Nat. Volatiles & Essent. Oils, 2021; 8(6): 2483-2492

NVEO
Natural Volatiles &
Essential Oils

Employment Principle Components Analysis In The Estimation Of Variation And Genetic Relationship In Trees Date Palm

Ahmed Yousef Lafta Hzaa* **and Aqeel Hadi Abdulwahid Al_Amiry**

Department of Horticulture, College of Agricultural, University of Basrah, Iraq

Abstract

This study was conducted using the main components analysis based on the results of six primers for Simple Sequence Repeat (SSR) markers to estimate the relationships and genetic variability thirteen date palm (Phoenix dactylifera L.) unique cultivars planted in central and southern Iraq. The primers gave a total of 208 bands, the 76 bands of them were polymorphous as percent was 100%. The analysis of the main components PCA showed that there were four components that gave the highest percentage of participation in the variance, as their total percentages were 58.34 %. The highest value of genetic similarity between cultivars was recorded between the two cultivars (Barben and Angasa) reach 0.545, while the lowest genetic similarity value was recorded between the two cultivars (Nbaiti and Emamt-Algaty). The analysis of genetic distance tree on the two orthogonal axes showed the formation of cultivars in two groups, the first group included seven cultivars, the two cultivars (Barben and Angasa) were recorded at the nearest distance while the two cultivars (Hamrawi and Barben) have a longer distance between them. As for the second group, the cultivars were more genetically related to each other, the two cultivars (Sakri and Abd-alhadi) , while the farthest genetically related cultivars were the two cultivars (Swadani and Ashger) Thus , it is through the analysis of PCA can be distributed the samples in the aggregates form include all groups of genetically closely related cultivars, and thus it is possible to obtain an easy way to study genetic relationship and evolutionary .

Keyword: PCA , SSR , Date Palm, cultivars.

Introduction

The date palm, Phoenix dactylifera L., is one of the most important evergreen fruit trees planted in Iraq, which began to be planted in Mesopotamia about 7000 years ago, and is believed to have originated in the Arabian Gulf and southern Iraq. Monocotyle done is propagated sexually by seeds and vegetative by cuttings or by tissue culture [1].

The Principle Components Analysis (PCA) is one of the effective techniques for summarizing large data and variables into small groups, as it compresses data in light of related measurements and expresses it with a new set of unrelated variables (orthogonal), which is known as dimensional reduction, and then the new variables can be major components, and this analysis was used in many agricultural research to express the information associated with the characteristics of phenotypic and genetic trait, the reduction of these different data and large variables into small groups and turn them into new variables

, and so as to achieve the best performance to determine the contrast between the new variables with each other through the analysis of factors to determine the key components that explain Correlation pattern within a set of known variables . (Leech et al. , 1991; Liu & Schisterman, 2004)

Abd et al. 2013 conducted a study on seedlings of date palms growing in Basra Governorate in order to identify their distinctive characteristics, where 59 traits of vegetative, flowering and fruit growth were studied on 25 cultivars. The analysis of the main components PCA was used to define the differences and symmetry between the cultivars studied, and it was found from the analysis of the main components that there is a group of phenotypic characteristics of leaves and fruits that can be used to distinguish the different cultivars. Three main components appeared that express the genetic variance, as the first component constituted 92.32% of the total variance. The cumulative between the breeds, the second component scored 75.25% of the total variance, while the third component was 51.33% of the total variance.

Al-Ghazii (2015) explained in his study on ISSR markers to estimate the genetic diversity of twenty-four Iraqi date palm cultivars using six primers. The result shows for these primers gave bands deference in numbers and location on gel, the PCA analysis showed that there are three components that explained 83.34% of the total variance, as the first component represented a percentage of 37.18% of the total variance, while the second and third components recorded participation rates in the total cumulative variance amounted to (29.19%). and 16.98%) respectively.

