DE GRUYTER

Biologia **71**/3: 305—315, 2016
Section Botany
DOI: 10.1515/biolog-2016-0042

Use of IRAP and REMAP markers to interpret the population structure of *Linum usitatissimum* from Iran

Hossein Abbasi Holasou[1], Babak Abdollahi Mandoulakani[1,2]*, Morad Jafari[1,2] & Iraj Bernousi[1]

[1]*Department of Plant Breeding and Biotechnology, Faculty of Agriculture, Urmia University, Urmia Iran; e-mail: b.abdollahi@urmia.ac.ir*
[2]*Department of Agricultural Biotechnology, Institute of Biotechnology, Urmia University, Urmia Iran*

Abstract: Flax (*Linum usitatissimum* L.) is the third largest natural fiber crop in the world. The flax genome shows environmentally induced heritable genomic changes. The activation of transposable elements has been proposed and identified as the mechanism behind this genotypic plasticity. Transposable elements, particularly the retrotransposons, generates genomic diversity by replication which makes them an excellent source of molecular markers. Inter-retrotransposon amplified polymorphism (IRAP) and the retrotransposon-microsatellite amplified polymorphism (REMAP) markers were used to assess the insertional polymorphism of LTR retrotransposons and genetic diversity in 80 genotypes of *L. usitatissimum* collected from Iran. A total of 77 and 133 loci were amplified using 7 IRAP and 13 REMAP primers, respectively. Percentage of polymorphic loci (PPL) for IRAP and REMAP markers were 53.25% and 58.92%, respectively. Average of expected heterozygosity (He), number of effective alleles (Ne) and Shannon's information index (I) for IRAP markers were slightly more than those of REMAP markers. A high level of intra-population genetic differentiation was found, which is supported by a moderate level of gene flow among populations. A model-based Bayesian approach and cluster analysis using Minimum Evolution (ME) algorithm distinguished genotypes collected from Alborz region from those collected from Zagros region. Mantel test between genetic and geographical distances of populations revealed low but significant correlation coefficient ($r = 0.36$, $P \leq 0.05$). The results demonstrated that molecular markers developed based on active LTR retrotransposons in flax could be used as relatively reliable tools to analysis population structure in *L. usitatissimum*.

Key words: *Linum usitatissimum*; Bayesian analysis; genetic diversity; IRAP markers.

Abbreviations: LTR, long terminal repeat; REMAP, retrotransposon-microsatellite amplified polymorphism; SSR, simple sequence repeat.

Introduction

Linum is the largest genus of the family *linaceae* and consists of nearly 200 species which are distributed throughout the temperate and subtropical regions of the world. Twenty-two species of this genus can be found in Iran (Rechinger 1974; Sharifinia & Assadi 2001). The most economically important species in the genus is cultivated flax (*L. usitatissimum*) and its wild progenitor (pale flax; *L. bienne* Mill.). For a long time flax has been cultivating as a dual-purpose crop grown for its fiber and linseed oil. Oil extracted from the flax seeds (40–45%) has high level of α-linoleic acid (an omega-3-fatty acids) content (Ivanova et al. 2011). The plant is sparking new interest for its sustainability and for novel pharmaceutical and nutritional uses (Haggans et al. 1999; Hilakivi-Clarke et al. 1999). Cultivated flax ($2n = 2x = 30$) is a self-fertile species (Fu et al. 2002) and has a large genome of 750 Mbp with an abundance of repeated sequences (Cullis 2005). In spite of the wide range of species diversity within the genus, along with diverse basic chromosome numbers (ranging from $n = 8$ to $n = 43$; Gill 1987), flax genetic improvement is hindered by insufficient genomic resources (Cloutier et al. 2011). The origin of flax (*L. usitatissimum*) is uncertain. Some authors consider *L. bienne* as the progenitor of small seeded flax, originating from Kurdistan and Iran, whereas others introduce *L. angustifolium* containing high oil content and seed weight, as the progenitor, originating from the Mediterranean region (Murre 1955). Recently, a study with molecular markers suggested that the three species originate from one common ancestor, *L. angustifolium* being most ancient (Muravenko et al. 2003). In another study the center of origin of cultivated flax is believed to be the Middle East, although secondary diversity centers were identified in the Mediterranean basin, Ethiopia, Central Asia, and India (Vavilov 1926; Zohary & Hopf 2000).

Due to the use of modern agricultural systems, which favor genetically uniform and high yielding vari-

* Corresponding author

eties, the genetic diversity of crop plants has generally been reduced relative to their wild ancestors (Laurenin 2009). Traditionally, genetic variation analyses relied on morphological and phenotypic markers (Barnes et al. 1960; Diederichsen & Hammer 1995), but these markers has been limited to a few phenotypic traits, which are under strong environmental influence and show little variation, especially for highly heritable traits, and have the disadvantage of being environmentally dependent (Rao 2004). Flax germplasm was first screened for molecular markers by Gorman et al. (1993) who developed polymorphic isoenzyme systems but these biochemical markers are limited in number and phylogenetic value (Lowe et al. 2004). Nowadays, molecular characterization of cultivated flax as well as studies of potential evolutionary pathways of wild flax species have been assessed using amplified fragments length polymorphism (AFLP) (Everaert et al. 2001), random amplified polymorphic DNA (RAPD) (Fu 2006), inter simple sequence repeat (ISSR) (Uysal et al. 2010; Wiesnerova & Wiesner 2004), inter-retrotransposon amplified polymorphism (IRAP) (Smykal et al. 2011) and sequence-specific amplified polymorphism (S-SAP) (Melnikova et al. 2014) markers. Taken together, these studies show that cultivated flax has low genetic diversity compare to wild relatives or some other crops (Melnikova et al. 2014; Smykal et al. 2008). As regards, various flax breeding programs in recent decades have reduced the genetic variability of flax cultivars, the reason can be found in the lack of useful alleles in genomes of modern cultivars (Diederichsen & Fu 2006; Tanksley & McCouch 1997).

In apparent contradiction to the lack of diversity indicated by marker studies, the flax genome shows environmentally induced yet heritable genomic changes, a phenomenon of interest for many years (Evans et al. 1966). Chen et al. (2005, 2009) stated that the activation of transposable elements, particularly the retrotransposons, has been proposed and identified as the mechanism behind this genotypic plasticity. Retrotransposons have been detected in *L. usitatissimum*, accounting for as much as 55% of the genome (Paterson et al. 2009). Some of the plant retrotransposons are structurally intact and transcriptionaly active and presented in many copies throughout plant genomes (Vicient et al. 1999). The dispersion, ubiquity, and prevalence of retrotransposons in plant genomes provide an excellent basis for the development of molecular marker systems based on the PCR technique. The new insertions promoted by replicationally active members of a retrotransposon family result in insertional polymorphism that can be detected by IRAP and REMAP markers (Kallendar et al. 1999; Waugh et al. 1997). IRAP has been used alone or in combination with REMAP in studies of genetic diversity and genetic mapping within several plant genera and species, including *Citrus* (Biswas et al. 2010), *Pisum* (Smykal 2006), *Medicago sativa* L. (Abdollahi Mandoulakani et al. 2012; Abdollahi Mandoulakani et al. 2015b), wheat (Abdollahi Mandoulakani et al. 2015c; Nasri et al. 2013) and melon (Gholamzadeh Khoei et al. 2015; Abdollahi Mandoulakani et al. 2015a). Smykal et al. (2011) reported the utility of retrotransposons, a dominant and ubiquitous part of eukaryotic genomes as well as model-based Bayesian analysis, for diversity studies in flax germplasm. Melnikova et al. (2014) indicated that S-SAP analysis is very useful for characterization of flax varieties and identification of accession belonging to different species or sections and provided new information about phylogenetic relationships within the genus *Linum*. The advantage of application of these marker systems consists in the ability to track an insertion event and its subsequent vertical radiation through a pedigree or phylogeny (Shimamura et al. 1997).

