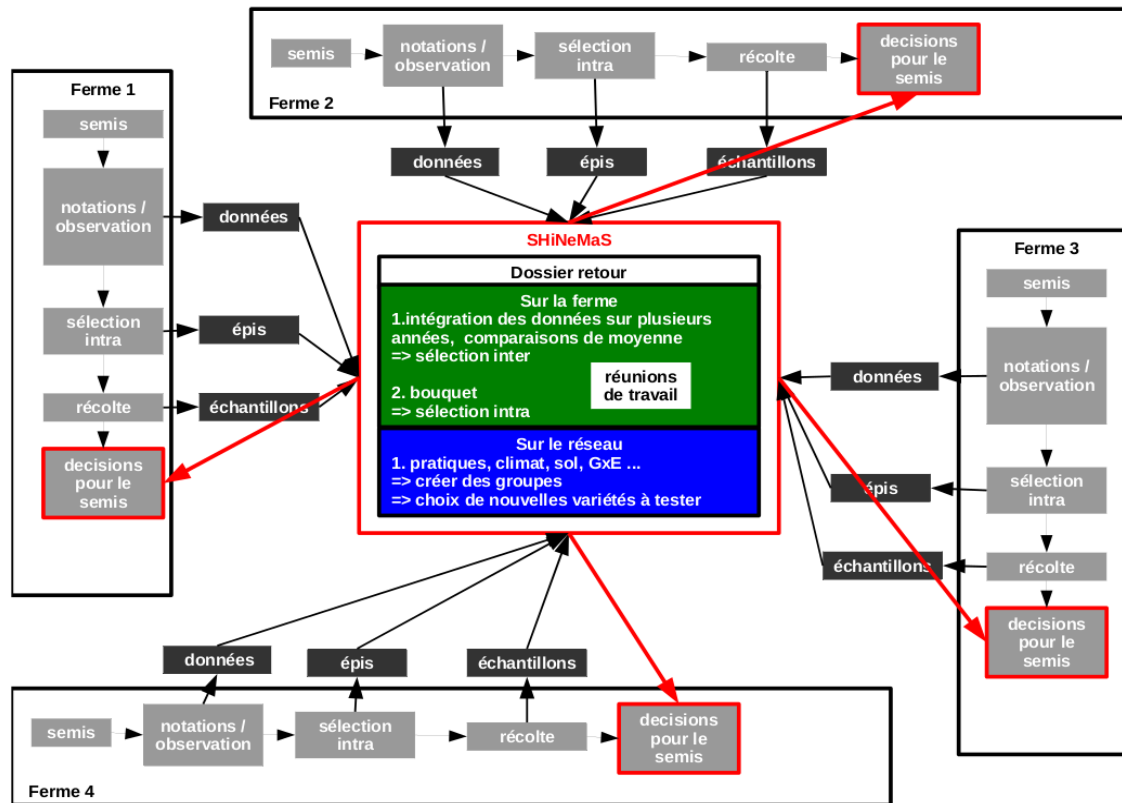


Figure 4 : Les différentes étapes du projet de Sélection Participative.

### 3.1 Semis

Les paysans choisissent les populations qu'ils souhaitent semer en plus du ou des témoins mis en place. Ils envoient la liste des lots de semences choisis à l'animateur qui coordonne les échanges. Certains échanges de paysan à paysan ont également lieu. L'animateur propose ensuite un plan de semis respectant le dispositif expérimental (fermes régionales ou satellites).

### 3.2 Notations et observations

Les notations sur des fiches sont faites au cours du cycle du blé par les paysans. L'animateur envoie les fiches ainsi qu'une notice « comment remplir la fiche » pour chaque étape : semis, tallage, épiage, maturité. Des informations sur le climat et d'éventuels accidents sur la culture sont recueillies également. Des notices concernant les maladies et les ravageurs sont également fournies. Quatre fiches de suivi sont donc envoyées au cours de l'année. Si nécessaire, les animateurs recontactent les paysans pour leur rappeler les dates d'observation.

### 3.3 Sélection intra variétale

Les paysans qui le souhaitent peuvent sélectionner des épis et envoyer les bouquets d'épis sélectionnés à l'équipe de recherche. Cela permet d'étudier le différentiel de sélection et la réponse à la sélection massale l'année suivante. Le différentiel de sélection (S) est la différence entre la moyenne

des épis sélectionnés et la moyenne du vrac (i.e. des épis pris au hasard). La réponse à la sélection (R) est la différence entre la moyenne des descendances issues du bouquet de sélection et des descendances issues du vrac. L'héritabilité au sens strict  $h^2_{ss}$  relie S et R tel que :  $R=h^2_{ss} S$ . Dans notre cas, nous avons accès à S et R, on en déduit l'héritabilité réalisée  $h^2_r$ .  $h^2_r$  est compris entre 0 et 1. Si elle vaut 0, alors le différentiel de sélection est dû uniquement à l'environnement. Si elle vaut 1, alors le différentiel de sélection est uniquement dû à la génétique.

### 3.4 Récolte

Des visites sont organisées sur les fermes régionales au moment des récoltes. Ces visites permettent de présenter le projet, et de rencontrer paysans, boulangers et représentants d'associations. Un « jeu de la sélection » est organisé. Dans un espace d'une trentaine de populations, chaque participant doit choisir cinq populations qu'il évalue positivement, et cinq populations qu'il évalue négativement. Une discussion est ensuite organisée sur les critères de choix des différents blés par les participants.

Après la récolte, les paysans envoient à l'équipe de recherche un échantillon de 50 épis par population (sauf pour les bouquets de sélection où le nombre d'épis est libre). L'équipe de recherche mesure, entre autres, le poids d'un épi, le poids de mille grains et le taux de protéines sur les grains des échantillons. Dans le cadre d'un travail de recherche sur l'impact du programme de SP sur l'évolution des caractères des populations, des mesures plus précises ont été faites (longueur de l'épi, nombre d'épillettes, etc). Mais celles-ci ne font pas parties d'un suivi « en routine » des populations.

### 3.5 Décision pour le semis

Les données sont rentrées dans la base de données ShiNeMas. Cela permet de générer un « dossier retour » qui est personnalisé pour chaque paysan. Ce dossier complète les observations des paysans pour choisir les populations à semer. Au cours de réunions de semis, une discussion autour des « dossiers retour » a lieu. Les paysans peuvent baser leur sélection (inter et intra variétale) sur les résultats obtenus. Ils peuvent également choisir de nouveaux génotypes sur le réseau. Ce dossier est structuré en deux parties.

- Une première partie présente les résultats obtenus sur la ferme : un bilan des notations qui ont été faites au cours du cycle et les comparaisons de moyennes sur les données quantitatives. Les résultats sur plusieurs années sont également présentés. Pour les paysans concernés, les résultats sur la réponse à la sélection et l'héritabilité réalisée sont présentés. Ces notions de génétique ont été discutées en réunion afin que les paysans s'approprient ces concepts. Ces résultats leur permettent d'évaluer l'impact de leur sélection sur les critères qu'ils considèrent d'intérêt.
- Une deuxième partie présente les résultats sur le réseau pour que les paysans puissent observer les regroupements des fermes et identifier de nouvelles populations à tester sur leurs fermes. Ils ont également accès aux résultats sur la sensibilité des populations à l'interaction.

## 4. Résultats et discussion

### 4.1 Diversité cultivée dans le réseau de fermes : nombre et types de populations

Depuis le début du projet de SP, 650 germplasm de différents types ont été cultivées par les paysans participants : variétés de pays, variétés anciennes, variétés modernes créées pour l'AB en Allemagne et en Suisse, populations issues de croisements, mélanges et sélections massales réalisées par les paysans dans les types précédemment cités. Sur ces 650 germplasm, 165 ont été créés par croisement par les paysans (avec l'aide technique de l'équipe de recherche).

Le nombre de lots de semences récoltés a évolué de 80 en 2006 pour atteindre plus de 900 en 2012 (Figure 5). Le nombre de germplasm mis en culture en 2012 était de 334. Il s'agit en grande partie de populations issues de croisements, de mélanges et de populations issues de sélections massales.

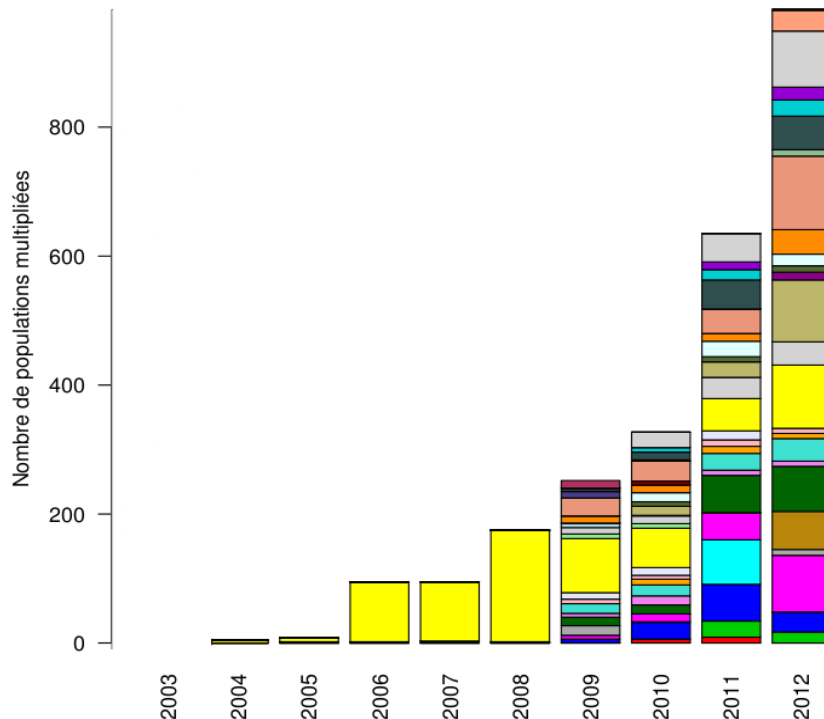


Figure 5 : Évolution du nombre de lots de semences récoltés sur le réseau. Une couleur représente un paysan.

## 4.2 Comparaison des populations sur les fermes et sélection intra et inter population.

### 4.2.1 Comparaison des populations

L'objectif est de différencier des populations entre elles. Le modèle Bayésien permet de faire des tests de comparaisons entre populations sur les fermes. L'erreur de type I (risque de rejeter l'hypothèse  $H_0$  : les populations sont égales) utilisée est de  $0.10/\text{nombre de génotypes sur la ferme}$  par défaut. La majorité des variables mesurées sur les fermes présentent au moins deux groupes distincts avec ce seuil. Il arrive que pour une variable donnée sur une ferme, un seul groupe soit détecté. Dans ce cas, nous augmentons l'erreur de type I jusqu'à pouvoir distinguer au moins deux groupes. Ceci permet aux paysans de différencier les populations. Un des objectifs des paysans du projet est de se (ré)approprier la sélection. Nous souhaitons donc leur apporter le maximum d'information pour qu'ils puissent sélectionner ; nous leur indiquons alors le risque associé au test.

### 4.2.2 Sélection entre populations

Il est possible de distinguer, donc de sélectionner, les populations. En moyenne, les paysans ont re-semé 75% (2009), 82% (2010), 76% (2011) et 66% (2012) des lots de semences récoltés l'année précédente. Au fur et à mesure que le projet avance, la sélection inter population semble en moyenne devenir plus forte.

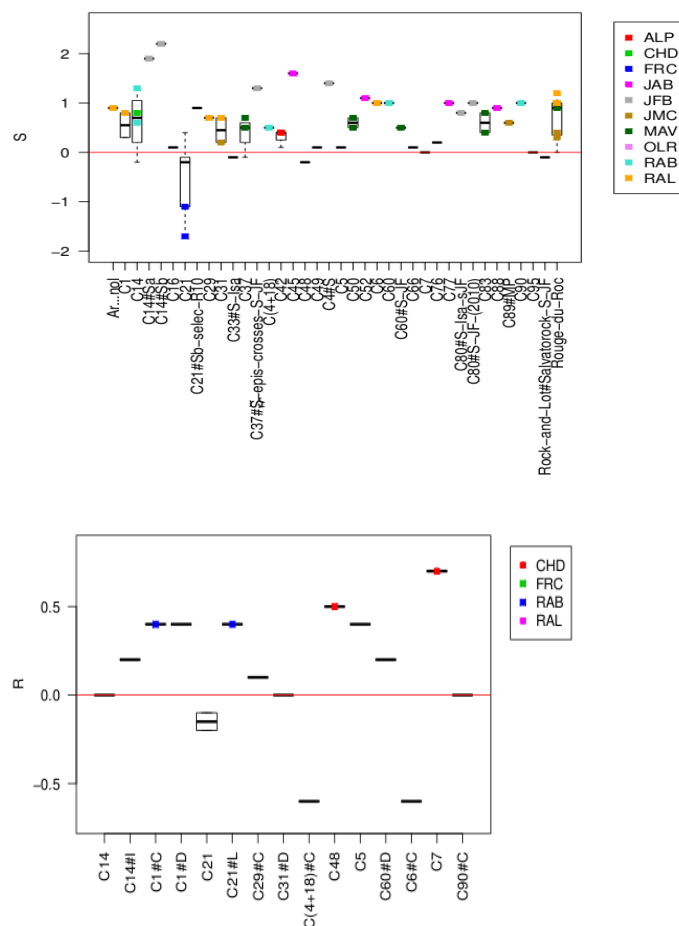
### 4.2.3 Sélection intra population et réponse à la sélection

Le nombre de sélection intra population sur le réseau était d'une trentaine en 2008-2009 pour passer à plus de 100 en 2011-2012 et 2012-2013. Ces sélections donnent naissance à de nouvelles populations

qui sont évaluées sur les fermes.

Tous les différentiels de sélection (différence entre sélection et vrac en 2010-2011) pour le poids de l'épi sont positifs sauf pour le paysan FRC (Figure 6). Cela s'explique car ce dernier n'a pas sélectionné sur le poids de l'épi mais sur la couleur uniquement. Parmi les différentiels positifs, 60% sont significativement supérieurs à 0.

Nous montrons également que quand la réponse à la sélection est significative, elle est positive (Figure 6). La réponse à la sélection sera d'autant plus forte que le caractère sera héritable. L'héritabilité du poids de l'épi ou de mesures s'en rapprochant est intermédiaire : inférieure à celle du poids de mille grains, de la hauteur ou de la précocité, mais supérieure à celle du nombre d'épis par plante ou du rendement (Goldringner *et al*, 1997; Sadras *et al*, 2013). C'est encourageant et donne aux paysans des résultats qui leur permettent de mieux orienter leur sélection. Cependant, tous les paysans n'ont pas suivi le protocole d'étude de la réponse à la sélection. Une majorité a semé directement la version sélectionnée sans le vrac, par manque de place sur la ferme ou suite à une mauvaise compréhension. Lors de la réunion de semis suivante, l'intérêt de suivre ce protocole a été abordé et de nombreux paysans le suivront cette année.



**Figure 6 :** Différentiel de sélection (gauche) et réponse à la sélection (droite) pour le poids de l'épi. Les génotypes sont en abscisse. Les valeurs de S (et respectivement R) sont en ordonnée. Pour une population, il peut y avoir plusieurs sélections. Les valeurs significativement différentes de 0 (au niveau 5%) sont marquées d'un carré de couleur qui correspond au paysan qui a fait la sélection.

