Pl. Syst. Evol. 264: 195–216 (2007)
DOI 10.1007/s00606-006-0478-x
Printed in The Netherlands

Plant Systematics and Evolution

Molecular phylogeny of the genus *Triticum* L.

K. A. Golovnina[1], **S. A. Glushkov**[1], **A. G. Blinov**[1], **V. I. Mayorov**[2], **L. R. Adkison**[2], and **N. P. Goncharov**[1]

[1]Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2]Mercer University School of Medicine, Macon GA, USA

Received April 11, 2006; accepted September 6, 2006
Published online: February 12, 2007
© Springer-Verlag 2007

Abstract. The genus *Triticum* L. includes the major cereal crop, common or bread wheat (hexaploid *Triticum aestivum* L.), and other important cultivated species. Here, we conducted a phylogenetic analysis of all known wheat species and the closely related *Aegilops* species. This analysis was based on chloroplast *mat*K gene comparison along with *trn*L intron sequences of some species. Polyploid wheat species are successfully divided only into two groups – Emmer (sections *Dicoccoides* and *Triticum*) and Timopheevii (section *Timopheevii*). Results reveal strictly maternal plastid inheritance of synthetic wheat amphiploids included in the study. A concordance of chloroplast origin with the definite nuclear genomes of polyploid species that were inherited at the last hybridization events was found. Our analysis suggests that there were two ancestral representatives of *Aegilops speltoides* Tausch that participated in the speciation of polyploid wheats with B and G genome in their genome composition. However, G genome species are younger in evolution than ones with B genome. B genome-specific PCR primers were developed for amplification of *Acc*-1 gene.

Key words: wheat, *Triticum*, *Aegilops*, molecular evolution, plasmon and B genome inheritance.

Introduction

Humans and wheat share a remarkably parallel evolutionary history. About 3 million years ago, mankind diverged from apes, and diploid A, B, and D genome ancestral wheat species diverged from a common ancestor (Gill et al. 2004). About 200000 years ago, at nearly the same time as modern humans originated in Africa, two diploid grass species hybridized to form tetraploid wheat in the Middle East. Humans domesticated wheat 15000 years ago in the Fertile Crescent (modern day Iraq and parts of Turkey, Syria and Iran), marking the dawn of modern civilization (Harlan 1992). Now many of these wheat species represent the world's most important food crops including hexaploid common wheat *Triticum aestivum* L. (genome BBAADD, $2n = 6x = 42$) that occupies 17% of the world crop area (FAO 2004). The polyploid wheat phylogeny has been the subject of research during the last hundred years (Larionov 1914, Percival 1921, Lilienfeld and Kihara 1934, McFadden and Sears 1946, Riley et al. 1958, Tsunewaki and Ogihara 1983, Feldman and Levy 2005). It is apparent

now that intergeneric hybridization and allopolyploidy, involving the most closely related genus *Aegilops* L., played the key role in the speciation of wheat species. Recent studies have shown that allopolyploidy accelerates genome evolution in wheat in two ways: allopolyploidization triggers rapid genome changes (revolutionary changes) through the instantaneous generation of a variety of cardinal genetic and epigenetic alterations, and the allopolyploid condition facilitates sporadic genomic changes during the species life (evolutionary changes) that are not attainable at the diploid level (Feldman and Levy 2005). However, the decreased success in levels of genetic diversity due to domestication that is associated with population bottlenecks was also observed (Jaaska 1980, Tanskley and McCouch 1997, Provan et al. 2004).

Thus, wheat evolution proceeded in a reticulate manner: accumulation of sequence divergence in the diploid species followed by convergence via allopolyploidy. This reticulation complicates taxonomic considerations in traditional cladistic analysis depending upon divergent branching patterns (Jauhar and Crane 1989, Kellogg 1989, Blake et al. 1998). However, several types of analyses have provided an insight into the ancestry of definite genomes in allopolyploid species. These include comparisons among polyploids and potential progenitors using morphological traits, chromosome pairing relationships and biochemical and molecular markers (Percival 1921, Riley et al. 1958, Kerby and Kuspira 1986, Blake et al. 1998, Hammer et al. 2000, Goncharov 2002, Tsunewaki et al. 2002, Zhang et al. 2002, Gu et al. 2004).

Species of the genus *Triticum* L. exist as a polyploid series of di-, tetra- and hexaploid wheat with a basic number $n = 7$ (Sakamura 1918). Four basic genomes designated as A, B, D and G assist in the genome constitution of all *Triticum* species (Kihara 1924, Lilienfeld and Kihara 1934). The role of polyploidy as a significant widespread evolutionary strategy in angiosperms is known. Research on wheat phylogeny has contributed to the understanding of this important phenomenon, but there are still discrepancies and deficiencies in information. The revised phylogenetic scheme of the genus *Triticum*, which contains all known evolutionary information of wild ANO cultivated wheat species, is presented in Fig. 1. There are two evolutionary lines of polyploid wheats, i.e. Emmer and Timopheevii.

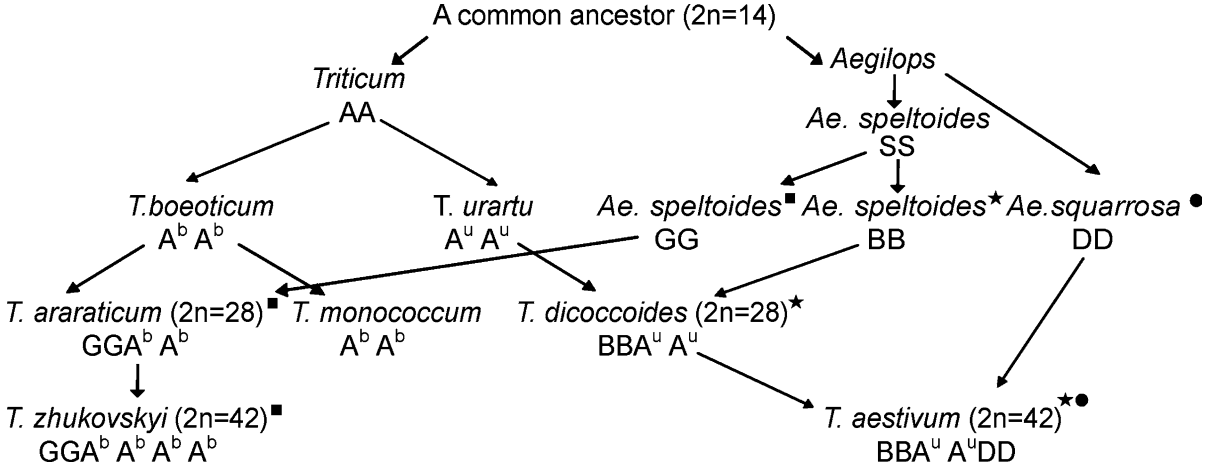

Fig. 1. The scheme represents the origin of polyploid *Triticum* species: *Dicoccoides* Flaksb. (*T. dicoccoides*), *Triticum* (*T. aestivum*) and *Timopheevii* A. Filat. et Dorof. (*T. araraticum*, *T. zhukovskyi*). The letters under representatives denote the genome constitution of corresponding species. The small black square, asterisk and circle indicate G, B and D genome inheritance, respectively

According to this reconstruction, at least four wheat and goatgrass diploid species contributed to the evolution of tetraploid and hexaploid domesticated wheats through amphiploidization. Now it is generally accepted that the donors of the A and D genomes were diploid wheats and goatgrass *Aegilops squarrosa* L. (=syn. *Ae. tauschii* Cosson), respectively (Kihara 1924, McFadden and Sears 1946). There are two different A genomes in the wheat species: the A^u genome in *T. urartu* Thum. ex Gandil. (Mandy 1970) and the closely related A^b genome in *T. boeoticum* Boiss. (Lilienfeld and Kihara 1934). The latter is proposed as an ancestor of cultivated *T. monococcum* L. A genome (Fig. 1). However, several questions related to allopolyploid evolution have not been addressed together. The identity of the donor(s) of the B and G genomes remains open. Many different species have been proposed as the original donor of these genomes, but it is now largely believed that the progenitor was a member of section *Sitopsis* of the genus *Aegilops*, namely *Ae. bicornis* (Forsk.) Jaub. et Spach, *Ae. longissima* Schweinf. et Muschl., *Ae. searsii* Feld. & Kis. or, most likely, *Ae. speltoides* Tausch (Tsunewaki and Ogihara 1983, Ogihara and Tsunewaki 1988, Provan et al. 2004). At the same time, it was hypothesized that the B-genome of polyploid wheat (Emmer group) is of a polyphyletic origin, i.e. it is a recombined genome derived from two or more diploid *Aegilops* species. This is supported by molecular data based on low-copy non-coding chromosome specific DNA sequences (Liu et al. 2003).

Recently, the development of molecular technologies, sequencing procedures and comparative DNA analysis, have become important instruments for evolutionary biologists. These approaches resolve many contradictions and provide significant insight into evolutionary events. The nuclear genome is inherited in a biparental manner, whereas chloroplast and mitochondrial genomes are generally inherited maternally (Korpelainen 2004). Biparental inheritance and polyploidy in nuclear genomes can cause difficulties for evolutionary analyses, whereas analyses of variations in chloroplast DNA (cpDNA) are effective for investigating relationships between plants because of maternal inheritance (Matsuoka et al. 2002). Uniparental inheritance lowers the impact of intermolecular recombination and helps simplify theories of chloroplast genome evolution in most plant taxa. Sequence comparison of cpDNA provides basic information supporting comparative evolutionary research at different taxonomic levels. To confirm evolutionary studies and to get more reliable results, it is fruitful to reconstruct phylogeny based on several data sets of different genomes (chloroplast, mitochondrial and nuclear). Polyploidy complicates such investigations that are based on nuclear DNA sequences. However it is possible to overcome the difficulties with a polyploid nuclear genomes by reconstructing molecular phylogeny based on separate nuclear genomes (Liu et al. 2003, Vakhitov et al. 2003). In these approaches, it is necessary to select specific primer sets for selective amplification of each genome.