In this study, we aimed to assess genetic diversity and activity of LTR retrotransposons in 80 genotypes of cultivated flax populations collected from Iran using IRAP and REMAP markers. The congruence between geographical distances and molecular markers-based distances of the populations was also investigated.

Material and methods

Plant material and DNA isolation

Plant materials consisted of 80 genotypes from nine flax populations (7–10 plants from each populations) (Table 1, Fig. 1) were kindly provided by the Research Center of Agriculture, West Azerbayejan, Urmia, Iran. Seeds were treated with 10% sulfuric acid for 5 to 10 min to break the dormancy. After washing the seeds with distilled water, they were kept in gibberllic acid 1500 ppm at 4 ℃ for two weeks. Seeds were planted in small pots with 10 cm diameter containing mixture of garden soil and vermiculite in the greenhouse with an ambient temperature 25 ℃. Genomic DNA was extracted from young leaves of 30-day seedlings using the method described by Ausubel et al. (1995) with minor modifications. The quality and concentration of the DNA were measured using a spectrophotometer and electrophoresis in a 0.8% (w/v) agarose gel.

IRAP and REMAP reactions

Four single and six IRAP primer combinations (Supplementary Tables 1 and 2) were used to analyze genetic diversity and integration events of retrotransposons in 80 flax genotypes. IRAP PCR amplifications were carried out in a Bio-Rad thermo cycler (Bio-Rad Laboratories, Hercules, CA, USA) in a total volume of 20 μL containing 45 ng of genomic DNA, 1× PCR buffer (10 mM Tris-HCl, 50 mM KCl, pH 8.3), 1.5 mM $MgCl_2$, 0.2 μ M of each dNTP, 0.5 U Tag DNA polymerase (CinnaGen, Iran), and 10 pmol of each primer. The cycling program were an initial denaturation of 4 min at 94 ℃ , 35 cycles of 45 s at 94 ℃ , 40 s at 55 ℃ , 2 min at 72 ℃ followed by a final extension step of 5 min at 72 ℃. The PCR products were resolved by electrophoresis (Bio-Rad) using 1.8% Resolute^{TM} line Biozyme agarose gel in 0.5× TBE buffer with constant voltage of 70 V for 3–4 h. Gels were stained by ethidium bromide. DNA fragments were then visualized under UV light and photographed using a gel documentation system.

Eighty-four REMAP primer combinations, derived from four single IRAP primers with 8 ISSR primers (Supplementary Tables 1 and 2) were applied. First, all IRAP and REMAP primers were tested on nine flax genotypes

Table 1. Collection site, code, geographical information and soil type of the *L. usitatissimum* populations used in the current study.

Code	Collection site	Number of plant	Latitude (°N)	Longitude (°E)	Elevation (m)	Soil Type*
TN97-33	Zabol	10	31°1′43″ N	61°30′4″ E	483	L
TN97-55	Bonab	10	37°20′25″ N	46°3′22″ E	1288	Si-C
TN97-92	Shahindezh	10	36°40′45″ N	46°34′1″ E	1370	S-C
TN97-106	Meshginshahr	10	38°23′56″ N	47°40′55″ E	1422	Si-C
TN97-246	Arak	8	34°5′30″ N	49°41′21″ E	1748	Si-S
TN97-273	Isfahan	9	32°39′16″ N	51°40′4″ E	1579	Si-C
TN97-27442	Zanjan	9	36°40′0″ N	48°28′59″ E	1644	–
TN97-27819	Ardabil	7	38°15′0″ N	48°16′60″ E	1349	–
TN97-907A	Karaj	7	35°50′24″ N	50°56′20″ E	1301	–

*Soil Type (S: sandy, Si: silt, C: clay, Si-S: silty sandy, S-C: sandy clay, Si-C: silty clay, L: loam).

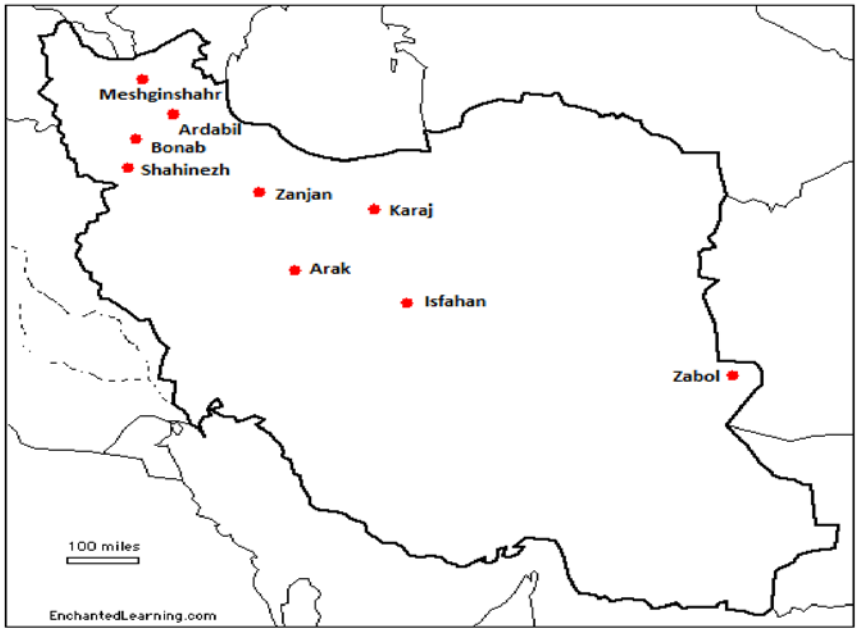

Fig. 1. Geographic distribution of the *L. usitatissimum* populations used in this study

to choose the primers producing scorable and discernible banding patterns. PCR amplification reactions and temperature profile, electrophoresis and visualization of REMAP markers were as stated for IRAPs.

Data analysis

The amplified fragments were scored independently as 1 or 0 for their presence or absence at each position, and the obtained binary data were used for analysis. Genetic similarity matrices between individual pairs of genotype were calculated for IRAP, REMAP and IRAP+REMAP data using Number of differences coefficients. These matrices were used to construct dendrograms using Minimum Evolution (ME) algorithm in MEGA 4.0 (Tamura et al. 2007). The statistical stability of the clusters was also estimated by a bootstrap analysis with 1000 replications using this software. Nei genetic distance between populations was measured based on IRAP, REMAP and combined data in GenAlEx 6.4 (Peakall & Smouse 2006). To verify the adjustment between similarity matrices and respective dendrogram-derived matrices (cophenetic matrix), the cophenetic correlation coefficient was estimated. To estimate the degree of correlation among the three cophenetic matrices derived from IRAP, REMAP, and combined data, Mantel test was performed with NTSYSpc (Rohlf 2000). An analysis of molecular variance (AMOVA) based on IRAP+REMAP data was performed using GenAlEx 6.4 (Peakall & Smouse 2006), to partition genetic variation within genotypes and among populations. The number of loci, percentage of polymorphic loci (PPL), number of alleles or loci with a frequency higher than or equal to 5%, number of private loci or alleles, number of less common loci with frequency lower than or equal to 25% and 50%, mean of expect heterozygosity (He), number of migrants between populations (Nm), number of effective alleles (Ne), Shannon's information index (I) and standard error of mean heterozygosity were also calculated for each population using GenAlEx 6.4 (Peakall & Smouse 2006). Polymorphic information content (PIC) was calculated for each primer using the formula $PIC = 1 - \sum P_{ij}^2$, (Smith et al. 1997).

Population structure was analyzed using a model-based Bayesian approach in the software Structure 2.3.4 (Pritchard et al. 2000). Ten independent runs were performed setting the number of subpopulations (k) from 1 to 20, burn in period and MCMC iterations, both to 100,000 and a model for admixture and correlated allele frequencies. The optimal K value (as true cluster number) was determined by the posterior probability [ln P(D)] and an ad hoc statistic ΔK based on the rate of change in [ln P(D)] between successive K (Evanno et al. 2005) using the software

Table 2. Characteristics of the IRAP and REMAP primers used in this study.

