Caryologia
International Journal of Cytology, Cytosystematics and Cytogenetics

Firenze University Press
www.fupress.com/caryologia

OPEN ACCESS

Citation: Zahra Noormohammadi, Masoud Sheidai, Seyyed-Samih Marashi, Somayeh Saboori, Neda Moradi, Samaneh Naftchi, Faezeh Rostami (2022). Identification of genetic regions associated with sex determination in date palm: A computational approach. *Caryologia* 75(3): 135-144. doi: 10.36253/caryologia-1780

Received: August 12, 2022

Accepted: November 24, 2022

Published: April 5, 2023

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

Competing Interests: The Author(s) declare(s) no conflict of interest.

ORCID
ZN: 0000-0003-3890-9001

Identification of genetic regions associated with sex determination in date palm: A computational approach

Zahra Noormohammadi[1,*], Masoud Sheidai[2], Seyyed-Samih Marashi[3], Somayeh Saboori[1], Neda Moradi[1], Samaneh Naftchi[1], Faezeh Rostami[1]

[1] *Department of Biology, Science and Research Branch, Islamic Azad University, Tehran, Iran*
[2] *Faculty of Biological Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran*
[3] *Date Palm & Tropical Fruits Research Center, Horticultural Sciences Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Ahwaz, Iran*
*Corresponding authors. E-mail: marjannm@yahoo.com, z-nouri@srbiau.ac.ir

Abstract. Sex determination of date palm seedlings is the challengeable effort for breeders. Different studies based on molecular markers and genome sequencing have provided some insight in to the genetic regions related to sex determination in date palms in general. But due to differences in cultivar population structure and also cost of whole genome sequencing, we may need a more suitable approach in developing countries for this task. Therefore, we suggest using a combination of different available molecular markers and a computational approach to identify the genetic regions involved in sex differentiation in date palm cultivars. In this study we used twenty-three cultivars including 7 male and 16 female cultivars that were examined by 30 different dominant and co-dominant molecular markers which deal with different genomic regions. Grouping of the tree samples based on 178 loci resulted in genetic differentiation of the studied male and female palm trees. Multiple correspondence analysis (MCA) bi-plot also showed genetic difference within male and female trees. Heatmap plot specified those markers which differentiate date palm trees. SSR (simple sequence repeats) and IRAP (inter retrotransposon amplified polymorphism) markers provided sex linked markers for male cultivars. In present study, we introduced sex specific alleles for Iranian male date palm cultivars as a fast track in seedlings. Different association studies performed identified the candidate genetic regions which are significantly associated with sex differentiation in date palm cultivars.

Keywords: DNA based markers, DAPC, LFMM, MCA, *Phoenix dactylifera*, sex determination.

1 INTRODUCTION

Date palm (*Phoenix dactylifera*, Arecaceae) is the one of the most important fruit products (more than 1.3 M tonnes, FAOSTAT, 2019) in

Iran. Date palm trees are propagated by both offshoots and seeds. Each female tree produces 0-3 offshoots per year. This kind of propagation is not in a way of classical breeding along with low diversity (Adawy et al. 2014). Seed germination is supposed to be easiest way for propagation with high heterozygousity while, it is a time consuming method and may take more than 5 years before flowering.

Date palm is dioecious and the only species of the Palmae with sex chromosomes with 2n =36. In fact, this phenomenon makes difficulty for identification of female trees at early stages of growth (Maryam et al. 2016). Using cytological and biochemical methods may not be used to discriminate date palm male trees from the female trees. However, recently, researchers reported the presence of the occurrence of sex-specific loci in date palm trees (Elmeer and Mattat 2012, Adawy et al. 2014, Al-Ameri et al. 2016, Hassanzadeh and Bagheri 2019). These authors used different molecular markers like random amplified polymorphic DNA (RAPD), inter simple sequence repeat (ISSR), simple sequence repeat (SSR), sequence characterized amplified region (SCAR) and start codon targeted polymorphism (SCoT).

Different genome sequencing approaches have been carried out on limited number of palm trees of different palm species (Torres et al. 2018) to date plan cultivars (Al-Dous et al. 2011, Hazzouri et. al. 2019). These studies have identified the genomic regions associated with sex determination and also a conserved two-locus system present in all palm species. However, due to possible interference of population genetic structure in different date palm cultivars cultivated throughout the world and the high expenses of whole genome sequencing, we suggest to use available molecular markers and analyze the results by using computational approaches to locate the genetic regions associated with dates palm sex determination.

Therefore, The present study was performed with following objectives:1-To identify sex-specific alleles for Iranian male date palm cultivars by using a combination of seven dominant and co-dominant molecular makers like, SSR, ISSR, SCoT, IRAP (inter retrotransposon amplified polymorphism), REMAP (retrotransposon microsatellite amplified polymorphismand SRAP (sequence related amplified polymorphism) for date palm sex determination and, 2- identify the candidate genetic regions involved in sex determination by applying different computational methods like, Fisher Exact test, DAPC (Discriminant analysis of principal components analysis), and LFMM (Latent factor mixed model, Collins and Jombbart 2015, Zhang et al. 2019).

2. MATERIALS AND METHODS

2.1. Plant materials and genetic analysis

In total we studied 23 date palm cultivars, of which 7 were male (including: Sabzparak, GhannamiSabz, Wardi, GhannamiSorkh 1, GhannamiSorkh 2, Foreign male 1 and Foreign male 2) and 16 female cultivars.

Two to three trees from each cultivar were randomly selected for molecular studies (totally 63 trees). Details of the cultivars with their accession numbers are provided in our previous study (Saboori et al. 2021). All cultivars are located Omol-tomair station of Date Palm & Tropical Fruits Research Center, Ahwaz, Iran.

