Genetic diversity and population structure of Triticum araraticum and Triticum timopheevii

Ting Peng
Sichuan Agricultural University
Xiaomei Jiang
Sichuan Agricultural University
Dinghao Wang
Sichuan Agricultural University - Chengdu Campus
Minghu Zhang
Sichuan Agricultural University - Chengdu Campus
Xin Liu
Sichuan Agricultural University - Chengdu Campus
Ming Hao
Sichuan Agricultural University - Chengdu Campus
Wei Li
Sichuan Agricultural University - Chengdu Campus
Dengcai Liu
Sichuan Agricultural University - Chengdu Campus
Bo Jiang
Sichuan Agricultural University - Chengdu Campus
Lin Huang
Sichuan Agricultural University - Chengdu Campus
Shunzong Ning
Sichuan Agricultural University - Chengdu Campus
Zhongwei Yuan
Sichuan Agricultural University - Chengdu Campus
Bihua Wu
Sichuan Agricultural University - Chengdu Campus
Zehong Yan
Sichuan Agricultural University - Chengdu Campus
Xuejiao Chen
Sichuan Agricultural University - Chengdu Campus
Xue Chen
Sichuan Agricultural University - Chengdu Campus
Lianquan Zhang (zhanglianquan1977@126.com)
Sichuan Agricultural University - Chengdu Campus https://orcid.org/0000-0001-9005-8705

Research Article

Keywords: T. araraticum, T. timopheevii, genetic diversity, population structure, FISH, GBS

Posted Date: May 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1418521/v1

Version of Record: A version of this preprint was published at Genetic Resources and Crop Evolution on January 23rd, 2023. See the published version at https://doi.org/10.1007/s10722-023-01537-4.

Abstract

Genetic diversity is an important resource to improve new wheat cultivars in the breeding program. In this study, fluorescence in situ hybridization (FISH) and genotyping by sequencing (GBS) were used to evaluate the genetic diversity and population structure of the timopheevii group (A^tA^tGG, 2n = 4x = 28) including 15 *Triticum timopheevii* Zhuk. accessions and 35 *Triticum araraticum* Jakubz. accessions. FISH analysis showed that there were 14 and 19 FISH signal variations in the A^t and G genome of *T. araraticum*, respectively. But there were only two signal variations in the G genome of *T. timopheevii*. In this study, 190,402 SNP markers were obtained from GBS, in which the lowest and highest frequencies of SNPs were found in the G and A^t genomes, respectively. Genetic diversity analysis of the 50 accessions indicated that the mean GD and PIC were 0.30 and 0.26, with the ranges of 0.1–0.5 and 0.1–0.4, respectively. The highest and lowest numbers of SNPs identified on the chromosomes 4G and $2A^t$ were 7,748 and 17,527, respectively. Structure and cluster analyses divided 50 accessions into two subpopulations (POP1 and POP2). POP1 mainly consisted of most *Triticum araraticum* accessions, and POP2 comprised all accessions of *Triticum timopheevii* and one *Triticum araraticum* accession AS273. In addition, the genetic variance showed that genetic variation was greater within populations (97%) than between populations (3%). POP1 and POP2 had high levels of genetic diversity and POP1 showed higher genetic diversity. The POP1 (*I* = 0.53, *He* = 0.35, *uh* = 0.35) showed higher genetic diversity than POP2. The present results provided important information for the improvement of cultivated durum and hexaploid wheat production in future breeding programs of China and other countries.

Introduction

The genus *Triticum* contains tetraploid emmer (2n = 2x = 28, AABB), diploid einkorn wheat (2n = 2x = 14, AA), timopheevii (2n = 2x = 28, A^tA^tGG) wheat, and hexaploid common wheat (2n = 6x = 42, AABBDD) and *Triticum. zhukovskyi* (2n = 6x = 42, A^tA^tGGAA) (Lilienfeld 1951; Menabde and Ericzjan 1960). The Timopheevii group of species *Triticum timopheevii* Zhuk. and *Triticum araraticum* Jakubz. belong to the secondary pool of common wheat, have partial homology with the common wheat genome (Timonova et al. 2013). The degree of sequence divergence between the B and G genomes is greater than that between the A and A^t genomes (Rodríguez et al. 2000; Huang et al. 2002; Salina et al. 2006). The single nucleotide polymorphisms (SNPs) derived from genotyping-by-sequencing (GBS) indicated that the *T. timopheevii* genome is clustered more closely to the AABB(DD) lineage than the species AA genome based on the B genome tree (Hyun et al. 2020).

T. timopheevii is a cultivated tetraploid wheat species endemic to western Georgia (Dorofeev et al. 1979), which was first discovered by the Russian botanist Piotr Zhukovsky in western Georgia in 1922 (Zhukovsky et al. 1924). Later, it was also found in other regions of western Georgia, where the local inhabitants consume a kind of gruel prepared by naturally mixing *Triticum monococcum* and *T. timopheevii* (Belea 1992; Zohary et al. 2012). Recently, ten accessions collected in Turkey and Iraq that were originally classified as *Triticum turgidum* were identified as *T. timopheevii* by *Ppd-1* typing (Czajkowska et al. 2019). Caferhöyük, Eastern Turkey is the oldest site where the “new” glume wheat (NGW, assumed to be *T. timopheevii*) has been identified (de Moulins 1993). The NGW of early Eurasian agriculture, grown by prehistoric farmers throughout western Asia and eastern Europe, is a cultivated member of the *T. timopheevii* group, as proven by ancient DNA typing (Czajkowska et al. 2020). The remains of NGW spikelet bases have been documented in two settlements in Turkey: Yenikapı (eastern Thrace) and Yumuktepe (Cilicia) (Ulaş and Fiorentino 2021). Currently, this species is held within seed banks and special collections ex situ, and is no longer cultivated as an agricultural crop (Dorofeev et al. 1979; Badaeva et al. 1994a; Czajkowska et al.

2020). *T. timopheevii* is resistant to fungal diseases such as leaf rust, powdery mildew, smut and stem rust, and therefore can be used for wheat breeding improvement by directly crossing it with common wheat or using synthetic amphiploids as a bridge (McIntosh and Gysuch 1971; Badaeva et al. 1991; Peusha et al. 1996; Järve et al. 2002; Mikó et al. 2015; Liu et al. 2018).

T. timopheevii was domesticated from the wild species *T. araraticum* (Dorofeev et al. 1979; Badaeva et al. 1994a; Kilian et al. 2009; Mori et al. 2009). *T. araraticum* was first collected in 1925–1930 in Armenia and classified as a separate species by Dr. Jakubziner in 1931 (Badaeva et al. 1994b). Geographical distribution of *T. araraticum* covers southeastern Turkey, western Iran, northern Iraq, and some regions of Transcaucasia: the Nakhichevan Autonomous Republic, Armenia, and northeastern Azerbaijan (Badaeva et al. 1994b; Zohary and Hopf 2012). The A-genome donor of *T. araraticum* and *T. timopheevii* is *Triticum urartu* (Dvorak et al. 1993; Edet et al. 2018; Hyun et al. 2020). However, the G-genome donor is *Aegilops speltoides* or *Aegilops searsii* (Mori et al., 1997; Golovnina et al. 2007; Kilian et al. 2007; Takahashi et al. 2010; Edet et al. 2018). Analysis of nucleotide sequence divergence in different cereal species revealed that the progenitor of *T. timopheevii* wheat might have emerged around 3–9 million years ago (Huang et al. 2002; Gornicki et al. 2014). To date, four possible domestication sites have been reported about the origin of *T. timopheevii* in Georgia, including the Transcaucasian mountains (Jakubziner 1933, 1950; Nesbitt and Samuel 1996), northern Iraq (Wagenaar 1966), northern Syria or southern Turkey (Mori et al. 2009), and the southeastern part of the Taurus region of Turkey (Ulaş and Fiorentino 2021)..

The karyotype of *T. araraticum* and *T. timopheevii* has been comprehensively characterized using chromosome banding (C-banding) (Badaeva et al. 1994b). Although C-banding technology can clearly reveal the distribution of constitutive chromatin on chromosomes, it is impossible to determine the composition of heterochromatin. Even though some chromosomes have similar C-banding pattern distributions, it cannot be guaranteed that they have the same origin. In recent years, fluorescence in situ hybridization (FISH) has been found as an effective tool for precisely identifying the chromosomes of the A^t and G genomes of *T. timopheevii* (Mikó et al. 2015; Badaeva et al. 2016; Liu et al. 2018). Oligo-pTa535-1 preferentially targets tandem repeats on A^t-genome chromosomes of *T. timopheevii* (Mikó et al. 2015; Badaeva et al. 2016; Liu et al. 2018). Oligo-pSc119.2-1 mainly hybridizes to the G-genome chromosomes of *T. timopheevii*, and Oligo-pTa71-2 mainly hybridizes to the subterminal regions of the short arms of $6A^t$ and 6G (Mikó et al. 2015; Liu et al. 2018). The combination of oligonucleotide probes OligopSc119.2-1, Oligo-pTa535-1, and Oligo-pTa71-2 can distinguish individual chromosomes originating from *T. timopheevii* (Liu et al. 2018).

SNP is the most common type of sequence variation in plant genomes (Batley and Edwards 2007) and is appropriate for the analyses of population structure, genetic variation, genomic selection, marker-trait association, QTL mapping and so on (Kumar et al. 2012). Numerous SNPs can be discovered using high-throughput sequencing technology (He et al. 2014). With the advent of next-generation sequencing (NGS) technologies, GBS has been commonly applied for biomarker disccovery and detection in plant breeding programs (Elshire et al. 2011; Poland et al. 2012; He et al. 2014). An improved GBS protocol has been successfully employed in cereal crops, including oat, wheat and barley (Poland et al. 2012; Huang et al. 2014; Alipour et al. 2017; Hyun et al. 2020).

Plant genetic diversity analysis is a crucial aspect for plant evolution, conservation, inheritance and breeding (Peterson et al. 2014). Intraspecific divergence of *T. araraticum* was generated by structural rearrangements using C-banding analysis (Badaeva et al. 1990; Badaeva et al. 1994b). Up to now, genetic diversity in *T. timopheevii* was only showed by C-banding patterns and chloroplast DNA fingerprinting (Badaeva et al. 1994a; Mori et al. 2009). In the present study, genetic diversity was evaluated by FISH and GBS using 15 *T. timopheevii* accessions and 35 *T.*

araraticum accessions. The objectives were to evaluate the population structure and genetic diversity of the 50 *timopheevii* accessions and to characterize genetic differentiation within and between the subpopulations. This study lays a foundation for the future genomic selection in wheat breeding program.

Material And Methods

Plant samples

Thirty-five *T. araraticum* and fifteen *T. timopheevii* accessions were employed in the present study (Table S1 and Table S2). *T. timopheevii* accessions were from nine countries and *T. araraticum* accessions were from four countries. Among them, the spike type of AS273 was different from the typical spike type of *T. timopheevii* and *T. araraticum*, although it was classified into *T. araraticum* because of its fragility of rachis similar to *T. araraticum* (Fig. 1). In addition, 10 different *T. turgidum* accessions were used as controls in FISH signal cluster analysis (Table S3). The lines with code AS and D were maintained at our institute, while those with code PI were supplied by the USDA-ARS germplasm bank (http://www.ars-grin.gov).

FISH karyotyping and data analysis

Cytological observation of the root tip cells was conducted as described previously (Zhang et al., 2007), and their chromosome number was also calculated. The FISH procedure was the same as that reported by Hao et al. (2013). Three probes, namely, Oligo-pSc119.2-1, Oligo-pTa535-1, and Oligo-pTa71-2, were used in this study. The probes were synthesized and labelled with TAMRA or FAM by TSINGKE Biological Technology Company (Chengdu, China). At least 3 split phases with good signals were selected in each wheat FISH image for observation. Olympus BX-63 epifluorescence microscope was used to observe hybridization signals, and the images were recorded using a Photometric SenSys Olympus DP70 CCD camera (Olympus, Tokyo, Japan). Photoshop ver. 7.1 (Adobe Systems Incorporated, San Jose, CA) was used to process raw images. The identification of A^t- and G-genome chromosomes was conducted as described previously (Liu et al., 2018). The signal sites were counted using Excel. NTSYS 2.10 software (Rohlf 2000) was employed to measure genetic similarity and distance between samples. Cluster analysis was conducted using the unweighted pair-group method with arithmetic mean (UPGMA) approach according to the matrix of genetic similarity coefficients, and a dendrogram was plotted.

