Adv. Hort. Sci., 2020 34(4): 381-395 DOI: 10.13128/ahsc-9005

AHS
Advances in Horticultural Science

Morphological and biochemical classification of Iranian mango germplasm collection by multivariate analysis: implications for breeding

D. Samsampour [1] (*), H. Kazemzadeh-Beneh [1], G.-R. Damizadeh [2], Z. Mirzaei [1]

[1] *Department of Horticulture Science, Biotechnology and Molecular Genetic Section, University of Hormozgan, Bandar Abbas, Iran.*

[2] *Agriculture and Natural Resources Research Center of Hormozgan, Bandar Abbas, Iran.*

OPEN ACCESS

Key words: cluster grouping, factor analysis, genetic variability, mango fruit, principal component.

(*) **Corresponding author:**
samsampoor@hormozgan.ac.ir

Citation: SAMSAMPOUR D., KAZEMZADEH-BENEH H., DAMIZADEH G.-R., MIRZAEI Z., 2020 - *Morphological and biochemical classification of Iranian mango germplasm collection by multivariate analysis: implications for breeding.* - Adv. Hort. Sci., 34(3): 381-395

Copyright:
© 2020 Samsampour D., Kazemzadeh-Beneh H., Damizadeh G.-R., Mirzaei Z. This is an open access, peer reviewed article published by Firenze University Press (http://www.fupress.net/index.php/ahs/) and distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement:
All relevant data are within the paper and its Supporting Information files.

Competing Interests:
The authors declare no competing interests.

Received for publication 9 June 2020
Accepted for publication 23 October 2020

Abstract: This study was conducted at southern Iran with the aim to evaluate the phenotypic diversity of 84 mango cultivars by using morphological and biochemical traits. Two other industrial cultivars 'Longra' and 'Senderi' used as control of the study. Descriptive results indicated that the value of both quantitative and qualitative variables all the cultivars were lower than the 'Senderi'. The variability among cultivars was highly significant in the measured traits. The fruit weight, fruit length, stone weight, stone length, fruit beak, dry matter, and petiole length showed highly discriminating power. Pearson's correlation analysis revealed high variability due to the existence of significant positive and negative correlations among traits. Agglomerative hierarchical clustering confirmed remarkable variation in the studied germplasm and identified three major clusters with several sub-clusters. The 80.20, 70.11, and 100% of the quantitative variability was explained by principal component analysis (PCA), factor analysis (FA), and liner discriminant function (LDF) where fruit descriptors contributed most of the total variation, respectively. However, multivariate analysis proved that fruit related characters were most powerful to differentiate cultivars. The cultivars displayed distinct grouping by FA and LDF compared to PCA. The results revealed that the Iranian mango germplasm has a high potential for specific breeding project regarding fruit size and quality that should be further completed by a molecular marker analysis.

1. Introduction

The Mango (*Mangifera indica* L.), from *Anacardiaceae* family, is a fruit species native to Asia and grown comprehensively in tropical and subtropical countries, that originated as an allopolyploid from eastern India, Assam and Burma region (Krishna and Singh, 2007). Mangoes are extensively cultivated in the orchards or often planted in the fruit-gardens man-

ner in the southern region of Iran. The main areas of this region which is famous for the mango production are Hormozgan, Sistan and Balouchestan, and Kerman (Kahnuj and Jiroft) provinces.

The role of germplasm characterization in varietal development of crops as genotypes with desirable traits has been well identified and utilized in the crop improvement programs and also to determine evolutionary relationships (Piyasundara *et al*., 2008; Donkor *et al*., 2019). However, the accessibility of the germplasm depends mostly on the information available on characterization and evaluation. The investigation of plant material with desired traits by means of the morphological characterization by multivariate statistical techniques is an essential step for the effective utilization of crop germplasm (Piyasundara *et al*., 2008). Multivariate statistical techniques, PCA and cluster analysis are the popular multivariate techniques that widely applied to identify genetic diversity in germplasm of olives (Hagidimitiou *et al*., 2005), strawberry (Lavin *et al*., 2005), tea (Piyasundara *et al*., 2008), mangoes (Sennhenn *et al*., 2013; Jamil *et al*., 2015) and sour cherry (Ganopoulos *et al*., 2016). Among them, multivariate techniques guarantee the accurate interpretation of the information generated through characterization studies.

The local farmers believe that the mango cultivars cultivated at southern Iran may have been derived from Pakistan and Indian germplasms over 300-400 years ago. From the beginning of mango cultivation in the southern regions of Iran, mangoes have a local name by the farmers and the native people living based on their appearance and farmer's interest. Beside this, the area under mango cultivation is increasing each year in the southern regions of Iran, but the basic information regarding the analysis of genetic diversity and identification cultivars in order to improve breeding programs are lacking. Before any breeding programs on mango, there is the need to collect and characterize the local cultivars as plant material that are available in Iran. Moreover, the evaluation of morphological traits through phenotyping constitutes is a quick method to characterize the mango germplasm and being provide useful qualitative information for the breeding.

Therefore, the main objectives of the present study were (i) to identify the phenotypic diversity in 86 local mango cultivars of the Hormozgan province of southern Iran using morphological approaches based on multivariate statistical techniques, (ii) to evaluate specific traits for breeding and to progress future genetic resource conservation strategies.

2. Materials and Methods

Experimental areas survey

The experiments were conducted at Rudan, Siyahu, and Minab germplasm collections (RSM-collection), located in Hormozgan province of Iran (Fig. 1). Samplings were collected from RSM-collection during growing season for leaf (October to November), during flowering for flower (February and April) and fruit harvesting (July to September). Rudan, Siyahu, and Minab lie within latitude 57° 6' N, 27° 7' N and 49° 49' N and longitude 27° 7' E, 57° 11' E, and 34° 4' E and altitude 100, 300 and 700 m above sea level, respectively. The RSM-collection has an average annual rainfall of 227, 250 and 200 mm in two seasons (October and January) and main daily temperature of 28, 29 and 18°C, respectively (Metrological Service, Bandar Abbas, 2014-2017).

Plant material

A total of 84 cultivars from an inspection on mango gardens or scatter manner planted in RSM-collection followed by two industrial cultivars including 'Longra' and 'Senderi' as controls were comprised in the sampling procedure for characterization (Table S1). To assay adult trees of RSM-collection germplasm, three trees of each cultivar were randomly chosen and were labeled to use data collection. The labeled trees were at fruit-bearing capacity, healthy and in crop condition at beginning of the study. Sampling and morphological assessment of 49 variables as mango descriptors were programmed in the experimental collection for three consecutive years (2014-2017) at the same time of harvest season. Different horticultural practices, including fertil-

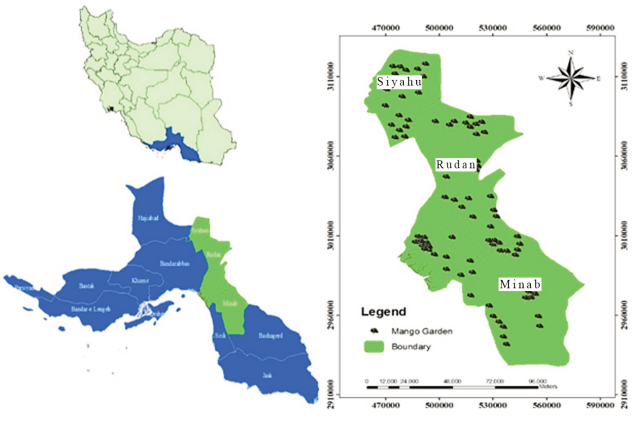

Fig. 1 - Geographical zones of the 84 mango cultivars followed by two industrial cultivars as controls used in this study from RSM- germplasm (Rudan, Siyahu, and Minab collection) located at southern of Iran (black trees indicate sampling locations).

izer application, spraying, irrigation and others, were performed at regular intervals each year.

Morpho-physiological analysis of traits

A total of 49 morphological and biochemical traits including 22 quantitative and 27 qualitative traits were scored following the guidelines system for mango descriptors published by International Plant Genetic Resources Institute (IPGRI), Rome for tree, leaf, flower and fruit traits (Mukherjee, 1989; PGRI, 2006). The quantitative traits were fruit length [FrLl], fruit breadth [FrBr], fruit skin thickness [FrSkTh], fruit weight [FrWe], flesh content (fruit flesh weight/fruit weight) [FrCo], fiber length [FrFi], stone length [StLe], stone weight [StWe], stone fiber length [StFiLe], pH [pH], TSS [TSS], titratable acidity [TiAc], total sugar [ToSu], moisture [MoSt], dry matter [DrMa] for fruit descriptors, leaf length [LeLe], leaf width [LeWi], leaf ratio [LeRa], petiole length [PeLe], petiole ratio [PeRa] for leaf descriptors, trunk circumference [TrCi] and canopy diameter [CaDi] for trunk descriptors. The qualitative traits analyzed were fruit shape [Frsh], flesh texture [FlCo], adherence [AdHe], fiber in pulp [FiPu], quantity of fiber [QuFi], stalk insertion [StIn], basal cavity [BaCa], beak [Be], beak type [BeTy], sinus [Si], sinus type [SiTy], groove stone [GrSt], shoulder [Sh], shoulder slope [ShSl], skin colour ripe fruit [SkCoRiFr] for fruit descriptors, inflorescence position [InPo], inflorescence shape [InSh], inflorescence colour [InCo], inflorescence hairiness [InHa], flower type [FlTy] for flower descriptors, leaf shape [LeSh], leaf colour [LeCo], leaf texture [LeTe], leaf tip [LeTi], leaf margin [LeMa] for leaf descriptors, trunk number [TrNu] and Tree habit [TrHa] for trunk descriptors.

