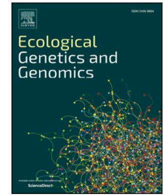
Ecological Genetics and Genomics 33 (2024) 100299

Contents lists available at ScienceDirect

Ecological Genetics and Genomics

journal homepage: www.elsevier.com/locate/egg



Genetic diversity and structure of Tunisian and Indian date palm (*Phoenix dactylifera* and *sylvestris*) cultivars and genotypes revealed by AFLP markers

Karim Kadri [a,b,**], Anoop Anand Malik [f], Hammadi Hamza [c], Salem Marzougui [d,e], Mohamed Aziz Elhoumaizi [g], Shyam Sundar Sharma [h], Mohammed Elsafy [i,*]

[a] *Date Palm Genetic Resources, Centre Regional de Recherches en Agriculture Oasienne (CRRAO), Tozeur Road, Km1, PB 62, 2260, Degache, Tunisia*
[b] *Laboratory of Biotechnology Applied on Agriculture, Institut National de La Recherche Agronomique de Tunisie, University of Carthage, Street Hedi Karray, 2049, Ariana, Tunisia*
[c] *Arid and Oases Cropping Laboratory, Arid Region Institute, Medenine, 4119, Tunisia*
[d] *Université de Carthage, INRAT, LR16INRAT02, Laboratoire des Grandes Cultures, 2049, Tunisia*
[e] *Université de Jendouba, Ecole Supérieure d'Agriculture du Kef, 7119, Le Kef, Tunisia*
[f] *TERI-Deakin Nano-Biotechnology Centre, The Energy and Resources Institute, 122 001, Gurugram, Haryana, India*
[g] *Laboratory for Agricultural Productions Improvement, Biotechnology and Environment (LAPABE), Faculty Sciences, University Mohammed First, BP:717, 60000, Oujda, Morocco*
[h] *Department of Biotechnology, TERI. School of Advanced Studies, Vasant Kunj, New Delhi, India*
[i] *Department of Plant Breeding, Swedish University of Agricultural Sciences, Växtskyddvägen1, P.O. Box 190, SE-234 22, Lomma, Sweden*

ARTICLE INFO

Keywords:
Date palm
Genetic diversity
AFLP markers
Genetic structure
Heatmap

ABSTRACT

The date palm breeding programs need to discover valid genetic fingerprints to characterize cultivars and assess their genetic diversity. This study assessed the genetic diversity among thirty-nine date palm cultivars from Tunisia (*Phoenix dactylifera*) and India (*Phoenix sylvestris*) by using six AFLP (Amplified Fragment Length Polymorphism) markers. 360 loci were amplified, with 127 loci polymorphic (34.35 %). The Jaccard's similarity coefficient ranged from 0.161 to 0.931, with the mean genetic distances of 0.568. AFLP's average marker index value was 7.28, with a resolving power of 10.91. The analysis of population structure showed two main clusters with a clear separation between Tunisian and Indian cultivars.

Furthermore, the heatmap analysis allowed the identification of 10 bands specific to the Indian accessions, which were not detected in Tunisian genotypes. These loci could be linked to genes involved in adapting the species in Indian lands, which allowed the study of the genetic diversity of date palm resources of different origins, confirming the existence of at least two origins of domestication. Additionally, identifying AFLP loci specific to *P. dactylifera* and *P. sylvestris* will significantly contribute to breeding programs by exploiting species-specific polymorphisms.

1. Introduction

The dioecious nature of the date palm (*Phoenix dactylifera* L.) has contributed to significant genetic variability when propagated through seeds. This inherent genetic diversity has facilitated the selection of numerous clones with distinct morphological and physiological traits. Consequently, countries with a tradition of date palm cultivation possess a rich genetic heritage. To comprehensively study this genetic wealth, it is crucial to distinguish between two forms: The diverse genetic pool comprising millions of hybrid date palms resulting from seed propagation and the 'varietal' heritage, consisting of cultivars maintained through vegetative reproduction.

The number of date palm cultivars recorded is estimated at more than 500 in Iraq, 400 in Iran, 300 in Libya, 450 in Morocco, and almost 250 in Tunisia [1]. Knowledge of the genetic resources of cultivated species is of great importance for improving productivity and the efficiency of any activity aimed at its preservation and enhancement. With the advancement of knowledge on date palm genetic resources in the

* Corresponding author.
** Corresponding author. Date Palm Genetic Resources, Centre Regional de Recherches en Agriculture Oasienne (CRRAO), Tozeur Road, Km1, PB 62, 2260, Degache, Tunisia.
E-mail addresses: kadri.karim@iresa.agrinet.tn (K. Kadri), mohammed.elsafy@slu.se (M. Elsafy).

https://doi.org/10.1016/j.egg.2024.100299
Received 21 August 2024; Received in revised form 4 October 2024; Accepted 7 October 2024
Available online 9 October 2024
2405-9854/© 2024 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

ancient world, two date palm genetic pools linked to geographic distribution have been described [2–5]. Unlike the countries of the Middle East, where the genetic resources of male date palm cultivars have been studied [6,7], the genetic heritage of this species is still unknown in northern countries.

In India, date palm is an important fruit crop of the arid and semiarid region and predominantly grown in the western border, especially in the Kutchh district of Gujarat with about 18286 ha. with a production of 171522 MT of fresh fruits. Among Rajasthan, Maharashtra, Tamil Nadu, Kerala, and other states Gujarat is the top producer. To our knowledge, a few studies on genetic diversity analysis were detected on Indian date palms, including a study on genetic diversity analysis of date palm germplasm conducted in the Kutch region of India using RAPD and ISSR markers by Srivashtav, Kapadia [8]. In this study, RAPD markers detected 39.77 % polymorphism among eight genotypes compared to 23.07 % with ISSR markers.

In Tunisia, and unlike female cultivars, work aimed at inventorying male date palm cultivars has not yet made it possible to list the entire Tunisian phoenicicol heritage, and a large majority of cultivars, especially in the palm groves of the southwest, have not yet been described. These male palms, locally named "Dokkars," form heterogeneous populations that are rarely cloned, and each individual has its own characteristics [9,10]. They are sometimes identified by the name of the female cultivar, which phenotypically resembles her. In Tunisia, the male feet "Dokkars" are poorly known, and their multiplication is often done by seeds [11], unlike other phoenicultural countries (Iraq) where "Dokkars" are selected from the best female cultivars and their multiplication is by offshoots and have known names. We are talking about "Fehls," "Siwi," "Sammani," "Zuegloul" [12]. A large number can disappear even before they are known, even if preliminary studies have shown some originality of this germplasm in Tunisia [13,14]. Information on genetic diversity is essential for the conservation and sustainable utilization of important plant genetic resources. The knowledge of genetic variation provides valuable insights into the taxonomy and phylogeny of plant species. In recent years, several molecular studies have been carried out in date palm aiming at a reliable assessment of genetic diversity using PCR-based markers RAPD [15] ISSR [14], SSR [16] and AFLP [17]. These markers have shown considerable performance in date palm marker-assisted conservation, selection, and improvement programs. AFLP markers are highly reliable since they overcome drawbacks associated with other dominant marker techniques, such as RAPDs and ISSRs. [17,18]; The AFLP techniques can be used as an ideal tool for genetic diversity studies owing to its criterion when there is no prior sequence information to develop sequence-based markers such as SSR, SSAP, CAPS, and SCAR for rapid generation of large data for genetic diversity studies [19,20]. AFLP markers are also advantageous due to their comprehensive coverage throughout the genome, high multiplex ratio, and high resolution power [18,21]. As part of our research work, we have focused our efforts on the analysis of genetic diversity to obtain a more in-depth analysis of date palm male pollinators [10,14] using different types of molecular markers (ISSR, RAPD, SSR). This work continues the analysis of this diversity in male date palm cultivars. However, the present study aims to explore the usefulness of AFLP markers as a first step toward the evaluation of relatedness in a collection of date palm cultivars from Tunisia and India without the aim of using them in future breeding programs and for improvement.