GBS library preparation and sequencing

Hi-DNAsecure Plant Kit DP350 (TIANGEN, Beijing, China) was used to extract total genomic DNA from fresh young leaves of 2-week-old seedlings. The quality and integrity of DNA extracts were assessed by 1% agarose gel electrophoresis with a standard DNA ladder of size λ, and the yields were measure using a NanoDrop 2000 (Thermo Scientific, MA, USA) and Qubit® 2.0 fluorometer (Invitrogen, Carlsbad, USA). GBS library was constructed according to a previous protocol (Poland et al. 2012). Genome sequencing was conducted on 50 samples using an Illumina HiSeq PE150 platform. Burrows-Wheeler Aligner v0.7.8 was used to map the paired-end reads to the *Triticum dicoccoides* reference genome (Li and Durbin 2009). The reference genome was from the NCBI database (*T. dicoccoides* WEW_v2.0). The comparison results were sorted by SAMTOOLS (Li et al. 2009). The obtained SNPs were filtered by eliminating those with MAF < 0.01 and missing rate > 30% (Danecek et al. 2011; Luo et al. 2019). The markers were delivered as HapMap and variant call format (VCF) (v0.1.10) files (Danecek et al. 2011).

Population structure analysis

ADMIXTURE software was used to estimate population structure (Alexander et al. 2009; Li et al. 2019). The best K with minimum cross-validation error was evaluated according to a previous method (Alexander et al. 2009). TreeBeST (http://treesoft.sourceforge.net/treebest.shtml) was used to constructed unrooted neighbor-joining phylogenetic trees. The confidence intervals of the genetic association among the accessions were measured by conducting 1,000 bootstraps. The value was expressed as a percentage at the main nodes of each branch. GCTA software (http://cnsgenomics.com/software/gcta/pca.html) was used to calculate the feature vectors and eigenvalues of the PCA. The distribution map of the PCA was constructed with R software v3.5.0.

Statistical analysis

PowerMarker 3.25 software was used to calculate the basic summary statistics, including polymorphism information content (PIC), gene diversity (GD), observed heterozygosity (*Ho*) and MAF (Liu and Muse 2005). All SNPs were mapped to the *Triticum dicoccoides* reference genome (*T. dicoccoides* WEW_v2.0). The SNPs with a PIC of 0.27–0.30 were chosen and 7,212 SNPs were analyzed by AMOVA. The number of subpopulations measured by STRUCTURE analysis was selected for AMOVA. Genetic indices containing number of effective alleles (*Ne*), number of alleles (*Na*), observed heterozygosity (*Ho*), expected heterozygosity (*He*), Shannon's information index (*I*) and unbiased diversity index (*uh*) were determined. GeneAlEx 6.41 was used to perform AMOVA and estimate genetic indices (Peakall and Smouse 2006).

Results

FISH karyotype polymorphism of *T. araraticum* and *T. timopheevii*

Probe Oligo-pTa535-1 generated signals predominately on the A^t-genome chromosomes. However, several G-genome chromosomes showed different hybridization locations, and only a small number of signals were distributed on chromosomes 4G, 6G, and 7G. Oligo-pSc119.2-1 mainly hybridized with the G-genome chromosomes, and signals were located in the interstitial and subterminal regions, which allow the identification of 9 out of 14 chromosomes of *T. timopheevii*. There were a small number of signals on 1A and $5A^t$. Oligo-pTa71-2 mainly hybridized with the $6A^t$ and 6G chromosomes, and the signal was on the short arm.

There were FISH signal polymorphisms in the 50 accessions used in this study (Fig. 2 and Fig. 3). There was no signal variation in the A^t genome but two signal variations in the G genome of *T. timopheevii* (Fig. 2). PI119442 was the most typical FISH karyotype among the 15 *T. timopheevii* accessions. Seven chromosomes of the A^t genome were monomorphic and very conserved in all 15 *T. timopheevii* accessions (Fig. 2a). Both chromosomes 2G and 7G had one kind of signal variation in the G genome of all 15 *T. timopheevii* accessions (Fig. 2b). However, there were 14 kinds of FISH signal variations in the A^t genome (Fig. 3a) and 19 kinds of FISH signal polymorphisms in the G genome of *T. araraticum* (Fig. 3b). Chromosome $1A^t$ had four kinds of signal variations, showing the highest polymorphism. Chromosomes $4A^t$ and $5A^t$ each had three kinds of signal variations, showing strong signal polymorphism. Chromosome $2A^t$ had two kinds of signal variations. Chromosomes $3A^t$ and $7A^t$ each had one kind of signal variation, showing the lowest signal polymorphism. Notably, chromosome $6A^t$ was monomorphic. Chromosome 2G had seven kinds of signal variations, showing the highest signal polymorphism. Chromosomes 6G and 7G each had four kinds of signal variations, showing strong signal polymorphism. Chromosomes 1G and

4G each had two kinds of signal variations, showing low signal polymorphism. Similarly, it is worth noting that both chromosomes 3G and 5G were monomorphic. However, compared with the FISH signal of *T. timopheevii* PI119442, additional sets of red signals were observed on the short arms of $5A^t$ and 2G in *T. araraticum* AS273 (Fig. 4), and on the long and short arms of the chromosomes $1A^t$ and 6G, respectively, in *T. araraticum* AS274 and PI427414 (Fig. 4).

FISH signal cluster analysis

Cluster analysis was performed according to the physical location of the FISH signals from 50 accessions using 10 different *T. turgidum* accessions as outgroup, with a similarity coefficient of 0.57 as the standard (Fig. 5). Sixty accessions were divided into two groups based on their different FISH signals. The 50 accessions with the A^tA^tGG genome were clustered together, and the 10 accessions with the AABB genome were clustered together. Fifty accessions with the A^tA^tGG genome were classified as 5 clusters according to a similarity coefficient of 0.89. Among the clusters, cluster 1 was mostly composed of *T. timopheevii*, and clusters 2, 3, 4, and 5 were mostly composed of *T. araraticum. T. araraticum* and *T. timopheevii* showed strong genetic differentiation, with *T. araraticum* showing strong genetic diversity. However, three *T. araraticum* accessions, AS273, AS274 and PI427414, were included in the *T. timopheevii* group, showing that these 3 accessions were closely related to this group. *T. araraticum* was concentrated in Iraq, and the geographical distributions of *T. araraticum* and *T. timopheevii* overlapped in Turkey (Fig. 6).

Chromosome distribution of SNPs

Illumina HiSeq PE150 generated a total of 7,141,920 raw single-end sequence reads from 50 samples. It is assumed that G genome SNPs from the *T. timopheevii* accessions mapped to the B genome of the *T. dicoccoides* reference sequence, in view due to a close similarity between these two genomes (Hyun et al. 2020). After quality filtering, 190,402 SNPs were identified in the A^t, G genomes (Table S4, https://doi.org/10.6084/m9.figshare.17032322.v1), including 104,180 SNPs from the A^t genome, 84,930 SNPs from the G genome and 1,292 SNPs from unanchored scaffolds (Fig. 7a, Table S5). In the A^t genome, the chromosomes $1A^t$ and $2A^t$ harbored the lowest (1,1982) and highest (17,527) number of SNPs, respectively. In the G genome, the chromosomes 4G and 5G had the lowest (7,748) and highest (14,637) numbers of SNPs, respectively (Fig. 7b). The highest and lowest numbers of SNPs identified on the chromosomes 4G and $2A^t$ were 17,527 and 7,748, respectively (Fig. 7b). The ratio of SNP numbers in the A^t to G genomes was 1.23. Thus, the SNP number of A^t genome exceed that of G genome.

Population structure

The resulting dataset of 190,402 high-quality SNPs from GBS was used for cluster analysis. *T. araraticum* and *T. timopheevii* were clearly divided into two subpopulations (Fig. 8). Subpopulation 1 (POP1) contained most *T. araraticum* accessions except for AS273. Subpopulation 2 (POP2) contained all *T. timopheevii* accessions including one *T. araraticum* accession AS273. Two *T. timopheevii* accessions, CItr15590 and CItr15205 from Greece were grouped together with PI94760 from Georgia, with bootstrap values reaching 100%. PI221421 from Serbia was grouped together with PI251018 from Hungary, with bootstrap values of 100%. Most *T. timopheevii* are related

to PI119442 from Turkey, with bootstrap values of 65%. *T. araraticum* accessions from Iraq were divided into 6 subgroups (Fig. 8). Two *T. araraticum* accessions PI352265 and AS272 from Azerbaijan were grouped together with PI427339 from Iraq. Similarly, two *T. araraticum* accessions PI427366 and PI427370 from Iran were grouped together with PI427363 from Iraq.

When K had an inflection point, the cross-validation (CV) errors were low, with K = 2 minimizing the CV errors (Fig. 9). Clustering information for 50 individuals was shown when K = 1 to 8 (Fig. 10). K = 2, *T. araraticum* was first distinguished from *T. timopheevii*. The separation of these two sub-populations was also well supported by PCA (Fig. 11). K = 3, *T. araraticum* were divided into two subgroups, one subgroup includes three taxa A2, A6, and A8, and the other subgroup includes three taxa A4, A7, and A10. Nearly 70% of the genetic information for A5 was originated from the ancestors A4, A7, and A10. Almost half of the genetic information for AS273 (A9) came from the ancestor A11, and the remaining information was derived from the ancestor A2, A6, and A8.

One hundred percent different SNPs corresponding to the same physical location were screened among 35 and 15 accessions of *T. araraticum* and *T. timopheevii* to ensure that the SNPs at each corresponding physical location were common to all accessions in the respective populations. A total of 4139 SNPs were screened out (Table S6). The different SNPs between *T. araraticum* and *T. timopheevii* were widely distributed over all of the chromosomes (Fig. 12), and the most different SNPs were distributed on chromosomes $2A^t$, $3A^t$, $5A^t$, and $6A^t$. A special *T. araraticum* accession AS273, having an intermediate spike type was grouped with *T. timopheevii*, corresponding to the result of FISH signal cluster analysis. The analysis of 3034 specific SNPs from AS273 indicated that chromosomes $2A^t$, $5A^t$, $7A^t$, 3G, 4G and 6G were from *T. araraticum* and chromosomes $1A^t$, $3A^t$, $4A^t$, $6A^t$, 1G, 2G, 5G, and 7G were from *T. timopheevii* (Fig. 13).

Genetic diversity

Genetic diversity analysis of the 50 accessions indicated that the mean GD and PIC were 0.30 and 0.26, with the ranges of 0.1–0.5 and 0.1–0.4, respectively (Fig. 14a, b). The *H*o values ranged from 0 to 1, and the value was 0.1 when the samples were genotyped on a large number of markers (Fig. 14c). The average MAF was 0.26 (Fig. 14d). Intra-population genetic diversity analysis demonstrated that mean *N*a and *N*e values were 1.97 and 1.50 for the two subpopulations, respectively (Table 1). The *Ne* value of Pop1 (1.54) was higher than that of Pop2 (1.46). The mean values of *He*, *Ho*, *I* and *uh* were 0.32, 0.35, 0.49 and 0.33, respectively. As shown in Table 1, the Pop1 (*I* = 0.53, *He* = 0.35, *uh* = 0.35) showed higher genetic diversity than Pop2 (*I* = 0.45, *He* = 0.29, *uh* = 0.31). In addition, the results of AMOVA showed that much greater variation within populations (97%) was observed than among the populations (3%) (Table 2). Notably, the N_m value was relatively high (8.39), suggesting high gene flow between subpopulations (Table 2). These findings demonstrate a high genetic diversity within subpopulations, but a low genetic diversity between subpopulations.

Table 1
Means of genetic parameters for each subpopulation of the 50 timopheevii accessions. Number of alleles (*Na*), number of effective allele (*Ne*), Shannon's index (*I*), observed heterozygosity (*Ho*), expected heterozygosity (*He*), and unbiased diversity index (*uh*).

Population	*Na*	*Ne*	*I*	*Ho*	*He*	*uh*
Pop1	2.00	1.54	0.53	0.35	0.35	0.35
Pop2	1.94	1.46	0.45	0.36	0.29	0.31
Mean	1.97	1.50	0.49	0.35	0.32	0.33

Table 2
Analysis of molecular variance using 7212 SNP markers of genetic differentiation among and within two subpopulations of the 50 timopheevii accessions.

