Australian Journal of Crop Science AJCS
AJCS 5(1):82-91 (2011) ISSN:1835-2707

A comparative study of morphological and molecular diversity analysis among cultivated almonds (*Prunus dulcis*)

S. Kadkhodaei[1,2*], M. Shahnazari[2], M. Khayyam Nekouei[3], M. Ghasemi[4], H. Etminani[2], A. Imani[5], Arbakariya B. Ariff[1]

[1]Department of Bioprocess Technology, Faculty of Biotechnology & Biomolecular Sciences, University Putra Malaysia, Serdang, 43400. Selangor, Malaysia
[2]Department of Genomics, Agricultural Biotechnology Research Institute of Iran (ABRII)- Central Region, Isfahan 85135-487, Iran
[3]Agricultural Biotechnology Research Institute of Iran, Seed & Plant Improvement Campus, Mahdasht Rd., Karaj, Iran
[4]Planning Department and Civil Affairs of Isfahan Province Governor, Isfahan, Iran
[5]Department of Horticulture, Faculty of Agriculture, Abhar Branch, Islamic Azad University, Abhar, Iran; Departments of Horticulture, Seed and Plant Improvement Campus, Mahdasht Rd., Karaj, Iran

***Corresponding author: gs24152@mutiara.upm.edu.my, skad@abrii.ac.ir**

Abstract

In this study, a total of 39 morphological traits and 9 simple sequence repeat (SSR) loci were used (i) to study the morphological and genetic diversity among 53 selected almond (*Prunus dulcis*) cultivars/landraces with different geographical origins, (ii) to assess the level of correlation between phenotypic and nuclear genetic distance matrices, and (iii) to classify the accessions into groups based on molecular profiles and morphological traits. The analysis of the morphological data revealed significant differences among genotypes for all measured traits. The number of alleles per locus varied from 8 to 17 with an average of 12.86. The UDA-002 and UDA-005 markers were the most informative revealing 17 and 15 polymorphic alleles distributed among all genotypes, respectively, but considering polymorphism information content (PIC), UDA-023 was the most informative one. According to the results, the average values for number of effective alleles (5.59), Shanon index (1.97), expected heterozygosity (0.8) and PIC (0.89, ranging from 0.69 to 0.95) were relatively high. The mantel matrix correspondence test was used to compare the molecular and morphological similarity matrices. Although the correlation coefficient was low (r=0.02266, p=0.6926) but both morphological and molecular dendrogram clustered the genotypes into two main groups. The mean morphological similarity (0.59) was high in comparison to similarity calculated using SSR markers (0.23). The results will be useful for collections, conservation and various almond breeding programs.

Keywords: Agronomic traits, almond, genetic similarity, microsatellite, molecular marker, Morphology, phenotypic diversity, *Prunus*, SSR.

Abbreviations: SSR: Simple Sequence Repeats, PIC: Polymorphism Information content, QTL: Quantitative Trait Loci, MAS: Marker Assisted Selection, PCA: Principal Components Analysis.

Introduction

Prunus & Almond. The genus Prunus in Prunoideae, a subfamily of Rosaceae, includes several economically important fruit tree species such as apricot (P. armeniaca), cherry (P. avium), peach (P. persica), plum (P. mume), and common almond (P. dulcis Mill. syn. P. communis Fritsch.). Almond was probably domesticated during the 3rd millennium BC. It has been suggested that this domestication could have taken place in Central Asia where wild almond trees, can still be found (Shiran et al., 2007). Twenty-six almond species form a distinct and easily identified taxonomic group in the world (Browicz, 1996). In Iran 21 almond species and 6 natural hybrids have been described (Khatamsaz, 1992). The basic chromosome number (x) of almonds (common and wild types) is 8 (2n=16) and its DNA contents is relatively small (0.54–0.67 pg DNA/2C) (Dickson, 1992; Kadkhodaei, 2008). Iran is amongst the most important countries in the world producing Prunus fruits, as its ranking regarding to production of cherries, apricot and almond is the first, second and fourth, respectively. In 2007, worldwide annual production of almonds is exceeded 2065489 tons, which the contribution of Iran is 5.32% (110000 tons). Almond harvesting area in Iran is about 172000 ha and annual yield in 2007 has been 639.5 kg/ha (FAOSTAT Data Sources, 2007). Almond cultivation in Iran has a long historical background, and because of its self-incompatibility nature there are many genotypes growing in different regions of the country. These genotypes include a vast range of diversity in many characteristics such as blooming time. One of the major concerns of modern agriculture is the conservation and utilization of valuable genetic resources of crop plants. The need for correct identification applies to cultivars and accessions, independently of their mode of conservation, i.e. whether they are maintained in an in situ or ex situ field gene bank or an in vitro gene bank. Tools developed for the characterization for biodiversity may allow clarifications of synonyms and detection of the origin of species and cultivars (Maghuly, 2004). Traditionally, the identification and characterization of cultivars

and species has been based on morphological and physiological traits which have some limitation. Since morphological markers are prone to equivocal interpretations and time consuming, not always available for analysis and are affected by changing environmental conditions, molecular marker technology offers several advantages over the sole use of conventional markers in cultivar identification and breeding programs. Molecular markers help to distinguish labeling mistakes, identification of the genuine owner of the cultivar in question, routine identification of cultivars in nurseries. Further it simplifies work in breeding programs by accelerating the breeding process by allowing a selection before the first fruit crop, by tracking certain genes or genotypes among offspring of crosses. Molecular markers developed for Prunus also offer a powerful tool to study the evolution of the genome, and for understanding of genome structure and determinants of genetic diversity (Maghuly, 2004; Martínez-Gómez, 2005). Microsatellites or simple sequence repeat (SSR) are sequences of a few repeated and adjacent basepairs, well distributed over the eukaryote genome (Powell, 1996). PCR-based, SSR markers are becoming the marker of choice for fingerprinting and genetic diversity studies for a wide range of plants (Gupta, 1996). Because of their high polymorphism, abundance, and codominant inheritance, they are well suited for the assessment of genetic variability within crop species, and of the genetic relationships among species. In the case of Prunus species, primer pairs flanking SSRs have been cloned and sequenced in peach (Dirlewanger, 2002), apricot (Decroocq, 2003), cherry (Schueler, 2003) and almond (Messina, 2004; Testolin, 2004; Xu, 2004). These SSR markers have been used primarily for the molecular characterization and identification of cultivars in the related Prunus species. Also SSRs developed from different Prunus species have high potential for transportability across each other (Fathi, 2008; Mnejja, 2005; Shiran, 2007). Molecular markers have been applied efficiently in almond genus for development of genetic linkage maps (Gregory, 2004; Joobeur, 2000), mapping genes and QTL analysis (Sánchez-Pérez, 2007; Ballester, 2001), identification of the specific genotypes and marker assisted selection (Channuntapipat, 2003; Sánchez-Pérez, 2010) and DNA fingerprinting (Kadkhodaei et al., 2010). In this work, microsatellite markers (SSR) besides some morphological traits were used in the characterization of 53 almond genotypes which most of them are Iranian along with some foreign cultivars, aiming to produce a molecular identification key for them. The other objectives of the present study were investigation of the genetic diversity of Iranian cultivated almond germplasm, identification of their relationship to important foreign cultivars, and introduction of informative markers for important agronomical traits using recently developed microsatellite markers for almond. Also assessment to the genetic diversity of Iranian almond germplasm can be considered as a valuable basic study for future breeding programs.