Jaskani et al., 2016 used the SSR technique to estimate genetic diversity and determine sex in some Pakistani date palm trees. A total of 12 primers were used with 14 cultivars., where the SSR primers showed a large number totaling 124 bands and an average of 9.08 bands for each primer, and the number of bands produced by each primer ranged from (5-16) bands, and the sizes of the amplified bands were between 100 bp and 500 bp, and the primers produced 15 polymorphism, and analysis of the coordinates of the main sequence (PCoA) showed that the varieties included four independent groups with a variance of 29% of the first component and 14% of the second component.

Use Al-Faifi et al. (2016) SSR markers to assess the genetic diversity of 32 common date palm cultivars that grow in different geographical regions in the Kingdom of Saudi Arabia, as 93 primers were tested, all of which succeeded in the ability to amplify the date palm DNA fragment, the 22 primers were randomly selected to distinguish between the studied cultivars, and the results showed that the primers gaves a 91 amplified bands with an average of 4.14 for each primer. The values of the general genetic similarity of the cultivars ranged from 0.06 to 0.89, with an average of 0.41, and PCA was used to verify the genetic relationships between the cultivars which explained 40% of the variance. In addition, Hazaa (2019) estimated the genetic diversity among twenty female cultivars of uniqe date palm growing in central and southern Iraq by using six primer of SSR markers wich gave different bands reached 308 bands and an average of 51.3 bands for each primers. The study, show the percentage of total polymorphism was 100%. The results of the cluster analysis showed the distribution of cultivars in two main groups and according to their geographical location .

In view of the role of the main components analysis tools PCA in reducing the sources of variation and employing them in an easy way through which it is possible to reduce the size of the data and come up with important conclusions, and because the genetic relationships of the Iraqi date palm was little been

researched, therefore this research was conducted to show the role of these major components in reducing the size of the differences and finding the genetic relationships between the varieties studied.

Materials and working methods

For the purpose of achieving the objective of the study, thirteen female date palm cultivars were selected in several regions of southern and central Iraq during the 2020-2021 season. All agricultural operations were according to the methods used in palm orchards, and the cultivars were as homogeneous as possible in age and vegetative growth stage. The work of the molecular study was carried out in the Laboratory of Genetic Engineering and the Laboratory of biotechnologies - Department of Horticulture and Landscaping design- College of Agriculture - University of Basra.

The plant samples were collected from the fresh and white leaves nearest from heart of date palm and free from diseased and other cantamination. The leaves were washed several times with sterile distilled water, to clean them of dust and plankton, and then wiped with a medical cotton dipped in alcohol at a concentration of 70% for the purpose of sterilization, then they were cut into small pieces (1 cm2).) Using clean, sterile and sharp scissors, then placed the pieces in a ceramic mortar and Liquid Nitrogen was added, then the samples were ground well until they turned into a white powder, the powder was saved in a sample vial of 10ml.

DNA was extracted by CTAB method as reported by Doyle (1991) and Aitchitt et al. (1993). The quality and efficiency of DNA was estimated using a Nano Drop ND-2000 Spectrophotometer (THERMO SCIENTIFIC, USA) at 260 nm and the purity was checked by A260/A280 ratio.Six primers produced by BIONEER were used as noted Table (1) shows the primers, their sequences, sources, melting temperature for each primer and GC ratio.

Table 1 - Primers characteristics of Technology SSR from the BIONEER company Korea **.**

Primers	Sequence	Annealing temp (°C)	GC %	Source
PDCAT6	F: AATCAGGGAAACCACAGCCA R : GTTTAAAGCCTTCTCAAGATAGCCTCAG	53	46	Akkak et al. , (2009)
PDCAT18	F : CCTAAACCTGAATGAATCAAAGCA R : ACTAACATAAGGACAGTGCTATGTGATTG	54	38	Akkak et al. , (2009)
MPDCIR70	F : CCATTTATCATTCCCTCTCTTG R: CTTGGTAGCTGCGTTTCTTG	51.8	45	Billotte et al., (2004)
MPDCIR78	F : CCCCTCATTAGGATTCTAC R: GCACGAGAAGGCTTATAGT	49.3	47	Billotte et al., (2004)
DP157	F : TGGACAATGACACCCCTTTT R: GCCCACACAACAACAACCTCTCT	54.6	50	Elmeer et al .,2011
DP175	F: ACACACACACACACACACACC R : GTGGCTTCTTTTTGGCTGTC	57.6	51	Elmeer et al .,2011