For genetic analysis, genomic DNA was extracted based on Saboori et al. (2021). Seven different molecular markers as SSR (6 loci), EST-SSR (3 loci), SCoT (4 loci), ISSR (5 loci), SRAP (5 loci), IRAP (3 loci) and REMAP (4 loci) were examined for sexual determination (Table S1).

The touch-up PCR program was used for SSR markers; 94°C for 5 min, initial 10 cycles at 95°C for 30 sec, annealing step for 1 min at 51°C, 47.5 °C, 62 °C, 47.5°C, 46.9 °C and 45 °C for MPdCIR078, MPdCIR085, PdCUC3-ssr2, MPdCIR090, MPdCIR048 and MPdCIR025 loci respectively, 72°C for 1 min and 30 sec. Then 30 cycles were set at 95°C for 30 sec, annealing step (MPdCIR078 52°C, MPdCIR085 49.9 °C, PdCUC3-ssr2 65 °C, MPdCIR090 49.9 °C, MPdCIR048 48.8 °C and MPdCIR025 48 °C) for 1 min. the extension segment was at 72°C for 1 min and 30 sec following final extension at 72 °C for 15 min.

For EST-SSR the following procedure was used; 5 min at 94 °C, 30 sec at 95 °C, 30 sec at 50 °C, 52 °C- and 60 °C for EST-PDG3119-rubisco, EST-GTE and EST-DPG0633- Laccase loci respectively and 1 min and 30 sec at 72 °C. These segments repeated for 35 cycles. Final extension was for 5 min at 72 °C.

The SCoT loci were amplified based on our previous study (Saboori et al. 2020) and its data used for sexual determination.

The PCR reaction for the ISSR markers were performed according to the following thermal program: 5 minutes at 94°C, 40 cycles of 30 seconds at 94°C, 60 seconds at 50-52.5°C (50°C for (GA)9T, (CA)7AT, and (GA) 9A, 52/5°C for (GT)8YA and (AG) 8YT) and 90 seconds at 72°C and final extension for 7 minutes at 72°C. Molecular marker reactions were carried out in 25 µL volume with 20 ng genomic DNA and 5 U of *Taq* DNA polymerase (Bioron, Germany), 2X PCR buffer (50 mM KCl; 10 mM Tris-HCl, pH;8), 1.5-2 mM MgCl2; 0.2 mM of each dNTP (Bioron, Germany); 0.2 µM of each primer.

For IRAP and REMAP markers, we used thermal program: 3 min of initial denaturation at 94°C, followed

by 40 cycles of 94°C for 3 min and annealing temperature 51-54 °C for IRAPs and 51-56 °C for REMAPs for 30 s, 72°C for 3 min, and 10 min at 72°C for final extension. All reactions were set up at BIORAD Thermocycler, USA.

The band profile of each locus was visualized by using 2.5 % agarose gel electrophoresis. Staining of gels were performed by the sybergreen dye. We used the 100 base pair (bps) molecular size ladder for estimation of fragment size (Fermentas, Germany).

2.2. Data analysis

All obtained bands of 30 loci were scored as a binary data. Multiple correspondence analysis (MCA) of molecular data was performed to group the studied palm trees based on sex differentiation as performed in R-Package 4.1(Abdi and Williams 2010). AMOVA test was performed for differentiation of two male and female groups by using GenAlex ver. 6.5.

For Association studies, different statistical and bioinformatic approaches are available which have different assumptions. We used DAPC (Discriminant analysis of principal components), and Bayesian based method of LFMM (Latent factor mixed model), as suggested by Frichot et al (2013), and Collins and Jombbart (2015). These analyses were followed by Bonferroni correction as well as false discovery rate (FDR) tests, to avoid both type I and type II errors of rejecting true association results. These were done in R package 4.1. Similarly, for accuracy of the model presented by. DAPC we used cross validation method as implemented in R. 4.1.

3. RESULTS

In total we obtained 178 loci/ bands by different molecular markers studied. Ward dendrogram (Fig. 1), differentiated between the studied male and female palm trees as the samples of either group were placed in a separate cluster. Therefore, it shows that the molecular markers used in present study can differentiate palm trees based on different sexes.

Multiple correspondence analysis (MCA) of molecular data obtained revealed that about 60% of total genetic variation of the studied samples may be explained by about 10 PCA axis (Fig. 2).

The grouping of the studied palm trees by MCA biplot by using the first two PCA axes (Fig. 3), not only separated male and female trees from each other but also revealed some degree of genetic difference within male and female samples studied.

In general, two different genetic groups can be seen in either of these sexes. A heat-map was constructed