For all three years, measurements performed by the same two persons to avoid errors due to individual variation. In addition, to diminish the environmental effects, all parameters were averaged over three years. Some measurements such as leaf, fruit and tree descriptors were performed in garden-head and others like flower, biochemical and physiological traits of fruit were measured at the laboratory of plant biotechnology of Hormozgan University. Total soluble solid and pH were determined using a digital refractometer (ATC1E-ATAGO, Japan), and pH meter (PL-500, Taiwan), respectively. Titratable acidity of freshly extracted juices and total sugar the fruits were calculated by using standard process procedure to AOAC (2000) and Omokolo *et al*. (1996), respectively. For morpho-metrical flower data collection, ten recently opened flowers from tree of each cultivar were randomly collected, pooled and conserved in ethanol (70%) until measurements. Afterward, morpho-metrical of flower was analyzed by using microscope (model: IX3).

Data analysis

Descriptive analysis (minimum, maximum, mean, standard deviation and coefficient of variation) for each of 49 studied traits were calculated and the test of normality was accomplished on data to approve ANOVA assumptions. Statistical differences in the Variations of each trait among the cultivars was computed by one-way analysis of variance (ANOVA) at $p \leq 0.01$ using SPSS 21.0 (SPSS, Inc., Chicago, IL, USA) after verifying normal distribution of dependent variables by Kolmogorov-Smirnov test. Within correlation analysis, the Pearson coefficient (parametric) was used to measure the correlation among quantitative traits. The principal component analysis (PCA), factor analysis (FA), liner discriminant function (LDF), and agglomerative hierarchical clustering (AHC) were directed to analyze data in order to visualize possible differences among the mango cultivars. In each case, a biplot was drawn based on most important components to facilitate the visualization of the results. For dendrogram construction, the combined data from both the quantitative and qualitative traits were considered. To estimate the genetic dissimilarity component, the Euclidean, and Ward's method was selected as chosen distance for the agglomerative hierarchical clustering. All multivariate analysis was performed using XLSTAT software (version 2016.2).

3. Results

Variance analysis of traits

The high morphological variation observed among studied cultivars. The mean squares of mango cultivars was significant for the LeLe, LeWi, LeCo, LeTe, FrLe, FrWe, FrSh, StLe, StWe, StFiLe, FiPu, FlFiLe, ToSu, DrMa, MoSt, Brix, InHa, CaDi, TrNu, and TrCi, but they were similar in the rest of the characteristics (data not shown). The variance variation in quantitative traits was more than qualitative traits.

Descriptive statistics

The result of descriptive analysis demonstrated the high variation of CV for SiTy (95%), Si (94%), StFiLe (67%), pH (64%), BeTy (62%), TrNu (62%), LeCo (55%), InHa (52%), FrSh (52%), TrCi (42%), QuFi (41%), TrHa (41%), LeTu (40%) Be (36%), FlTy (35%),

FlFiLe (34%), StIn (34%), InSh (33%) and InCo (31%), respectively. In contrast, the other variables revealed low coefficient of variation (<30%) (Supplementary Table S2). In the traits mentioned above, the high they produced the CV, the greatest was the variability regarding the fruit descriptor was proportional to the CV. For leaf descriptor, the most relevant traits were: LeCo, and LeTu whereas InHa, InCo, InSh, and FlTy were the most relevant character for the flower descriptor. The studied cultivars exhibited high morphological diversity, with some quantitative and qualitative traits.

Correlations for quantitative variables

Pearson's correlation (parametric) among 22 morphological of quantitative characters were presented in Table 1. Results demonstrated that the 28 positives and 11 negative significant correlations. A significant positive correlation was obtained between FiLe and StLe (r= 1.00), and FrBr and FrWe (r= 0.734). In contrast, highest negative correlation was resulted among MoSt and DrMa (r= -0.996). Regarding fruit descriptors, the results of the paired linear correlation indicated that FrWe was positively correlated with FrBr, FrLe, StWe, and TSS while was negatively correlated with CaDi.

In the present study, significant correlation was found between fruit weight and other quality variables specially TSS. Additionally, the Pearson correlation of FrBr with FrWe, FlCo, and StWe was positive significant whereas was negative significant with TSS (Table 1). Furthermore, significant positive correlations was also observed among: FlCo and FrSkTh (r= 0.244), TSS and FrSkTh (r= 0.347), FlCo and FrWe (r= 0.461), FrWe and StLe (r= 0.247), FrWe and StWe (r= 0.558), StLe and TSS (r= 0.277), StWe and MoSt (r= 0.229), TiAc and ToSu (r= 0.294), and ToSu and DrMa (r= 0.294) while significant negative correlation was detected between FlCo and StWe (r= -0.431).

Principal component analysis of quantitative variables

According to apply Kaiser's criterion ("Eigenvalue" >1) (Kaiser, 1958), PCA analysis of the 22 quantitative traits resulted in 9 components for explaining total of variation among cultivars. PCA revealed the first nine

Table 1 - Correlation coefficients (Pearson) among 22 quantitative traits in 86 mango cultivars

	FrLe	FrBr	FrSkTh	FrWe	FlCo	FiLe	StLe	StWe	StFiLe	pH	TSS	TiAc	ToSu	MoSt	DrMa	LeLe	LeWi	LeRa	PeLe	PeRa	TrCi	CaDi
FrLe	1																					
FrBr	**0.426**	1																				
FrSkTh	0.44	-0.148	1																			
FrWe	**0.456**	**0.734**	0.022	1																		
FlCo	**0.243**	**0.372**	**0.244**	**0.46**	1																	
FiLe	-0.088	-0.087	-0.054	0.015	0.035	1																
StLe	**0.591**	0.094	-0.025	**0.247**	0.1	1	1															
StWe	**0.239**	**0.381**	-0.189	**0.558**	**-0.434**	0.035	0.187	1														
StFiLe	0.63	0.011	0.115	0.112	0.056	0.103	0.087	0.027	1													
pH	-0.213	-0.113	-0.105	-0.107	0.023	-0.364	-0.188	-0.115	-0.064	1												
TSS	**0.304**	**-0.32**	**0.347**	-0.18	0.029	-0.029	**0.277**	-0.155	-0.044	-0.062	1											
TiAc	0.114	-0.015	-0.123	-0.039	0.026	0.055	0.112	-0.074	0.106	**-0.383**	0.028	1										
ToSu	-0.65	**-0.23**	0.08	-0.138	0.105	-0.071	-0.002	-0.191	0.005	0.098	**0.471**	**0.294**	1									
MoSt	0.091	0.116	0.113	0.165	-0.094	-0.099	0.013	**0.229**	0.041	-0.193	-0.189	-0.155	**-0.29**	1								
DrMa	-0.073	-0.105	-0.11	-0.16	0.082	0.099	0.005	-0.211	-0.044	0.201	0.195	0.149	0.294	-0.996	1							
LeLe	0.205	0.142	0.108	0.069	0.099	-0.219	0.018	-0.033	0.047	**-0.236**	0.138	0.089	-0.03	-0.013	0.016	1						
LeWi	0.101	0.14	0.023	0.124	0.106	-0.361	-0.092	0.005	-0.067	-0.209	-0.041	0.079	-0.024	-0.023	0.039	**0.681**	1					
LeRa	0.113	0.048	0.113	-0.019	0.019	0.221	0.125	-0.047	0.132	-0.039	**0.215**	0.004	0.001	0.018	-0.033	**0.429**	**-0.342**	1				
PeLe	0.103	0.159	0.072	0.048	0.108	-0.12	-0.025	0.008	0.058	-0.198	0.03	0.093	-0.018	0.026	-0.022	**0.651**	**0.388**	**0.347**	1			
PeRa	-0.032	0.076	0.009	0.018	0.084	0.026	-0.017	0.033	0.013	-0.048	-0.094	0.031	-0.021	0.047	-0.044	-0.034	-0.092	0.055	**0.722**	1		
TrCi	-0.055	0.004	-0.084	0.118	0.046	-0.262	-0.138	0.128	-0.225	0.078	0.029	**-0.275**	-0.075	-0.011	0.018	-0.148	0.077	-0.056	-0.069	0.057	1	
CaDi	**-0.22**	-0.048	-0.085	0.067	-0.099	0.121	-0.152	0.118	-0.012	0.031	-0.54	-0.113	0.011	-0.095	0.102	-0.09	-0.132	0.105	-0.041	0.014	**0.645**	1

For full names of traits see Morpho-physiological analysis of traits head in Materials and Method section.