### 4.3 Évolution des pratiques des paysans, des associations et de l'équipe de recherche.

Les pratiques des paysans ont évolué depuis le début du projet: augmentation des échanges (Figure 7), des sélections (inter- et intra-population) et de la création de variabilité par croisement.

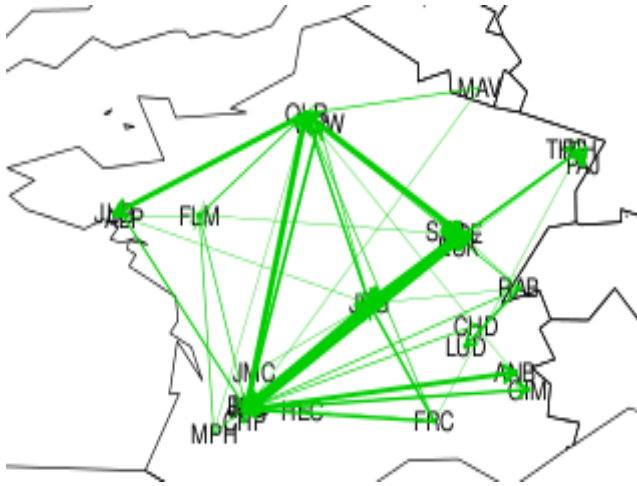


Figure 7 : Échanges entre paysans sur le réseau en 2011.

Les pratiques des chercheurs ont également évolué. Travailler avec les paysans demande plus de temps et d'organisation qu'un travail classique sur la station de recherche. Par exemple, le recueil des données implique une grande logistique et de nombreuses communications téléphoniques. La base de données ShiNeMas et les fonctions qui ont été développées sont d'une grande aide pour s'organiser.

### 4.4 Une approche qui permet une innovation génétique et sociétale

La stratégie mise en place est issue d'une co-construction entre paysans, associations et équipe de recherche. La réussite de ce type de projets demande du temps et requiert la confiance des différents acteurs (Storup et De Kochko, 2011).

Ce projet a trois échelles : nationale, régionale et locale (sur la ferme). Dans notre cas, les échelles nationale et régionale ne sont pas différenciées jusqu'en 2012. Depuis 2013, le processus est en cours de décentralisation dans chaque association régionale.

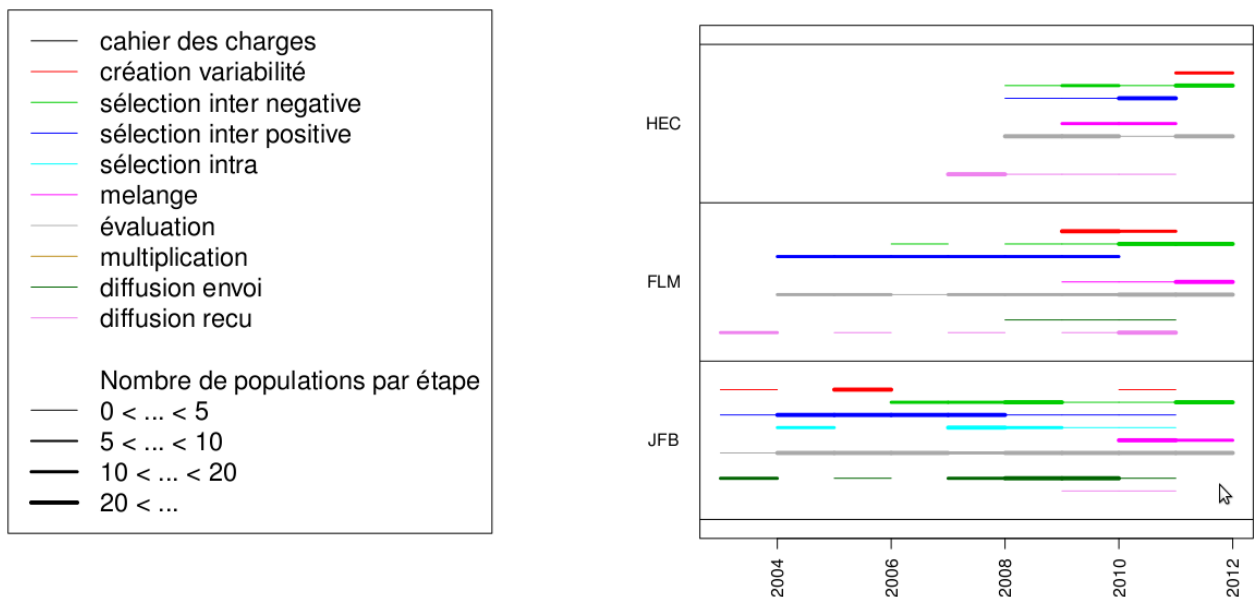
Dans un premier temps, le groupe a discuté les objectifs à atteindre lors de réunions. Chaque personne s'exprime lors de tours de table. Le groupe s'est mis d'accord sur les mesures à réaliser et les dispositifs expérimentaux à mettre en place. Certains paysans ont demandé des mesures supplémentaires comme la force boulangère (W). Cette demande provient de paysans intégrés dans des filières plus longues (avec un meunier qui achète les grains pour faire de la farine qui est ensuite vendue à des boulangers).

Les outils créés permettent de répondre aux objectifs de départ tout en permettant aux associations au niveau régional de s'adapter aux besoins des paysans et de répondre à leurs objectifs.

ShiNeMas permet de prendre en compte les variables que l'on souhaite. Ainsi, les groupes régionaux peuvent enregistrer des variables qui leur sont propres. Par exemple, dans les fermes d'élevage, la quantité de paille sera plus importante que dans les fermes céréalnières où la qualité du grain sera davantage prise en compte. Dans tous les cas, il est important que le groupe se mette d'accord pour avoir certaines mesures en commun afin que tout le monde puisse profiter des résultats des autres fermes. Par exemple, le poids de mille grains, la verse, la hauteur, la distance entre la dernière feuille et

la base de l'épi, sont des critères partagés par tous les paysans. Une analyse de ces choix ainsi que de l'évolution des critères dans le groupe reste à conduire. Le « dossier retour » illustre comment le projet s'adapte individuellement à chaque paysan.

Dans un programme de sélection conventionnelle, la sélection se fait de manière très linéaire : mise en place d'un cahier des charges, création de variabilité, sélection dans cette variabilité, évaluation, multiplication, inscription au catalogue et diffusion sur le marché. Dans notre cas, nous avons repéré dix types d'activités dans notre processus de sélection : mise en place d'un cahier des charges (d'une charte dans notre cas), création de variabilité, sélection entre populations négative (i.e. choisir une population qui ne sera pas ressemée), sélection entre populations positive, sélection intra population, mélange, évaluation, multiplication, envoi de lots de semences et réception de lots de semences. Ces étapes ne s'enchaînent pas de manière successive sur les fermes (Figure 8). En effet, grâce au travail du collectif, les paysans ne passent pas forcément par toutes les étapes, notamment ils commencent rarement par l'étape de création de variabilité mais utilisent plutôt au début les populations déjà créées. De plus, toutes les fermes ne sont pas dans les mêmes activités au même moment. Par exemple, la création de variabilité par croisement a été effectuée par 10 paysans sur 25. Certains paysans se sont échangés des lots de semences, etc.



**Figure 8** : Mise en place des différentes activités liées à la sélection dans le projet de SP. Trois dynamiques différentes sont représentées pour 3 paysans : HEC, FLM et JFB.

Les paysans sont acteurs de la sélection car ce sont eux qui choisissent ce qu'ils vont semer sur leurs fermes. Cette stratégie permet l'innovation au niveau individuel. Les paysans peuvent aller à leur rythme (Figure 8) et disposer des résultats concernant leurs essais sans être dépendants des autres paysans. Cependant, la mise en réseau est primordiale. C'est la mise en réseau qui permet de réaliser les statistiques et de récupérer de nouvelles populations à tester. Dans ce sens, les paysans sont dépendants entre eux. C'est cet équilibre entre indépendance/dépendance qu'il convient de trouver pour mener à bien le projet.

La diversité des pratiques de sélection permet au réseau une grande capacité de création variétale et de décentralisation de la prise de contrôle (qui fait quoi, qui sait quoi, qui envoie les semences, qui crée les nouvelles populations, etc.). Ce point est important car tous les paysans ont la possibilité de jouer le même rôle dans ce projet, ainsi les sources d'innovations sont multipliées. Deux types d'innovations émergent.

L'innovation variétale tout d'abord : les croisements qui ont été réalisés, les mélanges et les pressions

évolutives sur chaque ferme permettent de créer de nouvelles populations adaptées aux pratiques, au terroir et au débouché de chacun. Cette diversité peut être utilisée comme nouvelle base de sélection. Certains paysans utilisent ces nouvelles variétés-populations en mélange. Elles viennent en général compléter un mélange pré-existant ou créer un nouveau mélange.

L'innovation méthodologique ensuite. Le fait de décentraliser le processus de sélection permet de faire profiter le groupe des nouvelles idées venant des fermes (acteurs). Par exemple, une paysanne a proposé d'ajouter sur les fiches les informations issues des fiches précédentes. La fiche d'été mentionne maintenant les notations prises sur la fiche d'hiver et la fiche de printemps (Figure 2). Cela a été pris en compte par l'équipe de recherche, qui a pu ensuite en faire profiter tous les paysans. Un questionnaire permet de demander aux paysans leurs avis sur le « dossier retour » et sur les fiches. Pour la saison 2012-2013, il en est ressorti que l'information devait être plus synthétisée. L'équipe de recherche a donc fait évoluer le dossier.

## 5. Conclusion et perspectives

Notre travail est le résultat d'une interaction, d'une co-construction entre paysans, associations et équipe de recherche. Ce travail est basé sur la décentralisation de la sélection qui permet la valorisation des interactions génotypes-environnements. Ce travail a permis de développer des outils et méthodes innovants de sélection et de gestion à la ferme de la biodiversité cultivée. Un travail est en cours pour caractériser la diversité génétique produite/gérée par le programme de SP et l'impact de ces pratiques sur la diversité. Des variétés-populations sont adoptées par certains participants. Le paysan (re)prend toute sa place dans les différentes étapes du processus de sélection. Une telle approche rend les paysans plus autonomes et renforce l'appartenance à un réseau. L'équilibre entre dépendance/indépendance du paysan au réseau est important. Tout en étant très libre dans ses choix sur sa ferme, son activité se replace dans un contexte plus global. Les innovations génétiques (i.e. les variétés-populations) et organisationnelles reposent sur la mise en réseau et la possibilité pour les acteurs de s'exprimer. Depuis cette année, les associations s'autonomisent, ce qui est nécessaire pour décentraliser la sélection et l'innovation au maximum.

La méthode que nous avons créée n'est pas une méthode universelle qu'il faut appliquer telle quelle dans un autre contexte. Cette méthodologie est adaptable à chaque contexte local et transposable sur d'autres espèces. Chaque groupe doit discuter et mettre au point ses propres méthodes. Il faut beaucoup de temps de discussion et de rencontre entre les différents acteurs pour mettre en place et faire perdurer ce type de programme (Storup et De Kochko, 2011).

C'est dans cette idée qu'ont été développés les différents outils (base de données et statistiques). Les outils statistiques utilisés permettent une grande souplesse dans la mise en place des dispositifs expérimentaux à condition que le nombre d'essais reste important et que les essais partagent le même dispositif expérimental. La base de données ShiNeMas permet également une grande souplesse. Les variables et les méthodes ne sont pas fixées, chaque groupe peut donc utiliser ShiNeMas selon ses besoins.

Tout ces outils ont été développés en langage libre et open source (PostgreSQL, python, Django, R, LaTeX). Cela permet une meilleure diffusion et une meilleure adaptabilité pour les spécificités locales. En ce moment, l'équipe de recherche améliore la convivialité de ces outils et méthodes ainsi que la portabilité des approches à d'autres groupes de producteurs et sur d'autres espèces.

Ces approches sont à replacer dans le concept des maisons de la semence. Les maisons de la semence sont des lieux qui permettent de conserver les semences, mais aussi d'organiser la gestion des ressources, des programmes de SP, de fournir un soutien technique, etc. Cela ouvre de nouvelles questions concernant la place des jardiniers, de l'espace dédié à la conservation et à la sélection chez les paysans.



## Remerciements

P. Rivière a été financé par une bourse DIM ASTREA (Région Ile de France). J. Dawson a été financée par une bourse postdoctorale INRA et pendant 5 mois par une bourse post-doctorale du European Community's Seventh Framework Programme (FP7/ 2007-2013) avec le Grant Agreement n245058-Solibam (Strategies for Organic and Low-input Integrated Breeding and Management). Les recherches qui ont conduit à ces résultats ont reçu un financement du programme FP7 Solibam mentionné ci-dessus.

Nous tenons à remercier les paysans qui ont participé plus ponctuellement à ce projet. Merci à L. Hazard et E. Demeulenaere pour leurs conseils sur la partie analyse de la méthodologie. Merci à F. Gauthier pour son aide dans la parallélisation des calculs sous condor. Merci enfin à l'équipe DEAP pour les nombreuses discussions enrichissantes autour de délicieux gâteaux.

## Références bibliographiques

- Ali Brac de la Perrière R., de Kochko P., Neubauer C., Storup B., 2011. Visions paysannes de la Recherche dans le Contexte de la Sélection Participative: Comment co-construire et mutualiser les connaissances sur les plantes? Pour l'Emergence D'une Université Du Vivant. Document essentiellement issu des travaux du séminaire « Retour d'expériences en sélection participative » qui s'est déroulé en février 2011 à Angers, dans le cadre du projet REPERE coordonné par la Fondation Sciences Citoyennes (<http://sciencescitoyennes.org/>)
- Annicchiarico P., Chiapparino E., Perenzin M., 2010. Response of Common Wheat Varieties to Organic and Conventional Production Systems Across Italian Locations, and Implications for Selection. *Field Crops Research* 116(3): 230–238.
- Bonneuil C., Demeulenaere E., 2007. Vers Une Génétique de pair à pair ? L'émergence de la sélection participative. *In Des sciences citoyennes? La question de l'amateur dans les sciences naturalistes*. Edition de l'Aude. pp.122-147.
- Bonneuil C., Goffaux R., Bonnin I., Montalent P., Hamon C., Balfourier F., Goldringer I., 2012. A new integrative indicator to assess crop genetic diversity. *Ecological Indicators* 23, 280–289.
- Ceccarelli S., Grando S., 2007. Decentralized-Participatory Plant Breeding: An example of demand driven research. *Euphytica* 155(3), 349-360.
- Ceccarelli S., Grando S., Tutwiler R., Baha J., Martini A.M., Salahieh H., Goodchild A., Michael M., 2000. A methodological study on participatory barley breeding I. Selection phase. *Euphytica* 111, 91-104.
- Chable V., Conseil M., Serpolay E., Le Lagadec F., 2008. « Organic varieties for cauliflowers and cabbages in Brittany: from genetic resources to participatory plant breeding ». *Euphytica* 164, 521-529.
- Chateil C., Goldringer I., Tarallo L., Kerbirou C., Le Viol I., Ponge J-F, Salmon S., Gachet S., Porcher E., 2013. Crop genetic diversity benefits farmland biodiversity in cultivated fields. *Agriculture, Ecosystems & Environment* 171, 25-32.
- Dawson J., Rivière P., Berthelot J.F., Mercier F., de Kochko P., Galic N., Pin S., Serpolay E., Thomas M., Giuliano S., Goldringer I., 2011. Collaborative plant breeding for organic agricultural systems in developed countries. *Sustainability* 3(8), 1206-1223.
- Desclaux D., Nolot J.M., Chiffolleau Y., Leclerc C., Gozé E., 2008. Changes in the concept of Genotype X Environment interactions to fit agriculture diversification and decentralized participatory plant breeding: A pluridisciplinary point of view. *Euphytica* 163: 533-546.
- Falconer D.S., 1952. The Problem of Environment and Selection. *American Naturalist*(86): 293–298.
- Finckh M.R., Gacek E.S, Goyeau H., Lannou C., Merz U., Mundt C.C., Munk L., Nadziak J., Newton A.C., de Vallavieille-Pope C., 2000. Cereal variety and species mixtures in practice, with emphasis on disease resistance. *Agronomie* 20, 813-837.
- Gauch H.G., Piepho H.S., Annicchiarico P., 2008. Statistical analysis of yield trials by AMMI and GGE: further considerations. *Crop Sci* 48, 866-889.

