H

EGYPTIAN ACADEMIC JOURNAL OF
BIOLOGICAL SCIENCES
BOTANY

ISSN 2090-3812

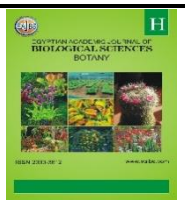
www.eajbs.com

Vol. 16 No.2 (2025)

Citation: Egypt. *Acad. J. Biolog. Sci.* (H. Botany) *Vol.16 (2) pp* 15-32 *(2025)*
DOI: 10.21608/EAJBSH.2025.451075

Egypt. Acad. Journal Biology. Sci., 16 (2):15-32 (2025)

Egyptian Academic Journal of Biological Sciences
H. Botany
ISSN 2090-3812
www.eajbsh.journals.ekb.eg.



Morphological and Microsatellite-Based Molecular Diversity in Egyptian Germplasm Collection of Peach (*Prunus persica* L.)

Atef M.H. Moatamed[1], Rania A. Mahmoud[*2], Abeer A. Dahab[3], Gehan A. Mahmoud[4] and Ehab M.B. Mahdy[5]

[1]Deciduous Fruits Research Department, Horticulture Research Institute (HRI), Agricultural Research Center (ARC), 9 Algamaa street, 12619 Giza, Egypt.
[2]Deciduous Fruits Research Department, Biotechnology Central Lab, Horticulture Research Institute (HRI), Agricultural Research Center (ARC), 9 Algamaa street, 12619 Giza, Egypt.
[3]Medicinal and Aromatic Plants Research Department, Biotechnology Central Lab, Horticulture Research Institute (HRI), Agricultural Research Center (ARC), 9 Algamaa street, 12619 Giza, Egypt.
[4]Fruit Crops Handling Research Department, Chemical Analysis Central Lab, Horticulture Research Institute (HRI), Agricultural Research Center (ARC), 9 Algamaa street, 12619 Giza, Egypt.
[5]National Gene Bank (NGB), Agricultural Research Center (ARC), 9 Algamaa street, 12619 Giza, Egypt.
*E. Mail: Atef_moatamed2000@yahoo.com ; raaniaahmed@arc.sci.eg

ARTICLE INFO
Article History
Received:27/7/2025
Accepted:4/9/2025
Available:9/9/2025

***Keywords*:**
Egyptian peach, morphological, molecular, germplasm collection, global warming.

ABSTRACT

Peaches [*Prunus persica* (L.) Batsch] are ranked third in the world as one of the most important temperate fruit tree crops from an economic standpoint. Selecting low-chill genotypes suitable for the climate of a warm winter growing region, which is expected to worsen as global warming causes reduction of accumulated winter chill, is essential for maintaining peach production and quality. Local Egyptian "Mit-Ghamr" peach is the principal cultivar grown in Dakahlia governorate, which includes many genotypes as they are usually propagated by seeds, where such genotypes which generally have a special taste and aroma vary greatly in phenological and yield traits within a single orchard. Therefore, the present study aims to assess phenotypic variability and molecular diversity of a native distinct peach germplasm collection adapted under Egyptian conditions for optimal utilization of genetic resources in confronting the negative impacts of climate change. Nine selected Egyptian peach genotypes; eight representatives of genotypes commonly known as Sultani "early, intermediate, and late maturing", Shamy, Hegazy, Mawy, Fark, and Neely, in addition to a newly selected one named "Wardy", were tested. A total of forty phenological traits were used for differentiating the nine selected Egyptian peach germplasm. For molecular characterization using microsatellite, ten specific SSRs primer pairs were analyzed for genotyping of the Egyptian peach collection. Data showed a broad sense of variation at the level of measured morphological traits combined with genetic criteria, which is useful for purposes of improvement through breeding programs, preservation, optimal utilization and management of Egyptian peach germplasm collection.

INTRODUCTION

Peaches [*Prunus persica* (L.) Batsch] are ranked third in the world as one of the most economically valuable temperate fruit tree crops, after apples and pears, with global production about 26 million tons/year (FAOSTAT, 2022). As an important agronomic trait, chilling requirement controls the floral bud break for proper flowering in peaches, which has recently gained popularity, as global warming leads to warmer winters, with insufficient chilling and early bloom, which is threatening the production of peaches (Demirel *et al*., 2023).

"Mit-Ghamr", Dakahlia governorate, is the most productive city over time. Local peach trees usually propagated sexually by seeds, so several genotypes were produced, where such genotypes vary greatly in growth habits, date of maturity, yield and fruit characteristics within a single orchard (Eliwa, 2005). These genotypes have locally common names known to farmers, where the most famous of which is Sultani. The fruits of Egyptian peach genotypes have a distinct, desirable flavor and are available in the market throughout the season, compared to early and mid-season peach cultivars (Mansour *et al*., 1999).

Morphological information, including plant descriptors, is vital for preservation, utilization and management of germplasm collections (FAO, 2006). Additionally, it makes characterization assist specialists improve their strategies and capacities to manage genebank, and facilitate genetic improvement of various species (Gotor *et al*., 2008). Since single use of traditional methods, based on morphological and phenological traits, are time-consuming, blurred by influencing environmental factors and often unable to accurately identify the individual genotype, so the present study also involved detection of genetic variability to explore Egyptian peach germplasm, where genetic aspects is an effective mean of improving crops and developing molecular markers, an efficient concise, time, and cost-saving evaluation tool, that could assist selection and breeding programs. Microsatellites or simple sequence repeats (SSRs) are considered an ideal molecular marker for assessing genetic diversity and relationships among various plant species as well as DNA fingerprinting (Hussein, 2017) because they are reliable, highly polymorphic, efficient in detecting a single locus, transportable between peach genotypes and across closely related species, and abundance with codominant inheritance (Gupta *et al*., 1996; Sosinski *et al*., 2000).

Since global warming reduces accumulated winter chill in fruit growing regions, particularly those with warm winter like Egypt, selecting low-chill cultivars is crucial for successful cultivation of deciduous fruit trees which depend on cool to break dormancy such as peaches (Farag *et al*., 2010). Parker and Abatzoglou (2019) reported that, future warming is expected to have minimal impact on the low-chill peach 'Gulfprince', especially when compared to moderate- and high-chill cultivars like 'Juneprince' and 'Elberta'. They suggested that shifting to cultivars with lower chilling requirements, such as 'Gulfcrest' or other varieties bred for warmer climates, could help mitigate or even eliminate the negative effects of reduced chill accumulation due to climate change. Therefore, the present study aims to assess phenotypic variability and molecular diversity of a native distinct peach germplasm collection adapted under Egyptian conditions for optimal utilization of genetic resources in confronting the negative effects of unfavorable changing climate.

MATERIALS AND METHODS

Prospecting and Sampling of the Studied Materials:

Nine selected Egyptian peach genotypes (*Prunus persica* L.), were collected from its recognized area; eight representatives of genotypes commonly known as Sultani "early, intermediate, and late maturing", Shamy, Hegazy, Mawy, Fark, and Neely, in addition to a

newly selected one named “Wardy”, of low chill local "Mit-Ghamr" peach as calculated by Chilling Hours Model (Weinberger, 1950), were located at Dakahlia Egyptian governorate. Samples were collected following the procedures of Bennett (1970), Harlan (1975), Marshall and Brown (1975), Hawkes (1976, 1980), Arora (1981), and Chang (1985).

