International Journal of Horticultural Science and Technology
Vol. 6, No. 1; June 2019, pp 137-150
Print ISSN: 2322-1461 Online ISSN: 2588-3143
DOI: 10.22059/ijhst.2019.276425.282
Web Page: https:// ijhst.ut.ac.ir, Email: ijhst@ut.ac.ir

A Comparative Study of Genetic Diversity, Heritability and Inter-relationships of Tree and Nut Attributes between *Prunus scoparia* and *P. elaeagnifolia* using Multivariate Statistical Analysis

Akram Ansari[1], Ali Gharaghani[1, 2*]

1. Department of Horticultural Science, College of Agriculture, Shiraz University, Shiraz, Iran
2. Drought Research center, College of Agriculture, Shiraz University, Shiraz, Iran

(Received: 20 February 2019, Accepted: 26 May 2019)

Abstract
By applying multivariate statistical analysis, this research aimed to estimate the heritability and find relationships between the vegetative and reproductive characteristics of *Prunus scoparia* and *Prunus elaeagnifolia*. Twenty genotypes of each species were selected randomly from cultivated populations and twenty-two traits including the tree, leaf, flower, nut and kernel attributes were measured. Results showed that there were high levels of genotypic and phenotypic variations among the genotypes of both species. Many of the measurements pertaining to the leaf, flower, nut and kernel, showed very high heritability (H^2 >90%) in both species, whilst some traits such as shoot diameter in *P. scoparia* and kernel moisture in both species had very lower heritability (H^2 <50%). Generally, the heritability of measured traits in *P. elaeagnifolia* were higher than those of *P. scoparia*, especially for economically important traits including yield (H^2 = 94 and H^2 = 54.61, respectively in *P. elaeagnifolia* and *P. scoparia*), nut weight (H^2 = 97.83 and H^2 = 85.39.61, respectively in *P. elaeagnifolia* and *P. scoparia*) and oil percentage (H^2 = 75.55 and H^2 = 61.43, respectively in *P. elaeagnifolia* and *P. scoparia*). Stepwise regression analysis revealed that the most influential factors on yield of *P. scoparia*, were the fruit set, flower diameter and leaf length, whilst for the *P. elaeagnifolia*, the yield was mostly determined by fruit set and leaf area. The high genetic diversity and heritability of the studied traits, indicates high genetic potential of this germplasm to be utilized in future breeding programs.

Keywords: Wild almond, Breeding, Stepwise regression, Cluster analysis, Bi-plot.

Introduction

Almond, as a nut crop, is among the most profitable crops in the world. It can thrive in semi-arid areas and calcareous soils, and the nut has high nutritional value (Kiani et al., 2015). Almond breeding programs face many challenges due to the narrow genetic background of commercial cultivars. Therefore, in many almond breeding programs across the world, wild almonds considered as valuable genetic resources which can be used to broaden the genetic background and to introduce new traits into commercial relatives (Gharaghani et al., 2017). Late bloom, early maturity, adaption to drought and salinity, resistance to low temperatures in winter, reduced insect infestation and fungal attacks and having nectary flowers for honeybees are among the useful traits which can be

* Corresponding author, Email; agharghani@shirazu.ac.ir

considered in this regard (Sorkheh et al., 2009). Therefore, knowledge about genetic diversity of wild genetic resources of almond is an essential prerequisite for involvement of native germplasm in almond breeding programs. On the other hand, assessment of genetic diversity is necessary to evaluate the existed levels of genetic variability, which is considered as a guarantee for conservation management of natural populations (Cohen et al., 1991).

Morphological evaluations are a common method for the assessment of genetic diversity with respect to different traits. Even though new tools such as molecular markers are deemed useful, they are still costly (Atanasov et al., 2015). Advanced statistical methods and multivariate techniques, including cluster analysis, principle components analysis or regressions could be described as efficient tools for the evaluation of genotypes, cultivars, or even the screening of populations (Anumalla et al., 2015). Considering structural relationships among different traits of plants, however, advanced statistical methods can provide reliable procedures for the complex process of selection in breeding programs (Anumalla et al., 2015).

Nearly 20 wild almond species have been reported in Iran, indicating that Iran is a center of diversity and even a center of origin for almond (Gharaghani et al., 2017). These valuable genetic assets can be used productively in almond breeding programs, however a precise identification of genetic relationships is required before the incorporation of wild almonds into breeding programs (Pinar et al., 2016).

Prunus scoparia Spach. and *Prunus elaeagnifolia* Spach. are two important wild almond species in Iran, where they have been widely scattered across the country since ancient times. Albeit both species have been used as rootstocks for almond cultivars, *P. elaeagnifolia* has been further used as a rootstock for plum cultivars in Iran, especially where the availability of water is scarce (Gharaghani et al. 2017). *P. scoparia* is a multi-purpose species in Iran. It has the potential of becoming a crop of choice in arid and semi-arid areas and can adapt to adverse climatic conditions. The kernel and gum of this species could be used in industries that produce pharmaceuticals, chemicals and food. Such opportunities make the species not only a good choice for reforestation but also a horticultural candidate for multipurpose nut production (Gharaghani and Eshghi, 2015).

Although limited number of scientific reports pertaining to the genetic diversity of wild almond species is available in Iran (Zeinalabedini et al., 2016, Sorkhe et al., 2009), there is a substantial demand for a deeper knowledge of genotypic and phenotypic diversity of wildly scattered almond species in Iran, especially those on the southern region of Zagros mountains (Rahimi Dvin et al., 2017). Accordingly, this comparative study sought to explore the available genotypic and phonotypic diversity of two wild almond species, i.e. *P. scoparia* Spach. and *P. elaeagnifolia* Spach. In this regard the heritability of various plant and nut traits was measured and their structural relationships were also assessed. A linear relationship was also determined between the yield and other measured traits using stepwise regression analysis.

Materials and methods

Plant materials

This research was carried out by using cultivated populations of two wild almond species, i.e. *P. scoparia* and *P. elaeagnifolia* (Fig. 1) located in the campus at the School of Agriculture, Shiraz University, Badjgah, Shiraz, Fars province, Iran. The geographic coordinates are 38° 29' north latitude and 35° 52' eastern longitude, with an altitude of 1810m. These populations are in the same age and were established more than 30 years ago for research and conservation purposes. A

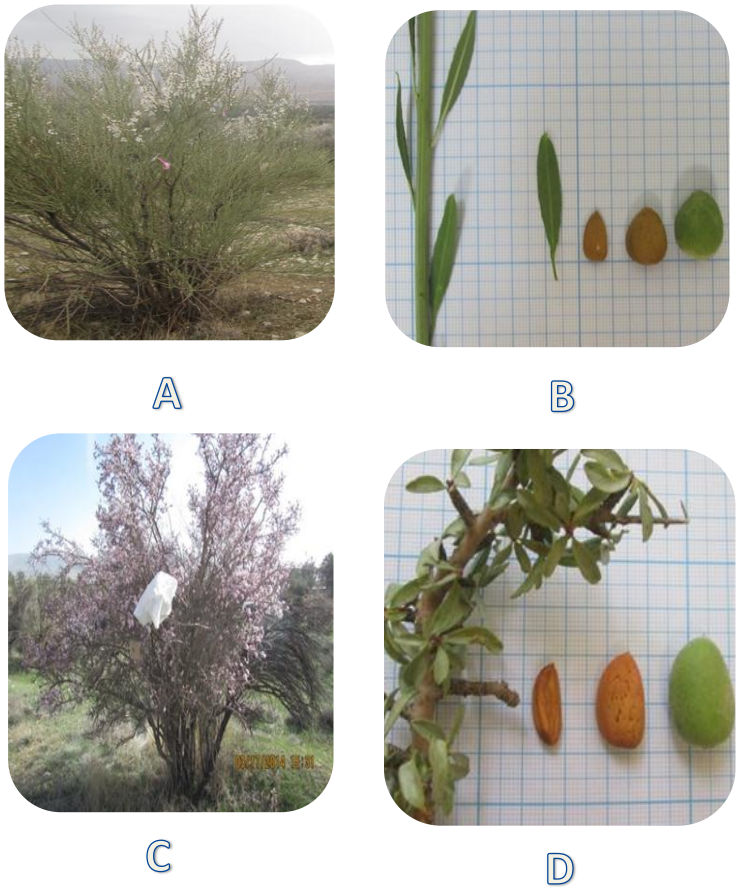

Fig. 1. Shrub, shoot, leaf, kernel, nut and fruit of *P. scoparia* (A and B) and *P. elaeagnifolia* (C and D).

wild collection of these species had been brought from Dasht-e Muk area of Firuz Abad county (in the Fars province) which is considered as one of the hotspots for the diversity and evolution of these two wild almond species in Iran. Twenty genotypes of each species were randomly selected for this study.