Primer name	TL	PL	PPL	He	Ne	I	PIC	Band size (bp)
IRAP								
LTR1833	12	6	50	0.155	1.271	0.229	0.43	100–2000
LTR1854	12	6	50	0.134	1.246	0.194	0.34	300–2000
LTR1868	9	3	34	0.145	1.271	0.206	0.32	75–2000
LTR1886	12	7	58	0.157	1.286	0.226	0.37	300–3000
LTR1833-LTR1868	8	4	50	0.069	1.103	0.114	0.31	75–1000
LTR1854-LTR1868	13	6	46	0.086	1.149	0.130	0.29	75–1500
LTR1868-LTR1886	11	9	82	0.170	1.299	0.254	0.51	200–3000
Total	77	41						
Average	11	5.86	53.25	0.13	1.23	0.19	0.37	
REMAP								
LTR1886-425	10	5	50	0.145	1.262	0.211	0.37	75–1500
LTR1854-443	13	9	69	0.072	1.124	0.110	0.27	200–2000
LTR1886-443	10	4	40	0.122	1.223	0.176	0.31	300–1500
LTR1854-UBC815	11	5	45	0.099	1.188	0.142	0.23	300–2000
LTR1868-UBC825	11	4	36	0.093	1.171	0.134	0.23	75–2000
LTR1868-UBC826	10	5	50	0.113	1.207	0.163	0.28	75–3000
LTR1854-UBC848	8	6	75	0.088	1.143	0.143	0.39	300–1500
LTR1868-UBC848	9	5	55	0.089	1.160	0.133	0.29	100–1500
LTR1868-UBC855	12	6	50	0.084	1.149	0.125	0.25	100–1500
LTR1886-UBC855	10	8	80	0.141	1.254	0.205	0.37	100–3000
LTR1886-A13	12	9	75	0.145	1.258	0.212	0.39	200–3000
LTR1868-A13	10	7	70	0.121	1.208	0.181	0.37	100–1500
LTR1854-A13	7	5	71	0.090	1.161	0.133	0.29	400–1500
Total	133	78						
Average	10.23	6	58.92	0.11	1.19	0.16	0.31	

TL: total loci, PL: polymorphic loci, PPL: percentage of polymorphic loci, He: expected heterozygosity, Ne: number of effective alleles, I: Shannon's information index

Structure Harvester. Inferred ancestry estimates of genotypes (Q-matrix) were derived for the selected subpopulation (Pritchard et al. 2000).

Results

IRAP analysis

Out of 10 IRAP primers tested (Table 2), four single primers (LTR1833, LTR1854, LTR1868, LTR1886) and three primer combinations (LTR1833-LTR1868, LTR1854-LTR1868, LTR1868-LTR1886) produced 77 distinguishable and scorable loci, which 41 loci (53.25%) were polymorphic. Length of the amplified fragments ranged from 75 to 3,000 bp. Primer combination LTR1868-LTR1886 generated the maximum amplified and polymorphic loci (Fig. 2). The maximum and minimum values of PIC was achieved by primers LTR1868-LTR1886 (0.51) and LTR1854-LTR1868 (0.29), averaging 0.37. Primer LTR1868-LTR1886 showed the highest values of He, Ne, I and PPL (Table 2).

To assess and partition total genetic variation among and within populations, AMOVA was performed based on the nine populations using IRAP data. Significant differences ($P< 0.05$) was detected within populations. The level of genetic variation was higher within populations (88%) compared to among populations (12%). Nei genetic distance between populations based on IRAP markers ranged from 0.027 (Shahindezh and Arak) to 0.095 (Shahindezh and Karaj) with a mean value of 0.053. Nm ranged from 0.80 (between populations Shahindezh and Karaj) to 49.75 (between populations Ardabil and Zanjan) (Table 4). Cluster analysis using IRAP markers placed 80 genotypes in five main groups (Fig. 3).

REMAP analysis

Thirteen out of 48 REMAP primers tested, amplified 133 distinguishable and scorable loci, which 78 (59%) were polymorphic. All single retroelement-based primers (expect for LTR1833) generated distinguishable and polymorphic banding patterns in combination with ISSR primers, with the size of amplified loci varied from 75 to 3000 bp (Table 2). The average of REMAP polymorphic loci was six per primer. Primers LTR1854-443 and LTR1886-UBC855 amplified the maximum number of loci (13) and the maximum amount of PPL, respectively (Table 2). PIC values varied from 0.23 (LTR1868-UBC825 and LTR1854-UBC815) to 0.39 (LTR1854-UBC848 and LTR1886-A13), averaging 0.31.

Similar to the results obtained by IRAP method, the level of genetic variation was higher within populations (86%) compared to among populations (14%). REMAP-based genetic distance between populations ranged from 0.009 (Meshginshahr and Ardabil) to 0.075 (Karaj and Isfahan), averaging 0.046. The least genetic differentiation (PhiPT = 0.003) was between populations Ardabil and Meshginshahr, while the maximum of that (PhiPT = 0.286) was between populations Zabol

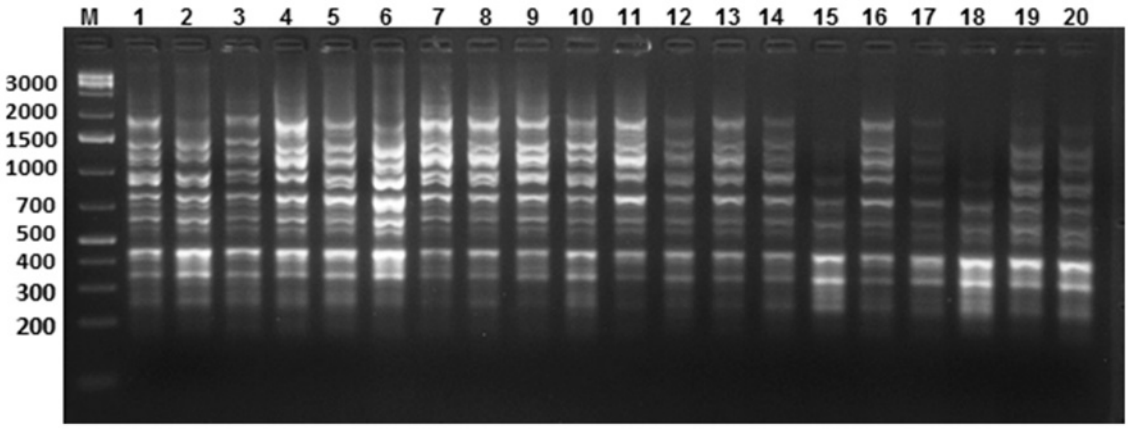

Fig. 2. Polymorphism detected by IRAP primer LTR1868-LTR1886. Lane M: 1kb O'GeneRuler^{TM} DNA ladder (Fermentas) in base pair, Lanes 1 to 10: genotypes from population Shahindezh, lanes 11 to 20: genotypes from population Zabol.

