ISSN 1479-2621
Plant Genetic Resources: Characterization and Utilization (2015) **13**(1); 1–8
doi:10.1017/S1479262114000549

Genetic structure and diversity of *Triticum monococcum* ssp. *aegilopoides* and *T. urartu* in Iran

Fatemeh Nasernakhaei, Mohammad Reza Rahiminejad*, Hojjatollah Saeidi and Manoochehr Tavassoli

Department of Biology, Faculty of Sciences, University of Isfahan, Isfahan, Iran

Received 17 December 2013; Accepted 22 February 2014 – First published online 8 April 2014

Abstract

To preliminarily evaluate the genetic diversity of the Iranian diploid *Triticum* L. gene pool, in this study, a total of 176 individuals belonging to *T. monococcum* L. ssp. *aegilopoides* (Link) Thell. and *T. urartu* Thum. ex Gandil. were pre-screened using single-strand conformation polymorphism (SSCP) analysis of the *Acc-1* and *Pgk-1* loci. A selected set of 76 DNA samples corresponding to the observed SSCP variants were sequenced for both loci and evaluated for nucleotide diversity associated with the taxonomic groups and geographical regions. We found three haplotypes, including one that was new for Iran, at each locus. Population structure and analysis of molecular variation results proved that the collection evaluated could be genetically divided into two distinct groups, which to a great extent was in accordance with the taxonomic classification. A genetic leakage from *T. monococcum* ssp. *aegilopoides* into *T. urartu* was observed during structure analysis, which was inferred on the basis of occasional outcrossing, despite their inbreeding nature. The results revealed that there was no variation within the populations belonging to *T. urartu*, while a meaningful variation was found between the geographical regions for *T. monococcum* ssp. *aegilopoides*. The median-joining networks revealed a conflict between morphology and haplotype variation, which was interpreted on the basis of introgressive hybridization, recombination signature and rapid speciation. In conclusion, we suggest that SSCP analysis is a useful tool in regions where thorough sequencing of an enormous number of DNA samples is time consuming and not affordable.

Keywords: *Acc-1*; diploid; genetic diversity; *Pgk-1*; population structure; *Triticum*

Introduction

It is generally acknowledged that wheat was domesticated across the Near East about 10,000 years ago (Nesbitt, 2001). The putative donor ancestors of wheat are still growing in this area and have been accumulating genetic diversity more than what had been inherited by cultivated *Triticum* taxa (Chantret *et al*., 2005; Haudry *et al*., 2007). These genetic resources are highly desirable and precious for improving the wheat gene pool and maintaining sustainable agroecosystems that are endangered by increasing climatic changes (Maxted *et al*., 2008). Therefore, the natural genetic resources of diploid *Triticum* including *T. monococcum s. lat.* and *T. urartu* have been frequently studied (Heun *et al*., 1997; Kilian *et al*., 2007b; Dvorak *et al*., 2011; Peng *et al*., 2011) in the Fertile Crescent, which is well known as the origin of wheat domestication. However, the above-mentioned gene pool in one of the oldest centres of wheat domestication, i.e. the Iranian part of the Fertile Crescent (Riehl *et al*., 2012, 2013) and areas very close to it, although not

* Corresponding author. E-mail: mrr@sci.ui.ac.ir

completely but to a great extent, has been ignored and can be considered as the lost ring of the chain of investigations.

Recently, DNA sequencing of nuclear gene loci, including *Acc-1* and *Pgk-1*, which are the focus of the present study, has emerged as an important tool for assessing the inter- and intra-relationships among the Triticeae taxa (Huang *et al.*, 2002a, b; Fan *et al.*, 2007; Kilian *et al.*, 2007a, b; Takenaka *et al.*, 2010). *Acc-1* and *Pgk-1* are predominantly known as single-copy genes in grasses and encode plastid acetyl-CoA carboxylase and plastid 3-phosphoglycerate kinase, respectively (Huang *et al.*, 2002a, b). The significant role played by these genes in the ecogenetic balance of plants (Longstaff *et al.*, 1989; Sasaki and Nagano, 2004) has made their sequencing data important for determining plant genetic structure and relationships across geographical areas (Avise, 2000). The variability of these genes was fundamentally screened in a vast area by Kilian *et al.* (2007b) and Golovnina *et al.* (2009), albeit the Iranian part has mostly remained untouched.

The main goal of this study was to preliminarily evaluate the molecular diversity of the wild diploid *Triticum* in Iran, using two functionally important genes, *Acc-1* and *Pgk-1*. In addition, the genetic structure and geographical distribution patterns of the above-mentioned loci were the focus of this study.