Measurements

Six vegetative traits (including shoot length and diameter, leaf length, width, area and length/width ratio) and 17 reproductive traits (including flowering date, flower diameter, fruit set, nut weight, length, width and diameter, nut ripening date, kernel weight, length, width and diameter, kernel moisture, oil and protein content) were measured in the growing season of 2014. Almond descriptors were used for characterizing the leaf, flower, nut and kernel attributes. The descriptors were developed by the International Plant Genetic Resources Institute (IPGRI) (Gulcan, 1985). Measurements were made on three sub-samples (i.e. three branches of each tree were treated as three replications for each genotype in each species). Regarding the measurement of shoots, leaves, flowers, nuts and kernels each sub-sample was comprised of 20 shoots, leaves, flowers, nuts and kernels per tree. Variables were measured by using proper tape measure, digital caliper and electronic balance (0.001 g precision). Phenological traits including flowering date (50% open bloom) and fruit ripening date (50% husk split) were recorded as date. Fruit set was calculated on three branch per tree by dividing the number of fruit to number of previously counted flowers on the same branches*100. The protein content was determined by using a method described by Ahmed and Schmidt (1979), and the total oil content was measured by using a Soxhlet

extractor with n-hexane as solvent, according to descriptions by Venkatachalam and Sathe (2006).

Statistical analyses

The experiment was conducted based on a completely randomized design (CRD) with twenty genotypes per species and three replications per genotype. Analysis of variance (ANOVA) and mean comparison of values for genotypes were performed in each species separately, according to the GLM procedure of SAS 9.3 software. Descriptive statistics such as means and standard deviations in addition to analysis of variance (ANOVA) and estimated genetic parameters were calculated by the software SAS v.9.3 (SAS Institute, 2003). The genotypic coefficient of variation (GCV), phenotypic coefficient of variation (PCV), broad sense heritability (H^2) and response to selection (RS) were calculated according to the following formulas (Falconer & Mackay, 1996);

$$GCV = (\sqrt{VG}/\mu x) \times 100$$

$$PCV = (\sqrt{VP}/\mu x) \times 100$$

$$H^2 = (VG/VP) \times 100$$

$$RS = 2.06 \times \sqrt{VP} \times (H/100)$$

where VP, VG and μx represent the phenotypic variance, genotypic variance and grand mean.

To illustrate the similarity and/or dissimilarity among the genotypes in the two studied species, a cluster analysis of the genotypes was made based on all measured traits. Squared Euclidian distance and the Ward's method for grouping were applied by using the R 3.5.0 statistical software (core package of the software) (R Development Core Team, 2008). Principal component analysis and regression analysis were performed by using SPSS 25 software (SPSS Inc., Chicago, IL, USA).

Results

Genetic diversity and heritability estimates

Considering the results, a high amount of variability was clearly observed in most of the measured traits in both *P. scoparia* and *P. elaeagnifolia* species. The genotypic and phenotypic parameters of *P. scoparia* and *P. elaeagnifolia* were estimated (Tables 1 and 2). According to the results of genetic estimation, all measured traits showed different degrees of genotypic and phenotypic variations. The highest genotypic coefficient of variation (GCV) (61.95%) and phenotypic coefficient of variation (PCV) (78.75%) were estimated in relation to the leaf area and the moisture content of kernels in *P. scoparia*, respectively, while the highest GCV and PCV in *P. elaeagnifolia* were obtained in the shoot length (52.23%) and the moisture content of kernels (80.32%), respectively (Table 1 and 2).

In this research, many of the measured traits showed very high levels of heritability (H^2 > 90%) in both species. These measured parameters included the leaf, flower, nut and kernel attributes, as well as yield (Table 1 and 2). On the other hand, some traits showed lower levels of heritability. These were the shoot diameter (H^2 = 57.7 and H^2 = 44.24 in *P. scoparia* and *P. elaeagnifolia*, respectively) and the moisture content of kernels (H^2 = 38.58 and H^2 = 29.88 in *P. scoparia* and *P. elaeagnifolia*, respectively), (Table 1 and 2). Nonetheless, their GCV and PCV variations showed high values in relation to the moisture content. High levels of heritability were recorded for kernel oil (H^2 = 75.55 and H^2 = 61.43 in *P. scoparia* and *P. elaeagnifolia*, respectively) and for protein (H^2 = 89.54 and H^2 = 90.00 in *P. scoparia* and *P. elaeagnifolia*, respectively), (Table 1 and 2).

Table 1. Genetic parameters calculated in *P. scoparia* using the expected values of the ANOVA table

Parameter	Mn	Mx	M	SE	GV	EV	PV	H^2	PCV	GCV	RS	NM
Shoot length (cm)	14	43	26.7	1.41	32.47	8.53	41	79.2	23.98	21.34	10.45	37.15
Shoot diameter (mm)	2.2	4.32	3.09	0.13	0.19	0.14	0.33	57.47	18.61	14.11	0.68	3.77
Leaf length (cm)	1.9	5.08	3.19	0.18	0.65	0.02	0.67	97.04	25.75	25.36	1.64	4.83
Leaf width (cm)	0.2	0.85	0.49	0.03	0.02	0	0.02	95.82	31.2	30.55	0.3	0.79
Leaf area (cm^2)	0.16	2.25	0.98	0.14	0.37	0.02	0.39	94.74	63.64	61.95	1.22	2.2
Leaf length/Leaf width	4.33	9.5	6.76	0.24	1.08	0.14	1.22	88.86	16.31	15.38	2.02	8.78
Flowering date	12th March	2th April	-	-	-	-	-	-	-	-	-	-
Flower diameter (cm)	1.87	3.19	2.45	0.08	0.13	0.01	0.14	94.57	15.3	14.88	0.73	3.18
Fruit set (%)	1.45	25.13	10.15	1.12	18.35	7.23	25.58	71.74	49.86	42.23	7.47	17.62
Nut Length (cm)	0.82	1.5	1.19	0.04	0.03	<0.01	0.04	93.23	16	15.45	0.37	1.55
Nut Width (cm)	0.5	0.83	0.67	0.02	0.01	<0.01	0.01	97.96	13.59	13.45	0.18	0.85
Nut weight (g)	0.08	0.25	0.17	0.01	<0.01	<0.01	<0.01	97.83	30.1	29.77	0.1	0.27
Nut diameter (cm)	0.37	0.58	0.44	0.01	<0.01	<0.01	<0.01	84.39	8.26	7.59	0.06	0.51
Nut ripening date	16th June	4th July	-	-	-	-	-	-	-	-	-	-
kernel length (cm)	1.04	1.82	1.51	0.06	0.06	<0.01	0.06	98.68	16.83	16.72	0.52	2.02
Kernel width (cm)	0.75	1.3	1	0.03	0.02	<0.01	0.02	99.04	14.97	14.9	0.3	1.3
Kernel weight (g)	0.28	1.07	0.6	0.05	0.04	<0.01	0.04	98.67	34.61	34.38	0.43	1.03
Kernel diameter (cm)	0.62	0.89	0.75	0.02	0.01	<0.01	0.01	96.92	10.77	10.6	0.16	0.91
Moisture (%)	0	17.65	4.49	0.78	4.83	7.69	12.53	38.58	78.75	48.92	2.81	7.31
Protein (%)	13.59	26.52	19.41	0.74	10.14	1.18	11.32	89.54	17.34	16.41	6.21	25.61
Oil (%)	20.88	39.63	32.93	0.94	13.79	4.46	18.25	75.55	12.97	11.28	6.65	39.58
Yield (kg)	0.1	2.22	1.52	0.15	0.11	0.05	0.115	94	7.24	7.57	0.67	2.01

