



ISSN: 0067-2904

Analysis of Sequences and Molecular Evolution in the Iraqi Date Palm Cultivars (*Phoenix dactylifera L.*) Propagated by Tissue Culture, Based on ITS1 Region.

Mohammed M. Hawash, Laith M.J. Al-Shamma

Department of Biology, Collage of Science, University of Baghdad, Baghdad, Iraq

Received: 27/10/2023

Accepted: 6/6/2024

Published: 30/5/2025

Abstract

The study involved seven Iraqi date palm (*Phoenix dactylifera L.*) cultivars propagated through tissue culture (TC) and two cultivars propagated by offshoots (Offs). The internal transcribed spacer 1 (ITS1) region of nuclear ribosomal DNA (nrDNA) was used as the marker for investigation. In an effort to institute genetic relations between various date palm cultivars and uncover their molecular evolution, trees were constructed using the "neighbor-joining" (NJ) and "maximum parsimony" (MP) approaches. The variation in GC content in the ITS1 region was recorded at 52–53%. The data suggests that transitions have a lower frequency compared to transversions in this region, where the overall bias R for transition/transversion has been estimated at 0.751. The aligned sequences of the cultivars were identified to be 637 bp, and the divergence values were observed to range between (0.0000%) and (0.0063%). Out of the nine cultivars, the aligned sequences permitted us to pinpoint eight haplotypes. The outcomes refer to a gradual demographic expansion for an Iraqi date palm population that will arrive at equilibrium after 0.6N generations. The research findings revealed that the Mir alhajj TC cultivar exhibited the highest susceptibility and sensitivity among the examined cultivars when propagated through the tissue culture technique. The primary 2 constituents' multi-variate PCA analysis revealed that PC1 had a variability of 0.9599 among cultivars and an Eigen value of 7.7099, whereas PC2 had a variation of 0.9766 and an Eigen value of 6.7500. Analysis of these ITS1 sequences revealed that they offer adequate variability in characteristics to investigate the relationships among genotypes and explore the evolutionary lineage between the cultivars studied. Moreover, the observed diversity among closely related genotypes provides robust evidence for the efficacy of ITS1 sequences in figuring out relationships and identifying the evolutionary history of those cultivars.

Keywords. Iraqi date palm, sequences Analysis, molecular evolution, ITS, tissue culture, (*Phoenix dactylifera L.*)

تحليل التتابع والتطور الجزيئي في أصناف نخيل التمر العراقي (*Phoenix dactylifera L.*)

المتكاثرة بزراعة الأنسجة على أساس منطقة ITS1

محمد مخلف هواش , ليث محمد جواد الشماع

قسم علوم الحياة، كلية العلوم، جامعة بغداد، بغداد، العراق

*Email: laith.alshamma@sc.uobaghdad.edu.iq

الخلاصة

تتضمن الدراسة سبعة أصناف من نخيل التمر العراقي (*Phoenix dactylifera* L.) تم تكاثرها عن طريق زراعة الأنسجة (TC) وصنفيين تم تكاثرهما عن طريق الفسائل (Offs), على أساس منطقة فاصل النسخ الداخلي (ITS1) من الريبوسوم النووي (DNA) كمؤشر للتحقيق, في محاولة لإقامة علاقات وراثية بين أصناف نخيل التمر المختلفة والكشف عن تطورها الجزيئي, تم إنشاء أشجار التطور باستخدام أسلوب (NJ) و (MP), إذ تم تسجيل التباين في محتوى GC في تلك المنطقة بنسبة 52-53% و تشير البيانات إلى أن الانتقالات أقل تردد مقارنة بالتبدلات في هذه المنطقة, حيث تم تقدير التحيز الإجمالي R للانتقال/التبدال كان 0.751, كما تم تحديد تسلسل المحاذاة للأصناف ليكون 637 زوج قاعدي, ولوحظ أن قيم المسافة تراوحت بين (0.0000%) و (0.0063%), ومن بين الأصناف التسعة, سمحت لنا المحاذاة بتحديد ثمانية أنماط فردية, وتشير النتائج إلى توسع ديموغرافي تدريجي لسكان نخيل التمر العراقي الذي سيصل إلى التوازن بعد 0.6 جيل, كما أظهرت النتائج أن الصنف مير الحاج TC أظهر أعلى حساسية للزراعة النسيجية من بين الأصناف المدروسة, كما كشف تحليل PCA متعدد المتغيرات للمكونين الأساسيين أن PC1 كان له تباين قدره 0.9599 بين الأصناف وقيمة Eigen تبلغ 7.7099, في حين كان لدى PC2 تباين قدره 0.9766 وقيمة Eigen تبلغ 6.7500, وقد كشف التحليل ل ITS1 أنها توفر تبايناً كافياً في الخصائص لدراسة العلاقات بين الأنماط الجينية واستكشاف النسب التطوري بين الأصناف التي تمت دراستها, فضلاً عن التنوع الملحوظ بين الأنماط الجينية ذات الصلة الوثيقة والذي يوفر دليلاً كافياً على فعالية منطقة ال ITS1 في اكتشاف العلاقات وتحديد التاريخ التطوري لتلك الأصناف .