Different genes from nuclear and chloroplast genomes have been useful for deducing the phylogenetic relationships of plant species. Sequences of internal transcribed spacers (ITS) are the most often used nuclear DNA sequences for revealing relationships at the inter- and intrageneric levels in plants (Barker et al. 2001). The coding sequences of the plastid genome (*rbc*L, *rps*4, *ndh*F, *atp*B) have lower rates of nucleotide substitutions than noncoding sequences (*atp*B-*rbc*L intergenic spacer, *trn*L intron, *trn*L-*trn*F intergenic spacer) and were successfully used in phylogenetic studies at the higher taxonomic level (Taberlet et al. 1991, Lewis and Doyle 2001, Lia et al. 2001, Mason-Gamer et al. 2002, Nishikawa et al. 2002, Kato et al. 2003, Makarevich et al. 2003, Neel and Cummings 2004). Nevertheless, coding sequences, such as those of the *mat*K gene that is the most rapidly evolving plastid gene, can provide sufficient information for reconstructions of phylogenetic relationships at the intrageneric level (Young and dePamphilis 2000, Yang et al. 2004).

Previous evolutionary analyses of the genus *Triticum* based on molecular data were restricted to diploid and some tetra- and hexaploid wheat species (Blake et al. 1998, Faris et al. 2001, Huang et al. 2002, Zhang et al. 2002, Liu et al. 2003, Yamane and Kawahara 2005). Therefore, there was no comprehensive phylogenetic study of wheat that included all representative species of the genus *Triticum*. Such data would provide a better understanding of wheat evolutionary history.

To elucidate this problem, the results of a molecular phylogenetic analysis including all *Triticum* species currently described and putative donors of polyploid wheat genomes of the genus *Aegilops* based on sequence comparison of chloroplast and nuclear (B genome) DNA are presented.

Materials and methods

Plant materials. Accessions of different wheat species and subspecies, goatgrass and rye *Secale cereale* L. and intergeneric synthetics were obtained from N.I. Vavilov All-Russian Institute of Plant Industry (St-Petersburg, Russia), Plant Germ Plasm Institute of Kyoto University (Kyoto, Japan), the National Small Grains Collection (Aberdeen, USA), International Centre for Agriculture Research in the Dry Areas (Aleppo, Syria) and Instituut voor Plantenveredeling, Landbouwhogeschool (Wageningen, the Netherlands) (Table 1). These species and subspecies include all known wheat species biodiversity. The botanical names of wheat species and their genomic formulas are given according to Goncharov (2002, 2005). Authenticity of artificial amphiploids (synthetic wheats) was being studied in other experiments (unpublished data).

Total DNA isolation. Total DNA was isolated from 50–170 mg of one to three acrospires using a standard CTAB method (Rogers and Bendich 1985). DNA were checked by electrophoresis in a 1% agarose gel containing ethidium bromide (0.5 mg/mL) in 1xTAE.

PCR amplification. The *trn*T-*trn*L intergenic spacer and the *trn*L intron were amplified using a, b, and c, d pairs of primers (Taberlet et al. 1991), respectively. Three pairs of primers for the amplification of the fragments of chloroplast *trn*K intron via the polymerase chain reaction (PCR) were designed using *trn*K gene sequence together with the intron of interest of *Triticum aestivum* (GenBank accession number NC002762) Pr1S (5'-GGGTTGCTAACTCAATGGTA-3') and Pr1A (5'-GTTCAGAACCATTTAATCCA-3'), Pr2S (5'-CACTTCTCTTTCAGGAATAT-3') and Pr2A (5'-CATAAAATCGAAGCAAGAGT-3'), Pr3S (5'-TTCTGTTTTTGGACTCAGCC-3') and Pr3A (5'-CGGAACTAGTCGGATGGAGT-3'). In order to amplify a fragment of nuclear DNA belonging to B genome, a multiple alignment program was developed using sequences of the *Acc*-1 (plastid acetyl-CoA carboxylase) gene (GenBank accession numbers AF343496 – AF343536) for different genomes of available *Triticum* and *Aegilops* representatives. A pair of B genome specific primers was designed for the Acc1 T sense (5'-GGT-ATATTATGTTCCTTTTTC- 3') and Acc1 T antisense (5'-TTTAGGCACAGAAATAACAT-3') strands based on the presence in this alignment B genome specific 11 bp deletion (Fig. 2). The selection of the appropriate oligonucleotide sequences was provided with the help of Vector NTI Suite 8.0 program (Lu and Moriyama 2004). All PCR reactions were performed in a 20 μl volume containing 65 mM Tris-HCl (pH 8.9), 16 mM $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, 200 μM of each dNTP, 0.5 μM of each primer, 20–50 ng genomic DNA template, and 1 U of Taq DNA polymerase. The PCR program had an initial strand separation step at 94°C for 3 min followed by 30 cycles of denaturation at 94°C for 30 s, annealing at 42°C for 42 s for chloroplast sequences and at 52°C for 42 s for nuclear sequence, and elongation at 72°C for 1 min. The PCR products were analyzed by agarose electrophoresis and extracted from gel with a Qiaquick Gel Extraction Kit (Qiagene; according to manufacturer protocol).

DNA sequencing. 200 ng of the PCR product was used in a 10 μl cycle sequencing reaction with the ABI BigDye Terminator Kit on an ABI 377 DNA sequencer. The obtained sequences were deposited in GenBank under accession numbers DQ419971 - DQ420055, DQ436342 (Table 1).

Phylogenetic analysis. The nucleotide sequences were aligned using ClustalX software (Thompson et al. 1997). A phylogenetic tree was generated by the neighbor joining (NJ) method using MEGA 3.1 program (Kumar et al. 2004). Statistical

Table 1. Plant accessions used in this study

Genus, section and species	Genome	Accessions	GenBank Ac. No.
***Aegilops* L.**			
section *Sitopsis* (Jaub. et Spach) Zhuk.			
Ae. speltoides Tausch	*SS*	PI 46593	*trn*K *intron*- DQ419987 *mat*K *3'*- DQ420035 *mat*K *5'*- DQ420009
Ae. longissima Schweinf. et Muschl.	S^lS^l	K-569137	*mat*K *3'*- DQ420042
Ae. sharonensis Eig	S^lS^l	Ae47098	*mat*K *3'*- DQ420041
Ae. bicornis (Forsk.) Jaub. et Spach	S^bS^b	Ae47573	*mat*K *3'*- DQ420040
Ae. searsii Feld. & Kis.	S^sS^s	Ae47388	*mat*K *3'*- DQ420055
***section Vertebrata* Zhuk. emend. Kihara**			
Ae. squarrosa L.	*DD*	K-992	*trn*K *intron*- DQ419986 *mat*K *3'*- DQ420045 *mat*K *5'*- DQ420008
***Triticum* L.**			
section *Urartu* Dorof. et A. Filat.			
T. urartu Thum. ex Gandil.	A^uA^u	K-28244	*trn*K *intron*- DQ419991 *mat*K *3'*- DQ420039 *mat*K *5'*- DQ420005 *trn*T-*trn*L spacer- DQ419996 *trn*L *intron*-DQ420001
section *Monococcon* Dum.			
T. boeoticum Boiss.	A^bA^b	K-25811	*trn*K *intron*- DQ419990 *mat*K *3'*- DQ420038 *mat*K 5'- DQ420006 *trn*T-*trn*L spacer- DQ419997 *trn*L *intron*-DQ420002
T. sinskajae A. Filat. et Kurk.	A^bA^b	K-48993	*trn*K *intron*- DQ419988 *mat*K *3'*- DQ420037 *mat*K *5'*- DQ420004
T. monococcum L.	A^bA^b	K-20970	*trn*K *intron*- DQ419992 *mat*K *3'*- DQ420054 *mat*K *5'*- DQ420003 *trn*L *intron*-DQ420000
section *Dicoccoides* Flaksb.			
T. dicoccoides (Körn. ex Aschers. et Graebn.) Schweinf.	BBA^uA^u	K-41965	*Acc*1- DQ419971 *trn*K *intron*- DQ419993 *mat*K 3'- DQ420011 *mat*K *5'*- DQ420010 *trn*T-*trn*L spacer- DQ419995 *trn*L *intron*- DQ419999
T. dicoccum (Schrank) Schuebl.	BBA^uA^u	cv. Dichter	*mat*K *3'*- DQ420020

Table 1. (Continued)

Genus, section and species	Genome	Accessions	GenBank Ac. No.
T. turgidum L.	BBA^uA^u	K-11597	*Acc*1- DQ419972 *mat*K *3'*- DQ420015
T. turanicum Jakubz.	BBA^uA^u	K-15993	*mat*K *3'*- DQ420024
T. polonicum L.	BBA^uA^u	K-22697	*mat*K *3'*- DQ420021
T. aethiopicum Jakubz.	BBA^uA^u	K-19059	*Acc*1- DQ419975 *mat*K *3'*- DQ420027
T. carthlicum Nevski	BBA^uA^u	K-13768	*mat*K *3'*- DQ420017
T. durum Desf.	BBA^uA^u	K-17769	*Acc*1- DQ419976 *mat*K *3'*- DQ420050
T. karamyschevii Nevski	BBA^uA^u	KU 190-1	*Acc*1- DQ419983 *mat*K *3'*- DQ420051
T. ispahanicum Heslot	BBA^uA^u	C1 N 2001	*Acc*1- DQ419973 *mat*K *3'*- DQ420012
section *Triticum*			
T. macha Dekapr. et Menabde	BBA^uA^uDD	K-38547	*mat*K *3'*- DQ420022
T. spelta L.	BBA^uA^uDD	K-20579	*Acc*1- DQ419982 *mat*K *3'*- DQ420019
T. spelta ssp. *tibetanum* (Shao) N. Gontsch.	BBA^uA^uDD	KU 510	*Acc*1- DQ419978 *mat*K *3'*- DQ420014
T. spelta ssp. *yunnanense* (King) N. Gontsch.	BBA^uA^uDD	KU 505	*mat*K *3'*- DQ420033
T. compactum Host	BBA^uA^uDD	WAG 8326	*Acc*1- DQ419974 *mat*K *3'*- DQ420018
T. aestivum L.	BBA^uA^uDD	cv. Chinese Spring	*mat*K *3'*- DQ420013
T. aestivum ssp. *petropavlovskyi* (Udacz. et Migusch.) N. Gontsch.	BBA^uA^uDD	K-43351	*Acc*1- DQ419984 *mat*K *3'*- DQ420028
T. sphaerococcum Perciv.	BBA^uA^uDD	K-14967	*Acc*1- DQ419985 *mat*K *3'*- DQ420016
T. dimococcum Schieman et Staudt	$BBA^uA^uA^bA^b$	KU 229-1	*mat*K *3'*- DQ436342
T. vavilovii (Thum.) Jakubz.	BBA^uA^uDD	KU 192	*Acc*1- DQ419977 *mat*K *3'*- DQ420052
T. antiquorum Heer ex Udacz.*	BBA^uA^uDD	K-56397	*Acc*1- DQ419981 *mat*K *3'*- DQ420049
section *Timopheevii* A. Filat. et Dorof.			
T. araraticum Jakubz.	GGA^bA^b	K-28244	*trn*K *intron*- DQ419989 *mat*K *3'*- DQ420031 *mat*K *5'*- DQ420007 *trn*T-*trn*L spacer- DQ419994 *trn*L *intron*- DQ419998
T. timopheevii (Zhuk.) Zhuk.	GGA^bA^b	K-29347	*mat*K *3'*- DQ420030