Mn: minimum, Mx: maximum, SE: standard error, M: mean, GV: genotypic variation, EV: environmental variation, PV: Phenotypic variation, H^2, broad sense heritability, PCV: phenotypic coefficient of variation, GCV: genotypic coefficient of variation, RS: response to selection (under 5% screening from the population), NM: mean of the next generation after the selection of 5% of the population.

Table 2. Genetic parameters calculated in *P. elaeagnifolia* using the expected values of the ANOVA table

Parameter	Mn	Mx	M	SE	GV	EV	PV	h^2	PCV	GCV	RS	NM
Shoot length (cm)	0.58	10	4.84	0.58	6.38	0.64	7.02	90.93	54.77	52.23	4.96	9.8
Shoot diameter (mm)	1.1	4.31	2.39	0.13	0.15	0.19	0.34	44.24	24.44	16.25	0.53	2.92
Leaf length (cm)	1.01	2.13	1.53	0.06	0.07	0	0.07	94.47	17.66	17.17	0.53	2.05
Leaf width (cm)	0.38	0.73	0.55	0.02	0.01	0	0.01	94.83	17.53	17.08	0.19	0.73
Leaf area (cm^2)	0.26	1.24	0.67	0.05	0.04	0.01	0.05	89.46	33.63	31.81	0.41	1.08
Leaf length/Leaf width	2.19	3.43	2.83	0.06	0.07	0.01	0.09	86.13	10.41	9.66	0.52	3.35
Flowering date	7th March	19th March	-	-	-	-	-	-	-	-	-	-
Flower diameter (cm)	1.32	2.32	1.7	0.05	0.05	0.01	0.06	87.45	14.27	13.35	0.44	2.13
Fruit set (%)	0.94	15.98	7.46	0.78	11.72	0.88	12.59	93.03	47.54	45.86	6.8	14.27
Nut Length (cm)	1.06	1.57	1.26	0.03	<0.01	<0.01	<0.01	97.35	10.94	10.8	0.28	1.54
Nut Width (cm)	0.55	0.82	0.68	0.01	<0.01	<0.01	<0.01	96.87	8.39	8.25	0.11	0.79
Nut weight (g)	0.12	0.27	0.17	0.01	<0.01	<0.01	<0.01	85.39	18.18	16.8	0.06	0.23
Nut diameter (cm)	0.34	0.52	0.41	0.01	<0.01	<0.01	<0.01	84.88	10.81	9.96	0.08	0.49
Nut ripening date	4th June	27th June	-	-	-	-	-	-	-	-	-	-
kernel length (cm)	1.35	2.03	1.63	0.04	<0.01	<0.01	<0.01	97.82	11.82	11.69	0.39	2.02
Kernel width (cm)	0.94	1.33	1.08	0.02	<0.01	<0.01	<0.01	97.53	8.38	8.27	0.18	1.27
Kernel weight (g)	0.45	1.1	0.69	0.03	<0.01	<0.01	<0.01	97.47	21.86	21.58	0.3	0.99
Kernel diameter (cm)	0.67	0.95	0.77	0.01	<0.01	<0.01	<0.01	86.86	7.83	7.3	0.11	0.88
Moisture (%)	0	14.29	3.9	0.69	2.94	6.89	9.82	29.88	80.32	43.91	1.93	5.83
Protein (%)	12.59	23.46	16.49	0.66	8.3	0.82	9.12	91	18.31	17.47	5.66	22.15
Oil (%)	20.5	48.13	36.53	1.24	19.56	12.29	31.85	61.43	15.45	12.11	7.14	43.67
Yield (kg)	0.13	1.28	0.234	0.07	0.06	0.02	0.11	54.61	25.64	47.01	0.47	0.345

Mn: minimum, Mx: maximum, SE: standard error, M: mean, GV: genotypic variation, EV: environmental variation, PV: Phenotypic variation, h^2, heritability, PCV: phenotypic coefficient of variation, GCV: genotypic coefficient of variation, RS: response to selection (under 5% screening from the population), NM: mean of the next generation after the selection of 5% of the population.

Cluster analysis

The ultimate clustering diagrams of *P. scoparia* and *P. elaeagnifolia* were portrayed based on all of the measured traits (Figures 3 and 4). The results showed that *P. scoparia* genotypes can be clustered into three distinct groups. The most closely related pairs of genotypes among all those of *P. scoparia* were the genotypes 1 and 19 in the third cluster. In this study, on one side, there were genotypes 10 and 16, and on the other side, genotypes 6, 17, 3, and/or 14, which can be recommended for parenting future crosses that could make new generations with high variations in almost all of the measured traits.

Clustering the genotypes of *P. elaeagnifolia* resulted in four separated groups of genotypes (Figure 5). Highly similar pairs in the second, third and fourth clusters were genotypes 9 and 16, genotypes 8 and 13, and genotypes 12 and 20, respectively. Highly distant pairs included genotypes 9 and 16 vs. genotypes 8 and 13, which can be used as dissimilar and distanced parents when aiming at producing highly variable genotypes in new generations.

Principle component analysis

The first two components (PCs) explained 64% and 51% of the total variability among data pertaining to *P. scoparia* and *P. elaeagnifolia*, respectively (Table 3). Figure 4 and Figure 5 are showing the genotypes points of the two species of *A. scoparia* and *A. eleagnifolia* versus the first two components of the principal component analysis in two-dimensional plot. According to the PC results, the genotypes of both species were grouped well, and the two graphs indicated separated groups as previously determined by the cluster analysis in both species (Fig. 2 and 3).

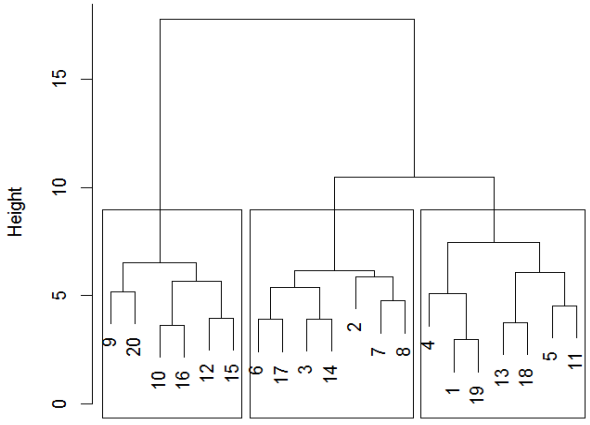

Fig. 2. Cluster dendrogram for 20 genotypes of *P. scoparia* based on all measured traits.