Source	df	SS	MS	Est. Var.	%	*P* value
Among Pops	1	9167.11	9167.11	167.97	3%	0.001
Within Pops	48	270708.25	5639.76	5639.76	97%	0.001
Total	49	279875.36		5807.73	100%	0.001
Nm	8.39					

Discussion

Plant genetic diversity offers opportunities for plant breeders to create new and improved varieties with desirable character (Govindaraj et al. 2015). *T. timopheevii* and *T. araraticum* are important genetic resources for improving wheat production. In the present study, FISH and GBS were used to evaluate genetic diversity in 35 and 15 accessions of *T. araraticum* and *T. timopheevii*, respectively, which were collected from 13 countries. In the FISH karyotype analysis of *T. timopheevii*, all 15 accessions had quite limited FISH signal variation, which was relatively stable throughout the distribution area. In the FISH karyotype analysis of *T. araraticum*, 33 FISH signal variations were found among the 35 accessions. However, chromosomes 6A^t, 3G, and 5G showed only one signal type, indicating that these three chromosomes remained more stable than other chromosomes during domestication, which is inconsistent with the findings of C-banding (Badaeva et al. 1990; Badaeva et al. 1994b). Cluster analysis indicated that the 50 *timopheevii* accessions with the A^tA^tGG genome were clustered, and the 10 *T. turgidum* accessions with the AABB genome were clustered, indicating different evolutionary lineages of *T. timopheevii* and *T. turgidum* (Oliveira et al. 2020). However, 50 accessions with the A^tA^tGG genome were divided into 5 clusters based on a similarity coefficient of 0.89.

In this study, 190,402 SNP markers were obtained from GBS, in which the lowest and highest frequencies of SNPs were found in the G and A^t genomes, respectively. The highest and lowest numbers of SNPs identified on the chromosomes 4G and 2A^t were 7,748 and 17,527, respectively. Genetic diversity and PIC values are greatly useful for assessing the polymorphism levels among genotypes in populations (Salem and Sallam 2016). In the present study, the mean PIC and GD were 0.26 and 0.3, respectively, (Fig. 14A, B). The PIC values were classified into three categories by Botstein et al. (1980): (1) if PIC > 0.5, the SNP is highly informative, (2) if 0.25 < PIC < 0.5, the SNP is a moderately informative, and (3) if PIC < 0.25, then the SNP is slightly informative. Yang and co-workers (2020)

analyzed 180 Asian and European common wheat accessions, and found a mean PIC of 0.27, indicating that the population has moderate genetic diversity. Our results showed that the mean PIC value was 0.26, suggesting that the 50 accessions have moderate polymorphism.

In the current study, the findings of AMOVA revealed high genetic diversity within populations (97%), while the variation between the populations was low (3%). The low genetic diversity between subpopulations was caused by high gene flow (Arora et al. 2014). It has been shown that the *Nm* (haploid) values of < 1 indicate low gene flow between subpopulations (Wright 1965). In this study, the N_m value was considerably high (8.39), suggesting high gene flow between subpopulations. These findings demonstrate high genetic diversity within subpopulations, while low genetic diversity between subpopulations. Population structure analysis is useful in understanding genetic diversity. The 50 timopheevii accessions were classified as 2 subpopulations according to the population structure data (Fig. 10). The PCA results (Fig. 11) were in agreement with STRUCTURE results.

Cluster analysis of 190,402 high-quality SNPs from GBS showed high genetic differentiation between *T. araraticum* and *T. timopheevii*, which agrees with previous studies revealed by C-banding (Badaeva et al. 1990; Badaeva et al. 1994b; Badaeva et al. 2016) and the study of storage proteins (Jakobashvili 1989). Based on the different SNPs derived from GBS, two main subpopulations were identified. In total, 4139 different SNPs between *T. araraticum* and *T. timopheevii* were obtained. These specific chromosomal regions may have been involved in the domestication of *T. timopheevii*. Most of the different SNPs were distributed on chromosomes $2A^t$, $3A^t$, $5A^t$, and $6A^t$. A special *T. araraticum* accession AS273, having an intermediate spike type, was clustered with all *T. timopheevii* accessions as revealed by cluster analysis. The analysis of specific SNPs from AS273 indicated that chromosomes $2A^t$, $5A^t$, $7A^t$, 3G, 4G and 6G were from *T. araraticum* and chromosomes $1A^t$, $3A^t$, $4A^t$, $6A^t$, 1G, 2G, 5G, and 7G were from *T. timopheevii.* It was showed that more than half of the genetic information of AS273 came from *T. timopheevii*, and the remaining information was derived from *T. araraticum*. We speculated that there were two possibilities about the origin of AS273. One possibility is that AS273 may have different evolutionary pathway of having A^t and G genome from the same parentage. The other possibility is that AS273 may result from the gene flow between *T. araraticum* and *T. timopheevii* because of cross-pollination, which does occur in bread wheat when florets open and self-pollination fails (Keydal 1978; Barradas and Romano 1979-80; Griffin 1987; Martin 1990; Hucl 1996; Okada et al. 2018).

To date, *T. timopheevii* was mainly used for wheat disease resistance breeding by directly crossing it with common wheat or using synthetic amphiploids as a bridge because of its resistance to fungal diseases, including leaf rust, powdery mildew, smut and stem rust (McIntosh and Gysuch 1971; Badaeva et al. 1991; Peusha et al. 1996; Järve et al. 2002; Mikó et al. 2015; Liu et al. 2018). Our results indicated that high genetic differentiation existed within *T. araraticum* and *T. timopheevii* populations collected from 35 countries. *T. timopheevii* was cultivated more extensively in regions with a humid climate (average rainfall per year 900–1200 mm) at altitudes of 300–1200 m (Badaeva et al. 1994b), and adapted to the colder and wetter regions of European countries (Ulaş and Fiorentino 2021). *T. araraticum* was distributed in regions with a drought climate, including southeastern Turkey, western Iran, northern Iraq, and some regions of Transcaucasia: the Nakhichevan Autonomous Republic, Armenia, and northeastern Azerbaijan (Badaeva et al. 1990). Therefore, both *T. araraticum* and *T. timopheevii* can be used as potential genetic sources for biotic and abiotic stress tolerance. Breeders can use them for overcoming resistance to biotic and abiotic stresses in future wheat breeding programs of China and other countries.

Conclusion

In this study, FISH and GBS were utilized to evaluate the population structure and genetic diversity of 50 timopheevii wheat accessions. Structure and cluster analyses divided 50 accessions into two subpopulations (POP1 and POP2). POP1 mainly consisted of most *Triticum araraticum* accessions, and POP2 comprised all accessions of *Triticum timopheevii* and one *Triticum araraticum* accession AS273. In addition, the genetic variance showed that genetic variation was greater within populations (97%) than between populations (3%). POP1 and POP2 had high levels of genetic diversity and POP1 showed higher genetic diversity. This study showed that the 50 timopheevii accessions having high genetic diversity could be employed for future breeding programs to establish new durum and hexaploid wheat cultivars.

Declarations

Availability of data and materials

The data that support the findings of this study are available in 'figshare' with the identifier data DOIs, including two dataset (Tetraploid wheat-GBS.vcf, https://doi.org/10.6084/m9.figshare.16692490.v1 and https://doi.org/10.6084/m9.figshare.17032322.v1)

Authors' contributions

TP and LQZ designed the research and wrote the manuscript. TP, XMJ, DHW, MHZ and XL performed the experiments. MH and XC helped with data analysis, and WL, BJ, LH, SZN, ZWY, DCL, BHW, ZHY and XJC supervised the study. All authors read and approved the final manuscript.

Funding

This work was supported by the Key Research and Development Program of Sichuan Province, China (2021YFYZ0002) and the National Natural Science Foundation of China (91935303 and 31671682).

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Research involving human participants and/or animals None.

Informed consent None.

References

1. Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19(9):1655–1664. doi: 10.1101/gr.094052.109
2. Alipour H, Bihamta MR, Mohammadi V, Peyghambari SA, Bai G, Zhang G (2017) Genotyping-by-sequencing (GBS) revealed molecular genetic diversity of Iranian wheat landraces and cultivars. Front Plant Sci 8:1293. doi: 10.3389/fpls.2017.01293
3. Arora A, Kundu S, Dilbaghi N, Sharma I, Tiwari R (2014) Population structure and genetic diversity among Indian wheat varieties using microsatellite (SSR) markers. Aust J Crop Sci 8:1281–1289

4. Badaeva ED, Boguslavsky RL, Badaev NS, Zelenin AV (1990) Intraspecific chromosomal polymorphism of *Triticum araraticum* (Poaceae) detected by C-banding technique. Pl Syst Evol 169: 13-24.10. doi: 1007/BF00935980
5. Badaeva ED, Filatenko AA, Badaev NS (1994a) Cytogenetic investigation of *Triticum timopheevi* (Zhuk.) Zhuk. and related species using C-banding. Theor Appl Genet 89: 622–628. doi: 10.1007/BF00222457
6. Badaeva ED, Badaev NS, Gill BS, Filatenko AA (1994b) Intraspecific karyotype divergence in *Triticum araraticum* (Poaceae). Pl Syst Evol 192:117–145. doi: 10.1007/BF00985912
7. Badaeva ED, Budashkina EB, Badaev NS, Kalinina NP, Shkutina FM (1991) General features of chromosome substitutions in *Triticum aestivum* × *T. timopheevi* hybrids. Theor Appl Genet 82:227–232. doi: 10.1007/BF00226218
8. Badaeva ED, Ruban AS, Zoshchuk SA, Surzhikov SA, Knu¨pffer H, Kilian B (2016) Molecular cytogenetic characterization of *Triticum timopheevi* chromosomes provides new insight on genome evolution of *T. zhukovskyi*. Plant Syst Evol 302:943–956. doi: 10.1007/s00606-016-1309-3
9. Barradas MC, Romano MCS (1979) -80) Hybrid wheat: I. Investigation of the capacity for outcrossing in *Triticum aestivum* subsp. *vulgare*. Melhoramento 28: 117–126
10. Batley J, Edwards D (2007) "SNP applications in plants," in Association Mapping in Plants, Oraguzie NC, Rikkerink EHA, Gardiner SE and De Silva HN, Eds Springer: New York NY pp95-102. doi: 10.1007/978-0-387-36011-9-6
11. Belea A (1992) Interspecific and intergeneric crosses in cultivated plants. Akademiai Kiado, Budapest pp 54–66
12. Berg EE, Hamrick JL (1997) Quantification of genetic diversity at allozyme loci. Can J For Res 27(3):415–424. doi: 10.1139/x96-195
13. Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331. doi: 10.1016/0165-1161(81)90274-0
14. Czajkowska BI, Oliveira HR, Brown TA (2019) A discriminatory test for the wheat B and G genomes reveals misclassified accessions of *Triticum timopheevi* and *Triticum turgidum*. PLoS ONE 14(4):e0215175. doi: 10.1371/journal.pone.0215175
15. Czajkowska BI, Bogaard A, Charles M, Jones G, Brown TA (2020) Ancient DNA typing indicates that the "new" glume wheat of early Eurasian agriculture is a cultivated member of the *Triticum timopheevi* group. J Archaeol Sci 123:105258. doi: 10.1016/j.jas.2020.105258
16. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158. doi: 10.1093/bioinformatics/btr330
17. De Moulins D (1993) Les restes des plantes carbonisées du Cafer Höyük. Cahiers de l'Euphrate 7:191–234
18. Dorofeev VF, Filatenko AA, Migushova EF, Udachin RA, Jakubziner MM (1979) In: Brezhnev DD (ed) Cultivated flora of the USSR, vol 1. Kolos (Leningrad branch), Leningrad. (in Russian).
19. Dvorak JDI, Terlizzi P, Zhang HB, Resta P (1993) The evolution of polyploid wheats: identification of the A genome donor species. Genome 36:21–31. doi: 10.1139/g93-004
20. Edet OU, Gorafi YSA, Nasuda S, Tsujimoto H (2018) DArTseq-based analysis of genomic relationships among species of tribe Triticeae. Sci Rep 8:16397–16397. doi: 10.1038/s41598-018-34811-y

21. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES et al (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6:e19379. doi: 10.1371/journal.pone.0019379
22. Golovnina K, Gluskov S, Blinov A, Myorov V, Adkison L, Goncharov N (2007) Molecular phylogeny of the genus *Triticum* L. Plant Syst Evol 264:195–216. doi: 10.1007/s00606-006-0478-x
23. Gornicki P, Zhu H, Wang J, Challa GS, Zhang Z, Gill BS et al (2014) The chloroplast view of the evolution of polyploid wheat. New Phytol 204:704–714. doi: 10.1111/nph.12931
24. Govindaraj M, Vetriventhan M, Srinivasan M (2015) Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Genet Res Intl 431487:1–14. doi: 10.1155/2015/431487
25. Griffin WB (1987) Outcrossing in New Zealand wheats measured by occurrence of purple grain. New Zeal J Agr Res 30:287–290. doi: 10.1080/00288233.1987.10421885
26. Hao M, Luo JT, Zhang LQ, Yuan ZW, Yang YW, Wu M et al (2013) Production of hexaploid triticale by a synthetic hexaploid wheat-rye hybrid method. Euphytica 193:347–357. doi: 10.1007/s10681-013-0930-2
27. He J, Zhao X, Laroche A, Lu ZX, Liu HK, Li Z (2014) Genotypingby-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. Front Plant Sci 5:484. doi: 10.3389/fpls.2014.00484
28. Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R et al (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum*/*Aegilops* complex and the evolutionary history of polyploid wheat. Proc Natl Acad Sci USA 99:8133–8138. doi: 10.1073/pnas.072223799
29. Huang YF, Poland JA, Wight CP, Jackson EW, Tinker NA (2014) Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. PLoS ONE 9:e102448. doi: 10.1371/journal.pone.0102448
30. Hucl P (1996) Out-crossing rates for 10 Canadian spring wheat cultivars. Can J Plant Sci 76:423–427. doi: https://doi.org/10.4141/cjps96-075
31. Hyun DY, Sebastin R, Lee KJ, Lee GA, Shin MJ, Kim SH et al (2020) Genotyping-by-sequencing derived single nucleotide polymorphisms provide the first well-resolved phylogeny for the genus *Triticum* (*Poaceae*). Front. Plant Sci 11:688. doi: 10.3389/fpls.2020.00688
32. Jakobashvili ZA (1989) Determination of phylogenetic relationships between wheat species by the analysis of polymorphism and inheritance of storage proteins (in Russian). PhD thesis, Institute of General Genetics, the USSR Academy of Sciences
33. Jakubziner MM, Priroda (1933) (8–9):98–102(in Russian).
34. Jakubziner MM (1950) New data on the wild wheat *Triticum araraticum* Jakubc. Bull Appl Bot Genet Pl-Breed 28:199–200 (in Russian)
35. JÓ“rve K, Jakobson I, Enno T (2002) Tetraploid wheat species *Triticum timopheevi* and *Triticum militinae* in common wheat improvement. Acta Agron Hung 50:463–477. doi: 10.1556/AAgr.50.2002.4.9
36. Keydel F (1978) Floral biology and its influence on varietal maintenance and multiplication in autogamous cereals. Saatgutwirtschaft 30:246–250
37. Kilian B, Özkan H, Deusch O, Effgen S, Brandolini A, Kohl J et al (2007) Independent wheat B and G genome origins in outcrossing *Aegilops* progenitor haplotypes. Mol Biol Evol 24:217–227. doi: 10.1093/molbev/msl151
38. Kilian B, Özkan H, Pozzi C, Salamini F (2009) Domestication of the *Triticeae* in the Fertile Crescent. In: Muehlbauer GJ, Feuillet C (eds) Genetics and genomics of the *Triticeae*. Springer, New York, pp 81–119. doi: 10.1007/978-0-387-77489-3_3

39. Kumar S, Banks TW, Cloutier S (2012) SNP discovery through next generation sequencing and its applications. Int J Plant Genomics 2012: 831460. doi: 10.1155/2012/831460
40. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760. doi: 10.1093/bioinformatics/btp324
41. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al (2009) The sequence alignment/map format and SAM tools. Bioinformatics 25(16):2078–2079. doi: 10.1093/bioinformatics/btp352
42. Li R, Li CQ, Chen HY, Liu XH, Xiao H, Chen SY (2019) Genomic diversity and admixture patterns among six Chinese indigenous cattle breeds in Yunnan. Asian-Australas J Anim Sci 32(8):1069–1076. doi: 10.5713/ajas.18.0605
43. Lilienfeld FA, Kihara H (1951) Genome-analysis in *Triticum* and *Aegilops*. X. Concluding review. Cytologia 16:101–123
44. Liu XJ, Zhang MH, Liu X, Li HY, Hao M, Ning SZ et al (2018) Molecular cytogenetic identification of newly synthetic *Triticum kiharae* with high resistance to stripe rust. Genet Resour Crop Evol 65(6):1725–1732. doi: 10.1007/s10722-018-0649-y
45. Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129. doi: 10.1093/bioinformatics/bti282
46. Luo Z, Brock J, Dyer JM, Kutchan T, Schachtman D, Augustin M, Ge Y, Fahlgren N et al (2019) Genetic diversity and population structure of a *Camelina sativa* spring panel. Front Plant Sci 10:184. doi: 10.3389/fpls.2019.00184
47. Martin TJ (1990) Outcrossing in 12 hard red winter-wheat cultivars. Crop Sci 30:59–62. doi: 10.2135/cropsci1990.0011183X003000010013x
48. McIntosh RA, Gyárfás J (1971) *Triticum timopheevi* as a source of resistance to wheat stem rust. Z Pflflanzenzüchtg 66:240–248
49. Menabde VL, Ericzjan AA (1960) To the investigation of Georgian wheat Zanduri. Proc Georg Acad Sci 25:731–732 [in Georgian]
50. Mikó P, Megyeri M, Farkas A, Molnár I, Molnár-Láng M (2015) Molecular cytogenetic identification and phenotypic description of a new synthetic amphiploid, *Triticum timococcum* ($A^tA^tGGA^mA^m$). Genet Resour Crop Evol 62:55–66. doi: 10.1007/s10722-014-0135-0
51. Mori N, Miyashita NT, Terachi T, Nakamura C (1997) Variation in cox*II* intron in the wild ancestral species of wheat. Hereditas 126:281–288. doi: 10.1111/j.1601-5223.1997.00281.x
52. Mori N, Kondo Y, Ishii T, Kawahara T, Valkoun J, Nakamura C (2009) Genetic diversity and origin of timopheevi wheat inferred by chloroplast DNA fingerprinting. Breed Sci 59:571–578. doi: https://doi.org/10.1270/jsbbs.59.571
53. Nesbitt M, Samuel D (1996) From staple crop to extinction? The archaeology and history of the hulled wheats. In: Padulosi S, Hammer K, Heller J. (eds) Hulled wheats. Proceedings of the first international workshop on hulled wheats. Castelvecchio Pascoli, Tuscany Italy pp 41–100
54. Okada T, Jayasinghe J, Nansamba M, Baes M, Warner P, Kouidri et al (2018) Unfertilized ovary pushes wheat flower open for cross-pollination. J Exp Bot 3:399–412. doi: 10.1093/jxb/erx410
55. Oliveira HR, Jacocks L, Czajkowska BI, Kennedy SL, Brown TA (2020) Multiregional origins of the domesticated tetraploid wheats. PLoS ONE 15:e0227148. doi: org/10.1371/journal.pone.0227148

56. Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in excel, Population genetic software for teaching and research. Mol Ecol Notes 6:288–295
57. Peterson GW, Dong Y, Horbach C, Fu YB (2014) Genotyping-by sequencing for plant genetic diversity analysis: a lab guide for SNP genotyping. Diversity 6:665–680. doi: 10.3390/d6040665
58. Peusha HO, Enno TM, Priilinn O (1996) Genetic analysis of disease resistance in wheat hybrids, derivatives of *Triticum timopheevi* and *Triticum militinae*. Acta Agron Hung 44:237–244
59. Poland JA, Brown PJ, Sorrells ME, Jannink J (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE 7:e32253. doi: 10.1371/journal.pone.0032253
60. Rodríguez S, Maestra M, Perera E, Díez M, Naranjo T (2000) Pairing affinities of the B- and G-genome chromosomes of polyploid wheats with those of *Aegilops speltoides*. Genome 43:814–819. doi: org/10.1139/g00-055
61. Rohlf FJ (2000) NTSYS-pc: Numerical taxonomy and multivariate analysis system version 2.1. Exeter Software, Setauket, NY
62. Salem KFM, Sallam A (2016) Analysis of population structure and genetic diversity of Egyptian and exotic rice (*Oryza sativa* L.) genotypes. C R Biol 339:1–9. doi: 10.1016/j.crvi.2015.11.003
63. Salina EA, Leonova IN, Efremova TT, Röder MS (2006) Wheat genome structure: translocations during the course of polyploidization. Funct Integr Genomics 6:71–80. doi: org/10.1007/s10142-005-0001-4
64. Takahashi H, Rai B, Kato K, Nakamura I (2010) Divergent evolution of wild and cultivated subspecies of *Triticum timopheevi* as revealed by the study of *PolA1* gene. Genet Resour Crop Evol 57:101–109. doi: 10.1007/s10722-009-9454-y
65. Timonova EM, Leonova IN, Röder MS, Salina EA (2013) Marker-assisted development and characterization of a set of *Triticum aestivum* lines carrying different introgressions from the *T. timopheevii* genome. Mol Breed 31:123–136. doi: org/10.1007/s11032-012-9776-x
66. Ulaş B, Fiorentino G (2021) Recent attestations of “new” glume wheat in Turkey: a reassessment of its role in the reconstruction of Neolithic agriculture. Veg Hist and Archaeobot 30:685–701. doi: org/10.1007/s00334-020-00807-w
67. Wagenaar EB (1966) Studies on the genome constitution of *Triticum timopheevi* Zhuk., II. The *timopheevi* complex and its origin. Evolution 20:150–164. doi: 10.1111/j.1558-5646.1966.tb03351.x
68. Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating author (s): sewall wright reviewed work (s): published by: society for the study of evolution stable. Evolution 19:395–420. doi: 10.1111/j.1558-5646.1965.tb01731.x
69. Yang X, Tan B, Liu H, Zhu W, Xu L et al (2020) Genetic diversity and population structure of Asian and European common wheat accessions based on genotyping-by-sequencing. Front Genet 11:1–14. doi: 10.3389/fgene.2020.580782
70. Zhang LQ, Yen Y, Zheng YL, Liu DC (2007) Meiotic restriction in emmer wheat is controlled by one or more nuclear genes that continue to function in derived lines. Sex Plant Reprod 20:159–166. doi: 10.1007/s00497-007-0052-x
71. Zhukovsky PM (1924) Wild emmer in Georgia-*Triticum dicoccum* Schrank *dicoccoides* Koern. Notes Appl-Sci. Dept Tiflflis Bot Garden 3:1–3 (in Russian)

72. Zohary D, Hopf M, Weiss E (2012) Domestication of plants in the Old World, 4th edn. Oxford University Press, Oxford

Figures

Figure 1

Spike type of *T. timopheevii* and *T. araraticum*. *T. timopheevii* accessions PI94761 (1), PI266850(2), PI119442(3), PI282932(4), PI272530(5), PI221421(6) and *T. araraticum* accessions PI427410(7), PI427414(8), PI427421(9), PI427413(10), PI427415(11) and AS273 (12).