(PCs) with eigen values greater than value 1, which could explain 80.20% of the total variation (Table 2). PC1, which accounted for 15.91% of the total variation, was strongly associated with fruit traits, such as fruit length, fruit weight, fiber length, stone length, and fruit breadth. Hence, the cultivars with high value of PC1 have lower biochemical traits of fruit as well as smaller tree size (canopy diameter). PC2 accounting for 13.33% of the total variation was positively correlated with total solid soluble, fruit dry matter, trunk circumference and leaf traits, while negatively correlated with the most of fruit parameters. PC3 had high contributing factor loading from petiole length, leaf length, leaf width, petiole ratio, trunk circumference and contributed 11.37% of the total variation. PC3 suggested that leaf traits could be located in one index. PC4, accounted for 8.79% of the total variation, was most determined by the traits of fruit dry matter, fruit breadth, fruit weight, flesh content and pH. PC5 up to PC9 explained 7.28%, 7.08%, 5.93%, 5.51%, and 5.21% of the total variation, respectively. However; PC5 represents mainly flesh content, fruit skin thickness, fruit weight, leaf ratio, and petiole ratio; PC6 explains canopy diameter, TSS, pH, petiole length, and petiole ratio; PC7 describes leaf length, pH, and stone weight; PC8 illustrates stone fiber length, leaf ratio, trunk circumference, and titratable acidity; PC9 most demonstrates total sugar, canopy diameter, TSS, and stone weight.

In addition, PCA-biplot based on PC1 and PC2 exhibited that the cultivars had wider variation for quantitative traits and thus the biplot did not produce a distinct grouping among cultivars (Fig. 2A). According PCA biplot, negative values for PC1 indicate cultivars with high content of dry matter and total sugar as well as lower pH and smaller canopy diameter; however, the cultivars which were placed in the green line rectangle belong this group (Fig. 2A).

Table 2 - First 9 components from the PCA of 22 quantitative traits in 86 mango cultivars

Quantitative traits	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
FrLe	0.77	0.08	-0.32	0.04	0.07	-0.06	0.08	-0.08	-0.03
FrBr	0.61	-0.32	0.14	0.51	0.1	0.13	-0.02	0.14	-0.02
FrSkTh	0.07	0.23	0.06	-0.33	0.63	0.00	0.11	0.00	0.28
FrWe	0.68	-0.37	-0.05	0.44	0.20	0.12	0.05	0.13	0.22
FlCo	0.31	0.21	0.01	0.39	0.71	0.12	-0.27	0.08	-0.07
FiLe	0.64	0.15	-0.61	-0.16	-0.16	0.02	-0.03	-0.15	-0.18
StLe	0.64	0.15	-0.61	-0.16	-0.16	0.02	-0.03	-0.15	-0.18
StWe	0.40	-0.51	-0.06	0.08	-0.46	0.06	0.27	-0.04	0.30
StFiLe	0.17	0.07	0.01	-0.13	0.03	0.17	-0.2	0.54	0.06
pH	-0.43	-0.10	-0.16	0.22	0.23	0.33	0.21	-0.17	-0.33
TSS	0.08	0.58	-0.38	-0.30	0.18	-0.03	0.29	-0.20	0.30
TiAc	0.15	0.37	-0.03	0.06	-0.36	-0.26	-0.53	0.30	0.25
ToSu	-0.19	0.51	-0.26	0.07	0.09	-0.03	-0.11	-0.04	0.50
MoSt	0.30	-0.58	0.27	-0.6	0.18	-0.13	-0.08	-0.01	0.07
DrMa	-0.29	0.58	-0.28	0.61	-0.19	0.13	0.09	-0.01	-0.06
LeLe	0.43	0.47	0.56	0.01	-0.05	-0.18	0.42	0.06	-0.09
LeWi	0.25	0.21	0.49	0.31	0.00	-0.62	0.21	-0.18	0.03
LeRa	0.24	0.33	0.12	-0.34	-0.05	0.56	0.31	0.34	-0.12
PeLe	0.39	0.39	0.68	-0.02	-0.15	0.3	-0.05	-0.32	0.06
PeRa	0.15	0.09	0.37	-0.03	-0.13	0.55	-0.45	-0.52	0.13
TrCi	0.15	0.51	0.29	-0.16	-0.15	0.02	-0.09	0.32	-0.31
CaDi	-0.21	-0.14	0.01	0.14	-0.18	0.37	0.34	0.21	0.44
Eigenvalue	3.50	2.89	2.50	1.93	1.60	1.56	1.30	1.21	1.15
Variability (%)	15.91	13.13	11.37	8.79	7.28	7.08	5.93	5.51	5.21
Cumulative	15.91	29.04	40.41	49.20	56.47	63.65	69.48	74.99	80.20

For full names of traits see Morpho-physiological analysis of traits head in material and method section.

The highest positive values for PC1 displayed cultivars with high fruit weight, fruit length, fruit breadth, stone weight and high of fruit moisture as well as lower total sugar and dry matter, which were positioned in the green line oval, as shown in figure 2A. These cultivars can be applied as parents in the mango breeding program as a source of genes for a greater fruit size in a second cycle of recurrent selection.

Principal component analysis of qualitative variables

According to PCA, the cultivars were quantitatively distinct based on qualitative variables. Considering the Kaiser's criterion ("Eigenvalue" >1), nine significant components were obtained that explained 64.26% of the total variation (Table 3). PC1 was associated with tree habit, basal cavity, inflorescence shape, inflorescence colour and leaf shape accounted for 12.70% of the total variation. PC2 had high contributing factor loading from fruit traits such as stalk insertion, beak, beak type, shoulder slope and contributed 10.43% of the total variation. The more PC3 value was positively correlated to sinus and sinus

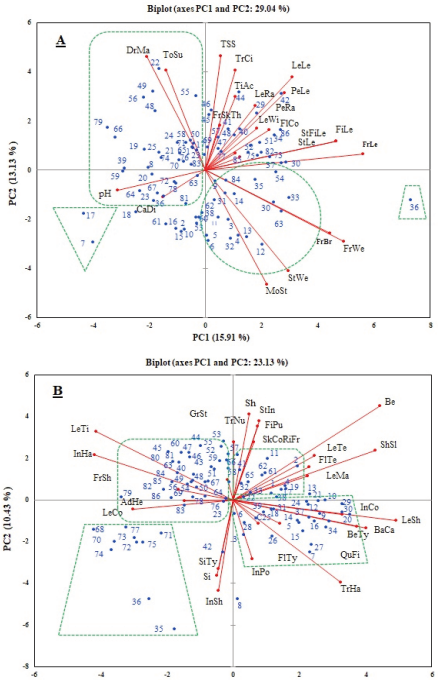

Fig. 2 - A) PCA biplot of the 22 quantitative traits with regard to the first two principal components. B) PCA biplot of the 27 qualitative traits with regard to the first two principal components among 86 mango cultivars.

Table 3 - First 9 components from the PCA of 22 quantitative traits in 86 mango cultivars

Quantitative traits	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
FrSh	-0.13	-0.05	0.05	-0.19	0.04	-0.09	-0.02	-0.56	-0.14
FtTe	-0.09	0.17	0.02	0.00	0.20	-0.04	-0.79	-0.10	-0.08
AdHe	0.05	0.06	-0.25	-0.20	-0.18	0.47	0.11	-0.37	0.04
FiPu	-0.06	0.13	-0.07	0.70	0.12	0.01	0.10	0.33	-0.02
QuFi	0.28	0.06	0.01	0.00	-0.01	0.05	-0.17	0.07	-0.80
StIn	-0.10	0.80	0.01	0.08	-0.29	0.13	0.06	0.01	-0.08
BaCa	0,60	-0.04	-0.19	0.05	0.08	0.09	-0.04	0.11	0.09
Be	0.12	0.66	-0.23	0.06	0.22	-0.06	-0.23	0.05	0.22
BeTy	0.25	0.40	0.23	-0.48	0.11	-0.11	-0.06	0.13	0.30
Si	-0.04	-0.06	0.95	-0.09	0.05	0.04	-0.02	-0.01	0.00
SiTy	-0.07	-0.02	0.95	-0.09	0.04	0.03	-0.06	0.00	0.03
GrSt	-0.22	-0.14	0.03	-0.04	0.59	0.12	-0.20	-0.04	-0.10
Sh	0.01	-0.02	-0.11	0.79	0.05	-0.09	-0.12	-0.24	0.05
ShSl	0.21	0.44	0.17	0.28	0.09	0.01	-0.49	-0.07	0.18
SkCoRiFr	0.01	0.29	0.11	0.17	0.27	-0.22	0.34	-0.07	-0.42
InPo	0.03	-0.04	0.04	-0.23	-0.06	-0.07	0.05	0.70	-0.17
InSh	0.26	-0.65	0.17	0.17	-0.08	0.14	0.03	0.07	0.39
InCo	0.51	0.11	-0.01	-0.22	0.15	-0.18	0.16	-0.31	0.02
InHa	-0.73	-0.13	-0.21	-0.14	0.09	-0.07	-0.13	-0.05	0.12
FlTy	0.24	0.13	0.13	0.10	0.01	0.89	-0.10	-0.03	0.02
LeSh	0.47	0.11	-0.21	-0.19	0.25	-0.02	-0.22	0.28	0.10
LeCo	-0.33	-0.09	-0.05	-0.16	0.25	0.68	0.17	0.17	-0.09
LeTe	0.00	0.04	0.01	0.05	0.72	-0.13	0.00	0.11	0.13
LeTi	-0.74	0.06	0.07	0.10	0.14	-0.02	0.11	-0.02	0.09
LeMa	0.23	-0.04	0.07	0.11	0.67	0.08	0.12	-0.17	-0.06
TrNu	-0.03	0.25	-0.12	0.01	0.25	0.03	0.52	-0.10	0.17
TrHa	0.65	-0.33	-0.02	-0.08	0.07	-0.07	-0.01	0.08	-0.08
Eigenvalue	3.43	2.82	2.44	1.77	1.59	1.53	1.37	1.27	1.14
Variability (%)	12.70	10.43	9.05	6.54	5.89	5.68	5.06	4.70	4.21
Cumulative	12.70	23.13	32.17	38.71	44.61	50.28	55.35	60.04	64.26