The SSR-PCR program was used depend up on primers: one cycle of 5 minutes at 95°C for the initial denaturation of the DNA strand, followed by 35 replication cycles comprising each cycle: 30 seconds at 95°C for the template denaturation and 45 seconds at (49-57). °C) for binding primers to template DNA and 1 minute at 72°C to elongate the bound primers with one final cycle at 72°C for 7 minutes as the final elongation cycle.

Molecular data analysis for the study

the Gel Analyzer 2010a was used as digital documentation program to allocate and identify the DNA packets resulting from the polymerase chain reaction (PCR), based on the DNA Ladder 100 bp produced by Promega Company, the results were transmitted to the Excel program for statistical processing, extracting the results from them and calculating the number of bands, The same, different and unique (Fazekas et al., 2014 ; Abou-Elwafa, 2018).

The results were analyzed according to the pre-prepared tables using the biostatistical program Past software ver 3, (Hammer et al., 2001). Genetic distances were measured after creating the Euclidendistandes matrix between the varieties under study using the Jaccard coefficient. Principal coordinate analysis (PCA) was also performed (Cumming & Wooff, 2007)

Results and discussion

Table (2) shows some of the characteristics of the six primers were used, the use of these primers resulted in a total of 208 amplified bands an average of 34.6 bands for each primers, the primer PDCAT6 produced the largest number of bands reached 41, while the primer MPDCIR70 gave the lowest number of produced bands reached 21. The primers gave polymorphic bands reached 76 bands an average of 12.6 bands for each primers. The primer DP157 was the most productive, giving 17 polymorphic bands, while the lowest number of multiform band was 8 that appeared in the results of the primer MPDCIR70. As it appears from the table 2 that all primers were distinguished by their results by the presence of unique bands distinct for the variety, totaling 26 bands and an average of 4.33 The primer DP157 produced nine unique packets and each of the primers (PDCAT18, DP175) produced three unique packets, while the primers (PDCAT6, MPDCIR78 and MPDCIR70) produced unique packets (5, 4, 2) respectively.

The six primers also proved their effectiveness in giving polymorphism among the female cultivars under study, the percentage of total polymorphism reached 100%, which indicates the great genetic differences between the cultivars, and these results differ with the results of (Racchi et al., 2014; Khierallah et al. , 2011). The efficiency and diagnostic ability of SSR primers were also estimated. The primer PDCAT6 was the most efficient, with an efficiency rate of 19.71%, while the primer MPDCIR70 recorded the lowest efficiency rate of 10.1%, because the primer produced the least number of amplified bands. As for the diagnostic ability of the primers, which depends on the bands packages produced by the primers, the primer DP157 scored the highest percentage of discriminatory ability, which amounted to 22.37%, while the primer MPDCIR70 had the lowest primer discriminatory ability, which amounted to 10.52%, because it recorded the lowest number of bands with polymorphism .

Table 2- Amplification products of SSR primers from the bands, their efficiency ratios and their discriminatory ability in date palm cultivars.

primers	no total bands	no of bands	Mono morphism bands	Poly morphism bands	Unique bands	Poly morphism %	Primer Efficiency %	Primer Discrimination power %
PDCAT 6	41	15	0	15	5	100	19.71	19.74
PDCAT 18	38	11	0	11	3	100	18.27	14.47
MPDCIR 70	21	8	0	8	2	100	10.1	10.52
MPDCIR 78	35	12	0	12	4	100	16.82	15.79
DP157	38	17	0	17	9	100	18.27	22.37
DP175	35	13	0	13	3	100	16.82	17.10
total	208	76	0	76	26			
Mean	34.6	12.6	0.0	12.6	4.33	100	16.66	16.66

Principal component analysis (PCA) for SSR primers to date palm cultivars

The PCA analysis was used based on the variance matrix for the results of the six primers(Fig. 1). Several main components appeared, four of which gave the highest participation rate in the variance, with a total percentage of 58.34% of the total variance. The first component was the most, with a percentage of 22.83%, while the second component recorded The participation rate amounted to 14.38%, and the third component participated by 11.72%, while the contribution rate of the fourth component was 9.41%.