- Goldringer I., Brabant P., Gallais A., 1997. Estimation of Additive and Epistatic Genetic Variances for Agronomic Traits in a Population of Doubled-haploid Lines of Wheat. *Heredity* 79(1): 60–71.
- MEA, 2005. Ecosystems and Human Well-being : General Synthesis. Island Press. Washington DC, USA: Millennium Ecosystem Assessment.
- Mendes Moreira P.M.R., Pêgo S.E., Vaz Patto C., Hallauer A.R., 2008. Comparison of selection methods on “Pigarro”, a Portuguese improved maize population with fasciation expression. *Euphytica* 163, 481-499.
- Morris M.L., Bellon M.R., 2004. Participatory plant breeding research: opportunities and challenges for the international crop improvement system. *Euphytica* 136, 21-35.
- Murphy K.M., Campbell K.G., Lyon S.R., Jones S.S., 2007. Evidence of Varietal Adaptation to Organic Farming Systems. *Field Crops Research* 102(3): 172–177.
- Nabugoomu F., Kempton R.A., Talbot M., 1999. Analysis of series of trials where varieties differ in sensitivity to locations. *Journal of Agricultural, Biological and Environmental Statistics* 4, 310–325.
- Papaix J., Goyeau H., Du Cheyron P., Monod H., Lannou C., 2011. Influence of Cultivated Landscape Composition on Variety Resistance: An Assessment Based on Wheat Leaf Rust Epidemics. *New Phytologist* 191(4): 1095–1107.
- Réseau Semences Paysannes, 2008. Voyage autour des blés paysans. Témoignages.
- Rivière P., Goldringer I., Berthelot J-F., galic N., Pin S., de Kochko P., Dawson J.C., 2013. Response to Farmer Mass Selection in Early Generation Progeny of Bread Wheat Landrace Crosses. *Renewable Agriculture and Food Systems FirstView*: 1–12.
- Robert C.P., 2001. The Bayesian choice. Springer Texts in Statistics. 2de édition. Springer ed.
- Rolland B., Le Champion A., Oury F-X., 2012. Pourquoi Sélectionner de Nouvelles Variétés de Blé Tendre Adaptées à L’agriculture Biologique ? *Courrier de L’environnement de l’INRA*(62): 71–85.
- Roussel V., Branlard G., Vezine J.C., Bertrand D., Balfourier F., 2005. NIRS analysis reveals temporal trends in the chemical composition of French bread wheat accessions cultivated between 1800 and 2000. *J Cereal Sci.* 42, 193-203.
- Sadras V.O., Slafer G.A., 2012. Environmental Modulation of Yield Components in Cereals: eritabilities Reveal a Hierarchy of Phenotypic Plasticities. *Field Crops Research* 127: 215–224.
- Smith M.E, Fernando Castillo G., Gomez F., 2001. Participatory plant breeding with maize in Mexico and Honduras. *Euphytica* 122, 551-565.
- Storup B., de Kochko P., 2011. Suivi d’une expérience de sélection participative en cours : L’exemple du Projet Européen SOLIBAM. Association pour une Fondation Science Citoyenne.
- Thomas M., Dawson J.C., Goldringer I., Bonneuil C., 2011. Seed exchanges, a key to analyze crop diversity dynamics in farmer-led on-farm conservation. *Genetic Resources and Crop Evolution* 58, 321-338.
- Tooker J.-F., Steven D.-F., 2012. Genotypically diverse cultivar mixtures for insect pest management and increased crop yields. *Journal of Applied Ecology* 49, 974-985.
- Virk D.S., Chakraborty M., Ghosh J., Prasad S.C., Witcombe J.R., 2005. Increasing the client orientation of maize breeding using farmer participation in eastern India. *Experimental Agriculture* 41, 413-426.
- Wolfe M., Baresel J., Desclaux D., Goldringer I., Hoad S., Kovacs G., Löschenberger F., Miedaner T., Ostergard H., Lammerts van Bueren E., 2008. Developments in breeding cereals for organic agriculture. *Euphytica* 163, 323-346.

# Annexe N : Les statistiques bayésiennes

## N.1 Modélisation bayésienne

Soit le modèle linéaire suivant :

$$y = X\theta + E$$

Avec,  $y$  un vecteur de  $n$  observations :  $y = (y_1 + y_2 + \dots + y_i + \dots + y_n)$ ; chaque  $y_i$  est la réalisation d'une variable aléatoire  $Y_i$  normale centrée en  $(X\theta)$  et de variance  $\sigma^2$  :  $Y_i \sim N((X\theta)_i, \sigma^2)$ . Les  $n$  variables aléatoires sont indépendantes.

$X$ , la matrice de dimension  $n \times k$ , contenant les valeurs observées de  $k$  variables explicatives disposées en colonnes ;

$\theta$ , un vecteur de  $k$  paramètres :  $\theta = (\theta_1, \theta_2, \dots, \theta_k)$  ;

$E$ , le vecteur des erreurs du modèle de longueur  $n$ , qui suit une loi normale centrée sur 0 :  $E \sim N_n(0, \sigma^2 I_n)$  avec  $I_n$  la matrice identité de dimension  $n \times n$ .

Pour estimer les paramètres du vecteur  $\theta$ , deux approches sont possibles : fréquentiste et bayésienne.

En statistique fréquentiste, les paramètres du modèle sont considérés comme des quantités fixes et inconnues. On cherche à maximiser la vraisemblance  $Pr(y|\theta)$ , c'est à dire à trouver les valeurs des paramètres qui permettent de coller au mieux aux données. Cela revient à résoudre un système avec  $k$  équations qui correspondent à la dérivé de la fonction de vraisemblance par rapport à chaque paramètre de  $\theta$ <sup>i</sup>. La dérivé seconde est également utilisée pour vérifier s'il s'agit d'un maximum ou d'un minimum. Ce qui ne pourra être expliqué par  $\theta$  sera dans la résiduelle  $E$ .

En statistique bayésienne, les paramètres sont considérés comme aléatoires, c'est à

---

i. En pratique, on passe par le *log* pour s'affranchir des produits et utiliser des sommes. On parle de log-vraisemblance.

dire qu'ils ont une certaine densité de probabilité. La philosophie des statistiques bayésienne est d'estimer la densité de probabilité (i.e. la distribution) de chaque paramètre. L'inférence bayésienne apporte une analyse probabiliste directe. Elle donne directement accès aux distributions des paramètres, c'est à dire la probabilité associée à la valeur d'un paramètre. Pour un paramètre donné, cette distribution est appelée distribution a posteriori ou posterior car elle résulte de deux types d'informations :

- une distribution *a priori* sur le paramètre : c'est le prior
- l'information issue des données : c'est la vraisemblance

C'est le théorème de Bayes qui fait le lien entre le posterior, le prior et la vraisemblance :

$$Pr(\theta|y) = \frac{Pr(y|\theta)Pr(\theta)}{Pr(y)}.$$

Où  $Pr(\theta|y)$  correspond à la distribution de paramètres du vecteur  $\theta$  sachant les données : c'est le posterior,  $Pr(y|\theta)$  est l'information issue de données : c'est la vraisemblance,  $Pr(\theta)$  est la distribution a priori des paramètres. Comme  $Pr(y)$  est une constante par rapport à  $\theta$ , il vient :

$$Pr(\theta|y) \propto Pr(y|\theta)Pr(\theta).$$

La distribution a posteriori dépend du produit de la vraisemblance et de la distribution a priori. Si le prior  $Pr(\theta)$  est non informatif, le posterior sera proche de la vraisemblance. Si les données ne sont pas informatives, le posterior sera proche du prior. Plus il y a de données et plus la vraisemblance aura du poids dans l'estimation du posterior par rapport au prior.

## N.2 Calcul de la distribution a posteriori

L'inférence bayésienne passe par des calculs d'intégrales souvent complexes. On peut faire ces calculs de manière explicite lorsque le prior et la vraisemblance sont conjugués : le prior et le posterior appartiennent alors à une même famille de distributions.

Par exemple, si les données suivent une loi normale de moyenne  $\mu$  inconnue et de variance  $\sigma^2$  connue, une loi a priori normale pour  $\mu$  est conjuguée avec la vraisemblance. La loi a posteriori de  $\mu$  est également normale et se calcule de manière explicite.

Dans le cas général, on ne sait pas calculer la loi a posteriori de manière explicite. Pour s'affranchir de cette difficulté, on peut estimer des résumés de cette distribution à

l'aide de la méthode Monte Carlo par chaînes de Markov (MCMC). La méthode Monte Carlo permet d'estimer l'espérance a posteriori d'une fonction  $h$  de  $\theta$ ,  $E(h(\theta)|y)$ , à l'aide d'un échantillon de la loi a posteriori  $Pr(\theta|y)$ . La difficulté vient de la simulation de cet échantillon car la loi a posteriori n'est en général pas standard, elle n'est en général connue qu'à une constante près et  $\theta$  peut être de grande dimension. Pour cela, on utilise des chaînes de Markov dont la loi stationnaire est la loi a posteriori et qui convergent vers cette loi stationnaire. On rappelle que les chaînes de Markov sont des processus stochastiques où les futurs états sont indépendants des états passés conditionnellement à l'état présent. Cela revient à écrire :  $Pr(\theta^{(t+1)}|\theta^{(1)}, \theta^{(2)}, \theta^{(3)}, \dots, \theta^{(t)}) = Pr(\theta^{(t+1)}|\theta^{(t)})$ .

Dans notre cas, pour estimer les distributions a posteriori, on utilise **JAGS** qui repose sur l'échantillonneur de Gibbs. Cet échantillonneur permet de simuler des réalisations des paramètres  $\theta$  selon la loi a posteriori  $Pr(\theta|y)$  à l'itération  $t$  à partir de leurs valeurs à l'itération  $t - 1$  (chaînes de Markov).

L'échantillonneur de Gibbs fonctionne comme suit. Prenons un exemple où le vecteur  $\theta$  qui regroupe trois paramètres :  $\theta = (\theta_1, \theta_2, \theta_3)$  que l'on cherche à estimer. Dans notre cas, on fait 100 000 simulations.

- **Simulation 1** : On simule  $\theta^{(1)}$  en partant d'un vecteur de valeurs initiales :  $\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \theta_3^{(0)})$ .
  1. On simule une valeur de  $\theta_1^{(1)}$  selon la loi conditionnelle  $Pr(\theta_1|\theta_2^{(0)}, \theta_3^{(0)}, y)$ .
  2. On simule une valeur de  $\theta_2^{(1)}$  selon la loi conditionnelle  $Pr(\theta_2|\theta_1^{(1)}, \theta_3^{(0)}, y)$ . On utilise la valeur nouvellement créée  $\theta_1^{(1)}$ .
  3. On simule une valeur de  $\theta_3^{(1)}$  selon la loi conditionnelle  $Pr(\theta_3|\theta_1^{(1)}, \theta_2^{(1)}, y)$ . On utilise la valeur nouvellement créée  $\theta_2^{(1)}$ .
- **Simulation 2** : On simule  $\theta^{(2)}$  en utilisant  $\theta^{(1)}$  avec  $\theta^{(1)} = (\theta_1^{(1)}, \theta_2^{(1)}, \theta_3^{(1)})$ .
  1. On simule une valeur de  $\theta_1^{(2)}$  selon la loi conditionnelle  $Pr(\theta_1|\theta_2^{(1)}, \theta_3^{(1)}, y)$ .
  2. On simule une valeur de  $\theta_2^{(2)}$  selon la loi conditionnelle  $Pr(\theta_2|\theta_1^{(2)}, \theta_3^{(1)}, y)$ .
  3. On simule une valeur de  $\theta_3^{(2)}$  selon la loi conditionnelle  $Pr(\theta_3|\theta_1^{(2)}, \theta_2^{(2)}, y)$ .
- **Simulation 3** : On simule  $\theta^{(3)}$  en utilisant  $\theta^{(2)}$  avec  $\theta^{(2)} = (\theta_1^{(2)}, \theta_2^{(2)}, \theta_3^{(2)})$ ; on reproduit les étapes 1., 2. et 3.
- **Simulations ...**
- **Simulation 100 000** : On simule  $\theta^{(100\ 000)}$  en utilisant  $\theta^{(99\ 999)}$  avec  $\theta^{(99\ 999)} = (\theta_1^{(99\ 999)}, \theta_2^{(99\ 999)}, \theta_3^{(99\ 999)})$ ; on reproduit les étapes 1., 2. et 3.

Chaque itération va simuler une valeur pour chaque paramètre qui est stockée dans un vecteur. Par exemple, le paramètre  $\theta_1$  a 100 000 valeurs issues des 100 000 simulations

MCMC stockées dans un vecteur. Ce vecteur est un échantillon de la loi a posteriori du paramètre.

Pour utiliser l'algorithme de Gibbs, on a besoin de connaître et de pouvoir simuler facilement la loi d'un paramètre conditionnellement aux autres paramètres et aux données (par exemple  $Pr(\theta_1|\theta_2, \theta_3, y)$ ). C'est le cas lorsque les lois a priori présentent des propriétés de conjugaison. Cependant, dans la pratique, il se peut que les lois ne montrent pas de propriétés de conjugaison. Il existe alors des « aménagements » qui permettent d'utiliser l'algorithme de Gibbs dans ce cas, mais qui peuvent dégrader ses performances. En pratique, cet algorithme converge très lentement pour certains modèles.

Les valeurs successives d'une chaîne MCMC sont corrélées. Lorsque cette corrélation est importante, les valeurs simulées à deux itérations successives apportent presque la même information. Pour limiter la possible dépendance entre les éléments de la chaîne, on ne garde qu'un élément sur  $k$ .  $k$  correspond au « thinning ». Dans notre cas, on prend  $k = 10$ , ce qui permet de réduire la corrélation entre les éléments de la chaîne. Sur le vecteur initial de 100 000 éléments, on n'en garde que 10 000.

Selon la théorie des chaînes de Markov, il y aura à terme convergence de la chaîne MCMC vers la loi a posteriori des paramètres. En pratique, pour estimer la loi a posteriori, on s'affranchit des premières valeurs de la simulation MCMC pour ne conserver que les points les plus proches de la phase stationnaire.

Une chaîne peut se stabiliser vers des valeurs « fausses ». Afin de limiter ce risque, on simule plusieurs chaînes MCMC. Si les chaînes convergent vers la même loi stationnaire, on est plus confiant dans le résultat. Il existe des tests pour étudier la convergence des chaînes. Dans notre cas, on utilise le test de Gelman-Rubin basé sur la variance à l'intérieur et entre les chaînes.