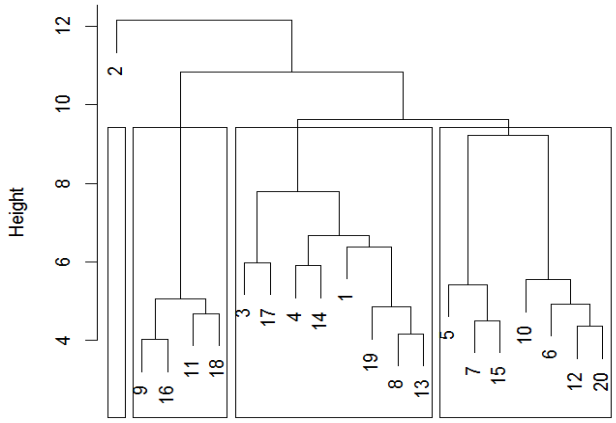

Fig. 3. Cluster dendrogram for 20 genotypes of *P. eleagnifolia* based on all measured traits.

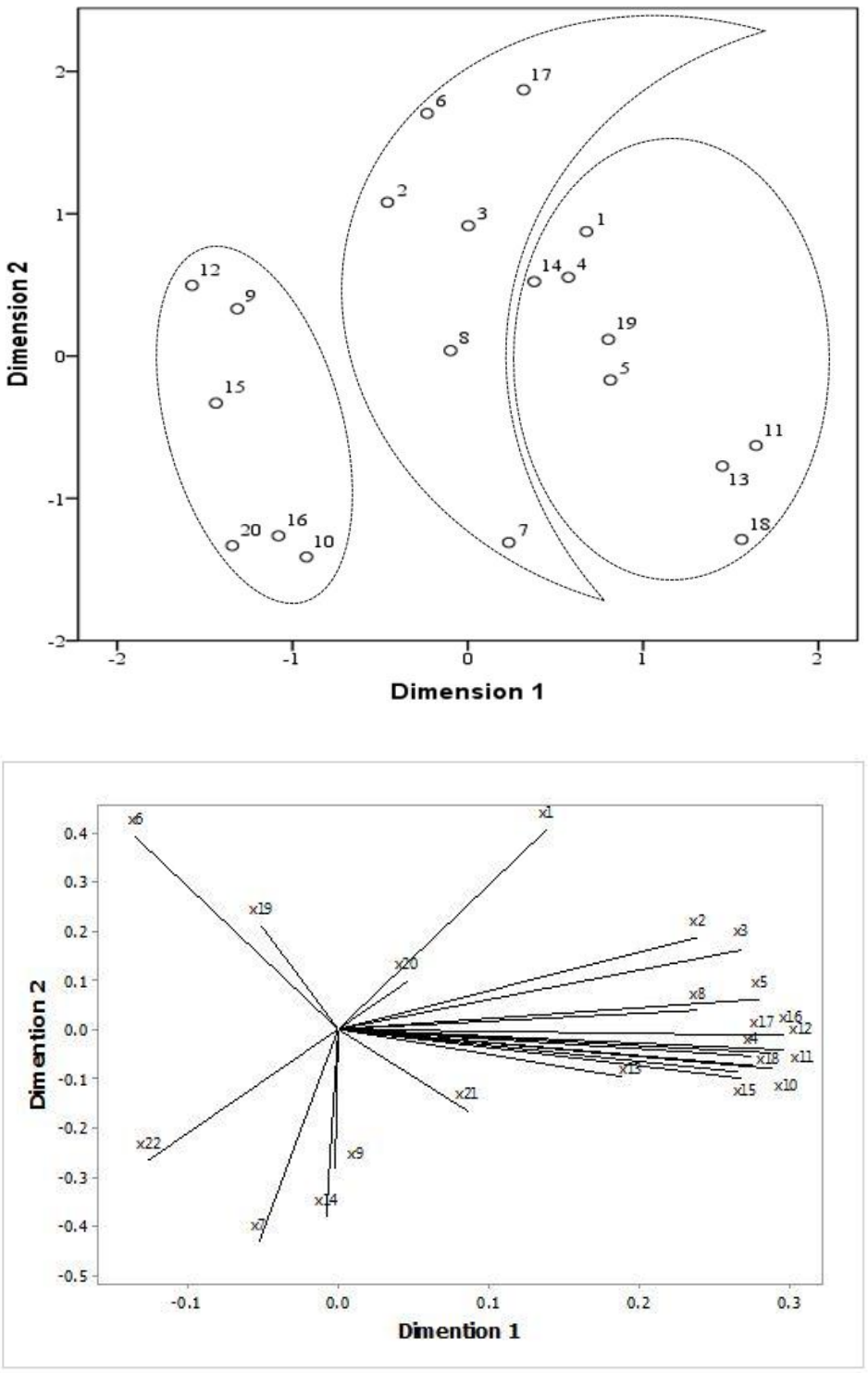

Fig. 4. Score (upper) and loading (lower) plots according to principal component analysis in *P. scoparia*. X1: Shoot length, X2: Shoot diameter, X3: Leaf length, X4: Leaf width, X5: Leaf area, X6: Leaf length/Leaf width, X7: Flowering date, X8: Flower diameter, X9: Fruit set, X10: Nut Length, X11: Nut width, X12:nut weight, X13:nut diameter, X14:Nut ripening date, X15: Kernel length, X16: Kernel width, X17: Kernel weight, X18: Kernel diameter, X19: Nut moisture, X20: Nut protein, X21: Nut oil, X22: Yield.

