Theor Appl Genet (2006) 113:407–418
DOI 10.1007/s00122-006-0304-4

ORIGINAL PAPER

Transferable bread wheat EST-SSRs can be useful for phylogenetic studies among the Triticeae species

L. Y. Zhang · C. Ravel · M. Bernard · F. Balfourier · P. Leroy · C. Feuillet · P. Sourdille

Received: 8 February 2006 / Accepted: 28 April 2006 / Published online: 31 May 2006

Abstract The genetic similarity between 150 accessions, representing 14 diploid and polyploid species of the Triticeae tribe, was investigated following the UPGMA clustering method. Seventy-three common wheat EST-derived SSR markers (EST-SSRs) that were demonstrated to be transferable across several wheat-related species were used. When diploid species only are concerned, all the accessions bearing the same genome were clustered together without ambiguity while the separation between the different sub-species of tetraploid as well as hexaploid wheats was less clear. Dendrograms reconstructed based on data of 16 EST-SSRs mapped on the A genome confirmed that *Triticum aestivum* and *Triticum durum* had closer relationships with *Triticum urartu* than with *Triticum monococcum* and *Triticum boeoticum*, supporting the evidence that *T. urartu* is the A-genome ancestor of polyploid wheats. Similarly, another tree reconstructed based on data of ten EST-SSRs mapped on the B genome showed that *Aegilops speltoides* had the closest relationship with *T. aestivum* and *T. durum*, suggesting that it was the main contributor of the B genome of polyploid wheats. All these results were expected and demonstrate thus that EST-SSR markers are powerful enough for phylogenetic analysis among the Triticeae tribe.

Communicated by M. Sorrells

Electronic Supplementary Material Supplementary material is available to authorised users in the online version of this article at http://dx.doi.org/10.1007/s00122-006-0304-4.

L. Y. Zhang · C. Ravel · M. Bernard (✉) · F. Balfourier · P. Leroy · C. Feuillet · P. Sourdille
UMR INRA-UBP Amélioration et Santé des Plantes, Domaine de Crouël, 234 Avenue du Brézet, 63100 Clermont-Ferrand, France
e-mail: michel.bernard@clermont.inra.fr

Introduction

The grass family (Poaceae or Gramineae) comprises the most important cultivated crops in the world. It includes all the cereals as well as forage crops, sugar cane, or bamboos. There are more than 10,000 grass species distributed among 651 genera which originated and diverged during about 60 million years. They are grouped in five sub-families among which the Bambusoideae, the Panicoideae, and the Pooideae which include the cultivated species such as wheat, barley, rice, and maize. Wheat species and their close relatives belong to the Pooideae sub-family and to the Triticeae tribe which comprises the genera *Triticum* (wheat), *Aegilops*, *Secale* (rye), and *Hordeum* (barley). Within this latter tribe, diploid as well as polyploid species can be found. In the Triticeae, the basic chromosome number is seven and the genomes are designated by letters. These genomes are not perfectly identical but show high levels of similarity and are thus named "homoeologous." Wheat polyploid species originated from hybridizations between either diploid species or diploid and polyploid species followed by natural genome doubling. For example, the hexaploid bread wheat (*Triticum aestivum* L.em.Thell. $2n = 6x = 42$) is an allopolyploid species (AABBDD) originated by hybridization of a AA-genome diploid species related to *Triticum urartu* ($2n = 2x = 14$, Dvorak et al. 1993; Huang et al. 2002) with a BB-genome diploid species from the *Sitopsis* section giving rise to *Triticum dic-*

occoides species (AABB), followed by an additional hybridization with a DD-genome diploid species, *Aegilops tauschii* ($2n = 2x = 14$, McFadden and Sears 1946). The origin of the B genome remains controversial but *Aegilops speltoides* seems to be the most likely living relative of an extinct or yet to be discovered B-genome donor species (Sarkar and Stebbins 1956; Riley et al. 1958; Rees and Walters 1965; Natarajan and Sharma 1974; Chen et al. 1975; Jaaska 1980; Hassan and Gustafson 1996; Maestra and Naranjo 1998).

Grasses taxonomy and phylogeny were initially based on morphological and physiological traits. However, discrimination between species was sometimes limited because of a lack of relevant descriptors, or divergence was overestimated because of the difficulty in scoring some characters. Morphological traits have thus been supplanted first by isozymes which have provided valuable insights into the phylogeny among the genera and species (McIntyre 1988) and later by DNA-based molecular markers. The results obtained with these tools were relatively consistent with the general taxonomic information provided by earlier morphological, physiological, and isozyme markers (Monte et al. 1993). Especially, SSRs were extensively explored in wheat evolutionary studies because of their high polymorphism level, their co-dominant inheritance, and their reproducibility (Plaschke et al. 1995; Donini et al. 2000; Prasad et al. 2000; Manifesto et al. 2001; Ben Amer et al. 2001; Leisova and Ovesna 2001; Röder et al. 2002; Zhang et al. 2002; Roussel et al. 2005). However, interspecific phylogenetic studies in wheat using SSRs were limited to exploration of the relationships between *Ae. tauschii* and the D genome of bread wheat (Lelley et al. 2000) because of the limited transferability of genomic SSRs to related species (Sourdille et al. 2001).

Recently, Zhang et al. (2005) demonstrated that bread wheat EST-derived SSR markers (EST-SSRs) developed from EST sequences showed a high level of transferability to close and wild relatives of wheat because they are derived from conserved coding regions. In addition, they are still a source of information for assessing genetic relationships (Eujayl et al. 2001, 2002) as suggested by Bandopadhyay et al. (2004), who studied DNA polymorphism among 18 species of *Triticum–Aegilops* complex and were able to construct a dendrogram separating the diploid and tetraploid species.

In this paper, we studied the genetic similarity among 150 accessions from 14 different Triticeae species representing 16 different genomes using a set of 73 EST-SSRs that we recently developed (Zhang et al. 2005) with the final aim as the evaluation of the potential of wheat EST-SSRs for phylogenetic studies among different diploid, tetraploid, and hexaploid sub-species of the Triticeae tribe.

Materials and methods

Plant material

A total of 150 accessions representing 14 species and 16 genomes of the Triticeae tribe were used (see Supplementary Table 1, seeds available on request). This included diploid and polyploid species and autogamous as well as allogamous species. Similar numbers of accessions (between two and five) for each species were randomly chosen among our collection, except for *T. aestivum* and *T. durum*, where 23 varieties were selected for each species. Seeds were obtained from the Centre of Biological Resources on Cereal Crops (INRA-Clermont-Ferrand, France) and from Jacques David (INRA Montpellier, France, tetraploid and *Aegilops* species). For each species, between five and ten seeds from self-pollinated ears (when available or possible) were sown for further DNA extraction.