A partir de l'échantillon simulé par MCMC pour un paramètre, on peut estimer divers résumés de la loi a posteriori : par exemple l'espérance a posteriori, la variance a posteriori, la probabilité a posteriori que ce paramètre appartienne à un intervalle donné.

# Annexe O : Décomposition des covariances entre individus dans une population consanguine

Cete partie est issue de Falconer et Mackay<sup>86</sup>; Gallais<sup>111</sup>; Goldringer<sup>118</sup>.

Les équations suivantes sont utilisées dans la suite.

$$\text{var}(X) = E(X^2) - E(X)^2 \quad (\text{O.1})$$

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2\text{cov}(X, Y) \quad (\text{O.2})$$

$$E(aX + bY) = aE(X) + bE(Y) \quad (\text{O.3})$$

## O.1 Le modèle génétique en population panmictique

On se place dans une population panmictique sous l'équilibre d'Hardy Weigber. La valeur génotypique d'un individu comportant les allèles  $i$  et  $j$  se décompose comme suit, c'est le modèle de Fisher<sup>101</sup> :

$$G_{ij} = \mu + \alpha_i + \alpha_j + \beta_{ij} \quad (\text{O.4})$$

avec  $\mu$  la moyenne générale de la population

$\alpha_i$  l'effet aléatoire du gamète  $i$ , appelé aussi effet additif du gène  $i$

$\alpha_j$  l'effet aléatoire du gamète  $j$ , appelé aussi effet additif du gène  $j$

$\beta_{ij}$  l'épistasie c'est à dire l'interaction entre le gène  $i$  et le gène  $j$ , appelé aussi l'effet de dominance entre  $i$  et  $j$

Ce modèle à plusieurs contraintes :

- indépendance de  $\alpha$  et  $\beta$
- $E_{ij}(G_{ij}) = 0$
- $E_i(\alpha_i) = 0$
- $E_{ij}(\beta_{ij}) = E_i(\beta_{ij}) = E_j(\beta_{ij}) = 0$

On peut écrire la variance génotypique comme :

$$\sigma_G^2 = var(\mu + \alpha_i + \alpha_j + \beta_{ij})$$

d'après les contraintes du modèle (indépendance des paramètres) et O.2, on a

$$\sigma_G^2 = var(\mu) + var(\alpha_i) + var(\alpha_j) + var(\beta_{ij})$$

$var(\mu) = 0$  car  $\mu$  est une constante

$$var(\alpha_i) = E(\alpha_i^2) - \underbrace{E(\alpha_i)^2}_{=0 \text{ d'après les contraintes}} = E(\alpha_i^2)$$

Comme on est en panmixie, il y a peu de chance d'avoir  $i = j$ , leurs covariances sont nulles. On fait également l'hypothèse que les allèle  $i$  et  $j$  sont identiques par descendance<sup>i</sup> (On verra dans la partie O.3 que cette hypothèse peut être pondérée par une probabilité). Il vient alors :

$$var(\alpha_i) + var(\alpha_j) = 2E(\alpha_i^2)$$

$$var(\beta_{ij}) = E(\beta_{ij}^2) - \underbrace{E(\beta_{ij})^2}_{=0 \text{ d'après les contraintes}}$$

On en déduit

$$\sigma_G^2 = 2E(\alpha_i^2) + E(\beta_{ij}^2)$$

Soit la variance additive  $\sigma_A^2$  tel que :  $\sigma_A^2 = 2E(\alpha_i^2)$

Soit la variance de dominance  $\sigma_D$  tel que :  $\sigma_D^2 = E(\beta_{ij}^2)$

ce qui donne :

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2 \tag{O.5}$$

à noter que ce modèle est vrai quand il n'y a pas d'effet d'épistasie.

---

i. désigne deux copies d'un même gène ancêtre.



## O.2 Le modèle génétique en population consanguine : effet du régime de reproduction

La consanguinité c'est la possibilité de retrouver chez un individu dit consanguin, à un locus, deux gènes dérivant sans mutation d'un ancêtre commun. La consanguinité est d'autant plus importante que la relation entre les parents est étroite.

Soit  $F$  la probabilité que deux individus aient un gène identique et  $(1 - F)$  sinon. On sépare la moyenne de la population consanguine ( $\mu_F$ ) en moyenne d'individus hétérozygotes ( $(1 - F)\mu$ ) et en moyenne d'individus homozygotes, c'est à dire des lignées ( $F\mu_L$ ) :

$$\mu_F = (1 - F)\mu + F\mu_L \tag{O.6}$$

de même, en reprenant le modèle O.4, la valeur d'un génotype homozygote s'écrit :

$$G_{ij} = \mu + \alpha_i + \alpha_j + \beta_{ii}$$

Ici, d'après le modèle O.4 :  $j = i$ , les gènes pris sont identiques, leurs covariances est de 1, on écrit alors :

$$G_{ii} = \mu + 2\alpha_i + \beta_{ii} \tag{O.7}$$

Ce point est important car il justifie par la suite que les valeurs des variances additives en lignées sont deux fois supérieures à celles des valeurs additives en panmixie.

Regardons plus en détails O.6.

$\mu_L$  est la valeur des lignées, c'est à dire la moyenne des  $G_{ij}$

$$\begin{aligned} \mu_L &= E(G_{ii}) \\ \mu_L &= E(\mu + 2\alpha_i + \beta_{ii}) \text{ d'après le modèle O.7 et d'après O.3} \\ \mu_L &= E(\mu) + 2 \underbrace{E(\alpha_i)}_{=0 \text{ d'après les contraintes}} + E(\beta_{ii}) \\ \mu_L &= \mu + E(\beta_{ii}) \end{aligned}$$

d'après O.6, on en déduit :

$$\begin{aligned} \mu_F &= (1 - F)\mu + F(\mu + E(\beta_{ii})) \\ \mu_F &= \mu - F\mu + F\mu + FE(\beta_{ii}) \\ \mu_F &= \mu + FE(\beta_{ii}) \end{aligned}$$

De même que pour O.6, on décompose la variance génétique totale de la population panmictique en individus homozygote ( $ii$ ) et hétérozygotes ( $ij$ )

$$\sigma_{GF}^2 = E(G_{ij}^2) - E(G_{ii})^2$$

On en déduit :

$$\sigma_{GF}^2 = (1 - F)E(G_{ij}^2) + FE(G_{ii}^2) - \mu_F^2 \quad (O.8)$$

On remarque que  $F$  est bien relié à  $G_{ii}$  qui correspond à un individu homozygote  $ii$

De même

$$\begin{aligned} \sigma_{GL}^2 &= E(G_{ii}^2) - E(G_{ii})^2 \\ \sigma_{GL}^2 &= E(G_{ii}^2) - \mu_L^2 \end{aligned}$$

et

$$\begin{aligned} \sigma_G^2 &= E(G_{ij}^2) - E(G_{ij})^2 \\ \sigma_G^2 &= E(G_{ij}^2) - \mu^2 \end{aligned}$$

On en déduit pour O.8, en remplaçant par les valeurs des espérances  $E$  :

$$\sigma_{GF}^2 = (1 - F)(\sigma_G^2 + \mu^2) + F(\sigma_{GL}^2 + \mu_L^2) - \mu_F^2$$

puis, d'après O.6

$$\begin{aligned} \sigma_{GF}^2 &= (1 - F)(\sigma_G^2 + \mu^2) + F(\sigma_{GL}^2 + \mu_L^2) - (\mu + F\mu_L - F\mu)^2 \\ \sigma_{GF}^2 &= (1 - F)\sigma_G^2 + (1 - F)\mu^2 + F\sigma_{GL}^2 + F\mu_L^2 - (\mu + F(\mu_L - \mu))^2 \end{aligned}$$

Pour simplifier l'écriture, on pose  $Z = (1 - F)\mu^2 + F\mu_L^2 - (\mu + F(\mu_L - \mu))^2$ , On a alors :

$$\begin{aligned} \sigma_{GF}^2 &= (1 - F)\sigma_G^2 + F\sigma_{GL}^2 + Z \\ Z &= (1 - F)\mu^2 + F\mu_L^2 - (\mu + F(\mu_L - \mu))^2 \\ Z &= \mu^2 - F\mu^2 + F\mu_L^2 - (\mu^2 + 2F\mu(\mu_L - \mu) + (F(\mu_L - \mu))^2) \\ Z &= \mu^2 - F\mu^2 + F\mu_L^2 - \mu^2 - 2F\mu(\mu_L - \mu) - (F(\mu_L - \mu))^2 \\ Z &= -F\mu^2 + F\mu_L^2 - 2F\mu(\mu_L - \mu) - F^2(\mu_L - \mu)^2 \\ Z &= F(-\mu^2 + \mu_L^2 - 2\mu(\mu_L - \mu)) - F^2(\mu_L - \mu)^2 \\ Z &= F(-\mu^2 + \mu_L^2 - 2\mu\mu_L + 2\mu^2) - F^2(\mu_L - \mu)^2 \\ Z &= F(\mu_L^2 - 2\mu\mu_L + \mu^2) - F^2(\mu_L - \mu)^2 \\ Z &= F(\mu_L - \mu)^2 - F^2(\mu_L - \mu)^2 \\ Z &= (F - F^2)(\mu_L - \mu)^2 \\ Z &= F(1 - F)(\mu_L - \mu)^2 \end{aligned}$$

D'où

$$\sigma_{GF}^2 = (1 - F)\sigma_G^2 + F\sigma_{GL}^2 + F(1 - F)(\mu_L - \mu)^2$$

Nous avons vu déjà précédemment que

$$\mu_F = \mu + F(\mu_L - \mu) = \mu + FE(\beta_{ii})$$

Pour un caractère sous contrôle de plusieurs locus  $l$ , on pose  $D_{0,l} = E(\beta_{ii})$ , la dépression de consanguinité pour le locus  $l$ , on a alors

$$\mu_F = \mu + F \sum_l D_{0,l}$$

le modèle O.8 devient alors

$$\sigma_{GF}^2 = (1 - F)\sigma_G^2 + F\sigma_{GL}^2 + F(1 - F) \sum_l D_{0,l}^2 \quad (\text{O.9})$$

On exprime donc la variance génétique totale comme la composante de la variance génétique de la population en panmixie (individus hétérozygotes), la variance génétique de la population fixée (individus homozygotes) et le carré de la dépression de consanguinité.

Pour aller plus loin dans l'exploration de la variance génétique totale, intéressons nous à  $\sigma_{GL}^2$ , la variance génétique de la population fixée.

On a vu avec O.7 que  $G_{ii} = \mu + 2\alpha_i + \beta_{ii}$  et que  $\mu_L = E(G_{ii}) = \mu + E(\beta_{ii})$

$$\begin{aligned} \sigma_{GL}^2 &= E(G_{ii}^2) && - E(G_{ii})^2 \\ \sigma_{GL}^2 &= E((\mu + 2\alpha_i + \beta_{ii})^2) && - E(\mu + E(\beta_{ii}))^2 \\ \sigma_{GL}^2 &= E(4\alpha_i^2 + 2\alpha_i\beta_{ii} + \beta_{ii}^2) && - E(\beta_{ii})^2 \\ \sigma_{GL}^2 &= 4E(\alpha_i^2) + 2E(\alpha_i\beta_{ii}) + E(\beta_{ii}^2) && - E(\beta_{ii})^2 \end{aligned}$$

On pose

$$4E(\alpha_i^2) = 2\sigma_A^2 \quad \text{la variance additive}$$

$$E(\alpha_i\beta_{ii}) = \sigma_{AD_O} \quad \text{la covariance entre les effets d'additivité et de dominance à l'état homozygote}$$

$$E(\beta_{ii}^2) - E(\beta_{ii})^2 = \sigma_{D_O}^2 \quad \text{la variance des résidus de dominance entre gènes identiques}$$

On a alors

$$\sigma_{GL}^2 = 2\sigma_A^2 + 2\sigma_{AD_O} + \sigma_{D_O}^2 \quad (\text{O.10})$$

A partir de l'expression O.10 et O.5, on peut réécrire le modèle O.9.

$$\begin{aligned}
 \sigma_{GF}^2 &= (1 - F)\sigma_G^2 && + F\sigma_{GL}^2 && + F(1 - F)\sum_l D_{0,l}^2 \\
 \sigma_{GF}^2 &= (1 - F)(\sigma_A^2 + \sigma_D^2) && + F(2\sigma_A^2 + 2\sigma_{AD_O} + \sigma_{D_O}^2) && + F(1 - F)\sum_l D_{0,l}^2 \\
 \sigma_{GF}^2 &= (1 - F)\sigma_A^2 + (1 - F)\sigma_D^2 && + 2F\sigma_A^2 + 2F\sigma_{AD_O} + F\sigma_{D_O}^2 && + F(1 - F)\sum_l D_{0,l}^2 \\
 \sigma_{GF}^2 &= (2F + 1 - F)\sigma_A^2 + (1 - F)\sigma_D^2 + 2F\sigma_{AD_O} + F\sigma_{D_O}^2 && + F(1 - F)\sum_l D_{0,l}^2
 \end{aligned}$$

$$\sigma_{GF}^2 = (1 + F)\sigma_A^2 + (1 - F)\sigma_D^2 + 2F\sigma_{AD_O} + F\sigma_{D_O}^2 + F(1 - F)\sum_l D_{0,l}^2 \quad (O.11)$$

L'expression de la variance génétique totale se décompose en :

- $(1 + F)\sigma_A^2$  la variance génétique additive
- $(1 - F)\sigma_D^2$  la variance génétique de dominance
- $2F\sigma_{AD_O}$  la covariance entre les effets d'additivité et de dominance à l'état homozygote
- $F\sigma_{D_O}^2$  la variance des résidus de dominance entre gènes identiques
- $F(1 - F)\sum_l D_{0,l}^2$  la somme des carrés la dépression de consanguinité pour le locus  $l$

Ces termes sont en fonction de  $F$ , la probabilité que deux individus aient un gène identique.

### 0.3 Décomposition de la covariance entre deux individus d'une population consanguine

Nous voulons modéliser la relation qu'il existe entre deux individus  $X$  et  $Y$  dans une population.