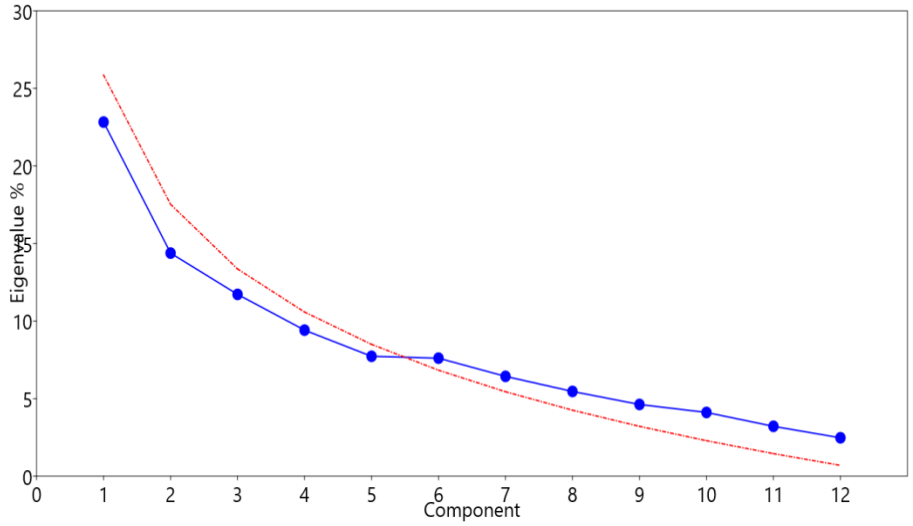

Figure 1- Imaginary roots of the cultivars under study due to the six primers

The genetic dimensions between the cultivars under study were estimated by constructing a matrix of concordance percentages resulting from the application of unbalanced pairwise group averages (UPGMA) depending on the products of primer, (Table 3).

Some cultivars recorded genetic similarity values close to one, and this indicates the conformity of the amplified bands between two cultivars, which appeared among a number of the cultivars under study. the greatest genetic similarity of 0.545 was recorded between the two cultivars (Barben and Angasa), in contrast, the lowest genetic similarity value of 0.000 was recorded between the two cultivars (Nbaiti and Emamt-Algaty), and this indicates the extent of the difference The genetic similarity values ranged between the highest and the lowest, and these results showed a high level of genetic diversity. The values of genetic similarity represent the genetic relationship between the cultivars, the more these values increase the amount of similarity of the genetic material and therefore similarity in phenotypic traits or others (Abd, 2015).

Table 3- Values of genetic dimensions for some date palm cultivars based on the results of amplification for six SSR primers**.**

	Barben	**Angasa**	**Emamt-Algaty**	**Filfila**	**Brtgala**	**Najdi**	**Hamrawi**	**Swadani**	**Sakri**	**Abd-alhadi**	**Habsi**	**Nbaiti**	**Ashger**
Barben	1	0.545	0.375	0.156	0.308	0.200	0.074	0.185	0.083	0.080	0.061	0.040	0.071
Angasa	0.545	1	0.370	0.281	0.407	0.308	0.179	0.161	0.071	0.107	0.083	0.071	0.097
Emamt-Algaty	0.375	0.370	1	0.250	0.370	0.179	0.143	0.129	0.074	0.071	0.056	0.000	0.065