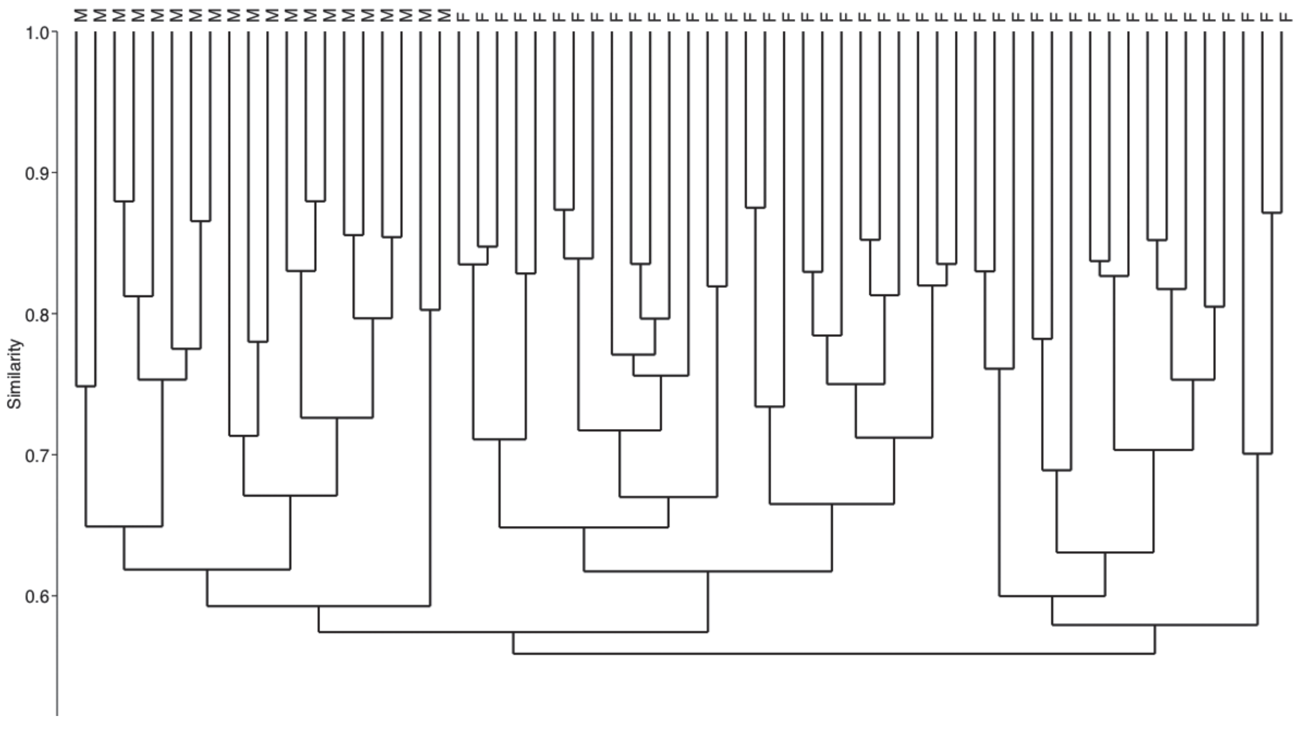


Figure 1. Ward dendrogram of male (M) and female (F) date palm trees based on combined molecular data.

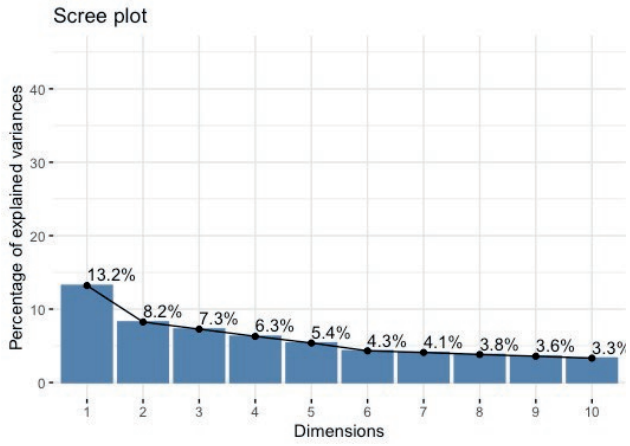


Figure 2. Multiple correspondence analysis (MCA) based on 30 loci studied on date palm trees.

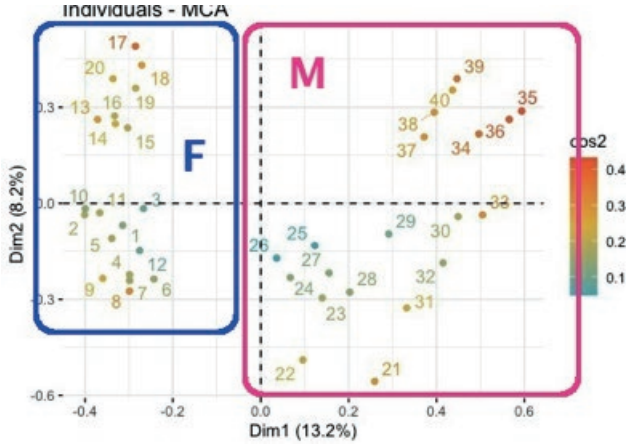
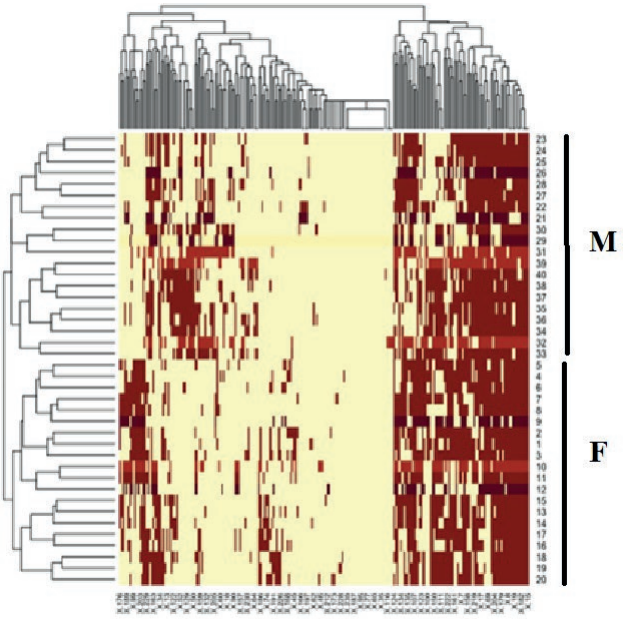


Figure 3. MCA biplot: Grouping date palm trees based on first two components. Numbers are individuals. Color gradient shows amount of variance based on MCAfrom high to low. M; male trees and F: female trees studied.