Materials and methods

Plant materials

In total, 176 individuals collected from ten locations across the distribution range of *T. urartu* and *T. monococcum* ssp. *aegilopoides* (Table S1, available online) were evaluated. The materials used in this study are deposited in the herbarium of the University of Isfahan, Iran. Taking into account the inbreeding nature of wheat and the topography and coordinates of the area under study, we subdivided the distribution range of the *Triticum* collection evaluated into three geographical regions, each representing one population (Fig. 1). The taxonomic identifications were made on the basis of the work of van Slageren (1994) and Nasernakhaei *et al.* (2013).

DNA extraction and polymerase chain reaction (PCR) amplification

DNA extraction was carried out using fresh leaves of a single plant following the Cetyltrimethylammonium bromide method of Gawel and Jarret (1991). The *Acc-1* and *Pgk-1* loci were amplified with the specific primers

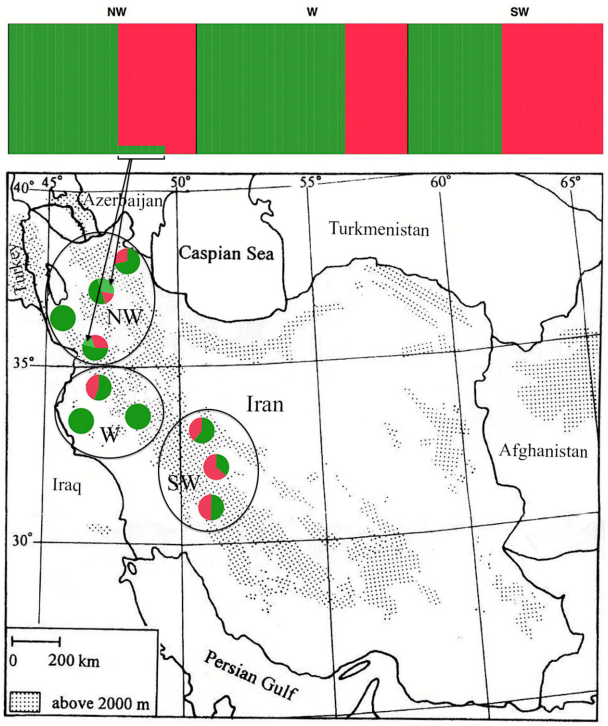

Fig. 1. Map of the collection sites, distribution of *Acc-1* and *Pgk-1* haplotypes, and population structure of the 76 sequenced samples belonging to the Iranian wild diploid *Triticum* gene pool. The coloured circles indicate *Acc-1* and *Pgk-1* haplotype members belonging to ten locations grouped into three geographical regions (North-west (NW), West (W) and South-west (SW)). The colour of each haplotype is the same as that of the genetic clusters identified using STRUCTURE. Haplotype A (green) and haplotype B (light green) are *Triticum monococcum* ssp. *aegilopoides* and haplotype C (red) is *Triticum urartu*. Structure analysis results revealed two clusters (green: cluster 1 and red: cluster 2) for three regions for the wild diploid *Triticum* gene pool. *T. monococcum* ssp. *aegilopoides* is completely made up by cluster 1, while *T. urartu* has contribution mainly from cluster 2, but it also has a small contribution from cluster 1. Each vertical column represents one individual.

Acc1T (2T)s (5′-GGA CTT AGT TTT TTG TCG TCA GTT-3′), Acc1Ta new (5′-CTT CCA AAC GTA AGG ACC AAT ACA-3′), Pgk4Ts (5′-GCT TGG CTC CCC TTG TGC CCC G-3′) and Pgk1Ta (5′-CAC ACT TCT CCA GCA GGG ATT CGA-3′) (Golovnina *et al.*, 2009). The PCR procedure was optimized according to thermal and $MgCl_2$ concentration gradient tests. DNA amplifications were carried out in a 25 µl reaction volume containing 40–50 ng of genomic DNA, 200 µM of each dNTPs, 0.1 µM of each primer, 2.5 µl of 10× buffer, 3 mM $MgCl_2$ and 1 U Taq DNA polymerase. The PCR program for each of the specific primers was implemented at the optimized conditions as follows: one cycle of 5 min at 94°C and 33 cycles of 45 s at 94°C, 45 s at 61°C (*Acc-1*)

https://doi.org/10.1017/S1479262114000549 Published online by Cambridge University Press

or at 67.5°C (*Pgk-1*) and 1 min at 72°C, followed by 10 min at 72°C.