2. Materials and methods

2.1. Plant material

This study was carried out on thirteen-nine germplasms. Leaves from thirty-three cultivars, including male and female of the Tunisian palm (*Phoenix dactylifera* L.), were collected; these palm trees are cultivated on sandy-loamy soil in the experimental station of the regional oasis agriculture research center, located in the district of Degache, governorate Tozeur, southwestern Tunisia (33.97846299469204, 8.20801877724391). In addition, apart from Tunisian cultivars, leaves from six date palm accessions (*Phoenix sylvestris*) were randomly sampled from India (Table 1S).

2.2. DNA Isolation

Total genomic DNA was isolated from lyophilized leaves by following a modified CTAB-based protocol [21,22]. Isolated genomic DNA was analyzed by electrophoresis in 1 % agarose gels for its qualitative and quantitative evaluation [23]. the gel was visualized under UV light and photographed using the gel documentation system (Azure™, C150 Biosystem). In addition, DNA aliquots were diluted to a final working concentration of 25–30 ng/μl.

2.3. AFLP Assays

In this work, the TE-AFLP procedure was based on Van der Wurff et al. (2000) and Sharma et al. (2016) with some modifications. Genomic DNA (250 ng) was digested using 5 units of *Eco*RI and *Pst*I and 2.5 units of *Mse*I, followed by an enzymatic inactivation step which consists of incubation at 70 °C for 10 min. The digested DNA was ligated with the *Eco*RI and *Pst*I adapters using 10 pmol of each adapter and 1 unit of T4 DNA ligase. The ligation was done at 20 °C. for 2 h. The adapter sequences were 5′-CTCG TAGACTGCGTACC and 5′-AATTGG-TACGCAGTCTAC for *Eco*RI and 5′-CTCGTAGACTGCGT ACATGCA and 5′-TGTACGCAGTCTA for *Pst*I, respectively. The ligation solution mixture was diluted 1:10 in buffer TE (10 mM Tris 0.1 mM EDTA). Primers specific to the *Eco*RI and *Pst*I adapters were used during the preamplification of the ligation products. PCR workflows for pre-amplification were identical to those followed in classical AFLP (Sharma et al., 2011). The selective amplification conditions were carried out using *Eco*RI primers labeled with γ32P-ATP with two selective nucleotides in combination with *Pst*I primers not labeled with one selective nucleotide. PCR conditions for selective amplification were 10 cycles consisting of a 30s denaturation step at 94 °C, a 30s annealing step at 70 °C and a 60s elongation step at 72 °C followed by 30 cycles of 94 °C for 30s, 60 °C for 30s, and 72 °C for 60s. A final extension of 5 min at 72 °C was programmed to allow the completion of the elongation of the products. DNA fragments were separated on 6 % denaturing polyacrylamide gels and visualized by silver staining using a silver staining kit (Biorad). Electrophoresis and autoradiography were performed as in the case of classic AFLP (Sharma et al., 2011).

2.4. Genetic Analysis

A preliminary selection was conducted on fewer palm accessions (Tunisian and Indian) using a set of 12 TE-AFLP primers. Finally, 6 combinations of primers (Table 1) were selected based on the number of loci generated, their polymorphic information content, their high-resolution power, and their degree of polymorphism. The fragments amplified on polyacrylamide gel were manually labeled as "1″ for the presence of bands and "0″ for their absence for each primer combination. All bands (monomorphic and polymorphic) were recorded, but only polymorphic bands were used to generate dissimilarity matrices. To estimate genetic dissimilarities, the binary matrix was used for determining the Jaccard's coefficient,

$$GDij = (b + c) / a + (b + c);$$

Where GD is the measure of genetic dissimilarity between individuals i and j, a is the number of polymorphic bands that i and j share, b is the number of bands present in i and absent in j, and c is the number of bands present in j and absent in i [24]. In order to create the phenetic dendrogram, the dissimilarity matrix was subjected to neighbor-joining clustering. The reliability and robustness of the phenograms were

Table 1
The statistics of the polymorphism obtained in the palm collection using the 6 combinations of AFLP primers.

Primer combination	Total bands	Polymorphic bands	Polymorphism (%)	Polymorphic Information Content (PIC)	Marker Index (MI)	Resolving Power (RP)
E-AG x M-CG	62	34	54.84	0.31	10.46	15
E-ACT x M-CG	56	30	53.57	0.36	10.78	16
E-AG x M-GA	72	11	15.28	0.43	4.73	7.8
E-ACT x M-GA	71	33	46.48	0.33	10.9	16
E-AT x M-GT	26	4	15.38	0.31	1.23	1.6
E-AA x M-CAT	73	15	20.55	0.37	5.58	8.7
	360	**127**	**34.35**	**0.36**	**7.28**	**11**

assessed using percentage values for each tree node, calculated by bootstrap analysis with 1000 iterations in R software version 4.3.3 [25].

2.5. Marker Attributes Analysis

The different marker attributes of each AFLP primer combination, like polymorphism information content (PIC), The PIC were computed using the following equation (Roldan-Ruiz et al., 2000): $PIC_i = 2f_i\,(1-f_i)$. Marker index (MI) was evaluated by using the equation as defined by Roland-Ruiz et al. (2000), according to the following formula: $MI = PIC \times EMR$, where the effective multiplex ratio (EMR) is the total number of polymorphic loci per primer. Each primer combination's resolution power (RP) was determined using an equation by Prevost and Wilkinson [26]. According to the following formula: $RP = \Sigma I_b$, Ib indicates the informative character of the band calculated according to the following method: $Ib = 1 - (2 \times |5 - p|)$, where p is the fraction of the total accessions in which the band is present. Dendrograms were performed using R software version 4.3.3.

2.6. Principal Coordinates Analysis (PCoA)

The PCoA plot was generated using the *vegan* package in R for calculating the principal coordinates analysis (PCoA) based on Euclidean distance. Clustering was performed with the *kmeans* function, and visualization was achieved using the *ggplot2* package for custom plotting.

2.7. Population Structure

The population structure of the studied accessions was analyzed using STRUCTURE software version 2.3.4, employing a model-based Bayesian clustering method [27]. We explored a range of putative subpopulation numbers from K = 1 to 10, running 320000 burn-in iterations followed by 320000 recorded Markov chain iterations for each K. To assess the sampling variance, we conducted 10 independent runs for each K. To estimate the optimal number of subpopulations, we calculated the logarithm of the likelihood for each K, denoted as ln P(D) = L (K). We then used the method proposed by Evanno et al. (2005) to find the most likely number of groups. This involved computing ΔK, where $\Delta K = [L''(K)]/SD$, to identify the peak value, indicating the optimal number of subpopulations.

2.8. Mantel Test

The Mantel test was performed to assess the correlation between genetic and geographic distance matrices using Pearson's product-moment correlation, with significance determined through 999 permutations. Geographic distances were calculated using the *Haversine* formula based on latitude and longitude data.

2.9. AMOVA

The percentage of molecular variance was calculated using GenALEx 6.51b2. Pairwise Population PhiPT Values were calculated between Tunisian and Indian populations with permutations of 999 following this equation:

PhiPT = *AP / (WP + AP) = AP / TOT*

where AP = Est. Var. Among Pops, WP = Est. Var. Within Pops.

2.10. Clone Assignment

The clone assignment was conducted using hierarchical clustering on date palm AFLP markers. *Jaccard* distance was calculated with the *vegan* package, and clustering was performed using the *hclust* function in the stats package. Clone groups were defined using a cut-off threshold with the *cutree* function, and results were visualized with a dendrogram plot.

3. Results

3.1. Assessment of Genetic Diversity

Fig. 1 shows the results of the AFLP profile of the 39 Date palm genotypes obtained with the primer combination E-ACT X M-CG. A total of 360 bands were generated by six primer combinations, with a mean number of 60 amplicons per primer combination. The number of polymorphic amplicons was 127 (34.35 %), and the one polymorphic band per primer varied between 4 for E-AT x M-GT to 34 for E-AG x M-CG with an average of 22 fragments primer of 22 (Table 1). The maximum genetic similarity (Jc = 0.931) was between accessions ABD2 and ABD3, which were collected from the same experimental plot, CRRAO Degache in Tunisia, noting that great morphological similarities have also been recorded between these two cultivars, mainly concerning the characters of the inflorescence. On the other hand, the minimum genetic similarity (Jc = 0.161) was found between accessions P171 collected from the experimental plot of CRRAO, Tunisia, and PD5 collected from Madhya Pradesh state, India. The mean genetic distance between the germplasm was 0.568, suggesting a high degree of genetic diversity between the Tunisian and Indian accessions.