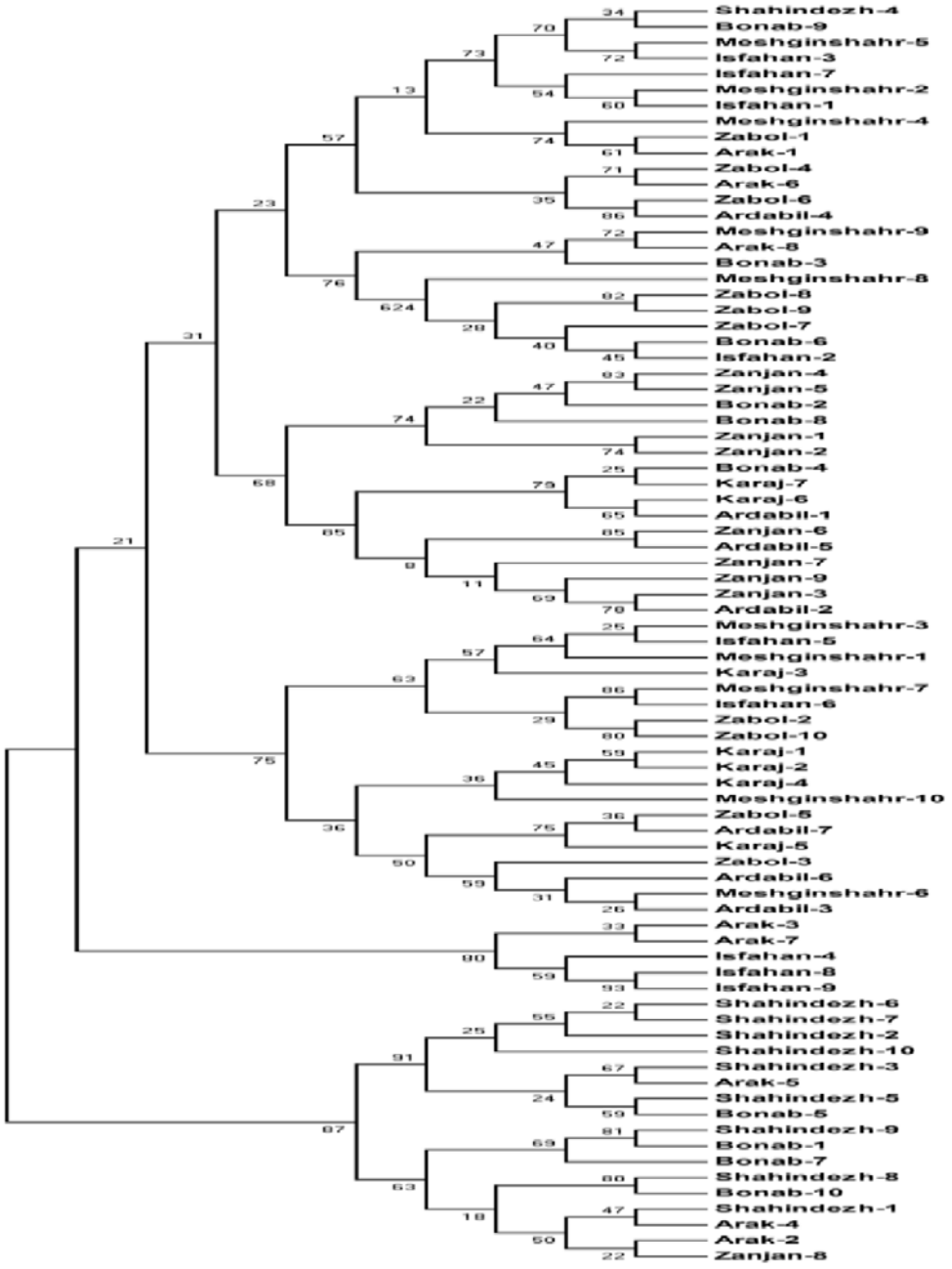

Fig. 3. IRAP dendrogram of 80 *L. usitatissimum* genotypes using Minimum evolution algorithm

Table 3. Characteristics of the IRAP and REMAP loci amplified in 80 genotypes of *L. usitatissimum*.

	Zagroz							Alborz			
	Shahindezh	Bonab	Zabol	Arak	Isfahan	Zanjan	Mean	Meshginshahr	Karaj	Ardabil	Mean
Number of loci	187	188	182	182	181	186	184.33	186	181	181	182.66
Percentage of polymorphic loci	34.29	38.57	33.33	31.90	31.90	32.38	33.73	32.38	27.62	29.05	29.68
No. loci freq ≥5% [a]	187	188	182	182	181	186	184.33	186	181	181	182.66
Number of private loci	0	0	0	0	0	0	0	0	0	1	0.33
No. L. Common loci (≤25%) [b]	2	0	0	1	0	1	0.67	2	0	0	0.67
No. L. Common loci (≤50%) [c]	13	13	9	9	11	9	10.66	14	9	8	10.33
Mean of He [e]	0.117	0.143	0.120	0.119	0.116	0.118	0.122	0.111	0.099	0.109	0.106
Shannon's information index (I)	0.174	0.209	0.177	0.174	0.171	0.174	0.180	0.165	0.147	0.160	0.157
Number of effective alleles (Ne)	1.207	1.259	1.209	1.214	1.210	1.214	1.219	1.195	1.174	1.196	1.188

[a]Number of loci with a frequency of ≥5%, [b]Number of loci with a frequency of ≤25%, [c]Number of loci with a frequency of ≤50%, [e]Mean of expected heterozygosity

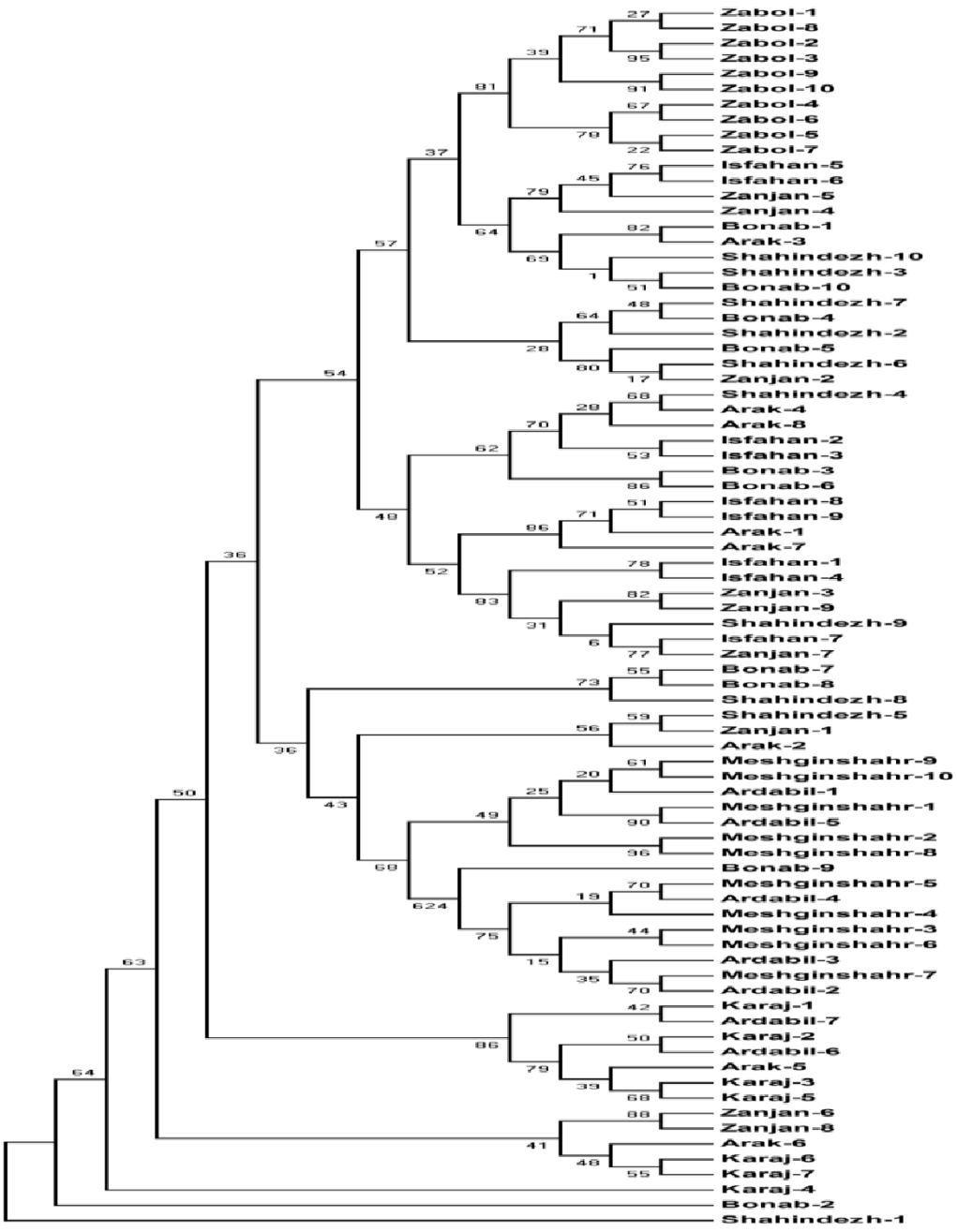

Fig. 4. REMAP dendrogram of 80 *L. usitatissimum* genotypes using Minimum evolution algorithm.

