



# Genotyping-by-Sequencing (GBS) Revealed Molecular Genetic Diversity of Iranian Wheat Landraces and Cultivars

Hadi Alipour<sup>1,2,3</sup>, Mohammad R. Bihamta<sup>2</sup>, Valiollah Mohammadi<sup>2</sup>, Seyed A. Peyghambari<sup>2</sup>, Guihua Bai<sup>4</sup> and Guorong Zhang<sup>3\*</sup>

<sup>1</sup> Department of Plant Breeding and Biotechnology, Faculty of Agriculture, Urmia University, Urmia, Iran, <sup>2</sup> Department of Agronomy and Plant Breeding, Faculty of Agriculture, University of Tehran, Karaj, Iran, <sup>3</sup> Agronomy Department, Kansas State University, Manhattan, KS, United States, <sup>4</sup> Hard Winter Wheat Genetics Research Unit, United States Department of Agriculture – Agricultural Research Service, Manhattan, KS, United States

## OPEN ACCESS

### Edited by:

Petr Smýkal,  
Palacký University, Olomouc, Czechia

### Reviewed by:

Tomáš Vyhnanek,  
Mendel University in Brno, Czechia  
Faheem Shehzad Baloch,  
Abant Izzet Baysal University, Turkey  
Moses Nyine,  
International Institute of Tropical  
Agriculture (IITA), Uganda

### \*Correspondence:

Guorong Zhang  
gzhang@ksu.edu

### Specialty section:

This article was submitted to  
Crop Science and Horticulture,  
a section of the journal  
Frontiers in Plant Science

**Received:** 28 April 2017

**Accepted:** 07 July 2017

**Published:** 29 August 2017

### Citation:

Alipour H, Bihamta MR,  
Mohammadi V, Peyghambari SA,  
Bai G and 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

**Background:** Genetic diversity is an essential resource for breeders to improve new cultivars with desirable characteristics. Recently, genotyping-by-sequencing (GBS), a next-generation sequencing (NGS) technology that can simplify complex genomes, has now be used as a high-throughput and cost-effective molecular tool for routine breeding and screening in many crop species, including the species with a large genome.

**Results:** We genotyped a diversity panel of 369 Iranian hexaploid wheat accessions including 270 landraces collected between 1931 and 1968 in different climate zones and 99 cultivars released between 1942 to 2014 using 16,506 GBS-based single nucleotide polymorphism (GBS-SNP) markers. The B genome had the highest number of mapped SNPs while the D genome had the lowest on both the Chinese Spring and W7984 references. Structure and cluster analyses divided the panel into three groups with two landrace groups and one cultivar group, suggesting a high differentiation between landraces and cultivars and between landraces. The cultivar group can be further divided into four subgroups with one subgroup was mostly derived from Iranian ancestor(s). Similarly, landrace groups can be further divided based on years of collection and climate zones where the accessions were collected. Molecular analysis of variance indicated that the genetic variation was larger between groups than within group.

**Conclusion:** Obvious genetic diversity in Iranian wheat was revealed by analysis of GBS-SNPs and thus breeders can select genetically distant parents for crossing in breeding. The diverse Iranian landraces provide rich genetic sources of tolerance to biotic and abiotic stresses, and they can be useful resources for the improvement of wheat production in Iran and other countries.

**Keywords:** Iranian wheat landraces, genetic diversity, genotyping-by-sequencing, single nucleotide polymorphism, population structure

## INTRODUCTION

Wheat (*Triticum aestivum* L.) is a staple food crop that feeds about 30% of the world population and provides over 20% of the calories consumed by humans (FAO, 2015). Due to a rapidly growing world population and climate changes, breeders and farmers are facing the challenge of increasing wheat production up to 70% by 2050 to meet future demands (FAO, 2009; Ray et al., 2013;

Marcussen et al., 2014), which needs a 2.4% of yield increase yearly. However, the current global average rate of crop yield increase is only 0.9% per year, which is far slower than the desired rate (Ray et al., 2013).

Wheat grain yield can be increased by improvement of both crop management practices and genetic improvement of cultivars for high yield potential (Sener et al., 2009). Genetic diversity is the foundation for such genetic improvement (Nielsen et al., 2014; Govindaraj et al., 2015). Iran is part of the wheat center of origin, and Iranian wheat landraces are important genetic resources for new alleles or genes to be used in breeding for new cultivars (Ciaffi et al., 1992). Wheat landraces are distinct and locally adapted accessions collected and grown by farmers may have a high level of tolerance to biotic and abiotic stresses hence, they are able to provide higher sustainable yields under low input agricultural conditions (Zeven, 1998; Skovmand et al., 2002). For example, Iranian landraces PI 1377397 (Toit, 1989) and PI 626580 (Valdez et al., 2012) have been used as the sources of resistance to Russian wheat aphid [*Diuraphis noxia* (Kurdjumov)]. ‘Turkey Red,’ a hard red winter wheat from Turkey, has been the foundation for hard winter wheat cultivars in the United States due to its cold tolerance (Olmstead and Rhode, 2002).

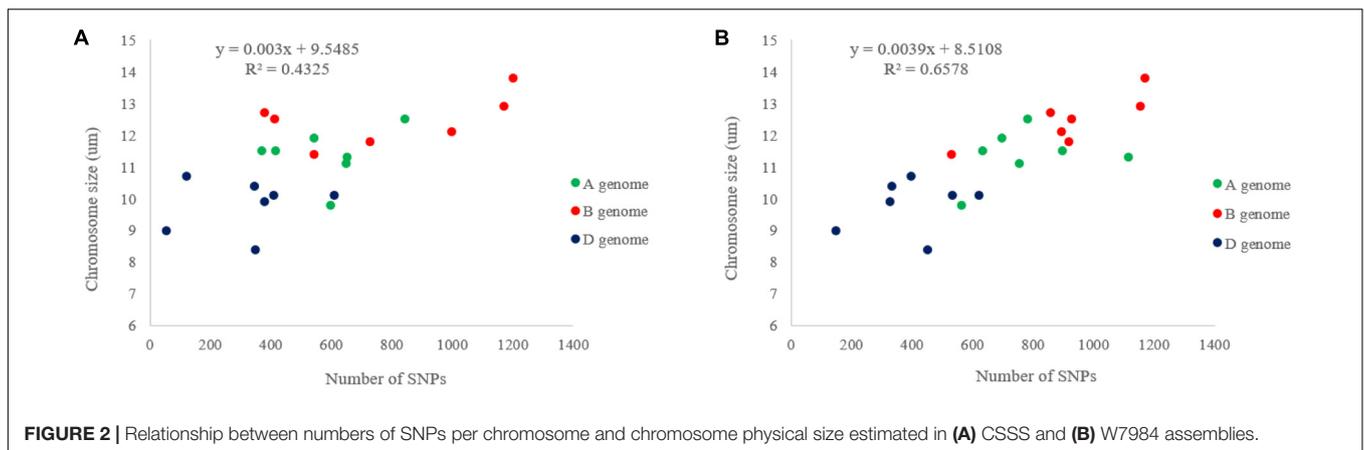
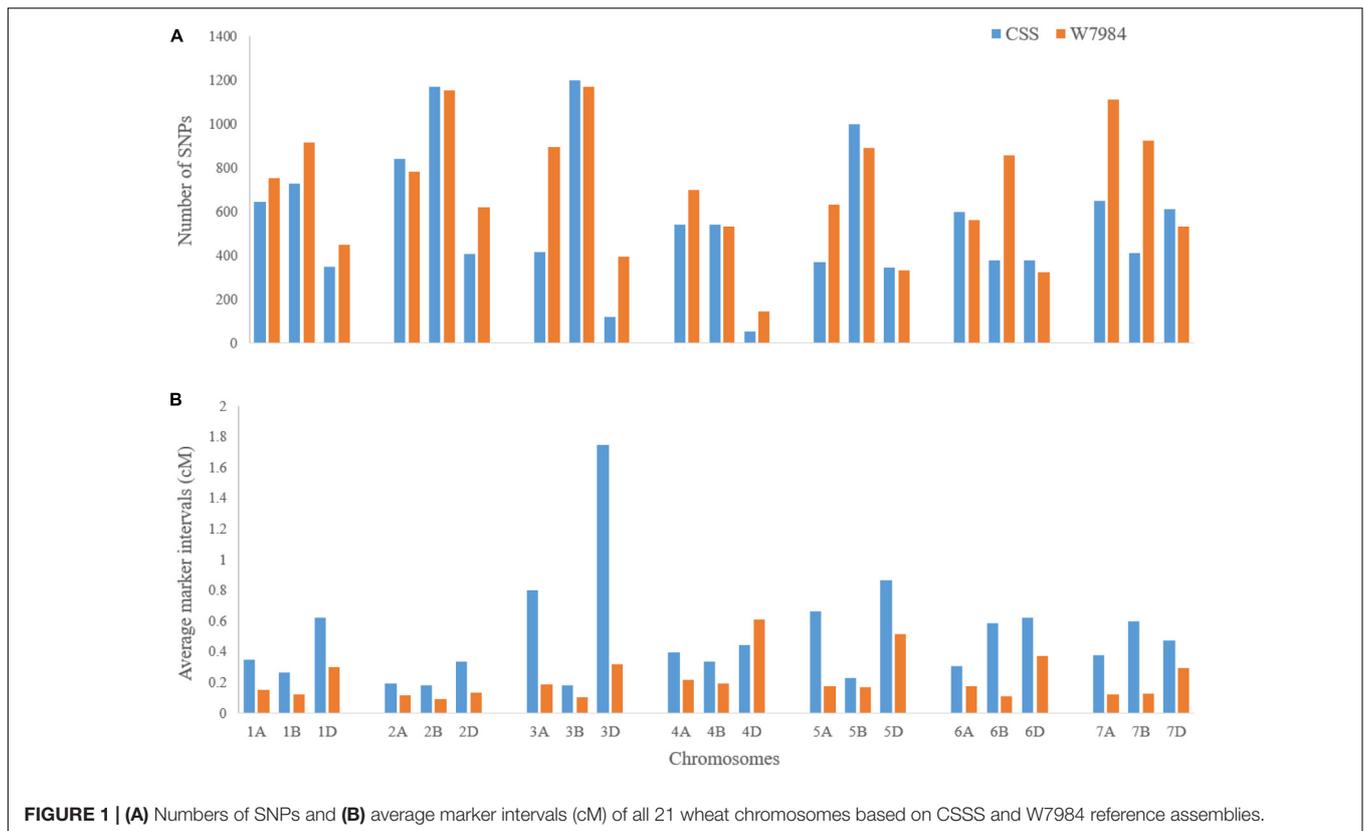
After the green revolution in the mid-20th century (Borlaug, 1968), wheat landraces have been widely replaced with modern semi-dwarf cultivars, which significantly narrowed genetic diversity. Although artificial selection performed during domestication and breeding processes has increased the frequency of favorable alleles controlling yield and input responses, some other desirable alleles such as biotic and abiotic stress resistance have been removed from breeding populations

and cultivars. Therefore, genetic diversity of bread wheat is diminishing in breeding programs (Sofalian et al., 2008). Conservation of landraces and their wild relatives becomes a critical measure to avoid genetic erosion and meet future need of wheat yield increase (Chen et al., 1994).