Pre-screening for polymorphisms by non-denaturing gel electrophoresis

PCR products were denatured for 5 min at 100°C and characterized by single-strand conformation polymorphism (SSCP) analysis, as described by Rodriguez *et al.* (2011). Electrophoresis was carried out at 300 V for 19–21 h at 8°C and visualized by silver staining as described by Sanguinetti *et al.* (1994).

Sequencing of PCR products

Based on the revealed SSCP patterns for each of the loci, a subset of 76 of the 176 individuals representing three geographical areas (Fig. 1) were sequenced in both directions (Bioneer, Korea).

Data analysis

The sequence data were aligned manually using BioEdit version 7 (Hall, 1999). The sequences were edited using ChromasPro version 1.41 (Technelysium Pty Ltd, Tewantin, Australia). For each species, the number of segregating sites, the number of haplotypes, the haplotype diversity (H_d; Nei, 1987) and the nucleotide diversity (P_i; Nei, 1987) were calculated using DnaSP version 5 (Liberado and Rozas, 2009). To check the concordance between taxonomic status and genetic structure, the combined haplotype data of *Acc-1* and *Pgk-1* were screened using a Bayesian, model-based clustering algorithm implemented in STRUCTURE version 2.3.4 (Pritchard *et al.*, 2000). File format conversion was done using PGDSpider version 2.0.5.1 (Lischer and Excoffier, 2012). The analysis was run for the range of genetic clusters from $K = 1$ to $K = 10$ with the admixture, LOCPRIOR and allele frequency-independent models, and it was replicated ten times for each K value. Each run was implemented with a burn-in period of 10,000 steps followed by 100,000 Markov chain Monte Carlo replicates. The most optimal value for K based on Evanno's ΔK (Evanno *et al.*, 2005) was selected using STRUCTURE HARVESTER (Earl and vonHoldt, 2012), and graphical representations of population assignments were obtained using CLUMPP (Jakobsson and Rosenberg, 2007) and DISTRUCT (Rosenberg, 2004).

Population structure was also investigated via analysis of molecular variation (AMOVA; Excoffier *et al.*, 1992), which was used to hierarchically partition genetic variation using 1000 permutations, and was calculated using ARLEQUIN version 3.1 (Excoffier *et al.*, 2005). The hierarchies were tested 'between species', 'among populations (different regions) within the species' and 'within populations' in the entire dataset. Subsequently, AMOVA models were analysed to test the distribution of genetic variance 'among' and 'within' the populations of *T. monococcum* ssp. *aegilopoides* and *T. urartu* separately.

The median-joining networks (Bandelt *et al.*, 1999) representing the relationships among the Iranian wild diploid *Triticum* haplotypes evaluated in this study and those taken from the NCBI (Kilian *et al.*, 2007a, b; Golovnina *et al.*, 2009) were drawn gene by gene using Network version 4.6.1.1 based on a maximum parsimony approach.

Results

Genetic variation

Totally, 791 and 679 bp were acquired for the *Acc-1* and *Pgk-1* loci, respectively. In sum, 44 sequences were deposited for both loci in the GenBank (accession numbers KC965047–KC965055 and KF023037–KF023071). All together, six haplotypes, i.e. three for each of the loci, were found and similarly named as A, B and C (Table 1). In both loci, haplotype B was new for Iran. The *Acc-1* and *Pgk-1* loci exhibited four *versus* 29 nucleotide substitutions and two *versus* four insertions/deletions (indels), respectively (Tables 2 and 3). Haplotype A (45/76), vastly distributed in an area all along the Zagros Mountains to Azerbaijan provinces, and haplotype B, restricted to only six individuals collected from North-west of Iran (6/76), of *T. monococcum* ssp. *aegilopoides* exhibited the maximum and minimum frequencies, respectively, and haplotype C of *T. urartu* exhibited a frequency of 25/76 (Table 1). Geographically, North-west of Iran representing all the three haplotypes appeared to be the most variable zone (Table 1 and Fig. 1). The greatest genetic diversity ($H_d = 0.416 \pm 0.090$ and $P_i = 0.007 \pm 0.001$) was found in the population belonging to *T. monococcum* ssp. *aegilopoides* from North-west of Iran (Table S2, available online), while genetic diversity in populations belonging

Table 1. *Acc-1* and *Pgk-1* haplotypes and their frequencies among the 76 sequenced samples belonging to the wild diploid *Triticum* gene pool in three geographical regions in Iran (see Fig. 1)

Haplotypes	North-west	West	South-west	Total
A	14	19	12	45
B	6	–	–	6
C	4	8	13	25
Sample size	24	27	25	76

https://doi.org/10.1017/S1479262114000549 Published online by Cambridge University Press