Cela revient à mesurer les covariances entre apparentés.

$$cov(X, Y) = cov(P_X, P_Y)$$

avec

$$\begin{aligned}
 P_X &= G_X + E_X \\
 P_Y &= G_Y + E_Y
 \end{aligned}$$

avec  $G$  l'effet lié au génétique et  $E$  l'effet lié à l'environnement. Soit la covariance entre  $X$  et  $Y$ , telle que :

$$\begin{aligned} \text{cov}(X, Y) &= \text{cov}(G_X + E_X, G_Y + E_Y) \\ \text{cov}(X, Y) &= \text{cov}(E_X + E_Y) + \text{cov}(G_X + G_Y) \end{aligned}$$

Avec comme hypothèse que les environnements pour  $X$  et  $Y$  sont indépendants (randomisation de l'expérience),  $\text{cov}(E_X + E_Y) = 0$ . Sachant que l'effet génétique  $G$  est la somme des effets additifs  $A$  et de dominances  $D$ .

$$G_X = A_X + D_X$$

$A$  et  $D$  ne sont pas corrélés (i.e. covariances nulles, cf partie O.1), il vient donc :

$$\text{cov}(X, Y) = \text{cov}(A_X, A_Y) + \text{cov}(D_X, D_Y)$$

Si on reprend le modèle génétique O.4, posons :

$$X : G_{ij} = \alpha_i + \alpha_j + \beta_{ij} \text{ et } Y : G_{i'j'} = \alpha_{i'} + \alpha_{j'} + \beta_{i'j'}$$

On a

$$\begin{aligned} \text{cov}(X, Y) &= \text{cov}(\alpha_i + \alpha_j + \beta_{ij}, \alpha_{i'} + \alpha_{j'} + \beta_{i'j'}) \\ \text{cov}(X, Y) &= E(\alpha_i \alpha_{i'}) + E(\alpha_i \alpha_{j'}) + E(\alpha_j \alpha_{j'}) + E(\alpha_j \alpha_{i'}) + E(\beta_{ij} \beta_{i'j'}) \\ &\quad + \text{termes de covariances nuls} \end{aligned}$$

Si on fait l'hypothèse que les allèles  $i$  et  $j$  sont identiques par descendance, il vient :

$$\begin{aligned} \text{cov}(X, Y) &= 4E(\alpha_i \alpha_{i'}) + E(\beta_{ij} \beta_{i'j'}) \\ \text{cov}(X, Y) &= 4E(\alpha_i^2) + E(\beta_{ij}^2) \end{aligned}$$

Soit la variance additive  $\sigma_A^2$  tel que :  $\sigma_A^2 = 2E(\alpha_i^2)$

Soit la variance de dominance  $\sigma_D^2$  tel que :  $\sigma_D^2 = E(\beta_{ij}^2)$ , il vient alors :

$$\text{cov}(X, Y) = 2\sigma_A^2 + \sigma_D^2$$

On remarque que la partie additive est doublée par rapport à ce qui a été démontré pour un individu (partie O.1) : deux individus apportent deux fois plus d'information sur les effets additifs (les  $\alpha$ ) alors que les effets de dominance (les  $\beta$ ) ne changent pas du fait de l'indépendance des effets d'interactions qui amènent des covariances nulles.

Il convient de pondérer cette hypothèse par des probabilités.

Soit  $\phi_A$ , le coefficient de simple parenté ; c'est la probabilité pour qu'à un locus donné, un gène de l'un des individus soit identique à un gène de l'autre. C'est à dire que ces deux

gènes dérivent par réplication (copie mendélienne) d'un même gène ancêtre. Si  $\phi_A = 0$  alors  $cov(A_X, A_Y) = 0$ .

Soit  $\phi_D$ , le coefficient de double parenté ; c'est la probabilité que  $X$  et  $Y$  aient reçus via leurs parents respectifs le même couple de gènes, c'est à dire que l'on tire le même couple de gène dans  $X$  et  $Y$ . Si  $\phi_D = 0$  alors  $cov(D_X, D_Y) = 0$ .

Pour une population non consanguine on a alors :

$$cov(X, Y) = 2\phi_{A(X,Y)}\sigma_A^2 + \phi_{D(X,Y)}\sigma_D^2 \tag{O.12}$$

Lorsqu'on considère un couple d'individus diploïdes et consanguins, sept états d'identité sont possibles, on les représente tel que :

$$\underbrace{(gene1gene2)}_{\text{individu 1}} \mid \underbrace{(gene3gene4)}_{\text{individu 2}}$$

On note  $i, j, k, l$ , quatre classes d'identités (ensemble des gènes identiques provenant d'un même ancêtre), deux lettres différentes sont des gènes non identiques.

Etat	combinaison	description
(1)	$(ij kl)$	aucun gène en commun
(2)	$(ii kl)$	aucun gène en commun et le même gène pour un des individus
(3)	$(ii jj)$	aucun gène en commun et le même gène pour chaque individu
(4)	$(ij ik)$	un gène en commun
(5)	$(ii ij)$	un gène en commun et le même gène pour un des individus
(6)	$(ij ij)$	deux gènes en commun
(7)	$(ii ii)$	deux gènes en commun et le même gène pour chaque individu

Les combinaisons (1) et (2) génèrent une covariance nulle entre les deux individus. Les autres états d'identité permettent de définir des coefficients d'apparentement à partir des probabilités associées à ces états. On pose  $\phi_1$ , la probabilité de tirage des gènes à un locus, à droite du trait (|) chez l'individu 1 et à gauche chez l'individu 2.  $\phi_{i|i}$  est le coefficient de simple apparentement et  $\phi_{ii|ii}$  est le coefficient de double apparentement déjà défini dans une population non consanguine.

Etat	Combinaison	Probabilité d'appartenance
(3)	( <i>ii</i>   <i>jj</i> )	$\phi_{ii jj}$
(4)	( <i>ij</i>   <i>ik</i> )	$\phi_{i i}$
(5)	( <i>ii</i>   <i>ij</i> )	$\phi_{ii i}$
(6)	( <i>ij</i>   <i>ij</i> )	$\phi_{ij ij}$
(7)	( <i>ii</i>   <i>ii</i> )	$\phi_{ii ii}$

En l'absence d'épistasie, la covariance entre  $X$  et  $Y$  sachant que  $X$  et  $Y$  sont consanguins s'écrit à partir des effets génétiques définis dans la population de référence<sup>118</sup>.

$$\begin{aligned} cov(X, Y) = & 4\phi_{i|i}E(\alpha_i^2) + 4\phi_{i|ii}E(\alpha_i\beta_{ii}) + \phi_{ij|ij}E(\beta_{ij}^2) + \phi_{ii|ii}E(\beta_{ii}^2) \\ & + (\phi_{ii|jj} - F_X F_Y)(E(\beta_{ii}^2))^2 \end{aligned} \quad (O.13)$$

$$\begin{aligned} cov(X, Y) = & 2\phi_{i|i}\sigma_A^2 + 4\phi_{i|ii}\sigma_{AD_0}^2 + \phi_{ij|ij}\sigma_D^2 + \phi_{ii|ii}\sigma_{D_0}^2 \\ & + (\phi_{ii|jj} + \phi_{ii|ii} - F_X F_Y) \sum_l D_{0,l}^2 \end{aligned} \quad (O.14)$$

Le fait que l'on soit dans une population consanguine a une influence sur les coefficients d'apparentements.

Soit :

$X$  un individu de génération  $g$  et issu de fécondation libre (état hétérozygote).

$Y$  un individu de génération  $g'$  et issu d'autofécondation (état homozygote).

$T$  un individu de génération  $t$  qui est le plus proche ancêtre commun entre  $X$  et  $Y$ .

On note  $cov(X, Y) = c_{tgg'}$

Le coefficient de consanguinité  $F_g$  à la génération  $g$  s'écrit  $F_g = 1 - (\frac{1}{2})^g$ .

En considérant qu'il n'y a pas d'épistasie et un déséquilibre de liaison entre les allèles nuls, la covariance entre  $X$  et  $Y$  peut s'écrire en fonction des coefficient de consanguinité  $F_g$ ,  $F_{g'}$  et  $F_t$ <sup>118</sup> :

$$\begin{aligned} c_{tgg'} = & (1 + F_t)\sigma_A^2 + \frac{(F_g + F_{g'} + 2F_t)}{2}\sigma_{AD_0}^2 + \frac{(1 - F_g)(1 - F_{g'})}{(1 - F_t)}\sigma_D^2 + \\ & \left[ F_t + \frac{(F_g - F_t)(F_{g'} - F_t)}{2(1 - F_t)} \right] \sigma_{D_0}^2 + F_t \left[ \frac{(1 - F_g)(1 - F_{g'})}{(1 - F_t)} \right] \sum_l D_{0,l}^2 \end{aligned} \quad (O.15)$$

A partir de l'expression de la covariance, on peut retrouver les différentes composantes de la variance génétique par régression entre  $X$  et  $Y$ , ou en analysant des structures en famille avec des analyses de variances.





# Annexe P : Marqueurs utilisés dans l'analyse moléculaire

## P.1 Marqueurs dans des zones neutres

Tableau P.1 - Marqueurs neutres (*NE*) utilisés dans les analyses génétiques.

Tableau P.1 – (partie 1/1)

nom des marqueur	ch	références	nom des marqueurs	ch	références
NE_143_1	3A	9K	NE_1460_15	5B	9K
NE_1644_49	1A	9K	NE_182_2	7B	9K
NE_210_3	3B	9K	NE_2328_17	NA	9K
NE_2353_18	7B	9K	NE_3044_79	5B	9K
NE_3630_24	5B	9K	NE_4465_58	5A	9K
NE_4662_51	4B	9K	NE_4816_27	1A	9K
NE_4929_28	6A	9K	NE_4961_29	6A	9K
NE_5071_86	7B	9K	NE_5083_30	1A	9K
NE_513_4	1A	9K	NE_5228_60	1B	9K
NE_5666_52	6B	9K	NE_578_5	1A	9K
NE_5904_36	7A	9K	NE_5912_59	7A	9K
NE_597_6	6B	9K	NE_605_7	1A	9K
NE_618_8	6B	9K	NE_6366_74	5B	9K
NE_6902_38	5B	9K	NE_6919_40	3B	9K
NE_7005_41	7A	9K	NE_7135_62	5A	9K
NE_7471_66	5B	9K	NE_7507_69	5B	9K
NE_7519_89	3B	9K	NE_7533_44	1D	9K
NE_7547_45	2A	9K	NE_987_48	5B	9K

Tableau P.1 – (partie 1/2)

nom des marqueur	ch	références	nom des marqueurs	ch	références
NE_22478	2B	Kaspar db	NE_26412	2B	Kaspar db
NE_21865	2D	Kaspar db	NE_23211	2D	Kaspar db
NE_60226	4A	Kaspar db	NE_65607	4A	Kaspar db
NE_64002	4D	Kaspar db	NE_68103	4D	Kaspar db
NE_22277	5D	Kaspar db	NE_85191	5D	Kaspar db
NE_80040	6D	Kaspar db	NE_87343	6D	Kaspar db
NE_6485_37 <sup>out</sup>	3D	9K	NE_7930_67 <sup>out</sup>	6B	9K
NE_2366_19 <sup>m</sup>	6A	9K	NE_96478 <sup>m</sup>	7D	Kaspar db
NE_77147 <sup>m</sup>	7D	Kaspar db			

La nomenclature des marqueurs est de la forme *NE\_lab\_label*. *NE* pour « neutre ». ch : chromosome. <sup>m</sup> : monomorphe. <sup>out</sup> : n'ont pas fonctionné. NA : pas disponibles. Kaspar db : Base de données Kaspar. 9K : 9K iSelect assay.

Les positions sur les chromosomes sont issues de Cavanagh et al.<sup>36</sup> et de la base de données en ligne.

## P.2 Marqueurs situés dans des gènes candidats

Tableau P.2 - Marqueurs dans les gènes candidats (*CA*) de la précocité de floraison.

Tableau P.2 – (partie 1/2)

nom du marqueur	caractere associé	polymorphisme	références
PHYA_2045_41	photorécepteur	SNP	9K
ZTL_2307_43	photorécepteurs	SNP	9K
VIL2_3677_25	vernalisation	SNP	9K
SMZ_44_33	photopériode	SNP	9K
VRN1.B_4509_32	vernalisation	SNP	9K
VRN1.B_4699_22	vernalisation	SNP	9K
FT_4805_29	floraison	SNP	9K
SMZ_4872_52	photopériode	SNP	9K
PHYA_5269_42	photorécepteurs	SNP	9K
C04_5440_36	photoperiode	SNP	9K
VRN1.B_5656_38	vernalisation	SNP	9K
CO1_5860_30	photopériode	SNP	9K
TaHd1A_6574_40	photopériode	SNP	9K
CO1_7108_23	photopériode	SNP	9K
SMZ_750_45	photopériode	SNP	9K
SOC1_7643_35	photopériode	SNP	9K
C04_923_27	photopériode	SNP	9K
ZTL_CH_47	photorécepteurs	SNP	9K
C01_4916_50 <sup>m</sup>	photopériode	SNP	9K
FTB_6905_28	vernalisation	SNP	9K
ZTL_7895_46 <sup>out</sup>	photorécepteurs	SNP	9K
VIL2_5042_26 <sup>out</sup>	vernalisation	SNP	9K
C03_4049_34 <sup>out</sup>	photopériode	SNP	9K

Tableau P.2 – (partie 1/2)

nom du marqueur	caractere associé	polymorphisme	références
VRN1.A_R_9612	vernalisation	SNP	<i>CLF</i>
VRN1.A_K_38	vernalisation	SNP	<i>CLF</i>
VRN1.B_M_9892	vernalisation	SNP	<i>CLF</i>
VRN1.B_R_11552	vernalisation	SNP	<i>CLF</i>
VRN1.A_Y_14418	vernalisation	SNP	<i>CLF</i>
VRN1.B_R_1853 <sup>m</sup>	vernalisation	SNP	<i>CLF</i>
VRN3.B <sup>m</sup>	vernalisation	SNP	<i>CLF</i>
VRN3.B_R_7683 <sup>out</sup>	vernalisation	SNP	<i>CLF</i>
VRN1.B_insert	vernalisation	6849bp indel	<i>CLF</i>
TaGI3_5396_37		SNP	201
LDD.A_6412_24		SNP	201
COAB_1		SNP	202
FTA_2	floraison	SSR	28
PPD.D1PROM_6	photopériode	2kb indel	19
VRN1.A_EX8_20	vernalisation	SNP	214
VRN1.D_IN1_7	vernalisation	4kb indel	109
TAGW2_15 <sup>m</sup>	associé à la largeur des grains et le poids de mille grains	SNP	225
PPD1.A_CADE_14	photoperiode	305bp indel	254
PPD1.B_SNP_GC_19	photopériode	SNP	19
PPDB.1_SNP_CT_18 <sup>m</sup>	photopériode	SNP	19

La nomenclature des marqueurs est de la forme *nom-gene\_lab\_label*.

<sup>m</sup> : monomorphe. <sup>out</sup> : n'ont pas fonctionné. NA : pas disponible. 9K : 9K iSelect assay<sup>4;36;46;238</sup>.

*CLF* : marqueurs suggérés par les collègues de l'INRA de Clermont Ferrand.

Les caractères associés sont tirés de<sup>129</sup>.

# **Annexe Q : Fiche sur les croisements**

Ces fiches ont été réalisés en 2010 pour la formation au Moulon par Nathalie Galic et Sophie Pin. Le protocole de semis des croisements a été réalisé en 2011.

## Fiche pratique pour croiser des blés tendres

Un programme de sélection participative a commencé en 2005 en France, avec notre équipe à l'INRA et le groupe blé du Réseau Semences Paysannes (RSP), groupe de paysans français qui s'intéresse à la culture et l'échange des variétés diversifiées et à l'autonomie de la production de semences.

La sélection et la gestion à la ferme des populations cultivées sont des moyens concrets de valoriser la diversité génétique. Ces approches permettent le maintien de la diversité dans chaque population et entre populations différentes, tout en augmentant l'adaptation locale et la stabilité des caractères agronomiques grâce à la résilience des populations hétérogènes.