DNA extraction, PCR amplification and SSR detection

DNA was extracted from fresh leaves ground in liquid nitrogen using a CTAB protocol as previously described (Tixier et al. 1998). A set of 73 EST-SSRs (Zhang et al. 2005; http://www.wheat.pw.usda.gov/ITMI/EST-SSR/) was selected (Supplementary Table 2) according to their ability to be transferable to the studied species. PCR reactions using the M13 protocol were carried out as described in Nicot et al. (2004) with an annealing temperature of 60°C for 30 cycles (30 s 94°C, 30 s 60°C, 30 s 72°C) and 56°C for 8 cycles. Amplification products were visualized using an ABI PRISM®3100 Genetic Analyzer (Applied Biosystems). Finally, fragment sizes were calculated using GENESCAN and GENOTYPER softwares (Applied Biosystems), where different alleles are represented by different amplification sizes for tandem repeats. Two alleles are considered as identical when they show the same fragment size. To discriminate between PCR failure and null allele, PCR reactions were done twice.

Estimation of genetic similarity

Because only a limited number of accessions was tested for each species (between two and five), it was not possible to accurately measure the allelic frequency for each SSR and to use statistical analyses such as AMOVA or *F* statistics. For phylogenetic studies, we

decided thus to generate a binary matrix as followed: presence of an amplified product of a given size was scored as "1" while the absence of the same amplification product was scored as "0." The binary data were used to compute the distance matrix as 1—the Jaccard's similarity coefficient (Jaccard 1908). Phylip software was used to identify the genetic similarity (Felsenstein 1993). As all the species studied belong to the Triticeae tribe, the assumption of a molecular clock was acceptable. Therefore, the dendrograms were obtained by the UPGMA clustering method. The reliability and goodness of fit of dendrograms obtained from EST-SSRs data were tested through bootstrapping based on 100 samples (Felsenstein 1985). This led to 100 dendrograms summarized in a consensus tree which indicated the proportion of bootstrapped trees showing that same clade.

Results

Transferability of the EST-SSRs to the different species was high and ranged from 100% for *T. aestivum* sub-species *compactum* and *petropavlovsky*, to 62.3% for barley (Zhang LY et al., unpublished results). Since transferability was not complete, null alleles were not considered and were quoted as missing data since there was a higher probability that the lack of amplification was due to the presence of numerous mutations in the flanking sequences, which are obviously different among the species rather than to a deletion of the genes which could have been considered as similar events. This also justifies the choice of the Jaccard distance index (Jaccard 1908) which does not consider as informative a shared absence of a given trait (here an amplification product). However, we only retained markers for which the percentages of missing data were low and never exceeded the maximum of one missing data/locus for each species. Thus, estimation of genetic distances was quite accurate.

Clustering of the diploid grass species

We selected the 46 accessions of diploid species (Supplementary Table 1), including the A-, B- (or S), and D-genome donors of hexaploid wheat, *Aegilops umbellulata* (UU), *Aegilops comosa* (MM), barley (HH), and rye (RR). The clustering of the lines was based on 1,170 informative fragments produced from 73 EST-SSRs. The Jaccard genetic distance coefficients (Jaccard 1908) ranged from 0.307 between *Triticum monococcum* accessions 68191 and 68212 to 0.978 between *Aegilops tauschii* accession 15 and rye accession SCW 3. The consensus tree obtained (Fig. 1) showed that all the accessions bearing the same genome were clustered together without ambiguity, the bootstrapped values for each group being generally higher than 85%. The A-genome sub-species were divided into two groups, one including *T. urartu* sub-species only, the other where *T. monococcum* and *T. boeoticum* sub-species were clustered together. This latter group presented two clusters, one mainly corresponding to *T. monococcum* sub-species, the other one to *T. boeoticum* sub-species. Within the A-genome sub-species, the *T. boeoticum* accessions 68182 and 68184 and the *T. urartu* accession 78096 were not clustered with the other accessions of the same sub-species. *T. boeoticum* accession 68182 was associated with *T. monococcum* sub-species while *T. boeoticum* accession 68184 was with *T. urartu* sub-species. On the contrary, the *T. urartu* accession 78096 was clustered with *T. boeoticum* sub-species. However, the bootstrap values of the nodes leading to these groups were, respectively, 23 and 26% indicating that these groups were not highly reliable.

Similarly, the B-genome sub-species were divided into two groups, one where *Aegilops searsii*, *Aegilops bicornis*, and *Aegilops longissima* were clustered together, while *Ae. speltoides* was apart. Within the B-genome sub-species, the *Ae. bicornis* accession 471323 (AEBIC-47) was associated with two of the *Ae. longissima* accessions while *Ae. longissima* accession 330486 (AELON-330) was associated with the *Ae. bicornis* accession 70 (AEBIC-70). Concerning the relationships among the three diploid ancestors of hexaploid wheat, the *Ae. tauschii* species was more closely related to the B-genome species than to the A-genome species. However, this grouping was found in less than 60% of the trees indicating that relative position of these species was not accurate. *Ae. umbellulata* (UU) and *Ae. comosa* (MM) were clustered together and formed a group which was associated with the one including B- and D-genome sub-species in 87% of the dendrograms. Barley was found to be closer to all these wild species compared to rye but in less than 60% of the trees. This indicated that the relative positions of barley, rye, and wheat related diploid wild species need to be clarified.

From this tree, we can conclude that wheat EST-SSRs are powerful enough to assess the genetic variability of wheat diploid relatives. Also, clustering of some accessions should be reconsidered according to the results we report.