Molecular markers have been widely used to study the population structure and genetic diversity of germplasm collections (Huang et al., 2002; Dreisigacker et al., 2005; Hao et al., 2006, 2008; Cavanagh et al., 2013). Single nucleotide polymorphisms (SNPs) are the most abundant type of sequence variations in plant genomes (Batley and Edwards, 2007). They are suitable for analysis of genetic variation, population structure, marker-trait association, genomic selection, QTL mapping, map-based cloning, and other plant breeding applications that need large number of markers to cover entire genomes (Kumar et al., 2012). High-throughput SNP arrays are available, but the high cost per sample limits their application in breeding. Recently, next-generation sequencing (NGS) provides a high-throughput and cost-effective molecular tool for breeding and has been widely used to speed up breeding processes (Poland and Rife, 2012; Edae et al., 2014). Rapid advances in NGS technology have driven the costs of DNA sequencing down to the point that genotyping-by-sequencing (GBS) can now be used for routine breeding screening in any crops (Elshire et al., 2011). GBS can be used for marker discovery and genotyping simultaneously and many samples can be multiplexed to reduce cost per sample (He et al., 2014). GBS uses restriction enzyme digestion to reduce the complexity of genomes, which makes it possible to analyze plant species with large and complex genomes such as wheat, and the low cost per sample makes it feasible for breeders to use it as a

**TABLE 1** | A summary of single nucleotide substitutions identified in the three homoeologous wheat genomes based on CSSS and W7984 reference assemblies.

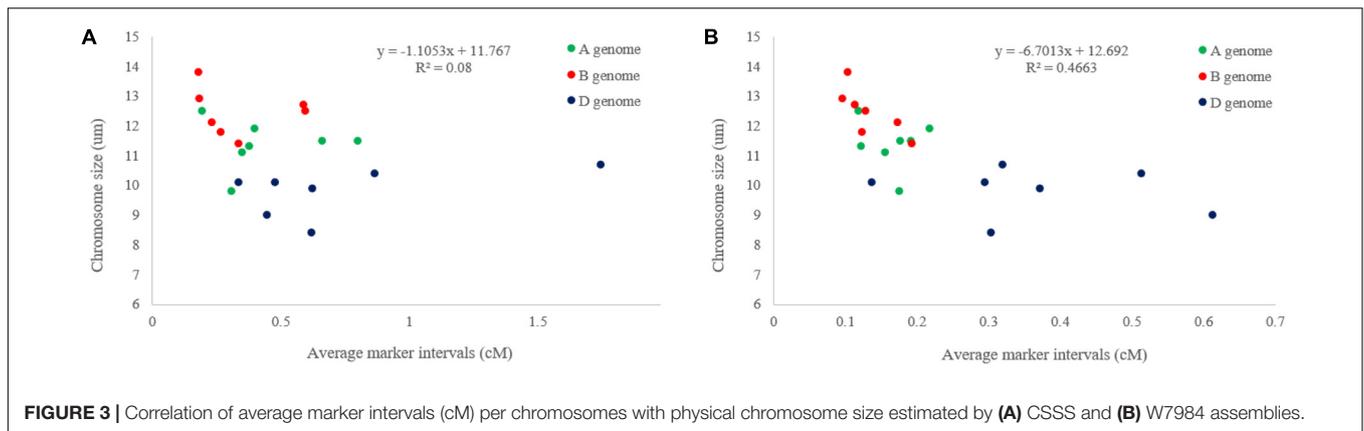
Genome	CSSS reference (International Wheat Genome Sequencing [IWGS], 2014)				W7984 reference (Chapman et al., 2015)				Total
	A	B	D	Unassigned	A	B	D	Unassigned	
No. of SNPs	4067	5425	2266	4748	5442	6449	2806	1809	16506
Chromosome size (Mbp)	5727	6274	4937	–	5727	6274	4937	–	16938
Marker density (SNP/Mbp)	0.710	0.865	0.459	–	0.950	1.028	0.568	–	0.974
A/G	1213	1564	651	1405	1648	1849	793	543	4833
C/T	1190	1527	582	1263	1576	1808	730	448	4562
G/A	131	203	80	173	175	238	108	66	587
T/C	125	171	84	141	165	212	90	54	521
Transition	2659	3465	1397	2982	3564	4107	1721	1111	10503
Ts %	65.38	63.87	61.65	62.81	65.49	63.68	61.33	61.42	63.63
A/T	195	253	136	223	273	309	150	75	807
A/C	310	411	192	367	410	481	234	155	1280
T/A	26	30	12	27	33	35	15	12	95
T/G	21	45	20	47	33	55	27	18	133
C/A	28	47	11	41	40	57	18	12	127
C/G	492	704	305	637	657	841	382	258	2138
G/T	280	383	162	332	363	457	201	136	1157
G/C	56	87	31	92	69	107	58	32	266
Transversion	1408	1960	869	1766	1878	2342	1085	698	6003
Tv %	34.62	36.13	38.35	37.20	34.51	36.32	38.67	38.59	36.37
Ts/Tv ratio	1.89	1.77	1.61	1.69	1.90	1.75	1.59	1.59	1.75



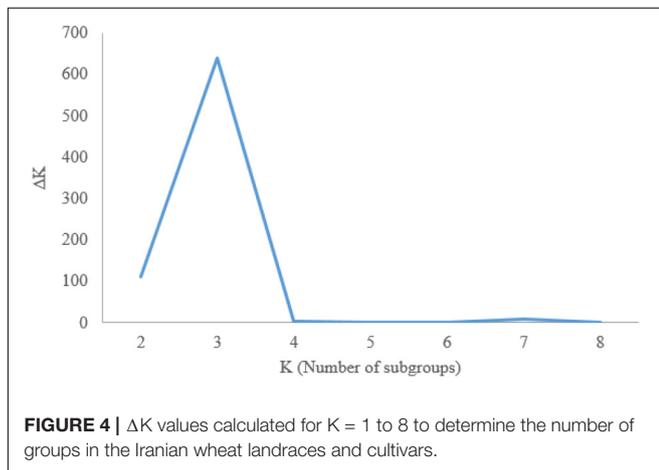
routine tool to aid breeding selection. An improved GBS protocol (Poland et al., 2012b) that uses two-enzyme (*PstI/MspI*) digestion provides a greater degree of complexity reduction and more uniform libraries for sequencing than the original single enzyme protocol (Elshire et al., 2011). The improved GBS protocol has been successfully used in cereal crops such as barley, wheat and oat (Poland et al., 2012a,b; Huang et al., 2014).

Iranian wheat landraces provide a rich source of genetic diversity and carry resistance genes to many different biotic stresses such as bunt diseases (Bonman et al., 2015), Russian wheat aphid (Ehdaie and Baker, 1999; Valdez et al., 2012; Bonman et al., 2015), leaf and stripe rusts (Kertho et al., 2015), stem rust

(Rouse et al., 2011; Newcomb et al., 2013), and abiotic stresses such as salinity (Jafari-Shabestari et al., 1995), drought and heat (Ehdaie et al., 1988). To date, most of the Iranian germplasm lines have not been characterized and used in modern plant breeding (Hoisington et al., 1999). These germplasm lines not only provide new sources of resistance to biotic and abiotic stresses, but also can enhance the biodiversity of the breeding materials (Huang et al., 2010). Therefore, assessing genetic variation and differentiation of Iranian wheat landraces and cultivars will facilitate the effective use of these valuable genetic resources in future breeding to broaden the genetic diversity of Iranian breeding materials and identify novel alleles that could be used



**FIGURE 3** | Correlation of average marker intervals (cM) per chromosomes with physical chromosome size estimated by (A) CSSS and (B) W7984 assemblies.



**FIGURE 4** |  $\Delta K$  values calculated for  $K = 1$  to 8 to determine the number of groups in the Iranian wheat landraces and cultivars.

by geneticists and breeders in Iran and other countries. To the best of our knowledge, the present study is the first to directly compare the population diversity of Iranian wheat landraces to a representative pool of Iranian wheat cultivars using SNP markers.

## MATERIALS AND METHODS

### Genetic Materials

A set of 369 Iranian hexaploid wheat accessions (**Supplementary Table S1**) were used in this study. They included 270 landraces collected between 1931 and 1968 in different climate zones (Supplementary Figure S1) and 99 cultivars released between 1942 and 2014. These were kindly provided by the University of Tehran (UT) and Seed and Plant Improvement Institute (SPII), Karaj, Iran, United States Department of Agriculture-Agricultural Research Service (USDA-ARS)-National Plant Germplasm System, United States, and International Center for the Improvement of Maize and Wheat (CIMMYT), Mexico.