Materials and methods

Plant material

Fifty three almond genotypes from different origins distributed throughout Iran besides some foreign cultivars and their hybrids with Iranian ones were used in this study. The genotypes maintained at an experimental orchard located at the Shahid Fozveh Agricultural Research Station in Isfahan, Iran (Table 1).

Morphological traits

Seventeen quantitative and 22 qualitative characters were recorded during 2006 and 2007 for 53 almond genotypes (Table 2). The traits include characteristics of kernel (Kernel Length-KL, Kernel Width-KW, Kernel Thickness-KT, Kernel Weight-KWG, Kernel/Nut W %-KNWP, Kernel Color-KC, Double Kernel-DK, Shriveling of Kernel-SK, Kernel Pubescence-KP), leaf (Petiole Length-PL, Leaf Width-LW, Leaf Length-LL, Leaf Area-LA, Leaf Basal Shape-LBS, Leaf Shape-LS, Leaves Arrangement), Pests and Diseases (Bacterial Canker-SR, Black Aphid-ABR, Twig Borer-SKR, Green Aphid-AR), Tree (Growth habit-GH), Flower (Bearing Habit-BH, Flower Density-FDe, Double Flower in buds-DF, Flower Color-FC, Initial Flowering-F5, Completed Flowering-F100, Duration of Flowering-FDu) and Nut (Nut Length-NL, Nut Width-NW, Nut Thickness-NT, Nut Weight-NWG, Nut Shape-NS, Nut Color-SC, Marking of outer Shell-MS, Suture Opening-SO, Shell Retention-SR, Ease of Hulling-EH, Softness of Shell-SS).

Molecular characterization

DNA extraction. DNA was extracted based on Kadkhodaei et al. (2006). The purified total DNA was quantified by gel electrophoresis, and its quality verified by spectrophotometry. DNA samples were stored at 4°C. Microsatellite analysis. Extracted almond genomic DNA was PCR-amplified using 9 pair flanking SSR sequences previously cloned and sequenced specifically in almond (Testolin, 2004). Details of the microsatellites analyzed and their origin are given in Table 3. As a parameter of for the selection of some of these microsatellites we have used the genetic informativeness and the technical aspects of the analysis. None of the microsatellites used, have not been studied extensively for almonds earlier, and almost all of the previous studies on almonds have been carried out using peach isolated SSRs. Amplification was conducted in a total volume of 15 µl with 10 ng DNA, 0.05 µM of both specific primers, 0.2mM of 10 mM dNTP mix, 0.1 U Taq polymerase, 1.5 mM $MgCl_2$, and 1×PCR buffer (100 mM Tris HCl, pH 8.3, 500 mM KCl). The amplification program consisted of 3 min at 94 °C, 34 cycles of 45 s at 94 °C, 45 s at the appropriate annealing temperature and 45 s at 72 °C, followed by a 5-min extension at 72 °C. Five micro-liters of the PCR product were separated on a 1% agarose gel to check the PCR amplification and determine approximated size of the amplified fragments. Then electrophoresis of PCR products was carried out on a 6-10 % polyacrylamid gel (depended on allele ranges in each locus) in 0.5 X TBE. The gels were run for 3 h at constant power (120 W) in ATTO Electrophoresis system. A 50-bp DNA Ladder (Fermentas) was used as a molecular size standard in each gel.

Data Analysis

Morphological data. The Gower general similarity coefficient (Gower, 1971) was used in cluster analysis of morphological traits. Unweighted Pair-Group Method using Arithmetic average (UPGMA) analysis was performed with the program MVSP (version 3.13b, Kovach Computing Services, Anglesy, Wales, UK). The software statistical analysis system (SAS, 1998 V.8) was used for Principal Components Analysis (PCA) to define Eigenvalues and Eigenvectors and also for comparison of the mean of groups to define effective traits in separation of the groups.

Table 1. The almond cultivars used in the present study.

No.	CODE	GENOTYPE	LOCATION	ORIGIN	No.	CODE	GENOTYPE	LOCATION	ORIGIN
1	SH-P1	Poostnazok1-1.2	Shiraz	Iran	32	H-13	Hamedan13	Hamedan	Iran
2	SH-M	MonaghaShirazi	Shiraz	Iran	33	H-8	Hamedan8	Hamedan	Iran
3	SH-8	Shiraz8	Shiraz	USA	34	H-22	Hamedan22	Hamedan	Iran
4	SH-21	Shiraz21	Shiraz	USA	35	H-20	Hamedan20	Hamedan	Iran
5	SH-27	Shiraz27	Shiraz	Spain	36	H-6	Hamedan6	Hamedan	Iran
6	SH-P2	Poostnazok2	Shiraz	Iran	37	H-27	Hamedan27	Hamedan	Iran
7	SH-18	Shiraz18	Shiraz	Spain	38	H-5	Hamedan5-39.4	Hamedan	Iran
8	SH-17	Shiraz17	Shiraz	Spain	39	H-30-1	Hamedan30-40.2	Hamedan	Iran
9	SH-7	Shiraz7	Shiraz	Spain	40	H-30-2	Hamedan30-40.3	Hamedan	Iran
10	Mamaei-D	MamaeiDobahreh	Najafabad	Iran	41	H-7	Hamedan7	Hamedan	Iran
11	Mamaei	Mamaei1	Najafabad	Iran	42	H-4	Hamedan4	Hamedan	Iran
12	Tajerei-1	Tajeri12.3	Najafabad	Iran	43	H-18	Hamedan18	Hamedan	Iran
13	Tajerei-2	Tajeri12.5	Najafabad	Iran	44	Sahand	Sahand46.2	Azarbayejan	Iran
14	Dob-1	Dobahreh13.2	Najafabad	Iran	45	SP-A-1	Spain47.2	Spain	Spain
15	Dob-2	Dobahreh13.5	Najafabad	Iran	46	SP-A-2	Spain47.3	Spain	Spain
16	Monagha	Monagha	Najafabad	Iran	47	Shekoufeh	Shekofeh	Ai * Nonpareil	Hybrid
17	SH-Sangi	SangiMahalli	Shiraz	Iran	48	Nonpareil	NonPareil	USA	USA
18	N103	103	Najafabad	Iran	49	A-54	Arak54	Arak	Iran
19	N101-1	101-1	Najafabad	Iran	50	H-55	Hamedan55	Hamedan	Iran
20	N101-2	101-2-19.2	Najafabad	Iran	51	*A.lycioides*-1	*A. lycioides*1	Isfahan	Iran
21	N101-3	101-2-19.3	Najafabad	Iran	52	*A.lycioides*-2	*A. lycioides*2	Isfahan	Iran
22	Rabee	Rabee	Najafabad	Iran	53	BT	*P. dulcis* var. amara	Isfahan	Iran
23	Dob-21	Dobahreh21	Najafabad	Iran					
24	H-12	Hamedan12	Hamedan	Iran					
25	H-16	Hamedan16	Hamedan	Iran					
26	H-17	Hamedan17	Hamedan	Iran					
27	H-11	Hamedan11	Hamedan	Iran					
28	H-21	Hamedan21-28.1	Hamedan	Iran					
29	H-14-1	Hamedan14-29.1	Hamedan	Iran					
30	H-14-2	Hamedan14-29.3	Hamedan	Iran					
31	H-10	Hamedan10	Hamedan	Iran					