Table 2. Nucleotide sequence variability of *Acc-1* (791 bp) in haplotypes A and B (*Triticum monococcum* ssp. *aegilopoides*) and haplotype C (*Triticum urartu*)[a]

Haplo-types	108	210	211	324	581–626	631	664
	e	i	i	e	i	i	i
A	A	–	–	A	A G A A A T T G G G C T G C G T T T T T T C G A A T T A G T T T C C A T T C G G G C A C
B	T	T	A	T	– G T
C	T	T	A	A	– A T

–, i and e correspond to insertion/deletion (indels), intron and exon, respectively.
[a]Only polymorphic sites are indicated.

to the other regions and also *T. urartu* was null ($H_d = 0$ and $P_i = 0$).

Population structure

Based on ΔK developed by Evanno *et al.* (2005), a sharp peak in ΔK was observed at $K = 2$, clearly dividing the collection evaluated into two clusters (Fig. 1), which was fully consistent with the diagnostic morphological features of *T. monococcum* ssp. *aegilopoides* and *T. urartu* (van Slageren, 1994; Nasernakhaei *et al.*, 2013). The probable *K* value could not be inferred based on the Ln*P*(*D*) value because there was no sudden change in it. *T. monococcum* ssp. *aegilopoides* was completely made up by cluster 1, while *T. urartu* had contribution mainly from cluster 2, but it also had a small contribution from cluster 1. The relationship between haplotype B (in light green) and the clusters is shown in Fig. 1.

AMOVA carried out using a concatenated sequence of both loci indicated that the highest variation (86.28%) occurred 'between species' (Table 4). Within species, most of the variation in *T. monococcum* ssp. *aegilopoides* was observed within regions (76%), though the overall F_{ST} value among regions was high ($F_{ST} = 0.24$). No variation was detected for *T. urartu*.

Haplotype networks (Fig. 2) visualized a global relationship combined of our materials and the relevant haplotype data taken from the NCBI for both loci.

Discussion

Genetic variation and population structure

The genetic diversity encountered among the Iranian wild einkorn wheat populations was screened by several researchers. Nejat-Boshehri and Fakhr Tabatabaei (2001) demonstrated high variability among the seed storage protein profiles. Maleki *et al.* (2006) using amplified fragment length polymorphism (AFLP) and Naghavi *et al.* (2009) analysing random amplified polymorphic DNA, AFLP and simple sequence repeat markers concluded that there is a vast variation in the above-mentioned gene pool. In the present study, finding many polymorphic characteristics in *Acc-1* and *Pgk-1* (Tables 2 and 3) can be considered as evidence for the presence of high nucleotide diversity within the Iranian wild einkorn.

Our observations are in agreement with those of Kilian *et al.* (2007b), who reported two haplotypes (equivalent to haplotypes A and C in the present study; Table S3, available online) from two localities (Kermanshah and Lorestan) for each of the two loci among their 15 Iranian

https://doi.org/10.1017/S1479262114000549 Published online by Cambridge University Press

Table 3. Nucleotide sequence variability of *Pgk-1* (679 bp) in haplotypes A and B (*Triticum monococcum* ssp. *aegilopoides*) and haplotype C (*Triticum urartu*)[a]

Haplo-types	54–59	80	125	132	155	206	260	405	424	452	454	465	466	470	473	475	476	480	493	497	504	509–517	518	524	527	531	553	556	557–561	563	589	622	631	652
	i		i	i	i	e	e	i	i	i	i	i	i	i	i	i	i	i	i	i	i	i	i	i	i	i	i	i	i		i	e	e	e
A	–	–	–	–	–	G	C	A	G	C	T	A	A	C	G	A	G	C	C	–	T	A	A	T	C	A	T	C	T	T	A	T	T	T	C	G	A	T	A	A	C	C	T	G	C	T	G
B	–	–	–	–	–	C	T	G	A	T	C	A	G	T	A	C	C	C	T	–	G	A	C	C	–	–	–	–	–	–	T	T	A	G	T	A	A	C	C	T	G	T	A	C	A	G
C	A	C	T	T	C	C	C	T	G	A	T	C	C	A	T	T	A	G	C	C	C	–	–	–	–	–	–	–	–	–	C	T	A	G	C	T	–	–	–	–	C	G	T	T	A

–, i and e correspond to insertion/deletion (indels), intron and exon, respectively.
[a]Only polymorphic sites are indicated.

lines. Furthermore, two haplotypes (both named B, equivalent to haplotypes *Acc-1* Tb1331-V and *Pgk-1* Tm15-V evaluated by Kilian *et al.* (2007b); Table 1, Table S3, available online, and Fig. 1) are reported herein for the first time for Iran.