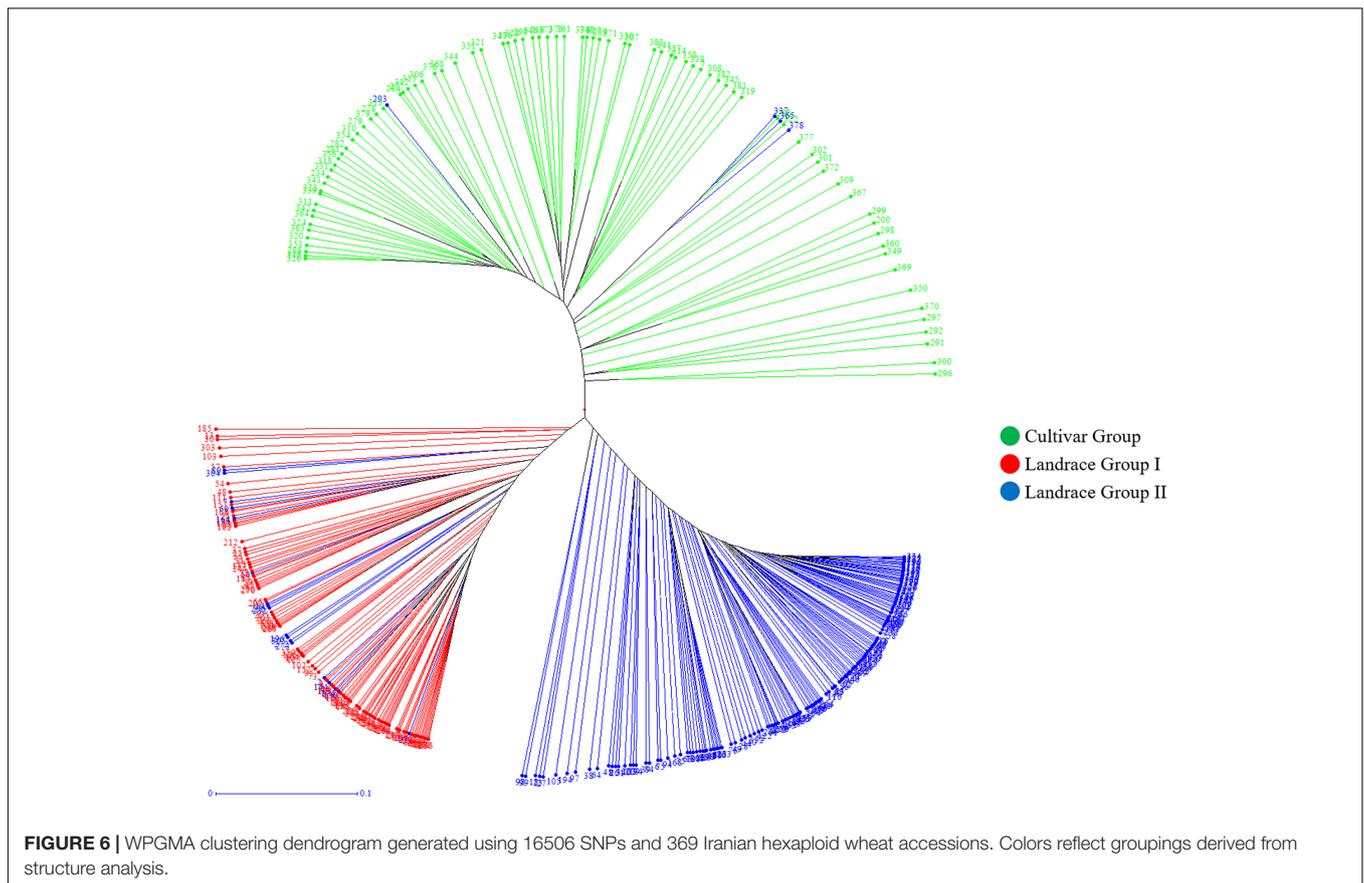
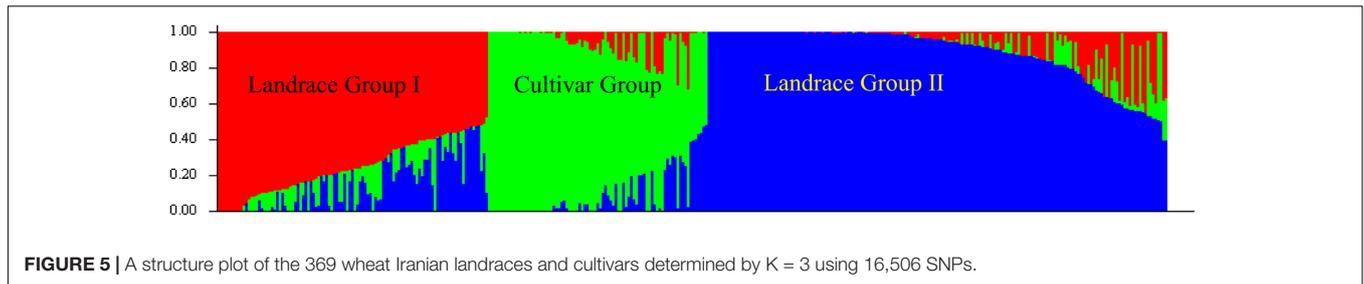
### GBS Library Preparation and Sequencing

Genomic DNA was extracted using a modified cetyltrimethyl ammonium bromide (CTAB) method (Saghai-Marouf et al.,

1984) from five 2-weeks-old seedlings. DNA concentration was quantified using the Quant-iT<sup>TM</sup> PicoGreen<sup>®</sup> dsDNA Assay (Life Technologies, Inc., Grand Island, NY, United States) and normalized to 20 ng/ $\mu$ l for library construction.

The GBS libraries were constructed following Poland et al. (2012a). In brief, genomic DNA was digested using the restriction enzymes *Pst*I and *Msp*I (New England BioLabs, Inc., Ipswich, MA, United States), and barcoded adapters were ligated to each DNA samples using T4 ligase (New England BioLabs, Inc.). All the ligated products from each plate were pooled and cleaned up using the QIAquick PCR Purification Kit (Qiagen, Inc., Valencia, CA, United States). Primers complementary to both adaptors were used for PCR. The PCR products were then cleaned up again using the QIAquick PCR Purification Kit and quantified using Bioanalyzer 7500 Agilent DNA Chip (Agilent Technologies, Inc.). After size-selection for 250–300 bp fragments in an E-gel system (Life Technologies, Inc.), the concentration of each library was estimated by the Qubit 2.0 fluorometer using Qubit dsDNA HS Assay Kit (Life Technologies, Inc.). The size-selected library was sequenced on an Ion Proton sequencer (Life Technologies, Inc.).

Sequence reads were first trimmed to 64 bp, and identical reads were grouped. Then, unique sequence tags were assigned to the sequence groups. The unique tags were aligned internally allowing mismatches of up to 3 bp to identify SNPs within the tags. SNP calling pipeline was employed as described by Poland et al. (2012b). This pipeline is implemented in TASSEL 3 and was functionally identical to UNEAK to the point of developing a binary presence/absence matrix for each tag across multiple lines. To identify putative SNPs, tags were internally aligned allowing up to 3 bp mismatch in a 64 bp tag. From aligned tags, SNP alleles were identified and the number of lines in the population with each respective tag was tallied in a  $2 \times 2$  table, counting the number of lines with one or the other tag, both, or neither. A Fisher Exact Test was then used to determine if the two alleles were independent, as would be expected for a single locus, bi-allelic SNP in a population of inbred lines. If the null hypothesis of independence for the putative SNP was rejected ( $p < 0.001$ ), we assumed that the tags were allelic in the population (and, therefore, that the putative SNP was a true single locus, bi-allelic SNP). A significance



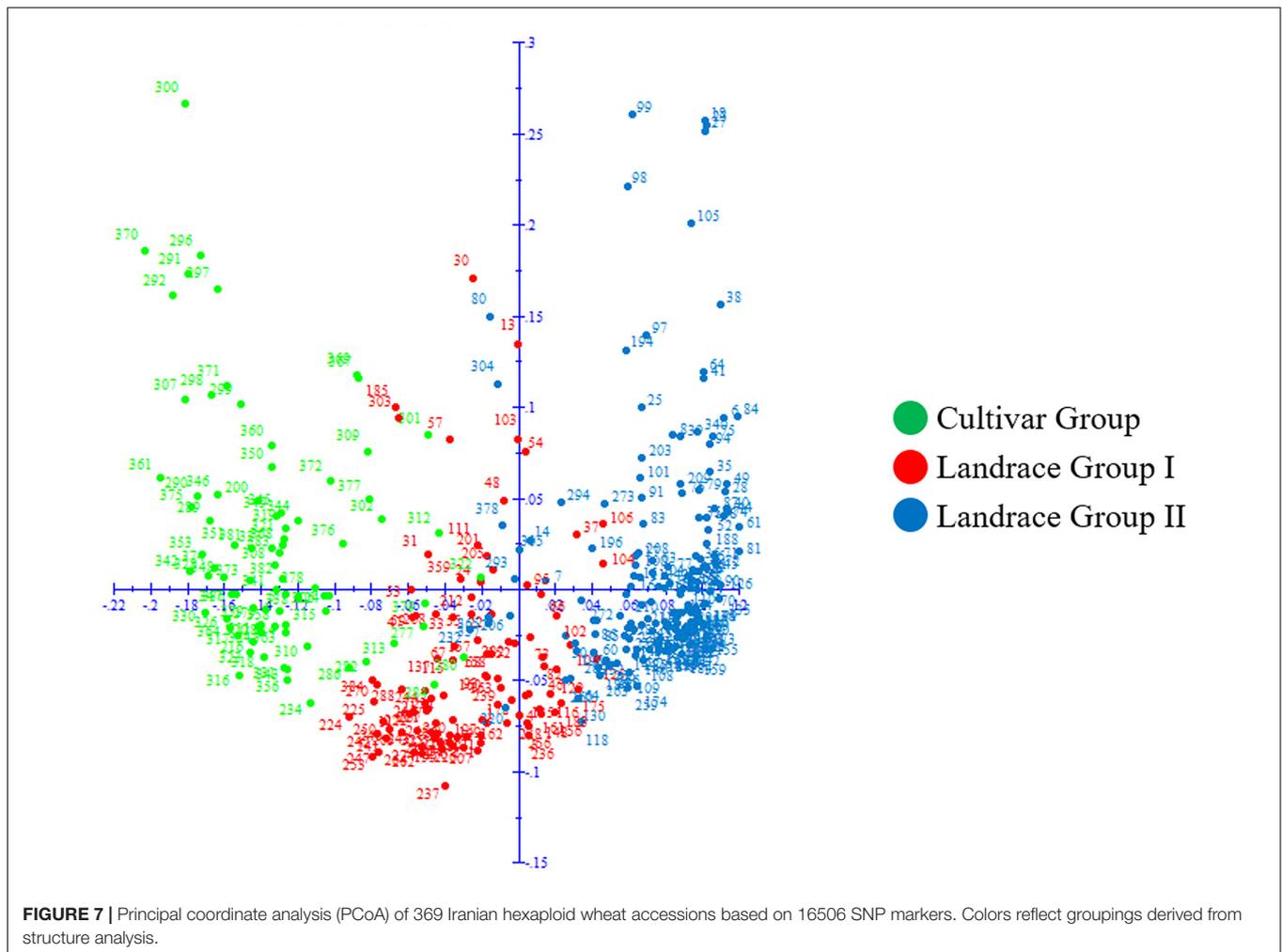
threshold of  $p < 0.001$  was selected for the size of population, based on previous work testing false discovery rates in duplicate samples.