Morphological Characterization:

Data was collected over the three growing seasons of 2021 to 2023. Peach descriptors (Table 2) were applied according to International Board for Plant Genetic Resources (IBPGR) and Commission of the European Communities (CEC), IBPGR, CEC (1984). Morphological terminology followed Stearn (1973). Traits were described on ten random healthy of each germplasm at a specified plant growth stage when the traits had expressed completely, and quantitative traits were measured as average.

Simple Sequence Repeat (SSR) Analysis:

Collected fresh leaf samples from the studied nine genotypes were kept frozen till extraction of total genomic DNA following the manufacturer protocol of QIAGEN DNeasy Plant Mini Kit. NanoDrop Spectrophotometry (MicroDigital, Korea) was used to determine DNA concentration. The extracted DNA was kept frozen till using in molecular analysis. For molecular characterization using microsatellite, the nine peach genotypes were analysed by ten SSRs primer pairs (Table 1). This group of SSRs was chosen considering their position in the linkage groups of the peach genome, highly polymorphic, clarity and reproducibility of their amplification patterns. These SSR primers were previously developed by different research groups; BPPCT001, BPPCT006, BPPCT008 (Dirlewanger *et al*., 2002), pchgms1, pchgms2, pchgms3 (Sosinski *et al*., 2000), PS9f8 (Joobeur *et al*., 2000), UDP98-022, UDP98-407 and UDP98-412 (Testolin *et al*., 2000)) and applied for genotyping of peach collection as described by Bouhadida *et al*. (2011).

Table 1: Characteristics of the ten SSR markers tested.

Locus	Sequence (5'-3')	Ta (°C)	Origin
BPPCT001	*F*: AAT TCC CAA AGG ATG TGT ATG AG *R*: CAG GTG AAT GAG CCA AAG C	60	*P. persica*
BPPCT006	*F*: GCT TGT GGC ATG GAA GC *R*: CCC TGT TTC TCA TAG AAC TCA CAT	58	*P. persica*
BPPCT008	*F*: ATG GTG TGT ATG GAC ATG ATG A *R*: CCT CAA CCT AAG ACA CCT TCA CT	59	*P. persica*
pchgms1	*F*: GGG TAA ATA TGC CCA TTG TGC AAT C *R*: GGA TCA TTG AAC TAC GTC AAT CCT C	57	*P. persica*
pchgms2	*F*: GTC AAT GAG TTC AGT GTC TAC ACT C *R*: AAT CAT AAC ATC ATT CAG CCA CTG C	58	*P. persica*
pchgms3	*F*: ACG GTA TGT CCG TAC ACT CTC CAT G *R*: CAA CCT GTG ATT GCT CCT ATT AAA C	58	*P. persica*
PS9f8	*F*: GGT TCT TGG TTA TTA TGA *R*: ACA TTT CTA TGC AGA GTA	48	*P. cerasus*
UDP98-022	*F*: CTA GTT GTG CAC ACT CAC GC *R*: GTC GCA GGA ACA GTA AGC CT	64	*P. persica*
UDP98-407	*F*: AGC GGC AGG CTA AAT ATC AA *R*: AAT CGC CGA TCA AAG CAA C	60	*P. persica*
UDP98-412	*F*: AGG GAA AGT TTC TGC TGC AC *R*: GCT GAA GAC GAC GAT GAT GA	57	*P. persica*

The PCR amplifications, in a final volume reaction mixture of 25 µl contained 3 µl DNA template, 1.5 µl of each primer (Willowfort), 12.5 µl Thermo Scientific DreamTaq Green PCR Master Mix and 6.5 µl nuclease free water, were performed on a T100 Thermal

Cycler (BioRad, USA). The thermal profile was programmed for one cycle of 3 min at 95°C (initial denaturation), followed by 35 repeated cycles each consisted of; 1 min at 94°C (denaturation), 45 s at the corresponding annealing temperature (Table 1), and 1 min at 72°C (extension), then final extension for 7 min at 72°C (Bouhadida *et al*., 2011), and stored at 4°C. Each amplification reaction was repeated twice. The PCR products were resolved electrophoretically on 2% agarose gel in 1x TBE (10x: 0.45 M Tris-borate, 0.01 M Na_2EDTA, pH 8.0) and ethidium bromide stain, then visualized using Gel Documentation System (UVITEC Cambridge, UK). Molecular size of the separated fragments was determined by reference to GeneDireX 100 bp DNA ladder.

Statistical Analysis:

Phenotypic data were described morphologically, while quantitative traits were calculated as mean values. Pearson correlation matrix, Kaiser-Meyer-Olkin (KMO) criterion, distance to central and centroids class, and variance decomposition were constructed using XLSTAT package "Addinsoft 2022". A heatmap was plotted according to Rokach and Maimon (2005) and Hahne *et al*. (2008) using the package of XLSTAT Version 2019.2.2 based on Euclidean distance as the similarity measure and hierarchical clustering with complete linkage.

For each SSR locus, calculations of both polymorphic information content (PIC) and heterozygosity value were done as described by Botstein *et al*. (1980) using PolyPICker online program. For all samples, DNA bands on the gel were expressed for presence as one '1' and zero '0' for absence, thus creating a binary statistical matrix. Calculating similarity matrices correlation between genotypes was done with Jaccard coefficient, then accordingly the phylogenetic dendrogram was constructed by UPGMA (unweighted pair group method) average linkage clustering through XLSTAT software.

RESULTS

Phenotypic Characterization and Variability:

A total of 40 morphological traits were used for differentiating the studied nine Egyptian peach germplasm (Table 2). Wide phenological variations are noticed among local germplasms. Twenty-two traits were polymorphic. We recorded three ecotypes of the cultivar Sultani that grows in the same orchard. Sultani was as early (E), intermediate (M), and late (L) in the traits of harvest maturity and season of flowering. Among them, morphological traits were noticed viz., ground color of mature fruit (early Sultani scores light orange), firmness of flesh (early Sultani scores medium firmness), flesh color (intermediate Sultani score white), and leaf size (early Sultani scores small). We recorded three more traits among all germplasms. The trait of flower color oscillated from white (Shamy) to rose (the rest genotypes). The seed is colored white (Shamy), pink (Hegazy), and red (the rest genotypes); and the flesh around the stone had color white (Shamy), light red (Fark and Mawy), red (intermediate Sultani, late Sultani, Neely, and Wardy), dark red (early Sultani), and dark pink (Hegazy) as shown in Table (2) and Figure (1). Some traits can be used as distinctness traits viz., unique markers, di-, or tri-polymorphism for distinguishing among Egyptian germplasm (Table 2). The coloration of the shoot tip, fruit size, and apex shape were unique traits for Shamy, leaf size was a unique trait for Mawy, and fruit shape and separation of stone were unique markers for Fark. The traits of eating quality noticed Fark (poor) and Wardy (good) and flesh juiciness Fark (juicy) and Wardy (intermediate juicy).