On the basis of the comprehensive study carried out by Kilian *et al.* (2007b), haplotypes A and B are the first and second most frequently found alleles in Turkey, and the presence of the latter (only six individuals) in Iran can be interpreted as a trace of its general distribution.

At the species level, individuals identified as *T. urartu* exhibited no genetic variability, which is in agreement with the findings of Kilian *et al.* (2007b) and Adderley and Sun (2014). This is parallel with the taxonomic homogeneity dominating in this species (Nasernakhaei *et al.*, 2013). The relatively high genetic diversity observed among the individuals belonging to *T. monococcum* ssp. *aegilopoides* collected from North-west of Iran (Table S2, available online, and Fig. 1) could be associated with the proximity of this area to the putative centre of diversity of diploid *Triticum* gene pool, i.e. South-east of Turkey according to Heun *et al.* (1997).

The genetic leakage from *T. monococcum* ssp. *aegilopoides* into *T. urartu* (Fig. 1), which is against the morphological distinctness, could clearly justify the occasional outcrossing, despite the fact that einkorn is a typical inbreeder (Zohary and Hopf, 2000; Kilian *et al.*, 2007b).

Relationship between morphology and haplotype variability

In this study, *T. monococcum* ssp. *aegilopoides* and *T. urartu* were recognized based on two morphological features: (1) loose long hairs on leaf and anther length of 3–5 mm and (2) dense short hairs on leaf and anther length of 1.9–2.8 mm, respectively (Nasernakhaei *et al.*, 2013). However, superimposing the revealed haplotypes on the above-mentioned classification revealed a conflict with morphological features (Fig. 2 and Fig. S1, available online). It is worth mentioning that there is no purpose herein to make any generalization between the evolutionary changes of only one or two genes and the evolutionary trend of one organism. Any genetic segment has its own story, and to make a rational comparison between the phylogeny of organisms and gene trees, sequencing of multiple genes (>50) is required (J. Dubcovsky, 2013, pers. commun.).

The 46 bp indel in *Acc-1* in haplotype A (Table 2), which is probably a consequence of a non-homologous double-strand repair (Puchta, 2005), was also utilized by Golovnina *et al.* (2009) to subdivide their *Acc-1* haplotypes into two variants.

https://doi.org/10.1017/S1479262114000549 Published online by Cambridge University Press

Table 4. Results of molecular variance analysis using concatenated sequence data of *Acc-1* and *Pgk-1* in *Triticum monococcum* ssp. *aegilopoides* and *Triticum urartu* for three geographical regions (population) in Iran

Grouping	Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	Fixation index
Complete wild diploid gene pool (*T. monococcum* ssp. *aegilopoides*, *T. urartu*)	Among species	1	1350.0	39.55	86.28	$F_{CT} = 0.86$
	Among populations Within species	4	88.624	1.432	3.12	$F_{ST} = 0.89$
	Within populations	70	340.20	4.86	10.60	$F_{SC} = 0.22$
	Total	75	1778.8	45.85		
T. monococcum ssp. *aegilopoides*	Among populations	2	88.62	2.238	24	$F_{ST} = 0.24$
	Within populations	48	340.2	7.087	76	
T. urartu	Among populations	NA	NA	NA	NA	NA
	Within populations	NA	NA	NA	NA	NA

d.f., Degrees of freedom; F_{CT}, variance among groups relative to total variance; F_{ST}, variance among populations; F_{SC}, variance among populations within groups (species); NA, statistics that could not be calculated due to a lack of variation.

In addition, some studies have shown that while the deletion form of this indel characterizes only the sequences equivalent to haplotypes B and C evaluated in the present study, its insertion form is exclusively present in a group of species including *Aegilops tauschii* Coss., *Ae. speltoides* Tausch, *Hordeum vulgare* L. and *Secale cereale* L., which are very similar to haplotype A evaluated in the present study (Huang *et al.*, 2002a, b; Kilian *et al.*, 2007a; Goncharov *et al.*, 2008; Golovnina *et al.*, 2009; Kang *et al.*, 2010). This conflict may be hypothesized based on the suggestion of *T. monococcum s. lat.* being a paraphyletic group, with respect to only the above-mentioned indel. This suggestion can be justified on the basis of a common *Triticum* diploid ancestor that on *Acc-1* gene diverged into two lineages, one carrying variant A and the other variant B–C. Taking into account the 'commonality concept' of 'polarization', variant A can be regarded as the plesiomorphic state of this gene. More recently, *T. urartu* has derived morphologically from the lineage of B–C variant; a hypothetical visualization of this suggestion is shown in Fig. S1 (available online). The relevant literature has shown that currently there is confusion regarding the taxonomy of *T. monococcum s. lat.* (van Slageren, 1994; Nasernakhaei *et al.*, 2013), which may be due to the paraphyletic origin of this species.