Molecular Data. Allele size was measured with UVDoc 99.02 analysis software (UVI Tech, Cambridge, UK) by manual editing to increase accuracy. This procedure carried out two times to exclude wrong scorings. We developed software ABRIISTAT30 for manipulating the large amount of data (Kadkhodaei, 2008c). We used the PowerMarker3.25 software package (Liu et al., 2005) to generate a similarity matrix based on Nei coefficient (Nei, 1983) and to produce a dendrogram using UPGMA method. Different genetic diversity parameters calculated as follow: Heterozygosity- (the proportion of heterozygous individuals in the population) at a single locus, Gene diversity- often referred to as expected heterozygosity (He), is defined as the probability that two randomly chosen alleles from the population are different, and Polymorphism information content (PIC) (Botstein, 1980).

Comparison of the molecular and morphological data. The mantel matrix correspondence test (Mantel, 1967) was used to compare the molecular and morphological similarity matrices. Genotype identification. A molecular identification key was developed using ABRIISTAT software (Kadkhodaei, 2008) for all of the genotype.

Table 2. Morphological traits used for almond characterization.

No.	Trait	Code	No.	Trait	Code
	Kernel			Tree	
1	Kernel Length	KL	21	Growth habit	GH
2	Kernel Width	KW		Flower	
3	Kernel Thickness	KT	22	Bearing Habit	BH
4	Kernel Weight	KWG	23	Flower Density	FDe
5	Kernel/Nut W %	KNWP	24	Double Flower in buds	DF
6	Kernel Color	KC	25	Flower Color	FC
7	Double Kernel	DK	26	Initial Flowering	F5
8	Shriveling of Kernel	SK	27	Completed Flowering	F100
9	Kernel Pubescence	KP	28	Duration of Flowering	FDu
	Leaf			Nut	
10	Petiole Length	PL	29	Nut Length	NL
11	Leaf Width	LW	30	Nut Width	NW
12	Leaf Length	LL	31	Nut Thickness	NT
13	Leaf Area	LA	32	Nut Weight	NWG
14	Leaf Basal Shape	LBS	33	Nut Shape	NS
15	Leaf Shape	LS	34	Nut Color	SC
16	Leaves Arrangement	LAr	35	Marking of outer Shell	MS
	Pests and Diseases		36	Suture Opening	SO
17	Bacterial Canker	SR	37	Shell Retention	SR
18	Black Aphid	ABR	38	Ease of Hulling	EH
19	Twig Borer	SKR	39	Softness of Shell	SS
20	Green Aphid	AR			

Results and discussion

Morphological traits

Cluster analysis. Genetic relationships among the almond genotypes have been shown in Figure 1. The almond genotypes clustered into two main groups: wild almonds and cultivated almonds. Mean value of groups for each trait presented in Table 4. The cultivated almonds group divided by 4 subgroups included two (II1), three (II2), forty four (II3) and two (II4) genotypes, respectively (Figure 1). The wild almonds group contained genotypes which had the minimum average for most of the studied traits, but average of this group for traits such as ease of hulling (EH), shell color (SC), marking of outer shell (MS), leaves arrangement (LAr), leaf basal shape (LBS), kernel color (KC) and sensitivity to Anarsia Lineatella (the peach twig borer) was higher than the cultivated group. Genotypes average in the first subgroup of cultivated almonds for some important traits such as kernel width (KW), thickness (KT) and weight (KWG), ease of hulling (EH), growth (tree) habit (GH), double flower in buds (DF), flower density (FD), bearing habit (BH), petiole length (PL), sensitivity to Anarsia Lineatella (peach twig borer), sensitivity to Pterochloroides persica (peach black aphid) and Myzus persicae (green peach aphid) was higher than the other subgroups. The second subgroup had the highest average for kernel length (KL), suture opening of the shell (SO), shriveling of kernel (SK), leaf length (LL), leaf width (LW), leaf area (LA) and leaf shape (LS), sensitivity to Pseudomonas syringae pv. syringae (Bacterial canker and blast of stone fruit trees) and duration of flowering (FDu). The average of third subgroup for the most of studied traits was in mid of the other subgroups but had the maximum value for nut length (NL), nut thickness (NT) and marking of outer shell (MS). The fourth subgroup showed the highest average for kernel width (KW), nut weight (NWG), Kernel/Nut W %, softness of shell (SS), kernel pubescence (KP), Initial flowering (F5) and completed flowering (F100). The first and the fourth subgroups had the highest average for the most important yield related traits, so these groups contain the best genotypes in this regard. The third subgroup divided by 5 minor groups. Minor group one had the highest average for kernel weight (KWG), suture opening (SO), shell retention (SR), ease of hulling (EH), leaf area (LA), leaf length (LL) and leaf arrangement (LAr), and the genotypes of second minor group showed the highest average for traits like nut width (NW) and thickness (NT), kernel thickness (KT), softness of shell (SS), shriveling of kernel (SK), Kernel Pubescence (KP), growth habit, flower density (FD), sensitivity to Anarsia Lineatella (peach twig borer) and Myzus persicae (green peach aphid). The highest average in third minor group referred to nut weight (NWG), bearing habit (BH), and sensitivity to Pterochloroides persica (peach black aphid) as well as the fourth minor group for nut length (NL), Kernel/Nut W %, ease of hulling (EH), sensitivity to Pseudomonas syringae pv. syringae (Bacterial canker and blast of stone fruit trees) and fifth minor group for percentage of double kernel (DK), double flower in buds (DF) and Initial flowering (F5), completed flowering (F100) and duration of flowering

Table 3. Almond SSR loci used in the present study and their characteristics.