(Fig. 4) to group those molecular markers which show similarity in differentiating palm tree samples.

AMOVA test also confirmed differentiation between two male and female groups (r = 0.129, P = 0.001, Table S2).

The primers with highest degree of contribution to genetic differentiation of male and female date palm trees have been shown in Fig. 5. Primer bands IRAP-Nikita_2000, 3'LTR_100, ISSR-(CA)7AT-400, SSR-MPdCIR048-120, are among the most contributing primers of the first MCA axis. Similarly, primer bands IRAP-Nikita_200, SSR-MPdCIR048-200, Nikita+SSR2_100, are among the primers which highly contributed in MCA axis 2.



Figue 4. Heatmap cluster based on 30 molecular loci and male (M) and female (F) date palm trees studied.

A much more detailed information can be obtained from MCA-biplot (Fig. 6), which shows the presence (Y), and absence (N), of the particular primer bands contributed in separating female (Numbers 1-20 in Fig. 5), and male (Numbers 21-40, in Fig. 5) palm trees. It is evident from this biplot that both presence and absence of different primer bands of the utilized molecular markers can together differentiate male palm trees from females.

We found some loci which are specific in male date palm cultivars (Table 1, Fig. S1). The MPdCIR048(GA)32-SSR locus, showed specific allele between 100-130 bps. The 110 bps band was unique for all male trees including SabzParak, GhannamiSabz, Wardi, GhannamiSorkh, GhannamiSorkh2, Foreign male1 and Foreign male2 cultivars. The 130 bps allele was observed only in GhannamiSorkh1, GhannamiSorkh2, Foreign male1 and Foreign male 2 while, allele in size of 120 bps was unique in two male cultivars (Sabzparak and Wardi). In EST-SSR data, one allele in EST-PDG3119-rubisco with 200 base pairs in size showed specificity in all male cultivars except SabzParak male cultivar.

Similarly, IRAP-3LTR locus produced alleles between 400 to 1600 bps which were specific at some of the male trees (Table 1). For instance, the band in 300 bps in length was amplified in all male trees except Sabzeparak male cultivar and were absent in all female date palm trees.

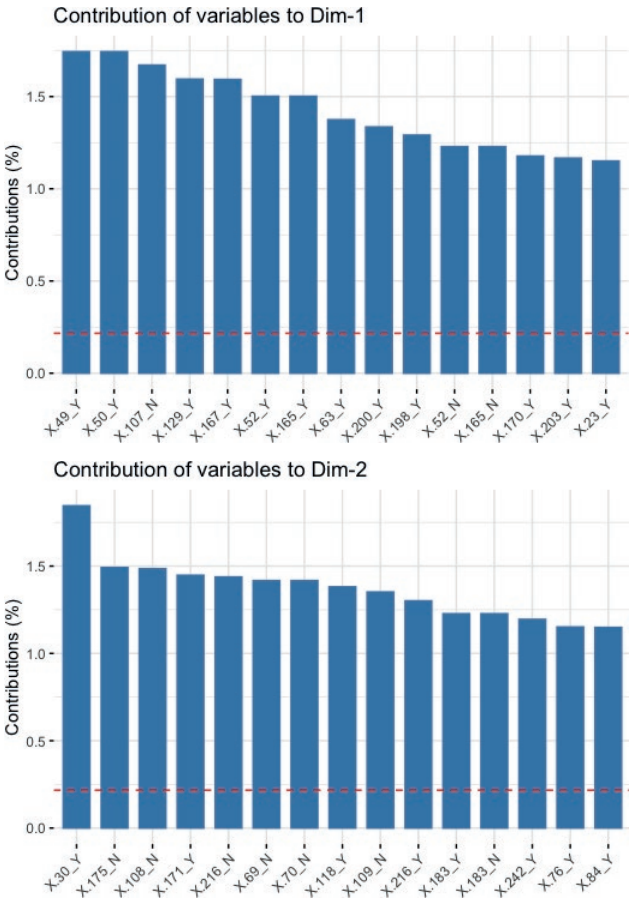


Figure 5. MCA axes. 1 and 2, showing primer bands with highest degree of contribution in date palm male and female differentiation.

3.1. Association studies molecular markers and sex differentiation

All association approaches produced almost similar results and identified the same loci as the genetic regions which are associated with sex differentiation. For example, Fisher exact test and DAPC identified three following marker loci after Bonferroni correction at p = 0.01: X18 (SRAP loci), X50 X51 X53 X64 (IRAP loci), X91 X98 X108 (REMAP loci), and X171 X174 X176 (SSR loci).

These loci can efficiently differentiate male versus female date palm trees studied (Fig. 7).

Similarly, LFMM analysis of both Lasso and Ridge methods produced almost similar results after Bonferroni correction and false discovery rate (FDR) test (Fig. 8). The loci which identified are in agreement with the results of Fisher exact test and DAPC presented before.

Therefore, a combination of molecular markers may be used in early-stages sex determination in date palm cultivars with focus on highly associated marker loci.