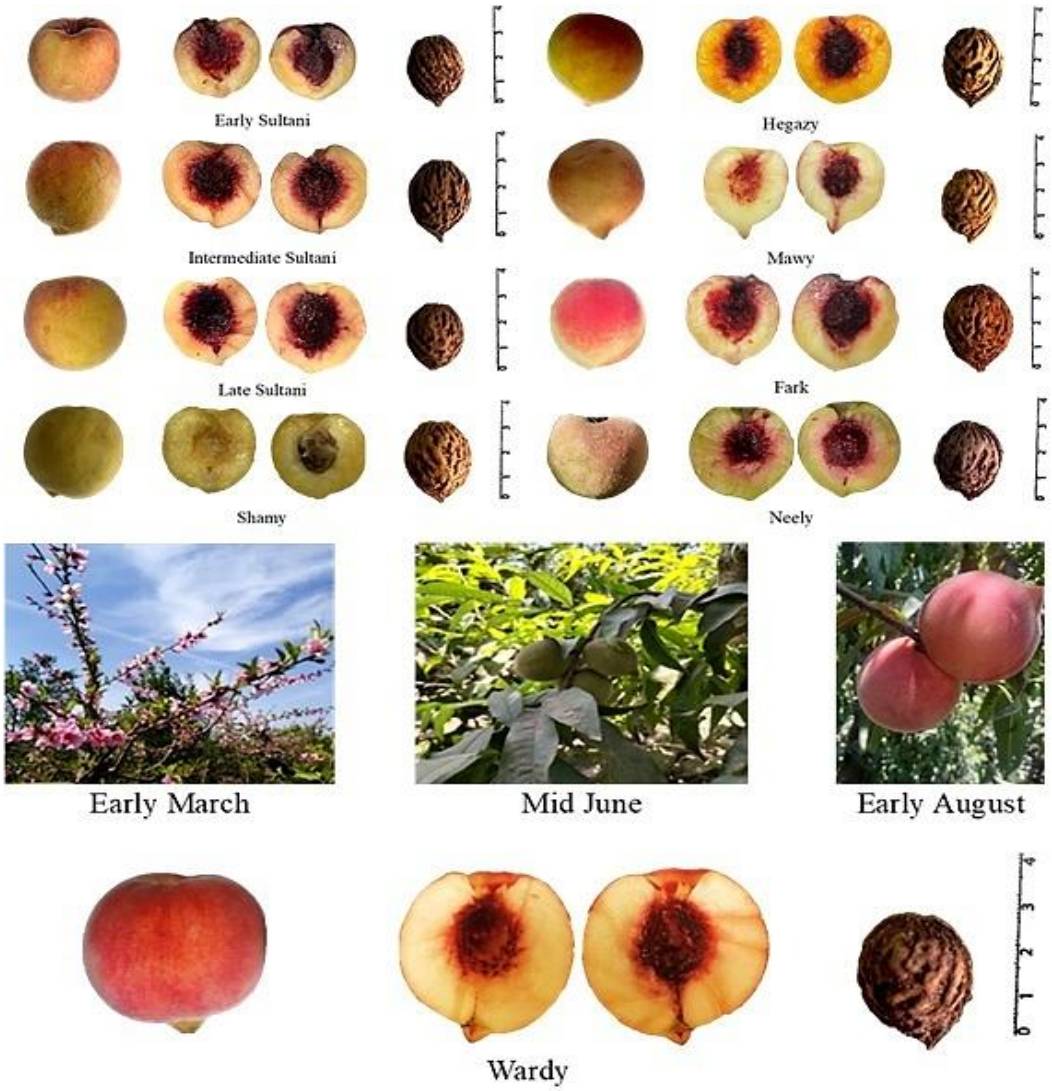


Fig. 1: Selected Egyptian peach germplasm; Sultani "early, intermediate & late maturity", Shamy, Hegazy, Mawy, Fark, and Neely, in addition to the newly introduced "Wardy".

Table 2: Index of characteristics for nine Egyptian peach germplasm.

Trait	E. Sultani	M. Sultani	L. Sultani	Shamy	Hegazy	Mawy	Fark	Neely	Wardy
Tree habit (1-Extremely upright; 3-Upright; 5-Spreading; 7-Drooping; 9-Weeping)	3	3	3	3	3	3	3	3	3
Tree vigor (3-Weak; 5-Intermediate; 7-Strong)	7	7	7	7	7	7	7	7	7
Rootstock compatibility (3-Poor; 5-Intermediate; 7-Good)	7	7	7	7	7	7	7	7	7
Coloration of shoot tip (3-Weak; 5-Medium; 7-Strong)	7	7	7	3	7	7	7	7	3
Tree chilling requirement (3-Low; 5- Medium; 7-High)	5	5	5	5	5	3	7	3	5
Leaf size	3	7	7	7	3	9	3	3	7

(1-Extremely small; 3-Small; 5-Medium; 7-Large; 9-Extremely large)									
Leaf blade shape (3-Oblate; 5-Rounded; 7-Elongated)	1	1	1	9	9	9	9	9	7
Petiole length (3-Short; 5-Medium; 7-Long)	7	7	7	3	3	5	3	5	3
Flower size (3-Small; 5-Intermediate; 7-Large)	5	5	5	7	7	5	3	3	7
Self-compatibility of flowers (1-Incompatible; 2-Special condition; 3-Compatible)	2	2	2	2	2	2	2	2	2
Bearing habit (1-On spurs; 2-On spurs and on one-year-old shoots; 3-On one-year-old shoots)	3	3	3	3	3	3	3	3	3
Cropping efficiency (1-Extremely low; 3-Low; 5-Intermediate; 7-High; 9-Extremely high)	4	4	4	4	4	4	4	4	4
Fruit size (1-Extremely small; 3-Small; 5- Intermediate; 7-Large; 9-Extremely large)	7	7	7	9	7	7	7	7	7
Fruit shape (1-Round; 2-Round/flat; 3-Elliptic; 4-Ovate; 5-Tringular/Heart-shaped; 6-Oblong/Rectangular)	1	1	1	1	1	1	4	1	1
Cavity depth (3-Shallow; 5- Intermediate; 7-Deep)	5	5	5	5	5	5	5	5	5
Suture depth (3-Shallow; 5-Intermediate; 7-Deep)	5	5	5	5	5	5	5	5	5
Symmetry along the surface (0-Predominantly asymmetric; 1-Predominantly symmetric)	1	1	1	1	1	1	1	1	1
Apex shape (1-Depressed; 2-Flat; 3-Rounded; 4-Pointed)	4	4	4	3	4	4	4	4	4
Fruit attractiveness (1-Extremely poor; 3-Poor; 5-Fair; 7-Good; 9-Excellent)	7	7	7	7	7	7	7	7	7
Skin pubescence (0-Absent or glabrous; 1-Present or pubescent)	1	1	1	1	1	1	1	1	1
Ground color (1-Green-yellow; 2-Light cream; 3-Cream; 4-Yellow; 5-Light orange; 6-Orange; 7-Dark orange)	5	4	4	2	7	6	6	6	6
Over color (0-None; 1-Trace; 2-Slight; 3/4-Mottled; 5-Intermediate red; 6/7-Mostly red; 8/9-Fully red)	5	5	5	1	5	5	5	5	7
Eating quality (1-Extremely poor; 3-Poor; 5-Fair; 7-Good; 9-Excellent)	4	4	4	4	4	4	2	4	7
Aroma (3-Little; 5-Intermediate; 7-Rich)	7	7	7	7	7	7	7	7	7
Firmness of flesh (1-Extremely soft; 3-Soft; 5-Medium; 7-Firm; 9-Extremely firm)	5	7	7	7	7	7	1	3	7
Flesh juiciness (3-Dry; 5-Intermediate; 7-Juicy)	3	3	3	3	3	3	7	3	5