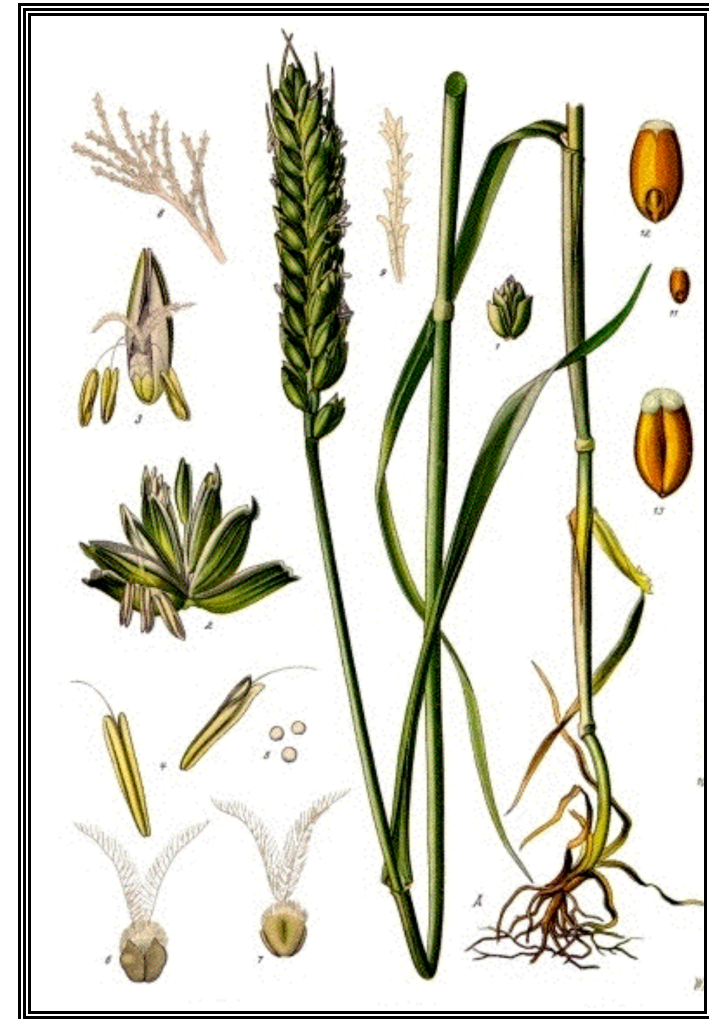
Cette fiche a été produite dans le cadre des journées de formation aux croisements et de discussions autour de la sélection participative. Elle est destinée aux paysans qui veulent réaliser eux-mêmes des croisements entre variétés de pays ou variétés anciennes pour créer des nouvelles populations dans l'esprit d'un travail collectif de sélection paysanne.

Auteurs : Sophie Jouanne-Pin et Nathalie Galic

Equipe Diversité, Evolution et Adaptation des Populations  
Isabelle Goldringer, Julie Dawson  
UMR de Génétique Végétale  
INRA Ferme du Moulon  
91190 Gif-sur-Yvette, France  
Tél. 01 69 33 23 70

Réseau Semences Paysans [www.semencespaysans.org](http://www.semencespaysans.org)  
Cazalens 81 600 BRENS • Tél. 05 63 41 72 86

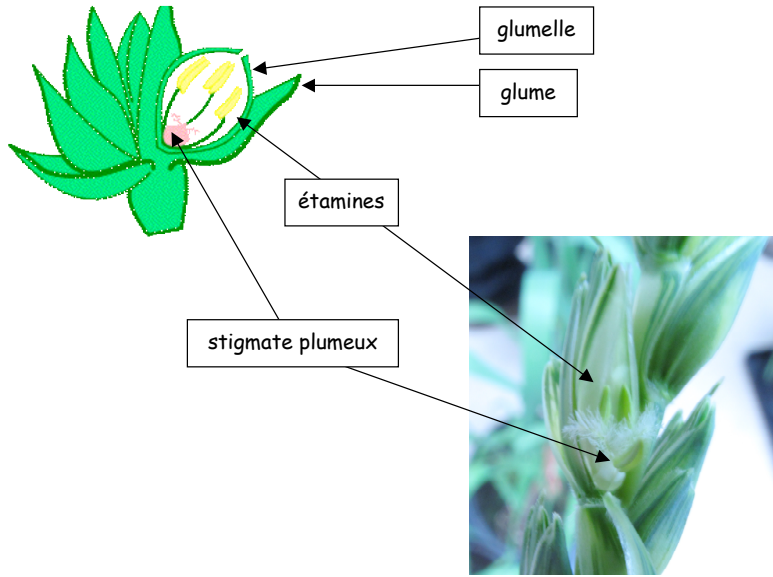
## Croisements de Blé



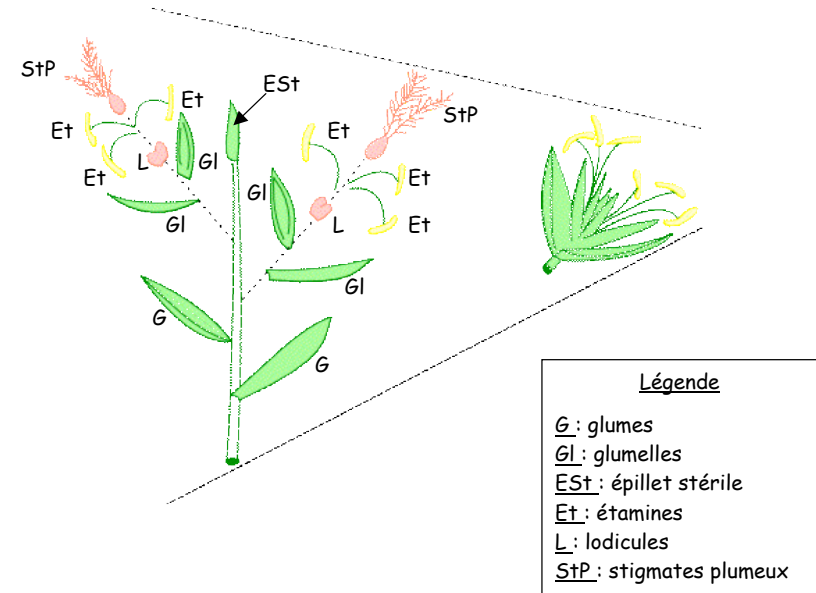
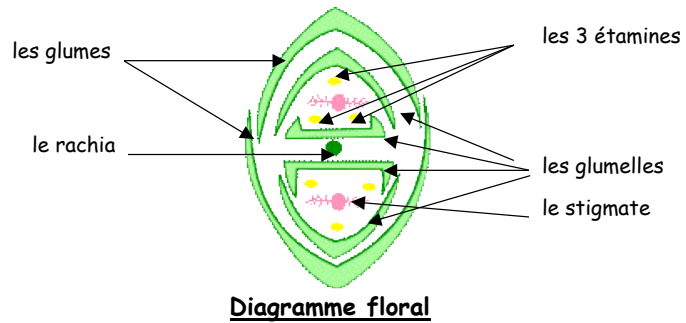
Source : List of Koehler Images



## La fleur du blé



Cliché : Mathieu THOMAS



**Schéma d'un épillet**



La castration

1



Choix de l'épi à castrer :  
il ne faut pas qu'il soit trop  
sorti du haut de la gaine

2



On enlève les épillets stériles  
à la base de l'épi

3



On enlève la fleur centrale  
de tous les épillets (Cet  
étape peut aussi se faire  
après l'étape 5).



4



On coupe le haut de l'épi, où  
il y a des épillets stériles

5



On coupe le haut des épillets

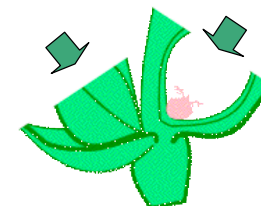
6



Source : GNIS pédagogique

On enlève les 3 étamines par fleur

Les 2 fleurs



**ATTENTION**  
Il y a 3 étamines par fleur  
et  
2 fleurs  
donc  
6 étamines par épillet à enlever





7



Une fois fini, on met un sac pour protéger les stigmates des pollens environnants, sur lequel on note la date de castration.

4 à 5 jours plus tard

Quand la femelle est réceptrice, et que le stigmate est bien plumeux

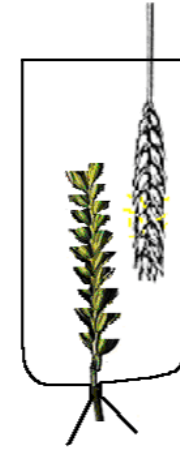
8



Pour choisir le mâle :  
Il faut que sa floraison (étamines sorties), représente environ 1 cm sur l'épi



9



On retourne le mâle au dessus de la femelle, on secoue et on referme le sac, en le laissant dans le sac.  
Il est possible d'enlever l'épi mâle 1 semaine après le croisement.

**ATTENTION**  
On laisse le sac jusqu'à la récolte

Informations à mettre sur le sac du croisement

Date de castration  
Nom de la femelle  
  
X  
  
Nom du mâle  
Date de la pollinisation

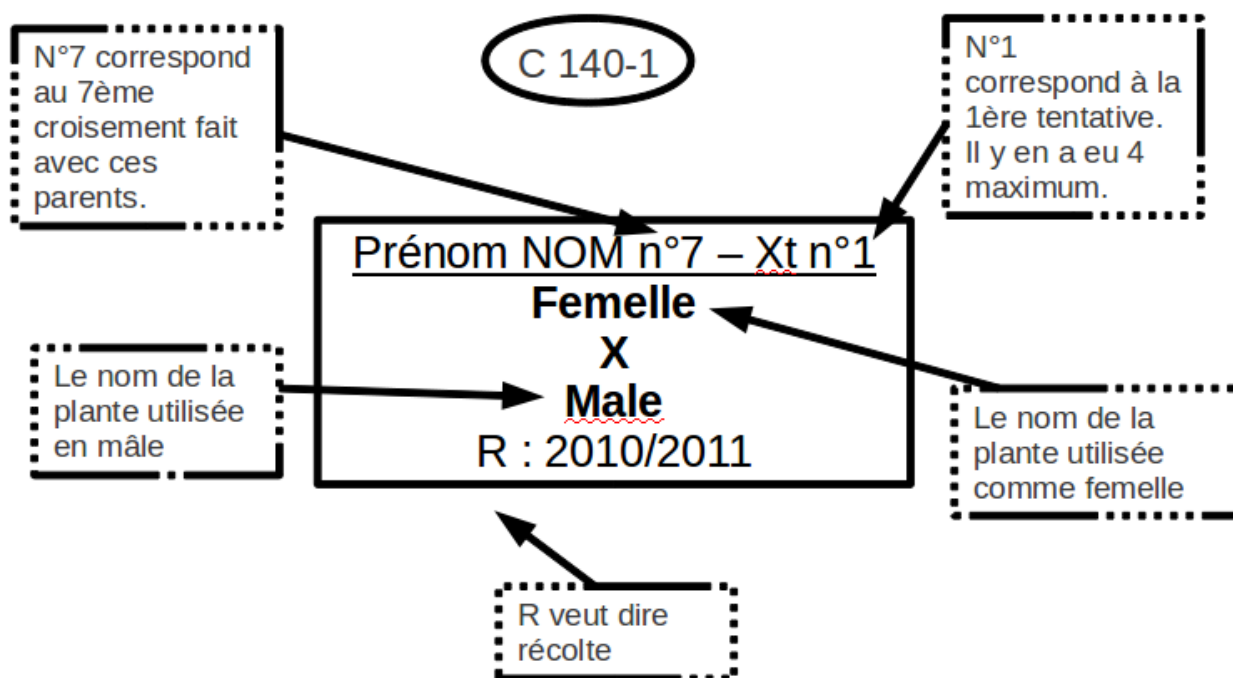
**ATTENTION**  
Utiliser un feutre indélébile

## Fiche semis des croisements faits au Moulon

Cette fiche fait le point sur la manière de cultiver les croisements qui ont été fait au Moulon et pourquoi.

Pour chacun des croisements que vous nous avez donné à faire, nous avons fait 4 tentatives. Certaines ont marché, d'autres pas ... Les graines sont dans des sachets. Il y a deux sachets pour chaque croisement.

Le 1<sup>er</sup> sachet contient les graines issues du croisement, les étiquettes se présentent sous la forme suivante :



**Ce qui est entouré est le nom du lot. A mettre sur vos étiquettes.** Ici c'est le lot C 140-1. Après avoir semé les lots, vous pourrez mélanger tous les lots de la famille C 140 (C 140-1, C 140-2, C 140-3 par exemple) pour créer une nouvelle population qui aura comme nom C 140.

2. Le 2<sup>ème</sup> sachet contient les graines de la plante utilisée comme femelle, les étiquettes se présentent sous la la forme suivante :

Pour voir si le croisement a bien fonctionné (c'est à dire qu'il n'y a pas eu d'autofécondation), **il est important de semer côte à côte les graines issues du croisement (les F1 = la 1ère génération) et les graines de la plante qui a été utilisée comme femelle.**

T SVP

2 cas possibles :

- **le croisement a bien fonctionné** : On observe au champ des différences entre la femelle et le croisement. Le croisement est assez homogène.

- ***le croisement n'a pas fonctionné : On N'observe PAS de différences au champ entre la femelle et le croisement ou le croisement est très hétérogène.***

Pour ceux qui veulent aller plus loin, continuer la fiche, sinon c'est ok !

## Pourquoi ?

### 1. Base génétique

Un individu diploïde a deux versions d'un gène. Une version qui vient de son père et une version qui vient de sa mère.

Par exemple, prenons une plante qui a la version « **A** » du gène, qui vient de sa mère et la version « **A** » du gène qui vient de son père. La plante est donc « **AA** »

On peut aussi avoir une plante qui a la version « **A** » du gène, qui vient de sa mère et la version « **a** » du gène qui vient de son père. La plante est donc « **Aa** »

On peut aussi avoir une plante qui a la version « **a** » du gène, qui vient de sa mère et la version « **a** » du gène qui vient de son père. La plante est donc « **aa** »

Quand la plante se reproduit, elle donne la moitié de ce qu'elle possède.

Une plante qui est « **AA** » va forcément donner « **A** » à ses descendants.

Une plante qui est « **aa** » va forcément donner « **a** » à ses descendants.

Une plante qui est « **Aa** » va soit donner « **A** », soit donner « **a** » à ses descendants.

Sur l'ensemble du génome, il y a beaucoup de locus (beaucoup de lettres), donc les différences vont se cumuler.

### 2. Nos croisements sur les blés.

Le blé est autogame, ce qui veut dire qu'après plusieurs générations, il va être fixé. Cela veut dire que les deux versions d'un gène sont les mêmes. Dans notre cas, prenons comme exemple la femelle (**AA**) avec le mâle (**aa**).

Quand la femelle s'autoféconde avec elle-même, cela donnera « **AA** » (il n'y a pas d'autres choix possibles).

Quand il y a un croisement, la femelle donne « **A** » (pas d'autres choix possibles) et le mâle donne « **a** » (pas d'autres choix possibles). Les graines issues de la fécondation sont donc « **Aa** ».

Quand nous faisons les croisements, nous castrons une plante qui sera la femelle et nous lui donnons du pollen de la plante mâle. Nous faisons donc le croisement sur la plante femelle. Cela veut dire que si il y a autofécondation, les graines seront « **AA** » (la plante femelle s'autoféconde) et si il y a eu croisement avec la plante mâle, les graines seront « **Aa** ».

Quand on castre les épis, il est possible que des étamines restent dans les épillets ... des étamines vont larguer du pollen qui va féconder les stigmates dans les épillets ... ce qui donne de l'autofécondation : ce n'est pas recherché dans notre cas !