3.2. Marker Attributes

The EMR, PIC, MI, and RP with polymorphic bands were calculated to evaluate the discriminatory power of each AFLP primer combination. The EMR value ranged from 4 (E-AT x M-GT) to 34 (E-AG x M-CG), PIC values ranged from 0.31 (E-AG x M-CG and E-AT X M-GT) to 0.43 (E-AG x M-GA) and MI value from 1.23 (E-AT x M-GT) to 10.90 (E-ACT x M-GA). RP values varied from 1.59 (E-AT x M-GT) to 16.15 (E-ACT x M-CG). The mean values for all marker attributes like effective multiplex ratio, polymorphic information content, marker index, and resolving power were 21.16, 0.36, 7.28, and 10.91, respectively (Table 1).

3.3. Genetic Relationships

The accession cluster analysis was performed using the neighbor-joining unweighted clustering technique and based on data obtained from AFLP markers (Fig. 2). Based on the genotypic data, three major clusters were identified. Cluster 1 showed the largest group, which includes seventy-seven genotypes composed of indigenous Tunisian

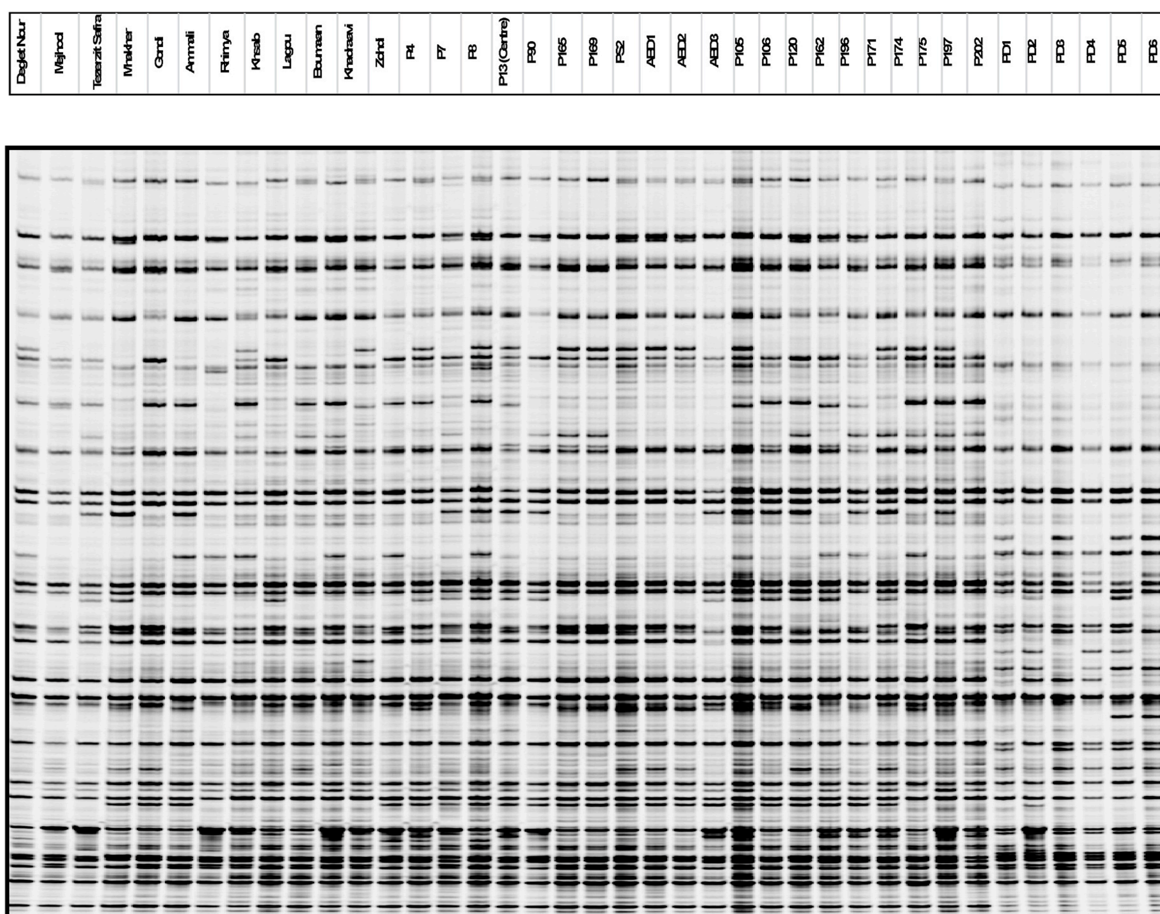


Fig. 1. A representative AFLP profile of Date palm using primer combination E-ACTxM-CG.

cultivars except for the Mejhoul. This group is subdivided within itself into 7 subgroups. Cultivars are composed of three sub-clusters. The first sub-group comprised three male pollinators "ABD1″, "ABD2″ and "ABD3". The second sub-cluster was formed by the pollinators" P7″," P4″," P175″ and the female cultivars "Deglet Nour". The third sub-cluster was composed of three pollinators "P90″, "P105″ and "P106". The fourth subgroup was formed by four pollinators, "P13″, "P8″," P165″," P120," and the female cultivars, "Gondi"," Lagou" and "Menaker". The five subclusters were comprised of three cultivars "Ammali"," P169″ and" PS2". The sixth subgroup was formed only by two female cultivars, "Tezerzit Safra" and "Mejhool". The last subgroup was formed only by male pollinators "P196″," P171″," P197″," P174″ and" P162".

3.4. Principal Coordinate Analysis (PCoA)

In order to confirm this conclusion, the results of pairwise similarity indices and dendrograms constructed for different cultivars from the AFLP were analyzed using principal coordinate analysis (PCoA). The absorption of the total variability is subdivided into 10 axes, with a maximum absorption for the first two axes (Fig. 3), contributing 33.9 % of the total variability. Considering the distribution of cultivars according to the PCoA (axis 1-axis 2). Three clusters were identified in the cultivar distribution, similar to those reported in the UPGMA dendrogram (Fig. 2). However, this confirmed the significant divergence of Indian cultivars from others, which correlates positively with axis 1 (Fig. 3).

In addition, this observed grouping is achieved depending on the geographical origin of the cultivars studied. This distribution shows that the second group is subdivided into two subgroups; the first is formed by female foreign cultivars introduced (Khadraoui, Khsab, Bouman, Zohdi, etc.), which show a negative correlation with axis 2 and significantly diverge from the native affiliations. The other clade is formed by the two varieties Mejhool, Tezerzit Sefra, and the two pollinators P171 and P197, which also characterize the negative part of axis 2. Regarding the third group, which was formed by Tunisian male trees and grouped with Indigenous female trees (Deglet Nour, Gondi, and Lagau), this cluster correlates positively with both axes (Fig. 3).

3.5. Population Structure

AFLP datasets were analyzed using a model-based Bayesian clustering approach implemented in STRUCTURE. Population structure was assessed using STRUCTURE v.2.3.4 software to analyze the set of 39 date palm cultivars. The optimal number of K was analyzed using ΔK, and our results showed strong, clear peaks for ΔK at K = 2, and secondary clustering at K = 5 (Figs. 4 and 2S), indicating the existence of two subpopulations in the studied date palm accessions. The four distinct groups observed at K = 2 represent two main genetic clusters corresponding to Tunisian-Middle Eastern and Indian date palm cultivars, respectively. At K = 5, sub-structuring within the main groups is evident, suggesting further genetic differentiation possibly related to specific geographic or cultivar-specific factors.

The collection is subdivided into two large groups. For the Indian cultivars, they formed the eastern group, which is significantly different from the second. The second group represent the western part and its formed by introduced female cultivars such as Rhimiya, "Kshab", "Kadrawi", "Zohdi", "Medjool" and the autochthonous Tunisian male and female cultivars (Fig. 2S).