Clustering of the tetraploid grass species

Similarly, we used the 48 accessions of tetraploid grass species (Supplementary Table 1) to construct a

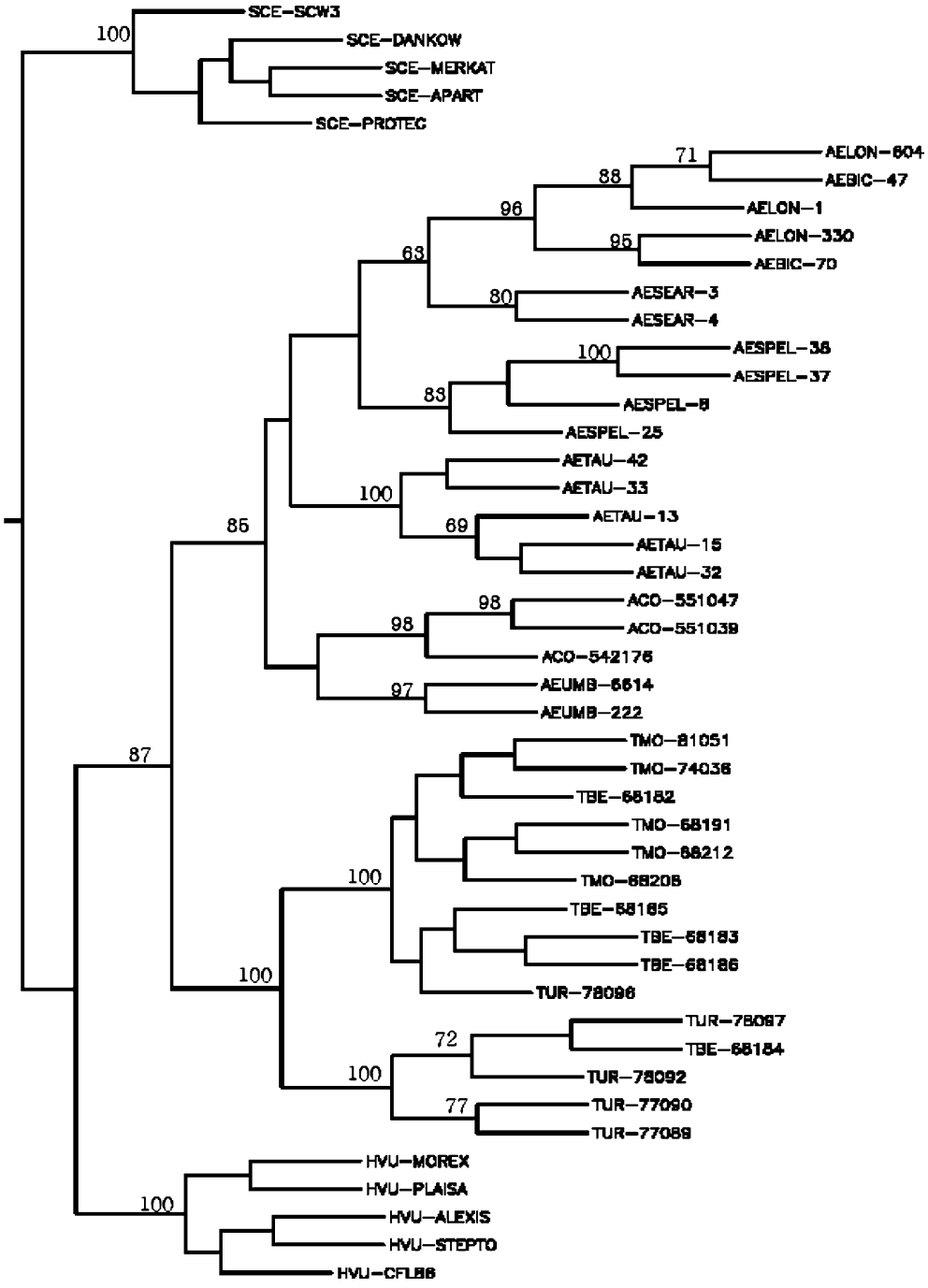

Fig. 1 Consensus tree of 46 accessions of diploid species reconstructed from 100 UPGMA trees obtained from data resampled in a set of 73 EST-SSRs. The accession codes are the same as those indicated in Supplementary Table 1 (SCE rye, AELON *Ae. longissima*, AEBIC *Ae. bicornis*, AESEAR *Ae. searsii*, AESPEL *Ae. speltoides*, AETAU *Ae. tauschii*, ACO *Ae. comosa*, AEUMB *Ae. umbellulata*, TMO *T. monococcum*, TBE *T. boeoticum*, TUR *T. urartu*, HVU barley). The branch lengths are proportional to the number of times that each group appeared. Additionally, *numbers* indicated bootstrap values larger than 60%

dendrogram. The clustering of the lines was elaborated based on 167 informative fragments produced from 73 EST-SSRs. The Jaccard genetic distance coefficients (Jaccard 1908) ranged from 0.322 between *Triticum turgidum* accessions 7786 and 387456 to 0.674 between *Triticum carthlicum* accession 94753 and *T. dicoccoides* accession 467014. The consensus dendrogram obtained from bootstrapped data is presented in Fig. 2. Separation between the different sub-species was less clear compared to the diploid species. However, all the

