Molecular Characterization of a Diagnostic DNA Marker for Domesticated Tetraploid Wheat Provides Evidence for Gene Flow from Wild Tetraploid Wheat to Hexaploid Wheat

Jan Dvorak, Eduard D. Akhunov, Alina R. Akhunov, Karin R. Deal, and Ming-Cheng Luo

Department of Plant Sciences, University of California, Davis

All forms of domesticated tetraploid wheat (*Triticum turgidum*, genomes AABB) are nearly monomorphic for restriction fragment length polymorphism (RFLP) haplotype *a* at the *Xpsr920* locus on chromosome 4A (*Xpsr920-A1a*), and wild tetraploid wheat is monomorphic for haplotype *b*. The *Xpsr920-A1a*/*b* dimorphism provides a molecular marker for domesticated and wild tetraploid wheat, respectively. Hexaploid wheat (*Triticum aestivum*, genomes AABBDD) is polymorphic for the 2 haplotypes. Bacterial artificial chromosome (BAC) clones hybridizing with PSR920 were isolated from *Triticum urartu* (genomes AA), *Triticum monococcum* (genomes A^mA^m), and *T. turgidum* ssp. *durum* (genomes AABB) and sequenced. PSR920 is a fragment of a putative ATP binding cassette (ABC) transporter gene (designated *ABCT-1*). The wheat *ABCT-1* gene is more similar to the *T. urartu* gene than to the *T. monococcum* gene and diverged from the *T. urartu* gene about 0.7 MYA. The comparison of the sequence of the wheat A genome BAC clone with that of the *T. urartu* BAC clone provides the first insight into the microsynteny of the wheat A genome with that of *T. urartu*. Within 103 kb of orthologous intergenic space, 37 kb of new DNA has been inserted and 36 kb deleted leaving 49.7% of the region syntenic between the clones. The nucleotide substitution rate in the syntenic intergenic space has been 1.6×10^{-8} nt^{-1} $year^{-1}$, which is, respectively, 4 and 3 times as great as nucleotide substitution rates in the introns and the third codon positions of the juxtaposed gene. The RFLP is caused by a miniature inverted transposable element (MITE) insertion into intron 18 of the *ABCT-A1* gene. Polymerase chain reaction primers were developed for the amplification of the MITE insertion site and its sequencing. The *T. aestivum ABCT-A1a* haplotype is identical to the haplotype of domesticated tetraploid wheat, and the *ABCT-A1b* haplotype is identical to that of wild tetraploid wheat. This finding shows for the first time that wild tetraploid wheat participated in the evolution of hexaploid wheat. A cline of the 2 haplotype frequencies exists across Euro-Asia in *T. aestivum*. It is suggested that *T. aestivum* in eastern Asia conserved the gene pool of the original *T. aestivum* more than wheat elsewhere.

Introduction

Wheat species form a classical polyploid series at 3 ploidy levels (fig. 1). At the diploid level, there are 2 einkorn wheats: *Triticum monococcum* and *Triticum urartu*. The latter contributed the A genome of tetraploid *Triticum turgidum* and *Triticum timopheevii* (Dvorak et al. 1988; Dvorak et al. 1993). The 2 tetraploids founded separate evolutionary lineages in the genus (fig. 1) that are reproductively isolated from each other by a strong sterility barrier and low levels of meiotic chromosome pairing. The lineage of *T. turgidum* and *Triticum aestivum* was founded by *T. turgidum* ssp. *dicoccoides* (henceforth wild emmer), which is the only truly wild polyploid wheat in this lineage.

The archaeological record shows that emmer was domesticated about 10 000 years ago (Willcox 1997). Ozkan et al. (2002) suggested the northern part of the Fertile Crescent as the site of emmer domestication, but the absence of wild emmer populations from many areas in that region precluded more precise identification of the site. Investigating a more complete sample of wild emmer, Mori et al. (2003) concluded that emmer was domesticated in the Kartal Dagi mountains northeast of Gaziantep in Turkey. Including these materials into the reassessment of their previous study, Ozkan et al. (2005) concluded that emmer was domesticated either in the Karaca Dag mountain region west of Dyiarbakir in southeastern Turkey and/or the Sulaimaniya region in Iran. Luo et al. (2006) showed that the Sulaimaniya region is an unlikely candidate site for emmer domestication and pinpointed emmer domestication to the Karaca Dag mountain region. Luo et al. (2006) further showed that the domesticated emmer (*T. turgidum* ssp. *dicoccon*) gene pool was enriched by gene flow from wild emmer in southern Levant (Lebanon, southwestern Syria, and Israel) either by separate domestication of emmer in that region and absorption of that gene pool into the gene pool of domesticated emmer diffusing from Turkey or by introgressive hybridization between domesticated and wild emmer in southern Levant. The gene flow from wild to domesticated emmer in southern Levant was probably the principle factor that caused the subdivision of domesticated emmer into 2 basic populations, northern and southern (Luo et al. 2006).

Triticum aestivum originated via hybridization of tetraploid wheat with *Aegilops tauschii* (fig. 1) (Kihara 1944; McFadden and Sears 1946) in Armenia—southwestern Caspian region (Dvorak, Luo, Yang, and Zhang 1998). *Triticum aestivum* comprises a number of forms that are either hulled or free threshing. Free-threshing bread wheat (*T. aestivum* ssp. *aestivum*) and club wheat (*T. aestivum* ssp. *compactum*) are the principal wheats of commerce. Hulled spelt (*T. aestivum* ssp. *spelta*) was an important cereal in Europe in Roman times and the Middle Ages but today is grown on a very limited scale in Europe and several places in Asia. The remaining hulled wheats are endemics of no economical significance. What is believed to be a free-threshing hexaploid wheat begins to appear in the archaeological record in Anatolia as early as 8500 years ago (for review, see Nesbitt and Samuel 1996).

Reproductive isolation between the tetraploid and hexaploid levels is weak in wheat and pentaploid hybrids

Key words: polyploidy, MITE, microsynteny, intergenic space, retrotransposon, divergence.

E-mail: jdvorak@ucdavis.edu.

Mol. Biol. Evol. 23(7):1386–1396. 2006
doi:10.1093/molbev/msl004
Advance Access publication May 4, 2006

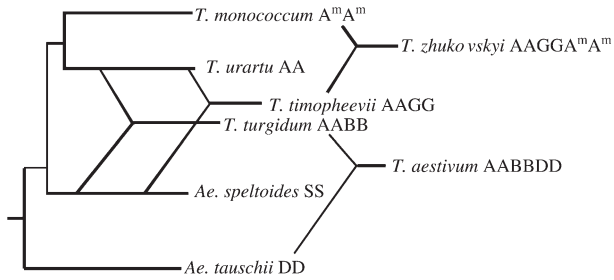

FIG. 1.—Wheat evolutionary lineages. The capital letters behind species names indicate genome formulas; sharing the same letters indicates sharing the same genomes. Note that S, B, and G are closely related in this scenario, and designation by different letters is maintained for historical reasons.

from $4x \times 6x$ hybridization backcross to either ploidy level, forming a bridge via which genes could potentially flow between the ploidy levels. It has been speculated that European spelt may have originated from hybridization of domesticated emmer with free-threshing club or bread wheat (Schiemann 1951; Mac Key 1966; Liu and Tsunewaki 1991; Yan et al. 2003). Tsunewaki (1968) speculated that *T. aestivum* ssp. *macha*, having hulled seeds and brittle spike rachis, may have originated from hybridization between tetraploid wild or domesticated emmer and hexaploid wheat.

Here we present molecular evidence revealing the existence of past gene flow from wild tetraploid wheat to cultivated hexaploid wheat based on genetic variation at the single-copy *Xpsr920* locus in the A genome. In polyploid wheat, the *Xpsr920* loci form an orthologous set on chromosomes 4A, 4B, and 4D, designated *Xpsr920-A1*, *-B1*, and *-D1*, respectively. A single orthologue was also found in *T. urartu*, *Xpsr920-A1*, and *T. monococcum*, *Xpsr920-A^{m}1* (Devos et al. 1995; Dvorak and Luo 1999). The *Xpsr920-A1* locus is in the proximal region of the long arm of chromosome 4A, which suffered several inversions and 2 translocations (Devos et al. 1995). Except for the 4A/5A translocation, which happened at the diploid level, the remaining structural rearrangements originated during the evolution of wild emmer. Because heterozygosity for an inversion is a potent recombination suppressor in the inverted region, fixation of 4A inversions during the evolution of wild emmer must have swept away genetic variation from most of this chromosome. Most of the polymorphism that exists in the inverted regions of this chromosome probably originated since the fixation of these inversions.