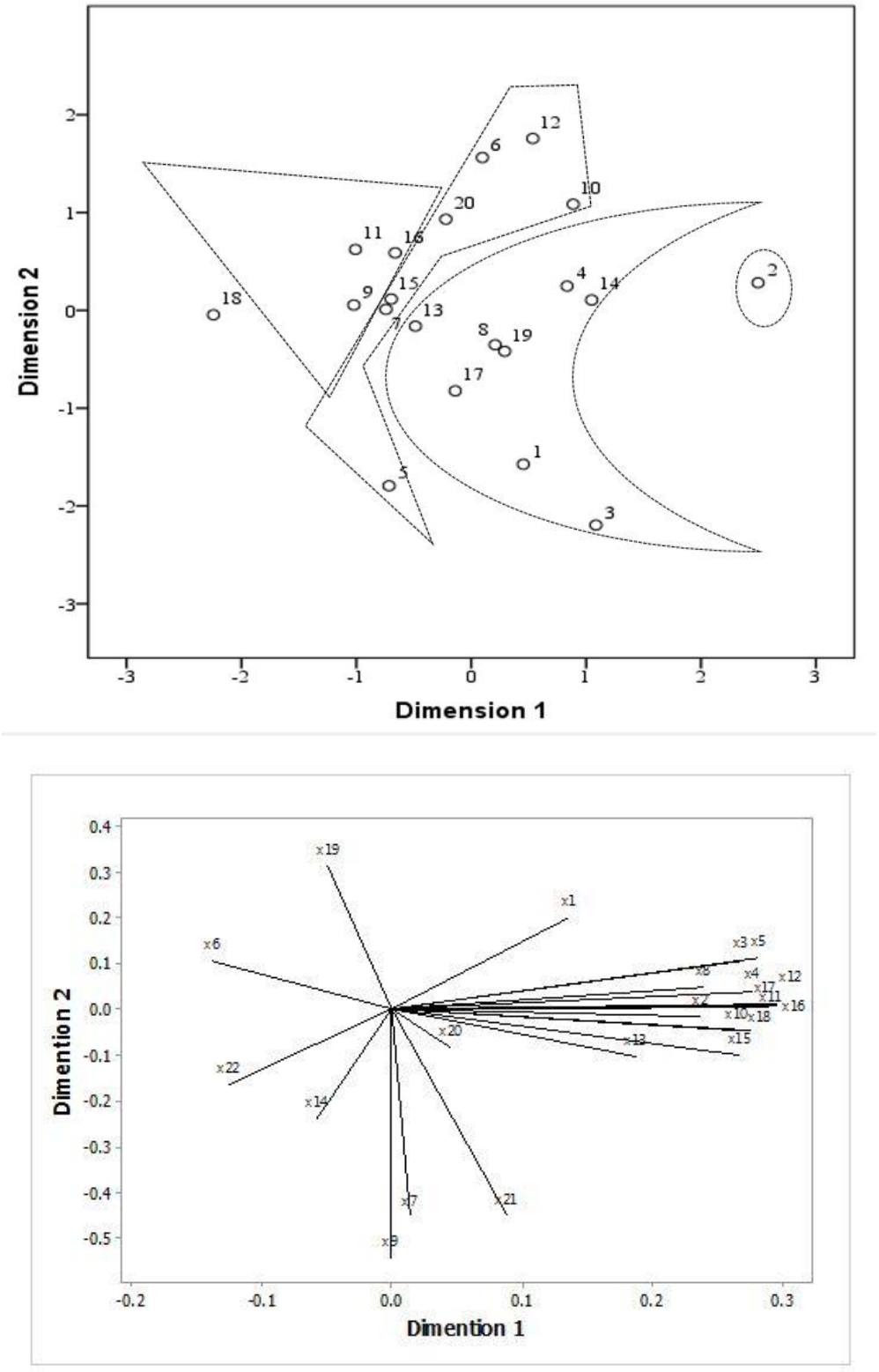

Fig. 5. Score (upper graph) and loading (lower graph) plots according to principal component analysis in *P. elaeagnifolia*.

X1: Shoot length, X2: Shoot diameter, X3: Leaf length, X4: Leaf width, X5: Leaf area, X6: Leaf length/Leaf width, X7: Flowering date, X8: Flower diameter, X9: Fruit set, X10: Nut Length, X11: Nut width, X12:nut weight, X13:nut diameter, X14: Nut ripening date, X15: Kernel length, X16: Kernel width, X17: Kernel weight, X18: Kernel diameter, X19: Nut moisture, X20: Nut protein, X21: Nut oil, X22: Yield.

Table 3. Model summary of the principal component analysis in *P. scoparia* and *P. elaeagnifolia*

Dimension	Cronbach's Alpha	Eigenvalue	Variance (%)
A. scoparia			
1	0.952	10.456	52.282
2	0.585	2.249	11.245
Total	0.97	12.705	63.527
A. eleagnifolia			
1	0.889947	6.470369	32.35184
2	0.765094	3.660857	18.30428
Total	0.949	10.13123	50.65613

Table 4. Pearson phenotypic correlation coefficients (above the main diagonal) and genotypic correlation coefficients (under the main diagonal) based on all measured traits for P. scoparia

Trait	Shoot length	Shoot diameter	Leaf length	Leaf width	Leaf area	Leaf length/Leaf width	Flowering date	Flower diameter	Fruit set	nut Length	nut Width	nut weight	nut diameter	Nut ripening date	Kernel length	Kernel width	Kernel weight	Kernel diameter	nut moisture	nut protein	nut oil	Yield
1	1	.62**	.61**	.42**	.45**	.10*	-.01	.26**	-.22	.27**	.33**	.36**	.241**	-.114	.30**	.41**	.35**	.37**	.30**	.06ns	-.37**	.02
2	.62**	1	.88**	.75**	.81**	-.10*	-.08	.63**	.06	.74**	.72**	.77**	.34**	-.19	.81**	.77**	.76**	.64**	-.23**	.28**	-.24**	.33
3	.58**	.78**	1	.87**	.94**	-.21	-.09	.74**	-.08	.67**	.76**	.78**	.44**	-.14	.65**	.78**	.72**	.69**	.02	.12	-.50*	.26
4	.41	.68**	.87**	1	.95**	-.63**	.02	.66**	-.01	.73**	.82**	.82**	.42**	-.12	.74**	.84**	.787*	.74**	-.08	.15	-.46*	.25
5	.43	.74**	.94**	.96**	1	-.44**	.009	.74**	-.10	.73**	.83**	.84**	.46**	-.05	.72**	.85**	.80**	.75**	-.03	.16	-.54*	.24
6	.09	-.11	-.22**	-.63**	-.42	1	.02	-.20	-.11	.09	-.46**	-.42**	-.18**	-.11	-.46**	-.46**	-.45**	-.42**	.21**	-.15	.12	-.08
7	-.016	-.08	-.09	.02	.01	.02	1	.20	-.05	.03	-.08	.24	.06	.009	.27	-.25	.04	-.17	-.18	.08	-.09	.01
8	.25	.57**	.75**	.67**	.75**	-.20**	.20	1	-.03	.54**	.71**	.72**	.57**	-.09	.55**	.73**	.70**	.74**	-.20**	.14	-.40	.32
9	-.251**	.11*	-.09	-.005	-.10*	-.16*	-.05	-.04	1	-.01	.005	-.07	-.05	-.05	.11*	-.002ns	-.04	.02	-.17**	.09	.12*	.84**
10	.24	.64**	.66**	.72**	.72**	-.41	.03	.54*	-.007	1	.87**	.90**	.43	.04	.97**	.88**	.91**	.75**	-.17	-.03	-.23	.15
11	.32	.65**	.763**	.81**	.82**	-.44*	-.08	.70**	-.007	.87**	1	.97**	.62**	.06	.83**	.95**	.93**	.88**	-.18	.09	-.29	.21
12	.36	.70**	.80**	.81**	.85**	-.38	.24	.72**	-.05	.92**	.97**	1	.67**	.29	.88**	.95**	.95**	.90**	-.14	.08	-.30	.19
13	.24	.33	.45*	.42	.46*	-.16	.06	.57**	-.05	.46**	.63**	.682**	1	-.15	.41**	.59**	.56**	.76**	-.24	-.02	-.06	.16
14	-.1	-.2	-.14	-.12	-.05	-.11	.009	-.09	-.05	.04	.06	.29	-.15	1	.37	.25	.20	.29	.05	.07	.14	.10
15	.29	.73**	.65**	.73**	.71**	-.45*	.27	.55*	.10	.96**	.83**	.87**	.40	.37	1	.87**	.91**	.75**	-.21	.03	-.17	.24
16	.4	.69**	.77**	.83**	.84**	-.45*	-.25	.72**	-.007	.87**	.95**	.95**	.58**	.25	.87**	1	.97**	.90**	-.16	.12	-.36	.22
17	.32	.67**	.71**	.77**	.79**	-.44*	.04	.69**	-.04	.90**	.93**	.94**	.56**	.20	.91**	.97**	1	.89**	-.21	.04	-.25	.14
18	.35	.57**	.69**	.74**	.74**	-.40	-.17	.73**	.001	.76**	.89**	.90**	.77**	.29	.75**	.91**	.9**	1	-.23	.03	-.27	.23
19	.25	-.13	.01	-.07	-.02	.16	-.18	-.16	-.1	-.20**	-.21**	-.19**	-.30**	.05	-.27**	-.20**	-.26**	-.31**	1	-.20	-.15	-.08
20	.05	.24	.13*	.14*	.17**	-.15*	.08	.14*	.08	-.04	.08	.06ns	-.02ns	.074	.03ns	.12*	.05	.04	-.27**	1	-.41	.19
21	-.36	-.22	-.53**	-.49**	-.58**	.12*	-.09	-.42**	.11	-.25**	-.29**	-.28**	-.04ns	.14	-.18**	-.38**	-.26**	-.28**	-.20**	-.43**	1	-.16
22	.07	.45**	.27**	.29**	.20**	-.11*	.01	.54**	.71**	.16*	.35**	.36**	.12*	.10	.44**	.29**	.13*	.38**	-.03	.19**	-.13*	1