SSR locus/ GeneBank no.	Sequence (5'-3')	Total No. of alleles	na[1]	Ne[2]	I[3]	Obs. Het	Exp.Het	Nei[4]	Ave.Het	fixation index (Fis)	PIC
UDA-002	5'-AAACGTGAGGTCTCACTCTCTC	85	17	6.83	2.26	0.33	0.86	0.85	0.85	0.6156	0.91
BV102479	5'-GCCATTTAAGGGTCTGGTCA										
UDA-005	5'-CATCACACACAAACACAAATGC	119	15	4.73	2.01	0.70	0.79	0.79	0.79	0.1123	0.93
BV102480	5'-GCATTGTGCTCTTCATGGAC										
UDA-008	5'-AGACGCTTTGCATACATACAAG	135	13	6.56	2.11	0.90	0.85	0.85	0.85	-0.0634	0.93
BV102481	5'-TGCAGGAACTGGGATTAGAGA										
UDA-009	5'-AAAACATCTCTCTCCTCCATGC	107	14	6.06	2.07	0.57	0.84	0.83	0.83	0.313	0.93
BV102482	5'-AGTTCTCTGGCAGCACAAGC										
UDA-015	5'-ACTCCATCGCTTGCATTTTC	81	9	2.48	1.35	0.25	0.60	0.60	0.60	0.5875	0.69
BV102474	5'-GCTCCGTGTGTGTTTGTGTG										
UDA-022	5'-GCCGTCTCATTTTCCCATTA	87	8	4.73	1.73	0.32	0.79	0.79	0.79	0.5964	0.88
BV102476	5'-GTGCGATGGAGGAGCACT										
UDA-023	5'-TTGCCGTGATACACTAACAACT	101	14	7.72	2.24	0.71	0.88	0.87	0.87	0.1822	0.95
BV102477	5'-ACCTGCCAAGTAAGTGCCTA										
Mean		102	12.86	5.59	1.97	0.54	0.80	0.80	0.80		0.89
St. Dev			3.24	1.75	0.32	0.25	0.09	0.09	0.09		

[1]na = Observed number of alleles; [2]ne = Effective number of alleles; [3]I = Shannon's Information index; [4]Nei's expected heterozygosity

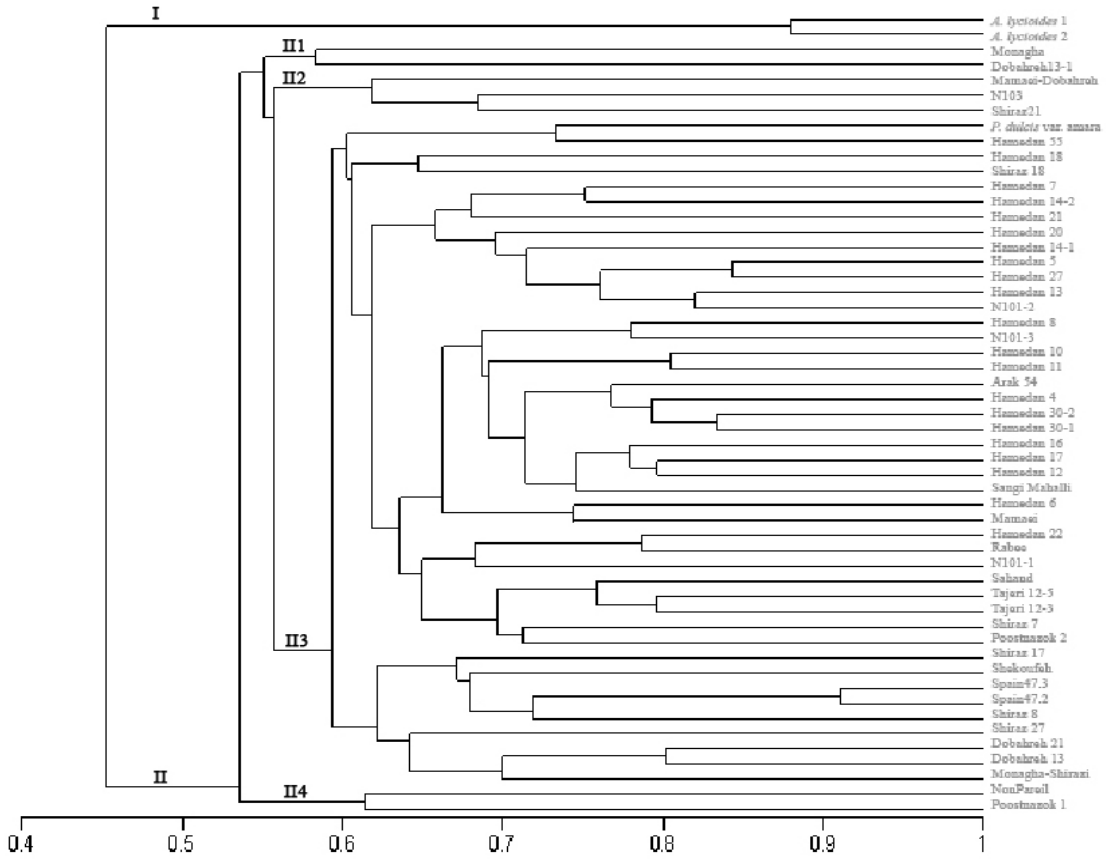

Fig 1. UPGMA dendrogram for morphological traits of 53 almond genotypes. Genetic distances were based on Gower General Similarity Coefficients.

Table 4. Cluster means for 39 characters estimated in 53 genotypes of almond.

TRAIT	CLUSTER						TRAIT	CLUSTER					
	1A	2A	1 B	2 B	3 B	4B		1A	2A	1 B	2 B	3 B	4B
Nut Length	1.09	2.88	1.78	2.75	2.87	3.48	Flower Density	5.00	7.50	9.00	5.66	7.73	4.00
Nut Width	0.91	1.86	1.65	1.76	1.88	1.86	Double Flower in buds	1.00	2.02	3.00	3.00	1.95	1.00
Nut Thickness	0.73	1.34	1.10	1.94	1.32	1.10	Flower Color	0.50	0.86	1.00	0.00	0.90	1.00
Nut Shape	1.00	2.62	4.00	2.00	2.50	5.00	Bearing Habit	1.00	2.11	3.00	2.00	2.09	2.00
Nut Color	4.02	7.00	3.00	3.00	4.00	6.00	Initial Flowering	.	5.96	2.00	4.30	6.09	9.50
Marking of outer Shell	9.00	5.19	5.00	3.00	5.00	7.00	Completed Flowering	.	15.96	13.50	16.00	15.81	20.00
Suture Opening	0.00	0.88	2.50	0.00	0.68	5.00	Duration of Flowering	.	9.94	11.50	11.66	9.72	10.50
Shell Retention	0.00	2.52	7.00	0.00	2.40	4.50	Bacterial Canker	0.50	0.25	0.00	0.33	0.27	0.00
Ease of Hulling	7.00	4.09	4.00	3.66	4.18	3.00	Twig Borer	0.00	0.68	1.00	0.33	0.35	1.00
Softness of Shell	5.00	4.64	7.00	3.60	4.40	9.00	Green Aphid	1.00	2.03	3.00	2.00	1.97	2.50
Kernel Length	0.98	1.91	1.74	2.53	1.88	2.18	Black Aphid	0.00	0.13	0.50	0.00	0.13	0.00
Kernel Width	0.66	1.13	1.41	1.11	1.12	1.12	Petiole Length	0.00	20.00	30.25	25.73	20.07	22.30
Kernel Thickness	0.50	0.62	0.98	0.75	0.60	0.69	Leaf Length	13.55	58.73	74.55	78.40	55.89	76.10
Kernel Weight	0.14	0.65	0.48	0.66	0.65	0.72	Leaf Width	3.35	20.43	24.80	28.03	19.54	24.05
Kernel/Nut W %	44.68	30.53	55.98	26.07	28.95	46.41	Leaf Area	26.20	817.00	1144.40	1224.00	760.30	1126.50
Kernel Color	7.00	6.13	7.00	5.00	6.18	6.00	Leaf Basal Shape	5.00	2.60	3.00	1.00	2.77	1.00
Kernel Pubescence	3.00	4.17	4.00	4.33	4.22	3.00	Leaf Shape	1.00	1.39	2.00	2.33	1.27	2.00
Double Kernel	0.00	0.14	0.00	0.08	1.52	0.28	Leaves Arrangement	1.00	0.86	0.50	0.00	0.90	0.50
Kernel Weight	0.31	2.48	0.86	2.62	2.57	1.88							
Shriveling of Kernel	4.00	4.98	4.00	5.00	5.02	5.00							
Growth habit	4.50	4.50	5.00	4.83	4.45	2.25							