Table 1. (Continued)

Genus, section and species	Genome	Accessions	GenBank Ac. No.
T. timopheevii ssp. *militinae* (Zhirov) N. Gontsch.	GGA^bA^b	K-46007	*mat*K *3'*- DQ420032
T. zhukovskyi Menabde et Erizjan	$GGA^tA^tA^bA^b$	K-43063	*mat*K *3'*- DQ420036
T. timococcum Kost.**	$GGA^tA^tA^bA^b$	K-43805	*mat*K *3'*- DQ420053
section *Compositum* N. Gontsch.			
T. palmovae G. Ivanov	DDA^bA^b	K-58045	*mat*K *3'*- DQ420044
T. kiharae Dorof. et Migusch.	GGA^tA^tDD	K-47897	*mat*K *3'*- DQ420029
T. soveticum Zhebrak	$BBA^uA^uGGA^tA^t$	KU 227	*mat*K *3'*- DQ420026
T. soveticum ssp. *fungicidum* (Zhuk.) N. Gontsch.	$BBA^uA^uGGA^tA^t$	K-38258	*mat*K *3'*- DQ420023
T. flaksbergeri Navr.	$GGA^tA^tBBA^uA^u$	K-55251	*mat*K *3'*- DQ420034
***Secale* L.**			
Secale cereale L.	*RR*	cv. Imperial	*mat*K *3'*- DQ420046
***Hordeum* L.**			
Hordeum pusillum Nutt.	*HH*		matK -AB078133****
***Bromus* L.**			
Bromus inermis Leyss.	*AABB***		matK - AF164398****
Intergeneric synthetics			
×*Aegilotriticum* Tschermak	$UUM^oM^oBBA^uA^u$	KU 224	*mat*K *3'*- DQ420043
×*Tritordeum* Aschers. et Graebn.	$HHBBA^uA^uDD$	HT109	*Acc*1- DQ419979
			*mat*K *3'*- DQ420048
×*Triticale* Muntzing (×*Secalotricum* Wittm.)	BBA^uA^uRR	cv. Altayskaya 5	*Acc*1- DQ419980
			*mat*K *3'*- DQ420025

Ae: International Centre for Agriculture Research in the Dry Areas (Aleppo, Syria),
K: N.I. Vavilov All-Russian Institute of Plant Industry (St.-Petersburg, Russia),
KU: Plant Germ Plasm Institute of Kyoto University (Kyoto, Japan),
PI: the National Small Grains Collection (Aberdeen, USA),
WAG: Instituut voor Plantenveredeling, Landbouwhogeschool (Wageningen, the Netherlands),
cv. - cultivars and lines HT109, C1 N2001 from collection of Wheat Genetics Laboratory of Institute of Cytology and Genetics (Novosibirsk, Russia).
* - *T. antiquorum* is analog of Neolithic compact wheat species discovered nowadays in Tajikistan by Prof. R. A. Udachin (Goncharov and Gaidalenok 2005).
** - *T. timococcum* Kost. is synthetic analog of natural *T. zhukovskyi* Menabde et Erizjan (Goncharov 2002).
*** - On the basis of previous cytogenetic studies tetraploid *B. inermis* has the genomic composition of AABB (Armstrong 1991). They are not identical to wheat A and B genomes, probably the author used the first alphabet letters.
****- sequences obtained not in the present study.

support for the tree was evaluated by bootstrapping (1000 replications) (Felsenstein 1985).

For completeness, we also constructed maximum likelihood phylogenies. Two separate analyses were conducted. The topology of the detailed tree constructed by PHYML algorithm (Guindon and Gascuel 2003) coincides with a consensus tree obtained using the "tree-puzzle" approach (Schmidt et al. 2002). Results of both were identical with the previous NJ topology. Here, only phylogenetic tree constructed by NJ algorithm is presented.

The alignment of the *Pgk*-1 gene, that includes the sequences of some wheats, and *Aegilops* genes is available from the corresponding author.

131
163
286
296
312
330
331
346
361
417
588
600
605
750
791
A_T.urartu (AF343518)
A_T.monococcum subsp. monococcum (AF343517)
A_T.timopheevii subsp. timopheevii (AF343497)
A_T.timopheevii subsp. armeniacum (AF343514)
A_T.turgidum subsp. dicoccoides (AF343499)
A_T.turgidum subsp. dicoccoides (AF343511)
A_T.boeoticum (AF343504)
A_T.aestivum (AF343519)
D_Ae.tauschii subsp. tauschii (AF343501)
D_Ae.tauschii subsp. tauschii (AF343496)
D_T.aestivum (AF343498)
S_Ae.searsii (AF343529)
S_Ae.bicornis (AF343521)
S_Ae.longissima (AF343531)
S_Ae.longissima (AF343532)
S_Ae.sharonensis (AF343525)
S_Ae.speltoides subsp. ligustica (AF343535)
S_Ae.speltoides subsp. ligustica (AF343523)
S_Ae.speltoides subsp. ligustica (AF343536)
S_Ae.speltoides subsp. speltoides (AF343508)
S_Ae.speltoides subsp. speltoides (AF343530)
S_Ae.speltoides subsp. speltoides (AF343533)
S_Ae.speltoides subsp. speltoides (AF343520)
S_Ae.speltoides subsp. speltoides (AF343534)
S_Ae.speltoides subsp. speltoides (AF343527)
G_T.timopheevii subsp. armeniacum (AF343503)
G_T.timopheevii subsp. timopheevii (AF343505)
B_T.aestivum (AF343510)
B_T.turgidum subsp. dicoccoides (AF343507)
B_T.turgidum subsp. dicoccoides (AF343506)
B_T.turgidum subsp. dicoccoides (AF343500)
B_T.dicoccoides
B_T.turgidum
B_T.spelta
Consensus
Acc1-T sense 5′GGTATATTATGTTTTAATCCTCCCCTTTTTTC3′
Acc1-T antisense 5′TTTAGGCACAGAAATAACATC3′

Fig. 2. Part of *Acc*-1 gene alignment with the positions and directions of the primers (Acc1-T sense, Acc1-T antisense) used for specific amplification of B genome sequences. Letters at the beginning of each species name depict the definite genome the corresponding sequence belongs to. GenBank accession numbers of each sequence are in brackets. Sequences of *T. dicoccoides*, *T. turgidum*, *T. spelta* obtained in this study are marked in bold letters. We show only three sequences because the remaining ones are identical to them. Numbers above nucleotide sequences depict the positions of the corresponding nucleotide in the whole alignment. Homologous nucleotides are represented as dots and the deletion of a nucleotide as dashes. Dashes in the beginning and at the end of sequences obtained in present study denote the absence of data due to incomplete sequences

Results

Analysis of the chloroplast *trn*T-*trn*L intergenic spacer and *trn*L intron regions. The *trn*T-*trn*L intergenic spacer and *trn*L intron regions of some *Triticum* species belonging to different sections *Monococcon* Dum. (*T. urartu*, *T. monococcum*, *T. boeoticum*), *Timopheevii* A. Filat. et Dorof. (*T. araraticum*), and *Triticum* (*T. dicoccoides*) were amplified and sequenced. These sequences were analyzed along with homologous sequences of *T. aestivum* (GenBank accession number AB042240) and other closely related species examined earlier by other researchers, namely *Ae. tauschii* (AF519113), *Ae. speltoides* (AF519112), and *Ae. uniaristata* Vis. (AF519114) (Mason-Gamer et al. 2002). However, out of the 406 aligned positions in sequences of the *trn*T - *trn*L intergenic spacer from *Triticum* species only two nucleotide substitutions in *T. boeoticum* and *T. urartu* and one in *T. aestivum* and *Ae. tauschii*, which are insufficient for phylogenetic analysis (data not shown) were revealed. The number of variable sites in the alignment of *trn*L intron sequences was similar but there was a principal 10 bp insertion (AAACTCATAA) in *Ae. speltoides*, *T. aestivum*, *T. araraticum* and *T. dicoccoides* that divides all studied species into two groups (Fig. 3). No other cpDNA of diploid *Aegilops* species and wild diploid wheats included in the analysis contained this insertion. Moreover, sequences of the diploid *Triticum* species were distinguished from the others based on three specific nucleotide substitutions (G217 - A217, G319 - T319, A474 - C474, Fig. 3).

Phylogenetic reconstruction based on sequences of *mat*K gene. The *mat*K gene encodes a maturase - like protein and is situated inside the *trn*K (tRNA-Lys) intron (Neuhaus and Link 1987). Recently, a function of *mat*K beyond its presumed host intron–specific maturase activity was speculated corresponding with the evolution of the *trn*K intron itself in plants (Hausner et al. 2006). In this study, the *trn*K intron was examined as a model for group II introns. Structure and function as well as evolutionary constraints of these introns were studied previously (Kelchner 2002).

The *trn*K region together with the *mat*K gene is variable due to a high rate of evolution and is widely used as a marker in molecular phylogenetics at the species and genus levels (Young and dePamphilis 2000). Previously, evolutionary implications of *mat*K indels were investigated in Poaceae Barnhart (Hilu and Alice 1999). We amplified and sequenced the *trn*K intron from chloroplast genomes of six *Triticum* (*T. boeoticum*, *T. monococcum*, *T. sinskajae*, *T. urartu*, *T. dicoccoides*, *T. araraticum*) and two *Aegilops* (*Ae. squarrosa*, *Ae. speltoides*) species belonging to different sections to find the most appropriate molecular marker. The positions and directions of the three pairs of primers designed are shown in Fig. 4.