** and * are signs for significant at the 0.01 and 0.05 levels.

Table 5. Pearson phenotypic correlation coefficients (above the main diagonal) and genotypic correlation coefficients (under the main diagonal) based on all measured traits for *P. elaeagnifolia*

Trait	Shoot length	Shoot diameter	Leaf length	Leaf width	Leaf area	Leaf length/Leaf width	Flowering date	Flower diameter	Fruit set	nut Length	nut Width	nut weight	nut diameter	Nut ripening date	Kernel length	Kernel width	Kernel weight	Kernel diameter	nut moisture	nut protein	nut oil	Yield
1	1	0.55*	0.3**	0.3**	0.4**	0	-0.3	0.1*	0.2	0	0.4**	0.6**	0.7**	0.07	0.01	0.2**	0.3**	0.7**	0.7**	0.01**	0.01**	0.09
2	0.6**	1	0.6**	0.6**	0.5**	0	0.25	0.08	0.05	0.3**	0.01	0.5**	0.3**	-0.1	0.4**	0.2**	0.3**	0.11*	0.5**	0.01**	0.01**	0.12
3	0.35	0.55*	1	0.83**	0.9**	0.25	-0.3	0.34	-0.1	0.3**	0.2**	0.6**	0.4**	0.01	0.4**	0.4**	0.6**	0.4**	0.3**	-0.1	-0.2	-0.1
4	0.38	0.56**	0.8**	1	0.93**	-0.3	-0.2	0.19	-0.2	0.1*	0.2**	0.4**	0.5**	-0.7	0.2**	0.3**	0.4**	0.4**	0.6**	0.01	-0.4*	-0.1
5	0.42	0.48*	0.9**	0.9**	1	0.01	0.01	0.3	-0.3	0.1**	0.3**	0.6**	0.6**	0.1	0.2**	0.4**	0.5**	0.5**	0.5**	0.01	-0.3	-0.2
6	0.01	0.01	0.2**	0.01**	0.01	1	0.24	0.22	0.16	0.3**	0.01	0.2**	0.01**	0.01	0.3**	0.06	0.2**	0.07	0.01**	-0.1	0.38	0.01
7	-0.3	0.25	-0.3	-0.2	0.01	0.24	1	0.01	0.01	0.12	0.01	0.4*	0.01	0.4*	0.09	-0.4	-0.1	0.36	0.23	0.01	0.3	0.21**
8	0.14	0.09	0.3**	0.1**	0.3**	0.2**	0	1	0.12	0.01	0.3**	0.2**	0.03	-0.3	0.08	0.5**	0.4**	0.3**	0.01	-0.2	0.28	0.1
9	0.2**	0.05	0.01*	0.01**	0.01**	0.1*	0	0.1*	1	0.05	0.1*	0.1*	0.01**	0.03	0.01	0.01	0.01	0.06	0.09	0.01**	0.4**	0.55**
10	0.01	0.33	0.35	0.13	0.17	0.28	0.12	0.01	0.05	1	0.09	0.57**	-0.2	0.4*	0.9**	0.4**	0.6**	-0.1	0.01	-0.1	0	0.01
11	0.46*	0.01	0.21	0.19	0.3	0.01	0.01	0.36	0.13	0.08	1	0.55*	0.24	0.42	0.02	0.7**	0.5**	0.58**	0.16	-0.3	0.36	-0.2**
12	0.62**	0.43	0.62**	0.46*	0.59**	0.22	0.4*	0.22	0.09	0.5**	0.5**	1	0.47*	0.31	0.5**	0.6**	0.8**	0.59**	0.33	-0.2	0.01	-0.2**
13	0.68**	0.3	0.4	0.55*	0.63**	-0.2	0.01	0.01	-0.1	0**	0.2**	0.4**	1	0.16	0.01**	0.03	0.1*	0.78**	0.64**	-0.1	-0.2	-0.2**
14	0.07	-0.1	0.01	-0.7	0.1	0.01	0.4*	-0.3	0.03	0.4*	0.42	0.31	0.16	1		0.34	0.02	-0.2	0.01	0.37	0.35	0.24**
15	0.01	0.37	0.47*	0.21	0.26	0.35	0.09	0.07	0.01	0.97**	0.03	0.55*	-0.2		1	0.51*	0.7**	-0.1	0.01	0.01	-0.1	0.01
16	0.24	0.17	0.47*	0.39	0.47*	0.05	-0.4	0.48*	0.01	0.48*	0.74**	0.65**	0.05	0.34	0.5**	1	0.85**	0.4	0.11	-0.1	0.08	-0.3**
17	0.33	0.3	0.62**	0.44	0.54*	0.22	-0.1	0.43	0.02	0.65**	0.54*	0.8**	0.16	0.02	0.7**	0.8**	1	0.48*	0.2	-0.1	0.01	-0.1
18	0.71**	0.14	0.45*	0.43	0.57**	0.06	0.36	0.3	0.03	0.01**	0.6**	0.6**	0.8**	-0.2	0.01**	0.4**	0.4**	1	0.57**	-0.2	0.06	-0.2**
19	0.53*	0.38	0.3	0.47*	0.41	-0.2	0.23	0.01	0.06	0.01*	0.1**	0.4**	0.8**	0.01	0.01*	0.1*	0.2**	0.7**	1	-0.3	-0.1	0.01
20	-0.4	-0.2	0.01*	0.01	0.01	0.01**	0.01	0.01**	-0.4*	0.01*	0.01**	0.01**	0.01**	0.37	0.01	0.01*	0.01*	0.01**	0.01**	1	-0.3	-0.2**
21	-0.1	-0.6**	0.01**	0.01**	0.01**	0.4**	0.3	0.3**	0.4	0.01	0.3**	0.01	0.01**	0.35	0.01*	0.1*	0.01	0.08	0.01**	0.01**	1	0.17*
22	0.07	0.2**	0.01**	0.01*	0.01**	0.01	0.21	0.1	0.55*	0.01	0.3**	0.2**	0.2**	0.24	0.01	0.5**	0.1*	0.5**	0.01	0.01**	0.2**	1

** and * are signs for significant at the 0.01 and 0.05 levels.

Phenotypic and genotypic correlations

Table 4 and Table 5 represent phenotypic and genotypic correlation coefficients for all pairs of the traits in *P. scoparia* and *P. elaeagnifolia*, respectively. The results showed that a high degree of association exists between the traits. Those in relation to the nut and kernel were closely correlated in a positive way according to both genotypic and phenotypic correlations. Moreover, traits that related to leaf characteristics showed highly positive correlations with one another and also with traits related to the nut and kernel in both species. Based on phenotypic correlations for yield, the only positive correlation of significance, occurred betweenthe yield and the percentage of fruit set in both of *P. scoparia* (r = 0.85) and *P. elaeagnifolia* (r = 0.55). According to genotypic correlations, many of the calculated correlations were significant. Specifically, however, the yield produced by *P. scoparia* was found to correlate positively with the kernel length (r = 0.44), kernel width (r = 0.29), kernel diameter (r = 0.38), flower diameter (r = 0.54), shoot diameter (r = 0.45), and percentage of fruit set (r = 0.72). Furthermore, the yield produced by *P. elaeagnifolia* correlated with shoot diameter (r = 0.45), leaf area (r = 0.39), the percentage of fruit set (r = 0.55), nut weight (r = 0.30), kernel width (r = 0.53), kernel diameter (r = 0.54), kernel protein (r = 0.46) and kernel oil (r = 0.21).

