© NIAB 2011
ISSN 1479-2621

Plant Genetic Resources: Characterization and Utilization (2011) 9(3); 445–453
doi:10.1017/S1479262111000645

AFLP-based molecular characterization and population structure analysis of *Silybum marianum* L.

Seyed Abolghasem Mohammadi[1,2]*, Majid Shokrpour[3], Mohammad Moghaddam[1,2] and Aziz Javanshir[1,2]

[1] *Department of Agronomy and Plant Breeding, Faculty of Agriculture, University of Tabriz, Tabriz, Iran,* [2] *Center of Excellence for Cereal Molecular Breeding, University of Tabriz, Tabriz, Iran and* [3] *Department of Agronomy and Plant Breeding, Faculty of Agriculture, University of Moghedas Ardabil, Tabriz, Iran*

Received 15 January 2011; Accepted 26 March 2011 – First published online 18 April 2011

Abstract

Amplified fragment length polymorphism analysis was employed to investigate the population structure of 32 Iranian *Silybum marianum* populations along with two commercial varieties. A total of 415 polymorphic marker loci were produced by 27 primer combinations with an average of 15.37 markers per combination. Polymorphic information content ranged from 0.24 to 0.44 with an average of 0.35 per primer combination, and marker index was in the range of 2.56–9.50 with an average of 5.37. The average Nei's genetic diversity (H_E) and Shannon's diversity index (I) were 0.201 and 0.296, respectively. The coefficient of differentiation among populations (G_{ST}) was 0.44, indicating that 44% of the total molecular diversity resulted from differences between populations. We identified three major groups based on cluster analysis and principal coordinate analysis, which were mostly in concordance with the geographical grouping of the populations. The molecular diversity estimate could be useful for selecting appropriate populations to improve *S. marianum* through conventional and molecular breeding.

Keywords: AFLP; analysis of molecular variance; genetic variation; *Silybum marianum*

Introduction

Iran is estimated to have as many as 7576 plant species (Ghahreman and Attar, 1999), and 12.5% of these are medicinal species (Schippmann *et al.*, 2002). This gene pool is currently under threat because of pasture exploitation, forest degradation and the introduction of intensive agriculture (Taeb, 1996). Bringing wild species into cultivation could help to reduce genetic erosion. This is necessary to develop improved varieties having useful attributes such as high yield, increased nutritional value and important metabolite composition. Comprehensive knowledge about the degree and patterns of within- and between-population genetic variation is an important prerequisite for initiating a breeding programme for medicinal plants.

Milk thistle or *Silybum marianum* (L.) Gaertn. (family *Compositae*), a medicinal plant with unique pharmaceutical properties for treating liver diseases, grows throughout various geographical areas in Iran. It is cultivated commercially for seed in Europe, Egypt, China and Argentina (Anonymous, 1995) and recently in a very limited area in Iran. Flavonolignan compounds, called silymarin, are the main basis of *Silybum* and are mainly found in seeds. Seeds contain 1–4% silymarin, which is a mixture of at least three flavonolignans: silybin,

* Corresponding author. E-mail: mohammadi@tabrizu.ac.ir

https://doi.org/10.1017/S1479262111000645 Published online by Cambridge University Press

silydianin and silychristin (Murphy *et al.*, 2000). Studies show that the active compounds, flavonolignans, have various effects such as anti-cancer (Davis-Searles *et al.*, 2005; Kroll *et al.*, 2007; Zi and Agarwal, 1999), anti-cholesterolaemia (Krecman *et al.*, 1998) and cytoprotection against hepatotoxins (Dvorak *et al.*, 2003). They are used especially in the treatment of several liver disorders (Morazzoni and Bombardelli, 1995).

In recent years, herbal remedies and supplements have represented a growing market worldwide. For example, in the United States, milk thistle was the 10th best-selling herbal dietary supplement in 2005 (Blumenthal *et al.*, 2006). Despite worldwide demand for silymarin (18–20 tons per year), research efforts on the domestication and breeding of this plant have been negligible (Ram *et al.*, 2005).

The determination of genetic diversity within and among populations is of great importance for the improvement of medicinal plants. Furthermore, the identification of genetic relationships among populations or genotypes is essential for the efficient utilization of plant genetic resources (Tang *et al.*, 2007). Molecular techniques, especially deoxyribonucleic acid (DNA)-based markers, provide effective tools for comprehensive genetic analysis of diversity and population structure. Among the various marker systems available at present, amplified fragment length polymorphism (AFLP) markers are widely used to study genetic diversity in medicinal plants (Rahimmalek *et al.*, 2009; Tang *et al.*, 2007; Tatikonda *et al.*, 2009; Zhu *et al.*, 2009). AFLP provides multi-locus and genomewide marker profiles. Some efforts have been made to analyse the genetic diversity and population structure in *S. marianum* using morphological (Shokrpour *et al.*, 2007) and biochemical (Shokrpour *et al.*, 2008) characteristics, but no study is available on the application of DNA-based markers for molecular characterization of *S. marianum* germplasm.