Uniformity ripening of fruit (1-Not uniform; 2-Uniform)	2	2	2	2	2	2	2	2	2
Texture of flesh (1-Extremely coarse; 3-Coarse; 5-Intermediate; 7-Fine; 9-Extremely fine)	9	9	9	9	9	9	9	9	9
Skin cracking susceptibility (1-Extremely low; 3-Low; 5-Medium; 7-High; 9-Extremely high)	1	1	1	1	1	1	1	1	1
Pit-burn susceptibility (1-Extremely low; 3-Low; 5-Intermediate; 7-High)	7	7	7	7	7	7	7	7	7
Season of flowering (1-Extremely early; 2-Very early; 3-Early; 4-Early/intermediate; 5-Intermediate; 6-Intermediate/late; 7-Late; 8-Very late; 9-Extremely late)	2	2	3	6	5	5	5	8	5
Harvest maturity (1-Extremely early; 2-Very early; 3-Early; 4-Early/mid-season; 5-Mid-season; 6-Mid-season/late; 7-Late; 8-Very late; 9-Extremely late)	3	3	7	6	6	6	6	8	6
Flesh color (1-White-greenish; 2-White; 3-Light cream; 4-Cream; 5-Yellow; 6-Light orange; 7-Orange; 8-Deep orange; 9-Red)	4	2	4	2	7	4	2	2	7
Stone size (1-Extremely small; 3-Small; 5-Medium; 7-Large; 9-Extremely large)	7	7	7	7	7	7	7	7	7
Stone shape (1-Round; 2-Ovate; 3-Oblong; 4-Elliptic; 5-Elongated)	2	2	2	1	2	2	2	2	3
Stone surface (1-Smooth; 2-Pitted)	2	2	2	2	2	2	2	2	2
Separation of stone (3-Clinging; 5-Semi-clinging; 7-Free)	3	3	3	3	3	3	7	3	3
Flower color (1-White; 2-Rose)	2	2	2	1	2	2	2	2	2
Seed color (1-White; 2-Pink; 3-Light red; 4-Red)	4	4	4	1	2	4	4	4	4
Flesh color around the stone (1-White; 2-Pink; 3-Red; 4-Deep red)	4	3	3	1	3	2	2	3	3

Heatmap Visualization of Peach Germplasm with Their Traits:

The heatmaps tool is classified and explored by a set of traits that match to germplasm according to Euclidian distances via ascendant hierarchical clustering, optionally went before by the algorithm of k-means based on the matrix size. It mirrors the commuted data of the matrix where the values are swapped by corresponding color intensities. A grouped data of traits was heat-mapped to illustrate a chromatic assessment of the germplasm under study (Fig. 2).

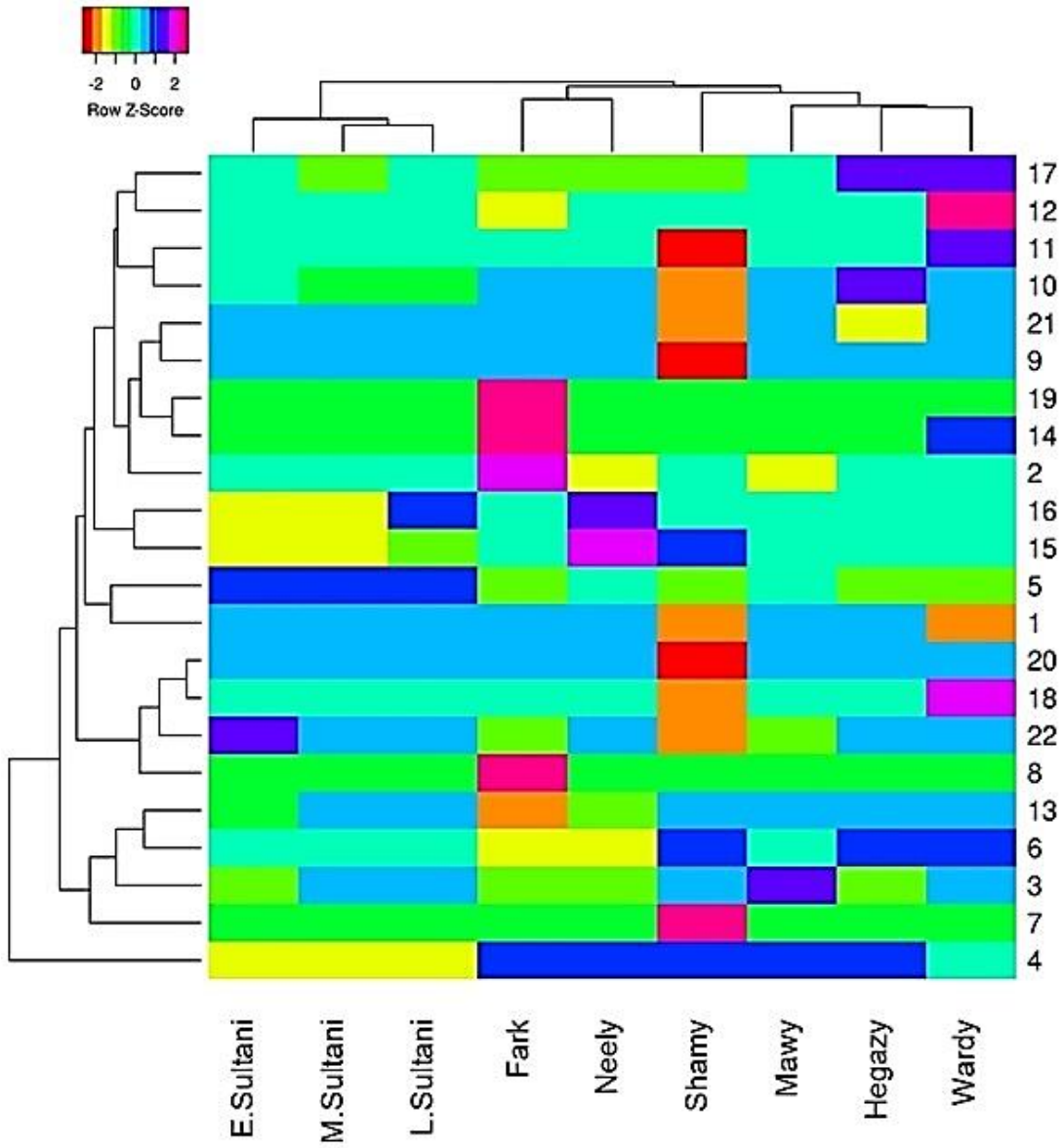


Fig. 2: Similarity matrix heatmap between 22 polymorphic traits of the nine studied genotypes; whereas 1- shoot-tip coloration, 2- tree chilling requirement, 3- leaf size, 4- blade shape, 5- petiole length, 6- flower size, 7- fruit size, 8- fruit shape, 9- apex shape, 10- ground color, 11- over color, 12- eating quality, 13- flesh firmness, 14- flesh juiciness, 15- flowering season, 16- harvest maturity, 17- flesh color, 18- stone shape, 19- separation of stone, 20- flower color, 21- seed color, 22- flesh color around stone (for more details about traits, see Table 2).

The right structure (cluster among traits) is separated into two main groups. The first one (group A) links with the trait of leaf blade shape only. The second group (group B) is split further into two sub-cluster. The first sub-group (right sub-cluster) correlates with the traits of fruit size, leaf size, flower size, and flesh firmness. The second sub-group (left sub-cluster) associates with the rest characteristics. It is worth noting that the most of traits in each cluster are quantitative traits. On the other side, the top structure (cluster among germplasm) is split into two main groups. The first group (left group) includes the germplasm Sultani only. Another group (right one) is divided into two main sub-clusters;

the first sub-cluster (left subgroup) consists of Neely and Fark only, while the rest germplasm falls into another subcluster (right subgroup).