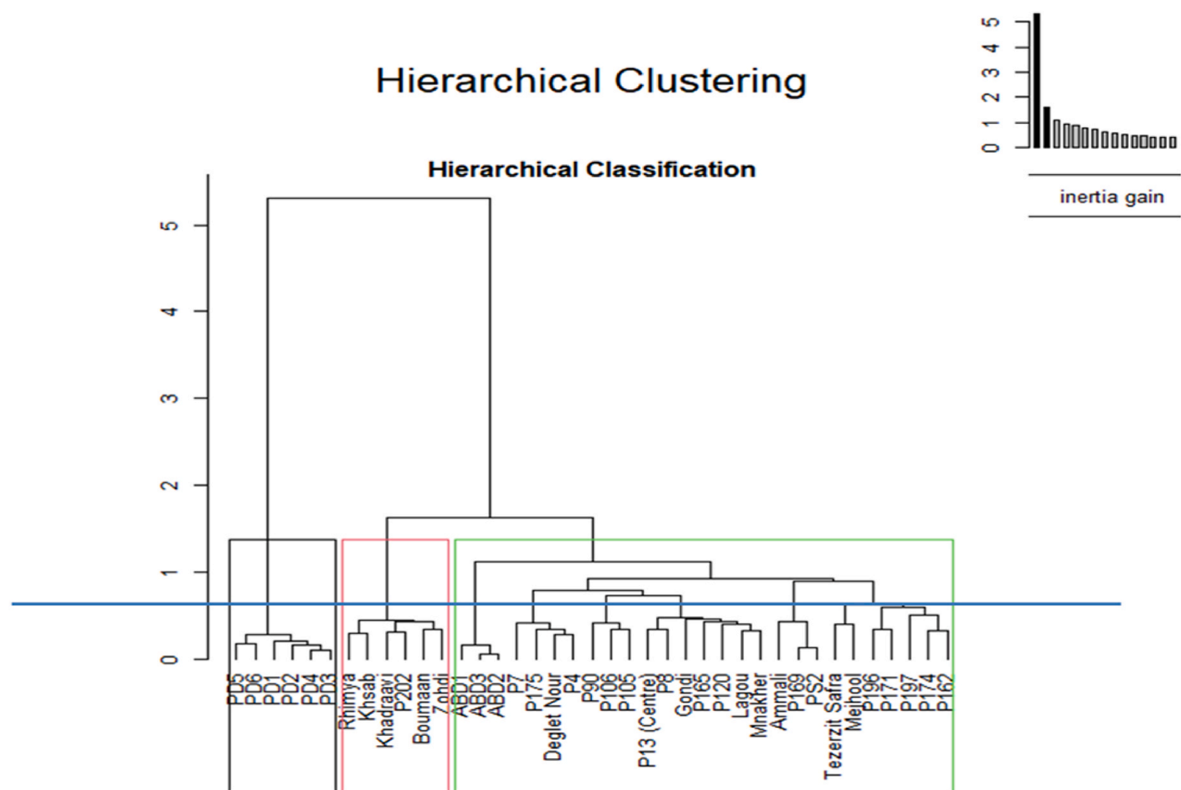


Fig. 2. The UPGMA dendrogram based on AFLP fingerprint according to Neighbor-joining (NJ) analysis of genetic similarities estimated among the 39 genotypes/ accessions of date palm.

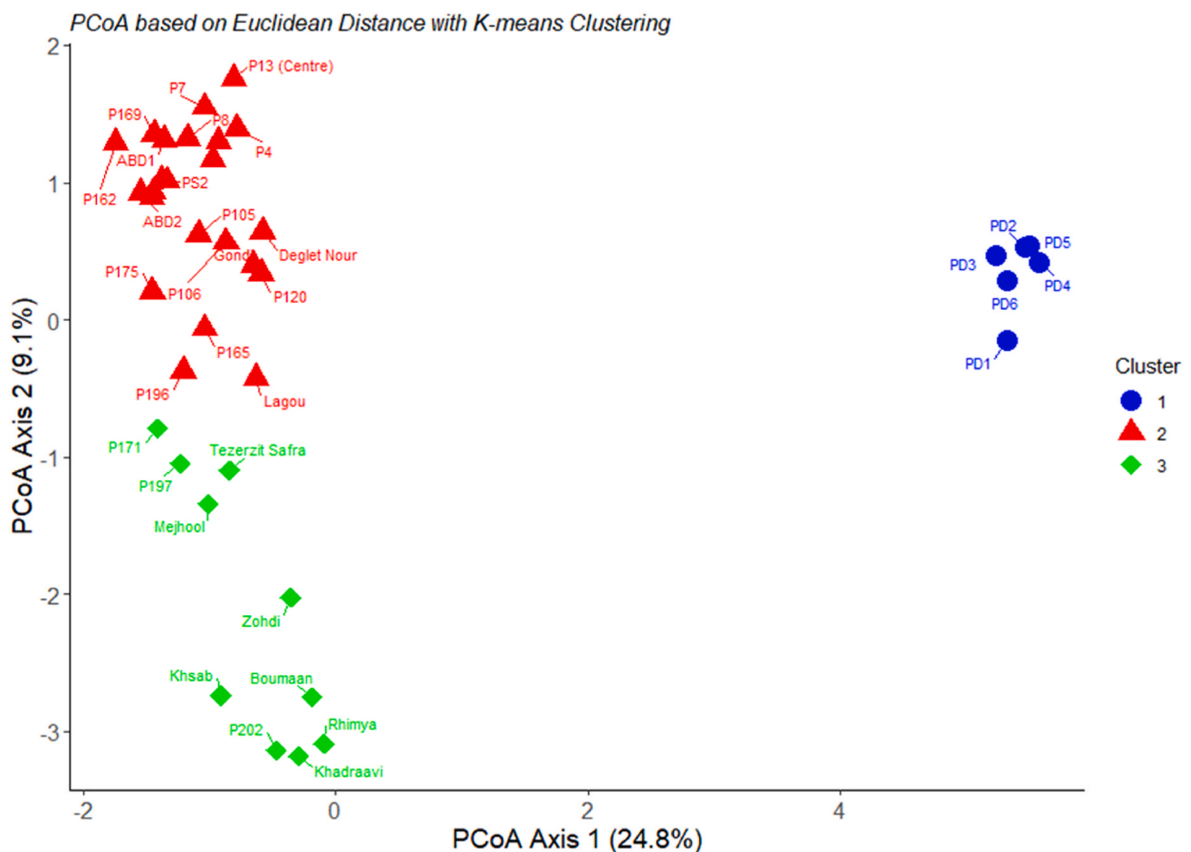


Fig. 3. Principle Cordinate Analysis (PCoA) plots based on amplified fragment length polymorphism of the 38 date palm cultivars.