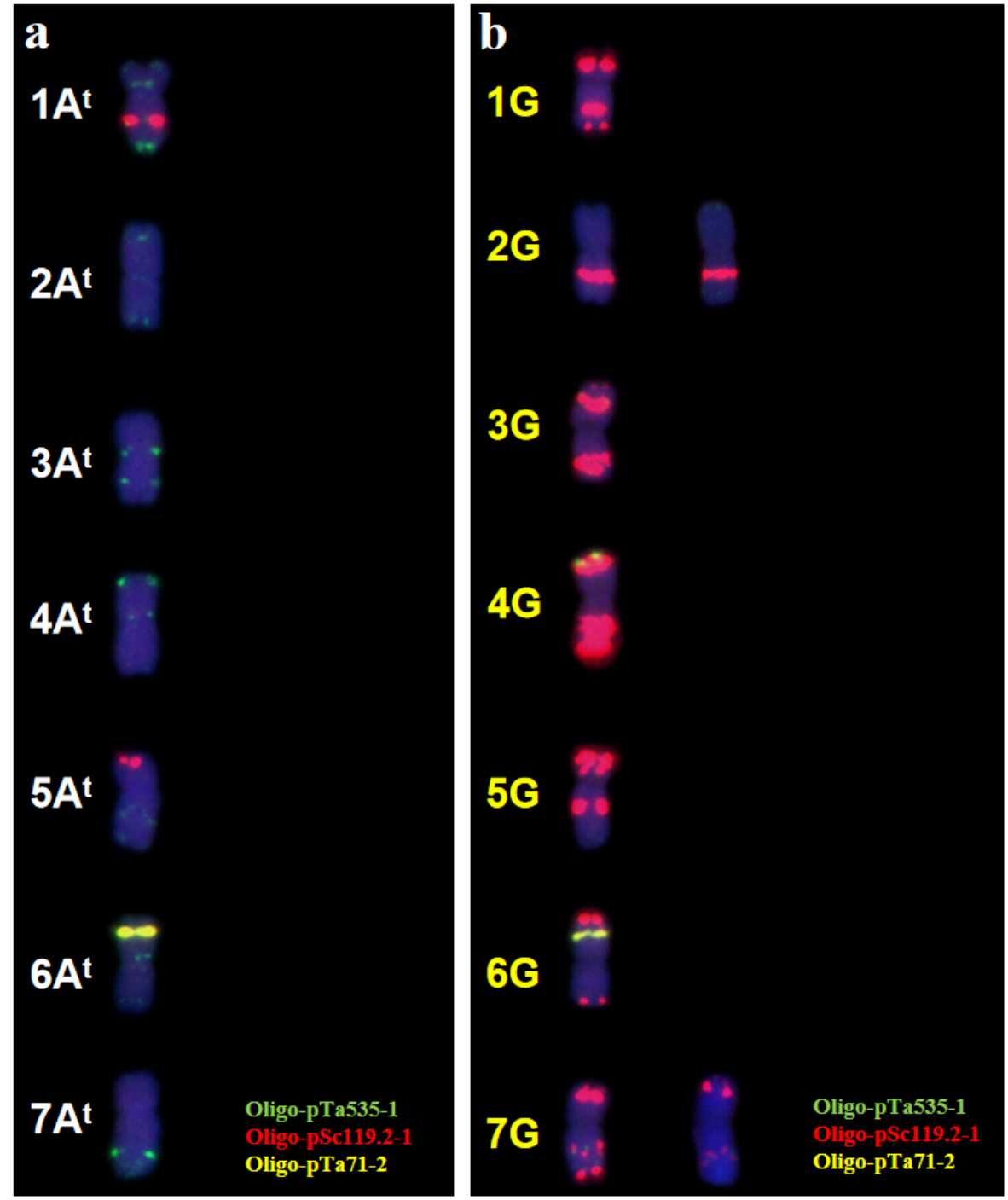

Figure 2

FISH karyotype diversity of the A^t (**a**) and G (**b**) genomes in *T. timopheevii* using Oligo-pSc119.2-1 (red), Oligo-pTa71-2 (yellow) and Oligo-pTa535-1 (green) as probes.

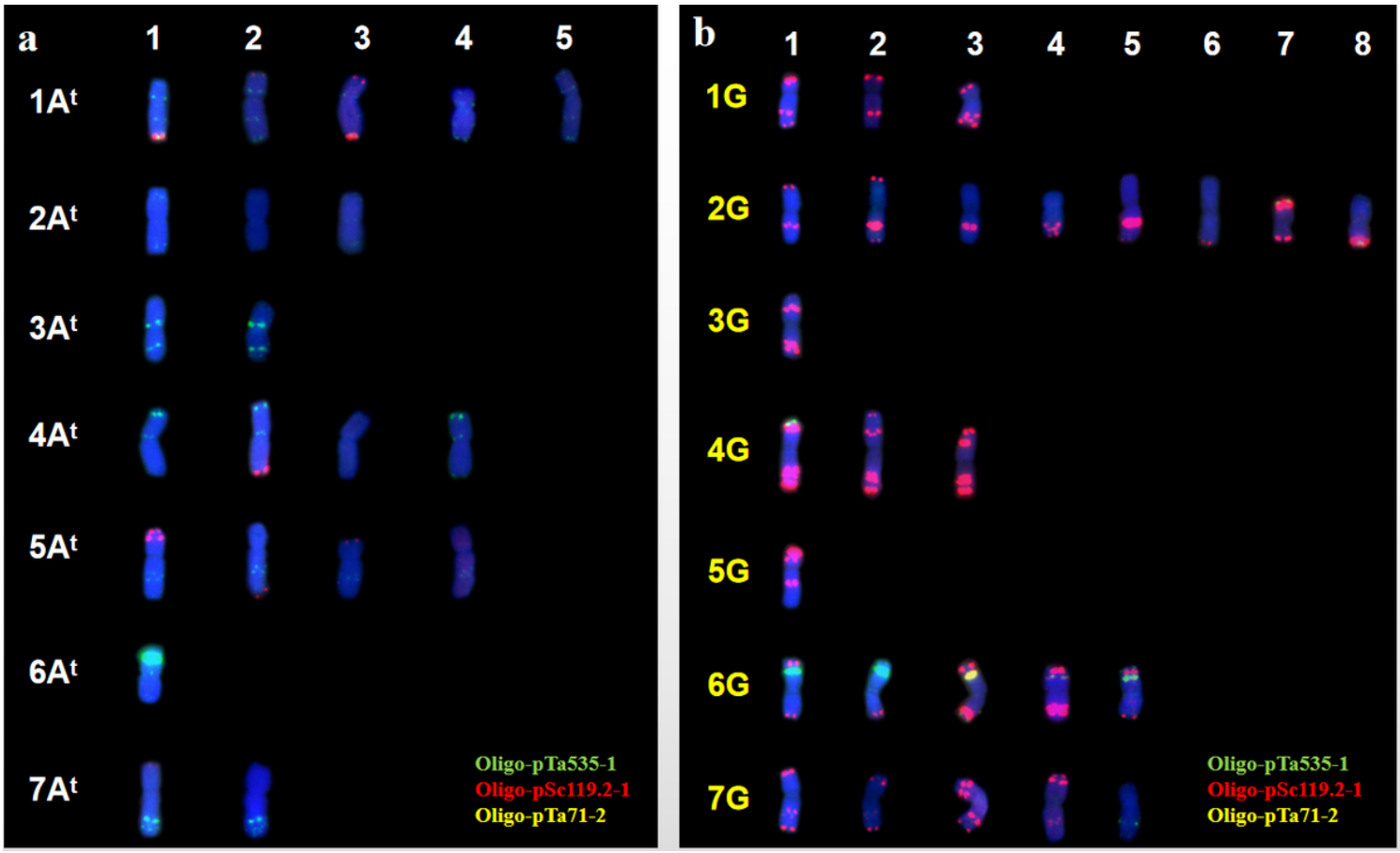

Figure 3

FISH karyotype diversity of the A^t (**a**) and G (**b**) genomes in *T. araraticum* using Oligo-pSc119.2-1 (red), Oligo-pTa71-2 (yellow) and Oligo-pTa535-1 (green) as probes.

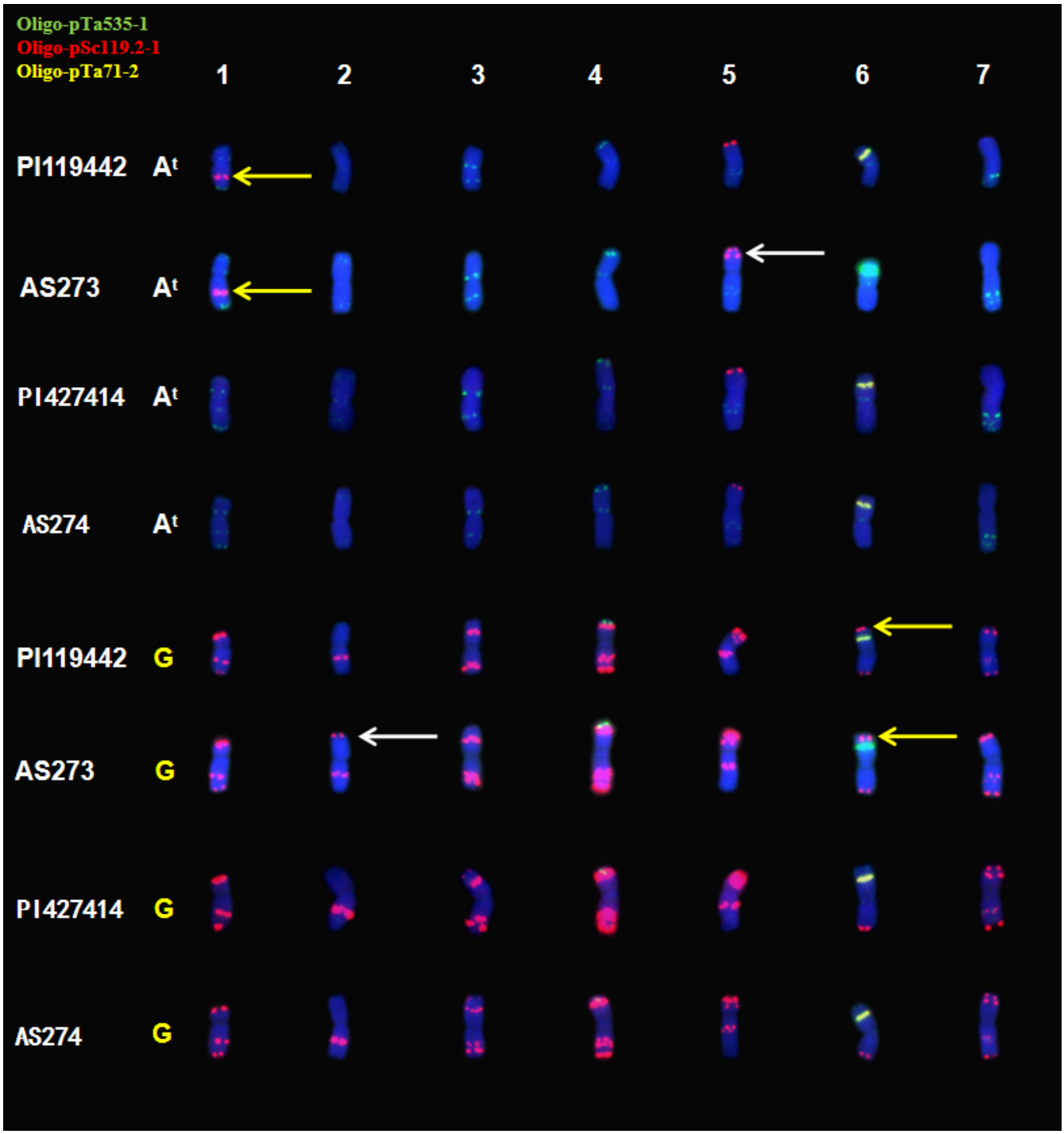

Figure 4

Comparisons of FISH signals between *T. timopheevii* PI119442 and *T. araraticum* AS273, AS274, and PI427414 using Oligo-pSc119.2-1 (red), Oligo-pTa71-2 (yellow) and Oligo-pTa535-1 (green) as probes. An extra pair of red signals was indicated by the white arrow in AS273 compared with *T. timopheevii* and *T. araraticum*, and a pair of the same red signals with *T. timopheevii* was indicated by the yellow arrow in AS273.