1. Introduction

The date palm, scientifically known as *Phoenix dactylifera* L., holds the distinction of being one of the most ancient cultivated fruit-bearing trees. Its significance in the historical region of Babylon, Iraq, dates back to approximately 4000 B.C. At this time, these palm trees were highly respected and celebrated for their durability and majesty. The date palm is one of the world's oldest cultivars of trees and belongs to the genus *Phoenix*. The Phoenician "*Phoenix*," which means date palm, and "*dactylifera*," which comes from the Greek word "daktulos," meaning a finger, are combined to give the botanical name (*Phoenix dactylifera* L.) for the date palm [1]. Oasis agriculture is believed to have flourished primarily during the earlier Bronze Age, about 3000 B.C., even though the oldest evidence of date palm planting in southern Mesopotamia belongs to the third millennium B.C. [2]. As well, the date palm also holds cultural significance for human beings in different nations of the world, especially in dry regions such as North Africa. Although its exact origin is not always acknowledged, evidence suggests that it probably happened near Iraq [2]. In addition to its cultural importance, the date palm plays a crucial role in oasis agriculture, contributing to the economic and ecological well-being of arid and semi-arid regions. There is proof that it was once used as a food source and a medication [3].

Date palm is a significant tree crop in arid regions, contributing significantly to the economies of many of these nations. The most common technique used for propagating date palms involves offshoots. However, the advent of tissue culture propagation techniques led to a significant expansion of date palm farms. While most tissue culture-grown trees are healthy and maintain their genetic identity, some aberrant phenotypes are observed [4]. The Internal Transcribed Spacer (ITS) region exhibits unique banding patterns [5] among the cultivars, surpassing the discriminatory power of other molecular markers. However, a low degree of genetic diversity is observed due to the similarity in banding patterns among offspring or plants propagated through tissue culture. The significance of the ITS region stems from its relatively high rate of nucleotide substitution, enabling the systematic comparison of cultivars

that have diverged relatively recently [6]. Therefore, ITS region band patterns were used to identify several phenotypically abnormal trees. Additionally, recent genetic variation can be observed in variegated trees propagated through tissue culture.

Comparatively less genetic diversity exists in date palms than in their perennial monocotyledon relatives. Despite previous findings indicating various chromosomes ($2n = 2x = 36$) linked to tissue culture multiplication and the innate cultivar type [7], the date palm is diploid. Duplication has been noted in conventional collections of genotypes, leading to a mere 10% of the total number of cultivars globally being classified as unique and commercially significant [8]. The marketing of preferred date palm cultivars has resulted in a significant reduction in genetic diversity, ultimately promoting monoculture through increased cultivation areas. High-density genomics techniques can address the urgent need for characterising, conserving, and utilising international date palm collections [9].