Filfila	0.156	0.281	0.250	1	0.414	0.321	0.241	0.114	0.179	0.063	0.135	0.065	0.088
Brtgala	0.308	0.407	0.370	0.414	1	0.214	0.138	0.125	0.154	0.107	0.114	0.111	0.133
Najdi	0.200	0.308	0.179	0.321	0.214	1	0.261	0.103	0.083	0.038	0.129	0.040	0.154
Hamrawi	0.074	0.179	0.143	0.241	0.138	0.261	1	0.000	0.042	0.040	0.097	0.087	0.036
Swadani	0.185	0.161	0.129	0.114	0.125	0.103	0.000	1	0.217	0.261	0.233	0.037	0.032
Sakri	0.083	0.071	0.074	0.179	0.154	0.083	0.042	0.217	1	0.353	0.192	0.158	0.083
Abd-alhadi	0.080	0.107	0.071	0.063	0.107	0.038	0.040	0.261	0.353	1	0.280	0.095	0.125
Habsi	0.061	0.083	0.056	0.135	0.114	0.129	0.097	0.233	0.192	0.280	1	0.192	0.129
Nbaiti	0.040	0.071	0.000	0.065	0.111	0.040	0.087	0.037	0.158	0.095	0.192	1	0.238
Ashger	0.071	0.097	0.065	0.088	0.133	0.154	0.036	0.032	0.083	0.125	0.129	0.238	1

The minimum spanning tree for the Date palm cultivars is shown in Figure (2). the cultivars are formed in two groups on the axis of the main components, the first group in which the cultivar Barben was the closest genetically related to the Angasa cultivar than the rest of the cultivars, while the cultivar Hamrawi was the furthest related to it. The second group was the two cultivars (Sakri and Abd-alhadi) are more genetically related to each other, while it is noted that the Swadani cultivar was one of the most genetically related cultivars to the Ashger cultivar, to which the closest was the Nbaiti cultivar. Genetic correlation between cultivars.