The present study assesses the molecular diversity in 32 populations of *S. marianum*, collected from seven provinces of Iran, as well as in two cultivars, using AFLP markers.

Materials and methods

Plant materials

A total of 32 populations were collected from natural habitats of *S. marianum* in Iran (Table 1), including seven from Ardabil, two from East Azarbaijan, six from Golestan, six from Mazandaran, eight from Khozestan, one from Lorestan and two from Fars province. In addition, two commercial varieties, CN seeds (Ltd., Ely, UK) and Budakalaszi (Budakalasz, Hungary), were used. For each population, fresh leaves from 15 field-grown plants were sampled, and individual plant sample was used for DNA isolation.

AFLP analysis

Total genomic DNA was extracted using the cetyltrimethylammonium bromide method (Saghai-Maroof *et al.*, 1984). The quantity and quality of the extracted DNA were evaluated by 0.8% agarose gel electrophoresis and a spectrophotometer. AFLP analysis was performed as described by Vos *et al.* (1995), except that 200 ng of genomic DNA was digested (3 h) with 1 U each EcoRI and Tru9I (an isoschizomer of MseI; Roche Diagnostic GmbH, Mannheim, Germany). The fragments were ligated with T4 DNA ligase to EcoRI (5 μM) and Tru9I (5 μM) adapters in a final volume of 50 ml at 37°C for 3 h. Pre-amplification reactions were performed with diluted DNA from the ligation reaction along with Tru9I/EcoRI primer pairs without selective nucleotides. Selective amplification was carried out using diluted DNA from the pre-amplification reaction and 64 primer pairs on a test panel of representative samples. After an initial denaturation step at 94°C for 2 min, selective amplification was done for 13 cycles of 30 s at 94°C, 30 s at 65°C with 0.7°C lowering for each cycle and 2 min at 72°C, followed by 24 cycles of 30 s at 94°C, 30 s at 56°C and 2 min at 72°C, and one final cycle of extension at 72°C for 10 min. The final PCR products were mixed with sequencing loading buffer, denatured for 5 min at 94°C and separated on a 6% denaturing polyacrylamide gel using the BioRad Sequi-Gen GT Sequencing Cell; the bands were visualized using silver staining (Bassam *et al.*, 1991). Finally, a set of 27 best primer combinations giving a reliable amplification and polymorphism were selected for analysis of all samples.

Data analysis

For statistical analysis, each polymorphic AFLP band was scored as a binary character for its absence (0) or presence (1). Each band was considered as a single and unique locus with Mendelian segregation. Individuals with more than 5% missing data were removed from analysis. Finally, data were compiled as a binomial matrix with 405 individuals and 415 polymorphic loci. The polymorphic information content (PIC) value for each AFLP primer combination was calculated as $PIC_i = 1 - f_i(1 - f_i)$, where PIC_i is the PIC of marker i, f_i is the frequency of the ith marker fragment when present and $1 - f_i$ is the frequency of the ith marker

https://doi.org/10.1017/S1479262111000645 Published online by Cambridge University Press

Table 1. Geographical location and environmental parameters for 32 populations of *S. marianum* collected across Iran

Province	Name	Altitude	Latitude	Longitude
Ardabil	ParsAbad	30	39°38′N	47°55′E
	GharaGhieh	240	38°30′N	47°46′E
	GhaleBabak	60	39°26′N	48°09′E
	Bilesavar	100	39°21′N	48°13′E
	RouhKandi	30	39°24′N	48°13′E
	Anjirlou	150	39°10′N	48°07′E
	GharaAghaj	600	39°01′N	47°40′E
East Azarbaijan	Tatar	320	39°30′N	46°58′E
	GharaChilar	420	38°52′N	46°32′E
Golestan	Gorgan	200	36°50′N	54°25′E
	Naharkhoran	420	36°44′N	54°28′E
	AghGhala	110	37°00′N	54°26′E
	AzadShahr	120	37°05′N	55°10′E
	Gonbad	120	37°16′N	55°09′E
	Kordkouy	100	36°47′N	54°06′E
Mazandaran	Behshahr	90	36°41′N	53°32′E
	Sari	400	36°33′N	53°03′E
	GhaemShahr	20	36°27′N	52°51′E
	Nour	30	36°34′N	51°57′E
	MahmoudAbad	20	36°37′N	52°14′E
	FereydounKenar	40	36°40′N	52°31′E
Khouzestan	Dezful	180	32°22′N	48°23′E
	Mollasani	20	31°35′N	48°52′E
	Andimeshk	200	32°27′N	48°20′E
	Abpakhsh	20	29°20′N	51°04′E
	Hamidieh	0	31°28′N	48°26′E
	Behbahan	300	30°35′N	50°14′E
	Shoush	85	32°11′N	48°14′E
	Ramhormoz	90	31°16′N	49°36′E
Lorestan	JolgeKhalaj	780	33°17′N	47°48′E
Fars	NourAbad	980	30°06′N	51°31′E
	Ghaemieh	800	29°36′N	51°39′E