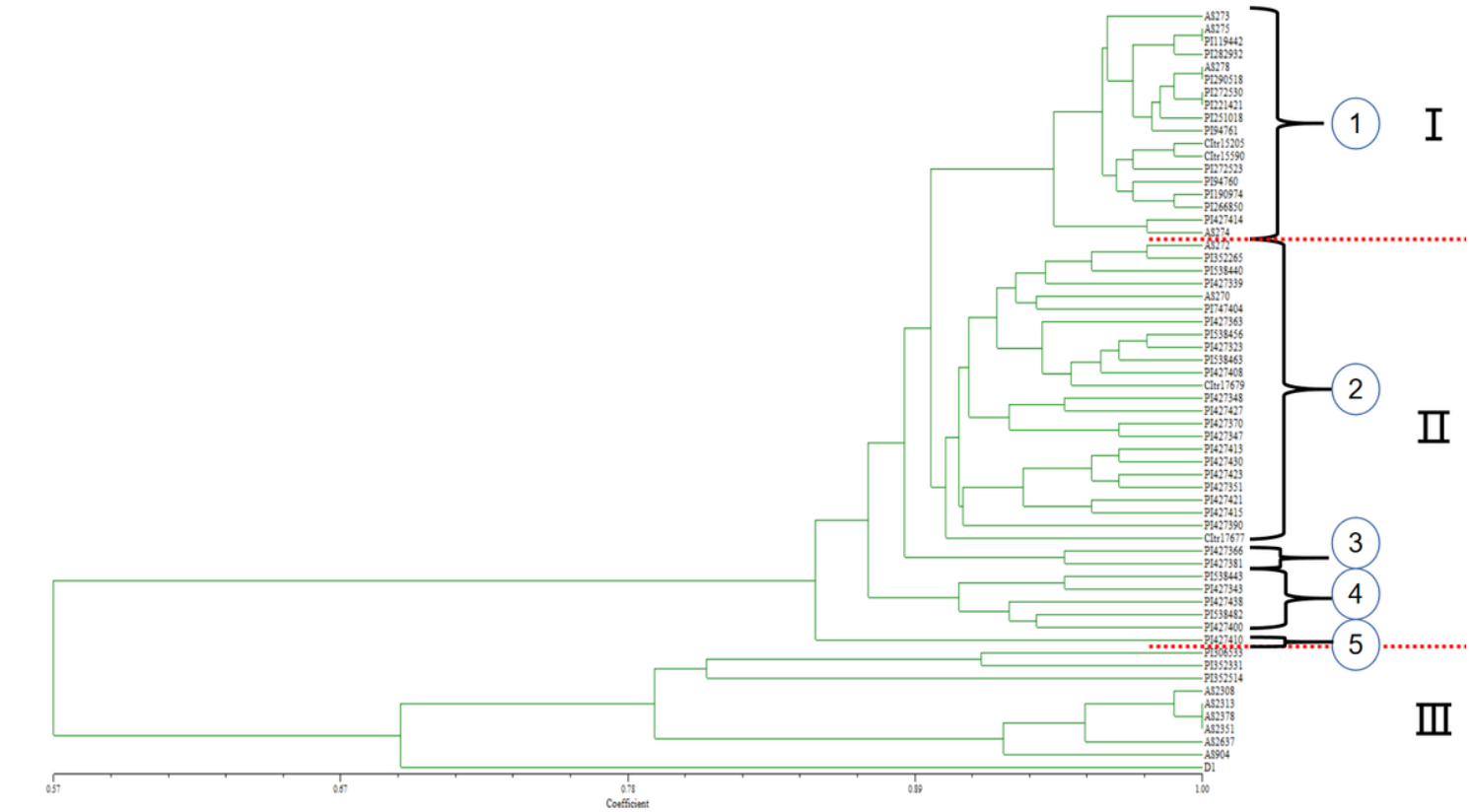

Figure 5

Genetic association among 60 tetraploid wheat accessions established according to the signal types of different chromosomes. Red, oligo-pSc119.2-1; yellow, oligo-pTa71-2; green, oligo-pTa535-1. I, II, and III represent the populations of *T. timopheevii*, *T. araraticum*, and *T. turgidum*, respectively.

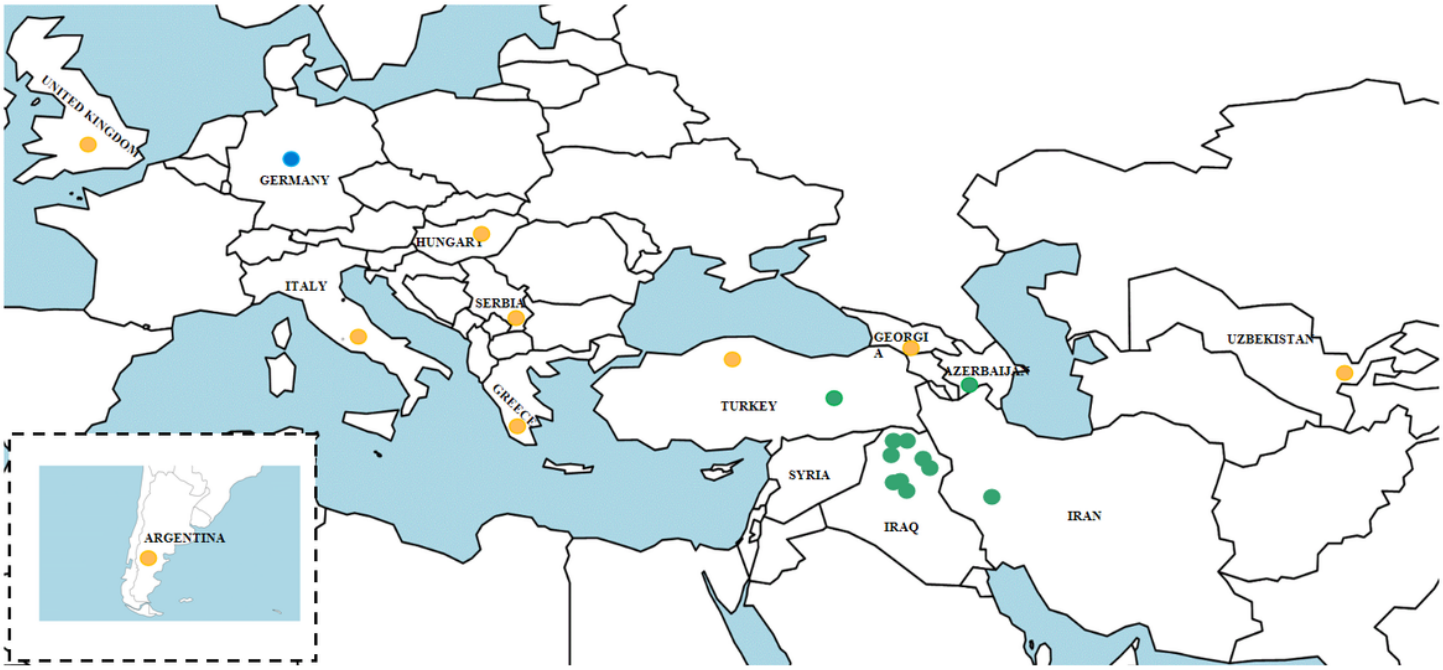

Figure 6

Geographical distributions of plastotypes observed in 32 and 13 accessions of wild and domesticated *T. timopheevii*, respectively. Species assignments from genetic analyses are indicated in different colours: *T. araraticum* in green, *T. timopheevii* in yellow, and AS273 in blue.

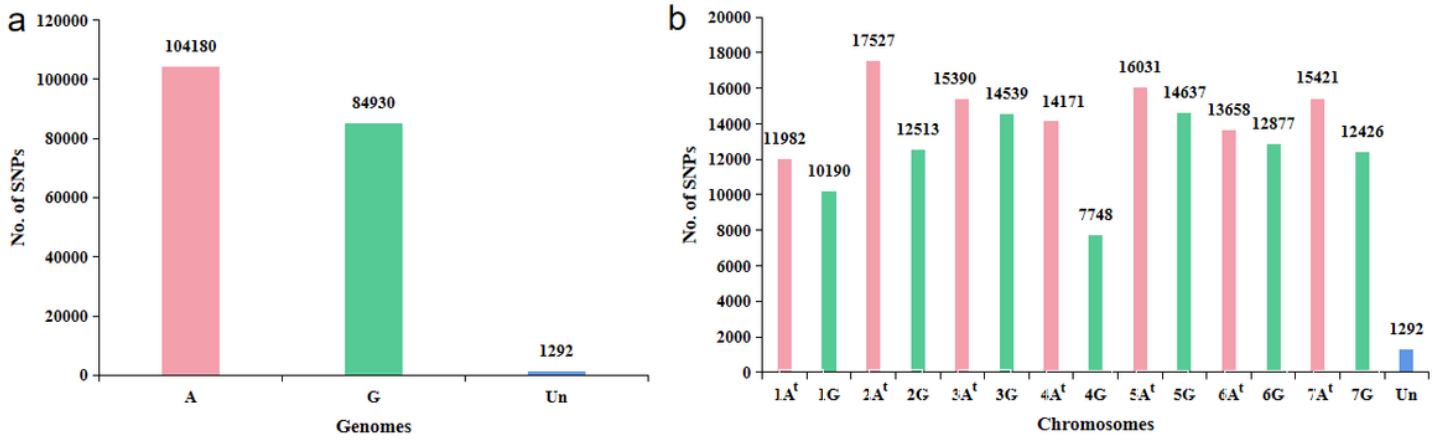

Figure 7

(a, b) SNP distribution on all chromosomes in the *timopheevii* wheat genomes A^t and G.

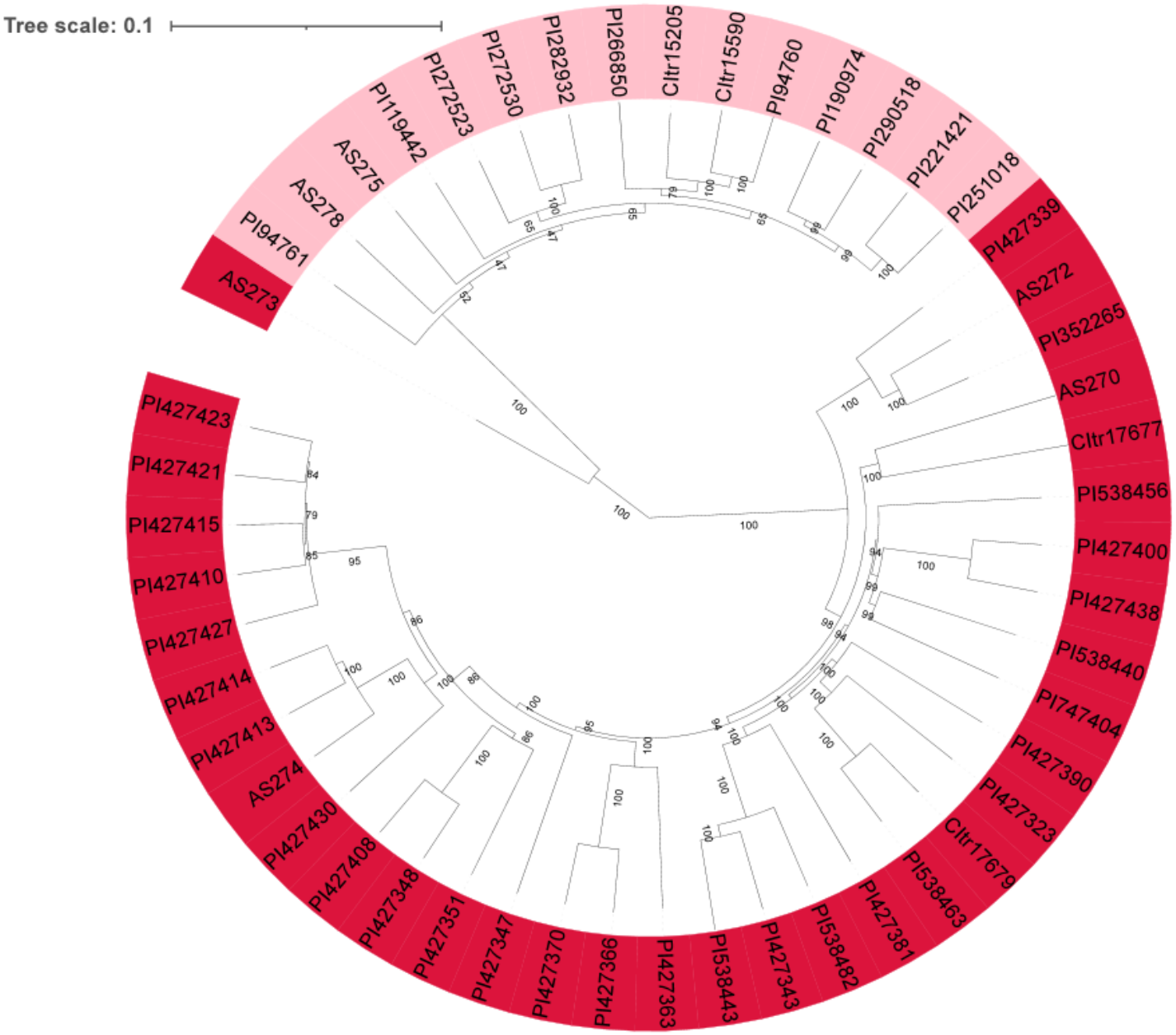

Figure 8

Neighbour-joining tree based on SNP markers of 35 *T. araraticum* accessions and 15 *T. timopheevii* accessions. *T. araraticum* is marked in dark red, while *T. timopheevii* is marked in light red.

