Genetic Diversity in Wild Diploid Wheat *T. urartu* Revealed by SSR Markers

M. AbbAsov

1Genetic Resources Institute of ANAS, Baku, AZ1106, Azerbaijan
2Baku State University, Baku, Azerbaijan

(Received 22 November 2017; Accepted 25 April 2018; Communicated by I. Molnár)

Genetic diversity of 74 *T. urartu* genotypes was studied using 11 SSR (simple sequence repeats) markers. The number of alleles ranged from 4 to 15, with an average of 8 alleles per primer. The mean values for the expected heterozygosity (He) and polymorphism information content (PIC) over all loci and populations were 0.56 and 0.52, respectively. From a geographic viewpoint the higher diversities were observed in Jordan, followed by Syria and Turkey. Diversity revealed within countries was higher than among them, even in the same regions of the relevant countries. The analysis of molecular variance revealed that most of the genetic variability was accounted for by differences within populations (90%), with less variability among them (10%). The dendrogram generated based on Nei’s dissimilarity matrix revealed three main clusters for which the grouping patterns were not clearly associated with the geographic origins, indicating the gene flow among different countries. Principal coordinate analysis (PCoA) confirmed subgrouping obtained by cluster analysis. In general, genetic distances between geographic regions were low or moderate, which was also supported by low values of pairwise Fst. Our findings can direct the sampling strategies on *T. urartu* in studied regions to find beneficial alleles. The heterotic groups detected by cluster and PCoA analysis in the present study can serve as effective candidates in crossing programs to broaden the genetic base in *T. urartu*.

**Keywords:** wheat, *T. urartu*, SSR, PIC, genetic diversity

**Introduction**

Wheat is the most widely grown crop in the world, demand for which is expected to increase by 60% in 2050. So, there is an endless need to increase the productivity and tolerance to abiotic and biotic factors to develop more sustainable wheat production. This improvement program should cover all wheat types (bread wheat, durum wheat, wild relatives) and all levels of variations from morphology to genomics. The use of wild relatives of wheat is also of strategic importance, as they possess rich allelic diversity for numerous important traits, as well as stress tolerance (Rouse and Jin 2011; Guzmán and Alvarez 2012).

*Corresponding author; E-mail: mehraj_genetic@yahoo.com*
Triticum urartu is a wild diploid wheat (2n = 2x = 14) first described in 1938 as a rare endemic species of Transcaucasia (Jakubizner 1959). Later some populations were also found in Turkey, Lebanon and Iraq (Johnson 1975).

The genus Triticum L. includes the genomes A\textsuperscript{u}, A\textsuperscript{b}, B, D, G. Of these, the only actual wheat genomes are A\textsuperscript{u} and A\textsuperscript{b}. Triticum urartu is considered to be the A\textsuperscript{u} genome donor for all polyploidy wheats (Dvorak et al. 1993). Several investigations supported that A genome carries important genes responsible for adult plant resistance (Helguera et al. 2000), flour color (Parker and Langridge 2000), sprouting resistance (Zanetti et al. 2000), cold tolerance (Vagujfalvi et al. 2000), heading date and plant height (Yuanwu and Yujun 1995; Ehtemam et al. 2010). Thus, significant importance should be given to T. urartu both as a rich source of important genes for broadening the genetic variation in cultivated wheats and for establishing the phylogeny of the genus Triticum. However, genetic diversity of T. urartu has not been well investigated as compared with other diploids (Wang et al. 2017).

Nowadays, DNA markers provide a good estimate of genetic diversity since they are not affected by the environment and can cover whole genome (Naghavi et al. 2009; Izzatullayeva et al. 2014; Hajiyev et al. 2015). Limited number of studies was carried out in T. urartu using various DNA markers such as restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and simple sequence repeats (SSR) (Vierling and Nguyen 1992; Dhaliwal et al. 1993; Moghaddam et al. 2000; Aliyev et al. 2007; Baum and Bailey 2013; Wang et al. 2017). Simple sequence repeats are allele-specific and co-dominant molecular markers, widely applied in cultivar identity verification, diversity studies and linkage map construction in wheat (Varshney et al. 2001; Huang et al. 2002).