fragment when absent (Roldain-Ruiz *et al.*, 2000). The marker index (MI) was calculated for each AFLP primer combination as $MI = PIC \times EMR$, where EMR is 'the effective multiplex ratio (E) defined as the product of the total number of loci/fragments per primer (n) and the fraction of polymorphic loci/fragments (β) ($E = n \cdot \beta$)' (Tatikonda *et al.*, 2009). Nei's (1978) unbiased genetic distance coefficient was used to estimate the genetic relationships between populations. The total genetic diversity ($H_T = H_S + D_{ST}$), as well as that within populations (H_S) and among populations (D_{ST}), was also calculated. The proportion of among-population total genetic diversity (G_{ST}) was calculated as the ratio D_{ST}/H_T. In addition, gene flow (N_m) was estimated by $N_m = 0.25 \times (1\text{-}G_{ST})/G_{ST}$. All genetic diversity parameters were estimated by PopGene software version 1.32. Analysis of molecular variance (AMOVA) was performed to partition the molecular genetics variance into components attributable to the variance between and within ecotypes (Excoffier *et al.*, 1992). AMOVA was carried out in Arlequin version 3.11 software (Excoffier *et al.*, 2005). The dendrogram was constructed based on a neighbour-joining algorithm and pairwise unbiased Nei's genetic distances using MEGA V3.0 (Kumar *et al.*, 2004). Principal coordinate analysis (PCoA) was also performed using GenAlEex v6.2 software (Peakall and Smouse, 2006).

Results

Marker polymorphism

The 27 primer combinations selected from 64 tested pairs produced a total of 415 polymorphic AFLP markers, in the range of 100–700 bp, across 32 *S. marianum* populations as well as two commercial varieties (Table 2). The number of polymorphic fragments per primer combination ranged from 6 (E-ACT/T-CAA) to 32 (E-ACA/T-CAG), with an average of 15.37. The PIC and MI values for each primer pair are shown in Table 2. The primer combinations E-ACA/T-CTA, E-ACA/T-CTG,

https://doi.org/10.1017/S1479262111000645 Published online by Cambridge University Press

Table 2. AFLP, number of polymorphic bands, PIC and MI per primer combination

Primer combination	Polymorphic bands	PIC	MI
E-ACA/T-CAG	25	0.38	9.50
E-ACA/T-CTT	32	0.29	9.12
E-ACC/T-CTC	21	0.24	5.03
E-ACA/T-CAA	15	0.34	5.13
E-ACA/T-CAT	19	0.33	6.27
E-ACC/T-CTT	12	0.38	4.51
E-ACC/T-CAG	12	0.33	4.00
E-ACA/T-CTA	19	0.40	7.66
E-ACA/T-CAC	22	0.34	7.55
E-ACA/T-CTC	31	0.29	9.10
E-ACA/T-CTG	15	0.41	6.14
E-ACC/T-CTA	13	0.30	3.89
E-ACC/T-CAA	12	0.34	4.05
E-ACC/T-CTG	10	0.28	2.81
E-ACC/T-CAT	8	0.37	2.98
E-ACG/T-CAG	18	0.39	6.96
E-ACG/T-CTG	10	0.40	4.04
E-ACG/T-CAT	11	0.34	3.72
E-ACG/T-CAC	18	0.39	7.05
E-ACG/T-CTC	17	0.44	7.48
E-ACG/T-CAA	8	0.36	2.89
E-AGG/T-CTA	10	0.28	2.82
E-AGG/T-CTC	17	0.42	7.23
E-AGG/T-CTT	11	0.35	3.86
E-ACT/T-CAA	6	0.43	2.56
E-ACT/T-CAC	12	0.28	3.41
E-AGC/T-CTA	11	0.31	3.43
Total	415	–	–
Mean	15.37	0.35	5.37

E-ACG/T-CTG, E-ACG/T-CTC, E-AGG/T-CTC and E-ACT//T-CAA showed high PIC values (over 0.40), with the highest value for E-ACG/T-CTC (0.44). Moreover, E-ACA/t-CAG, E-ACA/T-CTT and E-ACA/T-CTC showed high MI values (over 9.00). The average values for PIC and MI were 0.35 and 5.37, respectively.

Genetic diversity

The average Nei's genetic diversity index (H_E) and Shannon's diversity index (I) were 0.201 and 0.296, respectively. The highest and lowest within-population genetic diversity was observed in Sari (0.300 and 0.432) and Gonbad (0.096 and 0.140) populations, respectively, as revealed by Nei and Shannon's diversity indices (Table 3). The Hamidieh population also showed high genetic homogeneity.