Among all traits described, twenty-two traits were polymorphic. These traits were used for measuring more statistics, including the distance and variance decomposition, correlation, and Kaiser-Meyer-Olkin (KMO) test.

Distance and Variance Decomposition:

The distance to centroids (Tables 3 & 4), and the variance decomposition (Table 5) of germplasm classes were more important in the cluster. Peach genotypes were grouped into five classes (Table 3). Data revealed that all types of Sultani fall into a class-I with a variance of 14.3% and a distance range from 2.4 to 3.4. The class-II includes Shamy and Neely with a variance within-class of 66% and a distance of 5.7. The class-III comprises Hegazy and Wardy with a variance of 29.5% and a distance of 3.8. The class-IV and class-V consist of Mawy and Fark, respectively.

Table 3: Distance and variance of classes between the central traits and germplasm.

Class	**1**	**2**	**3**	**4**	**5**	**5**	**4**	**3**	**2**	**1**
I	0.0	12.8	12.4	10.1	14.4	13.3	9.5	8.9	8.0	0.0
II		0.0	10.5	9.2	13.5	5.7	6.1	6.3	0.0	
III			0.0	7.9	11.4	10.7	7.7	0.0		
IV				0.0	12.0	9.6	0.0			
V					0.0	0.0				
Objects	3	2	2	1	1	6	3	5	6	2
∑ weights	3	2	2	1	1	6	3	5	6	2
$V_{within\text{-}class}$	14.3	66.0	29.5	0.0	0.0	16.9	28.6	23.0	13.4	26.0
Min. distance to centroid	2.4	5.7	3.8	0.0	0.0	1.7	3.3	3.3	2.3	3.6
Av. distance to centroid	3.0	5.7	3.8	0.0	0.0	3.3	4.2	4.2	3.3	3.6
Max. distance to centroid	3.4	5.7	3.8	0.0	0.0	6.1	5.5	5.5	4.0	3.6
Weight	Sultani	Shamy, Neely	Hegazy, Wardy	Mawy	Fark	8, 17, 18, 20, 21, 22	4, 15, 16	3, 5, 6, 12, 13	2, 9, 10, 11, 14, 19	1, 7

Traits; 1- shoot-tip coloration, 2- tree chilling requirement, 3- leaf size, 4- blade shape, 5- petiole length, 6- flower size, 7- fruit size, 8- fruit shape, 9- apex shape, 10- ground color, 11- over color, 12- eating quality, 13- flesh firmness, 14- flesh juiciness, 15- flowering season, 16- harvest maturity, 17- flesh color, 18- stone shape, 19- separation of stone, 20- flower color, 21- seed color, 22- flesh color around stone.

As well, the traits were clustered into five classes (Table 3). The class-I consists of the traits of shoot-tip coloration (1) and fruit size (7) with a variance of 26%. The class-II includes the traits of tree chilling requirement (2), apex shape (9), ground color (10), over color (11), flesh juiciness (14), and separation of stone (19) with a variance of 13.4% and a mean distance 3.3. The class-III comprises leaf size (3), petiole length (5), flower size (6), eating quality (12), and flesh firmness (13) with a variance of 23% and an average distance of 4.22. The class-IV includes the blade shape (4), flowering season (15), and harvest maturity (16) with a variance of 28.67% and a mean distance of 4.27. The last class (class-V) includes fruit shape (8), flesh color (17), stone shape (18), flower color (20), seed color (21), and flesh color around stone (22) with a variance 16.9% and a distance of 3.37. The distance between the central traits and genotypes were also calculated. High value was

between class-I and class-V that is 14.42 between the central germplasm while that is 13.38 between the central traits. In contrast, low value registered between class-V and class-II (5.74) between the central traits and between class-IV and class-III (7.94) between the central germplasm.

For more statistic, distance to the class centroid of Egyptian genotypes and morphological characteristics were calculated as given in Table (4). Among the class centroids of genotypes, high value recorded between class-V and class-I (13.59), Class-V and class-IV (12.04), and class-V and class-III (11.82), while low value registered between class-IV and class-II (6.86), followed between class-IV and class-III (6.98). On the other side, the highest value scored between class-V and class-I (12.74), followed between class-V and class-IV (11.27), while that class-II and class-V (5.55) followed between class-II and class-III (5.89) recorded the lowest value among traits.

Table 4: Distance between the class centroids of germplasm and traits under study.

Class		Germplasm				
		I	**II**	**III**	**IV**	**V**
Traits	**I**		10.99	10.37	9.92	13.59
	II	7.83		7.95	6.86	10.00
	III	6.86	5.89		6.98	11.82
	IV	8.08	7.63	8.71		12.04
	V	12.74	5.55	8.61	11.27	

Table (5) shows the variance decomposition between and within classes for the optimal classification over all traits and Egyptian peach genotypes. The variance recorded 45% within the classes, and 55% between classes for the optimal classification through all genotypes, while that the variance recorded 51% within class and 49% between classes through all traits.

Table 5: Variance decomposition for optimal classification between traits and germplasm under study.

Parameter	Germplasm		Traits	
	Absolute	**Percent**	**Absolute**	**Percent**
$V_{\text{within-class}}$	19.245	45.40%	31.042	50.73%
$V_{\text{between-classes}}$	23.147	54.60%	30.153	49.27%
Total	42.392	100.00%	61.194	100.00%

Association and factor adequacy measurement

Correlation matrix and Kaiser-Meyer-Olkin (KMO) criterion were calculated as presented in Table (6). The correlation matrix directly calculated positive values between Egyptian peach germplasm except between Fark and intermediate Sultani (-0.05) and between Fark and late Sultani (-0.03). High correlation recorded between three types of Sultani, followed by Mawy with Shamy (0.74) and Wardy with Hegazy (0.7). It is notably that Mawy recorded a moderate correlation with Sultani, especially the types intermediate Sultani (0.51) and late Sultani (0.56). On this trend, Fark revealed lower correlations with others.

Table 6: Correlation matrix and Kaiser Meyer Olkin (KMO) among Egyptian peach germplasm.

Germplasm										KMO
E. Sultani	**1.00**									0.37
M. Sultani	0.84	**1.00**								0.41
L. Sultani	0.77	0.89	**1.00**							0.86
Shamy	0.10	0.32	0.39	**1.00**						0.36
Hegazy	0.34	0.22	0.34	0.63	**1.00**					0.52
Mawy	0.28	0.51	0.56	0.74	0.69	**1.00**				0.53
Fark	0.02	-0.05	-0.03	0.33	0.38	0.32	**1.00**			0.72
Neely	0.17	0.11	0.26	0.52	0.62	0.66	0.61	**1.00**		0.49
Wardy	0.22	0.30	0.40	0.64	0.70	0.67	0.09	0.35	**1.00**	0.85
KMO	0.37	0.41	0.86	0.63	0.52	0.53	0.72	0.49	0.85	0.55

The dataset subjected to correlation matrix and KMO to acquire more accurate results. Accordingly, the current study used was subjected to observe morphologically and weekly for a period of one and half year. The case study was superior to the recommended threshold, accordingly, being content with the stated standard. It achieved the KMO test to validate the dataset's suitability addressing the factor analysis and principal component analysis. KMO measures the sampling adequacy showing the degree of variability. KMO values are categorized as generally undesirable (below 0.5), sufficient (0.5-0.7), and exceptionally good (above 0.7). The obtained KMO result recorded 0.55, which refers to sufficiency. The KMO test defines how convenient data is for factor analysis and measures the sampling adequacy for each variable and the ratio of variability among variables that might vary in common. Data revealed that the highest value of KMO recorded on the germplasm of late Sultani (0.86), Wardy (0.85), and Fark (0.72). It reflects that the sampling was associated partly with in comparison to the total associations, which would be a bit of factor analysis. In contrast, the Egyptian genotypes of Mawy, Hegazy, Neely registered an average KMO value, which scored 0.53, 0.52, and 0.49, respectively. The rest local genotypes recorded the lowest value of KMO that means there are a large problem for factor analysis due to widespread correlations.