**Pour résumer :**

2 cas possibles :

***- le croisement a bien marché***

On a bien la moitié qui vient de la femelle et la moitié qui vient du mâle, donc les graines sont « *Aa* ». Donc la femelle « *AA* » et nos graines « *Aa* » sont différentes. **On observe au champs des différences entre la femelle et le croisement. Toutes les plantes du croisement se ressemblent.**

***- le croisement n'a pas marché***

Le croisement n'a pas bien marché et il y a eu autofécondation sur la femelle. On a donc des graines qui sont « *AA* », comme la femelle : la femelle « *AA* » et nos graines « *AA* » sont identiques. **Certaines plantes (voire toutes les plantes ) du croisement ressemblent à la femelle.**

Il se peut qu'une partie de la femelle ait été autofécondée et une partie ait été croisée ... donc dans un lot de grains issues de croisement, il peut y avoir des graines « *AA* » et des graines « *Aa* ». Il faut donc bien observer toutes les plantes.



# **Annexe R : Le dossier retour**

Il y a autant de dossiers retour que de paysans. Dans cette annexe, seuls les dossiers retour de FLM en 2011 et CHD en 2012 sont présentés avec quelques captures d'écran. Les résultats des analyses statistiques n'étaient pas encore implémentés dans le dossier.



**Projet de sélection participative entre le Réseau Semences Paysannes et l'INRA du Moulon : création de variétés de blé tendre de qualité adaptées à l'Agriculture Biologique et conservation de la biodiversité cultivée**

**Livret récapitulatif des résultats sur la ferme**

**Octobre 2011**



**INRA du Moulon:** UMR de Génétique Végétale, Équipe Diversité Évolution et Adaptation des Populations.

Ferme du Moulon, 91190 Gif-sur-Yvette, France.

Site internet : <http://moulon.inra.fr/index.php/fr/equipements/leap>

contact : [pierre.riviere@moulon.inra.fr](mailto:pierre.riviere@moulon.inra.fr) ; [isabelle.goldfinger@moulon.inra.fr](mailto:isabelle.goldfinger@moulon.inra.fr)

**Réseau Semences Paysannes:** Cazalens, 81600 Brens, France.

Site internet : <http://www.semencespaysannes.org/>

contact : [patrick@semencespaysannes.org](mailto:patrick@semencespaysannes.org)

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>FLM : Résultats sur la ferme</b>	<b>5</b>
	2.0.1 notes d'automne	5
	2.1 notes d'hiver	6
	2.2 notes de printemps	10
	2.3 notes d'été	14
	2.4 Mesures quantitatives	23
<b>3</b>	<b>Résultats par croisement</b>	<b>24</b>
	3.1 Croisement 1	24
	3.1.1 Mesures quantitatives	24
	3.2 Croisement 14	26
	3.2.1 Mesures quantitatives	26
	3.3 Croisement 16	26
	3.3.1 Mesures quantitatives	26
	3.4 Croisement 21	28
	3.4.1 Mesures quantitatives	28
	3.5 Croisement 21sb	28
	3.5.1 Mesures quantitatives	28
	3.6 Croisement 24	30
	3.6.1 Mesures quantitatives	30
	3.7 Croisement 29	30
	3.7.1 Mesures quantitatives	30
	3.8 Croisement 31	32
	3.8.1 Mesures quantitatives	32
	3.9 Croisement 32	32
	3.9.1 Mesures quantitatives	32
	3.10 Croisement 37	34
	3.10.1 Mesures quantitatives	34
	3.11 Croisement 42	34
	3.11.1 Mesures quantitatives	34
	3.12 Croisement 45	36
	3.12.1 Mesures quantitatives	36
	3.13 Croisement 46	36
	3.13.1 Mesures quantitatives	36
	3.14 Croisement 48	38
	3.14.1 Mesures quantitatives	38



## Chapitre 2

# FLM : Résultats sur la ferme

### 2.0.1 notes d'automne

Les données concernant le semis et le climat de 2009-2010 n'est pas noté. Des informations peuvent se trouver dans les remarques.

Information sur le climat 2010-2011

Pluie : faibles

Température :

Accident climatique : RAS, neige un peu plus que d'habitude

Accident sur la culture : non, protégé par une clôture

Info climat :

Les informations sont les suivantes :

taille micro parcelle 2009-2010 : 5 m2

date semis 2009-2010 : 3-4 février

methode semis 2009-2010 : semoir é essais

Commentaires 2009-2010 :

taille parcelle 2010-2011 : 1.5 ha

taille micro parcelle 2010-2011 : 5 m2

topographie 2010-2011 : plateau

type sol 2010-2011 : sableux

sol 2010-2011 : intermédiaire

battance 2010-2011 : oui

drainage 2010-2011 : non

labour 2010-2011 : oui

date semis 2010-2011 :

methode semis 2010-2011 : semoir

levées 2010-2011 : lente et régulière

precedent 2010-2011 : prairie

Commentaires 2010-2011 : Prairie 4 ans

Information sur le climat 2010-2011

Pluie : faibles

Température :

Accident climatique : ras

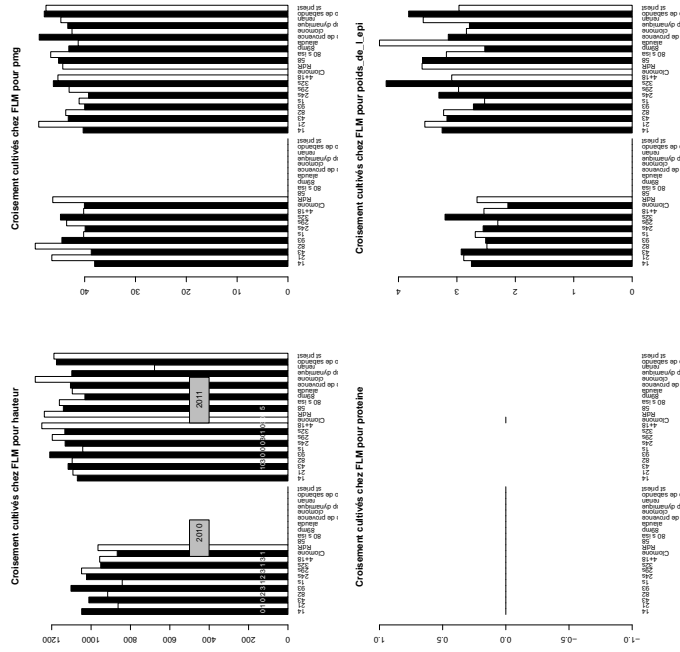
Accident sur la culture : faupins sur certaines zones

Info climat : développement irrégulier

Croisement	Date d'observation	Globale	Reprise	Densité de Peuplement (plantes/m <sup>2</sup> )	Port au tallage	Attitude des feuilles	Commentaires
14	2010						
21	2010						
43	08/03/11	2		100 à 200	demt-dresse		seule les lignes extéreur sont remplies
82	2010						
82	08/03/11	3		100 à 200	demt-dresse		densité moyenne
93	08/03/11	4		100 à 200	intermédiaire		
1s	08/03/11	4		200 à 300	demt-dresse		belle population
24s	08/03/11	4		200 à 300	intermédiaire		recouvre bien le sol et la ligne
29s	08/03/11	5		200 à 300	demt-dresse		forte densité avec une bonne hauteur
32s	08/03/11	4		200 à 300	demt-dresse		ligne bien garnies

2.4 Mesures quantitatives

Pour le graphique représentant la hauteur, le chiffre sur la barre représente la note de verse. Il est présent quand il n'y avait pas de répétition, sinon voir dans la partie mesures d'été. En 2010, c'est une note de 0 (droit) à 5 (versé). En 2011, c'est un pourcentage de verse. La hauteur est en mm, le poids de l'épis et le poids de mille grains (pmg) sont en grammes.



Croisement	Date	Globale	Biomasse	Crosses	Couleur	hétéro	Barbes	Verse	Commentaires
14	2010			pas crossé	blanc	4	barbus	0	
				intermédiaire	intermédiaire		nonbarbus		
14	2010			pas crossé	blanc	3	barbus	1	qs epis tres foncés mb
				intermédiaire	intermédiaire		intermédiaire		
21	2010			pas crossé	blanc	3	barbus		
				intermédiaire	intermédiaire		intermédiaire		
43	2010			pas crossé	blanc	3	barbus	0	
				intermédiaire	intermédiaire		intermédiaire		
43	28/07/11	3		pas crossé	blanc	2	barbus	10	
				intermédiaire	intermédiaire		intermédiaire		
43	28/07/11	3		pas crossé	blanc	2	barbus	10	
				intermédiaire	intermédiaire		intermédiaire		

2.3 notes d'été

Information sur le climat 2010-2011

Pluie : faibles

Température :

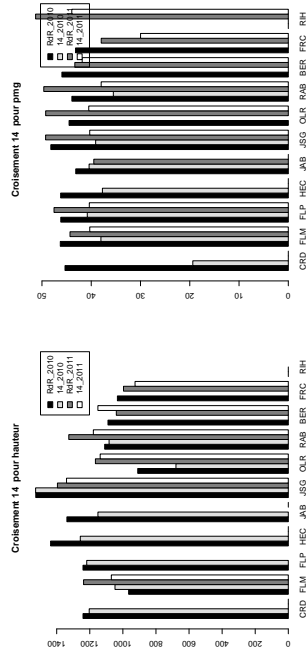
Accident climatique : pluie avant récolte

Accident sur la culture : taupin sur une partie + hévre

Info climat :

### 3.2 Croisement 14

#### 3.2.1 Mesures quantitatives



Croisement 14 pour hauteur

Croisement 14 pour pmg

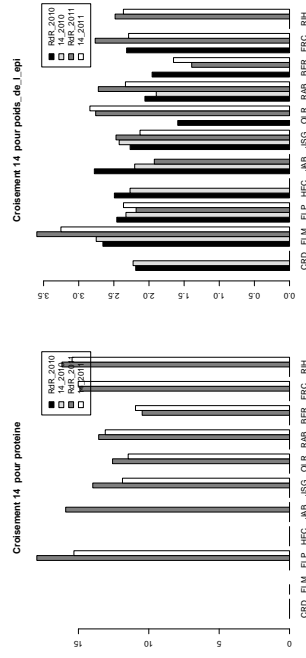
## Résultats par croisement

Dans cette partie, vous trouverez tout les résultats par croisement. Les croisements présents sur plus de deux fermes sont représentés. Pour les données qualitatives, les rep et Rouge du Roc ne sont pas présents, il faut aller voir dans la partie ferme. Ce dossier est en cours d'élaboration et est loin d'être parfait ...

Pour la verse, l'information n'est pas sur le graphique mais dans la partie mesures été, quand l'info est disponible.

### 3.1 Croisement 1

#### 3.1.1 Mesures quantitatives



Croisement 14 pour protéine

Croisement 14 pour poids\_de\_lipi

### 3.3 Croisement 16

#### 3.3.1 Mesures quantitatives



**Projet de sélection participative entre le Réseau Semences Paysannes et l'INRA du Moulon : création de variétés de blé tendre de qualité adaptées à l'Agriculture Biologique et conservation de la biodiversité cultivée**

**Livret récapitulatif des résultats sur la ferme  
Septembre 2012**



**INRA du Moulon:** UMR de Génétique Végétale, Equipe Diversité Evolution et Adaptation des Populations,

Ferme du Moulon, 91190 Gif-sur-Yvette, France.

Site internet : <http://moulon.inra.fr/index.php/fr/equipements/deap>

contact : [pierre.riviere@moulon.inra.fr](mailto:pierre.riviere@moulon.inra.fr) ; [isabelle.goldringer@moulon.inra.fr](mailto:isabelle.goldringer@moulon.inra.fr)

**Réseau Semences Paysannes:** 3 avenue de la gare 47190 Aquillon, France.

Site internet : <http://www.semencespaysannes.org/>

contact : [patrick@semencespaysannes.org](mailto:patrick@semencespaysannes.org)

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Comment utiliser ce dossier ?</b>	<b>5</b>
2.1	Caractérisation des fermes sur le réseau	5
2.2	Résultat sur la ferme	6
2.2.1	les variétés du programme de sélection participative	6
2.2.2	les bouquets de sélection : différentiel de sélection et réponse à la sélection (pour les concernés)	8
2.2.2.1	Théorie	8
2.2.2.2	Lecture des résultats	8
2.3	Résultats sur le réseau des différentes variétés	10
<b>3</b>	<b>Caractérisation des fermes sur le réseau</b>	<b>12</b>
3.1	Répartition géographique des fermes	12
3.2	Fermes regroupées selon leurs similitudes	13
3.2.1	Données de semi	13
3.2.2	Données sur différentes variables	20
<b>4</b>	<b>CHD : Résultats sur la ferme</b>	<b>23</b>
4.1	variétés du programme de sélection participative	23
4.2	notes d'automne	23
4.3	notes d'hiver	24
4.4	notes de printemps	27
4.5	notes d'été	27
4.6	Mesures quantitatives	28
4.7	bouquets de sélection : différentiel de sélection et réponse à la sélection	31
4.7.1	différentiel de sélection 2011 et réponse à la sélection 2012	31
<b>5</b>	<b>Vos Variétés qui sont cultivées sur au moins deux fermes</b>	<b>35</b>
5.1	Résultats pour la variété C14	36
5.1.1	Mesures quantitatives	36
5.2	Résultats pour la variété C16	38
5.2.1	Mesures quantitatives	38
5.3	Résultats pour la variété C20	40
5.3.1	Mesures quantitatives	40
5.4	Résultats pour la variété C21	42
5.4.1	Mesures quantitatives	42
5.5	Résultats pour la variété C21#5b	44

## Chapitre 2

### Comment utiliser ce dossier ?

Nous vous donnons dans cette partie des clés pour lire chaque graphique, chaque tableau. Avec le temps, vous vous y habituerez !

#### 2.1 Caractérisation des fermes sur le réseau

1. **Lire un graphique en boîte à moustache.** L'intérêt de ce graphique est qu'il permet de voir la **diversité au sein d'une variété**. On peut voir comment se répartit un caractère. Sur la Figure 2.1, on observe la répartition des données avec le maximum, le minimum et la médiane (valeur où j'ai la moitié des individus au dessus et la moitié en dessous). On observe aussi les données qui sont considérées comme extrêmes.

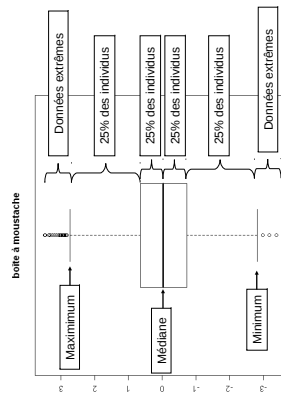


Figure 2.1 – Comment lire un graphique en boîte à moustache.

Dans le dossier, quand c'est possible (quand on a assez de données), on représente les données sous forme d'une boîte à moustache, sinon, c'est sous forme d'un histogramme classique.

## Chapitre 3

### Caractérisation des fermes sur le réseau

Plus d'informations sur la lectures des graphiques dans la partie 2 (page 5). Cette partie vous permettra de voir quelles fermes sont le plus proches de vous. Cela vous permet de mieux comparer et pourquoi pas, aller chercher des variétés qui se comportent bien sur des fermes proches des vôtres.

#### 3.1 Répartition géographique des fermes

Les différentes fermes sont classées par les initiales des paysans. La carte ci dessous montre qui est où.