Explanation for this conflict for *Pgk-1* is more complex. The closer relationship of *Pgk-1* haplotype B with haplotype C than with haplotype A (Table 3 and Fig. 2) may be interpreted based on a paraphyletic group, rapid speciation (unexpectedly short divergence times) and the genomic integration created by 'introgressive hybridization'. Wicker

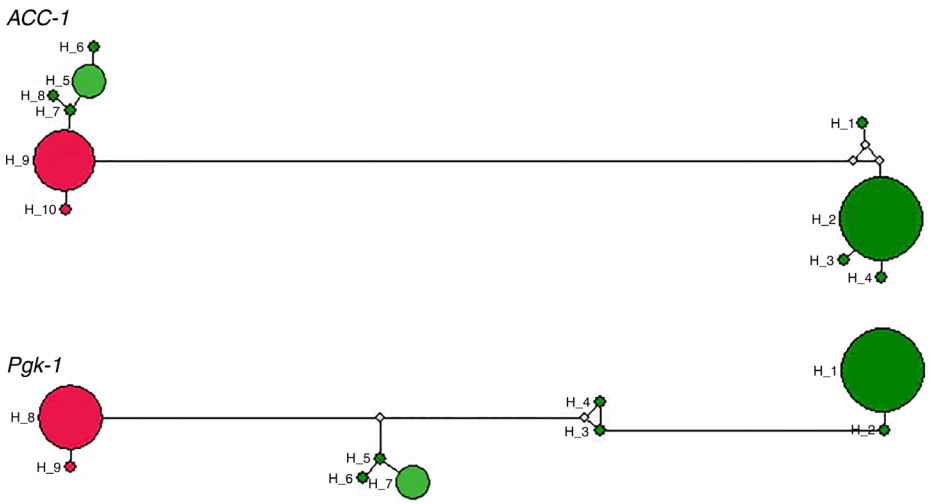

Fig. 2. Networks created based on the haplotype data obtained in this study and the relevant haplotype data obtained from the GenBank for *Acc-1* and *Pgk-1*. Green, light green and red colours correspond to *Triticum monococcum s. lat.* and *Triticum urartu*, respectively. For the name of each haplotype, see Table S3 (available online). The area of the circles is associated with the number of individuals.

https://doi.org/10.1017/S1479262114000549 Published online by Cambridge University Press

et al. (2009) demonstrated that between quite divergent subspecies and species there may be haplotype segments much older than those found during the divergence of the (sub)-species. They reported that in barley there were regions in the genome that diverged about one million years ago immediately next to regions that diverged only a few 100,000 years ago. They believe that these kinds of introgressions occur quite frequently both in wild species and in domesticated species. They also found that extremely divergent segments could be exchanged between rye and wheat (T. Wicker, 2013, pers. commun.).

Therefore when we found such divergent haplotype segments in our materials, this will produce phylogeny of single genes contradicts the overall phylogeny of the species (Doyle, 1992). This hypothesis is confirmed by the haplotype networks shown in Fig. 2 in which while the individuals having each of the haplotypes are identified carefully, their taxonomic relationships are overshadowed by the distribution of the haplotypes.

As a generalization, the observations made in this study stipulate that part of the haplotype variability observed in the local flora may be ignored in the general studies (haplotype B was not reported by Kilian *et al.* (2007b)) and also the urgent necessity of *in* and *ex situ* conservation of regional vegetations that are assumed to be the origin of domestication of some strategic plants. In addition, the present work will be completed by screening the important functional loci like those which mentioned in Kilian *et al.* (2007b). In conclusion, we suggest that SSCP analysis is an applicable molecular tool for pre-screening genetic variability in regions where thorough sequencing of an enormous number of DNA samples is time consuming and not affordable.

Supplementary material

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S1479262114000549

Acknowledgements

The authors are grateful to Dr Yuan Li for her valuable guidance, comments and corrections and to Drs Thomas Wicker, Jorge Dubcovsky, Benjamin Kilian, Nikolay P. Goncharov, Sidram Dhanagond, Piotr Gawronski, Heidi Lischer and Hong Chang Lim for their valuable help and guidance. They are also grateful to two anonymous reviewers and Dr Robert Koebner for improving the overall quality of the manuscript and Ms Faye Kalloniatis for her help. This study was carried financially supported by the University of Isfahan.