Molecular Diversity Based on SSR Analysis:

Twenty-eight alleles in total, with molecular sizes of 103 up to 222 bp, were detected for nine loci tested, representing 3.11 alleles per locus in average (Table 7). The amplification products of the nine SSRs loci were polymorphic among the nine peach genotypes, while the PS9f8 primer, originally developed from *P. cerasus*, was excluded from the analysis because it was null in Wardy and monomorphic across the other genotypes tested. Maximum total number of 6 alleles per locus was recorded by BPPCT006. The primer heterozygosity ranged from 0.19 (pchgms2, UDP98-022 and UDP98-412) to 0.76 (UDP98-407). The average polymorphic information content (PIC) of the nine SSR markers used is 0.38 with the highest values of 0.72 for UDP98-407, and 0.56 for both BPPCT001 and BPPCT006, while the lowest of 0.17 for pchgms2, UDP98-022 and UDP98-412.

Table 7: Allele variability calculated for the nine SSR markers in studied genotypes.

Locus	Allele size (bp)	No. of alleles	Primer heterozygosity	Polymorphic information content
BPPCT001	166 - 186	4	0.61	0.56
BPPCT006	103 - 134	6	0.61	0.56
BPPCT008	178 - 186	2	0.49	0.37
pchgms1	201 - 212	2	0.34	0.28
pchgms2	163 - 169	2	0.19	0.17
pchgms3	176 - 212	3	0.49	0.43
UDP98-022	164 - 171	2	0.19	0.17
UDP98-407	188 - 222	5	0.76	0.72
UDP98-412	110 - 117	2	0.19	0.17

Correlation indices of peach germplasm based on molecular criteria

According to similar indices, the dendrogram constructed based on SSR analysis (Fig. 3) separated the nine peach genotypes into two clusters. 1st cluster included intermediate Sultani in a group, and both late Sultani & Wardy in another one. 2nd cluster included two main groups; 1st group contained only early Sultani, while the 2nd group was composed of Shamy in a sub-group, May in another, and the three genotypes of Neely, Fark & Hegazy in a 3rd sub-group.

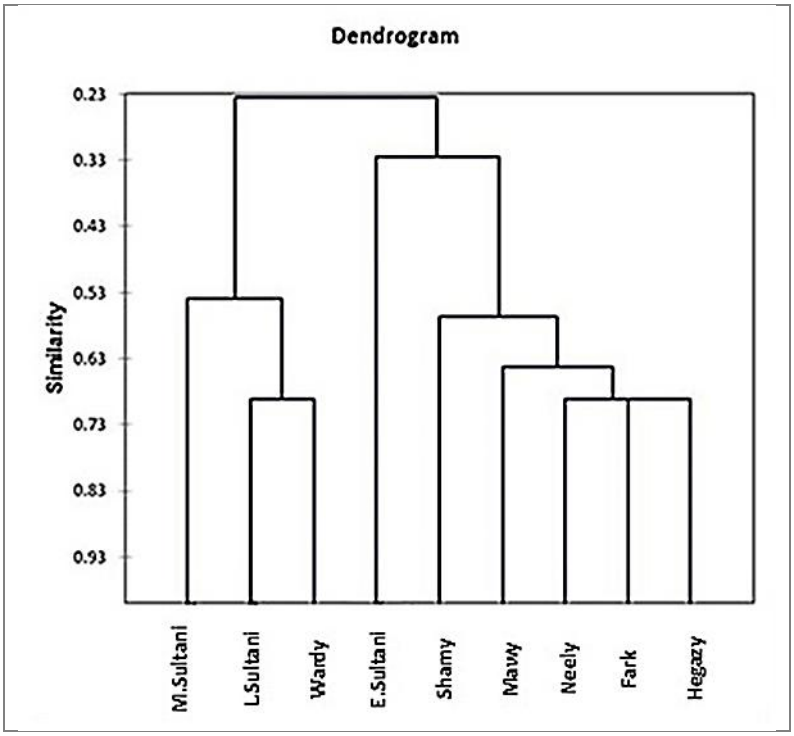


Fig. 3: Similarity phylogram generated by UPGMA using 9 microsatellite markers for 9 Egyptian peach genotypes, according to Jaccard coefficient.

Correlation Indices of Peach Germplasm According to Combined Phenotypic and Genotypic Data:

Calculated correlation matrix of nine peach germplasm (Table 8) showed a broad sense of variation at the level of measured morphological traits combined with genetic

criteria. The least similarity of 0.59 was recorded between Wardy and Mawy, while the highest scored did not exceed 0.88.

Table 8: Correlation matrix of peach genotypes according to combined phenotypic and genotypic data.

Genotype	E. Sultani	M. Sultani	L. Sultani	Shamy	Hegazy	Mawy	Fark	Neely	Wardy
E. Sultani	**1.00**								
M. Sultani	0.76	**1.00**							
L. Sultani	0.69	0.86	**1.00**						
Shamy	0.69	0.67	0.65	**1.00**					
Hegazy	0.69	0.63	0.65	0.78	**1.00**				
Mawy	0.67	0.61	0.63	0.86	0.86	**1.00**			
Fark	0.69	0.63	0.65	0.78	0.88	0.86	**1.00**		
Neely	0.78	0.71	0.69	0.88	0.88	0.86	0.88	**1.00**	
Wardy	0.65	0.76	0.88	0.61	0.61	0.59	0.61	0.65	**1.00**

DISCUSSION

Phenotypic Characterization and Variability:

Phenotypic data showed a broad sense of variation at the level of measured morphological traits, twenty-two out of forty in total were polymorphic, among the nine tested Egyptian peach genotypes. Morphological information, including plant descriptors, is vital for preservation, utilization and management of germplasm collections (FAO, 2006). Additionally, it makes characterization assist specialists improve their strategies and capacities to manage genebanks, and facilitate genetic improvement of various species (Gotor *et al*., 2008). The modernization of agricultural systems has situated plant genetic resources origin in a worrying situation (Mahdy and Ahmad, 2023). New peach varieties derived from European ones are gaining more space, to the detriment of Egyptian peach varieties, which have declined or even disappeared from the usual cultivation areas. A comprehensive valuation of morphological traits gives more engaging for more accurate information on genotyping (Mahdy *et al*., 2021). Germplasm characterization is a critical step in the authentication program, guaranteeing the reliability of breeding resources, conservation, and monitoring of the genetic makeup (El-Taher *et al*., 2023; Mahdy and Rizk, 2023). Morphological and pomological traits remain the first stage in categorizing, describing, and screening genetic resources and crop populations (Mahdy *et al*., 2021; Mahdy and Ahmad, 2023). It has been standardized and extended using plant scientists to extra precise authentication of peach genetic resources.

Heatmap Visualization:

The relationship of traits with Egyptian germplasm that shows significant variation may involve different approaches, viz., genetic improvement, genetic diversity, crop conservation, and agroecological and ecogeographical specialties. The heatmap findings reinforced the evidenced frequency of genetic diversity. Obtained results also revealed that, the traits significantly affected various variables. These outcomes are in line with those reported by Mahdy and Ahmad (2023). It is exploited and skipped direct and indirect

measurements of genetic diversity based on plant characterization and evaluation (Mahdy and Ahmad, 2023).