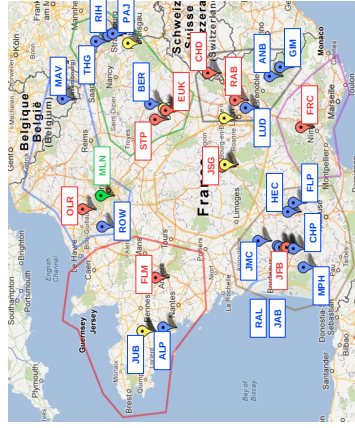
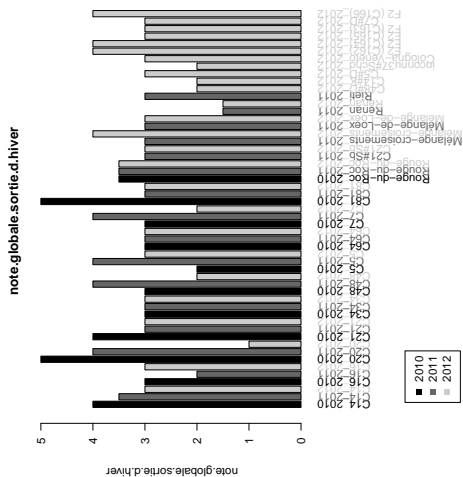


Figure 3.1 – Participants du programme : les fermes régionales (rouge), les fermes satellites (bleues), les associations (jaune) et l'INRA du Moulon (vert).

4.3 notes d'hiver



Chapitre 4

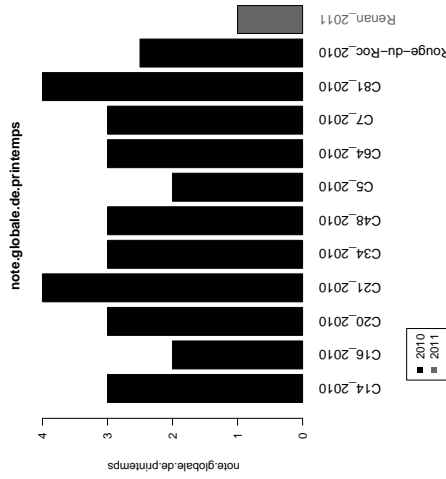
CHD : Résultats sur la ferme

4.1 variétés du programme de sélection participative

4.2 notes d'automne

Les informations sont les suivantes :  
 taille parcelle 2011-2012 : 0.5ha ha  
 taille micro parcelle 2011-2012 : 10m2 m2  
 topographie 2011-2012 : pente 3 à 10type sol 2011-2012 : argilo-calcaire  
 sol 2011-2012 : intermédiaire  
 batiance 2011-2012 : non  
 drainage 2011-2012 : non  
 labour 2011-2012 : oui  
 date semis 2011-2012 : 02/11/2011  
 methode semis 2011-2012 : volée  
 levee 2011-2012 : rapide et régulière  
 precedent 2011-2012 : prairie  
 Commentaires 2011-2012 :

4.4 notes de printemps



pas de retour cette année

4.5 notes d'été

pas de graphique pour note.globale.d.été car pas de données sur ces différentes années

pas de retour cette année

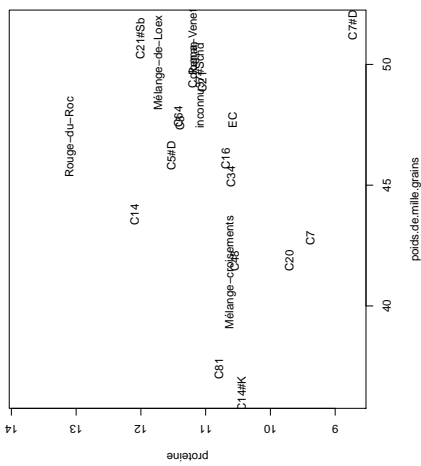
Variete	Date d'ob-	Reprise	Densité de peuplement (plantes/m²)	Port au tallage	Attitude des feuilles	Commentaires
Rouge-du-	26/03/2012	3	100 à 200	intermédiaire	de biais	peuplement irrégulier
Mélange-	26/03/2012	3	200 à 300	demi-dressé	horizontale	
Crosnements						
C48#D	26/03/2012	3	0 à 100	intermédiaire	de biais	
C14#K	26/03/2012	2	0 à 100	demi-dressé	de biais	densité faible : gel
Renan	26/03/2012	2	0 à 100	dressé	verticale	
C64	26/03/2012	3	100 à 200	intermédiaire	verticale	
C21#Sb	26/03/2012	4	200 à 300	demi-dressé	de biais	premières traces de maladie sur feuilles
C7	26/03/2012	2	0 à 100	dressé	verticale	faible peuplement : pb à la levée
C21	26/03/2012	2	100 à 200	intermédiaire	de biais	
C5#D	26/03/2012	3	100 à 200	demi-dressé	verticale	
incommu37#Sc	26/03/2012	2	0 à 100	demi-dressé	verticale	
C81	26/03/2012	2	200 à 300	demi-étalé	de biais	
Mélange-de-	26/03/2012	3	200 à 300	intermédiaire	de biais	
Loex						
C14	26/03/2012	3	200 à 300	intermédiaire	verticale	belle population
C20	26/03/2012	2	0 à 100	demi-étalé	horizontale	faible peuplement : pb à la levée
C5	26/03/2012	2	200 à 300	demi-dressé	de biais	
Cologna-	26/03/2012	4	100 à 200	demi-dressé	verticale	
Veneto						
F2 (C162)	26/03/2012	4	300 à 400	intermédiaire	de biais	
F2 (C164)	26/03/2012	5	300 à 400	demi-dressé	de biais	
F2 (C165)	26/03/2012	3	200 à 300	intermédiaire	de biais	
F2 (C163)	26/03/2012	3	100 à 200	étalé	horizontale	
C48	26/03/2012	2	100 à 200	demi-étalé	horizontale	

4.7 bouquets de sélection : différentiel de sélection et réponse à la sélection

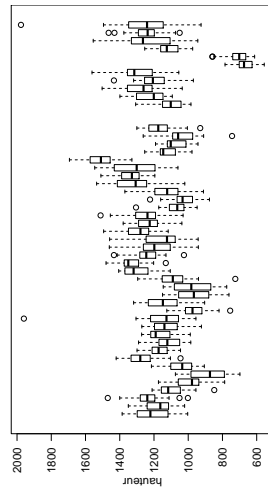
4.7.1 différentiel de sélection 2011 et réponse à la sélection 2012

4.6 Mesures quantitatives

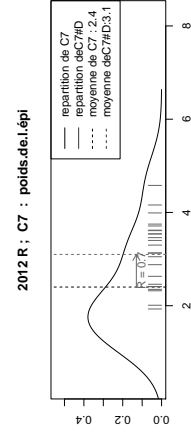
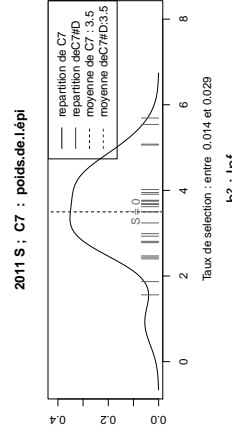
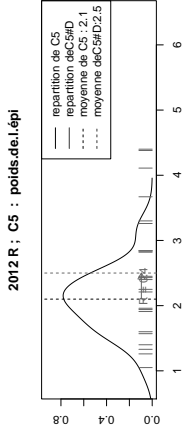
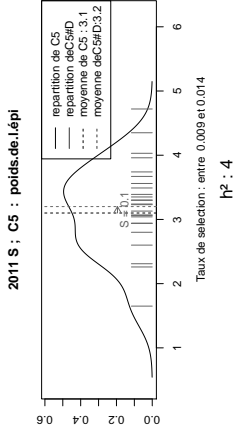
2012 : poids de mille grains et prot (g et %)



hauteur (mm)



Rouge-du-Roc-2010  
Rouge-du-Roc-2011  
Rouge-du-Roc-2012  
Renaud-2011  
Renaud-2012  
Mélange-de-Loex-2011  
Mélange-de-Loex-2012  
Inconnus7-2011  
Inconnus7-2012  
Colonna-Veneto-2011  
C8-2010  
C7-2010  
C5-2010  
C8-2010  
C2-2010  
C21-2010  
C22-2010  
C23-2010  
C24-2010  
C25-2010  
C26-2010  
C27-2010  
C28-2010  
C29-2010  
C30-2010  
C31-2010  
C32-2010  
C33-2010  
C34-2010  
C35-2010  
C36-2010  
C37-2010  
C38-2010  
C39-2010  
C40-2010  
C41-2010  
C42-2010  
C43-2010  
C44-2010  
C45-2010  
C46-2010  
C47-2010  
C48-2010  
C49-2010  
C50-2010  
C51-2010  
C52-2010  
C53-2010  
C54-2010  
C55-2010  
C56-2010  
C57-2010  
C58-2010  
C59-2010  
C60-2010  
C61-2010  
C62-2010  
C63-2010  
C64-2010  
C65-2010  
C66-2010  
C67-2010  
C68-2010  
C69-2010  
C70-2010  
C71-2010  
C72-2010  
C73-2010  
C74-2010  
C75-2010  
C76-2010  
C77-2010  
C78-2010  
C79-2010  
C80-2010  
C81-2010  
C82-2010  
C83-2010  
C84-2010  
C85-2010  
C86-2010  
C87-2010  
C88-2010  
C89-2010  
C90-2010  
C91-2010  
C92-2010  
C93-2010  
C94-2010  
C95-2010  
C96-2010  
C97-2010  
C98-2010  
C99-2010  
C100-2010





## 5.1 Résultats pour la variété C14

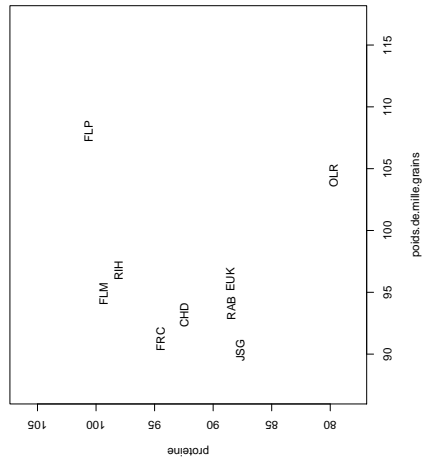
### 5.1.1 Mesures quantitatives

## Chapitre 5

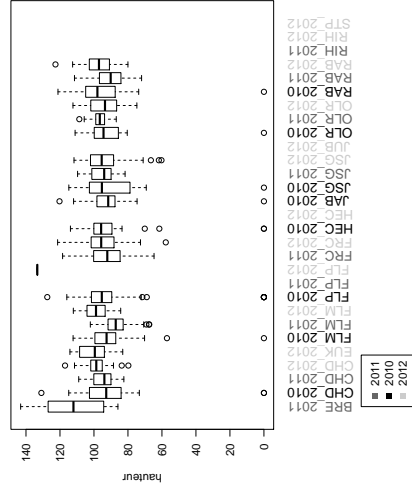
### Vos Variétés qui sont cultivées sur au moins deux fermes

Dans cette partie, vous trouverez les résultats concernant les variétés présentes sur vos fermes sur le réseau. Les variétés cultivées au moins sur deux fermes sont représentées. **ATTENTION, les données sont en % de Rouge-du-Roc notre témoin.**

2012 : poids de mille grains et prot (en % de Rouge-du-Roc)



hauteur (en % de Rouge-du-Roc)





# **Annexe S : Répartitions des tâches entre les paysans, les associations et le Moulon dès 2012**

Fiche réalisée par Guillemette Fruchaud lors de son stage de master 2 dans l'équipe DEAP pendant l'été 2012.

## Annexe : proposition de nouvelle organisation

Les associations à l'interface entre les paysans/nes et l'équipe de recherche : décentralisation du processus et collecte d'information par les assos avant de renvoyer des fichiers remplis au Moulon. (en jaunes les modifications de robin -CETAB-)

Date	Tache	Paysans/nes	↔	Association	↔	Moulon
Préparation de semis (Septembre)	Création dossier retour (résultats de l'année écoulée)					X
	Envoi dossier retour été + renvoi des bouquets de sélection par région				←	X
	Répartition aux différents paysans/nes de la région (lors d'une réunion préparation semis??) du dossier retour et des sélections.		←	X		
	Contacteur les paysans pour savoir ce qu'ils veulent semer (de l'an passé ou en +) (téléphone)		←	X		
	Gérer les diffusions de semences (par tel + trace écrite mail ou courrier)		←	X		
	Envoyer les pops demandées	X	Ou	X		
	Faire les plans de semis en fonction des réponses	X	Et	X		
Collecte d'infos sur les semis (Octobre-Novembre)	Faire le traitement de semence	(X)		X		
	Semer les pops	X				
	Envoi fiche semis + notice (mail)				←	X
	Impression et Envoi des plans de semis + fiche semis + notice + lettre + enveloppe timbrée de retour (poste)		←	X		
	Remplir les fiches semis	X				
	Renvoyer fiche semis remplie (poste)	X	→			
	Au besoin relancer (par tel.)		←	X		
	Remplir, au fur et à mesure des réponses, le tableur excel « fichier type semis »			X		
Renvoyer le tableur excel rempli complet (mail)			X	→		
Collecte d'infos sur le tallage (sortie d'hiver) Février	Création fiche tallage en fonction des pops semées.					X
	Envoi des fiches tallage + notice + fichier type pré-rempli avec les noms des pops (par mail)				←	X
	Impression et Envoi des fiches tallage + notice + lettre + enveloppe timbrée de retour (poste)		←	X		
	Remplir la fiche tallage	X				
	Renvoyer fiche tallage	X	→			
	Au besoin relancer par tel.		←	X		
	Remplir le tableur excel « fichier type tallage »			X		

	Renvoyer le tableur excel rempli (mail)			X	→	
Collecte d'infos sur l'épiaison (début Mai-Juin)	Création fiche épiaison en fonction des pops semées.					X
	Envoi de la fiche épiaison + fichier type pré-rempli avec les noms des pops (par mail)				←	X
	Impression et Envoi des fiches épiaison + notice + lettre + enveloppe timbrée de retour (poste)		←	X		
	Remplir la fiche épiaison	X				
	Renvoyer fiche épiaison (poste)	X	→			
	Au besoin relancer par tel.		←	X		
	Remplir le tableur excel « fichier type épiaison »			X		
	Renvoyer le tableur excel rempli (mail)			X	→	
Collecte d'infos à maturité (Juillet-Août)	Création fiche été en fonction des pops semées.					X
	Envoi de la fiche été + fichier type pré-rempli avec les noms des pops (mail)				←	X
	Impression et Envoi des fiches été + notice + lettre + enveloppe timbrée de retour (poste)		←	X		
	Remplir la fiche été	X				
	Renvoyer fiche été	X	→			
	Au besoin relancer par tel.		←	X		
	Remplir le tableur excel « fichier type été »			X		
	Renvoyer le tableur excel rempli (mail)			X	→	
Visite de fermes (été)	Organisation			X		(X)
	Intervention/présentation/jeu sur la sélection...			(X)		X
Mesures (Juillet-Août)	Mesurer sur les fermes régionales (1 ou 2 fermes par association)			X		(X)
	Préparer les sacs de récolte d'échantillon avec les étiquettes					X
	Envoyer le doc récolte + les sacs pour les échantillons		←	←	←	X
	Collecter les échantillons et faire les sélections (optionnel)	X				
	Renvoi des échantillons	X	→	→	→	
	Mesurer les échantillons					X
	Récolter les essais	X				
	Création dossier retour					X

Et la boucle est bouclée