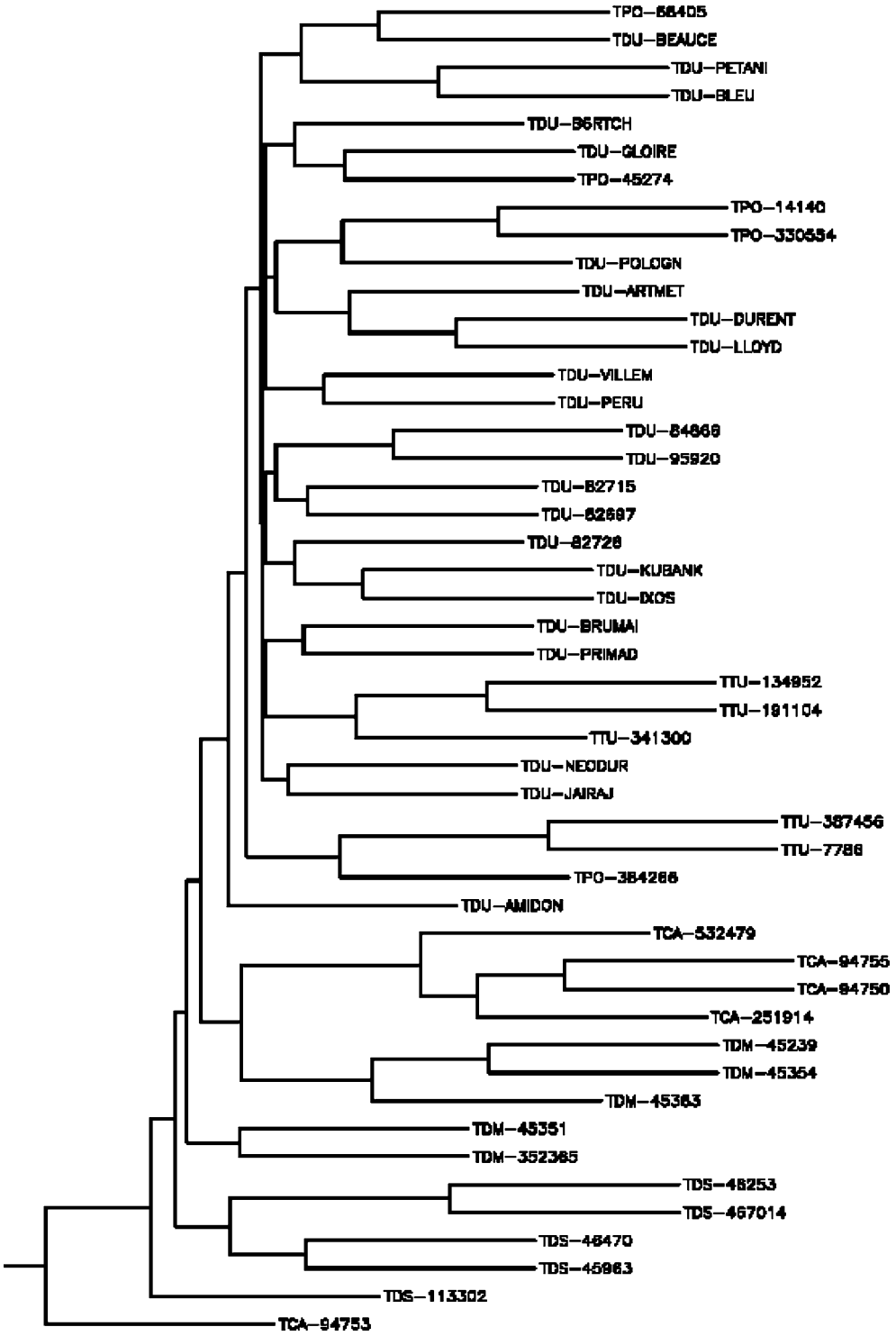

Fig. 2 Consensus tree of 48 accessions of tetraploid species reconstructed from 100 UPGMA trees obtained from data resampled in a set of 73 EST-SSRs. The accession codes are the same as those indicated in Supplementary Table 1 (TPO *T. polonicum*, TDU *T. durum*, TTU *T. turgidum*, TCA *T. carthlicum*, TDM *T. dicoccum*, TDS *T. dicoccoides*). The branch lengths are proportional to the number of times that each group appeared. Additionally, *numbers* indicated bootstrap values larger than 60%

T. durum accessions were grouped except the French variety "Amidonnier Blanc Barbu." Most of the *polonicum* (4/5) and the *turgidum* (3/5) sub-species were associated with the same group. The *carthlicum* sub-species were associated with the *dicoccum* sub-species while the *dicoccoides* sub-species formed a separate

cluster. The *carthlicum* accession 94753 was completely isolated.

From this tree, we can conclude that the relationships among all the wheat tetraploid species are closer compared to what was observed between the diploid species. This suggests that tetraploid species are more closely related than the diploid species indicating that they diverged more recently.

Clustering of the hexaploid grass species

Fifty accessions of hexaploid wheat species (Supplementary Table 1) were used to construct a dendrogram from 207 informative fragments produced from 73 EST-SSRs. The Jaccard genetic distance coefficients (Jaccard 1908) ranged from 0.231 between *Triticum aestivum* ssp. *macha* Landrace 4 and *T. ae.* ssp. *macha* accession 102V to 0.646 between *T. aestivum* cv Aurore and *T. ae.* ssp. *compactum* accession Lo To Mai. The consensus dendrogram obtained from bootstrapped data is presented in Fig. 3. Two groups were obtained. The first one was obtained in 48% of the bootstrapped trees and was made of most of the *T. aestivum* cultivars (15/22). On the contrary, the second one included all the others plus seven *T. aestivum* cultivars. Within this group, only *spelta* and *macha* sub-species were clustered together but UPGMA trees were found in less than 60% of the cases indicating that the clustering was not sure. The European accessions trended to cluster while the Asian ones were closer to each other.

Our results suggest that *T. aestivum* varieties seem to be split into two main groups: one includes most of the European and Asian (Korea, China, Afghanistan, Pakistan) varieties which is closer to other *T. aestivum* sub-species (*macha*, *spelta*...), the other includes the remaining varieties from other countries. This may indicate a common ancestral origin for each of the two groups.

Relationships between the A and B genomes for bread wheat and related diploid species bearing homoeologous genomes

Two sets of 16 and 10 markers evenly located on the A and B genomes of hexaploid wheat, respectively (Zhang et al. 2005), were selected to investigate the relationships among these two genomes of hexaploid, tetraploid, and diploid wheat species.

For the A genome, 244 informative fragments obtained from 16 EST-SSRs assigned to the A genome of bread wheat were used. The Jaccard genetic distance coefficients (Jaccard 1908) were calculated using these data and ranged from 0.052 between *T. durum* cv Ixos and Kubanka to 0.973 between *T. monococcum* accession 74036 and *T. aestivum* cv Seuseun 27. The consensus tree obtained from bootstrapped data is shown in Fig. 4. Tetraploid species (AB) were clustered together and were closely related to hexaploid species (ABD). Surprisingly, the French varieties Apache and Ornicar were closer to *durum* species than to *aestivum* species. Within the hexaploid species, the same separation into two groups as previously described (see previous section) was observed but the bootstrap values were not highly significant. On the contrary, the diploid species bearing the A genome were divided into two groups, one including all the accessions of *T. urartu*, the other grouping the sub-species *T. monococcum* and *T. boeoticum*. Within this latter group, *monococcum* and *boeoticum* sub-species were separated except *T. monococcum* accession 68191 which was linked to the *boeoticum* sub-group. *T. urartu* showed a closer relationship with *T. aestivum* and *T. durum* compared to *T. monococcum* and *T. boeoticum* confirming that *T. urartu* is the most probable ancestor of the A genome of polyploid wheats.