Distance and Variance Decomposition:

According to the observed data it can be concluded that, the tested Egyptian peach genotypes have significant variations based on their own morphological characteristics. It was also confirmed that, phenotypic heatmap analysis is a useful tool for distinguishing between the studied genotypes. It can be said that; it is major causes of variation in germplasm. The obtained results are in line with Yang *et al*. (2019) and Mahdy *et al*. (2021). The variance can be decomposed as between-classes (the variance between the means of classes) and within-class (the variance of the members of class). With this method, it can assess the extent to which between-clusters variance is larger or smaller than the variability within a cluster. In this way, an alternative can be utilized to correlate the variability between clusters to its summation with the variation within a cluster in the target object, which can thus confound the two variations in its denominator. When it is interesting to what extent the variability between classes outperforms variance within a class, or in reverse, the scientist therefore directly evaluates their relationship and straightforwardly interpret these two essential variances in terms of relative size (Muthén, 2002; Raudenbush and Bryk, 2002; Raykov, 2010). The rate of variation that discriminates between various traits is probably due to the variation in genetic resources. Conducting plant characterization and evaluation help in the first step towards the determination of existing variations that can be used to express the desired divergence. While morphological traits can be influenced by the environment, they are still widely used to study plant populations. This approach is common in crop improvement, breeding programs, studying genetic diversity, sustainable use of plant resources, and managing conservation efforts (Mahdy *et al*., 2021). Also, the richness of germplasm in a region of interest is the most used measurement of genetic diversity, monitoring of biological systems, and investigation of ecological associations. A survey of local genetic resources of a crop plays a critical role in defining many estimators aiming to collect and conserve the negative and positive bias of underdone main objects. The study of genetic diversity organizes the collection and long-term conservation of dominant and polydomain local genetic resources. Biodiversity also helps in defining the association between germplasm and better predicting the potential for time of threat in a region (Mahdy and Rizk, 2023).

Association and Factor Adequacy Measurement:

Correlation matrix assisted in evaluating relationships between local varieties shortlisted for statistical analysis. The adopted classification is defined as: (1) no relevance if $r < 0.3$, (2) less relevance if $0.3 \leq r < 0.5$, (3) median relevance if $0.5 \leq r < 0.8$, and (4) high relevance if $r \geq 0.8$ (Horn, 1965). Correlation data revealed positive and negative association values. Analogous associations were also noted in other studies (Eliwa, 2005; Anjum *et al*., 2018; Yang *et al*., 2019). The KMO test aims to determine the suitability of data for exploratory factor analysis based on the adequacy of sampling for each variable in the model and the amount of variation that might vary in common among variables. The higher the ratio, the greater the KMO value and the more adequate data are exploration factor analysis. It can be categorized as: (1) high values (KMO = 0.8-1) indicate that the obtained data may be useful through a factor analysis, (2) the values ranging from 0.5 to 0.8; a factor analysis may be moderate useful, and (3) low values (KMO = 0-0.5) refer to the factor analysis probably would not be very useful. The results revealed that, it can be utilized to determine the sampling adequacy for each germplasm. In despite of the low correlation registered by Fark with others, this variety has a moderate useful that refers to the sampling adequacy is good. Research samples are mostly not identical to the populations from which they are drawn, resulting from sampling selective samples (Fabrigar and Wegener, 2012). The difference of variables in a sample will also be restricted if the sample is more

obstructive than the population (Fabrigar *et al*., 1999). The association between qualitative observations and quantitative measurements proves this effect remarkably. It may be adequate to validate the domain restriction using sufficient data in such cases (Hunter *et al*., 2006).

Molecular Diversity:

This study using microsatellite markers has elucidated that, molecular characterization analysis for the tested SSR loci of *P. persica* origin able to distinguish between the nine selected Egyptian peach genotypes, where UDP98-407 is the most informative used primer owning the highest PIC value of 0.72 which is superior to be use as molecular marker (Zhang *et al*., 2007). In accordance, Hussein (2017) revealed that, SSR fingerprinting can provide sufficient resolving power to differentiate between peach genotypes derived from Egyptian Mit-Ghamr cultivar, despite the high similarity degree obtained in this study, and may contribute to constructing the molecular genetic database of peach germplasm. Moreover, late Sultani was proven as the most closely related to the newly introduced local Egyptian peach genotype "Wardy" as shown by SSR similarity indices, which was also confirmed by combined phenotypic and genotypic data. Undoubtedly, clustering of genotypes and its association with important agronomic traits may assist breeding programs for improvement purposes, preservation, optimal utilization and management of Egyptian peach germplasm collection (Hussein, 2017).

Declarations:

Ethical Approval: No animal model(s) or human participants were involved in the current study, so no ethical considerations are required.

Competing Interests: The authors declare no conflict of interest.

Author's Contributions: I hereby verify that all authors mentioned on the title page have made substantial contributions to the conception and design of the study, have thoroughly reviewed the manuscript, confirm the accuracy and authenticity of the data and its interpretation, and consent to its submission.

Funding: This research study was funded by STDF 'Science, Technology & Innovation Funding Authority', grant number 'PRIMA', within the framework of the project Freeclimb 'Fruit Crops Resilience to Climate Change in the Mediterranean Basin'.

Availability of Data and Materials: All datasets analysed and described during the present study are available from the corresponding author upon reasonable request.

Acknowledgements: Not applicable.

REFERENCES

Anjum, M.A., Rauf, A., Bashir, M.A., Ahmad, R. (2018). The evaluation of biodiversity in some indigenous Indian jujube (*Zizyphus mauritiana*) germplasm through physico-chemical analysis. *Acta Scientiarum Polonorum, Hortorum Cultus*, 17(4), 39-52. doi: 10.24326/asphc.2018.4.4

Arora, R.K. (1981). Plant genetic resources explorations and collection, planning and logistics. In: K.L. Mehra, R.K. Arora, S.R. Wadhi (eds). *Plant exploration and collection*. New Delhi: National Bureau of Plant Genetic Resources, p.46-54.

Bennett, E. (1970). Tactics of plant exploration. In: O.H. Frankel, E. Bennett (eds). *Genetic resources in plants - their exploration and conservation*. Oxford: Blackwell Scientific Publications, p.157-179.

Botstein, D., White, R.L., Skolnick, M., Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, 32(3), 314-331. doi: 0002-9297/80/3203-0013$01.58

Bouhadida, M., Moreno, M.Á., Gonzalo, M.J., Alonso, J.M., Gogorcena, Y. (2011). Genetic variability of introduced and local Spanish peach cultivars determined by SSR markers. *Tree Genetics & Genomes*, 7, 257-270. doi: 10.1007/s11295-010-0329-3

Chang, T.T. (1985). Germplasm enhancement and utilisation. *Iowa State Journal of Research*, 59, 399-424.

Demirel, G., Calle, A., Lawton, J.M., Atagul, O., Fu, W., Gasic, K. (2023). Ppe.CR.1 DNA test for predicting chilling requirement in peach. *Scientific Reports*, 13, 987. doi: 10.1038/s41598-023-27475-w

Dirlewanger, E., Cosson, P., Tavaud, M., Aranzana, M.J., Poizat, C., Zanetto, A. *et al*. (2002). Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). *Theoretical and Applied Genetics*, 105, 127-138. doi: 10.1007/s00122-002-0867-7

Eliwa, G.I. (2005). Approach to new peach cultivars by the aid of horticultural studies on Mit-Ghamr peach chosen strains. *Journal of Agricultural Science, Mansoura University*, 30(8), 4649-4663.