Single nucleotide polymorphism calling was conducted using the UNEAK (Universal Network Enabled Analysis Kit) GBS pipeline (Lu et al., 2013), which is part of the TASSEL 4.0 bioinformatics analysis package (Bradbury et al., 2007). Reads with the low-quality score ( $< 15$ ) were removed, and SNPs with heterozygotes  $< 10\%$ , a minor allele frequency  $> 1\%$ , and missing data  $< 20\%$  were used for further analysis. BLASTn analysis was carried out to align sequence reads to the flow-sorted Chinese Spring survey sequence (CSSS) (International Wheat Genome Sequencing [IWGS], 2014) and the Popseq W7984 sequence reference (Chapman et al., 2015) and their location on the genetic map (cM) were predicted through the comparison

with GenomeZipper, 90K consensus map and POPSEQ (International Wheat Genome Sequencing [IWGS], 2014).

## Data Analysis

Genetic diversity analysis was performed using DARwin version 6.010 software (Perrier and Jacquemoud-Collet, 2006) and the Jaccard index. The diversity tree was built using WPGMA and Neighbor-Joining algorithm (Saitou and Nei, 1987) that relaxes the assumption of equal mutation rates over space and time and produces an un-rooted tree. The confidence interval of the genetic relationships among the accessions was determined by performing 1,000 bootstraps, with the results expressed as percentages at the main nodes of each branch. Analysis of molecular variance (AMOVA) was used to partition the genetic variation into inter- and intra-gene pool diversities using Arlequin V3.5 software (Excoffier and Lischer, 2010).

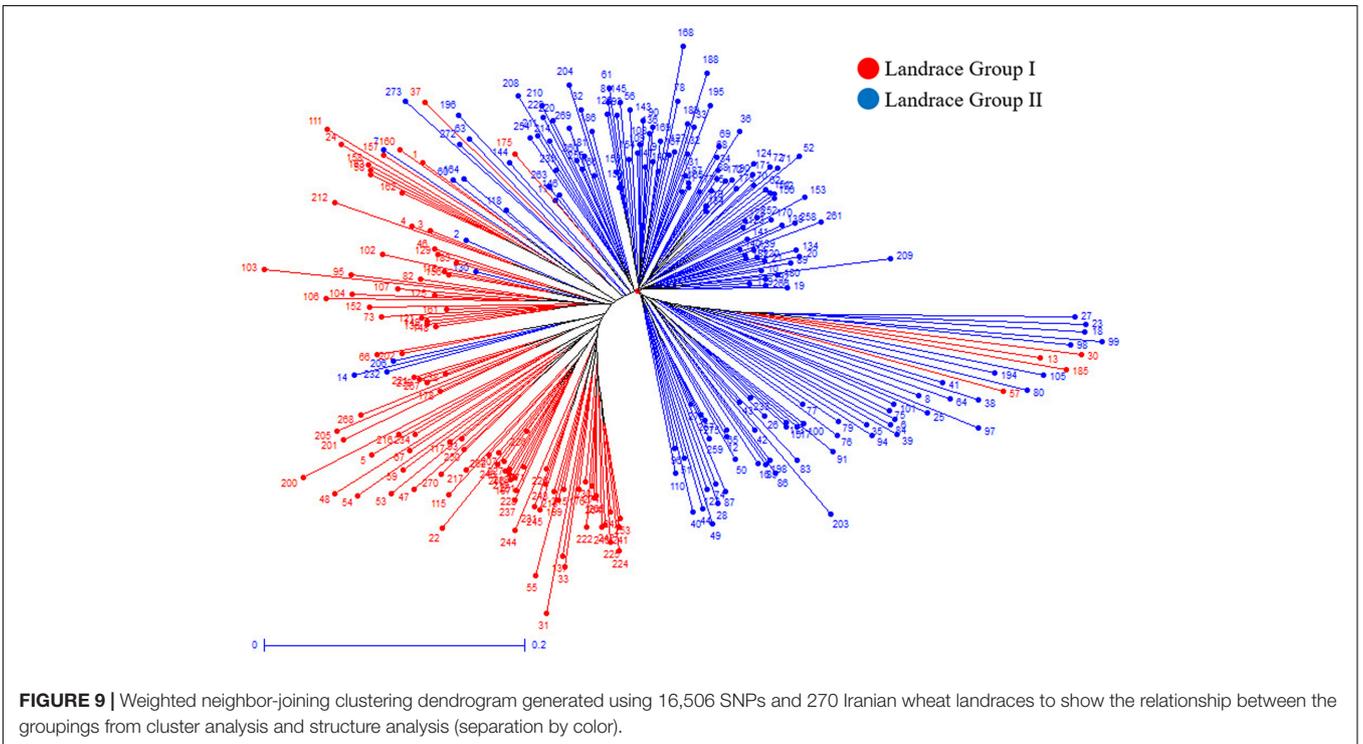
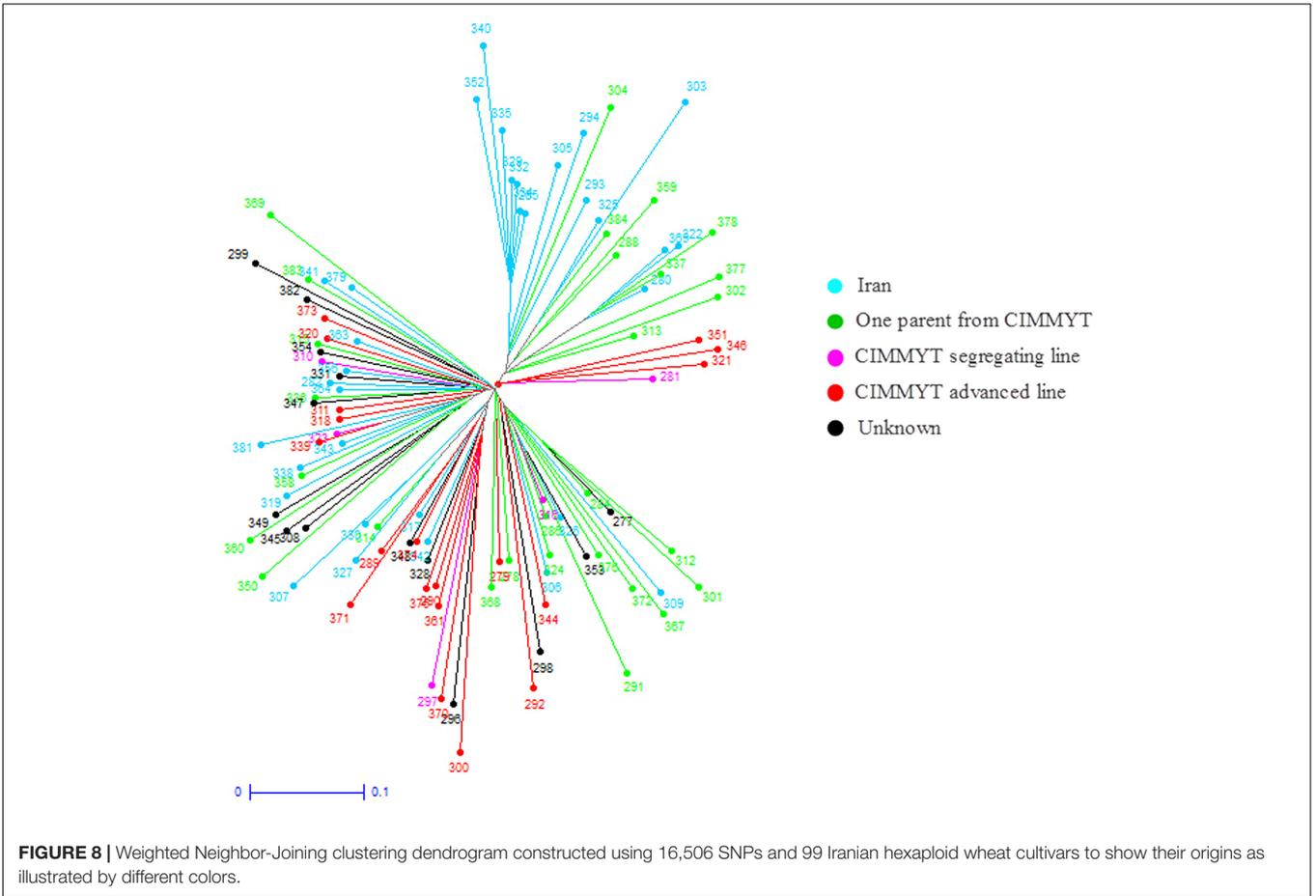


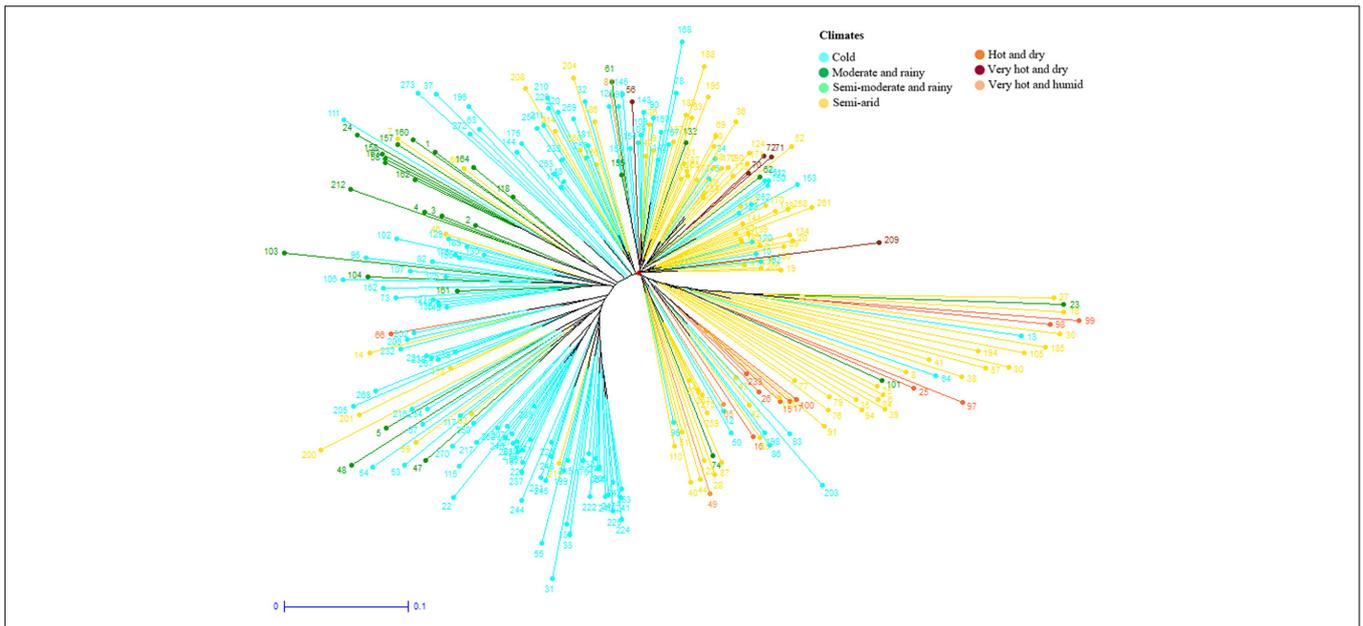
For analysis of population structure, a model-based Bayesian cluster analysis was performed using STRUCTURE version 2.3.4 (Pritchard et al., 2000). The structure analysis was run 10 times for each K value (K = 1 to 8) using a burn-in period of 10,000 steps and 10,000 MC steps and an admixture model. All parameters were set to default values recommended by the manufacturer (Pritchard et al., 2010). The probability of best fit into each number of assumed clusters (K) was estimated by an ad hoc statistic  $\Delta K$  based on the rate of change in the log probability of data between consecutive K values (Evanno et al., 2005).