Similarly, 117 informative fragments obtained from 10 EST-SSRs assigned to the B genome of bread wheat were used to investigate the relationships between the species bearing B or related to B genomes. In this case, Jaccard genetic distance coefficients (Jaccard 1908) ranged from 0.074 between *T. aestivum* cv Chortandinka and Coppadra to 0.952 between *Ae. longissima* accession 1 and *T. aestivum* cv Aifeng. The data were bootstrapped to obtain a consensus tree from 100 UPGMA trees (Fig. 5). Similar to the previous study, tetraploid species were clustered together and were closely related to hexaploid species. *Ae. speltoides* accessions were grouped together while other sub-species were more dispersed. In this tree, *Ae. searsii* accession 4 was clustered with the polyploid wheats but in less than 60% of the UPGMA trees indicating that this linkage was not highly significant. However, this was in accordance with the results from Feldman (1978) who proposed *Ae. searsii* as a potential candidate for the B-genome donor. *Ae. speltoides* sub-species had also close relationships with the polyploid wheats which may suggest that both *Ae. searsii* and *Ae. speltoides* may have contributed to the elaboration of the B genome of polyploid wheats. A larger number of *Ae. searsii* accessions should be tested to confirm or reject this hypothesis.

In both cases, tetraploid and hexaploid wheats were found to be closely related but with no clear and significant splitting between them. This is probably due to the fact that they diverged only recently about 250,000 years ago.

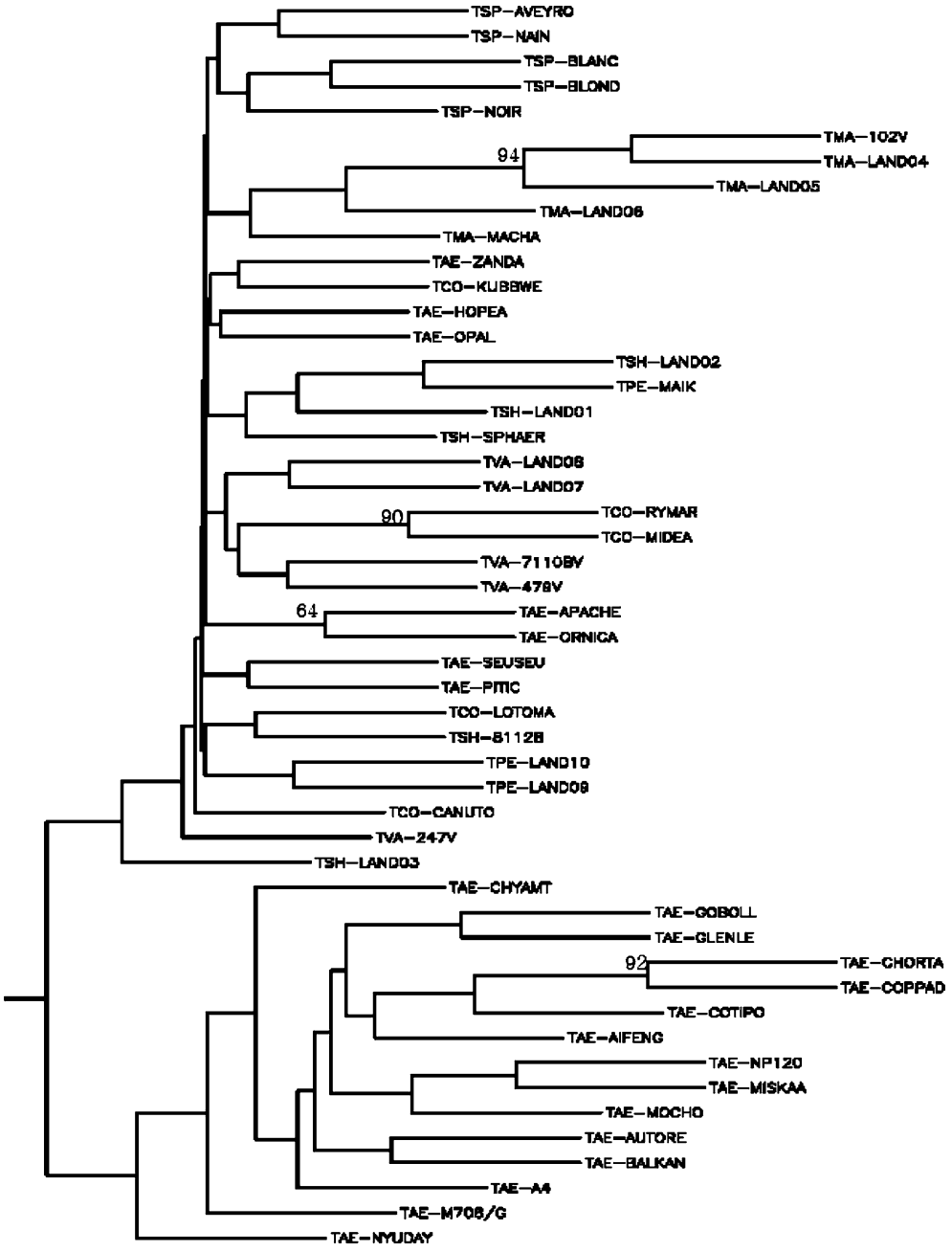

Fig. 3 Consensus tree of 50 accessions of hexaploid species reconstructed from 100 UPGMA trees obtained from data resampled in a set of 73 EST-SSRs. The accession codes are the same as those indicated in Supplementary Table 1 (TSP *T. spelta*, TMA *T. macha*, TAE *T. aestivum*, TCO *T. compactum*, TSH *T. sphaerococcum*, TVA *T. vavilovi*, TPE *T. petropavlovskyi*). The branch lengths are proportional to the number of times that each group appeared. Additionally, *numbers* indicated bootstrap values larger than 60%

Discussion

Bread wheat EST-SSRs were tested to explore their potential in phylogeny analyses of a large set of Triticeae species (14). We were thus able to compare the species according to their ploidy level (diploid species as well as tetra- and hexaploid species). We also compared the relationships between the diploid and the polyploid species according to the genome (A or B) using previously assigned EST-SSRs (Zhang et al. 2005).