For full names of traits see Morpho-physiological analysis of traits head in material and method section.

type belong to fruit descriptor that explained the 9.05% of the total variation. PC4 explained 6.54% of the total variation with highest positive correlations to fiber in pulp and shoulder. PC5 up to PC9 explained 5.89%, 5.68%, 5.06%, 4.70%, and 4.21% of the total variation, respectively. However; PC5 represents mainly leaf texture, leaf margin and groove stone; PC6 explains flower type, leaf colour and adherence; PC7 describes tree number and skin colour ripe fruit; PC8 illustrates inflorescence position and fiber in pulp; PC9 most demonstrates beak type, inflorescence shape and quantity of fiber (Table S5). Figure 2B represents PC1 and PC2 plotted on a bi dimensional plane. In contrast to quantitative traits, qualitative traits indicated clear cut differences between individuals and forms a spectrum of phenotypes, which means that they are highly heritable and the environment has very little influence on the phenotype of these traits. Hence, PCA-biplot exhibited that the cultivars scattering in all the quarters, and the association between traits and cultivars for qualitative traits was more discriminator than quantitative traits. The negative values for PC1 indicate cultivars with high content of leaf tip, inflorescence hairiness, fruit shape, and groove stone as well as lower leaf colour, inflorescence shape, fruit adherence, and fruit sinus. Cultivars were positioned in the green line rectangle belong this group. These could be target characteristics depending on different purposes in breeding. The highest positive values for PC1 illustrate cultivars with high tree habit growth, inflorescence position and colour, quantity fiber of fruit, leaf shape, and basal cavity of fruit, which were positioned in the green line parallelogram. In contrast, the highest positive values for PC2 show cultivars with high fruit traits, as shown in figure 2B.

Dendrogram using agglomerative hierarchical clustering

The dendrogram showed three main groups; the C1 is included of 31 cultivars, the C3 contained of 54 cultivars, and finally the C2 comprised of 4 cultivars (Fig. 3). Clustering analysis resulted in the high level of morphological and biochemical variation among cultivars, as confirmed by descriptive and variance analysis. It was considerable that there was no relationship between clustering pattern and geographical distribution. One of its reason can be related to synonyms, homonyms and misnames. Indeed at 89.66% dissimilarity cultivars, which were placed in C1, were separated from others by their small fiber length (centroid= 5.15), lowest TSS (centroid=10.64), lowest

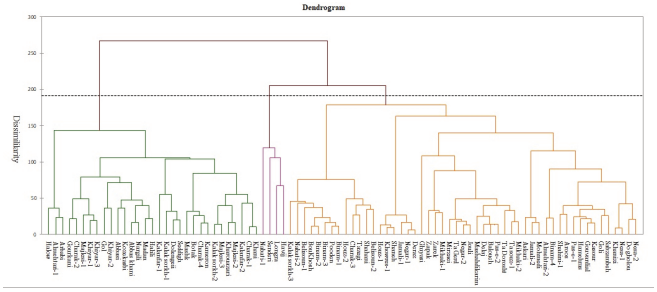

Fig. 3 - Dendrogram using AHC for 86 mango cultivars based on 22 quantitative and 27 qualitative traits.

leaf ratio (centroid= 4.23), their absence groove stone (centroid= 0) and also by their highest stone weight (centroid= 32.44). At 67.66% dissimilarity the other cultivars were divided into two groups: the cultivars, which were located in C2, characterized by their highest fruit dry matter (centroid= 17.93), larger trunk circumference (centroid= 1.88), presenting groove stone in fruit (centroid= 0.020), and having poor hairiness in flower inflorescence (centroid= 2.33), and the cultivars (C3) having highest fruit length (centroid= 9.39), fruit weight (centroid= 162.69), and also having highest value in the most of traits.

Factor analysis

The all dataset of both quantitative and qualitative variables was subjected to KMO (Kaiser-Meyer-Olkin) and Bartlett's test of sphericity. KMO (0.37) less than 0.5 and no significant Bartlett's test of sphericity showed that the all data were not very suitable for FA. KMO was found to be 0.58 for all dataset and the Bartlett's test of sphericity was significant (Chi-square = 342.459; df = 171, $p < 0.05$), when the dataset of two variables was separately used for FA. Based on FA, 22 quantitative traits were divided in eight factors (Fs) that had eigenvalue more than one and explained 70.17% of variance (Table 4). Percent of variance of each factor is showing importance of that factor. Factor loading more than 0.50 was considered as significant factor loading for each factor. F1 was included for fruit length, fruit breadth, fiber length, and stone length and it could explain 15.11% of total variation in the dependent structure for, and its suggested name is fruit yield. F2 accounted for 12.45% of total variability and was consisted of high total soluble solid, fruit dry matter as well as low fruit moisture which was named fruit taste. F3 justified 11.06% of total variability which is strongly influ-

enced by higher stone length and fiber length as well as lower petiole and leaf length, and it was suggested name fruit and weak leaf morphology. F4 accounted for 8.30% of variability and was mainly explained by moisture and dry matter which was named fruit dry matter. F5 was named leaf morphology and had high leaf width as well as low petiole ratio which contributed to 6.62% of total variation. F6 was introduced as fruit flesh and had high fruit flesh content contributed 6.44% of total variation. F7 and F8 had high contributing loading from titratable acidity which was named fruit titratable acidity.

It was noticed that the grate variation based on eight factors was observed among cultivars, but it could not produce very suitable or distinct grouping due to generate overlapping or same factors such as titratable acidity and leaf morphology in FA. This aspect has obviously been shown by FA-biplot based on first Fs extracted from quantitative variables in figure 4A. FA-biplot indicated that cultivars scattered in all the quarters and thus cannot generated proper grouping. To overcome this, quartimax rotation is a statistical technique used at one level of factor analysis as an attempt to clarify the relationship among factors. In Table 4, it can be seen that the quartimax as an oblique rotation displayed a higher efficacy to create clear patterns of results in FA where the factors are indeed correlated. However, in factor analysis by means of quartimax rotation and on the base of factor loading larger than 0.5, five rotation factors (RF) were identified and they all together justify 53.46% of existent variations among the traits. RF1 up to RF5 were called fruit size factor, dry matter factor, leaf size factor, fruit weight factor, petiole size and also were accounted for 12.62, 11.11, 10.11, 10.68, and 8.92% respectively. Result indicated that quartimax rotation minimized the complexity of the factor loadings to make the structure simpler to interpret, as shown in figure 4B. According to RFA-biplot cultivars were quantitatively separated to six distinct groups in comparison to FA-biplot. The cultivars 22 (‘Almehtari-2’) and 48 (‘Binam-4’) were separated into distinct group near the positive ends of the RF2 axis, and were correlated with respect to higher

Table 4 - Correlation coefficients explained by the first eight factor analysis (Fs) and quartimax rotation for 22 quantitative traits in 86 mango cultivars