Within three sequenced regions, the second contained part of the *mat*K gene and was the most variable one (Fig. 4). The expected lengths of the obtained fragments, which were calculated based on *T. aestivum* sequences,

100 base pairs
trnT
trnL 5′ exon
trnL 3′ exon
a
c
b
d
211 220 311 320 381 400 411 420 468 477
Ae.speltoides (SS)*
T.araraticum (GGA^bA^b)
T.dicoccoides (BBA^uA^u)
T.aestivum (BBA^uA^uDD)*
Ae.tauschii (DD)*
Ae.uniaristata (NN)*
T.monococcum (AA^b)
T.urartu (A^uA^u)
T.boeoticum (AA^b)

Fig. 3. Positions and directions of universal primers (a, b, c, d) used to amplify two non-coding regions of cpDNA (Taberlet et al. 1991). The length of the non-coding regions in the figure corresponds to the *T. aestivum* sequence. Below there is a part of the *trn*L intron sequences alignment where the principal variability was found. Numbers above nucleotides sequences depict the positions of the corresponding nucleotide in the whole alignment while arrows highlight the positions of substitutions. The regions of insertion in *Ae. speltoides* and polyploid wheats are marked grey. Nucleotide absence is shown as a dot. Letters in brackets denote the genome constitution of the corresponding species. The sequences that were obtained from GenBank database are marked with asterisks (accession numbers see in the text)

were about 850 bp, 960 bp, and 830 bp for the first, the second and the third fragments, respectively. The matrix made from the first and the third chloroplast regions of the analyzed species after partial sequencing was 872 bp and only two variable positions were found in the third region. Instead of two cytosine positions that were common for *Ae. speltoides*, *Ae. squarosa*, *T. araraticum*, *T. dicoccoides* and *T. aestivum* the cpDNA of diploid *Triticum* species (*T. boeoticum*, *T. monococcum*, *T. sinskajae*, *T. urartu*) had adenine and thymine nucleotides at that position (data not shown). There were three specific substitutions in the sequences of the second region that allowed separation of diploid wheats from the other species (Fig. 4, positions 161, 230, 520). The presence of two other substitutions allows combination of *Ae. speltoides*, *T. araraticum* and *T. dicoccoides* (Fig. 4, positions 302, 443). These data support similar results based on the *trn*L intron sequences (Fig. 3).

For further phylogenetic reconstruction, the nucleotide sequences of the second part of *mat*K gene for 45 samples of the *Triticum*, *Aegilops* and other closely related species were determined. Some of these species were wild and a few polyploid species synthetic (Table 1). Thirty-one of 523 aligned sites were variable and one indel character was coded. A phylogenetic analysis based on the obtained alignment using the neighbor-joining (Saitou and Nei 1987) and maximum likelihood (Schmidt et al. 2002, Guindon and Gascuel 2003) algorithms was performed. The topology of the tree retained was conserved in both approaches. The NJ phylogenetic tree is represented in Fig. 5. Based on the constructed tree topology, all analyzed species except *Secale cereale* L. and × *Tritordeum* Aschers. et Graebn. are subdivided into four related clades: I - *T. dicoccoides*, × *Triticale*, *T. vavilovii*, *T. dimococcum*, *T. spelta* ssp. *tibetanum*, *T. turanicum*, *T. dicoccum*, *T. soveticum* ssp. *fungicidum*, *T. sphaerococcum*, *T. antiquorum*, *T. soveticum*, *T. carthlicum*, *T. compactum*, *T. durum*, *T. aestivum*, *T. ispahanicum*, *T. polonicum*, *T. spelta ssp. yannanense*, *T. macha*, *T. aestivum* ssp. *petropavlovskyi*, *T. aethiopicum*, *T. karamyschevii*, *T. turgidum* and *T. spelta*; II - *T. kiharae*, *T. timoccocum*, *T. timopheevii*, *T. flaksbergeri*, *T. timopheevii* ssp. *militinae*, *Ae. speltoides*, *T. araraticum* and *T. zhukovskyi*; III - *T. boeoticum*, *T. urartu*, *T. sinskajae* and *T. monoccocum*; IV - *Ae. bicornis*, *T. palmovae*, *Ae. squarrosa* ssp. *strangulate*, × *Aegilotriticum*, *Ae. sharonensis*, *Ae. longissima*, *Ae. searsii* (Fig. 5). In the last case, ssp. *strangulata* of *Ae. squarrosa* was used because of its participation in the establishing of polyploid genome as shown earlier (McFadden and Sears 1946, Jaaska 1980, Pestsova et al. 2000).

The sequence of the *mat*K gene from *Bromus inermis* Leyss. was used as an outgroup. The bootstrap test was used for statistical support of the branch pattern of the phylogenetic tree.

Selective amplification of nuclear B genome. As mentioned, polyploidy of many *Triticum* species complicates phylogenetic studies based on their nuclear sequences. In this case, additional cloning procedures and knowledge about the specific characteristics of each genome are necessary to separate sequences of different genomes.

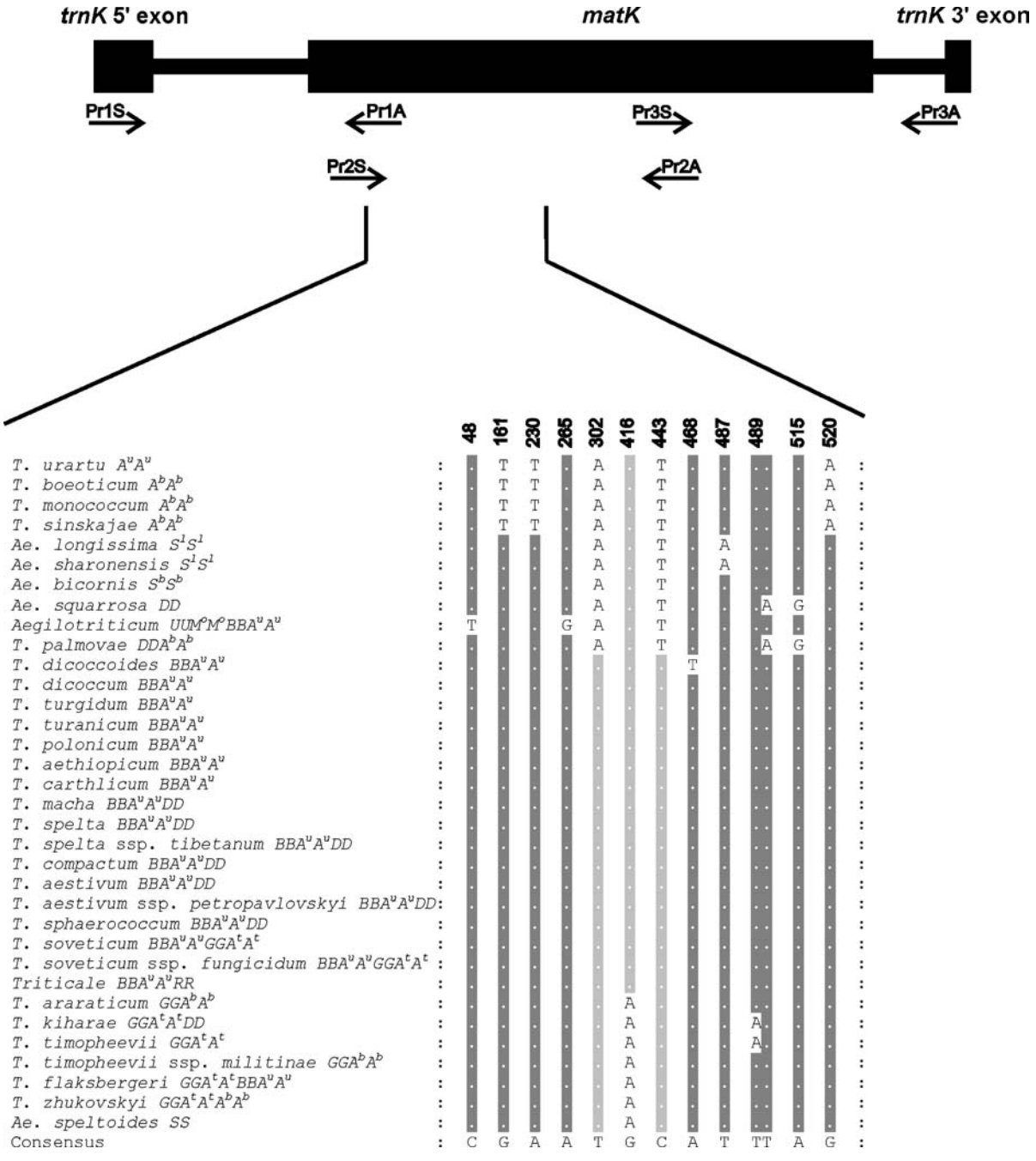

Fig. 4. The structure of the *trn*K gene with the internal *mat*K gene together with positions and directions of designed primers (Pr1S, Pr1A, Pr2S, Pr2A, Pr3S, Pr3A) used in this study. The scheme corresponds to *T. aestivum trn*K sequence. The divergent lines represent an alignment of variable sites inside the sequenced region of the *mat*K gene. Numbers above each column depict the positions of the corresponding nucleotide in the whole alignment. Nucleotides identical to the consensus are represented as dots. Letters next to the name of species denote its genome constitution

Information about different indel events and nucleotide substitutions in each genome may be fruitful for designing genome-specific PCR primers. To obtain this information, available sequences of *Acc*-1 (plastid acetyl-CoA carboxylase) and *Pgk*-1 (plastid 3-phosphoglycerate kinase) genes belonging to different genomes and wheat species were compiled from GenBank (accessions AF343496 - AF343536 and AF343474 - AF343495, respectively) and aligned. In the alignment of *Acc* -1 gene, there was an 11 bp deletion specific for the B genome of *Triticum* species (Fig. 2). A 6 bp insertion and a 6 bp deletion specific for the A genome of wheats was also observed in the alignment of the *Pgk*-1 gene (data are not shown). This alignment also included an 89 bp insertion which was present only in G and S genomes of *T. timopheevii* and *Ae. speltoides*, correspondingly. It is worth noticing that among six analyzed *Ae. speltoides* accessions from different populations three clones, two from one population and one from another, contained this big insertion in the *Pgk*-1 gene intron (data are not shown). No other *Aegilops* genomes have such an insertion. Moreover, in both alignments there were several nucleotide substitutions that were present only in a specific genome. This knowledge may help identify an appropriate genome to which a sequence belongs. Moreover, this alignment revealed a clear distinction of *Acc*-1 sequences of B and G genomes and supports results obtained in previous RFLP analyses of some tetraploid wheats (Mori et al. 1997).