Table 5. Eigenvectors, Eigenvalues, and proportions of variability for 8 principle components among 39 characters for 53 almond genotypes.

Character	Trait	Eigenvectors							
		Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
x1	Nut Length	0.0479	0.2725	0.2418	0.1670	0.1721	-0.1186	0.0809	-0.1863
x2	Nut Width	-0.0315	0.3718	0.1711	-0.0810	0.0274	-0.1072	-0.0746	-0.2253
x3	Nut Thickness	0.1047	0.1065	-0.1502	-0.2224	0.1674	-0.0795	0.0446	-0.3366
x12	Nut Shape	0.1107	-0.0759	0.2173	0.3675	0.0506	-0.0300	0.1551	-0.0060
x13	Nut Color	-0.1338	0.0846	0.0927	0.0292	0.4585	0.0920	0.0873	-0.0278
x14	Marking of outer Shell	-0.1221	-0.0363	0.1632	0.1388	0.1083	-0.1378	-0.3937	0.0710
x15	Suture Opening	-0.1187	-0.2686	0.2198	0.0237	0.2076	0.1781	-0.0501	-0.0717
x16	Shell Retention	-0.1440	-0.1323	0.0368	0.2154	0.0048	0.1333	0.2362	-0.2746
x17	Ease of Hulling	0.0167	0.0784	-0.1157	0.1767	-0.0566	-0.1598	-0.0579	0.0895
x18	Softness of Shell	0.1158	-0.2639	0.2073	0.1012	0.0612	0.0571	-0.2854	0.0710
x4	Kernel Length	0.1237	0.0754	0.2277	0.3447	-0.1055	0.0567	0.1134	-0.1532
x5	Kernel Width	0.1381	0.1667	0.2053	-0.0062	-0.2619	0.2647	-0.0702	-0.1931
x6	Kernel Thickness	0.1790	-0.1337	0.1570	0.0973	-0.2727	0.1250	0.0870	-0.0797
x7	Kernel Weight	0.1016	0.2243	0.2527	-0.1250	0.0219	0.0027	-0.1851	-0.1097
x9	Kernel/Nut W %	-0.0665	0.0801	-0.0375	0.2162	-0.1223	-0.0241	0.2367	0.1641
x10	Kernel Color	-0.1224	0.1801	-0.0089	-0.0680	-0.0732	0.2373	0.3634	0.0094
x20	Kernel Pubescence	0.0246	-0.2522	0.2649	-0.1155	-0.0367	0.1195	-0.1674	-0.0519
x11	Double Kernel	0.0513	0.1186	0.0704	-0.0591	0.0777	0.0743	0.1625	0.4106
x8	Kernel Weight	0.0476	0.4001	-0.0368	-0.0231	0.0418	-0.1374	-0.1302	-0.1122
x19	Shriveling of Kernel	-0.1808	0.1010	0.1719	-0.1128	-0.0323	0.1370	0.2516	0.1795
x22	Petiole Length	0.2813	0.1328	-0.0418	0.0489	0.0265	0.1610	-0.0275	0.1278
x24	Leaf Length	0.3728	0.0149	0.0785	0.0093	0.0658	0.0278	0.0915	0.0474
x23	Leaf Width	0.3386	-0.0004	-0.0129	-0.0366	0.0092	-0.1114	0.0000	0.1996
x21	Leaf Area	0.3883	0.0034	0.0628	0.0304	0.0248	-0.0877	0.0356	0.0761
x25	Leaf Basal Shape	-0.1752	0.1444	-0.0233	-0.1020	-0.0641	0.3878	-0.1230	-0.0368
x26	Leaf Shape	0.1697	-0.1075	-0.0008	-0.2450	0.2293	0.0693	0.0917	-0.1490
x27	Leaves Arrangement	-0.2420	0.0930	-0.1634	0.2473	-0.1225	-0.2158	-0.1295	0.1020
x28	Growth habit	-0.0442	-0.0729	0.0303	-0.1553	-0.3701	0.0375	-0.1209	-0.0665
x29	Flower Density	-0.0883	-0.0161	-0.1828	0.1481	-0.0069	0.2139	0.0697	0.0004
x30	Double Flower in buds	0.0422	-0.0136	0.0651	-0.0633	-0.3280	-0.2278	0.2919	-0.1241
x31	Flower Color	-0.1346	0.1102	0.1386	0.1119	-0.1878	0.0612	-0.2011	0.2153
x32	Bearing Habit	0.0100	-0.1539	-0.0286	0.2935	0.0366	-0.2784	0.0768	-0.2047
x37	Initial Flowering	-0.2110	-0.0962	0.2996	-0.1353	0.0019	-0.2088	0.1158	-0.0451
x38	Completed Flowering	-0.1447	-0.1289	0.3123	-0.2182	0.0104	-0.2434	0.1738	0.0963
x39	Duration of Flowering	0.1019	-0.0856	0.0795	-0.1978	0.0181	-0.1096	0.1421	0.2887
x33	Bacterial Canker	-0.0343	0.2138	0.1397	0.0403	0.1018	-0.1398	0.0043	0.2274
x35	Twig Borer	-0.0349	0.1588	0.2586	0.1373	0.0188	0.2115	-0.0367	0.1591
x36	Green Aphid	0.2156	-0.0387	-0.1560	0.0904	-0.0179	0.1581	-0.0300	0.0097
x34	Black Aphid	-0.0423	-0.0434	-0.0950	0.1684	0.3346	0.1590	0.0766	0.0383
Eigenvalue		5.1890	5.1877	4.3444	3.4213	2.9986	2.3687	2.2147	1.9375
Difference		0.8433	0.9231	0.4227	0.6298	0.1540	0.2773	0.1399	0.2329
Proportion		0.1330	0.1114	0.0877	0.0769	0.0607	0.0568	0.0497	0.0461
Cumulative		0.1330	0.1330	0.2444	0.3321	0.4090	0.4698	0.5265	0.5762

(FDu). Although considering geographical origin, the studied genotypes were from different regions (Hamadan, Shiraz, Najafabad and foreign countries), but the morphological dendrogram could not fit them correspondingly and cluster analysis only slightly reflected the geographical origin of populations. Similar finding were reported by Fikiru et al. (2010). Similarity. Similarity matrix for 53 almond genotypes was generated using Gower coefficient (data not shown). The maximum genetic similarity (0.91) was between two Spanish genotypes and then between two wild genotypes (0.88) which indicating the high similarity and close relationship between these genotypes. Hamadan genotypes (H-30-2 and H-30-1) and Tajeri genotypes grouped in the same sub clusters in the dendrogram with the similarities 0.833 and 0.796, respectively. Wild and cultivated almonds showed the least similarity (0.453) between each other. The mean morphological similarity was 0.59. These results confirm the report of Fathi et al. (2008) that indicated high level of variation in the studied almond cultivars.