Restriction fragment length polymorphism (RFLP) for the 9.5- and 10.0-kb *Dra*I restriction fragments at the *Xpsr920-A1* locus (Dvorak and Luo 1999) is an example of such polymorphism. A preliminary study suggested that wild tetraploid wheat is nearly monomorphic for the 9.5-kb restriction fragment, whereas domesticated tetraploid wheat is nearly monomorphic for the 10.0-kb restriction fragment (Dvorak and Luo 1999). *Triticum urartu* appeared to be monomorphic for the 9.5-kb restriction fragment suggesting that the 9.5-kb fragment was ancestral, and the 10.0-kb fragment was derived from it either by a mutation of 1 of the 2 *Dra*I restriction sites or by the insertion of about 0.5 kb of DNA. *Triticum monococcum* was monomorphic for a 14-kb *Dra*I fragment and wheat chromosome 4B for a 6.5-kb *Dra*I fragment (Dvorak and Luo 1999).

Puzzling variation at the *Xpsr920-A1* locus was observed at the hexaploid level. *Triticum aestivum* was expected to be monomorphic for the 10.0-kb *Dra*I fragment because it was believed that domesticated emmer was its parent (Jaaska 1978; Porceddu and Lafiandra 1986; Kimber and Sears 1987). However, a preliminary study suggested that *T. aestivum* was polymorphic for the 9.5- and 10.0-kb fragments (Dvorak and Luo 1999). Because evidence based on RFLP is intrinsically inconclusive, the 10.0-kb fragment could have been derived from the 9.5-kb fragment by a DNA insertion in tetraploid wheat and then reversed by a deletion of approximately the same size in *T. aestivum*; the logical conclusion that wild tetraploid wheat was one of the ancestors of hexaploid wheat could not be drawn.

Because of the potential significance of this finding for the understanding of wheat evolution, we cloned and sequenced *Xpsr920* haplotypes relevant to the resolution of this dilemma. Here we report comparisons of the *Xpsr920* genomic region in the genomes of diploid and tetraploid wheats and the use of these sequences for the assessment of the relationships of the *T. aestivum* haplotypes with those of tetraploid wheat. Sequencing of *T. urartu* and wheat bacterial artificial chromosome (BAC) clones facilitated the first comparison of microsynteny of the wheat A genome with that of its putative ancestor, *T. urartu*, and results of that comparison, as well as comparison of microsynteny of these clones with an orthologous B genome clone, are reported.

Materials and Methods

Restriction Fragment Length Polymorphism

A total of 281 accessions of wild emmer were used to estimate the frequency of *Xpsr920-A1a* and *Xpsr920-A1b* haplotypes by Southern hybridization of PSR920 with genomic DNAs. All geographic regions in which this species is found were sampled: Israel, Lebanon, southwestern Syria, southern Turkey (Gaziantep and Urfa regions), the Karaca Dag mountains in southeastern Turkey (Diyarbakir region), the Silvan region east of Dyiarbakir, northern Iraq, and northwestern Iran (table 1). Frequencies of the 2 haplotypes were also estimated in 196 accessions of domesticated emmer and 137 accessions of free-threshing tetraploid wheats (table 1). The frequencies of the 2 haplotypes were estimated in 77 accessions of hulled hexaploid wheats (*T. aestivum* ssp. *spelta* and *macha*) and 368 accessions of free-threshing wheat (table 1). Finally RFLP at the *Xpsr920* locus was assessed in 51 accessions of wild *T. monococcum* ssp. *aegilopoides* and 391 accessions of *T. urartu* across the distribution of these species.

Nuclear DNAs were isolated according to Dvorak et al. (1988). They were digested with *Dra*I restriction endonuclease and Southern blots were hybridized with the PSR920 *Pst*I clone (supplied by M. D. Gale, John Innes Institute, Norwich, United Kingdom) as described by Dvorak et al. (2004).

To determine the nature of the *Xpsr920-A1a* and *Xpsr920-A1b* haplotypes, durum wheat (*T. turgidum* ssp. *durum*) cv. Langdon and bread wheat cv. Chinese Spring bearing the *Xpsr920-A1a* haplotype and wild emmer PI428020 bearing the *Xpsr420-A1b* haplotype and Chinese Spring nullisomic-4A–tetrasomic-4B (henceforth

Table 1
Number of Accessions (*N*) in Populations of Diploid, Tetraploid, and Hexaploid Wheats and Estimated Frequencies of the *Xpsr920-A1a* and *Xpsr920-A1b* Haplotypes Generated by Southern Hybridization of Genomic DNAs with the PSR920 *Pst*I Clone

Species	Subspecies	Geographic Region	*N*	*Xpsr920-A1a*	*Xpsr920-A1b*
T. urartu	—	Armenia	7	0 (0.00)	7 (1.00)
		Iran	6	0 (0.00)	6 (1.00)
		Lebanon	82	0 (0.00)	82 (1.00)
		Syria	162	0 (0.00)	162 (1.00)
		Turkey	128	0 (0.00)	128 (1.00)
		Unknown	6	0 (0.00)	6 (1.00)
T. monococcum	*aegilopoides*	Armenia	1	0 (0.00)	0 (0.00)
		Azerbaijan	6	0 (0.00)	0 (0.00)
		Crimea	1	0 (0.00)	0 (0.00)
		Iran	5	0 (0.00)	0 (0.00)
		Iraq	9	0 (0.00)	0 (0.00)
		Israel	2	0 (0.00)	0 (0.00)
		Lebanon	4	0 (0.00)	0 (0.00)
		Syria	13	0 (0.00)	0 (0.00)
		Turkey	10	0 (0.00)	0 (0.00)
T. timopheevii	*armeniacum*	Transcaucasia	2	0 (0.00)	2 (1.00)
		Iran	4	0 (0.00)	4 (1.00)
		Iraq	13	0 (0.00)	13 (1.00)
		Turkey	1	0 (0.00)	1 (1.00)
T. turgidum	*dicoccoides*	Israel	73	0 (0.00)	73 (1.00)
		Lebanon	52	1 (0.02)	51 (0.98)
		Southwestern Syria	20	1 (0.05)	19 (0.95)
		Turkey Gaziantep	21	0 (0.00)	21 (1.00)
		Turkey Urfa	29	0 (0.00)	29 (1.00)
		Turkey Dyiarbakir	58	0 (0.00)	58 (1.00)
		Turkey Silvan	2	0 (0.00)	2 (1.00)
		Iraq	19	0 (0.00)	19 (1.00)
		Iran	7	0 (0.00)	7 (1.00)
T. turgidum	*dicoccon*	Ethiopia	39	39 (1.00)	0 (0.00)
		Oman	3	3 (1.00)	0 (0.00)
		Israel	7	7 (1.00)	0 (0.00)
		Jordan	1	1 (1.00)	0 (0.00)
		Lebanon	1	1 (1.00)	0 (0.00)
		Syria	2	2 (1.00)	0 (0.00)
		Italy	12	12 (1.00)	0 (0.00)
		Spain	5	5 (1.00)	0 (0.00)
		Central Europe	1	1 (1.00)	0 (0.00)
		Balkans	20	20 (1.00)	0 (0.00)
		Turkey	17	15 (1.00)	2 (0.00)
		Transcaucasia	23	23 (1.00)	0 (0.00)
		Ukraine–Russia	11	11 (1.00)	0 (0.00)
		Iran	46	46 (1.00)	0 (0.00)
		India	8	8 (1.00)	0 (0.00)
T. turgidum	*durum*	Egypt	21	21 (1.00)	0 (0.00)
		Western Europe	4	4 (1.00)	0 (0.00)
		Central Europe	3	3 (1.00)	0 (0.00)
		Eastern Europe	1	1 (1.00)	0 (0.00)
		Southern Europe	1	1 (1.00)	0 (0.00)
		Turkey	1	1 (1.00)	0 (0.00)
		Iran	20	19 (0.95)	1 (0.05)
		Syria	1	1 (1.00)	0 (0.00)
		Pakistan	1	1 (1.00)	0 (0.00)
		China	1	1 (1.00)	0 (0.00)
		North America	1	1 (1.00)	0 (0.00)
T. turgidum	*turgidum*	Egypt	3	3 (1.00)	0 (0.00)
T. turgidum	*turanicum*	Iran	54	54 (1.00)	0 (0.00)
T. turgidum	*polonicum*	Turkey–Iran	4	4 (1.00)	0 (0.00)
T. turgidum	*carthlicum*	Turkey	14	14 (1.00)	0 (0.00)
T. turgidum	*ispahanicum*	Iran	7	7 (1.00)	0 (0.00)
T. aestivum	*spelta*	Europe	51	2 (0.04)	49 (0.96)
		Asia	13	11 (0.85)	2 (0.15)
T. aestivum	*macha*	Georgia	13	0 (0.00)	13 (1.00)
T. aestivum	*aestivum*, *compactum*	China	144	134 (0.93)	10 (0.07)
		Pakistan	6	5 (0.83)	1 (0.17)
		Afghanistan	13	7 (0.54)	6 (0.46)
		Iran	92	46 (0.50)	50 (0.50)
		Turkey	57	16 (0.28)	41 (0.72)
		Europe	36	3 (0.08)	33 (0.92)
		North America	16	11 (0.69)	5 (0.31)
T. aestivum	*carthlicoides*	Turkey	4	0 (0.00)	4 (1.00)

N4AT4B) were digested with *Apa*I, *Bam*HI, *Bgl*II, *Dra*I, *Eco*RI, *Eco*RV, *Hin*dIII, *Kpn*I, *Sac*I, and *Xba*I, and Southern blots were hybridized with PSR920 and compared.