The total gene diversity (H_T) over the 34 *S. marianum* populations was 0.360, the mean diversity within populations (H_S) was 0.201 and the coefficient of differentiation among populations (G_{ST}) was 0.440, indicating that 44% of the total molecular diversity resulted from differences between populations. Gene flow among populations (N_m) was estimated at 0.318. The distribution of genetic diversity within and between *S. marianum* populations was explored using AMOVA, which indicated significant variance within and among milk thistle populations. Although, as with G_{ST}, the level of within-population genetic variation was higher than that of among-population genetic variation, and the AMOVA revealed that 72.71% of the total molecular variance is attributable to within-population genetic diversity. Among the populations studied, the maximum genetic distances were observed between Gonbadan from the northwest and Hamidieh, Andimeshk, Shoush as well as Behbahan from southern Iran. This result is an overall relationship with geographical distances of their collection sites. Ghaemshahr and Gorgan populations from northern Iran showed the lowest genetic distance.

Genetic relationships

The genetic separation of *S. marianum* populations based on their geographical regions supported the suitability of the AFLP technique for differentiating closely related populations. In the cluster analysis (Fig. 1), two major preliminary clusters could be identified for *S. marianum* populations. Group one consisted of two subclusters, including northern and northwestern populations. The populations from northwestern provinces (Ardabil and East Azarbaijan) were grouped in subcluster I, whereas the populations from the northern provinces Golestan and Mazandran were assigned to subcluster II. Group two included populations from the southern provinces. In this clustering, there were some inconsistence between molecular grouping and origin of population. Such that, Dezfoul and Mollasani populations (Khouzestan province from south of Iran) were grouped with the northern and northwestern populations, and Ghachillar (East Azarbayjan province), GhaemShahr (Mazandaran province) and Gorgan (Golestan province) populations were clustered with the southern populations, albeit in a separate subcluster.

To understand the genetic relationships of *S. marianum* populations as individuals, PCoA was also conducted based on the AFLP data matrix of 415 fragments for 34 populations. The scatter plot of the first and second principal components explaining 67% of the total molecular variation showed a clear genetic variation and differentiation pattern for *S. marianum* populations based on their geographical regions (Fig. 2). There were three separate clusters apparent, consistent with the grouping recovered by neighbour-joining analysis. One more, the two populations from the south (Dezfoul and Mallasani) were positioned outside of the cluster for

https://doi.org/10.1017/S1479262111000645 Published online by Cambridge University Press

Table 3. Within-population genetic diversity of 34 *S. marianum* L. accessions as revealed by Nei's and Shannon's diversity indices

Province/country	Name	Shannon's index	Nei's index
Ardabil	ParsAbad	0.341	0.231
	GharaGhieh	0.431	0.296
	GhaleBabak	0.358	0.247
	Bilesavar	0.284	0.195
	RouhKandi	0.256	0.175
	Anjirlou	0.393	0.272
	GharaAghaj	0.277	0.187
East Azarbaijan	Tatar	0.205	0.138
	GharaChilar	0.425	0.285
Golestan	Gorgan	0.392	0.262
	Naharkhoran	0.280	0.192
	AghGhala	0.401	0.275
	AzadShahr	0.417	0.285
	Gonbad	0.140	0.096
	Kordkouy	0.185	0.127
Mazandaran	Behshahr	0.268	0.184
	Sari	0.432	0.300
	GhaemShahr	0.323	0.216
	Nour	0.369	0.255
	MahmoudAbad	0.350	0.240
	FereydounKenar	0.428	0.295
Khouzestan	Dezful	0.334	0.228
	Mollasani	0.269	0.179
	Andimeshk	0.200	0.137
	Abpakhsh	0.229	0.156
	Hamidieh	0.149	0.101
	Behbahan	0.152	0.103
	Shoush	0.211	0.144
	Ramhormoz	0.234	0.157
Lorestan	JolgeKhalaj	0.229	0.157
Fars	NourAbad	0.183	0.124
	Ghaemieh	0.171	0.116
UK	CNseeds	0.356	0.242
Hungary	Budakalaszi	0.380	0.258
	Average	0.296	0.201

southern populations, which is indicative of possible introgression.