4. DISCUSSION

Sex determination is one of the main concerns of date palm breeders. Different studies have been reported to introduce proper biomarkers but they are restricted to few cultivars. Recent GWAS study revealed 112 SNPs related to sex determination in a region with ~6 Mb at LG 12 while other 43 SNPs dispersed on other date palm

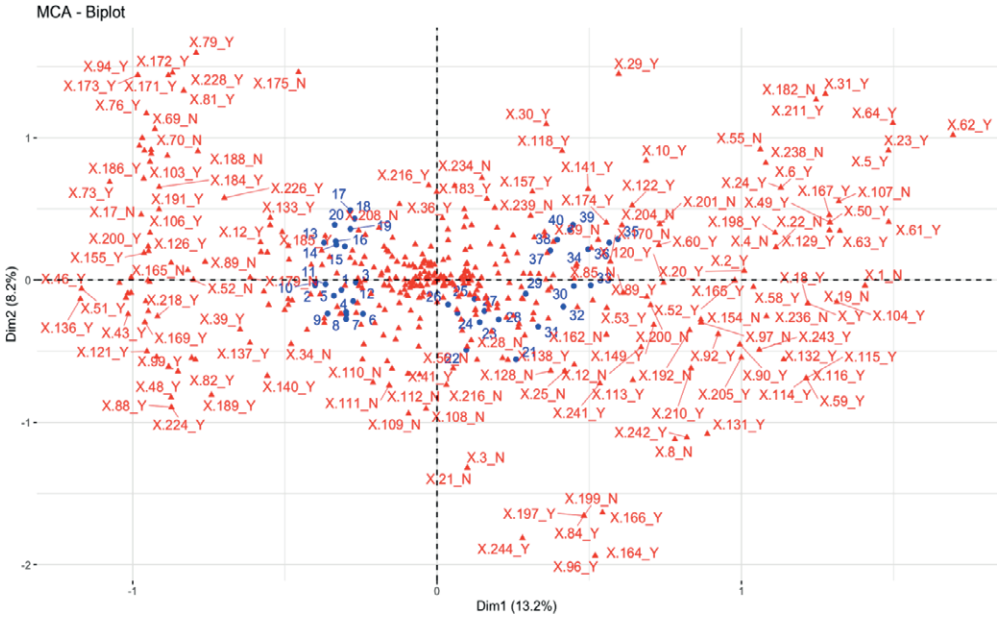


Figure 6. MCA-biplot based on alleles studied. Red numbers : alleles (Y=presence and N=absence), blue number: date palm trees numbers.

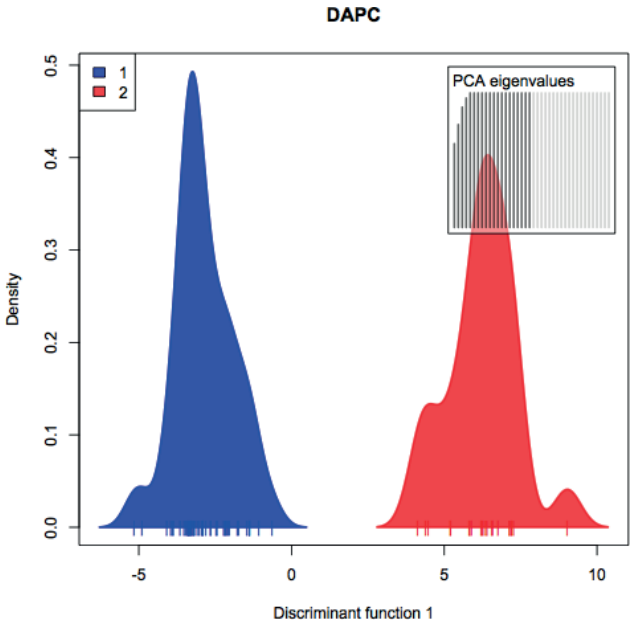


Figure 7. DAPC plot showing differentiation of female (1) versus male (2). date palm trees based on associated marker loci.

chromosomes (Hazzouri et al. 2019). Torres et al. (2021) also reported four conserved genes in *Phoenix* males. Some mutations in putative genes involved in sex determination (CYP703 and GPAT3), is supposed to repress recombination in these regions, leading to gynodioecy and therefore result in establishing male sex in palm tree (Torres et al. 2021). However, genetic variations have been observed in both sex linked and autosomal regions. It may call more effort to find proper sex biomarker for date palm cultivars.

In present study, we used 30 different loci may be located in different genome regions. Base on the allele data, 60% of alleles covered the most variable alleles. The first two MCA components showed the highest level of variation and enabled to differentiate male and female trees. Most of the identified alleles in these two components belong to SSR and IRAP loci (Fig. 5).

We used multiple correspondence analysis (MCA) as a statistical precise method for categorical data (binary data, Abdi and Williams, 2010)

MCA bi-plot provides a picture which illustrates the variable that can discriminates the studied individual MCA bi-plots (Fig. 6) depicted all 250 alleles of 30 loci studied. It clearly shows the contribution of each allele to discrimination of male and female trees. Moreover, different association studies in present paper identified genomic regions which are associated with sex differentiation in date palm cultivars.