To estimate gene diversity at loci along chromosomes 4A and 4B, Southern blots of DNAs of 58 accessions of wild emmer from the Dyiarbakir region and 117 landraces of the northern population of domesticated emmer (the Balkans, Russia, Turkey, Transcaucasia, and Iran) were hybridized with BCD734, BCD1006, BCD1262, BCD1302, BCD1652, PSR115, PSR153, PSR920, PSR921, PSR922, UCW39, WG464, WG876, and BE443449. The BCD clones were supplied by M. E. Sorrells, Cornell University; the PSR clones were supplied by M. D. Gale; the UCW clone was supplied by J. Dubcovsky, University of California, Davis; the WG clones were supplied by A. Kleinhofs, Washington State University, Pullman; and the BE clone was supplied by O. D. Anderson, US Department of Agriculture (USDA)/ARS Albany, California. Expected heterozygosity (*He*) was computed for each locus and was used as an estimate of gene diversity at a locus. Mean expected heterozygosity was used to estimate average gene diversity on chromosomes 4A and 4B in wild and domesticated emmer.

Cloning of *Xpsr920* Loci

The BAC library of durum wheat Langdon (Cenci et al. 2003) was used as the source of the *Xpsr920-A1a* haplotype, and the BAC library of *T. urartu* G1812 (Akhunov et al. 2005) was used as the source of the *Xpsr920-A1b* haplotype. The Langdon BAC library was also used as a source of the *Xpsr920-B1* haplotype. Finally, a *T. monococcum* BAC library (Lijavetzky et al. 1999) was used as a source of the *T. monococcum* haplotype. In all, 9, 11, and 28 high-density membranes, each containing 18 432 double-printed clones, of the *T. urartu*, *T. monococcum*, and Langdon libraries were respectively hybridized with radioactively labeled probe of PSR920 as described by Akhunov et al. (2005). DNAs of clones hybridizing with PSR920 were isolated using the standard alkaline–sodium dodecyl sulfate lysis technique. DNAs of the clones and genomic DNAs of *T. urartu* and Langdon were digested with *Dra*I restriction endonuclease. Restriction fragments were fractionated in a 1% agarose gel, denatured, blotted onto a Hybond N+ membrane (Amersham, Piscataway, NJ), and hybridized with PSR920. Clones corresponding to the *Xpsr920-A1a*, -A^m*1*, -*A1b*, and -*B1* haplotypes were identified by comparing the profiles of restriction fragments hybridizing with PSR920 in Southern blots of BAC clones with those of *T. urartu* G1812, *T. monococcum* ssp. *monococcum* DV92, and durum cv. Langdon genomic DNAs digested with *Dra*I, respectively.

DNA Sequencing

BAC DNAs were isolated using the QIAGEN Large Construct Kit (Qiagen, Valencia, CA). Shotgun libraries were prepared as described by Stein et al. (2000). DNA sequencing was performed using BigDye v 3.1 (Applied Biosystems), and products of the sequencing reaction were resolved with an ABI 3730xl fragment analyzer (Applied Biosystems, Foster City, CA). Base calling and assembly of contigs were performed using Phred/Phrap/Consed software (Gordon et al. 1998). Repeat DNA content of sequenced BAC clones was determined by comparison with the Triticeae Repeat Sequence (TREP) and GIRI databases. The coding potential of sequences was established by comparison of translated BAC sequences with the National Center for Biotechnology Information (NCBI) nonredundant database using the BlastX program and by comparison of BAC sequences with the NCBI database of expressed sequence tags (ESTs) using the BlastN program. Haplotype sequence comparison at the *Xpsr920* locus was performed with the BlastN program.

Direct sequencing of polymerase chain reaction (PCR) amplicons was used. PCR products were treated with Exonuclease I (USB, Cleveland, OH) and Shrimp Alkaline Phosphatase (USB) as described in the manual provided by the manufacturer. Three microliters of a purified PCR product were used in 10-μl sequencing reaction with BigDye v 3.1 (Applied Biosystems). Standard sequencing protocols provided by ABI were followed. Trace files were processed using the Sequencher program (Gene Codes, Ann Arbor, MI). Sequences were aligned using the ClustalW program followed by manual editing of alignments. The proportion of nucleotide sites at which 2 sequences were different (*p*-distance) and phylogenetic analysis of nucleotide sequences were performed with MEGA software (http://www.megasoftware.net).

PCR Conditions

A primer pair designated PSR920-UCD1 consisting of GCAAGTTCAAGCTGTGACCTGTG (left primer) and ATGCATAAATGGAGTTCAAAGTCT (right primer) was used to PCR amplify genomic DNA from the *Xpsr920-A1* locus. The following PCR regime was used: 94 °C for 5 min; 10 touchdown cycles with 0.5 °C decrease in annealing temperature starting with 94 °C for 20 s, 63 °C for 20 s, and 72 °C for 2 min followed by 35 cycles of 94 °C for 20 s, 58 °C for 30 s, and 72 °C for 2 min; and final extension at 72 °C for 10 min. An ABI 9700 thermocycler was used.

Results

Restriction Fragment Length Polymorphism

Genomic DNAs of durum wheat cv. Langdon and bread wheat cv. Chinese Spring, both having the *Xpsr920-A1a* haplotype, wild emmer PI428020, having the *Xpsr920-A1b* haplotype, and Chinese Spring N4AT4B were digested with *Apa*I, *Bam*HI, *Bgl*II, *Dra*I, *Eco*RI, *Eco*RV, *Hin*dIII, *Kpn*I, *Sac*I, and *Xba*I. Southern blots were hybridized with PSR920. Except for *Dra*I, no restriction endonuclease revealed RFLP associated with chromosome 4A, indicating that the *Xpsr920-A1a*/*b* polymorphism is caused either by a mutation of one of the *Dra*I sites or an insertion of DNA far away from the PSR920 hybridization site. *Dra*I restriction endonuclease was therefore used to discriminate between the *Xpsr920-A1a* and *Xpsr920-A1b* haplotypes.

The *Xpsr920-A1a*/*b* RFLP was investigated in 281 accessions of wild emmer, representing the distribution of the species from Israel to Iran and 196 accessions of domesticated emmer (table 1). Of 281 accessions of wild emmer, the *Xpsr920-A1b* haplotype was present in 278 accessions and the *Xpsr920-A1a* haplotype was present

Table 2
Mean Gene Diversity at 13 Loci on Chromosomes 4A and 4B in Wild Emmer in the Dyiarbakir Region in Southeastern Turkey and Northern Population of Domesticated Emmer

Chromosome	Domestication Status	Diversity
4A	Wild	0.111
4A	Domesticated	0.041
4B	Wild	0.215
4B	Domesticated	0.146

in one accession from Lebanon, one from southwestern Syria, and one from Azerbaijan. The latter was intermediate between wild and domesticated wheat and therefore disregarded. In contrast, of 196 domesticated emmer accessions, 194 had the *Xpsr920-A1a* haplotype and 2 had the *Xpsr920-A1b* haplotype, one from the Gaziantep region in southern Turkey and one from the Tosya region in northern Turkey. Of 137 accessions of free-threshing cultivated tetraploid wheats, 136 had the *Xpsr920-A1a* haplotype, whereas 1 durum accession from Iran had the *Xpsr920-A1b* haplotype (table 1).