The objective of the current study was the assessment of the genetic variation among and within different populations and accessions of T. urartu collected from different localities and countries using SSR markers.

**Materials and Methods**

**Plant material**

Seventy four accessions of T. urartu were used as a research material. Of 74 accessions 31 were from Syria (SYR), 18 from Turkey (TUR), 10 from Lebanon (LBN), 10 from Jordan (JOR), 2 from Armenia (ARM). Three countries – Iran (IRN), Iraq (IRG) and Saudi Arabia (SAR) were presented by one genotype each. Accession number, countries of origin and collection sites were given in Table S1*.

Genomic DNA isolation was conducted following a CTAB protocol. A set of 11 SSR markers mapped to A and B genomes (Table 1) were selected for genotyping. PCR conditions for fluorescent-dye labeled primers were as follows: initial denaturation at 95 °C for 3 min; 40 cycles of denaturation at 95 °C for 1 min, annealing at 50 °C for 1 min and elongation at 72 °C for 2 min; final elongation at 72 °C for 10 min. DNA fragments were

*Further details about the Electronic Supplementary Material (ESM) can be found at the end of the article.
Table 1. Summary statistics of the 11 SSR markers used in the study

<table>
<thead>
<tr>
<th>Marker</th>
<th>Forward 5' 3'</th>
<th>Reverse 5' 3'</th>
<th>Chromosome location</th>
<th>Allele No</th>
<th>Major Allele Freq.</th>
<th>Ho</th>
<th>He</th>
<th>PIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>barc-213</td>
<td>GCGTAGATTTCGC</td>
<td>CCGTCCCTCC</td>
<td>1AL</td>
<td>15</td>
<td>0.19</td>
<td>0.01</td>
<td>0.88</td>
<td>0.87</td>
</tr>
<tr>
<td>barc-15</td>
<td>ATGCAAAGGC CGGGTTATC</td>
<td>CACCTCTAGCCTA CGCCAACATT</td>
<td>2AL</td>
<td>8</td>
<td>0.60</td>
<td>0.24</td>
<td>0.60</td>
<td>0.56</td>
</tr>
<tr>
<td>barc-1021</td>
<td>GGAAGGACCTGA CTGAC'TGCA TCTG</td>
<td>GCGATCACAACCAA AT'TCTT TTAACTA</td>
<td>3AL</td>
<td>6</td>
<td>0.81</td>
<td>0.08</td>
<td>0.32</td>
<td>0.30</td>
</tr>
<tr>
<td>barc-206</td>
<td>GCTTTGCCAGG TGAGC'ACTCT</td>
<td>TGGCGGGGTATTT GAGTTGGGATTT</td>
<td>4AS</td>
<td>7</td>
<td>0.49</td>
<td>0.61</td>
<td>0.67</td>
<td>0.62</td>
</tr>
<tr>
<td>barc-117</td>
<td>TCA'GTGCGTGT AA'TGTGCTAA</td>
<td>GAGGGCAGGA AAAAAATGACT</td>
<td>5AS</td>
<td>6</td>
<td>0.53</td>
<td>0.59</td>
<td>0.60</td>
<td>0.52</td>
</tr>
<tr>
<td>barc-174</td>
<td>TGGCA'TTTTCT AGCAACAATACAT</td>
<td>GCGAAGCTGGACCAG CCTTCTATCTGTT</td>
<td>1BL</td>
<td>11</td>
<td>0.41</td>
<td>0.76</td>
<td>0.70</td>
<td>0.65</td>
</tr>
<tr>
<td>barc-200</td>
<td>GCGATATGATTG GAGCTGATTG</td>
<td>GCGATGACGT'AGG ATGCGGAATTTG</td>
<td>2 BS</td>
<td>4</td>
<td>0.79</td>
<td>0.32</td>
<td>0.34</td>
<td>0.29</td>
</tr>
<tr>
<td>barc-101</td>
<td>GCTCCTCTCAG ATACCGAAAG</td>
<td>GCGAATGCTACACA CTATGAGCCAATG</td>
<td>2BL</td>
<td>6</td>
<td>0.62</td>
<td>0.22</td>
<td>0.56</td>
<td>0.52</td>
</tr>
<tr>
<td>barc-142</td>
<td>CCGGTGAGA GGACTAAAA</td>
<td>GGCGTGTCA AT'ATGAGC</td>
<td>5 BL</td>
<td>8</td>
<td>0.48</td>
<td>0.01</td>
<td>0.72</td>
<td>0.69</td>
</tr>
<tr>
<td>gwm-361</td>
<td>GTAACCTTTTG CAAAGGGG</td>
<td>ACAAAAGTGGC AAAAGGAGACA</td>
<td>6BS</td>
<td>8</td>
<td>0.75</td>
<td>0.27</td>
<td>0.43</td>
<td>0.41</td>
</tr>
<tr>
<td>gwm-219</td>
<td>GATGAGCGA CACCTAGCTC</td>
<td>GGGGTCCGA GTCCAACAAC</td>
<td>6B</td>
<td>4</td>
<td>0.77</td>
<td>0.02</td>
<td>0.37</td>
<td>0.33</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>0.58</td>
<td>0.29</td>
<td>0.56</td>
<td>0.52</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>83</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
separated on an ABI 3130xl Genetic Analyzer (Applied Biosystems/Thermo Fisher Scientifics) (Chao et al. 2007). Fragment analysis and allele calling were performed using GeneMapper software v.3.7 (Applied Biosystems).