Stepwise regression analysis

The results of the stepwise regression for both species are shown in Table 6. In *P. scoparia*, the final model was established after three steps of screening for the percentage of fruit set, flower diameter, and leaf length. Meanwhile, the most influential factors on the yield of *P. elaeagnifolia* were the percentage of fruit set and leaf area.

Table 6. Stepwise regression model taking yield as response variable in *P. scoparia* and *P. elaeagnifolia* separately

Step	Model	Unstandardized Coefficients	Stoddard error	Standardized Coefficients	t	P-value
	A. scoparia					
1	Intercept	-0.519	0.205	-	-2.531	0.021
	Fruit set percentage	0.125	0.018	0.846	6.738	0
2	Intercept	-1.447	0.415	-	-3.487	0.003
	Fruit set percentage	0.126	0.014	0.858	8.831	0
	Flower diameter	0.642	0.177	0.352	3.623	0.002
3	Intercept	-2.109	0.467	-	-4.52	0
	Fruit set percentage	0.119	0.02	0.786	5.992	0
	Flower diameter	0.712	0.112	0.381	4.132	0
	Leaf length	0.291	0.103	0.37	2.819	0.012
	A. elaeagnifolia					
1	Intercept	0.211	0.143	-	1.475	0.157
	Fruit set percentage	0.049	0.017	0.554	2.822	0.011
2	Intercept	-1.447	0.415	-	-3.487	0.003
	Fruit set percentage	0.086	0.016	0.668	7.231	0
	Leaf area	0.532	0.133	0.355	2.891	0.009

Discussion

The higher the variation among genotypes of a plant species in traits of interest, the higher the possibility of finding new genes or QTLs to be applied in revealing new cultivars through conventional and molecular breeding (Cruz, 2013; Holsinger and Weir, 2009; Xiong et al., 2015). Considering the results of current study, a high amount of variability was clearly observed in most of the measured traits in both of *P. scoparia* and *P. elaeagnifolia* species (Tables 1 and 2). The high variations of tree and nut characteristics of *P. scoparia* confirmed in this study, are in agreement with previous reports by Khadivi-Khub and Anjam (2014)

evaluating the tree and nut attributes, and Dvin et al. (2017), investigating the nut characteristics of this species in very diverse plant materials. Furthermore, Nikoumanesh et al. (2011) reported that leaf area has the highest coefficient of variation in some almond species. Zeinalabedini et al. (2012) obtained a high range of variability that was based on the coefficient of variations in relation to morphological traits in the almond germplasm, which is consistent with the results obtained in the current study. There were higher values of phenotypic coefficients of variation, as compared to the genotypic coefficients of variation, which indicate the undeniable role of the environment in affecting the phenotypes (Farshadfar et al., 2001).

Many of the measured traits in both species including the leaf, flower, nut and kernel attributes, as well as yield in *P. elaeagnifolia* showed very high levels of heritability ($H^2 > 90\%$), however some such as shoot diameter and kernels' moisture content showed lower levels of heritability ($H^2 < 50\%$) in both species (Tables 1 and 2). Such high levels of heritability could be considered as genotypic capability of generating new populations with high quality kernels (Riasat et al. 2018). Chandrababu and Sharma (1999) also reported a high level of heritability for several morphological traits of almonds species (H^2 = 89%, 90%, 94%, 95%, 90%, and 92% for nut length, nut width, nut weight, kernel weight, kernel length, and kernel width, respectively). Similarly, Kester et al. (1977) obtained similar results when studying almonds. Such high levels of heritability in wild almond genotypes deserve more attention in breeding programs.

A cluster diagram assigns different branch lengths to each genotype, thereby representing the distance between them. The longer the branch length, the more distanced the genotypes are genetically (Saed-Moucheshi et al. 2013). Besides a high level of variation among the genotypes in terms of the measured traits, a high level of similarity could still be realized among the genotypes, which could be explored as an advantage in breeding programs. According to the results the genotypes of *P. scoparia* clustered into three distinct groups (Fig. 2), whilst, the genotypes of *P. elaeagnifolia* classified in four separated groups (Fig. 3). The dissimilarity between genotypes denotes the capability of generating new genotypes and producing different associations or segregations of genes, thereby facilitating a partial removal of former linkages or the creation of new ones that can be applicable in both classical and modern breeding methods (Dicenta and Garcia, 1992). In order to generate new genotypes in a subsequent generation (with new linkage groups or new population properties), it is a common practice to use distant genotypes (Sorkheh et al., 2009).

Further into the grouping of the genotypes and after reviewing the results of the clustering, principal component analysis (PCA) was performed as a statistical procedure to gauge the precision of clustering. According to the PC results, the clustering of the genotypes was performed well, and the two graphs indicated separated groups as previously determined by the cluster analysis in both species (Table 3, Figures 4 and 5). By using PCA, the individuals and even the traits could be categorized into groups, thereby saving time and effort when screening for germplasm and populations that are worthy of breeding, as well as parent selection (Khadivi-Khub and Anjam, 2014; Rahimi Dvin et al., 2017).

To illustrate the association between traits and to find effective traits in relation to important variables, correlation coefficients based on Pearson's method were used for both species. The results showed that a high degree of association exists between the traits (Tables 4 and 5). Generally, genotypic correlation coefficients were lower than the phenotypic ones, although there were some cases in which the genotypic correlation was higher

than the phenotypic (Tables 4 and 5). These results are in agreement with previous findings on wheat (Riasat et al., 2018), pomegranate (Mars and Marrakchi, 1999) and almond (Dicenta et al., 1993). However, the more the number of traits that exist, the more coefficients need to be considered, thereby complicating the capability of correlation coefficients in finding associations between traits (Hazel, 1943). To that end, multivariate analysis can reduce the number of indices and facilitate research when exploring the association between traits. One of these techniques is principal component analysis which intends to reduce the dimensions of the data and enables the assessment of more information through fewer indices (Richards et al., 2002). The loading plot for both of *P. scoparia* and *P. elaeagnifolia* are presented in Figures 4 and 5, whereby the first two components are used. The graphs clearly supported the results obtained in the correlation analysis where the percentage of fruit set appeared to be the most prominent trait that correlated with the yield in both species. Furthermore, the traits in relation to kernel weight showed higher associations with the yield than with the other traits in both species. Positive correlations between leaf properties (e.g. leaf area) and fruit set indicate a direct association between vegetative and reproductive parameters. This correlation can be considered as an appropriate route to improve the growth vigor of trees located in arid areas where plant survival demands a rapid and strong growth at the start of the growing season (Nikoumanesh et al., 2011). Previous reports have also indicated that leaf area has positive correlations with the characteristics of the nut and kernel (Khadivi-Khub and Anjam, 2014; Zeinalabedini et al., 2016; and Karimi et al., 2009). Due to genetic linkage and the existence of linkage groups among genes, the availability of strong correlations between traits might bring about more traits of interest in the next generation when selecting for other traits (Dicenta and Garcia, 1992). In both species, a significant negative genotypic correlation was found between oil content and protein content of the kernel ($r = -0.44$ and $r = -0.43$ in *P. scoparia* and *P. elaeagnifolia*, respectively) indicating that it is rather difficult to obtain high amounts of oil and protein in one genotype (Colic et al., 2012). Regarding principal component analysis, a previous study has shown that traits in relation to the leaf and vegetative growth have different dimensions and relationships in separate species (Nikoumanesh et al., 2011). In a study on wild almond species, the features of the nut showed a larger contribution to species differentiation than those of the leaf (Sorkheh et al., 2009).