Discussion

Analysis of genetic variation

Bringing medicinal plants into cultivation necessitates to determine the level of genetic diversity in the wild germplasm and the application of traditional and biotechnological genetic techniques, both to improve yield and uniformity and to modify potency or toxicity. Although *S. marianum* is among one of the most ancient of all known herbal medicines and its derivatives have been used as herbal remedies for almost 2000 years (Sánchez-Sampedro *et al.*, 2008), according to our knowledge, no reports exist on *S. marianum* genetic diversity at the DNA level on this species. As is the case for other medicinal plant species such as *Siraitia grosvenorii* (Tang *et al.*, 2007), *Jatropha curcas* (Tatikonda *et al.*, 2009), *Incarvillea younghusbandii* (Zhu *et al.*, 2009) and *Achillea* species (Rahimmalek *et al.*, 2009), AFLP analysis was effective in detecting genetic variation in the *S. marianum* genome. The efficiency of a molecular marker technique depends upon the amount of polymorphism it can detect among the genotypes under investigation (Tatikonda *et al.*, 2009). High levels of polymorphism were obtained with 27 AFLP primer combinations from the 32 populations collected in seven provinces of Iran as well as two introduced varieties. All populations studied were polymorphic, and AFLP revealed a large number of polymorphic DNA fragments. Schmidt and Jensen (2000) reported that the accuracy of genetic diversity analysis and genetic differentiation increases with the use of an increasing number of loci.

https://doi.org/10.1017/S1479262111000645 Published online by Cambridge University Press

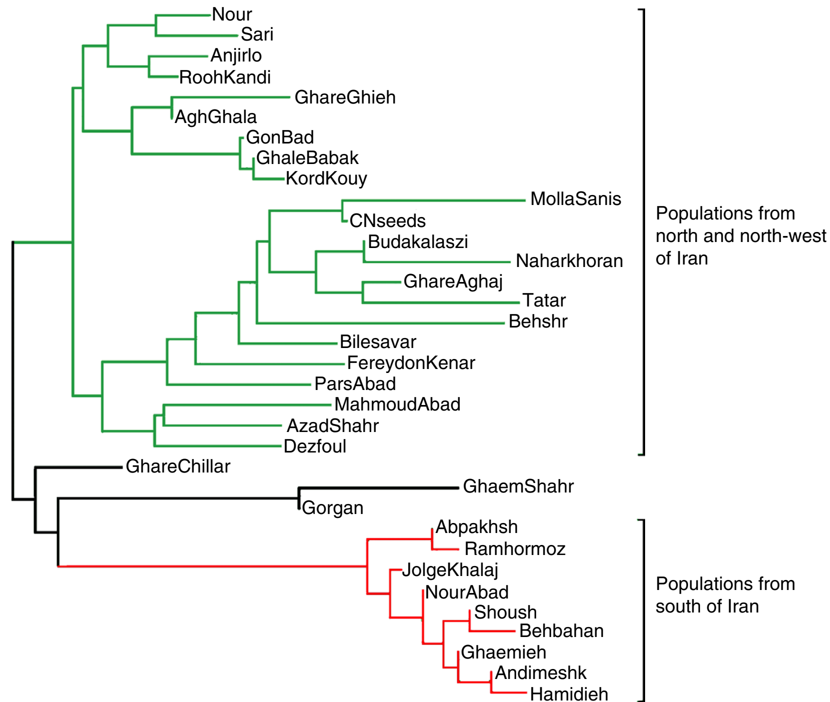

Fig. 1. Neighbour-joining tree for 32 populations and two commercial varieties of *S. marianum* based on Nei's unbiased pairwise genetic distance (A colour version of this figure can be found online at journals.cambridge.org/pgr).

Discriminatory power of AFLP primer combinations

A number of marker attributes such as PIC and MI have been used in studies to assess the informativeness or discriminatory power of AFLP markers for genetic diversity analysis in medicinal plant species (Rahimmalek *et al.*, 2009; Tang *et al.*, 2007; Tatikonda *et al.*, 2009; Zhu *et al.*, 2009). Although PIC has been used most extensively in the majority of marker-based diversity studies, MI is a convenient estimate for marker efficiency.

Calculating PIC values for different AFLP markers obtained by a particular primer combination revealed an average PIC value of 0.35 across 34 *S. marianum* accessions, with a range of 0.24–0.44. The maximum

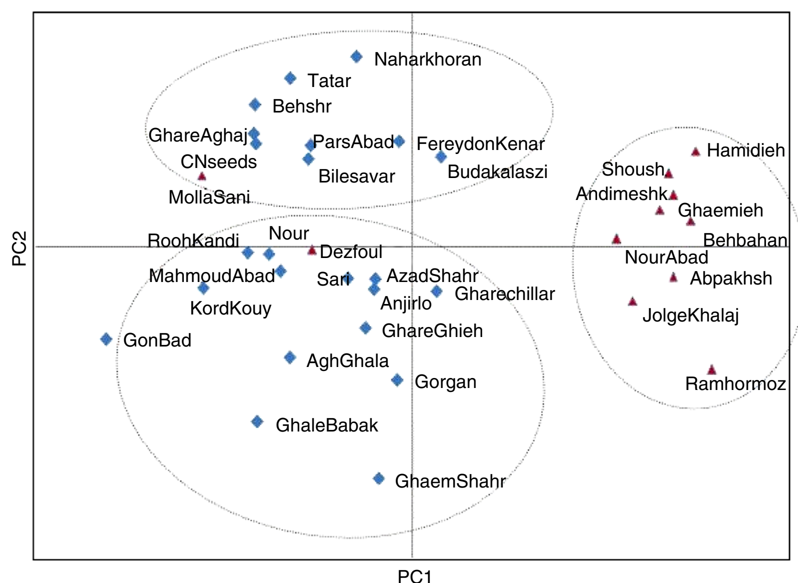

Fig. 2. PCoA of 32 populations and two commercial varieties of *S. marianum* based on 415 polymorphic AFLP markers (A colour version of this figure can be found online at journals.cambridge.org/pgr).