To evaluate the status of gene diversity along emmer chromosomes 4A and 4B, RFLP at 13 loci on chromosomes 4A and 4B was assessed in 58 accessions of wild emmer from the Dyiarbakir region and 117 landraces of the northern domesticated emmer population. Wild emmer from the Dyiarbakir region was used for this comparison because it was the putative source of domesticated emmer, and northern domesticated emmer was used because it was not affected by the gene flow from wild emmer that occurred in southern Levant. Chromosome 4A was less polymorphic than chromosome 4B in both wild and domesticated emmer (table 2, $P = 0.04$, 2×2 factorial ANOVA), presumably reflecting fixation of complex rearrangements during the evolution of wild emmer. Gene diversity was lower in domesticated emmer than in wild emmer at all loci in the long arm of chromosome 4B, but the reverse was true for loci in the short arm of chromosome 4B (fig. 2). In both wild and domesticated emmer, all centromeric loci linked to *Xpsr920-A1* were either monomorphic or nearly monomorphic (fig. 2). However, in domesticated emmer, chromosome 4A lacked variation along most of the long arm; only the most distal marker in the long arm was polymorphic (fig. 2).

Both the *Xpsr920-A1a* and *Xpsr920-A1b* haplotypes were present in hexaploid *T. aestivum*, and their frequencies formed an east–west cline in free-threshing bread and club wheats. The *Xpsr920-A1b* haplotype frequency was 0.07 in China, 0.17 in Pakistan, 0.46 in Afghanistan, 0.50 in Iran, 0.72 in Turkey, and 0.92 in Europe (table 1). The frequencies of the *Xpsr420-A1a* haplotype formed the reverse cline. The *Xpsr920-A1b* haplotype frequency was 0.15 in Asian spelt (Transcaucasia, Iran, Tadjikistan, and Afghanistan) and 0.96 in European spelt, paralleling the frequencies of the *Xpsr920-A1b* in free-threshing *T. aestivum* in these regions. The *Xpsr920-A1b* haplotype was fixed in *T. aestivum* ssp. *carthlicoides* but was absent from its tetraploid relative, *T. turgidum* ssp. *carthlicum*. It was also fixed in *T. urartu* and *T. timopheevii* (table 1). The accessions of *T. monococcum* were monomorphic for haplotype *Xpsr920-A^{m}1c*.

The *ABCT-1* Gene

Clones hybridizing with the PSR920 probe were isolated from *T. urartu*, *T. monococcum*, and durum cv. Langdon BAC libraries; their DNAs were digested with *Dra*I; and Southern blots were compared with those of *T. urartu* G1812, *T. monococcum* DV92, and durum cv. Langdon genomic DNAs. In 2 *T. urartu* and 3 *T. monococcum* BAC clones, the probe hybridized with the expected 9.5- and 14-kb *Dra*I restriction fragments, indicating that the clones harbored the *Xpsr920-A1b* and the *Xpsr920-A^{m}1* haplotypes, respectively. In 4 and 2 Langdon clones, the probe hybridized with 10.0- and 6.5-kb *Dra*I restriction fragments, indicating that the clones harbored the *Xpsr920-A1a* and *Xpsr920-B1* haplotypes, respectively.

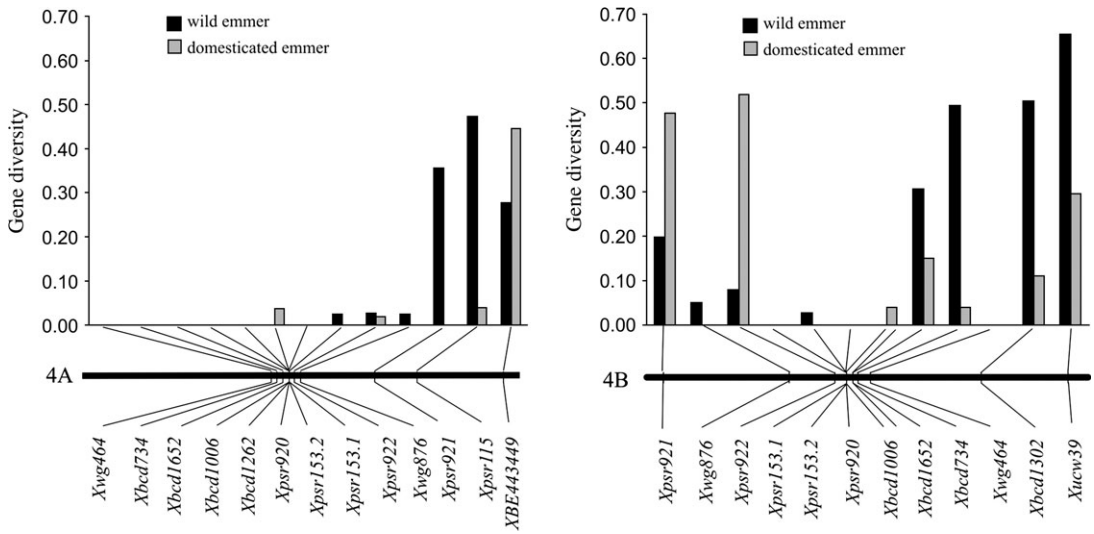

FIG. 2.—Genetic diversity at 13 loci along the 4A and 4B chromosomes in wild and domesticated emmer. The short arm is to the left and the long arm is to the right of each map. Note that several inversions and translocations involved chromosome 4A and the order of loci on that chromosome is not collinear with that in chromosome 4B.

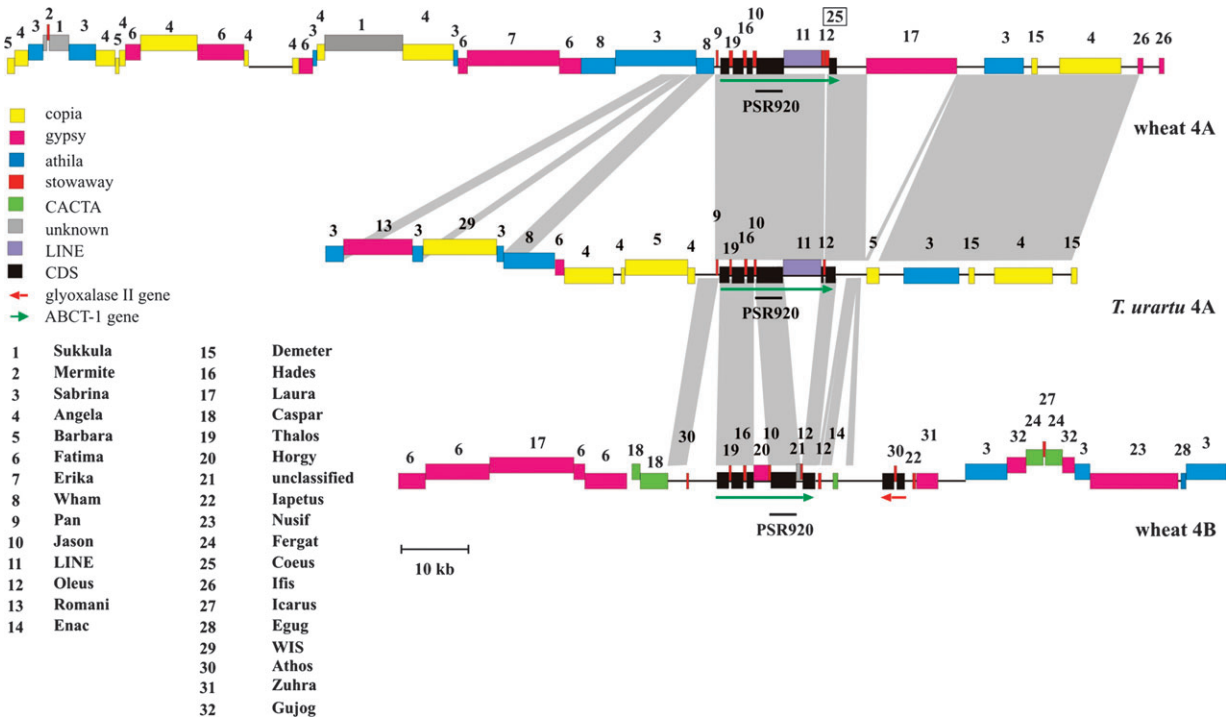

FIG. 3.—Structural comparison of *Triticum urartu* and durum wheat Langdon A and B genome genomic fragments cloned in *T. urartu* BAC 210J24 (middle), Langdon A genome BAC 643D12 (top), and B genome BAC 103H9 (bottom), respectively. Major groups of repetitive elements are shown as color-coded boxes. Numbers above the boxes correspond to different types of repetitive elements. More recent insertions of repetitive elements are shown as boxes raised above the elements into which they were inserted. The *ABCT-1* gene is indicated by a green arrow. Orthologous regions are joined with gray-shaded rectangles. Regions not showing any similarity to known coding or repetitive sequences are indicated as black lines. Element no. 25 (boxed) is the insertion of the Coeus MITE resulting in the *Xpsr920-A1a/b Dra*I RFLP.