Many methods exist for phylogenetic studies but there is a growing realization that the data on SSRs should be recorded as specific alleles in different genotypes and then used for analysis (Reif et al. 2005). This makes co-dominant SSRs better markers than other

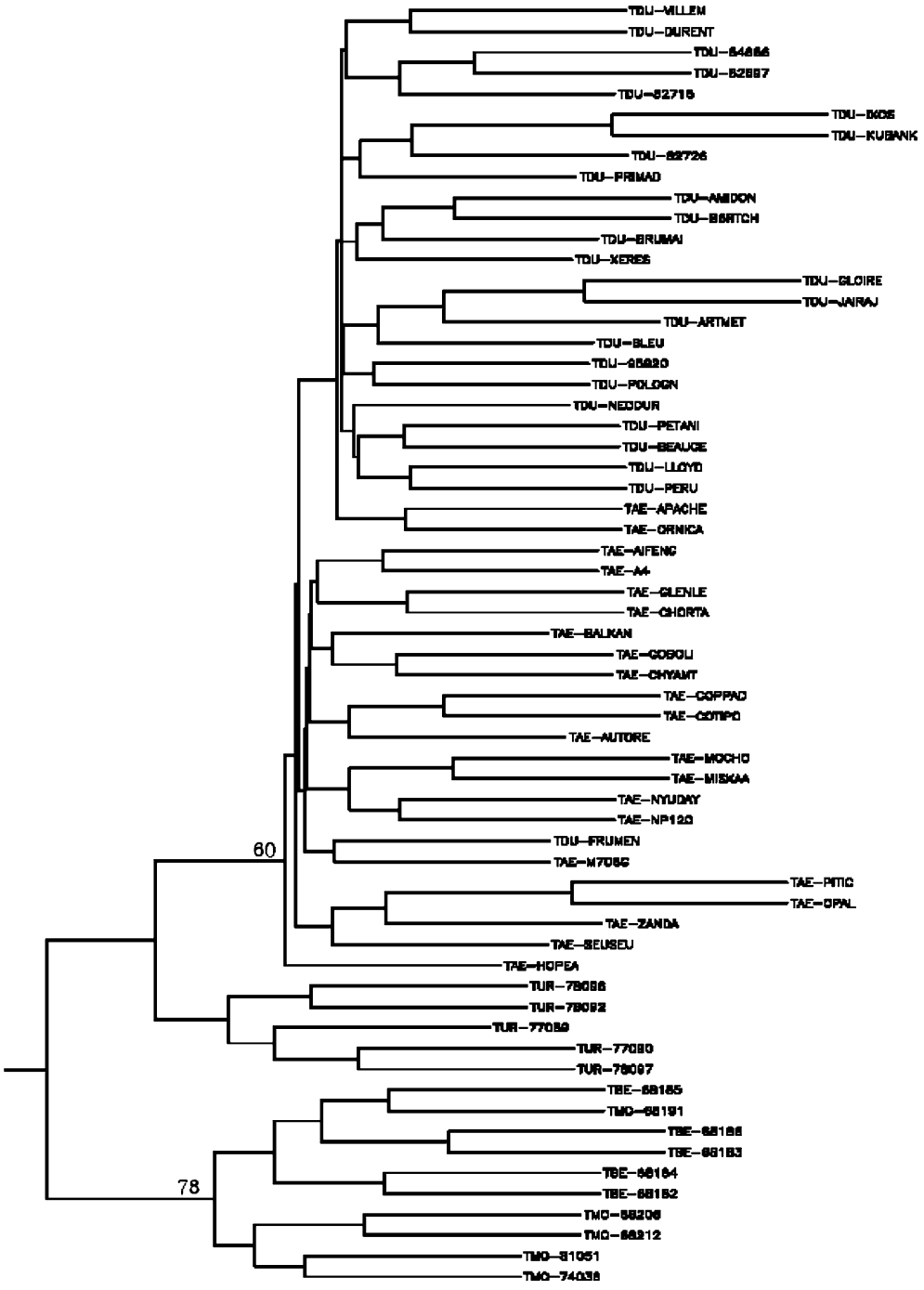

Fig. 4 Consensus tree of 62 Triticeae accessions (22 *T. aestivum*, 25 *T. durum*, 5 *T. monococcum*, 5 *T. urartu*, and 5 *T. boeoticum*) reconstructed from 100 UPGMA trees obtained from data resampled in a set of 16 EST-SSRs assigned to the A genome. The accession codes are the same as those indicated in Supplementary Table 1. The branch lengths are proportional to the number of times that each group appeared. Additionally, *numbers* indicated bootstrap values larger than 60%

dominant markers because if data are recorded as allelic variants, more powerful statistical analyses such as AMOVA or *F* statistics (Slatkin 1995) can be conducted. However, in this study, we analyzed a limited number of accessions for each species and because these species gave different number of alleles compared to wheat, the frequencies could not be estimated accurately. Moreover, some species were allogamous (*Lolium*, *Ae. speltoides*) and it was assumed that the diversity available within an accession was compensated by extracting DNA from a mix of several plants (see "Materials and methods"). It was thus inappropri-

Fig. 5 Consensus tree of 58 Triticeae accessions (22 *T. aestivum*, 25 *T. durum*, 4 *Ae. speltoides*, 2 *Ae. searsii*, 2 *Ae. Bicornis*, and 3 *Ae. longissima*) reconstructed from 100 UPGMA trees obtained from data resampled in a set of ten EST-SSRs assigned to the B genome. The accession codes are the same as those indicated in Supplementary Table 1. The branch lengths are proportional to the number of times that each group appeared. Additionally, *numbers* indicated bootstrap values larger than 60%

ate to use AMOVA and *F* statistics in our case which led us to score the data on SSRs in a binary format (presence/absence of alleles). In this study, we used a phenetic method based on the Jaccard's distance matrix (Jaccard 1908). As the divergence between all the species studied was quite recent, the molecular

clock was assumed which led us to choose UPGMA clustering method rather than the NJ method (Saitou and Nei 1987). Under this assumption, dendrograms based on UPGMA can be discussed from a phylogenetic point of view. Similarly, the inadequacy of one marker system over a number of marker systems and morphological traits was also demonstrated (Gupta and Varshney 1999). In order to solve both problems (type of markers and analyses), it would thus be interesting to test a larger number of accessions from each species together with a larger number of EST-SSRs and compare the results obtained with other types of markers (gSSRs, RFLPs, AFLPs, SNPs) in order to know which type and how many markers and samples give the best results for phylogenetic studies among members of the Triticeae tribe.

When genetic similarity among diploid relatives of hexaploid wheat was investigated, most of the accessions of each species were clustered into their corresponding group in our dendrogram. The three groups of ancestral diploid species were clustered which was consistent with other taxonomic studies of the Triticeae tribe (Appels et al. 1989). The close associations observed between the B-genome and the D-genome donors were in agreement with previous studies using RFLP markers (Monte et al. 1993). The same authors showed that the ancestral diploid species were more closely related to *Hordeum* species compared with *Secale* species, which was also consistent with what we found. Regarding the B-genome possible donors, we tested four species which were divided into two distinct groups in our dendrogram. This was in accordance with previous studies which mentioned similar classifications (Ogihara and Tsunewaki 1988; Sasanuma et al. 1996, 2004), and this is also in agreement with the classification of the Sitopsis section into two subsections, one for *Ae. speltoides* alone and the other including *Ae. longissima*, *Ae. searsii*, *Ae. sharonensis* and *Ae. bicornis* (von Eig 1929). Concerning the discrepancies that we reported in our results, this suggests that we should reconsider the classification of some of our accessions.