https://doi.org/10.1017/S1479262111000645 Published online by Cambridge University Press

PIC value for a biallelic marker such as AFLP is 0.50. In our study, six primer combinations showed PIC values >0.40. It was found that fragments occurring in 40–70% of individuals have high PIC values. For dominant markers such as AFLP, MI together with PIC has been used to assess the informativeness of the markers for medicinal plant species such as *Jatropha curcas* (PIC = 0.26, MI = 25.13; Tatikonda *et al.*, 2009) and *Valerianella locusta* (PIC = 0.25, MI = 4.47; Muminovic *et al.*, 2004). In our study of *S. marianum*, the MI varied from 2.56 to 9.50 with an average of 5.37. In general, the PIC and MI values were comparable with those reported for *Valerianella locusta* (Muminovic *et al.*, 2004), wheat (Bohn *et al.*, 1999) and soybean (Powell *et al.*, 1996). Various studies compared the utility of restriction fragment length polymorphism, simple sequence repeat (SSR), random amplified polymorphic DNA and AFLP markers for germplasm analysis and reported a high MI for AFLPs compared with other systems (Bohn *et al.*, 1999; Hongtrakul *et al.*, 1997; Powell *et al.*, 1996; Russell *et al.*, 1997). Therefore, AFLP markers have been recommended for fingerprinting genotypes and genetic diversity analysis especially where codominant markers such as SSRs are not available.

Population diversity and differentiation

The population differentiation in the analysed germplasm was assessed based on genotyping data obtained for all 27 primer combinations, using a neighbour-joining-based phenogram and PCoA. The majority of the populations were grouped in accordance with their geographical locations, with some minor deviation. Congruity between genetic distance and geographical distance has been reported for other plants such as Moringa (Muluvi *et al.*, 1999), Pistachia (Hormaza *et al.*, 1994) and barley (Pakniyat *et al.*, 1997). Furthermore, the grouping pattern of *S. marianum* populations based on molecular data were in agreement with the pattern reported for the same populations by Shokrpour *et al.* (2008) using morphological characteristics. Pairwise genetic distances between *S. marianum* populations also corresponded well with geographical distances. The maximum genetic distance was observed between Gonbad and Ramhormoz populations from northern and southern Iran, respectively, which also represent a large geographical distance.

The genetic structure patterns of *S. marianum* populations using both Nei's genetic diversity analysis and AMOVA suggested that within-population genetic variation was higher than the genetic differentiation among populations. Similar patterns of genetic variation were reported for *Valerianella locusta* (Muminovic *et al.*, 2004), Tunisian fig (*Ficus carica* L.; Baraket *et al.*, 2009) and *Caragana microphylla* (Chen *et al.*, 2009). It is widely accepted that breeding systems and seed dispersal mechanisms in particular are associated with levels of genetic variation within and among populations (Hamrick and Godt, 1996). Seed dispersal is the only mean by which *S. marianum* spreads. The seeds are equipped with a large pappus that allows effective spreading by wind. Although wind dispersal of seeds may be highly localized, the large pappus of *S. marianum* seeds allows them to disperse over long distances (kilometers). The long distance tail of seed dispersal distribution may cause gene flow among populations, and even small amounts of gene flow may have significant consequences for the homogenization of genetic variation among populations. At the within-population level, however, localized seed dispersal can generate significant fine-scale genetic structure, even in the face of evolutionarily significant rates of inter-populational gene flow (Chung *et al.*, 2004; Trapnell *et al.*, 2008), thereby minimizing the effects of isolation and population differentiation. The large seed pappus in *S. marianum* may also promote efficient seed exchange, hindering differentiation of populations.

Low within-population genetic diversity in the Gonbad and Hamidieh populations from northern and southern regions of Iran, respectively, could be explained by their specific geographical locations and the direction of dominant winds in the area. Gonbad is located in northeastern Iran and is surrounded by the Caspian Sea from the west, and the wind direction is mostly from west to east. Hamidieh is located at the southwestern border of Iran, and due to the direction of winds in this area, it could be considered as an isolated population such that seeds from nearby locations such as Molasani and Shoush do not infiltrate.