Triticum urartu BAC clone 210J24, *T. monococcum* BAC clone 9M6, and durum wheat Langdon BAC clones 643D12 (A genome) and 103H9 (B genome) were sequenced. The *T. urartu* clone contained a 111 912-bp genomic fragment and Langdon clones 643D12 and 103H9 contained 175 848- and 91 644-bp genomic fragments, respectively. The *T. monococcum* fragment was not sequenced completely, and hence its precise size was unknown. Also sequenced was the 852-bp PSR920 *Pst*I clone. The PSR920 sequence was homologous with exons 17 and 18 of a predicted gene within each BAC clone (fig. 3). The PSR920 clone showed 99% similarity with the sequence of the 103H9 BAC clone and 97% similarity with the sequences of 9M6, 210J24, and 643D12 clones, indicating that PSR920 harbors a B genome gene DNA fragment.

Repeated sequences in each BAC clone were annotated using the TREP and GIRI repeated sequence databases (fig. 3), and the repeats were masked in the BAC sequences. The BAC sequences were compared with the NCBI EST database, which revealed the existence of a putative 2381-bp open reading frame in the *T. urartu* clone. The gene contained 20 introns and was 16 864-bp long in the *T. urartu* BAC clone and 17 147 and 14 522 bp in the Langdon A and B genome clones, respectively. The sequence of only 13 312 bp of the *T. monococcum* clone was determined, and only an 8030-bp contig of the *T. monococcum* gene was of a sufficient depth to be useful for comparison with the gene in the other 3 clones.

Triticum urartu exonic sequences were concatenated, and the predicted mRNA was compared with the NCBI nonredundant protein database. The gene showed similarity to a putative rice mitochondrial ATP binding cassette (ABC) transporter. The rice gene encoded a protein of 733 amino acids, and the *T. urartu* gene encoded a protein of 737 amino acids. We propose to designate the wheat gene as *ABCT-1*. The wheat 4A and 4B genes will be designated as *ABCT-A1* and *ABCT-B1*, respectively, and the *T. monococcum* gene as *ABCT-A^{m}1*.

The *T. urartu* and wheat A genome genes were compared along their entire length. They were identical at the first and second codon positions. The remaining comparisons, which included all 4 genes, were made for the sequence of only 8030 bp. Along this large portion of the gene, the *T. urartu* and wheat A genome genes diverged by 0.5% nucleotides at the third codon positions and 0.4% nucleotides in introns (table 3). The *T. monococcum* nucleotide sequence was more diverged from that of the *T. urartu* gene and the wheat A genome gene than those of the latter 2 genes were diverged from each other (table 3). A maximum parsimony tree was constructed from nucleotide sequences of the 4 genes. The wheat A genome gene clustered with the *T. urartu* gene, and both genes were separated from the *T. monococcum* gene with 94% bootstrap confidence (not shown). The 3 genes clustered together relative to the wheat B genome gene.

Intergenic Space

Nucleotide sequences of the *T. urartu* BAC clone and the 2 Langdon clones were annotated and compared

Table 3
Proportion of Nucleotides in Which *Triticum urartu* 4A, Durum Wheat 4A, *Triticum monococcum* 4A^m, and Durum Wheat 4B *ABCT-1* Genes Differ in 6701 bp of the Intron Nucleotide Sequence (the triangle above the diagonal), 886 bp of the First and Second Codon Positions, and 443 bp of the Third Codon Positions (upper and lower rows in the triangle below the diagonal, respectively)

	T. urartu 4A	*T. durum* 4A	*T. monococcum* 4A^m	*T. durum* 4B
T. urartu 4A		0.004 ± 0.001[a]	0.006 ± 0.001	0.023 ± 0.002
T. durum 4A	0.000		0.007 ± 0.001	0.025 ± 0.002
	0.005 ± 0.001			
T. monococcum 4A^m	0.001 ± 0.001	0.001 ± 0.001		0.025 ± 0.002
	0.007 ± 0.001	0.002 ± 0.001		
T. durum 4B	0.003 ± 0.001	0.003 ± 0.001	0.002 ± 0.001	
	0.018 ± 0.002	0.023 ± 0.002	0.025 ± 0.001	

[a] Standard error.

(fig. 3). In each clone, intergenic space mostly comprised retroelements and other classes of repeated elements frequently nested into each other. Most of the elements belonged to already described repetitive sequences in the TREP and GIRI databases (fig. 3). The end points of the *T. urartu* BAC clone were nested within the wheat A genome clone. Therefore, the entire 119 912-bp sequence of the *T. urartu* clone could be compared with an A genome orthologous sequence. The 2 clones shared a long intergenic region downstream of the *ABC1* gene, but the structural similarity of the intergenic region upstream of the gene was greatly reduced by insertions and deletions of transposable elements (fig. 3). The proportion of nucleotide substitutions (*p*-distance) was 0.016 in the downstream region between the 2 haplotypes. This rate was 4 times as great as the nucleotide substitution rate in introns, in which it was 0.004 (table 2).

Upstream of the gene, the *T. urartu* haplotype contained a number of elements that were absent from the wheat A genome haplotype. Two intact long terminal repeat (LTR) retroelements, Romani (no. 13) and WIS (no. 29), were flanked by 5-bp target site duplications. Divergence (*p*-distance) between the LTRs of Romani was 0.017 and between the LTRs of WIS was 0.013, both values being similar to or smaller than divergence (*p*-distance) of 0.016 between the *T. urartu* and wheat A genome haplotypes. The absence of the elements from the A genome haplotype, the absence of the 5-bp duplication, and the LTR divergence similar to or smaller than that between the haplotypes suggest that the 2 retroelements were inserted into the *T. urartu* haplotype after its divergence from the wheat haplotype.

Microsynteny between the *T. urartu* and wheat haplotypes was further perturbed by the insertion of LTR retroelement Laura (no. 17) located downstream of the gene. The *p*-distance between the LTRs of this retroelement was 0.005, which is only a third of the divergence between the haplotypes. Applying similar reasoning as described above led to the conclusion that Laura (no. 17) was inserted into the wheat A genome haplotype after its divergence from the *T. urartu* haplotype.

A single 1758-bp LTR of retroelement Barbara (no. 5) was found downstream of the gene in the *T. urartu* haplotype. The presence of a 5-bp target site duplication flanking the LTR indicates that this is a solo-LTR generated by ectopic recombination between the LTRs of the Barbara retroelement. Using the complete Barbara elements in the vicinity to estimate the length of the original element, the element suffered an estimated deletion of 7331 bp.

The *T. urartu* haplotype contains a large 28 798-bp region located upstream of the gene that is absent from the wheat A genome haplotype. Divergence (*p*-distance) between the LTRs of Barbara element (no. 5) located in this region was 0.052, which is about 4-fold greater divergence than $P = 0.016$ between the haplotypes, indicating that this element was inserted before the divergence of the *T. urartu* haplotype and the wheat A genome haplotype. This reasoning suggests that the *T. urartu* haplotype contains the ancestral sequence and that this sequence was deleted from the wheat A genome haplotype.

In the span of 103 048 bp of orthologous intergenic space (gene sequences were excluded) shared by the *T. urartu* clone with the wheat A genome clone, 36 806 bp were inserted and 36 129 bp were deleted in one or the other haplotype since their divergence. A total of 49.7% of the intergenic sequence was shared by the 2 haplotypes.

4A and 4B Genome Comparison

The intergenic space of the B genome BAC clone 103H9 did not show any correspondence with the intergenic space of the wheat A genome clone or the *T. urartu* clone (fig. 3). Furthermore, clone 103H9 differed from the wheat A genome and *T. urartu* clones by the presence of another gene in the vicinity of the *ABCT-1* gene (fig. 3). The gene showed 70% similarity to an *Arabidopsis thaliana* glyoxalase II (hydroxyacylglutathione hydrolase) at the amino acid level. It contained 8 exons and 7 introns and was 3965-bp long from the translation initiation codon to the stop codon in the last exon. The coding sequence was 774-bp long and encoded 258 amino acids. The reading frame was open, suggesting that the gene is functional. It was transcribed in the opposite direction compared with the *ABCT-1* gene. A putative transcription start was identified 166 bp upstream of the translation start by searching promoter a database at http://www.softberry.com/berry.phtml. The failure to find this gene in the A genome clones could either mean that the gene was deleted from the A genome or duplicated and transposed in the B genome. It is also possible that the gene exists in the vicinity of the A genome *ABCT-1* gene but is too far to be included with the *ABCT-1* gene on a single BAC clone. More work is needed to determine which of these possibilities is true.