## RESULTS

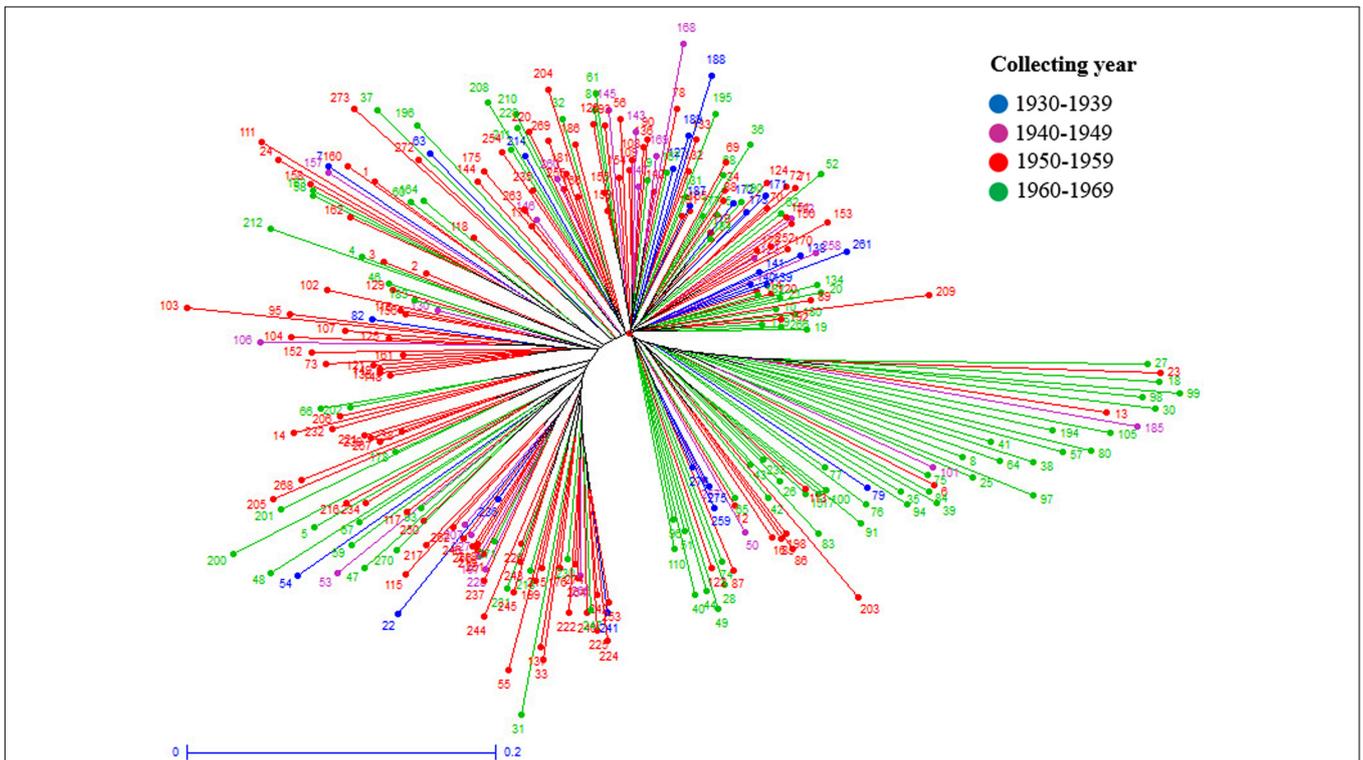
After eight Ion Proton runs of 369 samples, a total of 566,439,207 reads were identified with 458,363,607 (about 81%) unique reads. A total of 133,039 GBS-SNPs were called after filtering out duplicated reads. Among them, 16,506, 38,824, and 56,560 GBS-SNPs have <20, <50, and <80% missing data, respectively. Only the SNPs with <20% missing data were used to evaluate the genetic diversity of the diversity panel. In the BLASTn analysis, a total of 11,758 (~71.24%) and 14,697 (~89.04%)

SNPs were aligned to CSSS and W7984 reference genomes, respectively. The highest numbers of SNPs were mapped on the B-genome and the lowest on the D-genome in both CSSS and W7984 references. The highest values of transition-type SNPs were identified on the B genome with 3,465 and 4,107 SNPs while the lowest were on the D genome with 1,397 and 1,721 SNPs on the CSSS and W7984 references, respectively (**Table 1**). A similar chromosome distribution pattern was observed for transversion-type SNPs. More transition-type SNPs (63.63%) were observed than transversion-type SNPs (36.37%) with a transition/transversion (Ts/Tv) SNP ratio of 1.75 (10,503/6,003) over all three genomes. However, the ratio was significantly higher in the A genome than those in the B or D genomes. As expected, more A/G and C/T transitions were observed than G/A and T/C transitions. On the other hand, more C/G, A/C, G/T, and A/T transversions were observed than G/C, C/A, T/G, and T/A transversions. The SNP numbers in the most of the chromosomes of W7984 assembly were higher than those of CSSS, with only a few exceptions (**Figure 1A**). Similarly, the average marker density was lower in D-genome than in other two genomes of both references (**Figure 1B**). The maximum marker density per chromosome of W7984 assembly ( $\approx 10.4$  SNP/cM)





**FIGURE 10 |** Dendrogram to demonstrate the genetic relationships among 270 Iranian hexaploid wheat landraces based on 16506 SNP markers using Weighted Neighbor-Joining clustering. Climate factors were labeled by colors.



**FIGURE 11 |** A weighted neighbor-joining clustering dendrogram generated using 16506 SNPs and 270 Iranian hexaploid wheat landraces to show the years of accessions collected.

was much higher than that from CSSS assembly (5.5 SNP/cM). The largest marker intervals were observed for chromosome 3D (~1.77 cM) of CSSS and 4D (~0.62 cM) of W7984 assembly.

Highly significant correlation was observed between the number of SNPs per chromosome and their physical size for both CSSS (**Figure 2A**) and W7984 (**Figure 2B**), though this correlation

**TABLE 2** | Genetic variation among three groups identified by structure analysis on a diversity panel of 369 Iranian hexaploid wheat landraces and cultivars.

Diversity index	Landrace group I	Cultivar group	Landrace group II	Average
$N_a$	1.915 ± 0.003	1.779 ± 0.005	1.916 ± 0.003	1.870 ± 0.002
$N_e$	1.317 ± 0.004	1.310 ± 0.004	1.223 ± 0.003	1.283 ± 0.002
$H_e$	0.200 ± 0.002	0.192 ± 0.002	0.144 ± 0.002	0.179 ± 0.001
$I$	0.320 ± 0.003	0.301 ± 0.003	0.240 ± 0.002	0.287 ± 0.002
Private alleles	0.028 ± 0.002	0.012 ± 0.001	0.031 ± 0.002	0.026 ± 0.002
PPL	91.50	77.86	91.60	86.99 ± 40.57
PIC	Min_PIC = 0.003		Max_PIC = 0.375	Mean_PIC = 0.172

Values in tables are estimates ± standard error.  $N_a$ , average number of alleles per locus;  $N_e$ , effective number of alleles;  $H_e$ , average expected heterozygosity;  $I$ , Shannon diversity index; PPL, percentage of polymorphic loci; PIC, polymorphism information content.

was much higher for SNPs called using W7984 reference than CSS reference. The negative correlation between average marker intervals and chromosome sizes was significant for SNPs called using W7984 reference (Figure 3B), but not significant for SNPs called using the CSSS reference (Figure 3A).

## Population Structure

To assess the structure of the Iranian wheat diversity panel, delta K ( $\Delta K$ ) values were used to classify subgroups (K). The largest  $\Delta K$  was observed at  $K = 3$ , suggesting three groups in the panel (Figures 4, 5). Group I contains 104 accessions with 99 landraces and five cultivars, designated as 'Landrace Group I'; Group II consists of 84 accessions with 80 cultivars and four landraces, designated as 'Cultivar Group'; and Group III is the largest including 181 accessions with 167 landraces and 14 cultivars, designated as 'Landrace Group II.' Most of the cultivars that mixed with landrace group I and II, such as Azar, Biston, Dastjerdi, Deyhim, Homa, Ohadi, Shahi, Sefidak, Shahpasand, and Reyhani, were originally selected from Iranian landraces through continuous selection and purification during the breeding process.

Cluster analysis also classified the panel into three groups that matched with the results from structure analysis (Figure 6) and the principal coordinate analysis (PCoA) (Figure 7). In PCoA, the first and second coordinate explained 12.03 and 6.20% of the variation, respectively and the Landrace Group I was located between the Landrace Group II and Cultivar Group (Figure 7).

Further cluster analysis divided the cultivars into four subgroups (Figure 8). Most of the cultivars originated from Iran or contain one parent from Iran were separated from those originated from CIMMYT. Cluster analysis on landraces only identified the same two landrace groups as in structure analysis (Figure 9). The two landrace groups were grown in distinct

climates with Landrace Group I grown under cold or moderate, but rainy climates, and Landrace Group II grown under hot and dry or semiarid climates (Figure 10). Also, the two groups can be roughly separated by collection years with Landrace Group I collected before 1960 and Landrace Group II collected after 1960 (Figure 11).