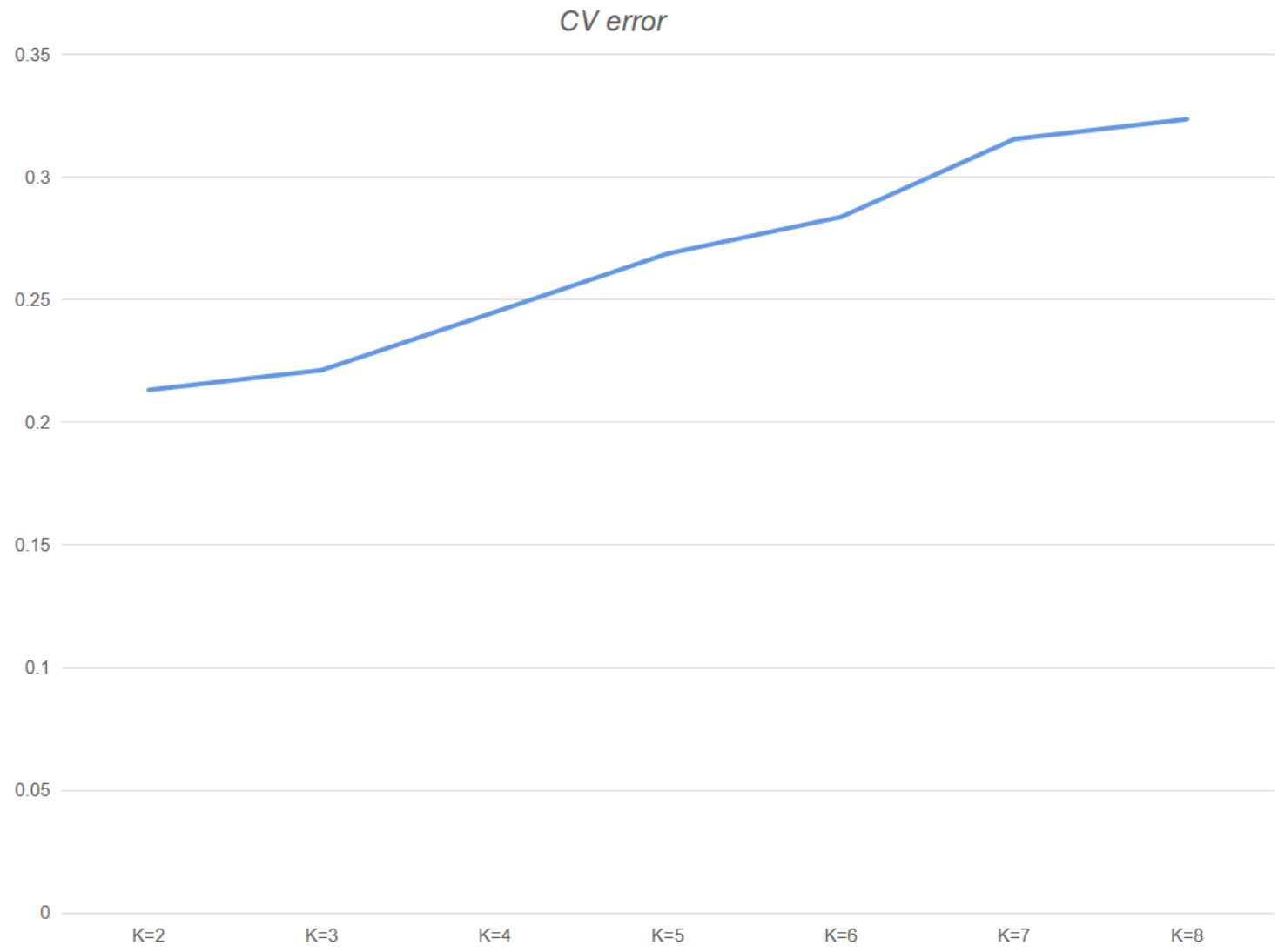

Figure 9

The CV errors vary among K values.

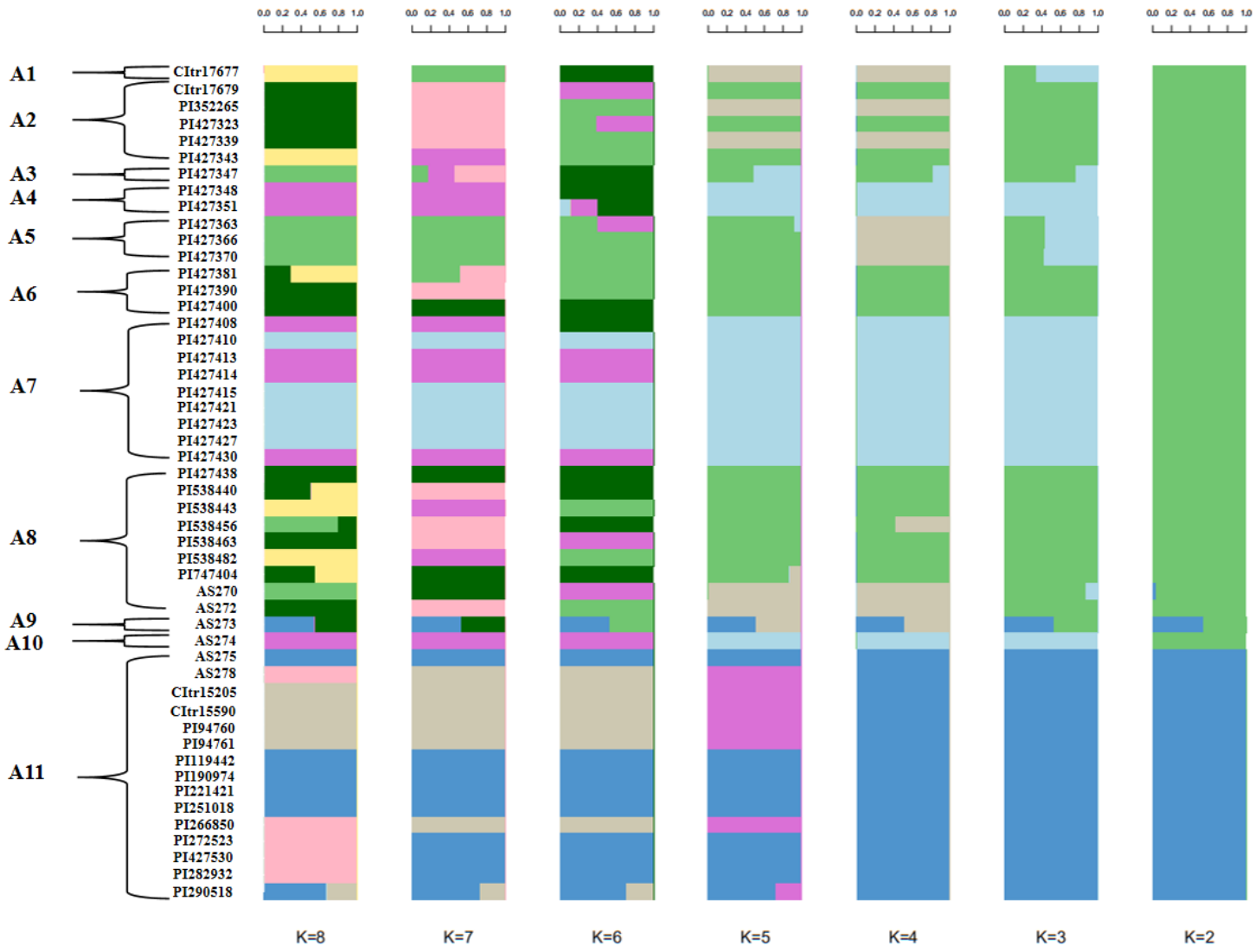

Figure 10

Analysis of population genetic structure according to the SNP markers of 50 germplasms (K = 1 to 8). Each individual is represented by a vertical bar, and the likelihood of it belonging to a particular population is indicated by the colors. When K=3, the population can be divided into the A1-A11 taxa.

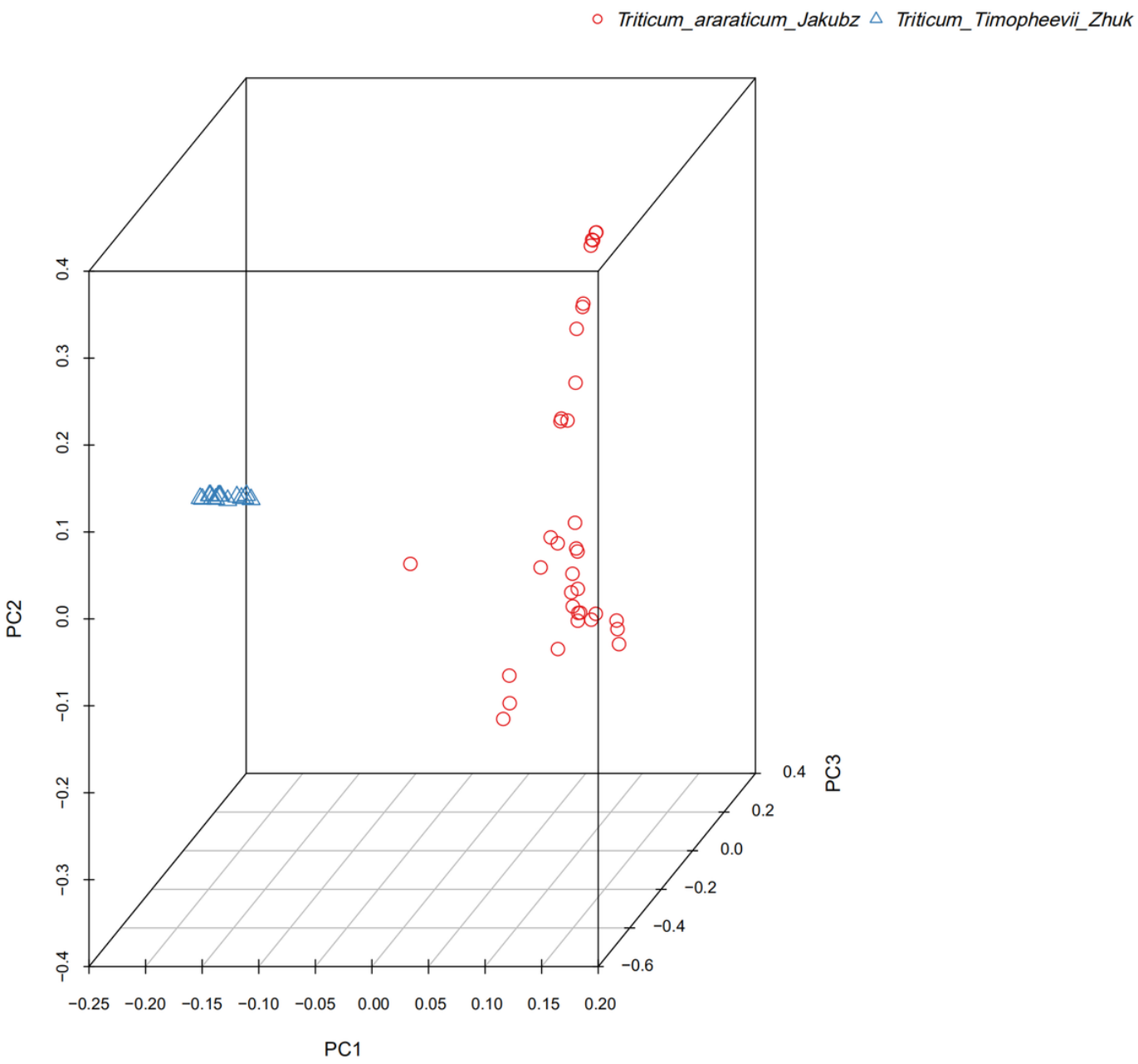

Figure 11

Principal coordinate analysis (PCA) based on SNP markers of 35 *T. araraticum* accessions and 15 *T. timopheevii* accessions.