The relationships between the tetraploid species were more confused. *T. durum*, *T. polonicum*, and *T. turgidum* sub-species were clustered together. This can be explained by the fact that these are three cultivated species or because not enough markers have been tested to clearly separate these species that diverged only recently. In addition, *polonicum* accessions were more or less dispersed among the *durum* accessions suggesting that they could have been involved in the pedigree of some current varieties. On the contrary, *carthlicum*, *dicoccum*, and *dicoccoides* sub-species were less related to cultivated accessions. The two former species were closer to all these tetraploid species while *dicoccoides* was found to be the less related, which confirms that this species could be the ancestor of all the tetraploid wheat species. The divergence could have occurred from *T. dicoccoides* to *dicoccum*, *carthlicum*, *turgidum*, and *durum*. This was consistent with the fact that the wild species *T. turgidum* ssp. *dicoccoides* was domesticated to form *T. t.* ssp. *dicoccum* and successive domestication steps generated durum wheat (*T. t.* ssp. *durum*), the most cultivated tetraploid wheat (Salamini et al. 2002). The origin and classification of *T. carthlicum* accession 94753, which undoubtedly roots the tree, should be confirmed prior to concluding on its position.

The hexaploid species were separated into two groups, one gathering most of the *aestivum* varieties, the other including the remaining sub-species and varieties. This latter group mainly contained European and Asian accessions indicating a possible common origin. In addition, the *aestivum* varieties from this group were dispersed throughout the cluster and were related to *compactum* as well as to *vavilovi* or *macha* accessions, suggesting that these species are either closely related or that they have been used as progenitors of wheat varieties in a course of European breeding programs.

In both tetraploid and hexaploid species, the difficulty in separating accurately the groups of sub-species can be explained by the high percentage of common bands between all the species, which probably reflect the high conservation of the coding sequences between all the Triticeae species. This problem would probably be solved by testing a larger number of either EST-SSRs or genomic SSRs. In this latter case, more diversity is supposed to be observed because most of genomic SSR markers detect polymorphism located in the non-coding regions of the genome which are supposed to evolve more rapidly (Brown et al. 2001). On the contrary, EST-SSRs derive from the expressed part of the genome and detect thus the genetic diversity appearing within the genes themselves. Using this strategy, Xu et al. (2004) reconstructed the phylogenetic tree for almond from China and the Mediterranean region. Two distinct groups were formed, one for Chinese cultivars and the other for the Mediterranean cultivars which agreed with their geographical origin. This suggests that a clear dendrogram based on a combination of both EST- and g-SSRs could be reconstructed for the wheat tetraploid and hexaploid species.

When the dendrogram was constructed based on data produced by a set of EST-SSRs assigned to the A genome of bread wheat, a clear phylogenetic relationship was obtained. All the polyploid wheats were

gathered in the same cluster and the relative positions of the tetraploid and hexaploid wheat accessions remained consistent to what was observed when each group was analyzed separately. This large group was more closely related to the group including *T. urartu* (A^uA^u) sub-species while the other A-genome diploid species [*T. monococcum* (A^mA^m) and *T. boeoticum* (A^bA^b)] were clustered together and were less related to polyploid wheats. Earlier cytogenetic studies suggested that the A genome of common wheat was contributed by *T. monococcum* (Sax 1922; Lilienfeld and Kihara 1934) but more recent evidence showed that *T. urartu* contributed the A genome of hexaploid wheat (Dvorak et al. 1993; Huang et al. 2002). Our results support the latter point of view.

Similarly, as for the analysis using A-genome EST-SSRs, the polyploid wheats were consistently clustered together using B-genome EST-SSRs while fewer markers were used (10 instead of 16 for the A-genome markers). This may be due to the fact that the B genome is frequently more polymorphic than the other two (A and D) and the markers that are located on this genome are thus more appropriate to separate the different species. For the origin of the B genome of polyploid wheats, our results indicate that *T. aestivum* and *T. durum* are more closely related to *Ae. searsii* accession 4 and *Ae. speltoides* (SS) sub-species than with other species bearing the S genome [*Ae. searsii* accession 3 (S^sS^s), *Ae. bicornis* (S^bS^b), *Ae. longissima* (S^lS^l)]. The origin of the B genome has been the subject of considerable speculations and investigations and it still remains largely unresolved. Earlier studies proposed lots of possible donors in the large range of the Poaceae family (Miller 1990), whereas other experiments have focused on *Ae. speltoides* (Natarajan and Sharma 1974; Chen et al. 1975; Jaaska 1980; Hassan and Gustafson 1996; Maestra and Naranjo 1998). Our results suggest that among the species bearing the S genome, *Ae. speltoides* and *Ae searsii* sub-species have probably largely contributed to the B genome of polyploid wheats, which supports its polyphyletic origin.

In conclusion, we have shown that common wheat EST-SSRs can be useful for phylogenetic studies in the Triticeae tribe especially between distant species such as rye, barley, and wheat. They can also be used in elaborating dendrograms agreeing with those obtained using data from morphological, physiological, RFLP, and nuclear sequences. It would now be interesting to scale up the analysis by increasing the number of accessions within each species as well as the number of markers tested to have more accurate data on the Triticeae phylogeny.

Acknowledgments The authors gratefully thank G. Gay and A. Loussert for growing the plants and J. David for providing with *T. turgidum* sub-species. G. Boutet and C. Pont are also greatly acknowledged for their help in managing the robotics of the genotyping platform. This work was funded by the China Scholarship Council (CSC). All these experiments comply with the current laws of France.