Data analysis

Measures of genetic diversity, i.e. total number of alleles, expected heterozigosity (He), observed heterozigosity (Ho) and polymorphism information content (PIC) were estimated using PowerMarker v. 3.51 (Liu and Muse 2005). PowerMarker software was also used to calculate allele frequencies, distances based on frequencies and the fixation indices (Fst) among different countries. Cluster analysis, PCoA analysis and Neighbour-joining tree were done using software package DARwin 6.0 (Perrier and Jacquemoud-Collet 2006).

Results

A total of 83 alleles were amplified among 74 T. urartu genotypes using 11 SSR primers. The number of fragments per primer ranged from 4 (barc-200 and barc-209) to 15

Figure 1. Dendrogram showing genetic relationship among 74 T. urartu accessions based on Nei’s genetic distance. Different colors indicate different countries (SYR – black, TUR – green, JOR – red, LBN – violet, ARM – blue, IRN – yellow, IRG – light blue, SAR – dark yellow)
(barc-213), with the average number being 8 (Table 1). Major allele frequency in the current study varied from 0.19 to 0.81, averaging 0.58. Twenty five alleles were found to be rare with a frequency of less than 1%. Observed heterozygosity ranged from 0.01 to 0.76, with the most heterozygous locus to be Xbarc-174. The microsatellite barc-213 had the highest He value (He = 0.88), whereas the lowest value was recorded for primer barc-1021 (He = 0.32). The mean observed (Ho) and expected heterozygosity (He) for the 11 loci were 0.29 and 0.56, respectively. PIC values of each marker locus ranged from 0.29

<table>
<thead>
<tr>
<th>Subsets</th>
<th>Sample size</th>
<th>Allele No</th>
<th>Major al freq.</th>
<th>HO</th>
<th>HE</th>
<th>PIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYR</td>
<td>31</td>
<td>5.6</td>
<td>0.58</td>
<td>0.28</td>
<td>0.56</td>
<td>0.51</td>
</tr>
<tr>
<td>TUR</td>
<td>18</td>
<td>4.5</td>
<td>0.64</td>
<td>0.29</td>
<td>0.51</td>
<td>0.47</td>
</tr>
<tr>
<td>LBN</td>
<td>10</td>
<td>3.3</td>
<td>0.65</td>
<td>0.34</td>
<td>0.45</td>
<td>0.40</td>
</tr>
<tr>
<td>JOR</td>
<td>10</td>
<td>4.0</td>
<td>0.57</td>
<td>0.28</td>
<td>0.58</td>
<td>0.53</td>
</tr>
<tr>
<td>ARM</td>
<td>2</td>
<td>1.5</td>
<td>0.80</td>
<td>0.32</td>
<td>0.24</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Figure 2. Scatter plot of the first two principal coordinate analysis axes of 74 T. urartu accessions generated from genetic distance calculations using 11 SSR markers