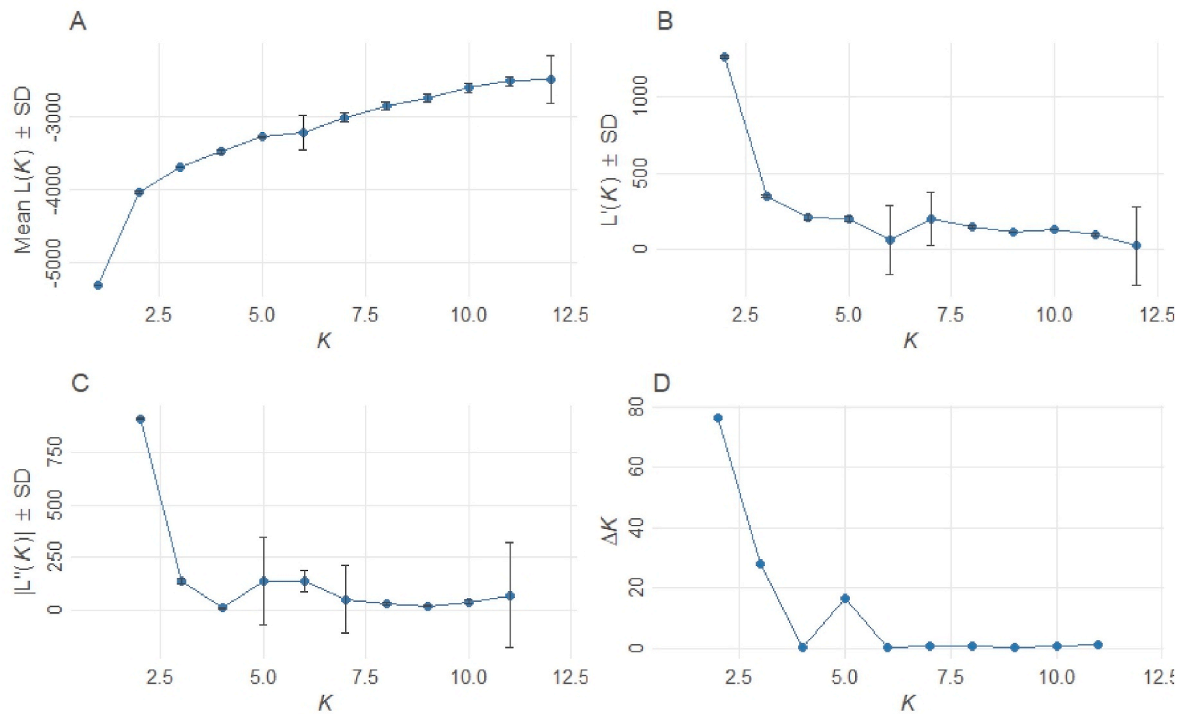


Fig. 4. Delta K values calculated according to STRUCTURE Harvester software within the complete data set of 39 cultivars. (A) Mean likelihood (LnP(D)) for each value of K. (B) Variance of LnP(D). (C) Mean value of the second-order rate of change (L''(K)) for each K. (D) Delta K (ΔK) indicating the most likely number of clusters.

3.6. Mentel Test

The Mantel test exhibited a strong positive correlation ($r = 0.7487$) between geographic and genetic distance. This correlation is highly significant ($P = 0.001$), suggesting that geographic separation is an important factor contributing to genetic differentiation between the populations.

3.7. AMOVA Analysis

According to AFLP data showed a significant genetic difference ($P = 0.001$) among the studied populations, with a total genetic diversity of 52 %. Pairwise comparisons of populations (Fig. 5) showed that the Tunisian population and the Indian one were significantly different, and the PhiPT was 0.518 ($P = 0.001$).

The hierarchical analysis of molecular variance supports a highly structured population, as evidenced by group differentiation, and the results were highly significant ($P < 0.001$).

3.8. Heat map graphic

Heatmap analysis was conducted based on the AFLP loci matrix using R software. (Fig. 1S). This analysis identified 10 AFLP loci specific to Indian cultivars, which are defined as follows: rr112, rr118, rr7, rr110, rr111, rr95, rr103, rr68, rr77, and rr36. 7 specific AFLP markers identify Tunisian cultivars: rr107, rr114, rr21, rr55, rr105, rr34, and rr41. The other markers were relatively common to all the cultivars analyzed. On the other hand, no specific sex-linked markers were defined. However, rr99, rr98, and rr29 bands were specific to PD5, PD6, and Khadrawi cultivars.

The hierarchical clustering dendrogram based on Clone Assignment Analysis reveals the genetic relationships among individuals in the

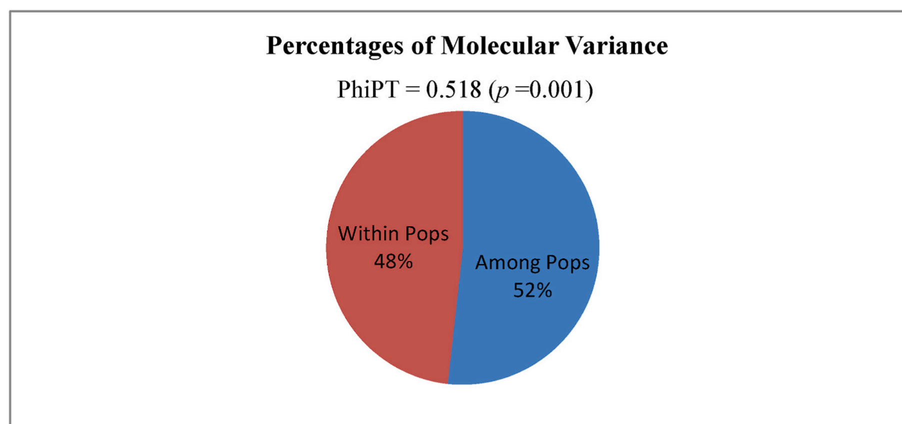


Fig. 5. AMOVA and PhiPT values between Tunisian and Indian date palm populations.

dataset, with distinct clone groups identified based on a cut-off threshold 0.1 (Fig. 6). Each branch represents a genotype, and clusters formed below the threshold indicate genetically similar individuals. For instance, 'ABD2′ and 'ABD3′ are grouped closely, suggesting they are clones or closely related genotypes. The varying branch heights suggest moderate genetic diversity within the collection, with unique genetic profiles observed for individuals like 'Lagou' and 'P120'. This analysis provides valuable insights into the genetic structure and diversity of the germplasm, aiding in effective management and breeding decisions.

4. Discussion

Studying the genetic diversity of date palm cultivars using morphological markers always gives unreliable results, especially when the cultivars are young and the fruits are not produced. It requires many phenotypic traits that are difficult to assess and depend on environmental conditions [28,29]. Molecular polymorphism has the advantage that environmental conditions do not influence the DNA sequence, the sample collection period, or the growth stage [10]. Several genetic fingerprinting techniques were used in date palm, but none met all the requirements regarding cost, ease of handling, and cultivar analysis. However, the AFLP technique meets the requirements and gives reliable results with few primer pairs. The present study used these promoted molecular markers to characterize and evaluate the diversity of male and female date palms from Tunisia and India. A combination of six AFLP primers generated a total of 360 amplicons, of which 127 were polymorphic, generating a polymorphism rate of 34.35 %. Khierallah, Bader [30] found similar results, with a polymorphism rate of 33 %. However, Sabir, Abo-Aba [17] and Cao and Chao [31] found a high polymorphism rate (75 and 85 %, respectively) within date palm cultivars from Saudi Arabia and the USA. However, the level of genetic diversity detected in the present study was very high (0.568). All germplasm cultivars constitute five major clusters congruent with geographical locations. Khierallah, Bader [30] found an average gene diversity of 0.31 using AFLP in an Iraq collection of 18 date palm cultivars (11 females and 7 males). A similar rate of genetic diversity was found by Jubrael, Udupa [32] by analyzing eighteen Iraqi date palm cultivars for AFLP markers, and similar to what Rhouma, Zehdi-Azouzi [7] revealed a diversity of 0.33 in a local Tunisian date palm collection (40 female ecotypes) using AFLP loci.

In addition, the mean value of the polymorphic information content (PIC) was 0.36, which is close to what Sabir, Abo-Aba [17] reported in a Saudi date palm collection using AFLP markers (0.37).