References

Appels R, Reddy P, McIntyre CL, Moran LB, Frankel OH, Clarke BC (1989) The molecular-cytogenetic analysis of grasses and its application to studying relationships among species of the Triticeae. Genome 31:122–133

Bandopadhyay R, Sharma S, Rustgi S, Singh R, Kumar A, Balyan HS, Gupta PK (2004) DNA polymorphism among 18 species of *Triticum–Aegilops* complex using wheat EST-SSRs. Plant Sci 166:349–356

Ben Amer IM, Börner A, Röder MS (2001) Detection of genetic diversity in Libyan wheat genotypes using wheat microsatellite markers. Genet Resour Crop Evol 48:579–585

Brown GR, Kadel EE III, Bassoni DL, Kiehne KL, Temesgen B, van Buijtenen JP, Sewell MM, Marshall KA, Neale DB (2001) Anchor reference loci in loblolly pine (*Pinus taeda* L.) for integrating pine genomics. Genetics 159:799–809

Chen K, Gray JC, Wildman SG (1975) Fraction I protein and the origin of polyploid wheats. Science 190:1304–1305

Donini P, Law JR, Koebner RMD, Reeves JC, Cooke RJ (2000) Temporal trends in the diversity of UK wheat. Theor Appl Genet 100:912–917

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

von Eig A (1929) Monographisch-Kretische Ubersicht der Gattung Aegilops. Repertorium specierum novarum regni vegetabilis 55:24–28

Eujayl I, Sorrells ME, Baum M, Wolters P, Powell W (2001) Assessment of genotypic variation among cultivated durum wheat based on EST-SSRs and genomic SSRs. Euphytica 119:39–43

Eujayl I, Sorrells ME, Baum M, Wolters P, Powell W (2002) Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theor Appl Genet 104:399–407

Feldman M (1978) New evidence on the origin of the B genome of wheat. In: Ramanujam S (ed) Proceedings of the 5th international wheat genetics symposium, New Delhi, India, pp. 120–132

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

Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle

Gupta PK, Varshney RK (1999) Molecular markers for genetic fidelity during micropropagation and germplasm conservation. Curr Sci 77:1308–1310

Hassan MD, Gustafson JP (1996) Molecular evidence for *Triticum speltoides* as a B-genome progenitor of wheat (*Triticum aestivum*). Genome 39:543–548

Huang SX, Sirikhachornkit A, Su XJ, 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(12):8133–8138

Jaaska V (1980) Electrophoretic survey of seedling esterases in wheats in relation to their phylogeny. Theor Appl Genet 56:273–283

Jaccard P (1908) Nouvelles recherches sur la distribution florale. Bull Soc Vaud Sci Nat 4:223–270

Leisova L, Ovesna J (2001) The use of microsatellite analysis for the identification of wheat varieties. Czech J Genet Plant Breed 37:29–33

Lelley T, Stachel M, Grausgruber H, Vollmann J (2000) Analysis of relationships between *Aegilops tauschii* and the D genome of wheat utilizing microsatellites. Genome 43:661–668

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

Maestra B, Naranjo T (1998) Homoeologous relationships of *Aegilops speltoides* chromosomes to bread wheat. Theor Appl Genet 97:181–186

Manifesto MM, Schlatter AR, Hopp HE, Suarez EY, Dubcovsky J (2001) Quantitative evaluation of genetic diversity in wheat germplasm using molecular markers. Crop Sci 41:682–690

McFadden ES, Sears ER (1946) The origin of *Triticum spelta* and its free-threshing hexaploid relatives. J Hered 37:81–89

McIntyre CL (1988) Variation at isozyme loci in Triticeae. Plant Syst Evol 160:123–142

Miller TE (1990) Systematics and evolution. In: Lupton FGH (ed) Wheat breeding. Chapman and Hall Ltd, London, pp 1–30

Monte JV, McIntyre CL, Gustafson JP (1993) Analysis of phylogenetic relationship in the Triticeae tribe using RFLP. Theor Appl Genet 83:649–655

Natarajan AT, Sharma NP (1974) Chromosome banding patterns and the origin of the B genome in wheat. Genet Res 24:103–108

Nicot N, Chiquet V, Gandon B, Amilhat L, Legeai F, Leroy F, Bernard M, Sourdille P (2004) Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). Theor Appl Genet 109:800–805

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

Plaschke J, Ganal M, Röder M (1995) Detection of genetic diversity in closely related bread wheat using microsatellite markers. Theor Appl Genet 91:1001–1007

Prasad M, Varshney RK, Roy JK, Balyan HS, Gupta PK (2000) The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. Theor Appl Genet 100:584–592

Rees H, Walters MR (1965) Nuclear DNA and the evolution of the wheat. Heredity 20:73–82

Reif JC, Melchinger AE, Frisch M (2005) Genetical and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management. Crop Sci 2005 45:1–7

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

Röder MS, Wendehake K, Korzun V, Bredemeijer G, Laborie D, Bertrand L, Isaac P, Rendell S, Jackson J, Cooke R, Vosman B, Ganal M (2002) Construction and analysis of a microsatellite-based database of European wheat varieties. Theor Appl Genet 106:67–73

Roussel V, Leisova L, Exbrayat F, Stehno Z, Balfourier F (2005) SSR allelic diversity changes in 480 European bread wheat varieties released from 1840 to 2000. Theor Appl Genet 111:162–170

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

Salamini F, Özkan H, Brandolini A, Schäfer-Pregl R, Martin W (2002) Genetics and geography of wild cereal domestication in the Near East. Nat Rev Genet 3:429–441

Sarkar P, Stebbins GL (1956) Morphological evidence concerning the origin of the B genome of wheat. Am J Bot 43:297–304

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

Sasanuma T, Chabane K, Endo TR, Valkoun J (2004) Characterization of genetic variation in and phylogenetic relationships among diploid *Aegilops* species by AFLP: incongruity of chloroplast and nuclear data. Theor Appl Genet 108:612–618

Sax K (1922) Sterility in wheat hybrids. II. Chromosome behavior in partially sterile hybrids. Genetics 7:513–552

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462

Sourdille P, Tavaud M, Charmet G, Bernard M (2001) Transferability of wheat microsatellites to diploid Triticeae species carrying the A, B and D genomes. Theor Appl Genet 103:346–352

Tixier MH, Sourdille P, Charmet G, Gay G, Jaby C, Cadalen T, Bernard S, Nicolas P, Bernard M (1998) Detection of QTLs for crossability in wheat using a doubled-haploid population. Theor Appl Genet 97:1076–1082

Xu Y, Ma RC, Xie H, Liu JT, Cao MQ (2004) Development of SSR markers for the phylogenetic analysis of almond trees from China and the Mediterranean region. Genome 47(6):1091–1104

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

Zhang LY, Bernard M, Leroy P, Feuillet C, Sourdille P (2005) High transferability of bread wheat EST-derived SSRs to other cereals. Theor Appl Genet 111:677–687