Iraq has long been recognized as the world's leading producer of dates, with a rich history of cultivating this crop. However, the country has faced a series of conflicts that have negatively impacted both the yield and genetic diversity of the date palm. These factors have significantly impeded Iraq's capacity to maintain its top-tier status and role as a key producer of dates. Nevertheless, Iraqi researchers and authorities are currently engaged in coordinated endeavours to mitigate the harm inflicted upon the crop during the preceding three decades [10]. Therefore, endeavours are directed towards the safeguarding of the Iraqi date palm gene pool.

Numerous investigations were performed with the aim of identifying Iraqi date palm cultivars using morphological traits and molecular markers. The analysis of DNA markers in Iraqi date palms is still in its developmental phase. Jubrael [11] pioneered using DNA markers to distinguish between cultivars at the (IPA) Centre for Agriculture Research in Baghdad; nine female cultivars have been identified using Random Amplified Polymorphic DNA (RAPD) markers, whereas eight male cultivars were discovered using the same approach [12]. Additionally, the genetic variability among 30 cultivars of Iraqi date palm was evaluated using microsatellite markers [13]. With the use of eight Iraqi date palm cultivars, the researcher examined the performance of over 1,000 simple sequence repeat (SSR) pairs of primers created by Hamwiah *et al.* [14], utilising genome sequencing data for this important crop. Inter simple sequence repeat (ISSR) markers were also used to assess the genetic relationships between 17 cultivars of Iraqi date palms [15]. A variety of molecular markers and other markers from these 17 cultivars collectively contributed insights to this study aimed at understanding their genetic associations. Successful characterization of Iraqi date palm genotypes has been achieved. These markers have proven to be somewhat suitable for identifying date palm varieties.

Given the potential constraints on phylogenetic interpretations resulting from the existence of polymorphisms, it is imperative to conduct a thorough investigation of this situation. Over the course of ten years, the utilisation of nuclear ribosomal DNA (nrDNA) sequences obtained from the ITS region in most investigations due to high conservation in this region, including ITS 1, 5.8S, and ITS 2, has served as the primary source of data for investigating the lower-level evolutionary relationships between plant species [16]. In preliminary investigations into ITS, variation yielded outcomes that were in accordance with the presence of uniform nrDNA arrays within individuals, as observed by Baldwin [16]; Ainouche, and Bayer [17]. This pattern was likely due to concerted evolution, which involves gene conversion and unequal crossing [18]. Nevertheless, certain studies have shown that a range of species exhibit intra-individual nrDNA variation. Although the rRNA genes were initially believed to be quite stable, discovery in the 1980s revealed that the 5' and 3' external

transcribed sequence ETS, non-transcribed spacer NTS, and ITS1 and 2 regions demonstrate a significant degree of intra- and interspecies variability. As a result, they are exceedingly valuable as markers in our phylogenetic and molecular evolution research at the population or interspecies levels. However, further studies are required to investigate the genetic variation and molecular features of different cultivars of date palm.

In this study, we have investigated the internal transcribed spacer 1 (ITS1) region of the nuclear ribosomal DNA (nrDNA) in Iraqi date palm cultivars. The aim was to differentiate among cultivars, some of which were propagated using tissue culture (TC) techniques, while others were propagated through offshoots (Offs). This highlights the utmost importance of preserving diversity to protect species that possess valuable traits. Therefore, we dedicated significant effort at the beginning of our investigation to carefully selecting superior cultivars with exceptional characteristics, specifically focusing on the most desirable among all the cultivars in Iraq. This was done in order to understand their genetic features and variations, with the aim of improving and perpetuating them

2. Materials and methods

2.1. Cultivars of the Date Palms

This study compared the genetic structure among nine important date palm cultivars overall. Seven of the cultivars were propagated via tissue culture (tc), while the other two were propagated from offshoots (offs). These nine significant cultivars were selected for analysis in this study. These cultivars were chosen based on their commercial value, with a focus on consumer preferences and fruit market demands. The original Arabian nomenclatures of date palm cultivars have led to a variety of spelling and pronunciation changes; for the sake of this study, the most often used spelling was employed. "Barhi tc", "Khalas tc", "Majhool tc", "Majhool offs", "Maktomi tc", "Mir alhajj tc", "Mir alhajj offs", "Showaithy tc", and Um alduhan tc were the chosen fruiting cultivars, listed in alphabetical order.