In details, we introduced male specific alleles for six important Iranian male date palm cultivars. The MPd-CIR048 SSR locus has two alleles (100, 130 bps) which could discriminate all male cultivars studied (Table 2). This SSR locus located in Cytosine-S-methyltransferase DRM2-like (DRM2). Jskani et al. (2016) also reported

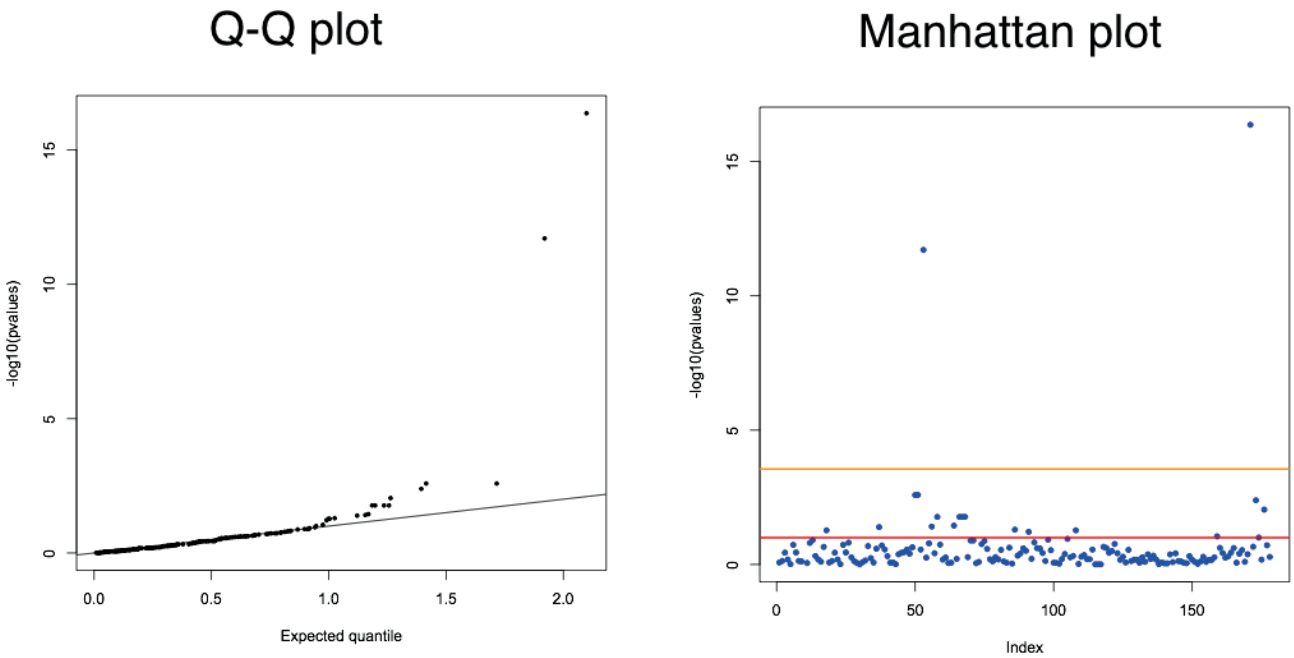


Figure 8. LFMM results showing associated marker loci with sex differentiation in date palm cultivars studied.

Table 1. Specific alleles in male date palm trees. M1 (SabzParak), M2 (GhannamiSabz), M3 (Wardi), M4 (GhannamiSorkh 1), M5 (GhannamiSorkh 2), M6 (Foreign male1), M7 (Foreign male 2)

Locus/Male	M1	M2	M3	M4	M5	M6	M7
SSR (bp)							
	110/120	110/110	110/110	110/130	110/130	110/130	110/130
MPdCIR048	100/120	110/110	110/110	110/110	110/130	110/130	110/130
	100/120	110/110	110/120	110/130	110/130	110/130	110/130
EST- PDG3119-rubisco	-	200	200	200	200	200	200
IRAP-3'LTR(bp)							
	400/500	300/500/600/	300/500/600/	100/200/300 400/500/600 700/800/900 1500	100/200/300 400/500/600 700/750/1500	100/200/300 400/600/700 900/1000/1500 1600	100/200/300 400/600/800 1500/1600

male specific alleles of this locus but in different size (250bps) while we detected that allele in both male and female trees. The 190/160 allele pattern for this SSR locus was reported by Elmeer and Mattat (2012) in Qatar male date palm trees. Maryam et al. (2016) also reported male specific alleles in this locus (250/250) for Pakestani cultivars. It clearly shows the pivotal role of this locus in sex determination although different allelic patterns may be obtained from different male cultivars due to genetic variations of the studied cultivars. On the other hand, Wang et al. 2020 reported that mPdIRDP52 was sex linked for cultivars collected from China.

ISSR markers as a dominant marker showed less discrimination power for sex discrimination. It is in agreement with the results of the study performed by Hassanzadeh Khankahdani and Bagheri (2019). In other plants like *Trichosantes dioica* Roxbi. and *Hippophae rhamnoides* ssp. *turkestanica* ISSR markers could show sex discrimination (Nanda et al. 2013, Adhikari et al. 2014,Das et al. 2017)

Retrotansposone based markers with their abundance, mode of amplification and insertion into the genome, have characteristics suitable for discriminate between species or genotypes (Biswas et al. 2010). About 21% of whole genome of *P.dactylifera* defined for retrotransposons, mostly including Ty1/Copia and Ty3/gypsy (Al-Dous et al. 2011, Al-Mssallem et al. 2013, Nouroz and Mukaramin 2019). We used IRAP and REMAPs' profiles between male and female date palm trees. IRAP loci are among loci with the high degree of contribution to total variance in our MCA analysis. However, only one of IRAP loci (3′LTR) showed sex discrimination in 5 male cultivars. this a first report on male- link IRAP markers in date palm trees.

Adawy et al. (2014) introduced two SCoT alleles (850 bps and 1150 bps) in SCoT36 and SCoT4 in Egyptian female date palm trees as sex-linked markers while in present study we observed these alleles in both gender. cDNA-SCoT allele based on flowering stage was also used for development of SCAR maker by Al-Ameri et al. (2016) for determination of male trees of Saudi Arabia date palm cultivars. They believed that the different gene expression patterns in flowering stage in male and female trees may provide a simple and inexpensive tool for sex determination. However, our literature reviews revealed that sex linked markers in date pale cultivars are specified in local orchards and restricted to the countries. With regard to multilocus nature of SSR markers, it may make different size of alleles for one locus in different cultivars with different geographical locations probably due to newmutations and allele adaptation.