El-Taher, A.M., Elzilal, H.A., Abd El-Raouf, H.S., Mady, E., Alshallash, K.S., Alnefaie, R.M. *et al*. (2023). Characterization of some *Cichorium* Taxa grown under Mediterranean climate using morphological traits and molecular markers. *Plants*, 12, 388. doi: 10.3390/plants12020388

Fabrigar, L.R., Wegener, D.T. (2012). *Exploratory factor analysis*. New York: Oxford University Press.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods,* 4(3), 272-299.

FAO (2006). Guidelines for soil description. 4th ed. Rome: Food and Agriculture Organization of the United Nations.

FAOSTAT (2022). Food and Agriculture Organization of the United Nations. FAOSTAT database: Data [online]. Available from: https://www.fao.org/faostat/en/#data/QCL [accessed December 27, 2023].

Farag, A.A., Khalil, A.A., Hassanein, M.K. (2010). Chilling requirement for deciduous fruits under climate change in Egypt. *Research Journal of Agriculture and Biological Sciences*, 6(6), 815-822.

Gotor, E., Alercia, A., Rao, V.R., Watts, J., Caracciolo, F. (2008). The scientific information activity of Bioversity International: the descriptor lists. *Genetic Resources and Crop Evolution*, 55, 757-772. doi: 10.1007/s10722-008-9342-x

Gupta, P.K., Balyan, H.S., Sharma, P.C., Ramesh, B. (1996). Microsatellites in plants: a new class of molecular markers. *Current Science*, 70(1), 45-54.

Hahne, F., Huber, W., Gentleman, R., Falcon, S. (2008). Bioconductor case studies. New York: Springer. doi: 10.1007/978-0-387-77240-0

Harlan, J.R. (1975). Practical problems in exploration: seed crops. In: O.H. Frankel, J.G. Hawkes (eds). *Crop genetic resources for today and tomorrow*. New York: Cambridge University Press, p.111-115.

Hawkes, J.G. (1976). Manual for field collectors (seed crops). Rome: International Board for Plant Genetic Resources, FAO.

Hawkes, J.G. (1980). Crop genetic resources field collection manual. Rome: International Board for Plant Genetic Resources (IBPGR) and the European Association for Research on Plant Breeding (EUCARPIA).

Horn, J.L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185. doi: 10.1007/BF02289447

Hunter, J.E., Schmidt, F.L., Le, H. (2006). Implications of direct and indirect range restriction for meta-analysis methods and findings. *Journal of Applied Psychology,* 91(3), 594-612.

Hussein, M.A.A. (2017). Analysis of genetic diversity and genetic fingerprint of some Egyptian strains of 'Meet-Ghamr' peach using SSR markers. *Journal of Agricultural Chemistry and Biotechnology, Mansoura University*, 8(5), 149-155. doi: 10.21608/jacb.2017.38494

IBPGR, CEC (1984). Peach descriptors. Rome: International Board for Plant Genetic Resources & Brussels: Commission of the European Communities.

Joobeur, T., Periam, N., de Vicente, M.C., King, G.J., Arús, P. (2000). Development of a second generation linkage map for almond using RAPD and SSR markers. *Genome*, 43(4), 649-655. doi: 10.1139/g00-040

Mahdy, E.M.B., Ahmad, H. (2023). A study of *Corchorus* L. diversity in Egypt using high-throughput phenotyping platform (HTPP): an Egyptian gene bank example. *Genetic Resources and Crop Evolution*, 70, 2009-2019. doi: 10.1007/s10722-023-01551-6

Mahdy, E.M.B., El-Shaer, H.F.A., Sayed, A.I.H., El-Halwagi, A. (2021). Genetic diversity of local cowpea (*Vigna* spp. (L.) Walp.) accessions cultivated in some regions of Egypt. *Jordan Journal of Biological Sciences*, 14(4), 775-789. doi: 10.54319/jjbs/140419

Mahdy, E.M.B., Rizk, R.M. (2023). Genetic variation of arta populations (*Calligonum polygonoides* subsp. *comosum*) in Egypt: genepools for biodiversity and afforestation. *Journal of Water and Land Development*, 56(I-III), 81-90. doi: 10.24425/jwld.2023.143748

Mansour, N.M., Fathi, M.A., Hassan, H.Z., Shahin, B.A. (1999). Horticultural evaluation and biochemical genetic identification of some selected local peach strains at Dakahlia. *Minufiya Journal of Agricultural Research, Minufiya University*, 24(5), 1741-1763.

Marshall, D.R., Brown, A.H.D. (1975). Optimum sampling strategies in genetic conservation. In: O.H. Frankel, J.G. Hawkes (eds). *Crop genetic resources for today and tomorrow*. New York: Cambridge University Press, p.53-80.

Muthén, B.O. (2002). Beyond *SEM*: general latent variable modeling. *Behaviormetrika*, 29, 87-117. doi: 10.2333/bhmk.29.81

Parker, L.E., Abatzoglou, J.T. (2019). Warming winters reduce chill accumulation for peach production in the southeastern United States. *Climate* , *7*, 94. doi:10.3390/cli7080094

Raudenbush, S.W., Bryk, A.S. (2002). Hierarchical linear models: applications and data analysis methods. 2nd ed. California: Sage Publications.

Raykov, T. (2010). Proportion of third-level variation in multi-level studies: a note on an interval estimation procedure. *British Journal of Mathematical and Statistical Psychology*, 63(2), 417-426. doi:10.1348/000711009X468004

Rokach, L., Maimon, O. (2005). Clustering methods. In: O. Maimon, L. Rokach (eds). *Data mining and knowledge discovery handbook*. New York: Springer, p.321-352. doi: 10.1007/0-387-25465-X_15

Sosinski, B., Gannavarapu, M., Hager, L.D., Beck, L.E., King, G.J., Ryder, C.D. *et al*. (2000). Characterization of microsatellite markers in peach [*Prunus persica* (L.) Batsch]. *Theoretical and Applied Genetics*, 101, 421-428. doi: 10.1007/s001220051499

Stearn, W.T. (1973). Botanical latin: history, grammar, syntax, terminology and vocabulary. 2nd ed. Newton Abbot: David & Charles.

Testolin, R., Marrazo, T., Cipriani, G., Quarta, R., Verde, I., Dettori, M.T. *et al*. (2000). Microsatellite DNA in peach (*Prunus persica* L. Batsch) and its use in fingerprinting

and testing the genetic origin of cultivars. *Genome*, 43(3), 512-520. doi: 10.1139/g00-010

Weinberger, J.H. (1950). Chilling requirements of peach varieties. *Proceedings of the American Society for Horticultural Science*, 56, 122-128.

Yang, X., Zhang, R., Zhai, Z., Pang, Y., Jin, Z. (2019). Machine learning for cultivar classification of apricots (*Prunus armeniaca* L.) based on shape features. *Scientia Horticultrae*, 256, 108524. doi: S0304423819304078

Zhang, Z., Deng, Y., Tan, J., Hu, S., Yu, J., Xue, Q. (2007). A Genome-wide Microsatellite Polymorphism Database for the Indica and Japonica Rice. *DNA Research*, 14(1), 37-45. doi: 10.1093/dnares/dsm005