Principal component analysis (PCA). PCA was performed to reduce the number of effective traits in group characterization. Eigenvectors, Eigenvalues and proportion of accounted variance for each variable has been shown in table 5. The first eight variables accounted for 62.23 percent of the total variability. The first component contains traits corresponding to leaf and sensitivity to Myzus persicae (green peach aphid); the second, nut width (NW), length (NL), kernel weight (KWG) and nut weight (NWG); and the third, kernel related traits, initial flowering (F5) and completed flowering (F100). This result is comparable with the report of Sorkhe et al. (2009) which in their report principal component analysis revealed that the nut weight and width, and the kernel weight had highest loading in the first component accounting for 45.8% of total variation. The fourth component stands for bearing habit (BH), nut thickness (NT) and shell retention (SR) in genotype characterization as well as in the fifth component which effect of suture opening (SO) and sensitivity to Pterochloroides persica (peach black aphid), in the sixth compon-

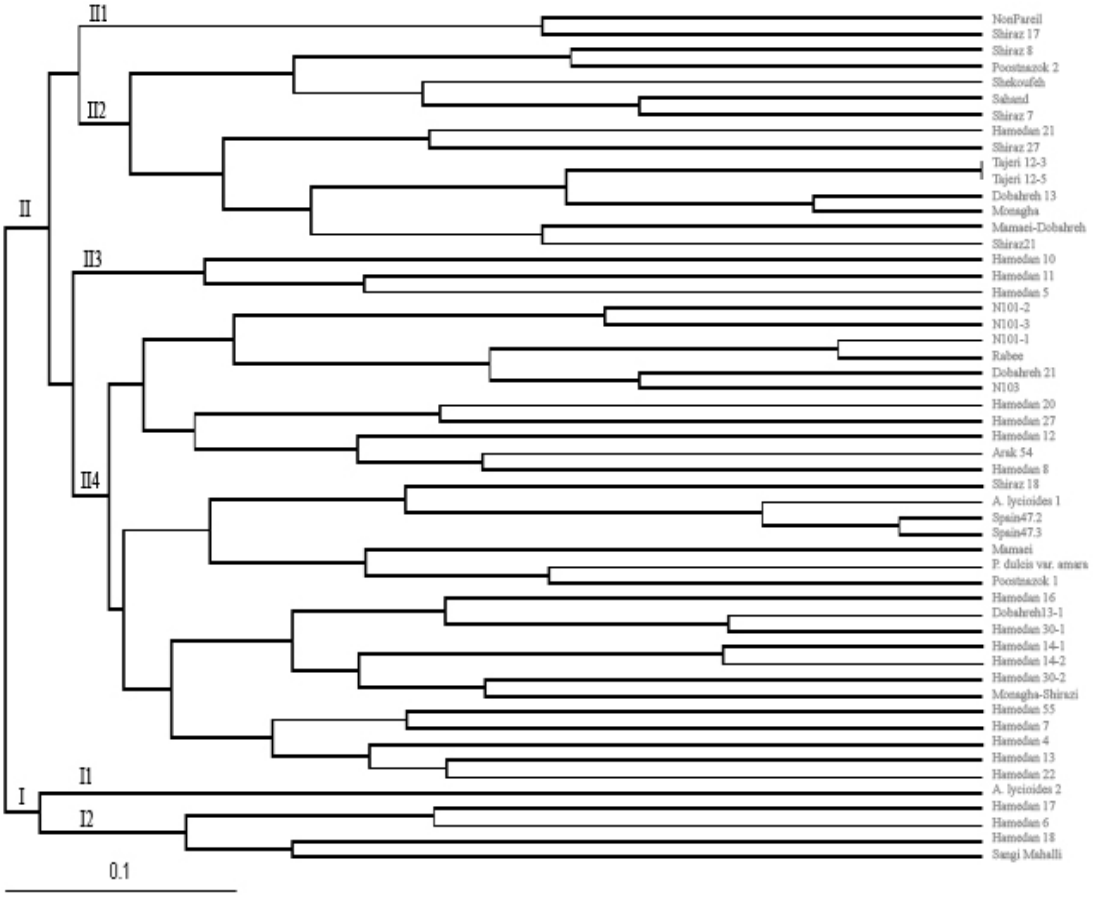

Fig 2. UPGMA dendrogram for molecular data relating 53 almond genotypes. Genetic distances were based on Nei Similarity Coefficients

ent flower density (FD), leaf basal shape and sensitivity to Anarsia Lineatella (peach twig borer) and in the seventh component shriveling of kernel (SK), shell retention (SR), kernel pubescence (KP), double flower in buds (DF) were highlighted, respectively. Combination of the traits double kernel% (DK), duration of flowering (FDu) and sensitivity to Pseudomonas syringae pv. syringae (Bacterial canker and blast of stone fruit trees) in eighth component showed an important role in characterization of the almond genotypes.