Cereal Research Communications 46, 2018
Out of 11 primer pairs, 7 were highly informative with PIC ≥ 0.5 and the remaining 4 were moderately informative with 0.25 < PIC < 0.50 (Table 1). The highest genetic variation was found in *T. urartu* accessions of Jordan origin (He = 0.58; PIC = 0.53) (Table 2).

Pair-wise dissimilarity (GD) among 74 accessions ranged from 0 to 1. The smallest genetic distance was noted between genotype pairs 45476 and 110835 of Lebanon origin and between 45484 (Syria) and 45472 (Lebanon). The dendrogram based on Nei’s dissimilarity matrix revealed three main clusters with a different number of subclusters in each (Fig. 1). Cluster I was further divided into two major subclusters with quite large inter subcluster distance (GD ≥ 0.6), indicating wider genetic divergence among them, while smaller distances were noted among groups of two other clusters. PCoA analysis revealed two main distinct groups for the entire population and confirmed subgrouping obtained by cluster analysis (Fig. 2). The first two coordinates explained 21% and 9% of the total variation, followed by 6.7%, 5.4% and 4.6% in third, fourth and fifth coordinates,

*Figure 3. Dendrogram of T. urartu accessions from 8 countries based on Nei’s genetic distance*
respectively. No significant association was found between the grouping of accessions and geographic origins based on both cluster and PCoA analyses.

The NJ tree was constructed based on Nei’s genetic distances that visually demonstrated the relationship among the eight geographic regions (Fig. 3). The genotypes from Turkey were very similar to Syrian genotypes with the lowest genetic distance (GD = 0.113), while the highest distance was noted between Turkey and Saudi Arabia (GD = 0.491) (Table 3). The Fst value ranged from 0.022 to 0.366 and was in accordance with the genetic distances.

### Discussion

SSR markers have been widely used to evaluate genetic diversity in different plant species (Babayeva et al. 2009; Ehtemam et al. 2010; Ganapathy et al. 2012). Several studies also demonstrated that SSR markers were more effective in detecting genetic variability and relationship than other marker systems (Naghavi et al. 2004; Medini et al. 2005). Investigation of genomic relationships of wheat relatives is extremely important for establishing the evolutionary way of the formation of cultivated forms of *Triticum*.

In the current study, the eleven SSR primers revealed 83 alleles among 74 *T. urartu* accessions representing 8 countries. On average, 8 alleles were amplified by each of SSR marker (Table 1). This was higher than the one reported for 49 *T. urartu* accessions using RAPD markers (Castagna et al. 1997). Other researchers reported averages of 8 to 19 alleles per locus in various *T. urartu* collections (Hammer et al. 2000; Wang et al. 2017) using SSR markers. Rare alleles were detected for all loci except *Xbarc-206*. These alleles could be used as molecular identity data for specific genotypes (Ren et al. 2014). The main parameters of genetic diversity He and PIC were higher for most of the studied loci. The mean values for He and PIC over all loci and populations were 0.56 and 0.52, respectively, which can be estimated as high values. Thus, SSR markers prove its efficiency in studying genetic diversity in *T. urartu* even with a relatively small number of primers. However, the diversity in the current study was relatively lower than that reported by

---

**Table 3.** Pairwise estimates of Nei’s genetic distance (a lower diagonal matrix) and Fst (an upper diagonal matrix) among studied regions