Quantitative variables	Factors without rotation								Factors after Quartimax rotation				
	F1	F2	F3	F4	F5	F6	F7	F8	RF1	RF2	RF3	RF4	RF5
FrLe	**0.699**	0.074	0.258	0.039	0.057	0.059	0.028	0.039	0.639	-0.047	0.203	0.326	0.048
FrBr	**0.563**	-0.276	-0.14	0.44	-0.149	0.018	0.084	-0.15	0.099	-0.119	0.092	**0.771**	0.055
FrSkTh	0.059	0.153	-0.046	-0.178	-0.007	0.382	0.013	0.088	0.079	-0.01	0.100	-0.141	0.158
FrWe	**0.695**	-0.357	0.03	0.485	-0.148	0.100	0.122	-0.087	0.275	-0.143	0.022	**0.880**	-0.026
FlCo	0.328	0.211	-0.03	0.443	-0.276	**0.705**	-0.224	-0.084	0.102	0.348	0.080	0.495	0.205
FiLe	**0.669**	0.150	**0.615**	-0.17	0.019	-0.121	-0.088	0.144	**0.928**	0.005	0.007	0.094	0.023
StLe	**0.669**	0.150	**0.615**	-0.17	0.019	-0.121	-0.088	0.144	**0.928**	0.005	0.007	0.094	0.023
StWe	0.404	-0.492	0.062	0.065	0.042	-0.518	0.269	0.061	0.200	-0.417	-0.066	0.415	-0.144
StFiLe	0.128	0.035	-0.001	-0.06	-0.058	0.014	-0.027	-0.151	0.100	-0.029	0.024	0.028	0.110
pH	-0.347	-0.064	0.11	0.133	-0.216	0.100	0.189	0.139	-0.207	0.182	-0.346	-0.028	-0.056
TSS	0.087	**0.574**	0.362	-0.33	0.100	0.234	0.156	0.267	**0.500**	0.296	0.146	-0.467	0.123
TiAc	0.140	0.366	0.035	0.50	0.219	-0.311	**-0.561**	**-0.567**	0.182	0.226	0.339	-0.079	-0.024
ToSu	-0.155	0.405	0.18	0.025	0.009	0.060	-0.103	-0.006	0.092	0.394	0.035	-0.235	0.010
MoSt	0.302	**-0.635**	-0.251	**-0.539**	0.150	0.250	-0.162	0.023	0.051	**-0.921**	0.028	-0.006	0.055
DrMa	-0.290	**0.638**	0.262	**0.552**	-0.143	-0.267	0.170	0.049	-0.038	**0.929**	-0.023	0.018	-0.065
LeLe	0.438	0.473	**-0.569**	-0.045	0.312	0.011	0.338	-0.044	0.008	0.001	**0.863**	0.035	0.312
LeWi	0.252	0.215	-0.500	0.331	**0.631**	0.010	-0.004	0.286	-0.187	0.024	**0.867**	0.203	-0.220
LeRa	0.260	0.327	-0.119	-0.447	-0.409	0.023	0.479	-0.447	0.234	-0.028	0.018	-0.174	**0.687**
PeLe	0.407	0.397	**-0.702**	-0.122	-0.270	-0.205	-0.066	0.221	-0.133	-0.019	**0.509**	0.158	**0.774**
PeRa	0.160	0.088	-0.389	-0.107	**-0.633**	-0.257	-0.410	0.362	-0.159	-0.025	-0.112	0.195	**0.723**
TrCi	0.119	0.381	-0.204	-0.138	0.026	-0.042	-0.019	-0.199	0.042	0.114	0.327	-0.147	0.282
CaDi	-0.156	-0.083	-0.009	0.057	-0.142	-0.108	0.187	-0.050	-0.145	0.033	-0.179	0.025	0.013
Eigenvalue	3.325	2.741	2.434	1.826	1.457	1.417	1.199	1.040	>1	>1	>1	>1	>1
Variability (%)	15.112	12.459	11.063	8.301	6.62	6.411	5.449	4.728	12.624	11.119	10.117	10.686	8.923
Cumulative	15.112	27.571	38.633	46.934	53.554	59.996	65.445	70.173	12.624	23.743	33.86	44.546	53.469

For full names of traits see Morpho-physiological analysis of traits head in material and method section.

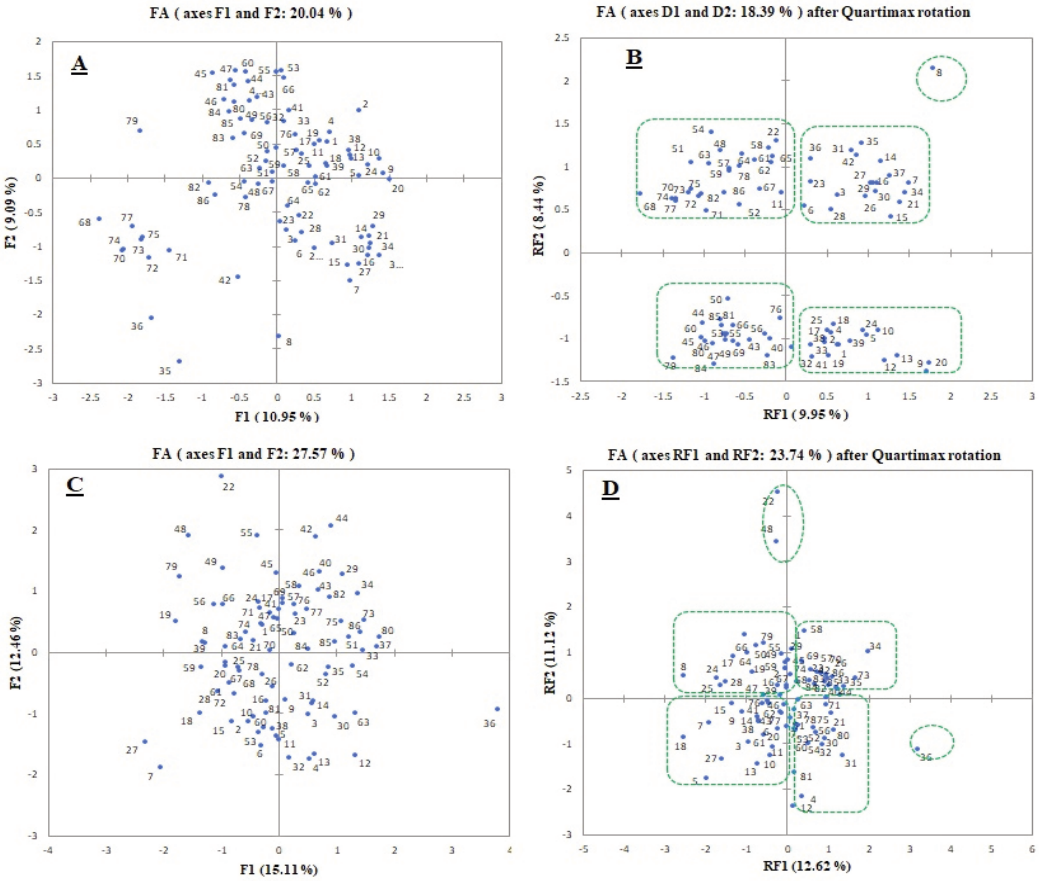

Fig. 4 - Factor analysis-biplot of the 22 quantitative (A) and 27 qualitative traits (C) with regard to the first two factors and with regard to the first two factors after quartimax rotation (RF) (B, D) among 86 mango cultivars, respectively.

fruit dry matter as well as lower fruit moisture which positioned in green line oval. Selection of these cultivars is a further advantage, as they have important postharvest characters for mango fruit breeding to extend their shelf-life. In contract, the 'Senderi' as control in green line circle was clearly diverged from the other cultivars that were located at the positive ends of RF1 axis, and was highly correlated with respect to fruit length, flesh content and stone length and the 'Kalanfar-2' was the most closet cultivar to control, as shown in figure 4B. Therefore, crosses between 'Kalanfar-2' and other cultivars with which 'Kalanfar-2' produced fertile F1 progeny would potentially allow the production of segregating populations for QTL analysis.

Additionally, FA was applied to investigate variation in qualitative variables. Based on FA, 27 qualitative traits were divided in six factors that had eigenvalue more than one and explained 41.39% of total variation (Table 5). Primary result of FA clearly proved that variation in quantitative variables is more complexity than qualitative variable among cultivars and thus reduced the number of dimensions and discovered simple patterns in the pattern of relationships among the variables in comparison to PCA. Also, FA confirmed that the variation in qualitative variables less than quantitative variables among cultivars while PCA wasn't able to detect it. The six factors of qualitative variables namely, (F1) fruit and leaf morphology, (F2 and 3) fruit sinus, (F4) flower type, (F5) fruit stalk insertion, and fruit flesh texture were accounted for 10.95, 9.08, 7.98, 5.13, 4.47, and 3.755% of total variation, respectively. Like FA for quantitative variables, same factors observed to separate cultivars base on the qualitative variables and thus it could not generate distinct grouping, as shown in figure 4C. FA-biplot indicated that cultivars scattered in all the quarters and thus cannot generated proper grouping. Hence, to clarify the relationship among factors, RFA was used. Base on the factor loading larger than 0.5, five RFs were identified and they all together justify 37.59% of existent variations (Table 5). RF1, called as fruit, flower, and growth morphology factor, which accounted for 9.94% of the

Table 5 - Correlation coefficients explained by the first six factor analysis (Fs) and quartimax rotation for 27 qualitative traits in 86 mango cultivars