Molecular markers

SSR polymorphism. Out of nine microsatellite primer pairs, 7 were polymorphic, which revealed sufficient alleles to characterize all genotypes. The number of alleles per locus varied from 8 (UDA022) to 17 (UDA002) with an average of 12.86 (Table 3). The UDA-002 and UDA-005 markers were the most informative revealing 17 and 15 polymorphic alleles distributed among all genotypes, respectively, but considering PIC, UDA-023 was the most informative one (table 3). According to the results, the average values for number of effective alleles (Ne) (5.59), Shanon index (1.97), expected heterozygosity (He) (0.8) and PIC (0.89, ranging from 0.69 to 0.95) were relatively high, which can be due to the use of almond specific SSRs rather than previous similar studies using other related Prunus species SSRs, and high performance of the SSRs as molecular tools for genetic identification. The results support those of Fathi et al. (2008) and Kadkhodaei et al. (2006). Of course, self-incompatibility behavior of almond species is the main reason for the higher polymorphism in almond rather than other Prunus species. Increasing the amount of Ne (and not necessarily the total number of alleles per locus) leaded to an increase in He and also indirectly in ability of the loci for separation of genotypes through increasing the number of allelic genotypes (table 3). Genetic similarity. Similarity matrix of the genotypes was generated using Nei coefficient (data not shown). The most similarity values belonged to Tajeri-1 and -2 (which were tested as duplicates), Rabee and N101 (possible budsport of Rabee), SH-P1 and 2 (possible duplicates) and also Monagha and Dob-2 (2 local cultivars in Najafabad). Contrarily, Dob-2 and H-14-1, H-10 and Mamaei-D and also A.lycioides2 and N103 revealed the most genetic distance between each other. The average for similarity among the genotypes was as low as 0.23. A few genotypes including Tajeri-1 and -2 and also soft shell SH-P1 and SH-P2 assumed to be mislabeled, but the study indicated that this hypothesis was not true for them and they are duplicates. Genotypes N101 and N103 which were found accidentally in an orchard located in NajafAbad (Esfahan –central region of Iran), showed high genetic similarity with Rabee cultivar (a native cultivar in this region), indicating that may be these two genotypes are a Bud Sport of this cultivar. Interestingly, among the studied genotypes these two genotypes have the highest double flower in buds (DF) and consequently high flower density (FD) which are important in for yield increase and using in breeding programs. Cluster analysis. In the molecular dendrogram (Figure 2), the genotypes were divided into 2 major groups: group I included mainly 2 subgroups where a wild genotype A.lycioides and 4 genotypes comprised the first and second subgroups, respectively. Group II were divided into 4 subgroups which the genotypes with more Kernel/Nut W %,

suture opening (SO), softness of shell (SS) and initial flowering (F5) and completed flowering (F100) comprised the first subgroup. All of the genotypes (except duplicates) could be distinguished using the studied SSR loci, indicating their high polymorphism and performance in this regard. This result agrees with the report of Fathi et al. (2008) which they revealed the same results using peach SSRs. Also we were able to detect some index genotypes using their specific SSRs profiles, which can be lead to rapid and direct detection of them using a low cost and time consuming method (data not shown). Homozygosity. Regarding to homozygosity for the studied loci, A.lycioides2 was heterozygote for all loci except one. Contrarily, N101-2 represented the most homozygosity for all of the loci.

Specific markers. The specific alleles, rare alleles and allelic genotypes (alleles and genotypes with frequency lower than 1%) were recorded among the germplasm (data not shown). The specific markers can be used efficiently in breeding programs through marker assisted selection (MAS).

Comparison of the molecular and morphological data

The mantel matrix correspondence test (Mantel, 1967) was used to compare the molecular and morphological similarity matrices. Although overall correlations between the distance coefficients was rather low (r=0.02266, p=0.6926) but both morphological and molecular dendrograms clustered the genotypes into two main groups. In both dendrograms, Tajeri-1 and -2, Hamadan14, Dobahre13-2, Monagha and the Spanish genotypes grouped in the same cluster. Although Hamadan genotypes in morphological dendrogram clustered together but in different groups, they showed more distribution in the molecular analysis dendrogram. It can be concluded that however these genotypes are similar in their origin they have a considerable diversity. Molecular dendrogram distinguished Najafabad genotypes in two groups distinctively and revealed their same origin. Morphological dendrogram (and not molecular dendrogram) could discriminate the wild almonds from the cultivated ones properly. Genetic similarity between Dobahre-2 and Monagha, N101 and Rabee, Spanish genotypes, Dobahre-1 and Hamadan30-1, Nonpareil and Shiraz17, Dobahre21 and N103 was 0.0745, 0.0627, 0.1098, 0.1117, 0.1882, and 0.149, respectively in molecular analysis. Most of the Hamadan genotypes clustered in one subgroup according to the molecular dendrogram but they placed in two subgroups in morphological dendrogram.

In our study, some of the agronomical traits mentioned in Table 2 showed positive correlation with SSR loci (data not shown). There are several studies on different fruit trees such as plums (Shimada et al., 1999), apple (Landry et al., 1994), pistachio (Hormaza et al., 1998) and grapevine (Vidal et al., 1999) reported correlation between molecular data and morphological and agronomical traits. However, there is no need to be a necessarily positive correlation between morphological and molecular markers especially SSRs because of their genetic nature, as there are studies (Zhang et al., 2010) revealing no or low correlation in this regard. Nevertheless, the genetic relationship observed using molecular markers may provide information on the history and biology of cultivars, but it does not necessarily reflect what may be observed with respect to agronomic traits (Métais et al., 2000).

Genotype identification

Using the ABRIISTAT software, input data (allele sizes) were processed to produce data format for different analysis softwares such as PowerMarker3.25 (Liu, 2005), PopGene (Yeh, 1997), NTSys (Rolf, 2000), and also different statistical parameters related to genetic diversity.

Another feature of ABRIISTAT software was development of a molecular identification key for the studied loci. We suggested a coding method for genotype identification using microsatellite markers. The method includes two parts, the first for locus characteristics or Gene Bank No. and second for allele(s) size, for example UDA002-125/115.

Conclusion

Almond cultivation in Iran has a long historical background, and because of its self-incompatibility nature there are many genotypes growing in different regions of the country. These genotypes include a vast range of diversity in many characteristics such as blooming time. In this study, SSRs revealed a powerful tool to quantify genetic diversity in almond. The mantel matrix correspondence test was used to compare the molecular and morphological similarity matrices. Although the correlation coefficient was low but both morphological and molecular dendrogram clustered the genotypes into two main groups. The mean morphological similarity was high in comparison to similarity calculated using SSR markers. Since genetic diversity of initial selection materials is essential for successful breeding and creation of new cultivars, this information will be useful for collections, conservation and various almond breeding programs.

Acknowledgments

This research was supported by the Agricultural Biotechnology Research Institute of Iran (ABRII). Isfahan Research Center for Agriculture and Natural Resources (H. A. Daneshvar and A. Nasr) kindly provided samples and information about cultivars indicated in this study. The authors would like to thank Zeinab Matinzadeh for her assistance in this project.

References

Ballester J, Company RSI, Arús P, de Vicente MC (2001) Genetic mapping of a major gene delaying blooming time in almond. Plant Breed 120:268-270

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Amer J Human Genet 32:314-331

Browicz K, Zohary D (1996) The genus Amygdalus L. (Rosaceae): species relationships, distribution and evolution under domestication. Genet Resour Crop Evol 43:229–247

Channuntapipat C, Wirthensohn M, Ramesh SA, Batlle I, Arús P, Sedgley M, Collins G (2003) Identification of incompatibility genotypes in almond (*Prunus dulcis* Mill.) using specific primers based on the introns of the S-alleles. Plant Breed 122:164-168

Decroocq V, Favé MG, Hagen L, Bordenave L, Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. Theor Appl Genet 106:912-922

Dickson EE, Arumuganathan K, Kresovich S, Doyle JJ (1992) Nuclear DNA content variation within the rosaceae. Am J Bot 79:1081–1086

Dirlewanger E, Cosson P, Tavaud M, Aranzana MJ, Poizat C, Zanetto A, Arus P, Laigret F (2002) Development of microsatellite markers in peach (*Prunus persica* (L.) Batsch) and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theor Appl Genet 105:127–138