Conclusions

Knowledge of genetic variation and the genetic relationships between genotypes is an important consideration for efficient utilization of germplasm resources. Furthermore, it is important for the optimal design of plant breeding programmes and influencing the choice of genotypes to cross for development of new populations (Russell *et al.*, 1997). The present data on the patterns of genetic variation suggest that milk thistle could be improved for various desirable traits. To date, studies of genetic variation in *S. marianum* relied on phenotypic and biochemical assays (Ram *et al.*, 2005; Shokrpour *et al.*, 2007; Shokrpour *et al.*, 2008). AFLPs have proven to be a robust and proficient tool to produce large numbers of informative markers that reveal intrapopulation

https://doi.org/10.1017/S1479262111000645 Published online by Cambridge University Press

diversity and that estimate the genetic distance between individuals and populations.

Acknowledgements

The present study was funded by the Center of Excellence for Molecular Plant Breeding, Department of Agronomy and Plant Breeding, Faculty of Agriculture, University of Tabriz, Tabriz, Iran, and the Iran National Science Foundation (INSF) (Grant No. 84153).

References

Anonymous (1995) Market report 1995. *HerbalGram* 39: 68.

Baraket G, Chatti K, Saddoud O, Mars M, Marrakchi M, Trifi M and Salhi-Hannachi A (2009) Genetic analysis of Tunisian fig (*Ficus carica* L.) cultivars using amplified fragment length polymorphism (AFLP) markers. *Scientia Horticulture* 120: 487–492.

Bassam BJ, Caetano-Anolles G and Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry* 196: 80–83.

Blumenthal M, Ferrier GKL and Cavaliere C (2006) Total sales of herbal supplements in United States show steady growth. *HerbalGram* 71: 64–66.

Bohn M, Utz HF and Melchinger AE (1999) Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs, and SSRs and their use for predicting progeny variance. *Crop Science* 39: 228–237.

Chen X, Gao Y, Zhao N, Zhao T and Zhu M (2009) An AFLP analysis of genetic diversity and structure of *Caragana microphylla* populations in Inner Mongolia steppe, China. *Biochemical Systematics and Ecology* 37: 395–401.

Chung MY, Nason JD and Chung MG (2004) Spatial genetic structure in populations of the terrestrial orchid *Cephalanthera longibrateata* (Orchidaceae). *American Journal of Botany* 91: 52–57.

Davis-Searles PR, Nakanishi Y, Kim NC, Graf TN, Oberlies NH, Wani MC, Wall ME, Agraval R and Kroll DJ (2005) Milk thistle and prostate cancer: differential effects of pure flavonolignans from *Silybum marianum* on antiproliferative end points in human prostate carcinoma cells. *Cancer Research* 65: 4448–4457.

Dvorak Z, Kosina P, Walterova D, Simanek V, Bachleda P and Ulrichova J (2003) Primary cultures of human hepatocytes as a tool in cytotoxicity studies: cell protection against model toxins by flavonolignans obtained from *Silybum marianum*. *Toxicology Letters* 173: 201–212.

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

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

Ghahreman A and Attar F (1999) *Biodiversity of Plant Species in Iran*. Vol. 1. Tehran: Tehran University Publications.

Hamrick JL and Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions: Biological Science* 351: 1291–1298.

Hongtrakul V, Huestis GM and Knapp SJ (1997) Amplified fragment length polymorphism as a tool for DNA fingerprinting sunflower germplasm: genetic diversity among oilseed inbred lines. *Theoretical Applied Genetics* 95: 400–407.

Hormaza JI, Dollo L and Polito VS (1994) Determination of relatedness and geographical movements of *Pistacia vera* L. (pistachio; Anacardaceae) germplasm by RAPD analysis. *Economic Botany* 48: 349–358.

Krecman V, Skottova N, Walterova D, Ulrichova J and Simanek V (1998) Silymarin inhibits the development of diet-induced hypercholesterolemia in rats. *Planta Medica* 64: 138–142.

Kroll DJ, Shaw HS and Oberlies NH (2007) Milk thistle nomenclature: why it matters in cancer research and pharmacokinetic studies. *Integrative Cancer Therapy* 6: 110–119.

Kumar S, Tamura K and Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics* 5: 150–163.

Morazzoni P and Bombardelli E (1995) *Silybum marianum* (*Carduus marianus*). *Fitoterapia* 66: 3–42.

Muluvi GM, Sprent JI, Soranzo N, Provan J, Odee D, Folkard G, McNicol JW and Powell W (1999) Amplified fragment length polymorphism (AFLP) analysis of genetic variation in *Moringa oleifera* Lam. *Molecular Ecology* 8: 463–470.

Muminovic J, Melchinger AE and Lubberstedt T (2004) Genetic diversity in cornsalad (*Valerianella locusta*) and related species as determined by AFLP markers. *Plant Breeding* 123: 460–466.

Murphy JM, Caban M and Kemper KJ (2000) Milkthistle (*Silybum marianum*). http://www.mcpedu/herbal/default.htm

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583–590.