Qualitative variables	Factors without rotation						Factors after Quartimax rotation				
	F1	F2	F3	F4	F5	F6	RF1	RF2	RF3	RF4	RF5
FrsH	-0.186	0.023	0.150	-0.112	0.138	-0.051	-0.182	0.120	-0.119	-0.088	0.149
FtTe	0.410	0.024	0.469	0.269	-0.075	**0.693**	0.088	0.224	**0.628**	0.111	0.190
AdHe	-0.175	0.136	-0.315	0.092	0.347	-0.017	-0.068	-0.195	-0.372	0.268	0.113
FiPu	0.136	0.353	0.115	0.245	-0.370	-0.155	-0.178	-0.225	**0.501**	0.049	-0.063
QuFi	0.148	-0.146	0.017	-0.070	0.076	-0.035	0.203	0.108	-0.013	-0.027	0.045
StIn	0.159	0.369	0.170	-0.008	**0.504**	-0.018	-0.040	-0.078	0.027	0.135	**0.646**
BaCa	0.482	-0.105	-0.339	0.051	-0.088	-0.080	**0.551**	-0.208	0.100	0.111	-0.137
Be	**0.734**	0.453	0.185	0.052	0.199	0.033	0.348	-0.253	**0.529**	0.099	0.625
BeTy	0.491	-0.272	0.101	-0.191	0.328	0.018	**0.564**	0.250	0.014	-0.042	0.315
Si	-0.141	**-0.780**	**0.603**	0.115	0.085	-0.146	-0.013	**0.966**	0.072	-0.005	-0.153
SiTy	-0.126	**-0.725**	**0.618**	0.122	0.096	-0.101	-0.031	**0.938**	0.100	0.001	-0.110
GrSt	-0.008	0.009	0.232	0.116	-0.144	-0.092	-0.137	0.116	0.245	-0.007	-0.034
Sh	0.122	0.429	0.154	0.269	-0.433	-0.110	-0.250	-0.294	**0.575**	0.036	-0.063
ShSl	**0.623**	0.085	0.306	0.247	-0.002	0.126	0.311	0.068	**0.620**	0.164	0.268
SkCoRiFr	0.101	0.174	0.306	-0.194	0.029	-0.374	-0.064	0.048	0.181	-0.223	0.310
InPo	0.037	-0.266	-0.121	-0.085	-0.030	0.114	0.208	0.098	-0.133	-0.050	-0.173
InSh	-0.129	-0.434	-0.382	0.279	-0.475	0.018	0.117	0.009	-0.053	0.514	**-0.815**
InCo	0.392	-0.119	-0.070	-0.257	0.111	-0.220	0.468	0.000	-0.015	-0.144	0.160
InHa	**-0.507**	0.278	0.188	0.015	-0.043	0.181	**-0.619**	-0.028	-0.065	-0.071	0.048
FlTy	0.114	-0.111	-0.298	0.853	0.443	-0.209	0.124	0.021	-0.025	**0.962**	-0.057
LeSh	**0.633**	-0.130	-0.207	-0.001	-0.037	0.016	**0.654**	-0.119	0.203	0.062	-0.007
LeCo	-0.396	0.045	-0.130	0.349	0.185	-0.149	-0.351	-0.006	-0.228	0.377	-0.097
LeTe	0.339	0.078	0.292	0.014	-0.272	-0.310	0.103	0.031	**0.502**	-0.141	0.054
LeTi	-0.515	0.345	0.389	0.082	-0.027	-0.095	**-0.743**	0.058	0.079	-0.056	0.155
LeMa	0.293	0.007	0.163	0.060	-0.174	-0.377	0.146	0.029	0.357	-0.038	0.010
TrNu	0.030	0.269	0.022	-0.091	0.089	-0.270	-0.073	-0.168	0.018	-0.053	0.238
TrHa	0.361	-0.400	-0.322	-0.082	-0.185	-0.049	**0.592**	-0.001	-0.033	-0.042	-0.345
Eigenvalue	2.957	2.453	2.156	1.387	1.209	1.014	>1	>1	>1	>1	>1
Variability (%)	10.952	9.086	7.985	5.139	4.474	3.755	9.947	8.443	7.57	5.1	6.478
Cumulative	10.952	20.038	28.023	33.161	37.64	41.394	9.947	18.39	25.961	31.12	37.59

For full names of traits see Morpho-physiological analysis of traits head in Materials and Methods section.

total variation. RF2 with the 8.44% of the variance reflected the fruit sinus dimension and therefore, being classified as fruit sinus. RF3 with the 7.57% of the total variance showed the high flesh texture, fiber in pulp, beak, shoulder, and leaf texture dimension and therefore, being classified as fruit morphology. RF4 with 5.16% of the variance classified as flower type factor due to high factor loading for flower type. Finally, RF5 demonstrated the stalk insertion and inflorescence shape factor which accounted for 6.47% of the variance. However, the quartimax rotation used in the factor analysis had excellent discrimination among cultivars based on factors criterion compared to FA without quartimax rotation. This aspect was confirmed by RFA-biplot where showed the cultivars separated into five distinct groups and thus offered the excellent grouping comparison with FA-biplot (Fig. 4D). The 'Nabati-1' (8) cultivar was qualitatively separated into single group near the positive ends of the RF2 and RF1axes, which positioned in green line circle, and was correlated with respect to higher fruit, flower, and growth morphology and fruit sinus classified-factors. So, the 'Nabati-1' cultivar can be considered as favorable genetic material for mango breeding via effective phenotypic selection of its correlated-traits (traits obtained by RF1 and RF2) and high expected genetic gain from selection for its correlated-traits can be achieved, as confirmed by cluster analysis.

Discriminant function analysis

The LDF from both quantitative and qualitative variables to classify the cultivars formed on RSM-collection population (Rudan, Siyahu and Minab) was

analyzed separately. Walk's lambda test was found to be 0.21 and 0.04 (p-value < 0.0001, alpha 0.05) for quantitative variables and qualitative variables, respectively. However, Walk's lambda test was significant for two discriminant function obtained from quantitative and qualitative variables, and thus there was low correlation between independent (measured variables) and dependent (population) variable to compute new directions (canonical variates or discriminant functions) in which the groups are best separated by LDF.

The LDF across the 22 quantitative variables originated from RSM-collection is shown in Supplementary Table 3. The first two discriminant functions were able to capture 100% of the total variance. Function 1 explains 82.37% of the total variance and function 17.62% of the total variance. Major contributors to discriminate among different cultivars in function 1 are the fruit length, fruit breadth, fruit weight, fiber length, stone length, stone weight, stone fiber length, and canopy diameter; meanwhile the leaf length, leaf ratio, petiole length, fruit weight, stone weight, TSS, fruit total sugar, fruit dry matter, and trunk circumference are major contributors in function 2. Therefore, these functions might represent the relationship of cultivars in each class (Rudan, Siyahu and Minab) with high efficiency at RSM-collection (Fig. 5A). Considering the 1st and 2nd discriminant functions, the cultivars presented at all three classes had closer genetic relationships; hence, there was a large overlap of the cultivars belonging to Minab and Rudan due to the stronger relationship between the classes. However, the results of confusion matrix for the cross-validation and Mahalanobis distance based on quantitative traits showed that the 55.81% (48 number) of total cultivars belonged to their original classes and the rest cultivars (44.19% = 38 number) were distributed among the three classes (Table S3).

Additionally, LDF across the 27 qualitative variables was performed (Table S3). The first two discriminant functions were able to capture 100% of the total variance. Function 1 explains 63.48% of the total variance and function 36.51% of the total variance. Function 1 explains 82.37% of the total variance and function 17.62% of the total variance. Major contributors to discriminate among different cultivars in function 1 are the leaf shape, tree habit. basal cavity, beak type, inflorescence hairiness, inflorescence colour, shoulder slope, and quantity of fiber; meanwhile the inflorescence shape, skin colour ripe fruit, beak fruit, stalk insertion, and flesh texture are major contributors in function 2. The cultivars presented at all three classes were qualitatively grouped well, with slight overlapping of between classes in compared grouping based on quantitative variables (Fig. 5B). However, the few cultivars distributed in among classes due to the slighter relationship

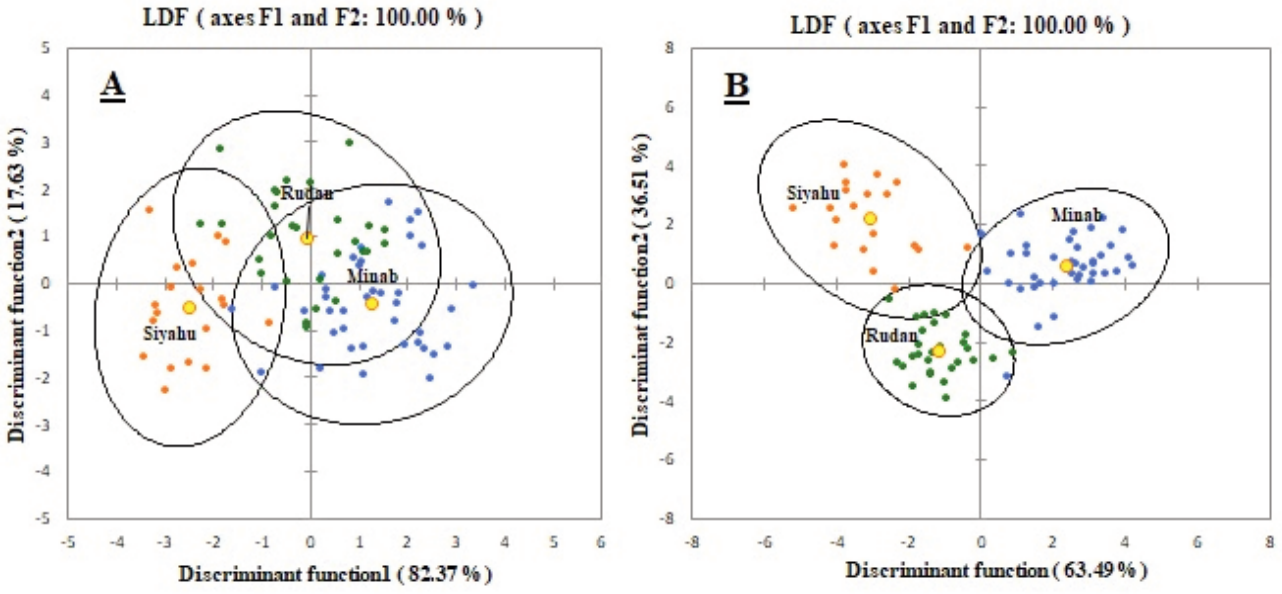

Fig. 5 - Liner discriminant functions-biplot of the 22 quantitative traits (A) and 27 qualitative traits (B) with regard to the two discriminate functions among 86 mango cultivars located at RSM-collection (Rudan, Siyahu and Minab).