Fathi A, Ghareyazi B, Haghnazari A, Ghaffari MR, Pirseyedi M, Kadkhodaei S, Naghavi MR, Mardi M (2008) Assessment of the genetic diversity of almond (*Prunus dulcis*) using microsatellite markers and morphological traits. Ir J Biotech 6(2):98-106

Fikiru E, Tesfaye K, Bekele E (2010) A comparative study of morphological and molecular diversity in Ethiopian lentil (Lens culinaris Medikus) landraces. Afr J Plant Sci 4(7):242-254

Gower JC (1971) A general coefficient of similarity and some of its properties. Biometrics 27:857-874

Gregory D (2004) The development of a genetic linkage map for almond based on molecular and agronomic markers. PhD thesis. Adelaide: The University of Adelaide; 2004

Gupta, PK, Balyan, HS, Sharma, PC, Ramesh, B (1996) Microsatellites in plants: a new class of molecular markers. Curr Sci 70:45-54

Hormaza JI, Pinney K, Polito VS (1998) Genetic diversity of pistachio (*Pistacia vera*, Anacardiaceae) germplasm based on randomly amplified polymorphic DNA (RAPD) markers. Econ Bot 52:78–87

Joobeur T, Periam N, Vicente MC, King GJ, Arús P (2000) Development of a second generation linkage map for almond using RAPD and SSR markers. Genome 43:649-655

Kadkhodaei S, Aghdaei SRT, Grigorian V, Moghadam, Hashemi SMM (2006) A study on genetic variation among some wild almond species using rapd markers. Acta Hort (ISHS) 726:93-98

Kadkhodaei S, Arabha A, Ebrahimi M (2008) Karyotype analysis of some Iranian wild almonds. 10th Iranian Genetics Congress

Kadkhodaei S, Elahy M, Khayyam Nekouei M, Imani A, Mardi M, Javanmard A, Ariff AB (2010) A panel of cultivate specific marker based on polymorphisms at microsatellite markers for Iranian cultivated Almonds (*Prunus dulcis*). Aust J. Crop Sci. 4 (9) 730-736

Kadkhodaei S, Khayyam M, Elahy M (2008c) ABRIISTAT 30: molecular data analysis software for genetic diversity studies. 18th International EUCARPIA General Congress Valencia, Spain

Kadkhodaei S, Tabaei-Aghdaei R (2005) A simple and inexpensive method for nucleic acid isolation from plants with high polysaccharides and polyphenolic compounds: development of DNA extraction procedure in case of Amygdalus spp. IV International Symposium on Pistachios and Almonds. Tehran, Iran

Khatamsaz M (1992) Rosaceae. In: Flora of Iran. Research Institute of Forests and Rangelands, pp1–352

Landry BS, Li RQ, Cheung WY, Granger RL (1994) Phylogeny analysis of 25 apple rootstocks using RAPD markers and tactical gene tagging. Theor Appl Genet 89:847–852

Liu K, Muse S (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics, 21(9):2128-2129

Maghuly, F, da Câmara Machado A, Ruthner Sz, Pedryc A, Bisztray G, Katinger H, Laimer M (2004) Biotechnology of temperate fruit trees and grapevines. IV CHB Symposium. Book of Abstracts, 204

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209-220

Martínez-Gómez P, Sánchez-Pérez R, Rubio M, Dicenta F, Gradziel TM, Sozzi GO (2005) Application of Recent Biotechnologies to Prunus Tree Crop Genetic Improvement. Cien Inv Agr 32(2):73-96

Messina R, Lain O, Marrazzo MT, Huang WG, Cipriani G, Testolin R (2004) Isolation of microsatellites from almond and apricot genomic libraries and testing for their transportability across different Prunus species. XIth Eucarpia Symp on Fruit Breed & Genetics. Acta Hort 663

Métais I, Aubry A, Hamon B, Jaluozot R (2000) Description and analysis of genetic diversity between commercial bean lines (*Phaseolus vulgaris* L.). Theor Appl Genet 101:1207-1214

Mnejja M, Garcia-Mas J, Howad W, Arús P (2005) Development and transportability across Prunus species of 42 polymorphic almond microsatellites. Mol Ecol Notes 5(3):531-535

Nei M, Takezaki N (1983) Estimation of genetic distances and phylogenetic trees from DNA anlysis. Proc 5th World Cong Genet Appl Livstock Prod 21:405-412

Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci 1:215-22

Rohlf FJ (2000) NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System. Version 2.1. Exceter Sofware, New York, USA

Sánchez-Pérez R, Howad W, Dicenta F, Arús P, Martínez-Gómez P (2007) Mapping major genes and quantitative trait loci controlling agronomic traits in almond. Plant Breed 126:310-318

Sánchez-Pérez R, Howad W, Garcia-Mas J, Arús P, Martínez-Gómez P, Dicenta F (2010) Molecular markers for kernel bitterness in almond. Tree Genet Genomes 6:237-245

Schueler S, Tusch A, Schuster M, Ziegenhagen B (2003) Characterization of microsatellites in wild and sweet cherry – markers for individual identification and reproductive processes. Genome 46:95-102

Shimada T, Hayama H, Haji T, Yamaguchi M, Yoshida M (1999) Genetic diversity of plums characterized by random amplified polymorphic DNA (RAPD) analysis. Euphytica 109:143–147

Shiran B, Amirbakhtiar N, Kiani S, Mohammadi SH, Sayed-Tabatabaei BT, Moradi H (2007) Molecular characterization and genetic relationship among almond cultivars assessed by RAPD and SSR markers. Sci Hortic 111:280-292

Socias I Company R (1998) Almond genetics at a turning point: the almond example. Theor Appl Genet 96, 588–601

Sorkheh K, Shiran B, Rouhi V, Asadi E, Jahanbazi H, Moradi H, Gradziel TM, Martínez-Gómez P, (2009) Phenotypic diversity within native Iranian almond (*Prunus* spp.) species and their breeding potential. Genet Resour Crop Ev 56(7):947-961.

Testolin R, Messina R, Lain O, Marrazzo MT, Huang W-G, Cipriani G (2004) Microsatellites isolated in almond from an AC-repeat enriched library. Mol Ecol Notes 4:459–461

Vidal JR, Coarer M, Defontain (1999) A Genetic relationships among grapevine varieties grown in different French and Spanish regions based on RAPD markers. Euphytica 109:161–172

Xu Y, Ma RC, Xie H, Liu JT, Cao MQ (2004) Development of SSR markers for the phylogenetic analysis of almond trees from China and the Mediterranean region. Genome 47:1091–1104

Yeh FC, Boyle TJB (1997) Popgene Version 1.32. Microsoft Windows-based software for population genetic analysis. pp 135

Zhang X, Zhang Y, Yan R, Han G, Hong F, Wang J, Cao K (2010) Genetic variation of white clover (Trifolium repens L.) collections from China detected by morphological traits, RAPD and SSR. Afr J Biotechnol 9(21):3032-3041