Table 4. Genetic differentiation (PhiPT), Nei genetic distance (GD) and number of migrants (Nm) values between each pair-wise of *L. usitatissimum* populations using IRAP, REMAP and IRAP+REMAP markers.

Population 1	Population 2	IRAP			REMAP			IRAP+REMAP		
		PhiPT	GD	Nm	PhiPT	GD	Nm	PhiPT	GD	Nm
Shahindezh	Bonab	0.052	0.033	4.56	0.004	0.015	62.25	0.008	0.022	31
Shahindezh	Meshginshahr	0.193	0.067	1.05	0.128	0.043	1.70	0.157	0.051	1.34
Bonab	Meshginshahr	0.126	0.057	1.73	0.135	0.048	1.60	0.131	0.051	1.66
Shahindezh	Zabol	0.167	0.059	1.25	0.149	0.045	1.43	0.156	0.050	1.35
Bonab	Zabol	0.147	0.054	1.45	0.136	0.039	1.59	0.141	0.045	1.52
Meshginshahr	Zabol	0.032	0.030	7.56	0.286	0.069	0.62	0.187	0.055	1.08
Shahindezh	Arak	0.059	0.027	3.98	0.036	0.028	6.69	0.045	0.028	5.30
Bonab	Arak	0.044	0.038	5.43	0.056	0.035	4.21	0.051	0.036	4.65
Meshginshahr	Arak	0.165	0.058	1.26	0.187	0.051	1.08	0.177	0.053	1.16
Zabol	Arak	0.162	0.047	1.29	0.203	0.055	0.98	0.186	0.052	1.09
Shahindezh	Isfahan	0.198	0.049	1.01	0.103	0.041	2.17	0.144	0.043	1.48
Bonab	Isfahan	0.120	0.043	1.83	0.128	0.045	1.70	0.124	0.044	1.76
Meshginshahr	Isfahan	0.098	0.037	2.30	0.268	0.071	0.68	0.196	0.059	1.02
Zabol	Isfahan	0.129	0.038	1.69	0.246	0.058	0.76	0.199	0.051	1.01
Arak	Isfahan	0.082	0.032	2.80	0.031	0.026	7.81	0.054	0.028	4.38
Shahindezh	Zanjan	0.173	0.082	1.19	0.024	0.028	10.16	0.093	0.047	2.44
Bonab	Zanjan	0.031	0.039	7.81	0.026	0.030	9.36	0.028	0.033	8.68
Meshginshahr	Zanjan	0.086	0.057	2.91	0.187	0.053	1.08	0.141	0.054	1.52
Zabol	Zanjan	0.115	0.064	1.92	0.175	0.052	1.18	0.149	0.056	1.43
Arak	Zanjan	0.144	0.068	1.48	0.050	0.030	4.75	0.096	0.044	2.35
Isfahan	Zanjan	0.139	0.058	1.55	0.075	0.028	3.08	0.106	0.038	2.11
Shahindezh	Karaj	0.237	0.095	0.80	0.094	0.044	2.41	0.153	0.062	1.38
Bonab	Karaj	0.109	0.062	2.04	0.123	0.046	1.78	0.116	0.052	1.90
Meshginshahr	Karaj	0.070	0.046	3.32	0.174	0.047	1.19	0.130	0.047	1.67
Zabol	Karaj	0.087	0.056	2.62	0.236	0.062	0.81	0.181	0.060	1.31
Arak	Karaj	0.161	0.066	1.30	0.163	0.048	1.28	0.162	0.054	1.29
Isfahan	Karaj	0.118	0.056	1.87	0.254	0.075	0.73	0.201	0.068	0.99
Zanjan	Karaj	0.021	0.030	11.65	0.162	0.052	1.29	0.105	0.044	2.13
Shahindezh	Ardabil	0.222	0.091	0.87	0.095	0.050	2.38	0.146	0.064	1.46
Bonab	Ardabil	0.073	0.058	3.17	0.104	0.050	2.15	0.090	0.053	2.53
Meshginshahr	Ardabil	0.026	0.036	9.36	0.003	0.009	83.8	0.003	0.018	83.08
Zabol	Ardabil	0.073	0.055	3.17	0.243	0.073	0.82	0.182	0.066	1.12
Arak	Ardabil	0.209	0.084	0.95	0.123	0.050	4.75	0.160	0.062	1.56
Isfahan	Ardabil	0.183	0.069	1.11	0.214	0.070	0.92	0.201	0.070	0.99
Zanjan	Ardabil	0.005	0.040	49.75	0.143	0.051	1.50	0.088	0.047	2.59
Karaj	Ardabil	0.068	0.039	3.43	0.112	0.046	1.98	0.096	0.044	2.35

and Meshginshahr (Table 4). Minimum evolution dendrogram using REMAP markers clustered 80 genotypes into 7 groups (Fig. 4).

Combined data analysis

Mantel test between IRAP and REMAP cophenetic matrices evidenced no significant correlation. Hence, IRAP+REMAP markers were used to construct a dendrogram to accumulate the power of both techniques. Cluster analysis based on ME algorithm divided all 80 genotypes into two major groups (Fig. 5) and separated genotypes distributed in Alborz region (North West and North of Iran) from those distributed in Zagros region (West and South of Iran). A model-based Bayesian approach (Fig. 6) confirmed the results obtained by ME cluster analysis. The mean Fst values for two groups were 0.24 and 0.12 respectively. IRAP+REMAP-based genetic distances ranged from 0.018 (Ardabil and Meshginshahr) to 0.070 (Ardabil and Isfahan). Nm varied from 0.99 (between Karaj and Isfahan and Ardabil and Isfahan populations) to 83.08 (between Ardabil and Meshginshahr) (Table 4).

AMOVA was performed to divide total genetic variation among and within populations. Similar to what was recorded by IRAP and REMAP markers, the level of genetic variation was higher within populations (87%) compared to among populations (13%). The PPL, He and I based on IRAP+REMAP data varied from 27 to 38, 0.099 to 0.143 and 0.147 to 0.209 in Karaj and Bonab populations, respectively (Table 3). The mentioned parameters for Zagros populations were more than those of Alborz populations (Table 3).