5. CONCLUSION

The urgent method for sex determination in date palm nursery orchards is pivotal for breeding industry. Our data could provide the highest variant alleles among 30 loci studied by MCA for sex determination. We also introduced sex specific alleles for male cultivars as a fast track in seedlings.

ACKNOWLEDGMENT

We acknowledge Science and Research Branch, Islamic Azad University for providing laboratory.

AUTHOR CONTRIBUTION STATEMENT

Z.N.: conceptualization of the project; M.Sh.: analyses of data; S.S, N.M., S.N., F.R.: data collection and lab

work; S.M.: providing samples; Z.N, M.Sh. S.M. and S.S. project design

DATA ARCHIVING STATEMENT

All tree samples used in this research are being archived in Herbarium of ShahidBeheshti University, Tehran, Iran. The accession numbers will be supplied once available.

REFERENCES

Abdi H, Williams LJ (2010) Principal component analysis Wiley interdisciplinary reviews: Comput Stat 2: 433-459.

Adawy SS, Jiang J, Atia MA (2014) Identification of novel sex-specific PCR-based markers to distinguish the genders in Egyptian date palm trees. Int J Agric Sci Res 4: 45-54.

Adhikari S, Saha S, Bandyopadhyay TK, Ghosh P (2014) Identification and validation of a new male sex-specific ISSR marker in pointed gourd (Trichosanthes dioica Roxb.). Sci World J;2014:216896. doi: 10.1155/2014/216896.

Al-Ameri AA, Al-Qurainy F, Gaafar ARZ, Khan S, Nadeem M (2016) Molecular identification of sex in *Phoenix dactylifera* using inter simple sequence repeat markers. BioMed Res Intl 2016: 1-5.

Al-Dous EK, George B, Al-Mahmoud M E, Al-Jaber MY, Wang H, Salameh YM, Malek J A (2011) De novo genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). Nature Biotechnol 29: 521.

Al-Mssallem I S, Hu S, Zhang X, Lin Q, Liu W, Tan J, Yu J (2013) Genome sequence of the date palm *Phoenix dactylifera* L. Nature Commun 4: 1-9.

Biswas MK, Xu Q, Deng XX (2010) Utility of RAPD, ISSR, IRAP and REMAP markers for the genetic analysis of *Citrus* spp. Sci Hort 124: 254-26.

Collins C, Jombbart T (2015) Genome-Wide association studies. Impeerial college London. MRC Center. For Outbreak Analysis and Modeling: 1-61.

Das K, Ganie SH, Mangla Y. et al. (2017) ISSR markers for gender identification and genetic diagnosis of Hippophae rhamnoides ssp. turkestanica growing at high altitudes in Ladakh region (Jammu and Kashmir). Protoplasma 254, 1063–1077. https://doi.org/10.1007/s00709-016-1013-8.

Elmeer K, Mattat I (2012) Marker-assisted sex differentiation in date palm using simple sequence repeats. 3 Biotech. 2: 241-247.

Frichot E, Schoville SD, Bouchard G, Francois O (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. Mol Biol Evol 30:1687-1699.

Hassanzadeh Khankahdani H, Bagheri A (2016) Identification of genetic variation of male and female date palm (*Phoenix dactylifera* L.) cultivars using morphological and molecular markers. Intl J Hort Sci Technol 6: 63-76.

Hazzouri KM, Gros-Balthazard M, Flowers JM, Copetti D, Lemansour A, Lebrun M, Purugganan MD (2019) Genome-wide association mapping of date palm fruit traits. Nature Commun 10: 1-14. https://doi.org/10.1038/s41467-019-12604-9.

Jaskani MJ, Awan FS, Ahmad S, Khan IA (2016) Development of molecular method for sex identification in date palm (Phoenix dactylifera L.) plantlets using novel sex-linked microsatellite markers. 3 Biotech 6:22.

Maryam Jaskani MJ, Awan FS, Ahmad S, Khan IA (2016) Development of molecular method for sex identification in date palm (*Phoenix dactylifera* L.) plantlets using novel sex-linked microsatellite markers. 3 Biotech 6(22): 1-7.

Nanda S, Kar B, Nayak S, Jha S, Kumar Joshi R (2013) Development of an ISSR based STS marker for sex identification in pointed gourd (Trichosanthes dioica Roxb.). Sci Hort 150: 11-15.

Nouroz F, Mukaramin A (2019) Genomic and evolutionary diversity of LTR retrotransposons in date palm (*Phoenix dactylifera*) Pakistan J Bot 51: 1637-1644.

Saboori S, Noormohammadi Z, Sheidai M, Marashi S (2020) SCoT molecular markers and genetic fingerprinting of date palm (*Phoenix dactylifera* L.) cultivars. Genet Resour Crop Evol 67: 73-82.

Saboori S, Noormohammadi Z, Sheidai M, Marashi S (2021) Insight into date palm diversity: genetic and morphological investigations. Plant Mol Biol Rep 39: 137-145.

Torres MF, Mathew LS, Ahmed I, Al-Azwani IK, Krueger R, Rivera-Nuñez D, Malek JA (2018) Genus-wide sequencing supports a two-locus model for sex-determination in Phoenix. Nature Commun 9: 1-9. DOI: 10.1038/s41467-018-06375-y.