Moreover, it is very close to Khierallah, Bader [30] found in 18 Iraqi date palm cultivars (0.31). On the other hand, Rhouma et al. (2011) found a higher PIC value (0.70) in the Tunisian collection of 40 date palm cultivars revealed by the AFLP markers, which approved the effectiveness of these markers in revealing molecular polymorphism. Indeed, the PIC values directly reflect the effectiveness of the molecular marker. The effective multiplex ratio, the marker index, and the resolving power were 21.16, 7.28, and 10.91, respectively. Rhouma-Chatti, Zehdi-Azzouzi [33] have detected a high value of resolving power and marker index, respectively 25.8 and 50.54 using AFLP markers and a value of RP (4.06) and MI (5.98) using RAMPO markers. Sabir, Abo-Aba [17], in another study on the genetic diversity of Saudi Arabian date palm using AFLP and ISSR markers, also found almost similar type of results except for the value of effective multiplex ratio and maker index detected by them were very high (534) and (266)

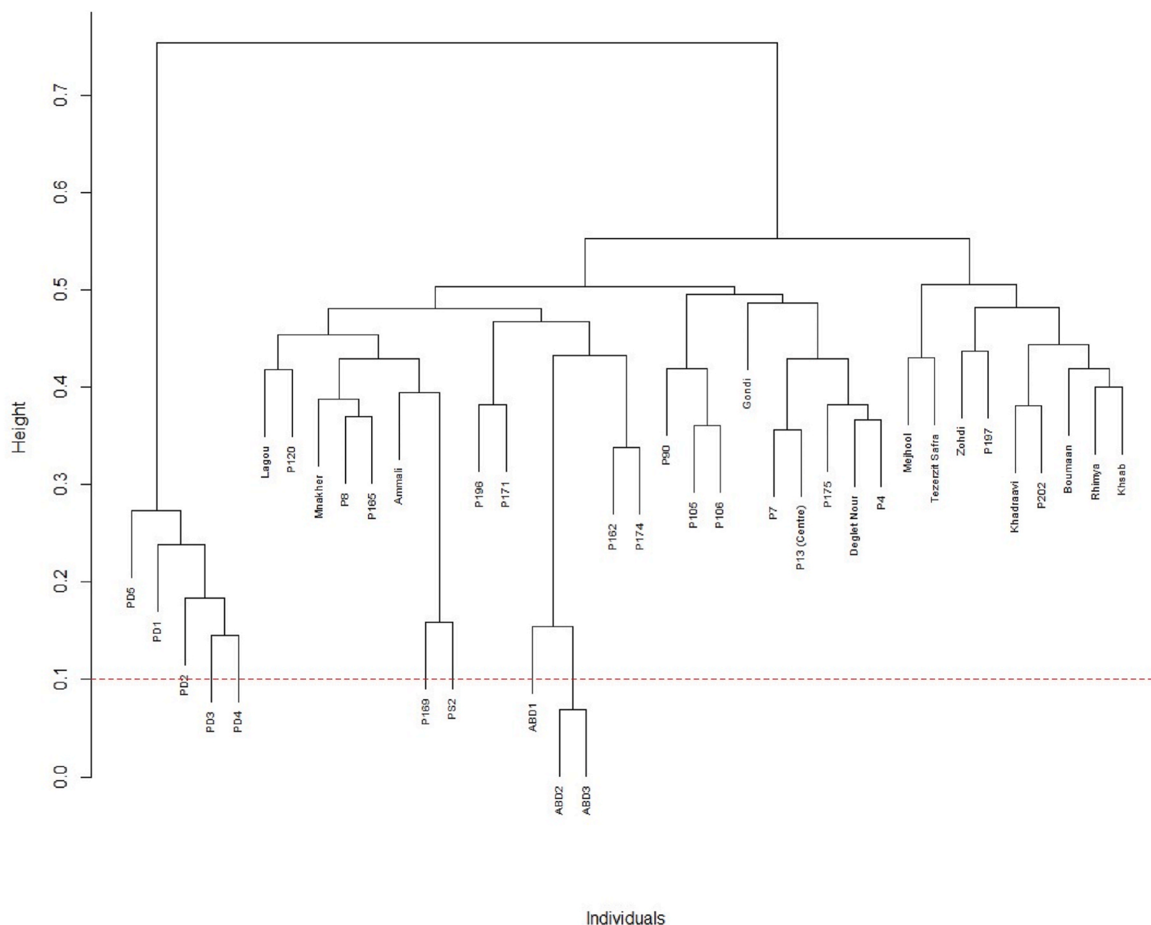


Fig. 6. Clone assignment analysis based on hierarchical clustering dendrogram.

using AFLP markers.

The PCoA analysis showed that the female foreign cultivars are grouped into a separate cluster, which implies the distribution of these cultivars according to geographical origin. On the other hand, the work of Rhouma, Zehdi-Azouzi [7] showed the distribution of native and foreign cultivars in the same group. Similarly, the multivariate analysis showed the segregation of Indian cultivars into a separate group; these results are consistent with those found by Sabir, Abo-Aba [17], who showed the grouping of Saudi date palm cultivars based on geographic location using AFLP markers. According to Khierallah, Bader [30], the variety distribution by PCoA analysis showed the general genetic diversity of Iraqi date palm germplasm. A tremendous genetic convergence of cultivars has been observed despite their grouping independently of their geographical origin despite their morphological diversity. The level of polymorphism among cultivars with AFLP markers ranged from 63 % to 84 % with an average of 76 % and emphasizes that the AFLP markers generated can be utilized in the future in breeding programs of date palm. Our results show that the AFLP markers presented a high level of polymorphism, allowing better discrimination of the cultivars compared to the other techniques previously used in the Tunisian date palm using isozyme and molecular markers such as RFLP, RAPD, RAMPO, ISSR, and SSR by Refs. [11,14,28,33,34]. The analysis of the genetic structure of the population confirmed the results obtained by the hierarchical analysis and the multivariate analysis by grouping this collection into three large groups, with a quite remarkable segregation of Indian cultivars (*Phoenix sylvestris*) from other cultivars (*Phoenix dactylifera*). Our results confirm the work of Gros-Balthazard, Newton [2]. In fact, according to these results, the difficulties in distinguishing the species of *Phoenix* and identifying close relatives of the date palm have long hampered research on its origin. Many hypotheses have been put forward to attribute a wild ancestor to it [35]. Some authors claim that the cultivated date palm comes from one or more wild forms of the same species [36]. Others stipulate that it derives from another species of the genus *Phoenix: P. sylvestris*, *P. canariensis*, *P. atlantica* and *P. reclinata* have been proposed [35]. Finally, a final hypothesis combines the previous two: the date palm comes from hybridization between wild date palms and another species of the Phoenix genus [35]. A genetic analysis based on nuclear microsatellite markers and a chloroplast minisatellite recently refuted the last two hypotheses [4]. Indeed, the allelic profile of the date palm appears strongly divergent from other Phoenixes, indicating that *P. dactylifera* is a distinct species that was domesticated independently of the others [4]. Several hypotheses have long coexisted concerning the center of domestication of the date palm [2]. It would be located in the north of Africa [2] or in India/Pakistan [37].

The primary clustering at $K = 2$ effectively distinguishes Tunisian and some Middle Eastern date palm cultivars (*Phoenix dactylifera*) from Indian cultivars (*Phoenix sylvestris*), indicating significant genetic divergence. The secondary clustering at $K = 5$ reveals finer substructuring within *P. dactylifera*, likely due to localized adaptation, breeding practices, or distinct genetic lineages among Tunisian and Middle Eastern cultivars. This aligns with a previous study by Saboori, Sheidai [38], who used SSR markers and *K-means* clustering to differentiate 36 date palm cultivars, revealing both broad and finer genetic groups. Their findings demonstrated that genetic clustering not only separates major population groups but also captures subtle differences due to adaptations to diverse environments and selective breeding practices. These insights emphasize the need to consider broad and nuanced genetic variations in breeding programs and conservation efforts for date palm populations.

The Mentel indicates a strong positive correlation ($r = 0.7487$) between geographic and genetic distance. This positive correlation observed in this study suggests that geographic separation is a significant factor contributing to genetic differentiation among these populations. Zehdi et al. [5] demonstrated that the genetic structure of date palms is geographically organized into distinct gene pools, primarily categorized into eastern and western groups. The studies indicate that populations from different geographic regions exhibit significant genetic divergence, with Fst values indicating strong differentiation between eastern (e.g., accessions from Asia and Djibouti) and western (e.g., accessions from North Africa) populations. This geographic isolation limits gene flow between populations, accumulating unique alleles and adaptations specific to local environments [39]. Using molecular markers, such as amplified fragment length polymorphisms (AFLP) in our study has revealed high levels of genetic diversity within date palm populations. However, this diversity is often structured according to geographic origin, with distinct allelic variations observed in populations from different regions. For instance, accessions from India have been noted to possess the richest allelic diversity compared to those from Tunisia. In conclusion, the positive correlation between geographic separation and genetic differentiation in date palm populations underscores the importance of both natural and anthropogenic factors in shaping the genetic landscape of this economically vital species [40]. This knowledge is essential for conservation strategies and breeding programs to maintain genetic diversity and improve crop resilience.