To detect a linear association and to find effective traits through yield, linear regression was used in both species, but since the regression coefficients may affect each other, some false results might be generated in the whole model of the multiple regression (Saed-Moucheshi et al., 2013). Accordingly, stepwise selection regression was applied to reveal and screen the most effective variables through the yield of these two almond species. The most influential factors on the yield of *P. elaeagnifolia* were the percentage of fruit set and leaf area, this is while in *P. scoparia*, the percentage of fruit set, flower diameter, and leaf length screened as the most effective factors on yield (Table 6). Imani and Shamili (2018) showed that nut length and nut width are the main variables that cause variations in the nut weight (i.e. the yield) of almonds per the regression model.

Conclusion

High levels of genotypic and phenotypic variations exist among the evaluated genotypes, and these include many of the vegetative and reproductive attributes. These observations denote the valuable nature of available genetic resources that originate from these two wild almond species. Furthermore, high levels of heritability for many of the traits pertaining to the growth

and nut-related characteristics suggest the feasibility of genetic improvement through conventional breeding. The vegetative growth, along with the features of the nut and kernel correlated strongly with the final amount of yield, whereas qualitative traits such as oil and protein contents did not share strong correlations with the amount of almond yield. Genotypes from different clusters could be used in crossbreeding programs to generate greater variations in vegetative and reproductive characteristics of wild almonds.

Compliance with Ethical Standards

Funding

This research was funded by Shiraz University (the affiliated institute of the authors) and there was no external funding for this research.

Conflict of interest

The authors declare that they have no conflict of interest.

Reference

1. Anumalla M, Roychowdhury R, Geda C.K, Mazid M, Rathoure A.K. 2015. Utilization of plant genetic resources and diversity analysis tools for sustainable crop improvement with special emphasis on rice. International Journal 3,1155-1175.
2. Atanasov A.G, Waltenberger B, Pferschy-Wenzig E.M, Linder T, Wawrosch C, Uhrin P, Temml V, Wang L, Schwaiger S, Heiss E.H, Rollinger J.M, Schuster D, Breuss J.M, Bochkov V, Mihovilovic M.D, Kopp B, Bauer R, Dirsch V.M, Stuppner H. 2015. Discovery and resupply of pharmacologically active plant-derived natural products: a review Biotechnology advances 33,1582-1614.
3. Chandrababu R, Sharma R. 1999. Heritability estimates in almond [Prunus dulcis (Miller) DA Webb]. Scientia Horticulturae 79, 237-243.
4. Cohen, J.I, Williams, J.T, Plucknett D.L, Shands H. 1991. Ex situ conservation of plant genetic resources: global development and environmental concerns. Science, 253(5022), 866-872.
5. Colic S, Rakonjac V, Zec G, Nikolic D, Aksic M.F. 2012. Morphological and biochemical evaluation of selected almond [Prunus dulcis (Mill.) DA Webb] genotypes in northern Serbia. Turkish Journal of Agriculture and Forestry 36, 429-438.
6. Cruz C.D. 2013. Genes: a software package for analysis in experimental statistics and quantitative genetics. Acta Scientiarum Agronomy 35, 271-276.
7. Dicenta F, Garcia J. 1992. Phenotypical correlations among some traits in almond. Journal of Genetics and Breeding 46, 241-246.
8. Dicenta F, García J, Carbonell E. 1993. Heritability of flowering, productivity and maturity in almond. Journal of Horticultural science 68,113-120.
9. Dvin S.R, Eshghi S, Avanzato D, Ansari A. 2017. Diversity in the nut and kernel characteristics of seven populations of Prunus scoparia from the central and southern Zagros regions of Iran by comparison with three other almond species. FRUITS 72, 370-381.
10. Falconer D.S, Mackay T.F.C. 1996. Introduction to quantitative genetics, Ronald Press, New York.
11. Farshadfar E, Farshadfar M, Sutka J. 2001. Combining ability analysis of drought tolerance in wheat over different water regimes. Acta Agronomica Hungarica 48, 353-361.
12. Gharaghani A, Solhjoo S, Oraguzie N. 2017. A review of genetic resources of almonds and stone fruits (Prunus spp.) in Iran. Genet. Resur. Crop Evol. 64, 611-640.
13. Gharaghani A, Eshghi S. 2015 Prunus scoparia, a Potentially Multi-Purpose Wild Almond Species in Iran. Acta Hortic. 1074, 67-72.
14. Hazel L.N. 1943. The genetic basis for constructing selection indexes. Genetics 28, 476-490.
15. Holsinger K.E, Weir B.S. 2009. Genetics in geographically structured populations: defining, estimating and interpreting F ST. Nature Reviews Genetics 10(9), 639-650.
16. Imani A, Shamili M. 2018. Almond nut weight assessment by stepwise regression and path analysis. International Journal of Fruit Science 18 (3), 1-6.
17. Institute S. 2003. SAS version 9.3. SAS Institute, Cary, NC.
18. Karimi H, Zamani Z, Ebadi A, Fatahi M. 2009. Morphological diversity of Pistacia species in Iran. Genetic Resources and Crop Evolution 56, 561-571.

19. Kester D, Hansche P, Beres V, Asay R. 1977. Variance components and heritability of nut and kernel traits in almond. Journal of American Society for Horticultural Science 102, 264-266.

20. Khadivi-Khub A, Anjam K. 2014. Morphological characterization of Prunus scoparia using multivariate analysis. Plant systematics and evolution 300,1361-1372.

21. Kiani S, Rajabpoor S, Sorkheh K, Ercisli S. 2015. Evaluation of seed quality and oil parameters in native Iranian almond (Prunus L. spp.) species. Journal of forestry research 26, 115-122.

22. Mars M, Marrakchi M. 1999. Diversity of pomegranate (Punica granatum L.) germplasm in Tunisia. Genetic Resources and Crop Evolution 46, 461-467.

23. Nikoumanesh K, Ebadi A, Zeinalabedini M, Gogorcena Y. 2011. Morphological and molecular variability in some Iranian almond genotypes and related Prunus species and their potentials for rootstock breeding. Scientia horticulturae 129, 108-118.

24. Pinar H, Unlu M, Ercisli S, Uzun A, Bircan M. 2016. Genetic analysis of selected almond genotypes and cultivars grown in Turkey using peroxidase-gene-based markers. Journal of forestry research 27, 747-754.

25. Riasat M, Pessarakli M, Niaz A.A, Saed-Moucheshi A. 2018. Assessment of different wheat genotypes with altered genetic background in response to different salinity levels. Journal of Plant Nutrition 33, 1-13.

26. Richards R, Rebetzke G, Condon A, Van Herwaarden A. 2002. Breeding opportunities for increasing the efficiency of water use and crop yield in temperate cereals. Crop science 42, 111-121.

27. Saed-Moucheshi A, Fasihfar E, Hasheminasab H, Rahmani A, Ahmadi A. 2013. A review on applied multivariate statistical techniques in agriculture and plant science. Int J Agron Plant Produc 4, 127-141.

28. Sorkheh K, Shiran, B, Rouhi V, Asadi E, Jahanbazi H, Moradi H, Gradziel T.M, Martínez-Gómez P. 2009. Phenotypic diversity within native Iranian almond (Prunus spp.) species and their breeding potential. Genetic resources and crop evolution 56, 947-961.

29. Xiong J.S, Ding J, Li Y. 2015. Genome-editing technologies and their potential application in horticultural crop breeding. Horticulture research 2, 15019. doi:10.1038/hortres.2015.19.

30. Zeinalabedini M, Majidian P, Dezhampour J, Khakzad M, Farsi M. 2016. First Report of a set of Genetic Identities in Prunus Rootstocks by SSR Markers. Journal of Plant Molecular Breeding 4, 17-25.

31. Zeinalabedini M, Sohrabi S, Nikoumanesh K, Imani A, Mardi M. 2012. Phenotypic and molecular variability and genetic structure of Iranian almond cultivars. Plant Systematics and Evolution 298, 1917-1929.