A specific pair of primers for selective amplification of B genome sequences of the *Acc*-1 gene from total DNA (Fig. 2) was designed. PCR analyses included all species from clade I together with total DNA of synthetic species from other clades that contain the B genome (octoploid *Aegilotriticum* (genome $UUM^{o}M^{o}BBA^{u}A^{u}$) and *Tritordeum* (genome $HHBBA^{u}A^{u}DD$). Some species such as *T. araraticum* (genome $GGA^{b}A^{b}$), *T. timopheevii* (genome $GGA^{b}A^{b}$), *T. zhukovskyi* (genome $GGA^{t}A^{t}A^{b}A^{b}$), *Ae. searsii* (genome $S^{s}S^{s}$), *Ae. speltoides* (genome SS) were analyzed to confirm the absence of the B genome or in case of *Aegilops* species to show genomic differences from the B region of interest (Fig. 6). Part of the *Acc*-1 gene of 13 wheat species and two synthetic ones (octoploid *Tritordeum* and hexaploid *Triticale*) was selectively sequenced using PCR-products. Sequences of all *Acc*-1 gene fragments from the B genome were identical in the analyzed polyploid species. The fragments were aligned together with other sequences from different genomes available from GenBank, all were identical and contained nucleotide sequences specific for the B genome (Fig. 2). This demonstrates the B genome origin of these sequences, the diversity of B and G genomes, and the possibility of selective amplification for certain genomic fragments.

Discussion

Our investigation of the phylogenetic relationships within the genus *Triticum* consisted of two principal parts: cpDNA sequences of wheats and their related species were analyzed followed by nuclear B genome analyses. It was necessary to find an appropriate molecular marker for phylogenetic studies because of reduced genetic diversity between *Triticum* representatives resulting from domestication and characterized by small population sizes together with intense selection for agronomic traits (Tanksley and McCouch 1997). Moreover, it was shown that nucleotide sequence diversity was 30-fold higher in *Ae. tauschii* than in the *T. aestivum* D genome and this is compatible with a genetic bottleneck created by recent polyploidization (Caldwell et al. 2004).

The majority of species within the genus *Triticum* are cultivated. The contribution of natural selection in their evolution is limited by man thus explaining decreases in DNA variability.

The principal indel and nucleotide substitution events in the chloroplast *trn*T-*trn*L region. Regions of the chloroplast *trn*T (tRNA - Thr) and *trn*L (tRNA - Leu) genes, along with their intervening non-coding regions, were reported useful in phylogenetic

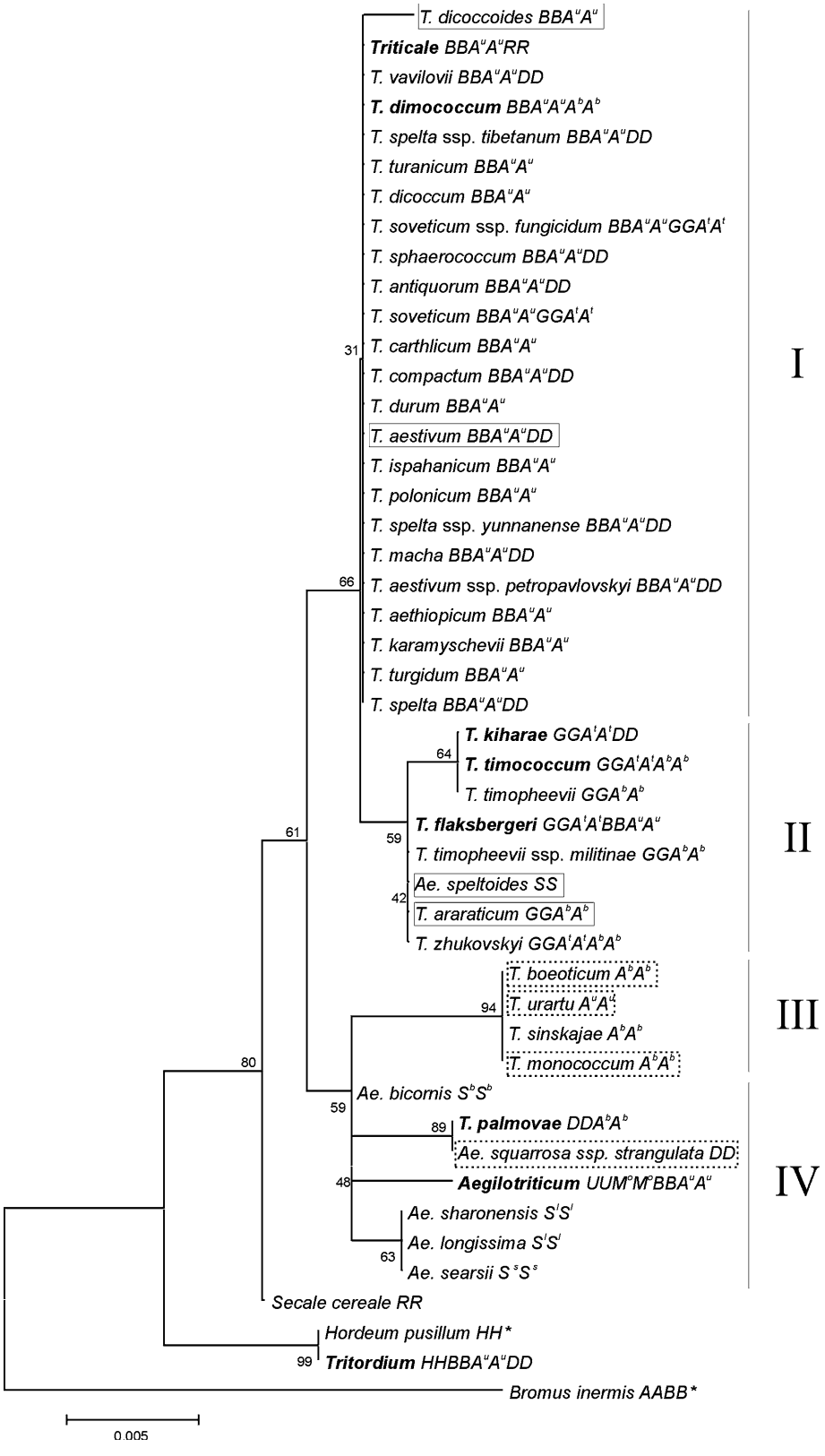
T. dicoccoides BBAuAu
Triticale BBAuAuRR
T. vavilovii BBAuAuDD
T. dimococcum BBAuAuAbAb
T. spelta ssp. tibetanum BBAuAuDD
T. turanicum BBAuAu
T. dicoccum BBAuAu
T. soveticum ssp. fungicidum BBAuAuGGAtAt
T. sphaerococcum BBAuAuDD
T. antiquorum BBAuAuDD
T. soveticum BBAuAuGGAtAt
T. carthlicum BBAuAu
T. compactum BBAuAuDD
T. durum BBAuAu
T. aestivum BBAuAuDD
T. ispahanicum BBAuAu
T. polonicum BBAuAu
T. spelta ssp. yunnanense BBAuAuDD
T. macha BBAuAuDD
T. aestivum ssp. petropavlovskyi BBAuAuDD
T. aethiopicum BBAuAu
T. karamyschevii BBAuAu
T. turgidum BBAuAu
T. spelta BBAuAuDD
T. kiharae GGAtAtDD
T. timococcum GGAtAtAbAb
T. timopheevii GGAbAb
T. flaksbergeri GGAtAtBBAuAu
T. timopheevii ssp. militinae GGAbAb
Ae. speltoides SS
T. araraticum GGAbAb
T. zhukovskyi GGAtAtAbAb
T. boeoticum AbAb
T. urartu AuAu
T. sinskajae AbAb
T. monococcum AbAb
Ae. bicornis SbSb
T. palmovae DDAbAb
Ae. squarrosa ssp. strangulata DD
Aegilotriticum UUMcMcBBAuAu
Ae. sharonensis SlSl
Ae. longissima SlSl
Ae. searsii SsSs
Secale cereale RR
Hordeum pusillum HH*
Tritordium HHBBAuAuDD
Bromus inermis AABB*
31
66
64
59
61
42
94
80
59
89
48
63
99
I
II
III
IV
0.005

Fig. 5. Neighbor-Joining phylogenetic tree based on the comparison of *mat*K sequences. Four observed clades are shown by solid lines on the right. There is a genome composition near each species. Synthetic wheats are represented in bold letters. Based on the indel event in the *trn*L intron sequence of some analyzed species, representatives with observed insertions are marked by solid boxes and the remaining ones by dotted boxes on the phylogenetic tree. Asteriscs denote species from which the *mat*K sequence was obtained from GenBank. Accession numbers are in the Table 1

reconstructions at the intrageneric taxonomic level (Mason-Gamer et al. 2002, Makarevich et al. 2003). In this study the sequences of the *trn*T-*trn*L intergenic spacer together with *trn*L intron were determined for five *Triticum* species (*T. urartu*, *T. monococcum*, *T. boeoticum*, *T. araraticum*, *T. dicoccoides*). The corresponding fragments of some other species (*T. aestivum*, *Ae. tauschii*, *Ae. speltoides*, *Ae. uniaristata*) available from GenBank were also included in the analysis. A 10 bp insertion was observed in the *trn*L intron sequences of *Ae. speltoides*, *T. aestivum*, *T. araraticum* and *T. dicoccoides* that divides all studied species into two groups (Fig. 3), and three specific nucleotide substitutions were observed in the sequences of diploid wheats (Fig. 3).

Based on the obtained results, it is concluded that the special insertion, probably together with the three described substitutions, appeared in *Ae. speltoides* chloroplast genome and were inherited in cpDNA of polyploid *Triticum* species (*T. aestivum*, *T. araraticum* and *T. dicoccoides*). Hence *Ae. speltoides* (genome SS) chloroplasts are related to chloroplasts of *T. aestivum* (genome BBA^uA^uDD), *T. araraticum* (genome GGA^bA^b) and *T. dicoccoides* (genome BBA^uA^u). These plastids may appear in wheat species only as a result of hybridization events involving *Ae. speltoides*, but not other *Aegilops* species of section *Sitopsis*. Parallel evolution towards acquiring the same insertion and substitutions between *Triticum* species and *Ae. speltoides* appears to be less likely.