According to many authors, the origin of the cultivated date palm is found in the Middle East [41,42]. This thesis seems accredited in the current state of archaeological data [43]. However, data are still rare in northern Africa, and it is not excluded that it was domesticated in several regions, including the Middle East and the western region of its distribution area. Resolving the geographical origins of the date palm requires cross-examination between different approaches. Based on diversity studies, genetics makes it possible to identify centers of domestication, such as corn [44]. For the date palm, diversity studies are still on a local scale, as for Tunisia [13], Sudan [45], and Marrocoo [46]. A study carried out on a global scale, including cultivars of diverse origins based on molecular markers of the SSR microsatellite type, confirmed the existence of two pools named Eastern and Western, suggesting that they each have their autochthonous origin [5].

The high genetic variability (48 %) observed within these populations of date palm (*Phoenix dactylifera* L.) can be attributed to several key factors, such as the palm reproduction mode, which favors cross-pollination, allowing greater genetic diversity by mixing genes between different varieties [47]. In addition, exchanges between farmers contribute to increased genetic mixing. Another factor is the adaptation of palms to environmental conditions, which exposes them to different abiotic stresses such as soil salinity and water stress. This exposure has led to the emergence of varieties well adapted to specific conditions, thus increasing genetic diversity [48]. Over time, the domestication of date palm has been influenced by natural and artificial selection, where favorable traits have been retained, leading to variability within cultivated populations [49]. These combined factors explain why date palms exhibit high genetic variability within their populations, which is essential for their adaptation and survival in the face of environmental changes and agricultural challenges.

Heatmap analysis allowed us to identify specific markers for Indian cultivars. These specific markers can be used in marker-assisted selection programs. The occurrence of these species-specific loci between the species *Phoenix dactylifera* and *Phoenix sylvestris* (wild palm), both ecologically and genetically, may be linked to various factors such as natural selection, domestication, and random mutations, thus contributing to the richness of genetic diversity within the genus *Phoenix*. In fact, these two species have distinct genetic structures, which may lead to species-specific loci. For example, studies have shown that genetic diversity within *Phoenix dactylifera* is high, with marked geographic variations that may influence the presence of particular alleles or loci in specific populations [5,39,49]. In addition, spontaneous mutations that occur in populations of these species may also lead to the occurrence of species-specific loci. The allelic diversity observed in microsatellite studies (SSR) shows that each species may have unique alleles that are not found in the other species [47]. These two species may occupy different ecological niches, resulting in adaptation to varied

environmental conditions. This favors the accumulation of specific loci related to disease resistance, drought tolerance, etc. [50,51].

The hierarchical clustering identified distinct clone groups, indicating a moderate level of genetic diversity within the germplasm collection. Unique genotypes, such as 'Lagou' and 'P120′, highlight the presence of valuable genetic variation. These findings support the need for targeted conservation efforts to preserve genetic diversity, as emphasized by Hoban, Hauffe [52].

These AFLP markers have been considered in several studies as an effective tool for analyzing genomic diversity and the genetic fingerprinting of cultivars [53,54]. In other studies, these AFLP markers have also been identified as economically important traits [55,56]. Al-Mssallem, Hu [57] interestingly demonstrated that *Phoneix dactylifera* identified specific AFLP loci linked to genes for stress resistance and sugar metabolism. In addition, the identified specific markers for Indian cultivars can also provide an idea about DNA sequences involved in the species-environment adaptation.

5. Conclusions

The results presented in this work show that the date palm presented genetic diversity characterized by geographical structuring. Indeed, the specific AFLPs loci confirmed the existence of two Eastern and Western gene pools, suggesting they each have their indigenous origin. The results indicated that AFLP is an efficient marker technology for unveiling genetic relationships between genotypes. This information can be used to establish basic collections valuable in genetic improvement strategies, selection of male pollinators with metaxenic effects such as ripening date or fruit quality or resistance to different stresses, and resilience to climate changes through correlation and linkage studies at the genome level.

CRediT authorship contribution statement

Karim Kadri: Writing – original draft. **Anoop Anand Malik:** Formal analysis. **Hammadi Hamza:** Formal analysis. **Salem Marzougui:** Writing – review & editing. **Mohamed Aziz Elhoumaizi:** Writing – review & editing. **Shyam Sundar Sharma:** Writing – review & editing, Visualization, Data curation. **Mohammed Elsafy:** Writing – review & editing, Visualization, Supervision, Conceptualization.

Ethics approval and consent to participate

All methods were carried out following relevant guidelines and regulations.

Funding

Not Applicable.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by The Agricultural Research and Higher Education Institute-Tunisia (Project CCVPP), The Energy and Resources Institute (ITEC cooperation), and the Department of Plant Breeding at the Swedish University of Agricultural Sciences (SLU).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.egg.2024.100299.

Data availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

[1] A. Rhouma, Le palmier dattier en Tunisie: Le patrimoine génétique, 1994. Arabesques.
[2] M. Gros-Balthazard, et al., Origines et domestication du palmier dattier (Phoenix dactylifera L.). État de l'art et perspectives d'étude, Revue d'ethnoécologie 4 (2013).
[3] K.M. Hazzouri, et al., Whole genome re-sequencing of date palms yields insights into diversification of a fruit tree crop, Nat. Commun. 6 (1) (2015) 8824.
[4] J.-C. Pintaud, et al., Biogeography of the date palm (Phoenix dactylifera L., Arecaceae): insights on the origin and on the structure of modern diversity, in: I International Symposium on Date Palm 994, 2011.
[5] S. Zehdi-Azouzi, et al., Genetic structure of the date palm (Phoenix dactylifera) in the Old World reveals a strong differentiation between eastern and western populations, Ann. Bot. 116 (1) (2015) 101–112.
[6] S. Hannachi, Inventaire variétal de la palmeraie algérienne, 1998 (No Title).
[7] S. Rhouma, et al., Genetic diversity in ecotypes of Tunisian date-palm (Phoenix dactylifera L.) assessed by AFLP markers, J. Hortic. Sci. Biotechnol. 82 (6) (2007) 929–933.
[8] V. Srivashtav, et al., Genetic diversity analysis of date palm (Phoenix dactylifera L.) in the Kutch region of India using RAPD and ISSR markers, Emir. J. Food Agric. 25 (11) (2013).
[9] N. Bouguedoura, et al., Date palm status and perspective in Algeria. Date palm genetic resources and utilization, Africa and the Americas 1 (2015) 125–168.
[10] K. Kadri, et al., Morphological and molecular evaluation of the genetic diversity of Tunisian local date palm pollinators, Acad. J. Biotechnol. 3 (2) (2015) 26–34.
[11] K. Kadri, M. Abdelhafidh, K. Aounallah, Analysis of the morphological diversity of inflorescence of six tunisian date palm (Phoenix dactylifera L.) pollinators, Eur. J. Agric. Food Sci. 3 (4) (2021) 56–62.
[12] M. Bacha, et al., Compatibility relationships in some date palm cultivars (Phoenix dactylifera L.), J. King Saud Univ. Agric. Sci 12 (2) (2000) 81–95.
[13] N. El Kadri, M.B. Mimoun, J.I. Hormaza, Genetic diversity of Tunisian male date palm (Phoenix dactylifera L.) genotypes using morphological descriptors and molecular markers, Sci. Hortic. 253 (2019) 24–34.
[14] K. Karim, R. Elkadri, M. Souhayla, Use of morphological and molecular markers (ISSR) in the analysis of the genetic diversity of a collection of date palm pollinators (Phoenix dactylifera L.) in the oases of Degache, Tunisia, Agriculture Oasienne et Développement Durable (2020) 19.
[15] S.H. Al Hadidi, H.N. Al Agha, A. Talaat, Genetic diversity of date palm in Qatar, in: Qatar Foundation Annual Research Forum Proceedings, Bloomsbury Qatar Foundation Journals Qatar, 2010.
[16] S. Zehdi, et al., Molecular polymorphism and genetic relationships in date palm (Phoenix dactylifera L.): the utility of nuclear microsatellite markers, Sci. Hortic. 148 (2012) 255–263.
[17] J.S. Sabir, et al., Characterization of ten date palm (Phoenix dactylifera L.) cultivars from Saudi Arabia using AFLP and ISSR markers, Comptes Rendus Biol. 337 (1) (2014) 6–18.
[18] S.S. Adawy, et al., Genomic diversity in date palm (Phoenix dactylifera L.) as revealed by AFLPs in comparison to RAPDs and ISSRs, Arab J Biotechnol 8 (1) (2005) 99–114.
[19] S. Freeman, et al., Isolation and characterization of highly polymorphic microsatellites in tea (Camellia sinensis), Mol. Ecol. Notes 4 (3) (2004) 324–326.
[20] P. Vos, et al., AFLP: a new technique for DNA fingerprinting, Nucleic acids research 23 (21) (1995) 4407–4414.
[21] S.S. Sharma, et al., Genetic diversity, population structure and association study using TE-AFLP markers in Pongamia pinnata (L.) Pierre germplasm, Tree Genet. Genomes 13 (2017) 1–10.
[22] A. Singh, et al., Assessment of genetic diversity in Azadirachta indica using AFLP markers, Theor. Appl. Genet. 99 (1999) 272–279.
[23] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: a Laboratory Manual, 1989.
[24] P. Jaccard, Nouvelles recherches sur la distribution florale, Bull. Soc. Vaud. Sci. Nat. 44 (1908) 223–270.
[25] X. Pierrer, J. Jacquemoud-Collet, Available from: Internet, DARwin Software, vol. 4, 2006, pp. 125–134, version 5.0. http://darwin.cirad.fr/Download.php [cited January 5, 2010].
[26] A. Prevost, M. Wilkinson, A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars, Theor. Appl. Genet. 98 (1999) 107–112.
[27] J.K. Pritchard, et al., Association mapping in structured populations, Am. J. Hum. Genet. 67 (1) (2000) 170–181.
[28] H. Hamza, et al., Investigation of genetic variation in Tunisian date palm (Phoenix dactylifera L.) cultivars using ISSR marker systems and their relation with fruit characteristics, Turk. J. Biol. 36 (4) (2012) 449–458.
[29] M. Sedra, Development of new Moroccan selected date palm varieties resistant to bayoud and of good fruit quality, Date palm Biotechnol. (2011) 513–531.