<table>
<thead>
<tr>
<th>OTU</th>
<th>ARM</th>
<th>IRG</th>
<th>IRN</th>
<th>JOR</th>
<th>LBN</th>
<th>SAR</th>
<th>SYR</th>
<th>TUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARM</td>
<td>0.138</td>
<td>0.129</td>
<td>0.061</td>
<td>0.247</td>
<td>0.138</td>
<td>0.166</td>
<td>0.168</td>
<td></td>
</tr>
<tr>
<td>IRG</td>
<td>0.437</td>
<td>0.366</td>
<td>0.063</td>
<td>0.151</td>
<td>0.267</td>
<td>0.093</td>
<td>0.118</td>
<td></td>
</tr>
<tr>
<td>IRN</td>
<td>0.375</td>
<td>0.329</td>
<td>0.037</td>
<td>0.150</td>
<td>0.243</td>
<td>0.131</td>
<td>0.113</td>
<td></td>
</tr>
<tr>
<td>JOR</td>
<td>0.357</td>
<td>0.429</td>
<td>0.371</td>
<td>0.097</td>
<td>0.119</td>
<td>0.047</td>
<td>0.034</td>
<td></td>
</tr>
<tr>
<td>LBN</td>
<td>0.240</td>
<td>0.412</td>
<td>0.322</td>
<td>0.188</td>
<td>0.139</td>
<td>0.155</td>
<td>0.161</td>
<td></td>
</tr>
<tr>
<td>SAR</td>
<td>0.466</td>
<td>0.429</td>
<td>0.438</td>
<td>0.376</td>
<td>0.386</td>
<td>0.050</td>
<td>0.124</td>
<td></td>
</tr>
<tr>
<td>SYR</td>
<td>0.257</td>
<td>0.379</td>
<td>0.295</td>
<td>0.146</td>
<td>0.123</td>
<td>0.423</td>
<td>0.113</td>
<td></td>
</tr>
<tr>
<td>TUR</td>
<td>0.250</td>
<td>0.364</td>
<td>0.335</td>
<td>0.217</td>
<td>0.202</td>
<td>0.491</td>
<td>0.113</td>
<td></td>
</tr>
</tbody>
</table>
Wang et al. (2017) for 238 *T. urartu* genotypes with the use of 62 SSR markers. The contradiction can be attributed to the collection size and number of studied loci. High genetic diversity in different *T. urartu* collections was also noted using IRAP markers (Farouji et al. 2015). On the contrary, Moghaddam et al. (2000) reported low gene diversity in 23 populations of *T. urartu* using isozyme markers at eight polymorphic loci.

From a geographic point of view, the highest number of alleles was found in the 31 genotypes of Syrian origin (5.6 alleles), followed by 18 genotypes of Turkey origin (4.5 alleles) (Table 2). However, He and PIC values were not correlated with sample and allele sizes. Thus Jordan subset represented by only 10 accessions exhibited the highest diversity. The higher diversities were also observed in Syria and in Turkey. Interestingly, most of the accessions of Lebanon and Turkey and all accessions of Jordan were collected from the same provinces within these countries, although from different altitudes. The high genetic diversity within these countries can be explained by the differences in soil-climatic conditions and altitude variation (Table S1). The high diversity in these regions is in agreement with previous reports of Wang et al. (2017), Moghaddam et al. (2000) and Hammer et al. (2000). *T. urartu* is endemic to the major geographic regions of the Fertile Crescent (Heun et al. 2008), the center of origin and diversity of most species, including wheat. All the mentioned countries are a part of Fertile Crescent, thus, high genetic diversity in the region was expected.

The analysis of molecular variance also revealed that most of the genetic variability was accounted for by differences within countries (90%), with less variability among them (10%) (data not shown). The less variability among countries indicates notable similarity in allele frequencies.

The level of genetic dissimilarity index in the current study ranged from 0 to 1 with an average of 0.50, suggesting abundant genetic variation in this collection. The studied *T. urartu* genotypes fell into three main clusters (Fig. 1). In this study, the grouping patterns were not clearly associated with the geographic origins. As a result, genotypes from different geographic regions were widely dispersed and did not form main clusters. Nevertheless, it is in accordance with Gashaw et al. (2007) who suggested that genotypes from the same region might have the different genetic background and those collected from the different region might have a similar genetic background. Mousavifard et al. (2015) also demonstrated that *T. urartu* accessions showed the lowest correlation between genetic diversity and geographic distance. Such occurrence could be due to the gene flow and germplasm movement (Gashaw et al. 2007). A similar finding was also noted in previous reports of Singh et al. (2003) in durum wheat. However, unlike our findings, some previous observations show that SSR markers can differentiate wheat accessions according to their region of origin (Adary 1978; Wang et al. 2017).