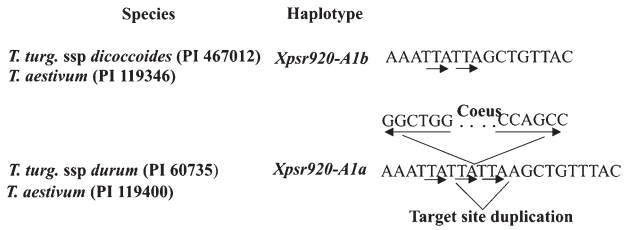

FIG. 4.—Nucleotide sequences surrounding the IVS-18 insertion site in the *ABC1-A1* gene in a sample of tetraploid wild emmer and durum wheat and hexaploid *Triticum aestivum* with the *Xpsr920-A1a* or *Xpsr920-A1b* haplotypes.

Equivalence of the *T. aestivum* and Wild Emmer *Xpsr920-A1b* Haplotypes

The only structural difference between the *T. urartu* gene and the durum wheat A genome gene was an insertion of 297 bp in intron 18 (henceforth IVS-18) in the durum wheat gene (element no. 25 in fig. 3). The insertion had the characteristics of a miniature inverted transposable element (MITE): it contained terminal inverted repeats 24-bp long and was flanked by a 3-bp target site direct duplication. The original sequence at the target site was a 3-nt tandem duplication TTATTA (fig. 4). The insertion duplicated the second TTA triplet of the original target sequence at the other side of the insertion (fig. 4). No homologous MITE was present in the TREP database or other redundant sequence databases. We name this MITE family "Coeus."

The nonredundant and EST sections of the NCBI database were searched for matches with the Coeus MITE. Ten independent matches were found in the wheat transcribed sequences having an e^{-10} *E* value or lower and at least one of the 24-bp MITE inverted repeats. The insertions were found in an H2A histone gene, a polyubiquitin (UBQ4) gene, and a gamma-type tonoplast intrinsic protein gene.

The Coeus MITE at the *Xpsr920-A1* was entirely contained within a restriction fragment delimited by *Dra*I sites, one in the PSR920 region of the gene and the other downstream of the last exon. The *Dra*I restriction fragment was 9609-bp long in the *Xpsr920-A1a* haplotype and 9312 bp in the *T. urartu* haplotype. The MITE was located near the far end of the *Dra*I restriction fragment relative to the PSR920 hybridization site, as predicted on the basis of restriction endonuclease digestions described earlier. The sizes of 9312- and 9609-bp restriction fragments predicted from nucleotide sequences were close to the 9.5 and 10.0 kb estimated from Southern blots for the *Xpsr920-A1b* and *Xpsr920-A1a* haplotypes, respectively. These facts suggested that the MITE is the causal agent of the RFLP.

To design PCR primers amplifying only the A genome IVS-18 sequence, including the MITE insertion site, primer pairs flanking the MITE were designed for each site at which the IVS-18 of the *ABCT-A1* gene differed from the orthologous site in the *ABCT-B1* gene. The A genome specificity of each primer pair was tested by comparing PCR amplification of Chinese Spring N4AT4B genomic DNA with that of Chinese Spring genomic DNA. A primer pair designated PSR920-UCD1 produced a 426-bp amplicon in Chinese Spring but no amplicon in N4AT4B,

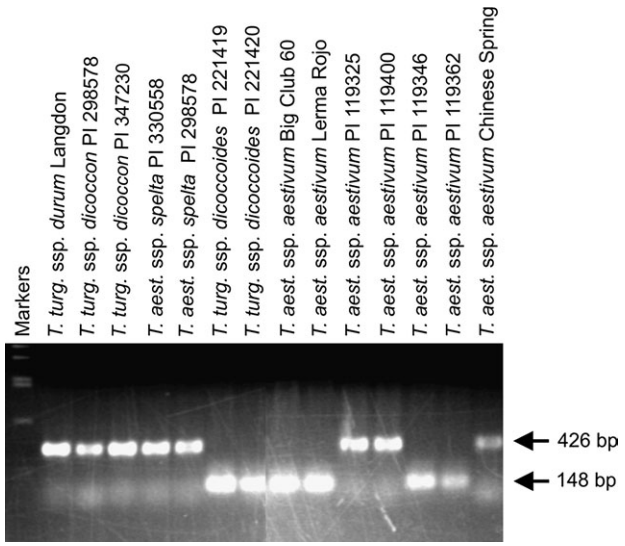

FIG. 5.—PCR amplification of tetraploid and hexaploid wheat genomic DNAs with the PSR920-UCD1 primers. Accessions that had shown an *Xpsr920-A1a* RFLP haplotype showed a 426-bp amplicon and those that had shown an *Xpsr920-A1b* RFLP haplotype showed a148-bp amplicon.

indicating that it specifically amplified the *ABCT-A1* sequence-tagged site (STS). The PSR920-UCD1 primer pair also amplified a product 426-bp long in durum wheat.

The PSR920-UCD1 primers were then used to amplify DNA from 14 tetraploid wheats representing all tetraploid taxa listed in table 1. All accessions were previously shown by RFLP to have the *Xpsr920-A1a* haplotype on the basis of Southern hybridization of the PSR920 clone. The lengths of the PCR amplicons were consistent with the lengths expected on the basis of the *Xpsr920-A1a* RFLP haplotype (fig. 5). Sequencing of the amplicons using the PSR920-1 primers as sequencing primers showed that all 14 accessions had the MITE insertion flanked by the TTATTA sequence on one side and the TTA sequence on the other side (fig. 4). Two accessions of wild emmer, previously shown to have the RFLP haplotype *Xpsr920-A1b*, did not have the MITE insertion and had only the TTATTA sequence as shown in figure 4. PSR920-UCD1 primers were used to amplify and sequence STSs from 32 *T. aestivum* accessions involving bread wheat, club wheat, and spelt, half representing the *Xpsr920-A1a* RFLP haplotype and half representing the *Xpsr920-A1b* RFLP haplotype. The 16 *T. aestivum* accessions with the *Xpsr920-A1a* haplotype had the same sequence as accession PI119346 in figure 4, and the 16 accessions with the *Xpsr920-A1b* RFLP haplotype had the same sequence as *T. aestivum* accession PI119400 in figure 4. On the basis of the equivalence of the *Xpsr920-A1b* sequences in wild emmer and *T. aestivum*, the hypothesis that the *T. aestivum Xpsr920-A1b* haplotype originated by a reverse mutation of the *T. aestivum Xpsr920-A1a* haplotype was rejected.

Discussion

Microsynteny between the A Genome of Wheat and the Genome of *T. urartu*

The *Xpsr920* locus is located in the low recombination region of chromosome 4 (Devos et al. 1995). Chromosome

regions with low levels of recombination show low levels of RFLP in the *Triticum–Aegilops* alliance (Dvorak, Luo, and Yang 1998). The absence of RFLP at this locus in *T. urartu*, *T. monococcum*, and wild emmer is consistent with this general pattern and suggests that representative haplotypes were sequenced.

The wheat A genome and *T. urartu ABCT-1* gene sequences were more closely related to each other than either was to the *T. monococcum* gene sequence. In a 100-kb region, microsynteny between the wheat A genome and the *T. urartu* genome was 50%. Only an average of 26% microsynteny was observed between wheat A genome BAC sequences and orthologous *T. monococcum* BAC sequences (Wicker et al. 2003). Although these observations report relationships in few genomic sites, they are nevertheless consistent with other evidence showing that the wheat A genome is more closely related to the *T. urartu* genome than to the *T. monococcum* genome (Nishikawa 1983; Dvorak et al. 1988; Tsunewaki et al. 1991; Dvorak et al. 1993; Huang et al. 2002).

Using an average intron divergence rate of 5.5×10^{-9} nt^{-1} $year^{-1}$ for genes in the *Triticum–Aegilops* alliance (Dvorak and Akhunov 2005), the 0.4% base pair divergence between the wheat and *T. urartu ABCT-1* introns translates into a divergence time of 0.7 MYA. In that time span, 3 retroelements and 1 MITE were inserted and 1 large deletion occurred in the *T. urartu* haplotype. One retroelement and one MITE were inserted and one 28-kb deletion occurred in the orthologous portion of the wheat haplotype. The rate of acquisition of large indels was comparable at the diploid and polyploid levels; the rate was 5.7 large indels 100 kb^{-1} Myr^{-1} at the diploid level and 4.3 large indels 100 kb^{-1} Myr^{-1} at the polyploid level. In view of this high rate of large indel acquisition by intergenic space, it is not surprising that no appreciable microsynteny was encountered in the intergenic space between the A and B genome haplotypes, which, judging from the number of intron nucleotide subsitutions, diverged 4 MYA.