with low overlap. Moreover, the results of confusion matrix for the cross-validation and Mahalanobis distance based on qualitative traits indicated that the 82.55% (71 number) of total cultivars belonged to their original classes and the rest cultivars (17.44% = 15 number) were distributed among the three classes (Table S4). Overall, result of LDF on quantitative traits for interpreting the genetic relationship among cultivars was better than LDF on qualitative traits.

4. Discussion and Conclusions

This study is the first assessment of the Iranian mango germplasm focused mainly the local cultivars. In this study, the cultivars were diverse having significant variation for morphological and biochemical traits. Variation in fruit and leaf traits is a sign of the presence of high degree of genetic variations among cultivars. Because all cultivar presented in RSM-collection have been cultivated through seed, traits segregation among them will be the reason generated genetic diversity in fruit and leaf characters. Conversely, the traits did not show variation would probably due to the same seed source propagated these cultivars.

Descriptive analysis displayed a high variation in both quantitative and qualitative traits among mango cultivars, but variability in qualitative was greater than quantitative traits. These results of this study are in the agreement with other studies, where descriptive analysis revealed a suitable genetic variability in mango germplasm (Sennhenn *et al*., 2013; Jamil *et al*., 2015), and as well as in some other crops, such as garlic (Panthee *et al*., 2006), melon (Lotti *et al*., 2008) and sour cherry (Ganopoulos *et al*., 2016). This broad genetic variability is the foundation for applied crop breeding that allows for selection of superior cultivars. Besides, information of the descriptive results of this study will be helpful to proceed plant breeding and conservation programs in the future. In the present research, generally, the descriptive results showed that 'Senderi' cultivar as control had high value of quantitative and qualitative variables compared with all cultivars. It was considerable that the value of FrLe (5.83 cm) and FrWe (232.50 gr) for 'Ta Dorosht' cultivar displayed closest or same value with controls, while the high value in TSS (19 Brix) for 'Jamali-1' and in ToSu (21.92%) for 'Shahani' was observed compared to controls.

Quantitative traits are the most significant traits of the majority of plants that are mainly influenced by the environment and hence have low heritability. Because of the response to direct selection for these traits may be unpredictable, therefore, plant breeders prefer to select for related traits that indirectly increase quantitative target traits. It was obvious that most significant Pearson's correlations were obtained by phenotypic traits particularly for the fruit yield and quality. In the breeding materials, the higher variation is the greater scope for its improvement through selection. For instance, cross combinations could be performed between cultivars with very large fruit length ('Senderi', 'Khiyar-1', 'Havij' and 'Mashk'), low stone weight ('Kozekasbi', 'Shahani' and 'Zapak') and high total soluble solids ('Jamali-1', 'Houz-1' and 'Kalak sorkh-3'). In order to achieve high yield and superior quality cultivars, output information of the morphological investigations might be quite helpful for any future mango breeding in RSM-collection. The 'Jamali-1', 'Houz-1', 'Kalak sorkh-3' and 'Shahani' will be more suitable cultivar for the fresh consumption cultivars because it showed the favorable level of soluble solids and titratable acidity. In open canopy trees light could penetrate well into the canopy causing increase of photosynthesis rate and transfer of carbohydrates from the leaves to the fruits. Similarly, Farrokhi *et al*. (2013) also showed the importance the correlation of suitable canopy volume and transferred carbohydrate to increase fruit weight in apple. This aspect will be very good option in selecting the candidate mango cultivar with suitable canopy volume and fruit quality traits, such as 'Ta Dorosht', for providing a source of material breeding. TSS is an important biochemical factor that its level is increased simultaneously with fruit development (Zarbakhsh *et al*., 2020). Pearson's correlations revealed that the cultivars with higher fruit weight tend to reveal relatively higher soluble solids. This is in agreement with the previous reports in apple (Farrokhi *et al*., 2013) and in sour cherry (Ganopoulos *et al*., 2016). It was considerable that no correlation was observed between soluble solids and titratable acidity, which has been confirmed previously in apricot (Ruiz and Egea, 2008).

In breeding programs, PCA-biplot is an important tool to identify and rank the superior cultivars and thus facilitating the mango selection process (Maia *et al*., 2016). PCA-biplot for quantitative variables showed that the highest PC1 values corresponded to 63 ('Ta Dorosht') with high fruit weight, fruit length and fruit breadth which was closeted to 36 ('Senderi').

Germplasm collection and conservation are important not only to preserve genetic resources, but also to enable breeders to exploit the genetic and phenotypic variation and develop superior cultivar. Therefore, the 'Ta Dorosht' could be introducing as superior cultivar and may be used as parent in backcrossing method with common cultivars. Moreover, regarding the 27 ('Kalanfar-1') and 7 ('Majlesi-1'), it showed lower values in all variables analyzed particularly for the canopy diameter and pH values, as shown in figure 2A (were positioned in the green line triangle). It seems that above-mentioned cultivars are exposed to highly endangered in the RMS- collection and should be consider a suitable conservation program to protect the total loss of them. Also, due to low vegetative vigor and tree size, these cultivars may be useful for the breeding program and being desirable as dwarfing rootstocks. Overall, the results of PCA for quantitative traits showed a good performance regarding fruit traits such as fruit weight, fruit length and fruit breadth, which are the most important for discriminating pomological traits, while it did not indicate distinct grouping in our studied cultivars. Additionally, our results are in accordance with the previous studies, which also documented that the weight and fruit size are useful parameters to discriminate cultivars in inter-specific almond × peach (Yaghini *et al*., 2013), and in sour cherry (Khadivi-Khub *et al*., 2013; Ganopoulos *et al*., 2016). The high negative for PC1 and PC2 for qualitative traits resulted in formatting a distinct group cultivar with strong adherence in fruit, broadly pyramidal inflorescence shape, sinus fruit, and light green leaf, which were positioned in the green line trapezoid. These findings are in agreement with various other studies that reported the maximum contribution of fruit and flower traits towards the genetic divergence in cherries (Khadivi-Khub *et al*., 2013; Ganopoulos *et al*., 2016) and mango (Maia *et al*., 2016). PCA could also permit the correlation of the phenotypic traits with the genetic linkages between the respective trait's loci in QTL mapping analysis. PCA previously has been used for germplasm evaluation in almond (Nikoumanesh *et al*., 2011), apple (Farrokhi *et al*., 2013), and mango (Krishnapillai and Wilson Wijeratnam, 2016).

In cluster analysis, the highest genetic distance was detected among C2 and C3 (25.18), followed from these among C1 and C2 (19.51), and among C1 and C3 (9.25). The such dendrogram which is able to show genetic relationship among the cultivars reported by Sennhenn *et al*. (2013) in Kenya's mango and by Krishnapillai and Wijeratnam (2016) in Sri Lanka mango. It is noticed that the 'Havij' and 'Nabati-1', as the most closely related with two control cultivars was confirmed by cluster analysis and thus, they genetically were grouped together. In this study, the most closely related pairs among the mango cultivars were 68 ('Negar-1') and 74 ('Deraz') in the C2 cluster, while the highest distance was obtained for 6 ('Halow') and 86 ('Nesa-2'). Based on the results, crossing between cultivars in distanced clusters like C1 and two clusters can provide much variation for the mango breeding purposes. Parental selection in breeding program is primarily dependent on the traits desired in the progeny and is best guided by the phenotypic expression of potential parents. In dendrogram, on one side in C2 cluster, there were cultivars 77 and 75, and on the other side in C3 cluster, cultivars 42, 35, which can be recommended for parenting future crosses that could make new generations with high variations in almost all of the measured traits. Moreover, information about the similarly or dissimilarly genetic relationship among the mango cultivars could be useful for grafting compatibility and improving rootstock, where rootstock and scion share the same genetic background.