Pakniyat H, Powell W, Baird E, Handley LL, Robinson D, Scrimgeour CM, Hackett CA, Forster BP, Nevo E and Caligari PDS (1997) AFLP variation in wild barley (*Hordeum spontaneum* C. Koch) with reference to salt tolerance and associated ecogeography. *Genome* 40: 332–341.

Peakall R and Smouse PE (2006) GENALEX 6: genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S and Refalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* 2: 225–238.

Rahimmalek M, Sayed Tabatabaei BE, Arzani A and Etemadi N (2009) Assessment of genetic diversity among and within *Achillea* species using amplified fragment length polymorphism (AFLP). *Biochemical Systematics Ecology* 37: 354–361.

Ram G, Bhan MK, Gupta KK, Thaker B, Jamwal U and Pal S (2005) Variability pattern and correlation studies in *Silybum marianum* Gaertn. *Fitoterapia* 76: 143–147.

Roldain-Ruiz I, Calsyn E, Gilliand TJ, Coll R, van Eijk MJT and De Loose M (2000) Estimating genetic conformity between related ryegrass (*Lolium*) varieties, 2. AFLP characterization. *Molecular Breeding* 6: 593–602.

Russell JR, Fuller JD, Macaulay M, Hatz BG, Jahoor A, Powell W and Waugh R (1997) Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theoretical Applied Genetics* 95: 714–722.

Saghai-Maroof MA, Soliman K, Jorgensen RA and Allard RW (1984) Ribosomal DNA spacerlength polymorphisms in

https://doi.org/10.1017/S1479262111000645 Published online by Cambridge University Press

barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proceeding of National Academic Science USA* 81: 8014–8018.

Sánchez-Sampedro MA, Peláez R and Corchete P (2008) An arabinogalactan protein isolated from medium of cell suspension cultures of *Silybum marianum* L. Gaernt. *Carbohydrate Polymers* 71: 634–639.

Schippmann U, Leaman DJ and Cunningham AB (2002) Impact of cultivation and gathering of medicinal plants on biodiversity: global trends and issues. In: *FAO, Biodiversity and the Ecosystem Approach in Agriculture, Forestry and Fisheries*. Rome: Inter-Departmental Work Group on Biological Diversity for Food and Agriculture, pp. 1–21.

Schmidt K and Jensen K (2000) Genetic structure and AFLP variation of remnant populations in the rare plant *Pedicularis palustris* (scrophulariaceae) and its relation to population size and reproductive components. *American Journal Botany* 87: 678–689.

Shokrpour M, Moghaddam M, Mohammadi SA, Ziai SA and Javanshir A (2007) Genetic properties of milk thistle ecotypes from Iran for morphological and flavonolignan characters. *Pakistan Journal of Biological Science* 10: 3266–3271.

Shokrpour M, Mohammadi SA, Moghaddam M, Ziai SA and Javanshir A (2008) Variation in flavonolignan concentration of milk thistle (*Silybum marianum*) fruits grown in Iran. *Journal of Herbs, Spices and Medicinal Plants* 13: 54–68.

Taeb M (1996) *Islamic Republic of Iran: Country Report to the FAO International Technical Conference on Plant Genetic Resource*. 17-23 June. Leipzig: FAO.

Tang SQ, Bin XY, Peng YT, Zhou JY, Wang L and Zhong Y (2007) Assessment of genetic diversity in cultivars and wild accessions of Luohanguo (*Siraitia grosvenorii* [Swingle] A.M. Lu et Z.Y. Zhang), a species with edible and medicinal sweet fruits endemic to southern China, using RAPD and AFLP markers. *Genetic Resources and Crop Evolution* 54: 1053–1061.

Tatikonda L, Wani SP, Kannan S, Beerelli N, Sreedevi TK, Hoisington DA, Devi P and Varshney RK (2009) AFLP-based molecular characterization of an elite germplasm collection of *Jatropha curcas* L., a biofuel plant. *Plant Science* 176: 505–513.

Trapnell DW, Schmidt JP and Hamrick JL (2008) Spatial genetic structure of the southeastern North American endemic, *Ceratiola ericoides* (Empetraceae). *Journal of Heredity* 99: 604–609.

Vos P, Hogers R, Bleeker M, Reijans M, Van Delee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407–4414.

Zhu Y, Geng Y, Tersing T, Liu N, Wang Q and Zhong Y (2009) High genetic differentiation and low genetic diversity in *Incarvillea younghusbandii*, an endemic plant of Qinghai-Tibetan Plateau, revealed by AFLP markers. *Biochemical Systematics Ecology* 37: 589–596.

Zi X and Agarwal R (1999) Silibinin decreases prostate: specific antigen with cell growth inhibition via G1 arrest, leading to differentiation of prostate carcinoma cells: implications for prostate cancer intervention. *Proceeding of National Academic Science USA* 96: 7490–7495.

https://doi.org/10.1017/S1479262111000645 Published online by Cambridge University Press