Nucleotide Substitution Rate in the Intergenic Space

The nucleotide substitution rate in the intergenic space was 1.6×10^{-8} nt^{-1} $year^{-1}$, but in the introns of the juxtaposed *ABCT-1* gene, it was only 4×10^{-9} nt^{-1} $year^{-1}$ ($P < 0.001$, *t*-test). Intergenic DNA is generally more heavily methylated than DNA of genes. Because deamination of 5-methylcytosine results in CpG to TpG transitions, intergenic space is expected to show higher mutation rates than introns (SanMiguel et al. 1998; Wicker et al. 2003). Recent comparison of the human and chimpanzee genome nucleotide sequences also revealed higher nucleotide substitution rates in intergenic space than in genes, and the difference was attributed to CpG to TpG transitions due to the deamination of 5-methylcytosine (Chimpanzee Genome Sequencing Consortium 2005). The greater mutation rate and faster nucleotide sequence divergence rate in intergenic DNA versus introns and silent codon positions must be taken into account in reconstructions of the evolutionary history of intergenic space and the timing of individual events.

Wheat Evolution

Of 281 accessions of wild emmer, representing its entire geographic range across the Fertile Crescent, only 2 accessions, 1 from Lebanon and 1 from southwestern Syria, had the *ABCT-A1a* haplotype. The *ABCT-A1a* haplotype in these accessions may testify to the existence of natural *ABCT-A1a*/*b* polymorphism in wild emmer or introgression from domesticated wheat into wild emmer. The latter possibility is supported by an allocation test in which both accessions appeared as an admixture and by evidence for extensive gene flow between wild emmer and cultivated wheat in southwestern Syria, Lebanon, and Israel (Luo et al. 2006). The existence of recent hybridization involving wild emmer and domesticated wheat in the Upper Jordan Valley was hypothesized on the basis of seed storage protein profiles (Blumler 1998).

Of 333 accessions of domesticated tetraploid wheat, 330 had the *ABCT-A1a* haplotype; 2 domesticated emmer accessions from Turkey and 1 durum wheat accession from Iran had the *ABCT-A1b* haplotype. The domesticated accessions came from 2 geographically distant regions, the Gaziantep region in southern Turkey and the Tosya region in northern Turkey, and both appeared to be an admixture (Luo et al. 2006). Hybridization with *T. aestivum* is the most likely cause of the presence of the *ABCT-A1b* haplotype in the exceptional durum from Iran.

The failure to detect the MITE insertion in wild emmer in the Dyiarbakir region in Turkey, the putative site of tetraploid wheat domestication, suggests that the insertion was either a rare polymorphism in the wild emmer population that founded domesticated emmer or actually occurred during the domestication process and was fixed. Domestication was very likely accompanied by a population bottleneck aiding fixation of rare alleles, such as the *ABCT-A1a* haplotype. Several confounding factors make it difficult to determine if selection, which would manifest itself by low levels of polymorphism in the vicinity of a selected locus, has also played a role in fixation of the *ABCT-A1a* haplotype. The gene is located in the proximal, low recombination region. Such regions naturally show low levels of polymorphism in wheat and related species (Dvorak, Luo, and Yang 1998). Fixation of an inversion involving the *ABCT-A1* locus prior to emmer domestication (Devos et al. 1995) must have also caused loss of variation from the vicinity of the locus. Therefore, the absence of variation in most of the long arm of domesticated emmer chromosome 4A, although resembling a selection sweep, could have had other causes. The brittle rachis phenotype of Chinese Spring plants with added and substituted wild *T. monococcum* chromosome $4A^m$ suggests that a gene affecting spike rachis disarticulation, which is an important component of the domestication syndrome in cereals, is located on chromosome 4 in the A genomes (RS Kota and J Dvorak, unpublished). An inactive allele of this gene must have been fixed during emmer domestication to obtain emmer plants with nonbrittle rachis.

Hitchhiking of the MITE insertion with a gene important for the domestication syndrome, such as the nonbrittle rachis allele, would satisfactorily account for the ubiquity of the MITE insertion in domesticated tetraploid wheat.

However, if domesticated tetraploid wheat were the parent of *T. aestivum*, hexaploid wheat should then also be monomorphic for the *ABCT-A1a* haplotype. The nucleotide sequences of IVS-18 in all sequenced accessions of *T. aestivum* that had the *ABCT-A1b* haplotype were identical to the nucleotide sequence of IVS-18 of this gene in wild emmer. If the *T. aestivum ABCT-A1b* haplotype originated by a reversion of *ABCT-A1a* haplotype, not only would the excision of the MITE have to be precise but also the target site duplication that was generated by the MITE insertion would have to be removed. These events are unlikely. A more likely source of the *ABCT-A1b* haplotype in hexaploid wheat is gene migration from another population. The source could have been either wild emmer or, conceivably, *T. timopheevii* or *T. urartu* because they also possess the *ABCT-A1b* haplotype. However, because the locus is in the inverted segment of chromosome 4A (Devos et al. 1995), recombination of the *T. aestivum* chromosome 4A with either the *T. urartu* or *T. timopheevii* homologous chromosomes and the introgression of *ABCT-A1b* haplotype into *T. aestivum* is unlikely, as evidenced by low chiasmate pairing of wheat 4A in interspecific hybrids of *T. aestivum* with wild *T. timopheevii* (Dvorak and Appels 1982) and *T. urartu* (Chapman et al. 1976). Hence, the most likely source of the *T. aestivum ABCT-A1b* haplotype is wild emmer.

The equivalence of the *ABCT-A1b* haplotype in wild emmer and *T. aestivum* represents the first line of evidence showing that wild emmer may have participated in the evolution of *T. aestivum*. There are 2 scenarios by which wild emmer could have participated in the evolution of *T. aestivum*: 1) wild emmer could have been the tetraploid parent, and the *ABCT-A1a* haplotype could have been contributed to *T. aestivum* by introgression from domesticated emmer or 2) domesticated emmer could have been the tetraploid parent, as has been hypothesized (Jaaska 1978; Porceddu and Lafiandra 1986; Kimber and Sears 1987), and the *ABCT-A1b* haplotype could have been contributed to *T. aestivum* by introgression from wild emmer. Hypothesis 2 is more likely than hypothesis 1 for the following reasons. *Aegilops tauschii*, the source of the *T. aestivum* D genome, consists of 2 gene pools: T (for *tauschii*) and S (for *strangulata*) (Dvorak, Luo, Yang, and Zhang 1998). The principal source of the *T. aestivum* D genome was the S gene pool (Dvorak, Luo, Yang, and Zhang 1998). The S gene pool and wild emmer are not sympatric, which requires postulating a shift in the distribution of these taxa after emmer domestication to accommodate hypothesis 1. Genetic distances of the A and B genomes of free-threshing *T. aestivum* populations in North America, Europe, Turkey, Iran, and China to wild and domesticated emmer were estimated on the basis of RFLP at 118 loci (J Dvorak and MC Luo, unpublished). Each *T. aestivum* population was closer to domesticated emmer than to wild emmer. The opposite would be expected if hypothesis 1 were true.

If the parent of the original hexaploid wheat were domesticated emmer, then the original *T. aestivum* would have to have the *ABCT-A1a* haplotype. If hexaploid wheat originated in Transcaucasia or southwestern Caspian (Dvorak, Luo, Yang, and Zhang 1998), the diffusion of *T. aestivum* northward and eastward would have conserved the high frequency of the *ABCT-A1a* haplotype. However, the diffusion westward to eastern Turkey would have resulted in sympatry with wild emmer and potential for gene flow from wild emmer to *T. aestivum*. The cline of the *ABCT-A1b* haplotype in *T. aestivum* is consistent with this scenario. This scenario suggests that hexaploid wheat in eastern Asia is more similar to the original *T. aestivum* than wheat in Transcaucasia, Turkey, and Europe.

Spelt has traditionally been viewed as the ancestor of free-threshing hexaploid wheat (McFadden and Sears 1946). If spelt were an ancestral form of *T. aestivum*, it should be monomorphic for the *ABCT-A1a* haplotype. Instead, European spelt is nearly monomorphic for the *b* haplotype. Indigenous European free-threshing wheats also show a very high frequency of the *b* haplotype. The high frequency of the *ABCT-A1b* haplotype in European spelt contrasts with the low frequency of this haplotype in Asian spelt, indicating separate origin of Asian and European spelt, which is consistent with other lines of evidence (Tsunewaki 1968; Jaaska 1978; Dvorak and Luo 1999). The fact that both types of spelt have the *ABCT-A1a* haplotype frequencies similar to those of indigenous free-threshing wheats from the area of their origin argues for gene flow between hulled and free-threshing wheats in the same geographic area and undermines the ancestral position of spelt in the *T. aestivum* evolution scheme.

Acknowledgments

We thank Jorge Dubcovsky for providing us with durum wheat BAC clones and Zu-Li Yang and Carolyn Thomas for technical assistance. This work was supported by grant 99-35301-7905 from the USDA National Research Initiative, Genetic Mechanisms Program, and grant DBI-0321757 from the National Science Foundation, Plant Genome Research Program.