## Genetic Diversity

Intra-population genetic diversity analysis revealed that mean observed ( $N_a$ ) and effective ( $N_e$ ) allele numbers were 1.87 and 1.28, respectively (Table 2). The lowest  $N_a$  was observed in the Cultivar Group (1.779), and the lowest  $N_e$  was in the Landrace Group II (1.223). The expected heterozygosity ( $H_e$ ) varied from 0.144 (Landrace Group II) to 0.200 (Landrace Group I). A similar order was observed for Shannon's diversity index, which varied from 0.240 (Landrace Group II) to 0.320 (Landrace Group I). The lowest private allele number was found in the Cultivar Group ( $0.012 \pm 0.001$ ), whereas both the landrace groups (Landrace Group I and II) showed a higher value of private alleles. Percentage of polymorphic loci per groups ranged from 77.86% (Cultivar Group) to 91.60% (Landrace Group II). Mean marker polymorphism information content (PIC) was low (0.172), ranging from 0.003 to 0.375.

Analysis of pairwise genetic differentiation and gene flow among three groups (Table 3) revealed that the highest value of genetic differentiation was between the Cultivar Group and the Landrace Group II ( $F_{st} = 0.309$ ), whereas the values were the same between the Landrace Group I to the Cultivar Group and between the two landrace groups. On the other hand, the lowest gene flow was between the Cultivar Group and the Landrace Group I ( $N_m = 0.559$ ) while the highest was between the two landrace groups ( $N_m = 1.166$ ).

The AMOVA on the landrace population vs. cultivar population showed a much greater variation within a population ( $67.36\% + 17.08\% = 84.44\%$ ) than among the populations ( $15.56\%$ ,  $p < 0.001$ ) (Table 4). Similar results were obtained using groups derived from structure analysis although slightly higher inter-population variation ( $22.98\%$ ,  $p < 0.001$ ) than within a population was observed. The  $F_{st}$  value of 0.16 between landraces and cultivars suggested a substantial degree of differentiation between them while a slightly higher  $F_{st}$  value (0.23) among three groups generated by structure analysis also suggests a high differentiation between the two landrace groups.

**TABLE 3** | Gene flow ( $N_m$ , upper right diagonal) and pairwise genetic differentiation ( $F_{st}$ , down left diagonal) among the two Iranian wheat landrace groups and the cultivar group grouped by structure analysis.

	Landrace group I	Cultivar group	Landrace group II
Landrace group I	–	1.166	1.031
Cultivar group	0.177	–	0.559
Landrace group II	0.195	0.309	–

## DISCUSSION

Using the POPSEQ approach, we ordered 11,758 (~71.24%) and 14,697 (~89.04%) SNPs to the W7984 and CSSS assemblies, respectively, which were much higher than these reported (22.5–46.3%) in previous studies (Würschum et al., 2013; Shavrukov et al., 2014; Wang et al., 2014). Edae et al. (2015) obtained 33,664 SNPs with up to 80% missing data from W7984 × Opata M85 RIL population out of which, 16,591 (49.3%) and 9709 (28.8%) SNPs were mapped to the W7984 and CSSS reference assemblies, respectively. Distribution of mapped SNPs among the A, B, and D genomes in this study was similar to these in previous reports (Akhunov et al., 2010; Poland et al., 2012a; Würschum et al., 2013; Marcussen et al., 2014; Shavrukov et al., 2014; Edae et al., 2015) with most SNPs mapped to B genome followed by A genome and D genome (Chao et al., 2009; Berkman et al., 2013; Lai, 2015). D genome is the youngest one among the three genomes in wheat evolutionary history. It is likely that older genomes underwent gene duplication and accumulated more mutations that led to sequence polymorphism. Substantial early gene flow could have occurred between *T. aestivum* and its tetraploid progenitor *T. turgidum* (AABB) but not between the hexaploid and *Aegilops tauschii* (DD). This could have resulted in greater sequence diversity in the A and B genomes than in D genome (Talbert et al., 1998; Caldwell et al., 2004; Dvorak et al., 2006; Berkman et al., 2013). In this study, the number of SNPs that mapped to A or B genome were twice as many as those that mapped to the D genome. These results agree with several other studies (Akhunov et al., 2003; Chao et al., 2009, 2010; Wang et al., 2013; Edae et al., 2014; Iehisa et al., 2014) but contradict with some previous reports which indicated that five times more SNPs mapped to A or B genome than to the D genome (Allen et al., 2011, 2013; Cavanagh et al., 2013). This result suggests that Iranian wheat landraces may have relatively higher SNP variation in the D genome than other sources. Higher diversity in the D genome may provide new elite and desirable alleles controlling agriculturally important traits to deal with global climate and environment changes (Trethowan and Mujeeb-Kazi, 2008; Jia et al., 2013).

We observed an average transition/transversion (Ts/Tv) ratio of 1.75 for all three genomes based on both CSSS and

W7984 references (1.89 for A genome, 1.76 for B genome and 1.59 on D genome), which reflects the high frequencies of A to G and C to T mutations following methylation. D genome has the smallest Ts/Tv ratio among the three genomes. Higher frequency of transition mutations has been observed in several species including hexaploid wheat (Lorenc et al., 2012; Winfield et al., 2012; Manickavelu et al., 2014) and barley (Turuspekov et al., 2016) with Ts/Tv SNP ratios ranging from 1.59 to 2.12, which agrees with the current study. Transition abundance in many species may result from the mutation of methyl cytosine to uracil and then to thymine (Coulondre et al., 1978). Bread wheat genome is highly methylated due to the two rounds of polyploidy. Therefore, highly repeated sequences and abundant transitions can be considered an 'evolutionary footprint' of methylation (Buckler and Holtsford, 1996; Feldman and Levy, 2012; Kalinka et al., 2017). It has been demonstrated that loss of genes occurred more frequently in the A and B genomes than in the D genome (Berkman et al., 2013; Pont et al., 2013). The observed Ts/Tv bias in wheat provides a high level of confidence in SNP prediction accuracy because such a bias was unlikely caused by erroneously called SNPs due to errors in sequencing or mapping (Lorenc et al., 2012).

Chromosome 4D had the least SNPs while chromosome 3B had the highest using both CSSS and W7984 assemblies. On the other hand, a relatively high positive correlation was observed between the number of SNPs mapped to a chromosome and size of the chromosome, which agrees with previous studies (Poland et al., 2012a; Saintenac et al., 2013; Edae et al., 2015). In addition, other factors such as the time of evolution of a genome also affect the number of SNPs mapped on chromosomes.

Structure analysis placed 369 Iranian wheat accessions into three groups with only 7% of accessions misplaced into an opposite group. The most misplaced accessions were the cultivars that were placed into one of the landrace groups because most of these cultivars were selected from Iranian landraces. Therefore they should belong to landraces. Cluster analysis generated a similar grouping pattern. When pedigrees, geographical regions of cultivation, years of accession released, growth habits and origins of cultivars were analyzed in cluster analysis, we found that accession pedigree is the main factor for separation of the Iranian cultivars. Most cultivars that originated from Iran or

**TABLE 4** | Analysis of molecular variance (AMOVA) result from the diversity panel of 369 Iranian hexaploid wheat landraces and cultivars.

	Source of variation	df	MS	Estimated variance	%	*F statistics (Fst)	Probability
Landraces vs. cultivars	Among population	1	6693.47	22.39	15.56	0.16	0.0000
	Among individuals within populations	367	218.02	96.74	67.36		
	Within individuals	369	24.54	24.54	17.08		
	Total of variation	737	129.54	143.63	100.00		
Structure grouping	Among population	2	7765.30	32.66	22.98	0.23	0.0000
	Among individuals within populations	366	194.47	84.96	59.76		
	Within population	369	24.54	24.54	17.26		
	Total of variation	737	129.93	142.17	100.00		

df, degrees of freedom; MS, mean of squares. \*Index considered as standardized variance of allele frequencies among subdivisions.

with one parent of Iranian wheat were clearly separated from those which originated from CIMMYT, suggesting that Iranian wheat may have a different genetic makeup from the CIMMYT wheat. That may explain why crossing made between Iranian and CIMMYT wheat genotypes produced high yielding cultivars. For instance, Pishgam and Parsi derived from a cross between Iranian and CIMMYT lines are currently the most widely planted cultivars in Iran (Mahfoozi et al., 2009; Amin et al., 2010). For landraces, a high level of genetic diversity was observed among those collected in different years, or from different geographical regions and climate zones. Genetic variation among the landraces collected from north of Iran was higher than the landraces from the south. For example, most of the landraces collected from Gilan and Mazandaran were grouped closely to each other. Landrace collection years can also separate the landraces into two groups based on whether they were collected before and after 1960. This may be attributed to the breeders at University of Tehran who initiated the purification of landraces after 1960. Therefore, genetic variability among the landraces can be affected by the years of collection, anthropogenic impact through dynamic storage practices, diverse needs, end-uses, seed exchange (gene flow) between farmers, and geographic and environmental conditions where wheat grows. It has been documented that genetic diversity of landraces could be related to their adaptability (Cooper et al., 2001) and farmers have played a pivot role in maintaining the genetic diversity (Zeven, 2002).

Genetic diversity for each of the predefined subpopulations was measured using Nei's and Shannon's genetic diversity indices, private alleles, the percentage of polymorphic loci and PIC in this study. The Landrace Group I showed higher values of Nei's and Shannon's gene diversity indices and the Cultivar Group demonstrated a lower average of private alleles and percentage of polymorphic loci than the two landrace subgroups. This result highlights the strong genetic separation among the groups and specific adaptation of the Iranian cultivars captured in the founder lines of the landraces. But whether the proportion of the genetic differentiation in the subgroups of Iranian landraces and cultivars are due to differentiation during different domestication events or introduction from CIMMYT remains unknown. It is possible that a single domestication event occurred with various hybridizations to endemic genotypes in each region, as has occurred in the domestication of rice (Londo et al., 2006; Huang et al., 2012). Further, the significant genetic differentiation ( $p < 0.001$ ) among the three groups as illustrated by pairwise  $F_{ST}$  and AMOVA analyses verified the differentiation of three groups in this panel. However, the Landrace Group I demonstrated a higher gene flow and lower  $F_{ST}$  value with the Cultivar Group, suggesting the Landrace Group I might play a bridge role between the Landrace Group II and the Cultivar Group. Most accessions in Landrace Group I originated from cold or moderate and rainy climates, and had been collected before 1960 while those in Landrace Group II were mainly collected from semiarid

or hot and dry climates after 1960. Since the landraces were collected from diverse climate regions and altitudes, many of these germplasm lines should be useful sources of genes to be used in breeding to address the challenge of climate change.