Phylogenetic tree topology represents the most probable B and G genome ancestor. The phylogenetic tree constructed represents evolutionary process in the genus *Triticum* corresponding with chloroplast origin. Each of four phylogenetic groups includes species that contain the same nuclear genome (Fig. 5). Because the phylogenetic analysis in this study was done based on sequence comparison of

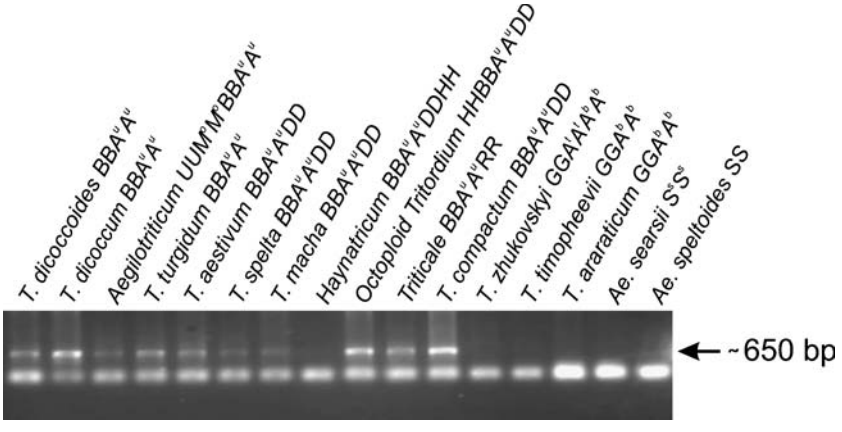

Fig. 6. B genome-specific amplification of *Acc*-1 gene among *Triticum* and closely related species. The size of corresponding band is shown on the right

cpDNA, the species are clustered according to the chloroplast origin. Therefore, species with identical nuclear genomes in their polyploid genome contribution and which form a common clade inherited the same plasmon. This means that there were maternal plants, which participated in speciation at the last hybridization events and which were the donors of the definite nuclear genome and plasmon at the same time. Moreover, it is possible to determine these donors. In support of a previous study, it was shown that the *T. timopheevii* chloroplast genome was inherited together with nuclear G genome as a result of a hybridization event between *Ae. speltoides* and *T. boeoticum* (Chen et al. 1975).

All diploid wheats can be robustly delineated and separated from *Aegilops* species, clade III and IV (Fig. 5). This is supported by unique substitutions in the *trn*L intron sequence separating the diploid *Triticum* species from other investigated plants (Fig. 3). Both groups originated from a common ancestor and constitute a cluster demonstrating close relatedness of diploid *Triticum* and *Aegilops* species.

The topology of our tree was supported also by the presence of unique substitutions and insertions identified in the *trn*L and *trn*K intron sequences of some *Triticum* and *Aegilops* species (Figs. 3 and 4). Based on the analysis, polyploid wheat species containing B and G genomic sequences were clearly divided into two separate groups (Fig. 5). As a result, it is concluded that this division occurred after the ancestor of B and G genomes were derived from the forefather of A, D and the other studied *Aegilops* genomes, namely U, S^b, S^l and S^s. Furthermore, *Ae. speltoides* genome S appears most closely related to the B and G genomes of wheats.

Phylogenetic differences between *Ae. speltoides*, wild diploid *Triticum* species and the rest of *Aegilops* species. An obvious separation of the *Ae. speltoides* chloroplast genome (II clade) from other *Aegilops* (IV clade) and diploid *Triticum* (III clade) species is observed. Additionally, this study included all five species from section *Sitopsis* of the genus *Aegilops* (Table 1), which were considered potential B and G genome donors to polyploid wheats by different authors (Breiman 1987, Gill and Appels 1988, Terachi and Tsunewaki 1992, Sasanuma et al. 1996). Differences between *Ae. speltoides* and the other *Sitopsis* species have been detected in previous works (Brody and Mendlinger 1980, Ogihara and Tsunewaki 1988, Giorgi et al. 2002, Yamane and Kawahara 2005). Eig was the first researcher who separated and placed this species in another subsection *Truncata* Eig, while the remaining species belonged to subsection *Emarginata* Eig (Eig 1929). Based on data presented, both *trn*L and *trn*K intron sequences of *Ae. speltoides* are more variable than the corresponding sequences from the other *Aegilops* and diploid *Triticum* species. This observation strongly coincides with previous results based on nucleotide variations of the four other chloroplast non-coding regions and microsatellite repeat motifs (Yamane and Kawahare 2005). The topology of the trees in both studies clearly demonstrates that the *Ae. speltoides* ancestor branched off before a separation of wild diploid *Triticum* and *Aegilops* species (Fig. 5). It supports a suggestion to separate *Ae. speltoides* into a new genus and probably combine wild diploid *Triticum* and *Aegilops* species into one genus (Yamane and Kawahare 2005). On the other hand, a problem the with classification of polyploid *Triticum* species occurs with this suggestion. In this case, it would be necessary to classify them with *Ae. speltoides*.

At the same time, a high degree of intraspecific variability in nuclear and mitochondrial DNA (mtDNA) of *Ae. speltiodes* accessions was shown based on RFLP markers (Breiman et al. 1991), protein electrophoretic characters (Asins and Carbonell 1986) and recent sequence comparison analysis of *Acc*-1 and *Pgk*-1 genes (Huang et al. 2002). Breiman and colleagues (1991) proposed both types of variation (nuclear and mitochondrial) are selectively neutral and that they are affected only by the effective population size, which is higher in outbreeding (*Ae. speltoides*) than inbreeding (*T. monococcum*, *Ae. squarrosa*)

species. Moreover, *Ae. speltoides* is a cross-pollinated species whereas all other *Aegilops* and *Triticum* species are self-pollinated (Hammer and Matzk 1993); therefore *Ae. speltoides* is the most divergent species within section *Sitopsis* (Giorgi et al. 2002).

Two *Ae. speltoides* ancestor accessions participated in the last polyploidization events of Triticum species from clades I and II. Among all analyzed species of section *Sitopsis*, *Ae. speltoides* is the most closely related to polyploid wheats, which have B or G genomes (i. e. Emmer and Timopheevii groups). A similar result was obtained in the previous plasmon analysis (PCR-SSCP) of *Triticum* and *Aegilops* species where six *Triticum* species were investigated (Wang et al. 1997). It was shown that only one of the *Aegilops* species, *Ae. speltoides*, was included in the Emmer group, supposing that this species is the donor of the plasmon and B and G nuclear genomes of all polyploid wheats. However, the *Ae. speltoides* plasmon is more closely related to the species of the Timopheevii group than to that of the Emmer group. This result also revealed that the polyploid wheat species constitute two evolutionary lineages (Fig. 5).

Recently, similarity of *Ae. speltoides* and *T. timopheevii* mitochondrial sequences and their difference from Emmer wheats was shown (Hedgcoth et al. 2002). These results and the phylogenetic tree topology that was constructed in the present study support the close relatedness of wheat G genome to the one of *Ae. speltoides*. At the same time, the tree shows that wheat species with B and G genomes in their genome are sister groups where *Ae. speltoides* constitutes the common clade with G genome species and is the most closely related *Aegilops* species to both sister branches.

Analysis of the *Pgk*-1 gene alignment revealed the presence of a specific 89 bp insertion in the intron of the *Pgk*-1 gene of three different *Ae. speltoides* representatives and its absence in another three accessions. This may confirm intraspecific variation in *Ae. speltoides* that is likely a reason for existence of two different ancestral *Ae. speltoides* forms, which gave rise to two evolutionarily close lineages of polyploid wheats (clade I and II). Moreover the mentioned 89 bp insertion was present only in G and S genomes of *T. timopheevii* and *Ae. speltoides* correspondingly and not in other *Aegilops* and *Triticum* genomes.

Based on our results and taking into account previous works, it is proposed that the evolutionary history of the last polyploidization event of *Triticum* species involved an *Ae. speltoides* ancestor. It is likely that there were two ancestral forms of *Ae. speltoides* involved in the two-step process of hybridization. The high degree of intraspecific variation observed among *Ae. speltoides* accessions and differentiation into B and G genomes of polyploid wheats supports this hypothesis.

Strictly maternal inheritance of cpDNA in *Triticum* and *Aegilops* species. Studying synthetic wheats provides the information about pathways (scheme of hybridization) which were used to produce them. It is known that the inheritance of cytoplasmic genomes is not universally maternal (Birky 1995, Korpelainen 2004). However, the phylogenetic tree developed in this study showed strictly maternal plastid inheritance in all studied synthetic species. Absence of paternal inheritance was also observed in all synthetic plants included in the study. Clustering of *T. palmovae* (genome DDA^bA^b), together with *Ae. squarrosa* (genome DD), is naturally due to a participation of *Ae. squarrosa* in speciation of *T. palmovae* as a maternal parent (Fig. 5). Octoploid *Tritordeum* HT 109 (genome $HHBBA^uA^uDD$) is a hybrid of *Hordeum vulgare* L. as the maternal plant and *T. aestivum* as the paternal one. Grouping of *Tritordeum* together with *H. pusillum* Nuttal. (GenBank accession number AB078133) demonstrates relatedness of their plastids and, as a consequence, their maternal inheritance (Fig. 5). The maternal plant of hexaploid *Triticale* (genome BBA^uA^uRR) is a *Triticum* species, whereas *Secale sereale* L. is the second parent. *Triticale* are included in clade I; this clade consists of polyploid wheat species with

genome constitution including B genome. This implies *Triticale* inherited both chloroplasts and nuclear genomes from its maternal parent. A significant pattern of inheritance in chloroplasts of octoploid *Aegilotriticum* KU224 (genome $UUM^{o}M^{o}BBA^{u}A^{u}$) was observed. This species was obtained by hybridization of *Ae. ovata* L. as the maternal parent and *T. durum* as the paternal one (Tschermak and Bleier 1926). The location of this synthetic cereal in clade III supports the statement about *Aegilotriticum* relationships with other investigated *Aegilops* species.