References

Adderley S and Sun G (2014) Molecular evolution and nucleotide diversity of nuclear plastid phosphoglycerate kinase (*PGK*) gene in Triticeae (Poaceae). *Gene* 533: 142–148.

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Cambridge: Harvard University Press.

Bandelt HJ, Forster P and Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16: 37–48.

Chantret N, Salse J, Sabot F, Rahman S, Bellec A, Laubin B, Dubois I, Dossat C, Sourdille P, Joudrier P, Gautier MF, Cattolico L, Beckert M, Aubourg S, Weissenbach J, Caboche M, Bernard M, Leroy P and Chalhoub B (2005) Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). *The Plant Cell* 17: 1033–1045.

Doyle JJ (1992) Gene trees and species trees: molecular systematic as one-character taxonomy. *Systematic Botany* 17: 144–163.

Dvorak J, Luo MC and Akhunov ED (2011) N.I. Vavilov's theory of centers of diversity in light of current understanding of wheat diversity, domestication and evolution. *Czech Journal of Genetics and Plant Breeding* 47: S20–S27.

Earl DA and vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.

Evanno G, Regnaut S and Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611–2620.

Excoffier L, Smouse P and Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479–491.

Excoffier L, Laval G and Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* 1: 47–50.

Fan X, Zhang HQ, Sha LN, Zhang L, Yang RW, Ding CB and Zhou YH (2007) Phylogenetic analysis among *Hystrix, Leymus* and its affinitive genera (Poaceae: Triticeae) based on the sequences of gene encoding plastid acetyl-CoA carboxylase. *Plant Science* 174: 701–707.

Gawel NJ and Jarret RL (1991) A modified CTAB DNA extraction procedure for *Musa* and *Ipomoea*. *Plant Molecular Biology Reporter* 9: 262–266.

Golovnina KA, Kondratenko EY, Blinov AG and Goncharov NP (2009) Phylogeny of the A genome of wild and cultivated wheat species. *Russian Journal of Genetics* 43: 1360–1367.

Goncharov NP, Golovnina KA, Kilian B, Glushkov S, Blinov A and Shummy K (2008) Evolutionary history of wheats – the main cereal of mankind. In: Dobretsov N, Kolchanov N, Rozanov A and Zavarzin G (eds) *Biosphere Origin and Evolution*. Berlin: Springer-Verlag, pp. 407–419.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.

Haudry A, Cenci A, Ravel C, Bataillon T, Brunel D, Poncet C, Hochu I, Poirier S, Santoni S, Glémin S and David J (2007) Grinding up wheat: a massive loss of nucleotide diversity since domestication. *Molecular Biology and Evolution* 24: 1506–1517.

Heun M, Schafer-Pregl R, Klawan D, Castagana R, Accerbi M, Borghi B and Salamini F (1997) Site of einkorn wheat

domestication identified by DNA fingerprinting. *Science* 278: 1312–1314.

Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R and Gornicki P (2002*a*) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. *Proceedings of the National Academy of Sciences of the United States of America* 99: 8133–8138.

Huang S, Sirikhachornkit A, Faris JD, Su X, Gill BS, Haselkorn R and Gornicki P (2002*b*) Phylogenetic analysis of the acetyl-CoA carboxylase and 3-phosphoglycerate kinase loci in wheat and other grasses. *Plant Molecular Biology* 48: 805–820.

Jakobsson M and Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801–1806.

Kang HY, Fan X, Zhang HQ, Sha LN, Sun GL and Zhou YH (2010) The origin of *Triticum petropavlovskyi* Udacz. et Migusch.: demonstration of the utility of the genes encoding plastid acetyl-CoA carboxylase sequence. *Molecular Breeding* 25: 381–395.

Kilian B, Ozkan H, Deusch O, Effgen S, Brandolini A, Kohl J, Martin W and Salamini F (2007*a*) Independent wheat B and G genome origins in out-crossing *Aegilops* progenitor haplotypes. *Molecular Biology and Evolution* 24: 217–227.

Kilian B, Özkan H, Walther A, Kohl J, Dagan T, Salamini F and Martin W (2007*b*) Molecular diversity at 18 loci in 321 wild and 92 domesticate lines reveal no reduction of nucleotide diversity during *Triticum monococcum* (Einkorn) domestication: implication for the origin of agriculture. *Molecular Biology and Evolution* 24: 2657–2668.