## CONCLUSION

Our studies demonstrated that GBS is a powerful tool for investigating population structure and genetic diversity of wheat landraces and cultivars. In this study, Iranian hexaploid wheat landraces and cultivars collected and released from different years, agro-climatic zones with different growth habits were grouped into three distinct groups. Iranian cultivars can be separated into four major groups with one group was mostly originated from Iranian ancestor(s). Large genetic variation was observed among landraces based on the years of collection, geographical and climate zones. We hope that this genetic diversity will help wheat breeders to select parents for crossing to improve wheat under different climate conditions in Iran and other countries.

## AUTHOR CONTRIBUTIONS

HA analyzed the data and wrote the paper. GB contributed reagents, Ion Proton sequencing and analysis tools. MB, VM, and SP provided seeds of the genotypes. HA, MB, VM, SP, GB, and GZ contributed to conceive the project and revise the paper. All authors read and approved the final manuscript.

## ACKNOWLEDGMENTS

This is contribution number 17-287-J from the Kansas Agricultural Experiment Station. This project is partly funded by the National Research Initiative Competitive Grants CAP project 2011-68002-30029 from the USDA National Institute of Food and Agriculture; and University of Tehran and Ministry of Science, Research and Technology, Iran. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. USDA is an equal opportunity provider and employer.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.01293/full#supplementary-material>

**TABLE S1** | List and information of the Iranian wheat landraces and cultivars.

## REFERENCES

- Akhunov, E. D., Akhunova, A. R., Anderson, O. D., Anderson, J. A., Blake, N., Clegg, M. T., et al. (2010). Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. *BMC Genomics* 11:702. doi: 10.1186/1471-2164-11-702
- Akhunov, E. D., Goodyear, A. W., Geng, S., Qi, L. L., Echalié, B., Gill, B. S., et al. (2003). The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. *Genome Res.* 13, 753–763. doi: 10.1101/gr.808603
- Allen, A. M., Barker, G. L., Berry, S. T., Coghill, J. A., Gwilliam, R., Kirby, S., et al. (2011). Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.). *Plant Biotechnol. J.* 9, 1086–1099. doi: 10.1111/j.1467-7652.2011.00628.x
- Allen, A. M., Barker, G. L., Wilkinson, P., Burridge, A., Winfield, M., Coghill, J., et al. (2013). Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat (*Triticum aestivum* L.). *Plant Biotechnol. J.* 11, 279–295. doi: 10.1111/pbi.12009
- Amin, H., Pazhmand, M. E., Dadaeen, M., Zakeri, A., Yasaie, M., Rajaie, S., et al. (2010). Parsi, a new bread wheat cultivar, resistant to stem rust (race Ug99) with good bread making quality for cultivation under irrigated conditions of temperate regions of Iran. *Seed Plant Improv. J.* 26, 289–292.
- Batley, J., and Edwards, D. (2007). “SNP applications in plants,” in *Association Mapping in Plants*, eds N. C. Oraguzie, E. H. A. Rikkerink, S. E. Gardiner, and H. N. De Silva (New York, NY: Springer), 95–102. doi: 10.1007/978-0-387-36011-9\_6
- Berkman, P. J., Visendi, P., Lee, H. C., Stiller, J., Manoli, S., Lorenc, M. T., et al. (2013). Dispersion and domestication shaped the genome of bread wheat. *Plant Biotechnol. J.* 11, 564–571. doi: 10.1111/pbi.12044
- Bonman, J. M., Babiker, E. M., Cuesta-Marcos, A., Esvelt-Klos, K., Brown-Guedira, G., Chao, S., et al. (2015). Genetic diversity among wheat accessions from the USDA national small grains collection. *Crop Sci.* 55, 1243–1253. doi: 10.2135/cropsci2014.09.0621
- Borlaug, N. E. (1968). *Wheat Breeding and Its Impact on World Food Supply*. Canberra, ACT: Australian Academy of Science, 1335.
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308
- Buckler, E., and Holtsford, T. P. (1996). Zea ribosomal repeat evolution and substitution patterns. *Mol. Biol. Evol.* 13, 623–632. doi: 10.1093/oxfordjournals.molbev.a025622
- Caldwell, K. S., Dvorak, J., Lagudah, E. S., Akhunov, E., Luo, M. C., Wolters, P., et al. (2004). Sequence polymorphism in polyploid wheat and their d-genome diploid ancestor. *Genetics* 167, 941–947. doi: 10.1534/genetics.103.016303
- Cavanagh, C. R., Chao, S., Wang, S., Huang, B. E., Stephen, S., Kiani, S., et al. (2013). Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proc. Natl. Acad. Sci. U.S.A.* 110, 8057–8062. doi: 10.1073/pnas.1217133110
- Chao, S., Dubcovsky, J., Dvorak, J., Luo, M.-C., Baenziger, S. P., Matnyazov, R., et al. (2010). Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (*Triticum aestivum* L.). *BMC genomics* 11:727. doi: 10.1186/1471-2164-11-727
- Chao, S., Zhang, W., Akhunov, E., Sherman, J., Ma, Y., Luo, M.-C., et al. (2009). Analysis of gene-derived SNP marker polymorphism in US wheat (*Triticum aestivum* L.) cultivars. *Mol. Breed.* 23, 23–33. doi: 10.1007/s11032-008-9210-6
- Chapman, J. A., Mascher, M., Buluc, A., Barry, K., Georganas, E., Session, A., et al. (2015). A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. *Genome Biol.* 16:26. doi: 10.1186/s13059-015-0582-8
- Chen, H., Martin, J., Lavin, M., and Talbert, L. (1994). Genetic diversity in hard red spring wheat based on sequence-tagged-site PCR markers. *Crop Sci.* 34, 1628–1632. doi: 10.2135/cropsci1994.0011183X003400060037x
- Ciaffi, M., Dominici, L., Lafandra, D., and Porceddu, E. (1992). Seed storage proteins of wild wheat progenitors and their relationships with technological properties. *Hereditas* 116, 315–322. doi: 10.1111/j.1601-5223.1992.tb00844.x
- Cooper, H. D., Spillane, C., and Hodgkin, T., (eds) (2001). “Broadening the genetic base of crops: an overview,” in *Broadening the Genetic Base of Crop Production*, (Wallingford: CAB International), 1–23. doi: 10.1079/9780851994116.0001
- Coulondre, C., Miller, J. H., Farabaugh, P. J., and Gilbert, W. (1978). Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature* 274, 775–780. doi: 10.1038/274775a0
- Dreisigacker, S., Zhang, P., Warburton, M., Skovmand, B., Hoisington, D., and Melchinger, A. (2005). Genetic diversity among and within CIMMYT wheat landrace accessions investigated with SSRs and implications for plant genetic resources management. *Crop Sci.* 45, 653–661. doi: 10.2135/cropsci2005.0653
- Dvorak, J., Akhunov, E. D., Akhunov, A. R., Deal, K. R., and Luo, M.-C. (2006). Molecular characterization of a diagnostic DNA marker for domesticated tetraploid wheat provides evidence for gene flow from wild tetraploid wheat to hexaploid wheat. *Mol. Biol. Evol.* 23, 1386–1396. doi: 10.1093/molbev/msl004
- Eadae, E. A., Bowden, R. L., and Poland, J. (2015). Application of population sequencing (POPSEQ) for ordering and imputing genotyping-by-sequencing markers in hexaploid wheat. *G3* 5, 2547–2553. doi: 10.1534/g3.115.020362
- Eadae, E. A., Byrne, P. F., Haley, S. D., Lopes, M. S., and Reynolds, M. P. (2014). Genome-wide association mapping of yield and yield components of spring wheat under contrasting moisture regimes. *Theor. Appl. Genet.* 127, 791–807. doi: 10.1007/s00122-013-2257-8
- Ehdaie, B., and Baker, C. (1999). Inheritance and allelism for resistance to Russian wheat aphid in an Iranian spring wheat. *Euphytica* 107, 71–78. doi: 10.1023/A:1003549512216
- Ehdaie, B., Waines, J., and Hall, A. (1988). Differential responses of landrace and improved spring wheat genotypes to stress environments. *Crop Sci.* 28, 838–842. doi: 10.2135/cropsci1988.0011183X002800050024x
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., 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
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x
- Excoffier, L., and Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x
- FAO (2009). *Global Agriculture Towards 2050*. Available at: [http://www.fao.org/fileadmin/templates/wsfs/docs/Issues\\_papers/HLEF2050\\_Global\\_Agriculture.pdf](http://www.fao.org/fileadmin/templates/wsfs/docs/Issues_papers/HLEF2050_Global_Agriculture.pdf)
- FAO (2015). *FAOSTAT, Statistics Division* FAO. Available at: <http://faostat3.fao.org/home/E>
- Feldman, M., and Levy, A. A. (2012). Genome evolution due to allopolyploidization in wheat. *Genetics* 192, 763–774. doi: 10.1534/genetics.112.146316
- Govindaraj, M., Vetriventhan, M., and Srinivasan, M. (2015). Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. *Genet. Res. Int.* 2015:431487. doi: 10.1155/2015/431487
- Hao, C., Dong, Y., Wang, L., You, G., Zhang, H., Ge, H., et al. (2008). Genetic diversity and construction of core collection in Chinese wheat genetic resources. *Chin. Sci. Bull.* 53, 1518–1526. doi: 10.1007/s11434-008-0212-x
- Hao, C., Zhang, X., Wang, L., Dong, Y., Shang, X., and Jia, J. (2006). Genetic diversity and core collection evaluations in common wheat germplasm from the Northwestern Spring Wheat Region in China. *Mol. Breed.* 17, 69–77. doi: 10.1007/s11032-005-2453-6
- He, J., Zhao, X., Laroche, A., Lu, Z. X., Liu, H., and Li, Z. (2014). Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Front. Plant Sci.* 5:484. doi: 10.3389/fpls.2014.00484
- Hoisington, D., Khairallah, M., Reeves, T., Ribaut, J.-M., Skovmand, B., Taba, S., et al. (1999). Plant genetic resources: what can they contribute toward increased crop productivity? *Proc. Natl. Acad. Sci. U.S.A.* 96, 5937–5943. doi: 10.1073/pnas.96.11.5937
- Huang, X., Börner, A., Röder, M., and Ganal, M. (2002). Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. *Theor. Appl. Genet.* 105, 699–707. doi: 10.1007/s00122-002-0959-4
- Huang, X., Kurata, N., Wei, X., Wang, Z.-X., Wang, A., Zhao, Q., et al. (2012). A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490, 497–501. doi: 10.1038/nature11532

- Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., et al. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42, 961–967. doi: 10.1038/ng.695
- Huang, Y.-F., Poland, J. A., Wight, C. P., Jackson, E. W., and Tinker, N. A. (2014). Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. *PLoS ONE* 9:e102448. doi: 10.1371/journal.pone.0102448
- Iehisa, J. C. M., Shimizu, A., Sato, K., Nishijima, R., Sakaguchi, K., Matsuda, R., et al. (2014). Genome-wide marker development for the wheat D genome based on single nucleotide polymorphisms identified from transcripts in the wild wheat progenitor *Aegilops tauschii*. *Theor. Appl. Genet.* 127, 261–271. doi: 10.1007/s00122-013-2215-5
- International Wheat Genome Sequencing [IWGS] (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* 345:1251788. doi: 10.1126/science.1251788
- Jafari-Shabestari, J., Corke, H., and Qalseth, C. O. (1995). Field evaluation of tolerance to salinity stress in Iranian hexaploid wheat landrace accessions. *Genet. Resour. Crop Evol.* 42, 147–156. doi: 10.1007/BF02539518
- Jia, J., Zhao, S., Kong, X., Li, Y., Zhao, G., He, W., et al. (2013). *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature* 496, 91–95. doi: 10.1038/nature12028
- Kalinka, A., Achrem, M., and Poter, P. (2017). The DNA methylation level against the background of the genome size and t-heterochromatin content in some species of the genus *Secale L. PeerJ* 5:e2889. doi: 10.7717/peerj.2889
- Kertho, A., Mamidi, S., Bonman, J. M., McClean, P. E., and Acevedo, M. (2015). Genome-wide association mapping for resistance to leaf and stripe rust in winter-habit hexaploid wheat landraces. *PLoS ONE* 10:e0129580. doi: 10.1371/journal.pone.0129580
- Kumar, S., Banks, T. W., and Cloutier, S. (2012). SNP discovery through next-generation sequencing and its applications. *Int. J. Plant Genomics* 2012:831460. doi: 10.1155/2012/831460
- Lai, K. (2015). *Genome Diversity in Triticum aestivum*. Ph.D. thesis, The University of Queensland, St Lucia, QLD.
- Londo, J. P., Chiang, Y.-C., Hung, K.-H., Chiang, T.-Y., and Schaal, B. A. (2006). Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9578–9583. doi: 10.1073/pnas.0603152103
- Lorenc, M. T., Hayashi, S., Stiller, J., Lee, H., Manoli, S., Ruperao, P., et al. (2012). Discovery of single nucleotide polymorphisms in complex genomes using SGSautoSNP. *Biology* 1, 370–382. doi: 10.3390/biology1020370
- Lu, F., Lipka, A. E., Glaubitz, J., Elshire, R., Cherney, J. H., Casler, M. D., et al. (2013). Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *PLoS Genet.* 9:e1003215. doi: 10.1371/journal.pgen.1003215
- Mahfoozi, S., Akbari, A., Chaichi, M., Sanjari, A. G., Nazeri, S. M., Abedi-Oskooee, S., et al. (2009). Pishgam, a new bread wheat cultivar for normal irrigation and terminal stage deficit irrigation conditions of cold regions of Iran. *Seed Plant Improv. J.* 25, 513–517.
- Manickavelu, A., Jighly, A., and Ban, T. (2014). Molecular evaluation of orphan Afghan common wheat (*Triticum aestivum* L.) landraces collected by Dr. Kihara using single nucleotide polymorphic markers. *BMC Plant Biol.* 14:320. doi: 10.1186/s12870-014-0320-5
- Marcusson, T., Sandve, S. R., Heier, L., Spannagl, M., Pfeifer, M., International Wheat Genome Sequencing Consortium, et al. (2014). Ancient hybridizations among the ancestral genomes of bread wheat. *Science* 345:1250092. doi: 10.1126/science.1250092
- Newcomb, M., Acevedo, M., Bockelman, H. E., Brown-Guedira, G., Goates, B. J., Jackson, E. W., et al. (2013). Field resistance to the Ug99 race group of the stem rust pathogen in spring wheat landraces. *Plant Dis.* 97, 882–890. doi: 10.1094/pdis-02-12-0200-re
- Nielsen, N. H., Backes, G., Stougaard, J., Andersen, S. U., and Jahoor, A. (2014). Genetic diversity and population structure analysis of European hexaploid bread wheat (*Triticum aestivum* L.) varieties. *PLoS ONE* 9:e94000. doi: 10.1371/journal.pone.0094000
- Olmstead, A. L., and Rhode, P. W. (2002). The red queen and the hard reds: productivity growth in American wheat, 1800–1940. *I. J. Econ. Hist.* 62, 929–966. doi: 10.1017/s0022050702001602
- Perrier, X., and Jacquemoud-Collet, J. (2006). *DARwin: Dissimilarity Analysis and Representation for Windows*. 5.0.158. Montpellier: CIRAD.
- Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J.-L. (2012a). 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
- Poland, J. A., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., et al. (2012b). Genomic selection in wheat breeding using genotyping-by-sequencing. *Plant Genome* 5, 103–113. doi: 10.3835/plantgenome2012.06.0006
- Poland, J. A., and Rife, T. W. (2012). Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome* 5, 92–102. doi: 10.3835/plantgenome2012.05.0005
- Pont, C., Murat, F., Guizard, S., Flores, R., Fouchier, S., Bidet, Y., et al. (2013). Wheat syntenome unveils new evidences of contrasted evolutionary plasticity between paleo- and neoduplicated subgenomes. *Plant J.* 76, 1030–1044. doi: 10.1111/tpj.12366
- Pritchard, J., Wen, X., and Falush, D. (2010). *Documentation for Structure Software: Version 2.3*. Chicago, IL: University of Chicago.
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Ray, D. K., Mueller, N. D., West, P. C., and Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS ONE* 8:e66428. doi: 10.1371/journal.pone.0066428
- Rouse, M. N., Olson, E. L., Gill, B. S., Pumphrey, M. O., and Jin, Y. (2011). Stem rust resistance in germplasm. *Crop Sci.* 51, 2074–2078. doi: 10.2135/cropsci2010.12.0719
- Saghai-Marouf, M. A., Soliman, K. M., Jorgensen, R. A., and Allard, R. (1984). Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 81, 8014–8018. doi: 10.1073/pnas.81.24.8014
- Saintenac, C., Jiang, D., Wang, S., and Akhunov, E. (2013). Sequence-based mapping of the polyploid wheat genome. *G3* 3, 1105–1114. doi: 10.1534/g3.113.005819
- Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.
- Sener, O., Arslan, M., Soysal, Y., and Erayman, M. (2009). Estimates of relative yield potential and genetic improvement of wheat cultivars in the Mediterranean region. *J. Agric. Sci.* 147, 323–332. doi: 10.1017/S0021859609008454
- Shavrukov, Y., Suchecki, R., Eliby, S., Abugalieva, A., Kenebayev, S., and Langridge, P. (2014). Application of next-generation sequencing technology to study genetic diversity and identify unique SNP markers in bread wheat from Kazakhstan. *BMC Plant Biol.* 14:258. doi: 10.1186/s12870-014-0258-7
- Skovmand, B., Rajaram, S., Ribaut, J., and Hede, A. (2002). *Wheat Genetic Resources. Plant Production and Protection Series*. Rome: FAO.
- Sofalian, O., Chaparzadeh, N., Javanmard, A., and Hejazi, M. (2008). Study the genetic diversity of wheat landraces from northwest of Iran based on ISSR molecular markers. *Int. J. Agric. Biol.* 10, 466–468.
- Talbert, L., Smith, L., and Blake, N. (1998). More than one origin of hexaploid wheat is indicated by sequence comparison of low-copy DNA. *Genome* 41, 402–407. doi: 10.1139/g98-037
- Toit, D. F. (1989). Components of resistance in three bread wheat lines to Russian wheat aphid (*Homoptera: Aphididae*) in South Africa. *J. Econ. Entomol.* 82, 1779–1781. doi: 10.1093/jee/82.6.1779
- Trethowan, R., and Mujeeb-Kazi, A. (2008). Novel germplasm resources for improving environmental stress tolerance of hexaploid wheat. *Crop Sci.* 48, 1255–1265. doi: 10.2135/cropsci2007.08.0477
- Turuspekoy, Y., Ormanbekova, D., Rsaliev, A., and Abugalieva, S. (2016). Genome-wide association study on stem rust resistance in Kazakh spring barley lines. *BMC Plant Biol.* 16(Suppl. 1):6. doi: 10.1186/s12870-015-0686-z
- Valdez, V. A., Byrne, P. F., Lapitan, N. L., Peairs, F. B., Bernardo, A., Bai, G., et al. (2012). Inheritance and genetic mapping of Russian wheat aphid resistance in Iranian wheat landrace accession PI 626580. *Crop Sci.* 52, 676–682. doi: 10.2135/cropsci2011.06.0331
- Wang, N., Thomson, M., Bodes, W. J., Crawford, R. M., Hunt, H. V., Featherstone, A. W., et al. (2013). Genome sequence of dwarf birch (*Betula nana*) and cross-species RAD markers. *Mol. Ecol.* 22, 3098–3111. doi: 10.1111/mec.12131
- Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183

- Winfield, M. O., Wilkinson, P. A., Allen, A. M., Barker, G. L., Coghill, J. A., Burrige, A., et al. (2012). Targeted re-sequencing of the allohexaploid wheat exome. *Plant Biotechnol. J.* 10, 733–742. doi: 10.1111/j.1467-7652.2012.00713.x
- Würschum, T., Langer, S. M., Longin, C. F. H., Korzun, V., Akhunov, E., Ebmeyer, E., et al. (2013). Population structure, genetic diversity and linkage disequilibrium in elite winter wheat assessed with SNP and SSR markers. *Theor. Appl. Genet.* 126, 1477–1486. doi: 10.1007/s00122-013-2065-1
- Zeven, A. C. (1998). Landraces: a review of definitions and classifications. *Euphytica* 104, 127–139. doi: 10.1023/A:1018683119237
- Zeven, A. C. (2002). Traditional maintenance breeding of landraces: 2. Practical and theoretical considerations on maintenance of variation of landraces by farmers and gardeners. *Euphytica* 123, 147–158. doi: 10.1023/A:1014940623838

**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.

Copyright © 2017 Alipour, Bihanta, Mohammadi, Peyghambari, Bai and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.