[30] H. Khierallah, et al., Assessment of genetic diversity for some Iraqi date palms (Phoenix dactylifera L.) using amplified fragment length polymorphisms (AFLP) markers, Afr. J. Biotechnol. 10 (47) (2011) 9570–9576.
[31] B.R. Cao, C.-C.T. Chao, Identification of date cultivars in California using AFLP markers, Hortscience 37 (6) (2002) 966–968.
[32] J.M. Jubrael, S.M. Udupa, M. Baum, Assessment of AFLP-based genetic relationships among date palm (Phoenix dactylifera L.) varieties of Iraq, J. Am. Soc. Hortic. Sci. 130 (3) (2005) 442–447.
[33] S. Rhouma-Chatti, et al., Exploitation of AFLP and RAMPO markers in the assessment of genetic diversity within Tunisian date palm (Phoenix dactylifera L.) cultivars, in: I International Symposium on Date Palm 994, 2011.
[34] M. Trifi, A. Rhouma, M. Marrakchi, Phylogenetic relationships in Tunisian date-palm (Phoenix dactylifera L.) germplasm collection using DNA amplification fingerprinting, Agronomie 20 (6) (2000) 665–671.
[35] D. Johnson, Worldwide dispersal of date palm from its homeland, Acta Hortic. 882 (2010) 369–375.
[36] D. Zohary, P. Spiegel-Roy, Beginnings of fruit growing in the old world: olive, grape, date, and fig emerge as important bronze age additions to grain agriculture in the near East, Science 187 (4174) (1975) 319–327.
[37] J.P. Hamilton, Travels through the Interior Provinces of Colombia, 1827.
[38] S. Saboori, et al., Genetic (SSRs) versus morphological differentiation of date palm cultivars: Fst versus Pst estimates, Caryologia 74 (3) (2021) 151–168.
[39] F.T. Purayil, et al., Genetic variability in selected date palm (Phoenix dactylifera L.) cultivars of United Arab Emirates using ISSR and DAMD markers, 3 Biotech 8 (2018) 1–8.
[40] S.R. Chaluvadi, et al., Phoenix phylogeny, and analysis of genetic variation in a diverse collection of date palm (Phoenix dactylifera) and related species, Plant Divers. 41 (5) (2019) 330–339.
[41] L. Costantini, P. Audisio, Plant and insect remains from the Bronze Age site of Ra's al-Jinz (RJ-2), Sultanate of Oman, Paleorient (2000) 143–156.
[42] D. Zohary, M. Hopf, E. Weiss, Domestication of Plants in the Old World: the Origin and Spread of Domesticated Plants in Southwest Asia, Europe, and the Mediterranean Basin, Oxford University Press, 2012.
[43] M. Tengberg, *Research into the origins of date palm domestication.* The date palm. From Traditional Resource to Green Wealth, 2003, pp. 51–62.
[44] J. Van Heerwaarden, et al., Genetic signals of origin, spread, and introgression in a large sample of maize landraces, Proc. Natl. Acad. Sci. USA 108 (3) (2011) 1088–1092.
[45] M. Elsafy, et al., Elucidating the genetic diversity of farmer cultivars of female date palms (Phoenix dactylifera L.) from Sudan by microsatellite markers, Genet. Resour. Crop Evol. 63 (2016) 975–986.
[46] M.A. Elhoumaizi, et al., Phenotypic diversity of date-palm cultivars (Phoenix dactylifera L.) from Morocco, Genet. Resour. Crop Evol. 49 (5) (2002) 483–490.
[47] M. Ibrahimi, et al., Analysis of genetic diversity and population structure of Moroccan date palm (Phoenix dactylifera L.) using SSR and DAMD molecular markers, J. Genet. Eng. Biotechnol. 21 (1) (2023) 66.
[48] H. Rahman, et al., Recent advances in date palm genomics: a comprehensive review, Front. Genet. 13 (2022) 959266.
[49] W. Ahmed, et al., Genetic diversity and population structure of date palms (Phoenix dactylifera L.) in Ethiopia using microsatellite markers, J. Genet. Eng. Biotechnol. 19 (1) (2021) 64.
[50] K. Elmeer, et al., Assessing genetic diversity of shishi date palm cultivars in Saudi Arabia and Qatar using microsatellite markers, Int. J. Hortic. Sci. Technol. 6 (1) (2019) 1–9.
[51] I.A. Mohammed, I.A. Mohamed, Validation of RAPD and ISSR markers used for sex determination in date palm grown under Sudan conditions, Adv. Res. Life. Sci. 3 (1) (2019) 17–22.
[52] S.M. Hoban, et al., Bringing genetic diversity to the forefront of conservation policy and management, Conserv. Genet. Res. 5 (2013) 593–598.
[53] G. Baraket, et al., Comparative assessment of SSR and AFLP markers for evaluation of genetic diversity and conservation of fig, Ficus carica L., genetic resources in Tunisia, Plant Mol. Biol. Rep. 29 (2011) 171–184.
[54] Y.H. Park, M.A. West, D.A. St Clair, Evaluation of AFLPs for germplasm fingerprinting and assessment of genetic diversity in cultivars of tomato (Lycopersicon esculentum L.), Genome 47 (3) (2004) 510–518.
[55] J. Kepiro, M. Roose, AFLP markers closely linked to a major gene essential for nucellar embryony (apomixis) in Citrus maxima× Poncirus trifoliata, Tree Genet. Genomes 6 (2010) 1–11.
[56] Y. Zhang, et al., Identification of amplified fragment length polymorphism (AFLP) markers tightly associated with drought stress gene in male sterile and fertile Salvia miltiorrhiza Bunge, Int. J. Mol. Sci. 14 (3) (2013) 6518–6528.
[57] I.S. Al-Mssallem, et al., Genome sequence of the date palm Phoenix dactylifera L, Nat. Commun. 4 (1) (2013) 2274.