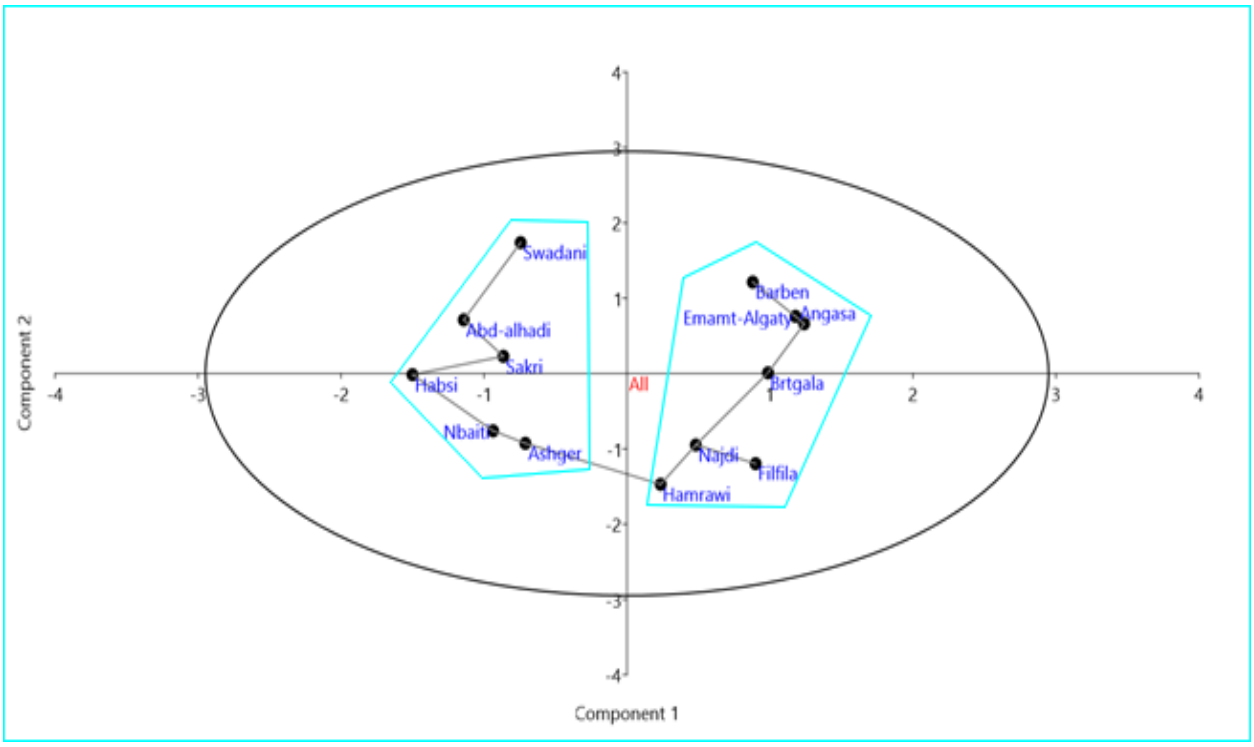

Figure 2- The minimum genetic extension tree for some date palm cultivars based on the results of amplification for six SSR primers.

Conclusion

The above results show the effectiveness of PCA in summarizing genetic data in the form of new variables to estimate genetic variances, finding genetic relationships between the cultivars under study, and contributing to understanding phylogenetic relationships.

Acknowledgments

We extend our thanks and appreciation to everyone who participated and contributed to this research paper from researchers and everyone who helped in preparing it and to the technicians in the laboratories of the College of Agriculture - University of Basra

References

[1] Abd, A-K. M.,(2015). Identification of Some Cultivars of Date Palm (Phoenix dactylifera L.)by RAPD – PCR Technique in Basrah. Basrah Journal of Agricultural Sciences,28 (1), 1-9. https://bjas.bajas.edu.iq/index.php/bjas/

Abd, A-K. M., Taha, A. & Mhoder, T. Y. (2013). A Morphological Study of Data Palm Seeds Strains (Phoenix dactylifera L.) Grown in Basra of Region Using Principal Component Analysis. Jordan Journal of Agricultural Sciences,9 (2), 259-279. https://journals.ju.edu.jo/JJAS/article/view/4297

Abou-Elwafa, S. (2018). Identification of associated genes with drought tolerance in barley. Biology Plantarum 62(2), 299-306. https://doi.org/10.1007/s10535-017-0765-0

Aitchitt, M., Ainsworth, C.C. & Thangavelu, M. (1993). A rapid and efficient method for the extraction of total DNA from mature leaves of the data palm (Phoenix dactylifera L.). Plant molecular biology reporter, 11(4), 317-319. https://doi.org/10.1007/bf02905332

Akkak, A., Scariot, V., Marinoni, D.T., Boccacci, P., Beltramo, C. & Botta, R. (2009). Development and evaluation of microsatellite markers in Phoenix dactylifera L. and their transferability to other Phoenix species. Biology Plantarum, 53(1), 164–166 https://doi.org/10.1007/s10535-009-0026-y

Al-Bakr A., (1972). The Date Palm. Its Past and Present Status. Alani Press, Baghdad:1085pp.

Al-Faifi, S.A., Migdadi, H.M., Algamdi, S.S., Khan, M.A., Ammar, M.H., Al-Obeed, R.S., Al-Thamra, M.I., El-Harty, E.H. & Jakse, J. (2016). Development, characterization and use of genomic SSR markers for assessment of genetic diversity in some Saudi date palm (Phoenix dactylifera L.) cultivars. Electronic Journal of Biotechnology, 21, 18-25. https://doi.org/10.1016/j.ejbt.2016.01.006

Al-Ghazii, N.O. (2015).The use of phenotypic and biochemical indicators and the technique of Inter Simple Sequence Repeat (ISSR) in estimating the genetic diversity of a number of Iraqi date palm cultivars, M.Sc. College of Agriculture - University of Basra. 149 pp.