Discussion

Retrotransposon insertional polymorphism in L. usitatissimum genome

To our knowledge, this is the first report of IRAP and REMAP-based assessment of retrotransposon insertional polymorphism in Iranian *L. usitatissimum* populations. In this study, 7 IRAP and 13 REMAP primers derived from *Linum* LTR retrotransposons were shown to amplify polymorphic and discernible banding patterns and used to evaluate genetic diversity in 80 genotypes of the *L. usitatissimum*. All single IRAP primers (LTR1833, LTR1854, LTR1868, and LTR1886) gener-

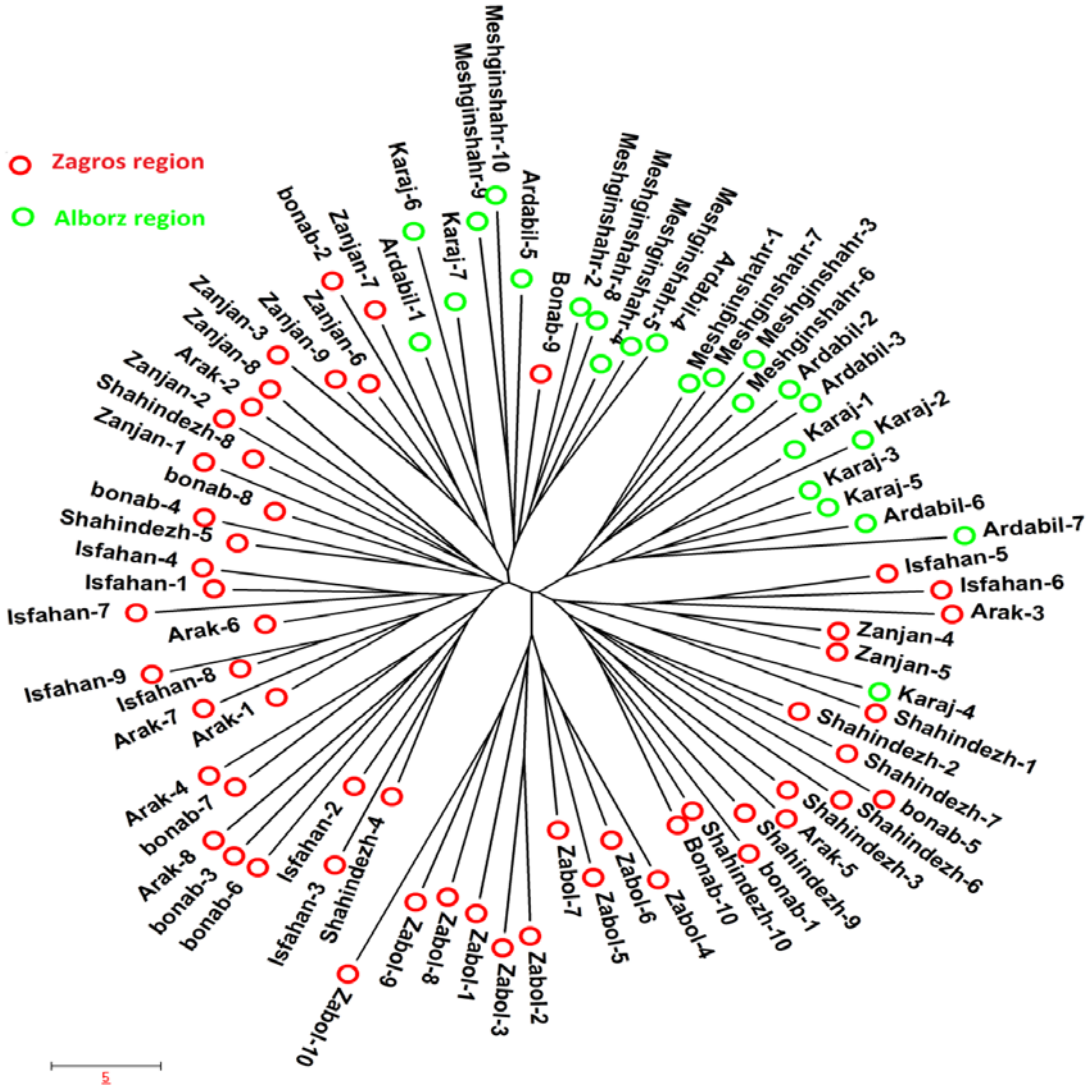

Fig. 5. Minimum evolution dendrogram of 80 *L. usitatisimum* genotypes based on IRAP+REMAP markers.

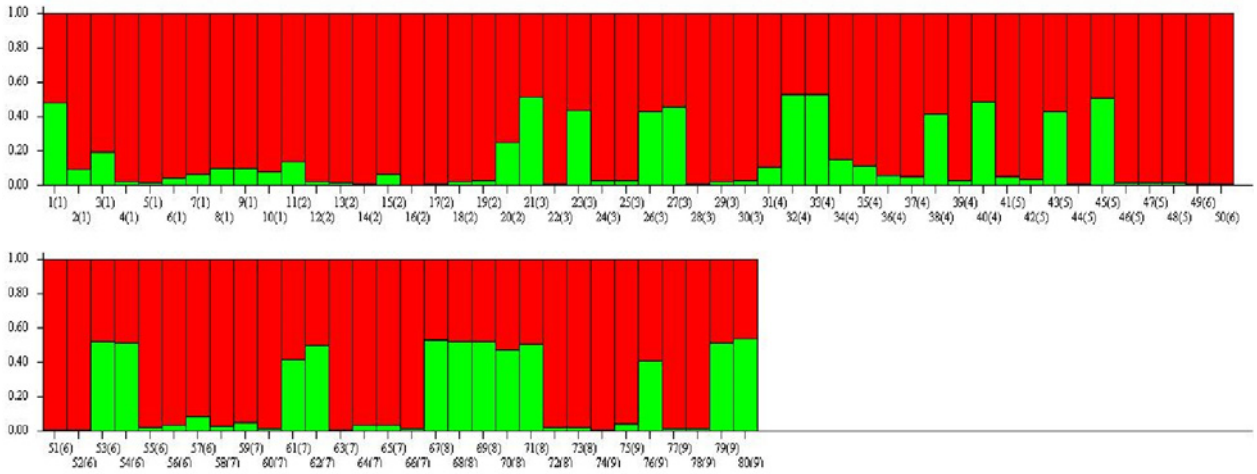

Fig. 6. A Bayesian model-based clustering of the 80 *Linum usitatissimum* genotypes. Bar plots show the membership coefficient estimate (Q) for each genotype for the inferred clusters with maximum log-likelihood probability. Bar colors and lengths represent inferred clusters and Q, respectively, identified by Structure for K = 2.

ated scorable banding patterns, indicating the presence and activity of these elements in Iranian flax genome. Out of tested IRAP primer combinations, three combinations (LTR1833-LTR1868, LTR1854-LTR1868 and

LTR1868-LTR1886) produced clear banding patterns, demonstrating nested insertion of the LTR retrotransposons in flax genome. In all IRAP primer combinations, LTR1868 was present indicating its high frequency and activity in the flax genome as well as a possible role of this retrotransposon family in the construction, organization, and evolution of the flax genome. Smykal et al. (2011) reported that different retrotransposon families integrate near or into each other in flax genome, and LTR retrotransposons especially LTR1868 family, has high integrated polymorphism in the *L. usitatissimum.* REMAP amplification of the used retrotransposon families indicated their insertion near the different SSRs in flax genome. In principle, retrotransposons may integrate in either orientation into the genome, and hence, any two members of one or different retrotransposon families may be found head to head, tail to tail, or head to tail (Abdollahi Mandoulakani et al. 2012; Kalendar et al. 1999). Moreover, retrotransposons belonging to different families may integrate into each other. Therefore, in our study, combined primers from LTR ends of different retrotransposon families were applied to increase the probability of finding bands.

Genetic relationship and characterization of flax populations

The estimated correlations between the three generated cophenetic matrices from IRAP, REMAP, and IRAP+REMAP dendrogram evidenced a relatively high and significant correlation of IRAP and REMAP with IRAP+REMAP. However, the matrices estimated by the techniques individually revealed a low and non-significant correlation (r = 0.22). These results are in agreement with findings in barley (Kalendar et al. 1999), rice (Branco et al. 2007), alfalfa (Abdollahi Mandoulakani et al. 2012) and wheat (Abdollahi Mandoulakani et al. 2015c; Nasri et al. 2013), since REMAP primers amplified DNA regions that are different from those covered by IRAP. Therefore, IRAP+REMAP data were used to reveal the association between studied populations and genotypes. The two first principal coordinates explained 51.7% of the total molecular variation, indicating relative clustering of retrotransposon insertion sites in some specific regions of the genome (data not shown). Although random distribution and insertional polymorphisms of retrotransposons in genome has been reported in different crop species (Abdollahi Mandoulakani et al. 2012, 2015c; Kalendar et al. 1999; Nasri et al. 2013), PCR amplification occurs only in location that retrotransposons are sufficiently near to each other. Therefore, it probably shows the high concentration of retrotransposon insertion sites in some parts of the flax genome.