Literature Cited

Akhunov ED, Akhunova AR, Dvorak J. 2005. BAC libraries of *Triticum urartu*, *Aegilops speltoides* and *Ae. tauschii*, the diploid ancestors of polyploid wheat. Theor Appl Genet 111:1617–22.

Blumler MA. 1998. Introgression of durum in to wild emmer and the agricultural origin question. In: Damania AB, Valkoun J, Willcox G, Qualset CO, editors. The origins of agriculture and crop domestication. Aleppo, Syria: ICARDA. p 252–68.

Cenci A, Chantret N, Kong X, Gu Y, Anderson OD, Fahima T, Distelfeld A, Dubcovsky J. 2003. Construction and characterization of a half million clone BAC library of durum wheat (*Triticum turgidum* ssp. *durum*). Theor Appl Genet 107:931–9.

Chapman V, Miller TE, Riley R. 1976. Equivalence of the A genome of bread wheat and that of *Triticum urartu*. Genet Res 27:69–76.

Chimpanzee Genome Sequencing Consortium. 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 437:69–87.

Devos KM, Dubcovsky J, Dvorak J, Chinoy CN, Gale MD. 1995. Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. Theor Appl Genet 91:282–8.

Dvorak J, Akhunov ED. 2005. Tempos of deletions and duplications of gene loci in relation to recombination rate during diploid and polyploid evolution in the Aegilops-Triticum alliance. Genetics 171:323–32.

Dvorak J, Appels R. 1982. Chromosome and nucleotide sequence differentiation in genomes of polyploid *Triticum* species. Theor Appl Genet 63:349–60.

Dvorak J, Luo M-C. 1999. Evolution of free-threshing and hulled forms of *Triticum aestivum*: old problems and new tools. In: Caligari PDS, Brandham PE, editors. Wheat taxonomy: the legacy of John Percival. London: Academic Press. p 127–36.

Dvorak J, Luo M-C, Yang Z-L. 1998. Restriction fragment length polymorphism and divergence in the genomic regions of high and low recombination in self-fertilizing and cross-fertilizing Aegilops species. Genetics 148:423–34.

Dvorak J, Luo M-C, Yang Z-L, Zhang H-B. 1998. The structure of *Aegilops tauschii* genepool and the evolution of hexaploid wheat. Theor Appl Genet 97:657–70.

Dvorak J, McGuire PE, Cassidy B. 1988. Apparent sources of the A genomes of wheats inferred from the polymorphism in abundance and restriction fragment length of repeated nucleotide sequences. Genome 30:680–9.

Dvorak J, di Terlizzi P, Zhang HB, Resta P. 1993. The evolution of polyploid wheats: identification of the A genome donor species. Genome 36:21–31.

Dvorak J, Yang Z-L, You FM, Luo MC. 2004. Deletion polymorphism in wheat chromosome regions with contrasting recombination rates. Genetics 168:1665–75.

Gordon D, Abajian C, Green P. 1998. Consed: a graphical tool for sequence finishing. Genome Res 8:195–202.

Huang S, Sirikhachornkit A, Su X, Faris J, Gill BS, Haselkorn R, Gornicki P. 2002. Genes encoding plastid acetyl-CoA carboxylase and 3-phopshoglycerate kinase of the *Triticum*/*Aegilops* complex and the evolutionary history of polyploid wheat. Proc Natl Acad Sci USA 99:8133–8.

Jaaska V. 1978. NADP-dependent aromatic alcohol dehydrogenase in polyploid wheats and their relatives. On the origin and phylogeny of polyploid wheats. Theor Appl Genet 53:209–17.

Kihara H. 1944. Discovery of the DD-analyser, one of the ancestors of *Triticum vulgare* (Japanese). Agric Horticulture (Tokyo) 19:13–4.

Kimber G, Sears ER. 1987. Evolution in the genus *Triticum* and the origin of cultivated wheat. In: Heyne EG, editor. Wheat and wheat improvement. Madison, WI: American Soceity of Agronomy. p 154–64.

Lijavetzky D, Muzzi G, Wicker T, Keller B, Wing R, Dubcovsky J. 1999. Construction and characterization of a bacterial artificial chromosome (BAC) library for the A genome of wheat. Genome 42:1176–82.

Liu Y-G, Tsunewaki K. 1991. Restriction fragment length polymorphism (RFLP) analysis in wheat. II. Linkage analysis of the RFLP sites in common wheat. Jpn J Genet 66:617–33.

Luo MC, Yang ZL, Kawahara T, You F, Waines JG, Dvorak J. 2006. The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. Theor Appl Genet. (forthcoming).

Mac Key J. 1966. Species relationship in *Triticum*. In: Mac Key JM, editor. Second International Wheat Genetic Symposium. Lund, Sweden: Hereditas. p 237–75.

McFadden ES, Sears ER. 1946. The origin of *Triticum spelta* and its free-threshing hexaploid relatives. J Hered 37:81–9, 107–16.

Mori N, Ishi T, Ishido T et al. (11 co-authors). 2003. Origins of domesticated emmer and common wheat inferred from chloroplast DNA fingerprinting. In: Pogna NE, Romano M, Pogna EA, Galterio G, editors. Tenth International Wheat Genetics Symposium, Paestum, Italy. Rome, Italy: Instituto Sperimentale per la Cerealicoltura. p 25–8.

Nesbitt M, Samuel D. 1996. From staple crop to extinction? The archaeology and history of hulled wheats. In: Padulosi S, Hammer K, Heller J, editors. Hulled wheats. Promoting the conservation and use of underutilized and neglected crops. Proceedings of the first International Workshop on Hulled wheats. International Plant Genetic Resources Institute, Rome, Italy. Tuscany, Italy: Castelvecchio Pacoli. p 41–100.

Nishikawa K. 1983. Species relationship of wheat and its putative ancestors as viewed from isozyme variation. In: Sakamoto S, editor. Sixth International Wheat Genetics Symposium; Plant Germplasm Institute, Kyoto University. Kyoto, Japan. p 59–63.

Ozkan H, Brandolini A, Pozzi C, Effgen S, Wunder J, Salamini F. 2005. A reconsideration of the domestication geography of tetraploid wheat. Theor Appl Genet 110:1052–60.

Ozkan H, Brandolini A, Schafer-Pregl R, Salamini F. 2002. AFLP analysis of a collection of tetraploid wheat indicated the origin of emmer and hard wheat domestication in southeastern Turkey. Mol Biol Evol 19:1797–801.

Porceddu E, Lafiandra D. 1986. Origin and evolution of wheats. In: Barigozzi C, editor. Origin and domestication of cultivated plants. Amsterdam: Elsevier. p 143–78.

SanMiguel P, Gaut BS, Tikhonov A, Nakajima A, Bennetzen JL. 1998. The paleontology of intergene transposons of maize. Nat Genet 20:43–5.

Schiemann E. 1951. New results on the history of cultivated cereals. Heredity 5:312–4.

Stein N, Feuillet C, Wicker T, Schlagenhauf E, Keller B. 2000. Subgenome chromosome walking in wheat: a 450-kb physical contig in Triticum monococcum L. spans the Lr10 resistance locus in hexaploid wheat (Triticum aestivum L.). Proc Natl Acad Sci USA 97:13436–41.

Tsunewaki K. 1968. Origin and phylogenetic differentiation of common wheat revealed by comparative gene analysis. In: Finley KW, Shepherd KW, editors. Third International Wheat Genetic Symposium. Canberra, Australia: Australian Academy of Science. p 71–85.

Tsunewaki K, Takumi S, Mori N, Achiwa T, Liu YG. 1991. Origin of polyploid wheats revealed by RFLP analyses. In: Sasakuma T, Kinoshita T, editors. Nuclear and organellar genomes of wheat species. Yokohama, Japan: Kihara Memorial Foundation. p 31–9.

Wicker T, Yahiaoui N, Guyot R, Schlagenhauf E, Liu ZD, Dubcovsky J, Keller B. 2003. Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and A^m genomes of wheat. Plant Cell 15:1186–97.

Willcox G. 1997. Archaeobotanical evidence for the beginnings of agriculture in Southwest Asia. In: Damania AB, Valkoun J, Willcox G, Qualset CO, editors. The origins of agriculture and crop domestication. Aleppo, Syria: ICARDA. p 25–38.

Yan Y, Hsam SLK, Yu JZ, Jiang Y, Ohtsuka I, Zeller FJ. 2003. HMW and LMW glutenin alleles among putative tetraploid and hexaploid European wheat (*Triticum spelta* L.) progenitors. Theor Appl Genet 107:1321–30.

William Martin, Associate Editor

Accepted April 19, 2006