Zhang F, Chen W, Zhu Z *et al.* (2019) OSCA: a tool for omic-data-based complex trait analysis. Genome Biol 20: 107 . https://doi.org/10.1186/s13059-019-1718-z

Wang Y, Ihase LO, Htwe YM, Shi P, Zhang D, Li D, et al. (2020) Development of sex-linked SSR marker in the genus Phoenix and validation in *P. dactylifera*. Crop Sci 60: 2452-2466.

SUPPLEMENTARY DATA

Table S1. Molecular marker primers, their names and sequences used in this study.

Marker	Locus name	Sequence (5′-3′)
SSR (Bodian et al. 2014)		
	MPdCIR025(GA)22-F	GCACGAGAAGGCTTATAGT
	MPdCIR025(GA)22-R	CCCCTCATTAGGATTCTAC
	MPdCIR048(GA)32-F	CGAGACCTACCTTCAACAAA
	MPdCIR048(GA)32-R	CCACCAACCAAATCAAACAC
	MPdCIR078(GA)13-F	TGGATTTCCATTGTGAG
	MPdCIR078(GA)13-R	CCCGAAGAGACGCTATT
	mPdCIR085(GA)29-F	GAGAGAGGGTGGTGTTATT
	mPdCIR085(GA)29-R	TTCATCCAGAACCACAGTA
	MPdCIR090(GA)26-F	GCAGTCAGTCCCTCATA
	MPdCIR090(GA)26-R	TGCTTGTAGCCCTTCAG
	PdCUC3-ssr2(GA)22-F	ACATTGCTCTTTTGCCATGGGCT
	PdCUC3-ssr2(GA)22-R	CGAGCAGGTGGGGTTCGGGT
EST-SSR		
(Zhao et al. 2012)	EST-PDG3119-rubisco-F	CATACTGATTATTGGCACACC
	EST-PDG3119-rubisco-R	GTACCATACCGTACCAGTTCA
(Zhao et al. 2012)	EST-DPG0633- Laccase -F	AGACTGGTTAAGTTGGTGGAG
	EST-DPG0633-Laccase-R	CTACAAAACTGATGTGGTGGT
	EST-GTE-F	GCTTGGCCATCTATGAAAC
	EST-GTE-R	ACTCTGAGCATCCATATCG
SCoT (Collard and Mackill 2009)		
	SCoT1	CAACAATGGCTACCACCA
	SCoT2	CAACAATGGCTACCACCC
	SCoT36	GCAACAATGGCTACCACC
	SCoT41	CAATGGCTACCACTGACA
ISSR		
	(GA)9T	GAGAGAGAGAGAGAGAGAT
	(CA)7AT	CACACACACACACAAT
	UBC849	GTGTGTGTGTGTGTGTYA
	(GA)9A	GAGAGAGAGAGAGAGAGAA
	UBC834	AGAGAGAGAGAGAGAGYT
SRAP (Feng et al. 2014)		
	ME1	TGAGTCCAAACCGGATA
	ME4	TGAGTCCAAACCGGACC
	EM3	GACTGCGTACGAATTGAC
	EM4	GACTGCGTACGAATTTGA
IRAP		
	Nikita	CGCATTTGTTCAAGCCTAAACC
	3'LTR	TGTTTCCCATGCGACGTTCCCCAACA
	5'LTR1	TTGCCTCTAGGGCATATTTCCAACA
REMAP		
	SSR_PdCIR025_R + NIKITA	
	SSR_PdCIR025_R + 3'LTR	
	SSR_PdCIR048_R + NIKITA	
	SSR_PdCIR048_R + 5'LTR1	

Table S2. AMOVA test based on date palm trees in two groups: male and female trees. df; degree of freedom, SS: sum of square, MA: mean of square, Var: variance.

Source	df	SS	MS	Est. Var.	var%
Among Pops	1	135.730	135.730	3.491	13%
Within Pops	68	1598.284	23.504	23.504	87%
Total	69	1734.014		26.996	100%
Stat	Value	P value			
PhiPT	0.129	0.001			

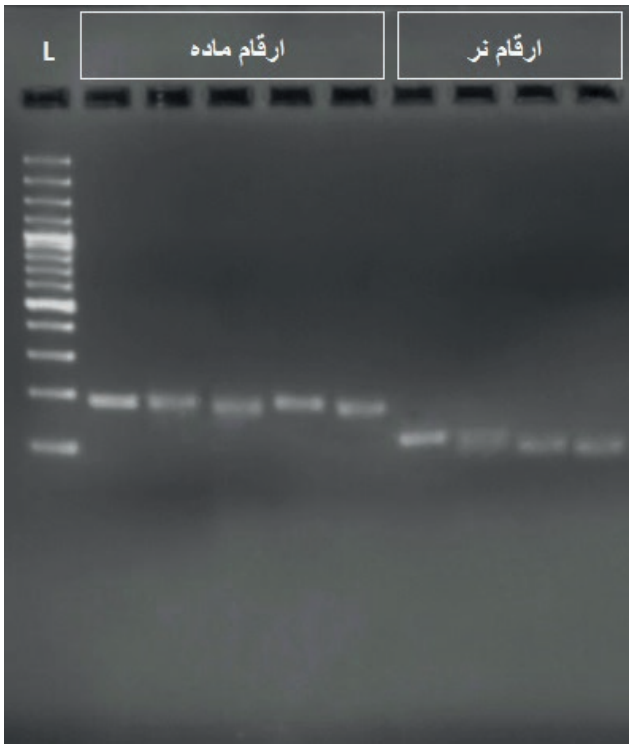


Figure S1. Allele pattern of some femal and male date palm trees. Left to right: first 5 samples are female and 6-9 male trees.