The low level of genetic variation detected among the populations (I = 0.172, He = 0.117, Ne = 1.209) might be ascribed to the self-pollinating nature of flax and low transposition activity or creation of bottleneck effect during flax selection. Also, various flax breeding programs in recent decades have significantly reduced the genetic variability of flax cultivars (Tanksley & McCouch 1997). Smykal et al. (2011) in a study of genetic diversity in a cultivated flax germplasm including 708 accessions reported that, the flax cultivars are less diverse than wild flax species and landraces. In our study, genetic diversity detected in populations collected from Zagros region was more than that in populations collected from Alborz region. This may reflect the high activity and therefore much insretional polymorphism of the used LTR retrotransposons in population collected from Zagros region. Geographical distribution of studied populations may help us to interpret the different genetic variability observed in Alborz and Zagros regions, science different climate conditions could induce different patterns of retrotransposon insertional polymorphism in populations (Kalendar et al. 2000). Since populations were collected from two regions with different climate conditions (Zagros; hot and dry, Alborza: cold and humid), therefore transpositional activity and insertional polymorphism of the retrotransposons in the genome might be different among populations. Zou et al. (2009) demonstrated that retrotransposon activity strongly depend on stresses such as high temperature. In our study, genetic diversity within populations was more than the diversity between populations. Vormans (2006) and Smykal et al. (2011) reported the same results. The pollen of flax can be spread over a long distance by the combination of insects and wind, this may facilitate gene flow among populations and consequently high level of genetic variation within populations.

ME cluster analysis and model-based Bayesian analysis (Figs 5 and 6) using IRAP+REMAP markers divided the genotypes in two main groups in concordance with geographical distribution of the genotypes. The geographic differences and genetic exchange should be mainly responsible for the observed genetic structure of the populations science populations collected from Alborz region (Ardabil, Karaj and Meshginshahr) were separated from those collected from Zagros region (Figs 5 and 6). Genotypes from the same populations were largely grouped in the vicinity of each other with a few exceptions. Uysal et al. (2010) reported that genetic distances among the pale flax accessions were significantly associated with their geographic distances and elevation differences.

In conclusion, the application of retrotransposon-based markers offer a promising potential for IRAP and REMAP analysis of *L. usitatissimum.* Our results showed that the genetic diversity of *L. usitatissimum* is low and it is necessary to extend the genetic base of Iranian flax germplasm. Introduction of new useful traits from the ancient primitive forms of cultivated flax and wild species could increase the polymorphism of modern flax varieties. Besides, the use of the strong and complementary statistical methods such as ME cluster analysis and Baysian methods proved to be useful for the determination of genetic relationships among flax genotypes and for the definition of the genetic structure of this collection.

Acknowledgements

The authors would like to thank the Institute of Biotechnology of the Urmia University (Urmia, Iran) for hosting the lab facilities. Dr A. Hassanzadeh Gorttapeh (Agricultural and Natural Research Center of West Azarbayjan, Urmia, Iran) appreciates for providing the seeds of flax populations.

References

Abdollahi Mandoulakani B., Piri Y., Darvishzadeh R., Bernoosi I. & Jafari M. 2012. Retroelement insertional polymorphism and genetic diversity in *Medicago sativa* populations revealed by IRAP and REMAP markers. Plant Mol. Biol. Rep. **30:** 286–296.

Abdollahi Mandoulakani B., Rahmanpour S., Shaaf S., Gholamzadeh Khoei S., Rastgou M. & Rafezi R. 2015a. Towards the identification of retrotransposon-based and ISSR molecular markers associated with populations resistant to ZYMV in melon. S. Afr. J. Bot. **100:** 141–147.

Abdollahi Mandoulakani B., Sadigh P., Azizi H., Piri Y., Nasri Sh. & Arzhangh S. 2015b. Comparative assessment of IRAP, REMAP, ISSR, and SSR markers for evalution of genetic diversity of alfalfa (*Medicago sativa* L.). J. Agr. Sci. Tech. **17:** 999–1010.

Abdollahi Mandoulakani B., Yaniv E., Kalendar R., Raats D., Bariana H.S., Bihamta M.R. & Schulman A.H. 2015c. Development of IRAP- and REMAP- derived SCAR markers for marker-assisted selection of the stripe rust resistance gene Yr15 derived from wild emmer wheat. Theor. Appl. Genet. **128:** 211–219.

Ausubel F.M., Brent R., Kingston R.E., Moore D.D., Seidman J.G., Smith J.A., Struhl K., Albright L.M., Coen D.M. & Varki A. 1995. Current Protocols in Molecular Biology. John Wiley, New York.

Barnes D.K., Culbertson J.O. & Lambert J.W. 1960. Inheritance of seed and flower colors in flax. Agron. J. **52:** 456–459.

Biswas M.K., Xu Q. & Deng X. 2010. Utility of RAPD, ISSR, IRAP and REMAP markers for the genetic analysis of *Citrus spp.* Sci. Hortic. **124:** 254–261.

Branco C.J.S., Vieira E.A., Malone G., Kopp M.M., Malone E., Bernardes A., Mistura C.C., Carvalho F.I.F. & Oliveira C.A. 2007. IRAP and REMAP assessments of genetic similarity in rice. J. Appl. Genet. **2:** 107–113.

Chen Y., Schneeberger R.G. & Cullis C.A. 2005. A site-specific insertion sequence in flax genotrophs induced by environment. New Phytol. **167:** 171–180.

Chen Y., Lowenfeld R. & Cullis C.A. 2009. An environmentally induced adaptive (?) insertion event in flax. Int. J. Genet. Mol. Biol. **3:** 038–047.

Cloutier S., Ragupathy R. & Niu Z. 2011. SSR-based linkage map of flax (*Linum usitassimum* L.) and mapping of QTLs underlying fatty acid composition traits. Mol. Breeding. **28:** 437–451.

Cullis C.A. 2005. Mechanisms and control of rapid genomic changes in flax. Ann. Bot. **95:** 201–206.

Diederichsen A. & Fu Y.B. 2006. Phenotypic and molecular (RAPD) differentiation of four infraspecific groups of cultivated flax (*Linum usitatissimum* L. subsp. *usitatissimum*). Genet. Resour. Crop Evol. **53:** 77–90.

Diederichsen A. & Hammer K. 1995. Variation of cultivated flax (*Linum usitatissimum* L. subsp. *usitatissimum*) and its wild progenitor pale flax (subsp. angustifolium (Huds.) Thell.). Genet. Resour. Crop Evol. **42:** 263–272.

Evanno G., Regnaut S. & Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. **14:** 2611–2620.

Evans G.M., Durrant A. & Rees H. 1966. Associated nuclear changes in the induction of flax genotrophs. Nature. **212:** 697–699.

Everaert I., de Riek J., de Loose M., van Waes J. & van Bockstaele E. 2001. Most similar variety grouping for distinctness evaluation of flax and linseed (*Linum usitatissimum L.*) varieties by means of AFLP and morphological data. Plant Var. Seeds. **14:** 69–87.

Fu Y.B. 2006. Redundancy and distinctness in flax germplasm as revealed by RAPD dissimilarity. Plant Genet. Resour. **4:** 117–124.

Fu Y.B., Diederichsen A., Richards K.W. & Peterson G. 2002. Genetic diversity within a range of cultivars and landraces of flax (*Linum usitatissimum* L.) as revealed by RAPDs. Genet. Resour. Crop Evol. **49:** 167–174.

Gholamzadeh Khoei S., Abdollahi Mandoulakani B. & Bernousi I. 2015. Genetic diversity in Iranian melon populations hybrids assessed by IRAP and REMAP markers. J. Agr .Sci. Tech. **17:** 1267–1277.

Gill K.S. 1987. Linseed. Publications and Information Division. Indian Council of Agricultural Research, New Delhi, India.