B genome similarity of all investigated wheat species. Each of five different *Aegilops* species was proposed to be a putative donor of B genome to tetra- and hexaploid wheat species: *Ae. bicornis*, *Ae. longissima*, *Ae. searsii*, *Ae. sharonensis* and *Ae. speltoides*. Independent assessment of the phylogenetic analysis based on cpDNA may be provided by comparison of nuclear sequences. Therefore, to support the topology of our phylogenetic tree and to improve its resolution within the group of polyploid wheats with B genome, the nuclear *Acc*-1 and *Pgk*-1 genes were analyzed. Nucleotide sequences of these genes were used to determine phylogenetic relationships between *Triticum* and *Aegilops* species of the wheat lineage and to establish a timeline of wheat evolution based on gene sequence comparisons (Huang et al. 2002).

The data suggest that the number of substitution is insufficient in the obtained sequences to reveal phylogenetic relationships among polyploid *Triticum* species. This may occur due to its relatively recent origin and because the B genome of wheat underwent a genetic bottleneck associated with human domestication. Based on alignment of the *Acc*-1 gene and PCR results (Fig. 6), the contemporary *Ae. speltoides* genome is more closely related to the G genome of wheats from section *Timopheevii* but not to other *Aegilops* species belonging to section *Sitopsis* (Fig. 2). Nevertheless, taking into account the presence of unique nucleotides and indels in the cpDNA and nuclear DNA regions of species from clade I and II, B and G genomes are closely related to each other; *Ae. speltoides* is the most likely donor of both plasmons and genomes.

This study also demonstrates the possibility of selective amplification of different genomes of polyploid species, particularly the B genome. Previously, successful selective amplification of D genome regions was performed in *T. aestivum, Ae. cylindrica* Host and *Ae. tauschii* (Caldwell et al. 2004).

Conclusion

The phylogeny of domesticated wheats has been reported (McFadden and Sears 1946, Mandy 1970, Jaaska 1980, Tsunewaki and Ogihara 1983, Kerby and Kuspira 1986, Goncharov 2002, Feldman and Levy 2005). Presence of two cytoplasm types in the polyploid species of the genus *Triticum* suggests a possibility of their diphyletic origin (Tsunewaki et al. 1976, Mori et al. 1995). This hypothesis was based on earlier hybridological and cytological analyses (Lilienfeld and Kihara 1934). However, opposite reasoning takes place (Tanaka et al. 1978). It is now apparent that DNA analysis is one of the most effective methods for studying phylogeny and evolution (Flavell et al. 1979). Results in the present study verify the previous hypothesis and reveal the existence of two ancestral *Ae. speltoides* forms, which were the donors of different but very closely related plasmons and genomes of polyploid wheat species.

Analysis of nuclear and chloroplast DNA allowed clear division of polyploid wheats Emmer (sections *Dicoccoides* and *Triticum*) and Timopheevii (section *Timopheevii*), i.e. into B and G genome species. Artificial amphiploids are related to a particular group depending on the maternal parents used. This result corroborates a previous suggestion of diphyletic origin of polyploid wheats based on chloroplast sequences and confirms that the chloroplast genome has coevolved and differentiated with the nuclear one in *Triticum* and *Aegilops*.

These data also correspond to the common classification: the genus *Triticum* consists of

three distinct groups: diploids (section *Monococcon*), tetra- and hexaploids (sections *Dicoccoides* and *Triticum*) and timopheevii (section *Timopheevii*). Each group includes both wild and cultivated wheat species. Section *Monococcon* includes wild species *T. boeoticum* and *T. urartu*, section *Dicoccoides* - *T. dicoccoides*, section *Timopheevii* - *T. araraticum*.

We are grateful to Dr. H. Bockelman (the National Small Grains Collection, Aberdeen, USA), Dr. T. Kawahara (Graduate School of Agriculture of Kyoto Univ., Kyoto, Japan), Drs. R.A. Udachin, O.P. Mitrofanova and N.A. Anfilova (N.I. Vavilov All-Russian Institute of Plant Industry, St-Petersburg, Russia) and D. J. Valkoun (International Centre for Agriculture Research in the Dry Areas, Aleppo, Syria) for supplying seeds of wheat species. The research was financed on the Subprogram II of Program basic research N25 of the Russian Academy of Sciences "The Origin and Evolution of Biosphere".

References

Armstrong K. C. (1991) Chromosome evolution in *Bromus*. In: Tsuchiya T., Gupta T. K. (eds.) Chromosome engineering in plants: genetics, breeding, evolution. Part B. Elsevier, Amsterdam, pp. 363–317.

Asins M. J., Carbonell E. A. (1986) A comparative study on variability and phylogeny of *Triticum* species. Theor. Appl. Genet. 72: 551–558.

Barker N. P., Clark L. G., Davis J. I., Duvall M. R., Guala G. F., Hsiao C., Kellogg E. A., Linder H. P. (2001) Phylogeny and subfamilial classification of the grasses (Poaceae). Ann. Missouri Bot. Gard. 88: 373–457.

Birky C. W. (1995) Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. Proc. Natl. Acad. Sci. USA 92: 11331–11338.

Blake N. K., Lehfeldt B. R., Lavin M., Talbert L. E. (1998) Phylogenetic reconstruction based on low copy DNA sequence data in an allopolyploid: The B genome of the wheat. Genome 42: 351–360.

Breiman A. (1987) Mitochondrial DNA diversity in the genera of *Triticum* and *Aegilops* revealed by Southern blot hybridization. Theor. Appl. Genet. 73: 563–570.

Breiman A., Bogher M., Sternberg H., Graur D. (1991) Variability and uniformity of mitochondrial DNA in populations of putative diploid ancestors of common wheat. Theor. Appl. Genet. 82: 201–208.

Brody T., Mendlinger S. (1980) Species relationships and genetic variation in the diploid wheats (*Triticum*, *Aegilops*) as revealed by starch gel electrophoresis. Pl. Syst. Evol. 136: 247–258.

Caldwell K. S., Dvorak J., Lagudah E. S., Akhunov E., Luo M.-C., Wolters P., Powell W. (2004) Sequence polymorphism in polyploid wheat and their D-genome diploid ancestor. Genetics 167: 941–947.

Chen K., Gray J. C., Wildman S. G. (1975) Fraction I protein and the origin of polyploid wheats. Science 190: 1304–1306.

Eig A. (1929) Monographisch-kritische Übersicht der Gattung *Aegilops*. Repertorium specierum novarum regni vegetabilis, Beihefte (Berlin) 55: 1–228.

FAO (2004) FAO Statistical Databases. http: //apps.fao.org

Faris J., Sirikhachornkit A., Haselkorn R., Gill B., Gornicki P. (2001) Chromosome mapping and phylogenetic analysis of the cytosolic acetyl-CoA carboxylase loci in wheat. Molec. Biol. Evol. 18: 1720–1733.

Feldman M., Levy A. A. (2005) Allopolyploidy – a shaping force in the evolution of wheat genomes. Cytogenet. Genome Res. 109: 250–258.

Felsenstein J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791.

Flavell R. B., O'Dell M., Smith D. B. (1979) Repeated sequence DNA comparisons between *Triticum* and *Aegilops* species. J. Heredity 42: 309–322.

Gill B. S., Appels R. (1988) Relationships between *Nor*-loci from different Triticeae species. Pl. Syst. Evol. 160: 77–89.

Gill B. S., Appels R., Botha-Oberholster A. et al. (15 co-authors) (2004) A workshop report on wheat genome sequencing: International genome research on wheat consortium. Genetics 168: 1087–1096.

Giorgi D., D'Ovidio R., Tanzarella O. A., Porceddu E. (2002) RFLP analysis of *Aegilops* species belonging to the *Sitopsis* section. Genet. Resour. Crop Evol. 49: 145–151.

Goncharov N. P. (2002) Comparative genetics of wheats and their related species. Novosibirsk,

Siberian University Press (in Russian with English Summary).

Goncharov N. P. (2005) Comparative-genetic analysis – a base for wheat taxonomy revision. Czech J. Genet and Plant Breed. 41. Special issue: 52–54.

Goncharov N. P., Gaidalenok R. F. (2005) Localization of genes controlling spherical grain and compact ear in *Triticum antiquorum* Heer ex Udacz. Russian J. Genet. 41: 1262–1267.

Gu Y. Q., Coleman-Derr D., Kong X., Anderson O. D. (2004) Rapid genome evolution revealed by comparative sequence analysis of orthologous regions from four Triticeae genomes. Pl. Physiol. 135: 459–470.

Guindon S., Gascuel O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52: 696–704.

Hammer K., Matzk F. (1993) Variation in breeding systems in the Triticeae. In: Damania A. B. (ed.) Biodiversity and wheat improvement. A Wiley-Sayce Publ., pp. 51–58.

Hammer K., Filatenko A. A., Korzun V. (2000) Microsatellite markers - a new tool for distinguishing diploid wheat species. Genet. Resour. Crop Evol. 47: 497–505.

Harlan J. R. (1992) Crops and man. 2nd ed. American Soc. of Agronomy, CSSA, Madison, Wisconsin.

Hausner G., Olson R., Simon D., Johnson I., Sanders E. R., Karol K. G., McCourt R. M., Zimmerly S. (2006) Origin and evolution of the chloroplast *trn*K (*mat*K) intron: a model for evolution of Group II intron RNA structures. Molec. Biol. Evol. 23: 380–391.

Hedgcoth C., El-Shehawi A. M., Wei P., Clarkson M., Tamalis D. (2002) A chimeric open reading frame associated with cytoplasmic male sterility in alloplasmic wheat with *Triticum timopheevi* mitochondria is present in several *Triticum* and *Aegilops* species, barley, and rye. Curr. Genet. 41: 357–365.

Hilu K. W., Alice L. A. (1999) Evolutionary implications of *mat*K indels in Poaceae. Amer. J. Bot. 86: 1735–1741.

Huang S., Sirikhachornkit A., Su X., Faris J., Gill B., Haselkorn R., Gornicki P. (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum-Aegilops* complex and the evolutionary history of polyploid wheat. Proc. Natl. Acad. Sci. USA 99: 8133–8138.

Jaaska V. (1980) Electrophoretic survey of seedling esterases in wheats in relation to their phylogeny. Theor. Appl. Genet. 56: 273–284.

Jauhar P., Crane C. F. (1989) An evaluation of Baum et al.'s assessment of the genomic system of classification in the Triticeae. Amer. J. Bot. 76: 571–576.