Bedjaoui, H. & Benbouza, H. (2020). Assessment of phenotypic diversity of local Algerian date palm (Phoenix dactylifera L.) cultivars. Journal of the Saudi Society of Agricultural Sciences, 19(1): 65-75. https://doi.org/10.1016/j.jssas.2018.06.002

Billotte, N., Marseillac, N., Brottier, P., Noyer, J.L., Jacquemoud - Collet, J.P., Moreau, C., Couvreur, T., Chevallier, M.H., Pintaud, J.C. & Risterucci, A.M. (2004). Nuclear microsatellite markers for the date palm (Phoenix dactylifera L.): characterization and utility across the Phoenix genus and in other palm genera. Molecular Ecology Notes, 4(2): 256–258. https://doi.org/10.1111/j.1471-8286.2004.00634.x

Cumming, J. & Wooff, D.A. (2007). Dimension reduction via principal variables. Computational statistics & data analysis, 52(1): 550-565. https://doi.org/10.1016/j.csda.2007.02.012

Doyle, J. (1991). DNA protocols for plants. In Molecular techniques in taxonomy , 283-293: Springer. https://link.springer.com/chapter/10.1007/978-3-642-83962-7_18

Elmeer, K., Sarwath, H., Malek, J., Baum, M. & Hamwieh, A. (2011). New microsatellite markers for assessment of genetic diversity in date palm (Phoenix dactylifera L.). 3 Biotech, 1(2): 91-97. https://link.springer.com/article/10.1007/s13205-011-0010-z

Fazekas, M., Madar, A., Sipiczki, M., Miklós, I. & Holb, I.J. (2014). Genetic diversity in Monilinia laxa populations in stone fruit species in Hungary. World Journal of Microbiology and Biotechnology, 30(6): 1879-1892. https://link.springer.com/article/10.1007/s11274-014-1613-4

Hammer, Ø., Harper, D.A. & Ryan, P.D. (2001). PAST: paleontological statistics software package for education and data analysis. Palaeontologia electronica, 4(1): 1-9. https://palaeo-electronica.org/2001_1/past/issue1_01.htm

Hazaa, A.Y.L. (2019). Employment of Molecular Markers to Study Genetic Diversity and Gender Determination of many Date Palm Phoenix dactylifera L. cultivars. PhD thesis, College of Agriculture - University of Basra.217 pp.

Jaskani, M.J., Awan, F.S., Ahmad, S. & Khan, I.A. (2016). Development of molecular method for sex identification in date palm (Phoenix dactylifera L.) plantlets using novel sex-linked microsatellite markers. 3 Biotech, 6(1): 22. https://link.springer.com/article/10.1007/s13205-015-0321-6

Khierallah, H.S., Bader, S.M., Baum, M. & Hamwieh, A. (2011). Genetic diversity of Iraqi date palms revealed by microsatellite polymorphism. Journal of the American Society for Horticultural Science, 136(4): 282-287. https://doi.org/10.21273/JASHS.136.4.282

Leech, J., Correll, R. & Myint, A.K. (1991). Use of principal-coordinate analysis to assist in aggregating species for volume-table construction. Forest ecology and management, 40(3-4): 279-288. https://doi.org/10.1016/0378-1127(91)90046-X

Liu, A. & Schisterman, E.F. (2004). Principal component analysis. Encyclopedia of Biopharmaceutical Statistics. New York: Marcel Dekker . https://onlinelibrary.wiley.com/doi/abs/10.1002/0470013192.bsa501

Racchi, M., Bove, A., Turchi, A., Bashir, G., Battaglia, M. & Camussi, A. (2014). Genetic characterization of Libyan date palm resources by microsatellite markers. 3 Biotech, 4(1): 21-32. https://doi.org/10.1007/s13205-013-0116-6

Wrigley, G. (1995). Date-palm (Phoenix dactylifera L.). The Evolution of Crop Plants. Longman, Essex : 399-403.

Zango, O. et al., (2017). Genetic diversity of Southeastern Nigerien date palms reveals a secondary structure within Western populations. Tree Genetics & Genomes, 13(4). Available at: http://dx.doi.org/10.1007/s11295-017-1150-z.