Liberado P and Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.

Lischer HEL and Excoffier L (2012) PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28: 298–299.

Longstaff M, Raines CA, McMorrow EM, Bradbeer JW and Dyer TA (1989) Wheat phosphoglycerate kinase: evidence for recombination between the genes for the chloroplastic and cytosolic enzymes. *Nucleic Acids Research* 17: 6569–6580.

Maleki M, Naghavi MR, Alizadeh H, Potki P, Kazemi M, Pirseyedi SM, Mardi M and Fakhr Tabatabaei M (2006) Study of genetic variation in wild diploid wheat (*Triticum boeoticum*) from Iran using AFLP markers. *Iranian Journal of Biotechnology* 4: 269–274.

Maxted N, Kell S and Ford-Lloyd B (2008) Crop wild relative conservation and use: establishing the context. In: Maxted N, Ford-Lloyd B, Kell S, Iriondo JM, Dulloo E and Turok J (eds) *Crop Wild Relative Conservation and Use*. Wallingford: CABI Publishing, pp. 3–30.

Naghavi MR, Malaki M, Alizadeh H, Pirseiedi M and Mardi M (2009) An assessment of genetic diversity in wild diploid wheat *Triticum boeoticum* from West of Iran using RAPD, AFLP and SSR markers. *Journal of Agricultural Science and Technology* 11: 585–598.

Nasernakhaei F, Rahiminejad MR, Saeidi H and Tavassoli M (2013) Taxonomic identity of the Iranian diploid *Triticum* as evidenced by nrDNA ITS analysis. *Phytotaxa* 143: 43–53.

Nei M (1987) *Molecular Evolutionary Genetics*. New York: Columbia University Press.

Nejat-Boshehri AA and Fakhr Tabatabaei M (2001) Protein fingerprinting of *Triticum thaoudar* Reut. population in Iran. *Iranian Journal of Agriculture Science* 3: 567–573 (in Persian).

Nesbitt M (2001) Wheat evolution: integrating archaeological and biological evidence. In: Caligari PDS and Brandham PE (eds) *Wheat Taxonomy: The Legacy of John Percival*. London: The Linnean Society of London, pp. 37–59.

Peng J, Sun D and Nevo E (2011) Domestication evolution, genetics and genomics in wheat. *Molecular Breeding* 28: 281–301.

Pritchard JK, Stephens P and Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

Puchta H (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *Journal of Experimental Botany* 56: 1–14.

Riehl S, Benz M, Conard NJ, Darabi H, Deckers K, Fazeli Nashli H and Zeidi-Kulehparcheh M (2012) Plant use in three Pre-Pottery Neolithic sites of the northern and eastern Fertile Crescent: a preliminary report. *Vegetation History and Archaeobotany* 21: 95–106.

Riehl S, Zeidi M and Conard NJ (2013) Emergence of agriculture in the foothills of the Zagros Mountains of Iran. *Science* 341: 65–67.

Rodriguez F, Cai D, Teng Y and Spooner D (2011) Asymmetric single-strand conformation polymorphism: an accurate and cost-effective method to amplify and sequence allelic variants. *American Journal of Botany* 98: 1061–1067.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* 4: 137–138.

Sanguinetti CJ, Dias Neto E and Simpson AJG (1994) Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. *Biotechniques* 17: 915–919.

Sasaki Y and Nagano Y (2004) Plant acetyl-CoA carboxylase: structure, biosynthesis, regulation, and gene manipulation for plant breeding. *Bioscience, Biotechnology, and Biochemistry* 68: 1175–1184.

Takenaka SH, Mori N and Kawahara T (2010) Genetic variation in domesticated emmer wheat (*Triticum turgidum* L.) in and around Abyssinian highlands. *Breeding Science* 60: 212–227.

van Slageren MW (1994) *Wild Wheats: A Monograph of* Aegilops *L. and* Amblyopyrum *(Jaub. & Spach) Eig. (Poaceae)*. Wageningen: Wageningen Agricultural University.

Wicker T, Krattinger SG, Lagudah ES, Komatsuda T, Pourkheirandish M, Matsumoto T, Cloutier S, Reiser L, Kanamori H, Sato K, Perovic D, Stein N and Keller B (2009) Analysis of intraspecies diversity in wheat and barley genomes identifies breakpoints of ancient haplotypes and provides insight into the structure of diploid and hexaploid Triticeae gene pools. *Plant Physiology* 149: 258–270.

Zohary D and Hopf M (2000) *Domestication of Plants in the Old World*. Oxford: Oxford University Press.