PCoA analysis was in concordance with the cluster analysis (Fig. 2). According to the PCoA analysis, accessions belonging to the first subgroup of Cluster I were clustered much apart from others in the right half part of the scatter plot and formed independent PCoA group (group I). This group was genetically well differentiated from the rest of collection. The distances between most genotype pairs within the group were high (GD ≥ 0.5), suggesting that this group is highly diverse. Group I can be used in future breeding pro-
gram, as a crossing of genotypes with high genetic distances may produce a broad-spectrum of variability in segregating populations (Gashaw et al. 2007). Group II in scatter plot contained rest of the accessions. Most of the accessions of cluster III were placed in the upper left, while the majority of the genotypes from cluster II were placed in below left part of the PCoA plot and was closer to the second subcluster of cluster I.

Table 3 represents Nei’s genetic distances (a lower diagonal matrix) and values of pairwise F-Statistics (Fst) (an upper diagonal matrix) between pairs of geographic regions included in the study. The highest genetic distance among 8 countries was recorded for Turkey and Saudi Arabia and the lowest for Turkey and Syria. Population from Lebanon showed smaller genetic distance with the population of Syria, followed by Jordan. In general, genetic distances between geographic regions were low or moderate (lower than 0.5), which was also supported by low values of pairwise Fst. This once again confirms low diversity among different populations.

In spite of the fact that the germplasms used in the current study were from different countries, most of them represented close geographic locations. For example, Jordan and Lebanon belonged to Eastern Mediterranean coast; longitude values for samples’ collection sites ranged from 35°37’ to 36°10’. All Syrian genotypes, except two, represented mostly southwestern and slightly northwestern part of Syria with longitude values between 36°01’ and 37°39’. Turkey belongs to Mesopotamia region, however, majority of accessions from this country was collected from Gaziantep with collection sites between 36°55’ and 37°35’. Relatively close locations of these regions confirm the high possibility of gene flow and migration among them. This, in turn, can result in high similarities and low diversity among these four countries as in case of our study. Similarly, Wang et al. (2017) reported low diversity between Lebanon and southwestern Syrian samples based on SSR data.

In the NJ tree, constructed based on Nei’s genetic distances, Turkey, Syria, Jordan and Lebanon were relatively close, followed by Armenian genotypes (Fig. 3). Iran, Iraq and Saudi Arabia were represented by only one genotype in our study, which makes further discussion of these samples difficult. Future analyses with larger collections of \textit{T. urartu} from these regions are needed to draw more precise conclusions about the relationship of these collections. However, it would be interesting to note that, longitude values for samples from Iran and Iraq were 41°47’ and 48°28’, respectively.

In conclusion, we have demonstrated the utility of SSR markers in revealing genetic diversity in \textit{T. urartu} accessions of diverse origin. High variability in the studied collection was detected by only 11 SSR markers, of which 7 markers were the most effective. Diversity revealed within countries was higher than among them, even in the same regions of the relevant countries. Our findings can direct the sampling strategies on \textit{T. urartu} in these regions to find beneficial alleles. On the contrary, diversity between pairs of countries was lower, which can be partially explained by the close location of these regions and indicate gene flow among them. The heterotic groups detected by cluster and PCoA analysis in the present study can serve as effective candidates in crossing programs to broaden the genetic base in \textit{T. urartu}. 

\textit{Cereal Research Communications 46, 2018}
References


Electronic Supplementary Material (ESM)

Electronic Supplementary Material (ESM) associated with this article can be found at the website of CRC at https://akademiai.com/loi/0806

Electronic Supplementary Table S1. T. urartu accessions used in the study