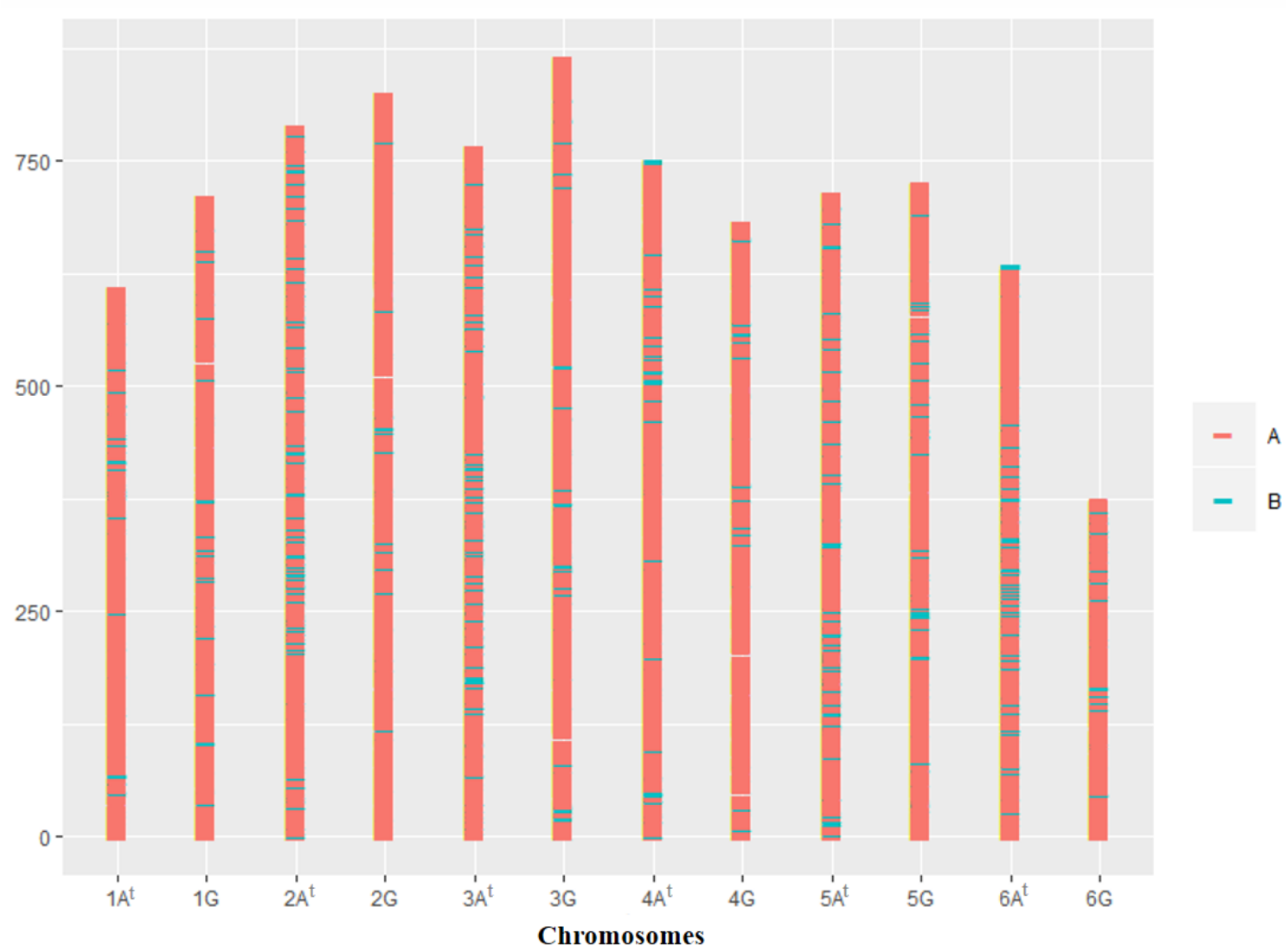

Figure 12

Chromosomal distribution of fixed nucleotide differences between *T. araraticum* and *T. timopheevii*. A represents the same SNP markers at the same physical location, and B represents different SNP markers at the same physical location.

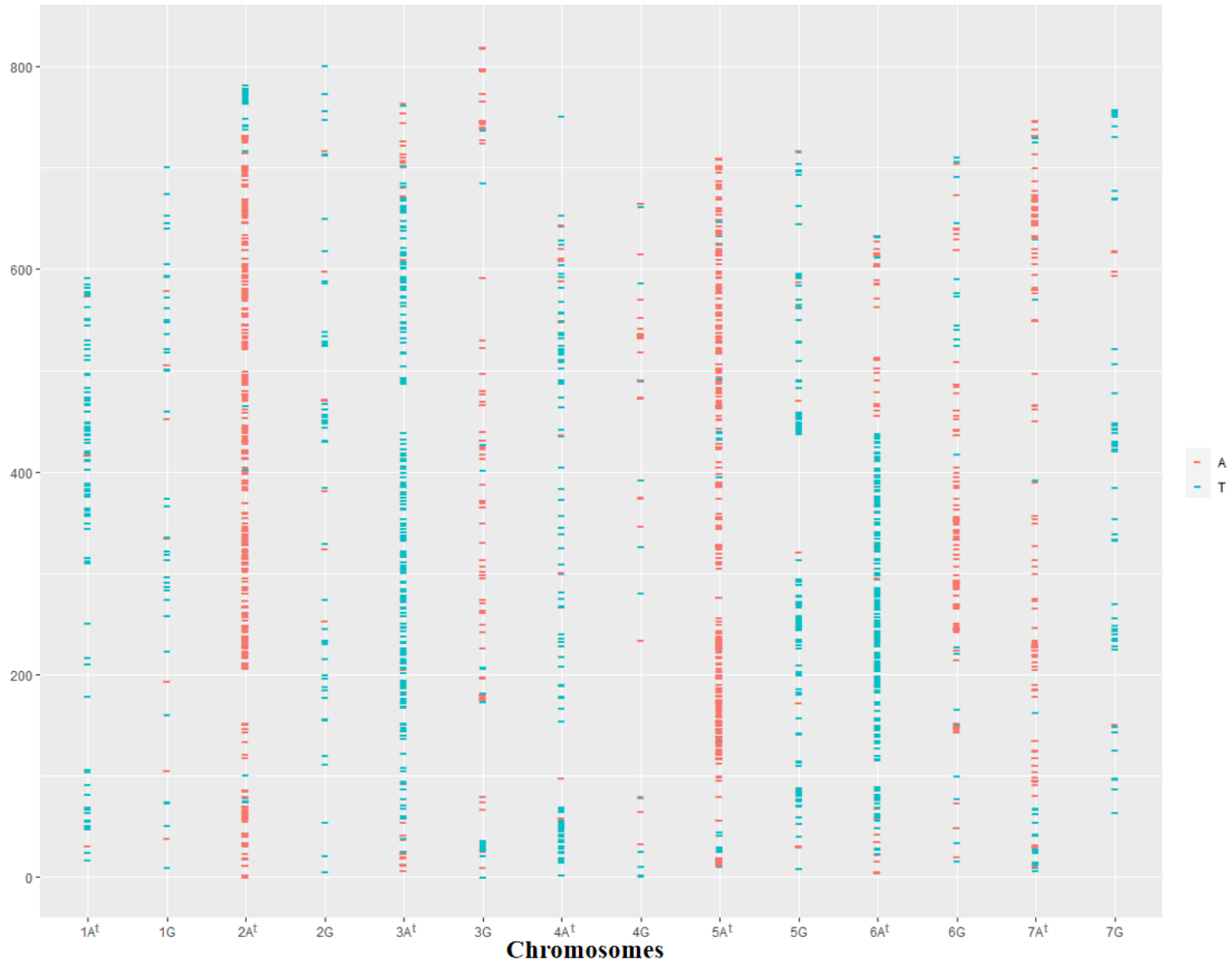

Figure 13

Different SNPs between *T. araraticum* and *T. timopheevii* were distributed on 14 chromosomes of AS273. A indicates the SNPs markers from *T. araraticum*. T indicates the SNPs markers from *T. timopheevii*.

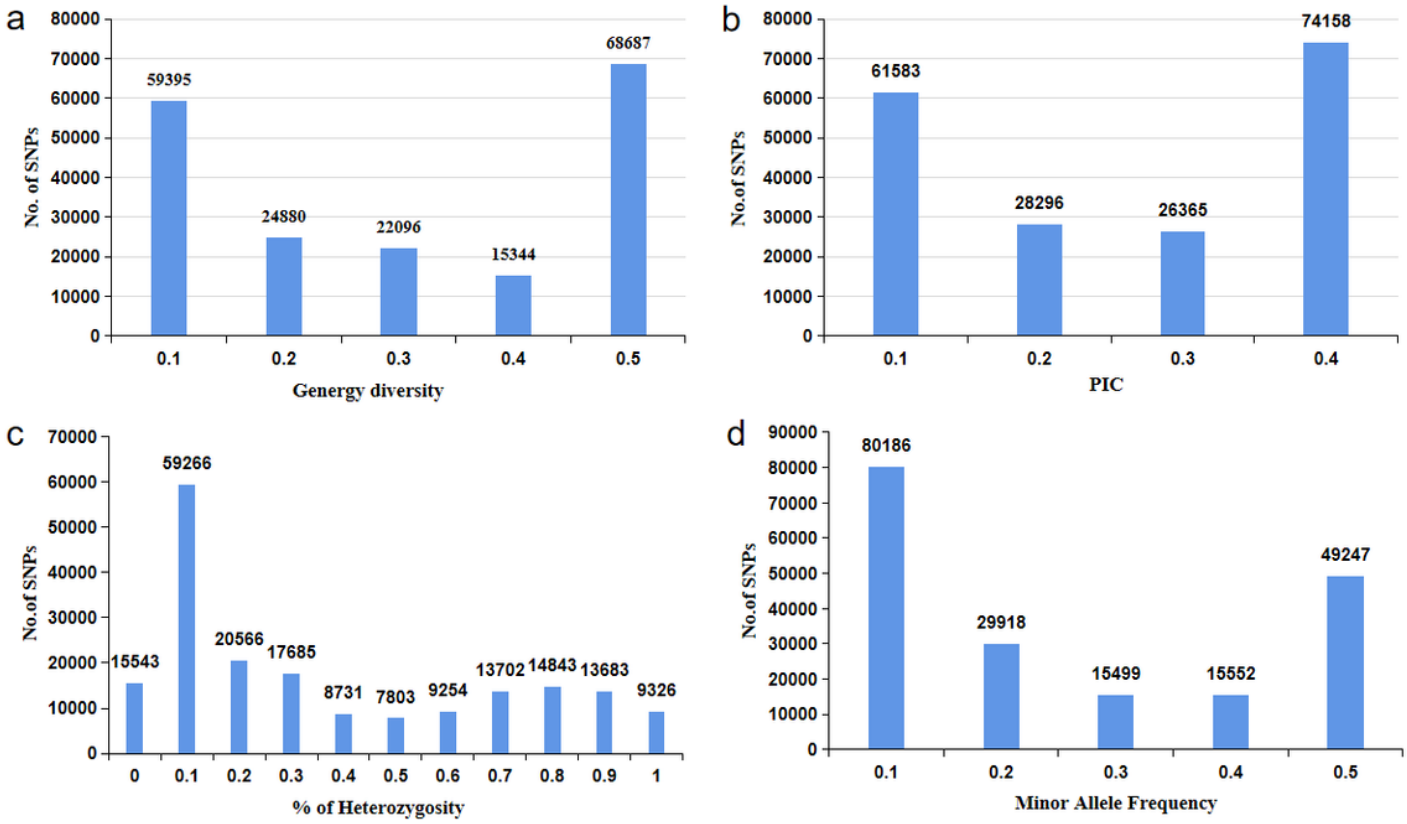

Figure 14

Distribution of the genetic diversity indices of 190,402 SNPs in the 50 *timopheevii* wheat accessions.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.xlsx
- TableS2.xlsx
- TableS3.xlsx
- TableS5.xlsx
- TableS6.xlsx