As an outcome from FA of quantitative and qualitative traits, the great variation was found in fruit morphology among cultivars. This result is in agreement with several other studies that reported high genetic diversity in mangoes (Maia *et al*., 2016). According to Martins *et al*. (2003), the variation observed in fruit morphology is very common even at intraspecific level. The variation may be attributed to environmental factors or genetic differences or both. In our study, the variations in the fruit traits can be attributed to differences in the age of the plant, fruit maturation stage, geographical sites, climatic condition, soil properties, and seed origin. Generally, this research supported that factor analysis is a useful tool for identification of the most significant variables in the biochemical and morphological data set of mangos. Factor analysis previously has been used for germplasm evaluation in several different plant (Felenji *et al*., 2011; Pour-Aboughadareh *et al*., 2017).

Discriminant functional analysis is particularly useful in defining groups of the cultivars as prior classification criteria. Moreover, it provides a graphical output illustrating the existence of groups. Generally, LDF of the variables produced better discrimination of the mango cultivars than the principal component analysis. Discriminant functional analysis previously

has been used for germplasm evaluation in several different plant (Rafiqul Islam *et al*., 2007; Sinkovič *et al*., 2017).

The present study exhibited that the presence of exploitable genetic diversity among cultivars introduced superior cultivar in which crossing of distant cultivar with desirable traits to develop cultivar for the study area and similar agro-ecology. We disclosed that many field traits have promise for Genome Wide Association Study analysis in the future, where combining molecular marker data with morphological can recognize the genes in mango controlling the main traits evaluated here. Overall, this diversity permits the effective selection of parents in various breeding programs, referring to fruit quality and aiming at different aspects of postharvest utilization, besides high yield and resistance to diseases. Hence, the present results provide important new information for the gene pool conservation, screening superior germplasm and emphasizes the importance of conservation of genetic resources for any fruit tree breeding program.

References

AOAC, 2000 - *Official methods of analysis of the association of official analytical chemists* (16th ed.). - AOAC, Washington, DC, USA.

DONKOR E., NYADANU D., DAPAAH H., 2019 - *Germplasm collection and morphological characterization of local accessions of tigernut (*Cyperus esculentus *L.) in Ghana for conservation and utilization.* - J. Plant Breed. Crop. Sci., 11(8): 196-205.

FARROKHI J., DARVISHZADEH R., HATAMI MALEKI H., NASERI, L., 2013 - *Evaluation of Iranian native apple (*Malus x domestica *Borkh) germplasm using biochemical and morphological characteristics.* - Agric. Conspec. Sci., 78(4): 307-313.

FELENJI H., AHARIZAD S., AFSHARMANESZ G.R., AHMADIZADEH M., 2011 - *Evaluating correlation and factor analysis of morphological traits in potato cultivars in fall cultivation of Jiroft area*. - American-Eurasian J. Agric. Environ. Sci., 11(5): 679-684.

GANOPOULOS I., MOYSIADIS T., XANTHOPOULOU A., OSATHANUNKUL M., MADESIS P., ZAMBOUNIS A., AVRAMIDOU E., ARAVANOPOULOS F.A., TSAFTARISC A., SOTIROPOULOS T., CHATZICHARISIS I., KAZANTZIS K., 2016 - *Morpho-physiological diversity in the collection of sour cherry (*Prunus cerasus*) cultivars of the Fruit Genebank in Naoussa, Greece using multivariate analysis.* - Sci. Hort., 207: 225-232.

HAGIDIMITIOU M., KATSIOTIS A., MENEXES G., PONTIKIS C., LOUKAS M., 2005 - *Genetic diversity of major Greek olive cultivars using molecular (AFLPs and RAPDs) markers and morphological traits*. - J. Amer. Soc. Hort. Sci., 130(2): 211-217.

JAMIL W., AHMAD S., AHMAD M., ALI S., ABBAS M.M., 2015 - *Morpho-physiological and biochemical profiling of some mango cultivars in Pakistan*. - J. Agric. Res., 53(3): 397-412.

KAISER H.F., 1958 - *The varimax criterion for analytic rotation in factor analysis*. - Psychometrika, 3: 187-200.

KHADIVI-KHUB A., JAFARI H.R., ZAMANI Z., 2013 - *Phenotypic and genotypic variation in Iranian sour and duke cherries*. - Trees, 27: 1455-1466.

KRISHNA H., SINGH S., 2007 - *Biotechnological advances in mango (*Mangifera indica *L.) and their future implication in crop improvement - A Review*. - Biotechnol. Adv., 25: 223-243.

KRISHNAPILLAI N., WILSON WIJERATNAM R.S., 2016 - *Morphometric analysis of mango varieties in Sri Lanka*. - Aust. J Crop. Sci., 10(6): 784-792.

LAVIN A., BARRERA C., RETAMALES J.B., MAUREIRA M., 2005 - *Morphological and phenological characterization of 52 accessions of* Fragaria chilonsis *(L.) Duch*. - Hort. Sci., 40(6): 1637-1639.

LOTTI C., MARCOTRIGIANO A.R., DE GC RESTA P., RICCIARDI A., ZONNO V., FANIZZA G., RICCIARDI L., 2008 - *Univariate and multivariate analysis performed on bio-agronomical traits of* Cucumis melo *L. germplasm*. - Genet. Resour. Crop. Evol., 55: 511-522.

MAIA M.C.C., DE ARAÚJO L.B., DOS SANTOS DIAS C.T., DE OLIVEIRA L.C., VASCONCELOS L.F.L., DE CARVALHO JÚNIOR J.E.V., SIMEÃO M., BASTOS M.Y.G., 2016 - *Selection of mango rosa genotypes in a breeding population using the multivariate-biplot method.* - Cienc. Rural. Santa Maria, 46(10): 1689-1694.

MARTINS M.R., OLIVEIRA J.C., DI MAURO A.O., SILVA P.C., 2003 - *Assessment of populations of sweet passion fruit (*Passilora alata *Curtis) obtained from open pollination*. - Rev. Bras. Frutic., 25(1): 111-114.

MUKHERJEE S.K., 1989 - *Descriptors for Mango*. - Inter. Board for Plant Genetic Resources, pp. 22.

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*. - Sci. Hortic., 129: 108-118.

OMOKOLO N.D., TSALA N.G., DJOCGOUE P.F., 1996 - *Changes in carbohydrate, amino acid and phenol content in cocoa pods from three clones after infection with* Phytophthora megakarya *Bra. and Grif*. - Ann. Bot., 77: 153-158.

PANTHEE D.K.R., REGMI H., SUBEDI P., BHATTARAI S., DHAKAL J., 2006 - *Diversity analysis of garlic (*Allium sativum *L.) germplasms available in Nepal based on morphological characters*. - Genet. Resour. Crop. Evol., 53(1): 205-212.

PGRI, 2006 - *Descriptors for mango (Mangifera indica).* - Inter. Plant Genetic Res. Inst., Rome, Italy.

PIYASUNDARA J.H.N., GUNASEKARE M.T.K., WICKRAMASIGHNE I.P., 2008 - *Identification of discriminating morphological descriptors for characterization of Tea (*Camellia sinensis *L.) germplasm in Sri Lanka*. - Trop. Agric. Res., 20: 193-199.

POUR-ABOUGHADAREH A., AHMADI J., MEHRABI A.A., MOGHADDAM M., ETMINAN A., 2017 - *Evaluation of agro-morphological diversity in wild relatives of wheat collected in Iran*. - J. Agr. Sci. Tech., 19: 943-956.

RAFIQUL ISLAM M., HAMID A., KHALIQ Q.A., AHMED J.U., HAQUE M.M., KARIM M.A., 2007 - *Genetic variability in flooding tolerance of mungbean (*Vigna radiata *L. Wilczek) genotypes.* - Euphytica, 156: 247-255.

RUIZ D., EGEA J., 2008 - *Phenotypic diversity and relationships of fruit quality traits in apricot (*Prunus armeniaca *L.) germplasm.* - Euphytica, 163(1): 143-158.

SENNHENN A., PRINZ K., GEBAUER J., WHITBREAD A., JAMNADASS R., KEHLENBECK K., 2013 - *Identification of mango (*Mangifera indica *L.) landraces from Eastern and Central Kenya using a morphological and molecular approach*. - Genet. Resour. Crop. Evol., 61(1): 1-16.

SINKOVIČ L., PIPAN B., MEGLIČ V., KUNSTELJ N., NEČEMER M., ZLATIĆ E., ZNIDARCIC D., 2017 - *Genetic differentiation of Slovenian sweet potato varieties (*Ipomoea batatas*) and effect of different growing media on their agronomic and nutritional traits*. - Ital. J. Agron., 12: 350-356.

YAGHINI H., SHIRANI M., ARCHANGI A., SORKHEH K., CHALESHTORI S.B., SANGI S.E., KHODAMBASHI M., TAVAKOLI F., 2013 - *Phenotypic diversity and relationships of fruit quality traits in inter-specific almond 3 peach backcrosses breeding progenies*. - Euphytica, 194: 305-324.

ZARBAKHSH S., KAZEMZADEH-BENEH H., RASTEGAR S., 2020 - *Quality preservation of minimally processed pomegranate cv. Jahrom arils based on chitosan and organic acid edible coatings*. - J. Food Safety, 40(2): 1-12.