Kato Y., Aioi K., Omori Y., Takahat N., Satta Y. (2003) Phylogenetic analysis of *Zostera* species based on *rbc*L and *mat*K nucleotide sequences: Implications for the origin and diversification of the seagrasses in Japanese waters. Genes Genet. Syst. 78: 329–342.

Kelchner S. A. (2002) Group II introns as phylogenetic tools: structure, function, and evolutionary constraints. Amer. J. Bot. 89: 1651–1669.

Kellogg E. A. (1989) Comments on genomic genera in the Triticeae (Poaceae). Amer. J. Bot. 76: 796–805.

Kerby K., Kuspira J. (1986) The phylogeny of polyploid wheats *Triticum aestivum* (bread wheat) and *Triticum turgidum* (macaroni wheat). Genome 29: 722–737.

Kihara H. (1924) Cytologische und genetische Studien bei wichtigen Getreidearten mit besonderer Rücksicht auf das Verhalten der Chromosomen und die Sterilität in den Bastarden. Mem. Coll. Sci. Kyoto. Imp. Univ. 1: 1–200.

Korpelainen H. (2004) The evolutionary processes of mitochondrial and chloroplast genomes differ from those of nuclear genomes. Naturwissenschaften 91: 505–518.

Kumar S, Tamura K., Nei M. (2004) MEGA3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment. Brief. Bioinform. 5: 150–163.

Larionov D. (1914) Several notes on the problem of genetic relationship between some *Triticum* genus species in connection with their taxonomy. Tr. Prikl. Bot. 7: 363–379.

Lewis C. E., Doyle J. J. (2001) Phylogenetic utility of the nuclear gene malate synthase in the palm family (Arecaceae). Molec. Phylogenet. Evol. 19: 409–420.

Lia V. V., Confalonieri V. A., Comas C. I., Hunziker J. H. (2001) Molecular phylogeny of *Larrea* and its allies (Zygophyllaceae): reticulate evolution and the probable time of creosote bush arrival to North America. Molec. Phylogenet. Evol. 21: 309–320.

Lilienfeld F., Kihara H. (1934) Genomanalyse bei *Triticum* und *Aegilops*. V. *Triticum timopheevi* Zhuk. Cytologia 6: 87–122.

Liu B., Segal G., Rong J. K., Feldman M. (2003) A chromosome-specific sequence common to the B genome of polyploid wheat and *Aegilops searsii*. Pl. Syst. Evol. 241: 55–66.

Lu G., Moriyama E. N. (2004) Vector NTI, a balanced all-in-one sequence analysis suite. Brief. Bioinform. 5: 378–388.

Makarevich I., Golovnina K., Scherbik S., Blinov A. (2003) Phylogenetic relationships of the Siberian *Iris* species inferred from noncoding chloroplast DNA sequences. Int. J. Pl. Sci. 164: 229–237.

Mandy G. (1970) New concept of the origin of *Triticum aestivum*. Acta Agronomica Hungarica. 19: 413–417.

Mason-Gamer R. J., Orme N. L., Anderson C. M. (2002) Phylogenetic analysis of North American *Elymus* and the monogenomic Triticeae (Poaceae) using three chloroplast DNA data sets. Genome 45: 991–1002.

Matsuoka Y., Yamazaki Y., Ogihara Y., Tsunewaki K. (2002) Whole chloroplast genome comparison of rice, maize, and wheat: implications for chloroplast gene diversification and phylogeny of cereals. Molec. Biol. Evol. 19: 2084–2091.

McFadden E. S., Sears E. R. (1946) The origin of *Triticum spelta* and its free-threshing hexaploid relatives. J. Heredity 37: 81–89.

Mori N., Liu Y.-G., Tsunewaki T. (1995) Wheat phylogeny determined by RFLP analysis of nuclear DNA. 2. Wild tetraploid wheats. Theor. Appl. Genet. 90: 129–134.

Mori N., Moriguchi T., Nakamura C. (1997) RFLP analysis of nuclear DNA for study of phylogeny and domestication of tetraploid wheat. Genes Genet. Syst. 72: 153–161.

Neel M. C., Cummings M. P. (2004) Section-level relationships of North American *Agalinis* (Orobanchaceae) based on DNA sequence analysis of three chloroplast gene regions. BMC Evol. Biol. 4: 15.

Neuhaus H., Link G. (1987) The chloroplast tRNALys(UUU) gene from mustard (*Sinapis alba*) contains a class II intron potentially coding for a maturase-related polypeptide. Curr. Genet. 11: 251–257.

Nishikawa T., Salomon B., Komatsuda T., von Bothmer R., Kadowaki K.-I. (2002) Molecular phylogeny of the genus *Hordeum* using three chloroplast DNA sequences. Genome 45: 1157–1166.

Ogihara Y., Tsunewaki K. (1988) Diversity and evolution of chloroplast DNA in *Triticum* and *Aegilops* as revealed by restriction fragment analysis. Theor. Appl. Genet. 76: 321–332.

Percival J. (1921) The wheat plant: A monograph. Duckworth and Co., London.

Pestsova E., Korsun V., Goncharov N. P., Hammer K., Röder M. S. (2000) Microsatellite analysis of *Aegilops tauschii* germplasm. Theor. Appl. Genet. 101: 100–106.

Provan J., Wolters P., Caldwell K. H., Powell W. (2004) High-resolution organellar genome analysis of *Triticum* and *Aegilops* sheds new light on cytoplasm evolution in wheat. Theor. Appl. Genet. 108: 1182–1190.

Riley R., Unrau J., Chapman V. (1958) Evidence on the origin of the *B* genome of wheat. J. Heredity 49: 91–98.

Rogers S. O., Bendich A. J. (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Pl. Molec. Biol. 5: 69–76.

Saitou N., Nei M. (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molec. Biol. Evol. 4: 406–425.

Sakamura T. (1918) Kurze Mitteilung über die Chromosomenzahlen und die Verwandtschaftsverhältnisse der *Triticum*-Arten. Bot. Mag. (Tokyo) 32: 151–154.

Sasanuma T., Miyashita N. T., Tsunewaki K. (1996) Wheat phylogeny determined by RFLP analysis of nuclear DNA. 3. Intra- and interspecific variation of five *Aegilops* Sitopsis species. Theor. Appl. Genet. 92: 928–934.

Schmidt H. A., Strimmer K., Vingron M., von Haeseler A. (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics. 18: 502–504.

Taberlet P., Gielly L., Pautou G., Bouvet J. (1991) Universal primers for amplification of three noncoding regions of chloroplast DNA. Pl. Molec. Biol. 17: 1105–1109.

Tanaka M., Kawahara T., Sano J. (1978) The evolution of wild tetraploid wheats. Proceed. 5th Intern. Wheat Genet. Symp. New Delhi. 1: 73–80.

Tanksley S., McCouch S. (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277: 1063–1066.

Terachi T., Tsunewaki K. (1992) The molecular basis of genetic diversity among cytoplasms of *Triticum* and *Aegilops*. VIII. Mitochondrial RFLP analysis using cloned gene as probes. Molec. Biol. Evol. 9: 917–931.

Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F., Higgins D. G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 15: 4876–4882.

Tschermak E., von Bleier H. (1926) Über fruchtbare *Aegilops*-Weizenbastarde (Beispiele für die Entstehung neuer Arten durch Bastardierung). Berichte der Deutsch. Bot. Gesellschaft 44: 110–132.

Tsunewaki K., Mukai Y., Tsuji S., Murata M. (1976) Genetic diversity of the cytoplasm in *Triticum* and *Aegilops*. V. Classification of 23 cytoplasm into eight plasma types. Japan. J. Genet. 51: 175–191.

Tsunewaki K., Ogihara Y. (1983) The molecular basis of genetic diversity among cytoplasms of *Triticum* and *Aegilops*. II. On the origin of polyploid wheat cytoplasms as suggested by chloroplast DNA restriction fragment patterns. Genetics 104: 155–171.

Tsunewaki K., Wang G.-Z., Matsuoka Y. (2002) Plasmon analysis of *Triticum* (wheat) and *Aegilops*. 2. Characterization and classification of 47 plasmons based on their effects on common wheat phenotype. Genes Genet. Syst. 77: 409–427.

Vakhitov V. A., Chemeris A. V., Sabirzhanov B. E., Akhunov E. D., Kulikov A. M., Nikonorov I. M., Gimalov F. R., Bikbulatova S. M., Baimiev A. Kh. (2003) The phylogeny of *Triticum* L. and *Aegilops* L. inferred from comparative analysis of nucleotide sequences in rDNA promoter regions. Russ. J. Genet. 39: 1–11.

Yamane K., Kawahara T. (2005) Intra- and interspecific phylogenetic relationships among diploid *Triticum-Aegilops* species (Poaceae) based on base-pair substitutions, indels, and microsatellites in chloroplast noncoding sequences. Amer. J. Bot. 92: 1887–1898.

Yang D.-Y., Fushimi H., Cai S.-Q., Komatsu K. (2004) Molecular analysis of *Rheum* species used as rhei rhizoma based on the chloroplast *mat*K gene sequence and its application for identification. Biol. Pharm. Bull. 27: 375–383.

Young N. D., dePamphilis C. W. (2000) Purifying selection detected in the plastid gene *mat*K and flanking ribozyme regions within a Group II intron of nonphotosynthetic plants. Molec. Biol. Evol. 17: 1933–1941.

Wang G.-Z., Miyashita N. T., Tsunewaki K. (1997) Plasmon analyses of *Triticum* (wheat) and *Aegilops*: PCR–single-strand conformational polymorphism (PCR-SSCP) analyses of organellar DNAs. Proc. Natl. Acad. Sci. USA 94: 14570–14577.

Zhang W., Qu L.-J., Gu H., Gao W., Liu M., Chen J., Chen Z. (2002) Studies on the origin and evolution of the tetraploid wheats based on the internal transcribed spacer (ITS) sequences of nuclear ribosomal DNA. Theor. Appl. Genet. 104: 1099–1106.

Addresses of the authors: Kseniya A. Golovnina (e-mail: ksu@bionet.nsc.ru), Sergei A. Glushkov, Alexander G. Blinov, and Nikolay P. Goncharov, Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Prospekt Lavrentyeva 10, Novosibirsk 630090, Russia. Vladimir I. Mayorov and Linda R. Adkison, Mercer University School of Medicine, 1550 College St., Macon GA 31207, USA.