Gorman N.B., Cullis C.A. & Aldridge N. 1993. Genetic and linkage analysis of isozyme polymorphism in flax. J. Hered. **84:** 73–78.

Haggans C.J., Hutchins A.M., Olson B.A., Thomas W., Martini M.C. & Slavin J.L. 1999. Effect of flaxseed consumption on urinary estrogen metabolites in postmenopausal women. Nutr. Cancer. **33:** 188–195.

Hilakivi-Clarke L., Clarke R. & Lippman M. 1999. The influence of maternal diet on breast cancer risk among female offspring. Nutrition. **15:** 392–401.

Ivanova S., Rashevskaya T. & Makhonina M. 2011. Flaxseed additive application in dairy products production. Procedia Food Sci. **1:** 275–280.

Kalendar R., Grab T., Regina M., Souniemi A. & Schulman A.H. 1999. IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. Theor. Appl. Genet. **98:** 704–711.

Kalendar R., Tanskanen J., Immonen S., Nevo E. & Schulman A.H. 2000. Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. Proc. Natl. Acad. Sci. USA. **97:** 6603–6607.

Laurentin H. 2009. Data analysis for molecular characterization of plant genetic resources. Genet. Resour. Crop Evol. **56:** 277–292.

Lowe A., Harris S. & Ashton P. 2004. Ecological Genetics: Design, Analysis and Application. Blackwell Science Ltd, Oxford, (UK), 326 pp.

Melnikova N.V., Kudryavtseva A.V., Zelenin A.V., Lakunina V.A., Yurkevich O.Y., Speranskaya A.N., Dmitriev A.A., Krinitsina A.A., Belenikim M.S., Uroshlev L.A., Snezhkina A.V., Sadritdinova A.F., Koroban N.V., Amosova A.V., Samatadze T.E., Guzenko E.V., Lemesh V.A., Savilova A.M., Rachinskaia O.A., Kishlyan N.V., Rozhmina T.A., Bolsheva N.L. & Muravenko O.V. 2014. Retrotransposon-based molecular markers for analysis of genetic diversity within the genus *Linum.* BioMed Res. Int. 231589.

Muravenko O.V., Lemesh V.A., Samatadze T.E., Amosova A.V., Grushetskaya Z.E., Popov K.V., Semenova O.Y.U., Khotyuleva L.V. & Zelenin A.V. 2003. Genome comparisons with chromosomal and molecular markers for three closely related flax species and their hybrids. Russ. J. Genet. **39:** 414–421.

Murre M. 1955. Vezelvlas. Uitgeverij Ceres, Meppel, The Netherlands: 112 pp.

Nasri S.H., Abdollahi Mandoulakani B., Darvishzadeh R. & Bernoosi I. 2013. Retrotransposon insertional polymorphism in Iranian bread wheat cultivars and breeding lines revealed by IRAP and REMAP markers. Biochem. Genet. **51:** 927–943.

Paterson A.H., Bowers J.E., Bruggmann R., Dubchak I., Grimwood J., Gundlach H., Haberer G., Hellsten U., Mitors T., Poliakov A., Schmutz J., Spannagl M., Tang H., Wang X., Wicker T., Bharti A.K., Chapman J., Alex Feltus F., Gowik U., Grigoriev I.V., Lyons E., Maher C.A., Martis M., Narechania A., Otillar R.P., Penning B.W., Salamov A.A., Wang Y., Zhang L., Carpita N.C., Freeling M., Gingle A.R., Hash C.T., Keller B., Klein P., Kresovich S., McCann M.C., Ming R., Peterson D.G., Rahman M., Ware D., Westhoff P., Mayer

K.F.X., Messing J. & Rokhsar D.S. 2009. The *Sorghum bicolor* genome and the diversification of grasses. Nature **457:** 551–556.

Peakall R. & Smouse P.E. 2006. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes. **6:** 288–295.

Pritchard J.K., Stephens M. & Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics **155:** 945–959.

Rao N. 2004. Plant genetic resources: Advancing conservation and use through biotechnology. Afr. J. Biotechnol. **3:** 136–145.

Rechinger K.H. 1974. Linaceae. In: Rechinger KH (ed.) Flora Iranica. vol, 106. Graz: Akademische Druck-und Verlagsansatalt.

Rohlf F.J. 2000. NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, Version 2.1. Exeter Software, New York.

Sharifinia F. & Assadi M. 2001. Flora of Iran, No. 34: *Linaceae*. Research Inst, Forests and Rangelands, Ministry of Jahad-e-Sazandegi, Iran (In Persian).

Shimamura M., Yasue H., Ohshima K., Abe H., Kato H., Kishiro T., Goto M., Munechika I. & Okada N. 1997. Molecular evidence from retroposons that whales form a clade within eventoed ungulates. Nature **388:** 666–670.

Smith J.S.C., Chin E.C.L., Shu H., Smith O.S., Wall S.J., Senior M.L., Mitchell S.E., Kresovich S. & Ziegle J. 1997. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.) comparisons with data from RFLPS and pedigree. Theor. Appl. Genet. **95:** 163–173.

Smykal P. 2006. Development of an efficient retrotransposon-based fingerprinting method for rapid pea variety identification. J. Appl. Genet. **47:** 221–230.

Smykal P., Hybl M., Corander J., Jarkovsky J., Flavell A.J. & Griga M. 2008. Genetic diversity and population structure of pea (*Pisum sativum* L.) varieties derived from combined retrotransposon, microsatellite and morphological marker analysis. Theor. Appl. Genet. **117:** 413–424.

Smykal P., Bacova-Kerteszova N., Kalendar R., Corander J., Schlman A.H. & Pavelek M. 2011. Genetic diversity of cultivated flax (*Linum usitatissimum* L.) germplasm assessed by retrotransposon-based markers. Theor. Appl. Genet. **122:** 1385–1397.

Tamura K., Dudley J., Nei M. & Kumar S. 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. **24:** 1596–1599.

Tanksley S.D. & McCouch S.R. 1997. Seed banks and molecular maps: unlocking genetic potential from the wild. Sci. **277:** 1063–1066.

Uysal H., Fu Y.B., Kurt O., Peterson G.W., Diederichsen A. & Kusters P. 2010. Genetic diversity of cultivated flax (*Linum usitatissimum* L.) and its wild progenitor pale flax (*Linum bienne* Mill.) as revealed by ISSR markers. Genet. Resour. Crop Evol. **57:** 1109–1119.

Vavilov N.I. 1926. Studies on the origin of cultivated plants. Leningrad, Russia: Institute of Applied Botany and Plant Breeding (Leningrad).

Vicient C.M., Kalendar R., Anamthawat-Jonsson K., Suoniemi A. & Schulman A.H. 1999. Structure, functionality, and evolution of the BARE-1 retrotransposon of barley. Genetica. **107:** 53–63.

Vromans J. 2006. Molecular genetic studies in flax (*Linum usitatissimum L.*). PhD Thesis, Wageningen University, The Netherlands.

Waugh R., McLean K., Flavell A.J., Pearce S.R., Kumar A., Thomas B.B.T. & Powell W. 1997. Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequences-specific amplification polymorphisms (S-SAP). Mol. Genet. Gen. **253:** 687–694.

Wiesnerova D. & Wiesner I. 2004. ISSR-based clustering of cultivated flax germplasm is statistically correlated to thousand seed mass. Mol. Biotechnol. **26:** 207–214.

Zohary D. & Hopf M. 2000. Domestication of plants in the Old World: the origin and spread of cultivated plants in West Asia, Europe and the Nile Valley, 3th ed. Oxford University Press, Oxford.

Zou J., Gong H., Yang T.-J. & Meng J. 2009. Retrotransposons-a major driving force in plant genome evolution and a useful tool for genome analysis. J. Crop Sci. Biotech. **12:** 1–8.

Received September 4, 2015
Accepted February 19, 2016