2.2. Plant materials

Samples of contemporary leaflets were collected from a range of sources for each of the nine chosen cultivars. These samples were derived from four different locations in Baghdad and Al-Anbar (table 1). For each of the nine cultivars, five leaves measuring 10 to 15 cm were randomly selected. Only healthy, well-established, and well-characterized trees that were between 10 and 15 years old were used as the source of these samples. To determine which cultivar these chosen trees belonged to, experts from the stations of date palm from the Ministry of Agriculture, Baghdad University, and Janet Al-Nakheel Company Laboratory in Iraq provided assistance. The collection of samples took place during the summer growth season, precisely in the middle of September 2022.

2.3. DNA Extraction

Total genomic DNA was isolated from the whole young leaflets, with a size of 20–100 mg, that were purified from impurities and frozen under liquid nitrogen using the FavorPrep™ Plant Genomic DNA Extraction Mini Kit (Korea).

2.4. Electrophoresis of DNA sequence on agarose gel

The fundamental concept of agarose gel electrophoresis encompasses the process of segregating nucleic acids based on their charge and size. In this methodology, the electrophoresis gel is treated with a red-light compatible staining solution to visualize the separated DNA fragments. Subsequently, the gels underwent the process of being exposed to

70 volts and 65 amperes for an hour. The DNA was observed by viewing it under a UV transilluminator in 1X TBE buffer [19].

2.5. Amplifying PCR and the (nr ITS1) region

A PCR premix kit (Intron/Korea) was used for PCR amplification, which contained 2.5U I-Taq DNA Polymerase, 2.5 mM DNTPs, 1X reaction buffer (10X), and 1X gel loading buffer. The PCR amplification mixture of the specific reaction for diagnosis gene was performed in a total volume of 25 μ l containing 1.5 μ l of DNA, 5 μ l Taq PCR PreMix, and 1 μ l of each primer (10 pmol), then 16.5 μ l of distilled water was added into the tube for a total volume of 25 μ l. The primers used for the amplification of PCR reactions were performed using universal ITS5 (5'-ATGATAACTCGACGGACCGC -3') and ITS2 (5' TCTTCGAGCCCCCAACTTTC -3') primers, designed by [20]. The following procedures for thermal cycling were used: Using a thermal cycler (Gene Amp, PCR system 9700; Applied Biosystem), the reaction was first denaturated at 95 °C for 5 min, then it went through 35 cycles of 95 °C for 45 sec., 57 °C for 1 min., and 72 °C for 1 min., with the final extension 2 occurring at 72 °C for 5 min. Following staining with red safe dye (Intron Korea), all PCR products were separated using 1.5% agarose gel electrophoresis and visualised under ultraviolet light (302 nm). Sequences were compared with the GenBank database using the Basic Local Alignment Search Tool (BLAST) at the National Centre for Biotechnology Information (NCBI) site (<http://www.ncbi.nih.gov>).

2.6. Sequence analysis

The sequences of the nine annotated accessions were submitted to the National Centre for Biotechnology Information (NCBI) Gen Bank (Accessions numbers: OQ911644.1, OQ911646.1, and from OQ911648.1 to OQ911655.1 for the small subunits of 18S nrDNA of the ITS region; Table 1). MEGA version 11.0.13 [21] was used to evaluate the nucleotide sequences after they had been aligned using the DAMBE programme version 6 [22]. The verification of alignment was conducted manually, and an estimation of pairwise sequence divergence among cultivars in the ITS1 region was made utilising a "Maximum Composite Likelihood" method, as outlined by Tamura, *et al.* [23]. GC content was calculated for each sequence online by Biologics Corp. at (<https://www.com/tools/GCcontent/>).

The distance matrix that ensued was subsequently calculated in order to produce phylogenetic trees utilising the "Neighbour-Joining" (NJ) method, as according to [24]. Utilising 1000 bootstrap replications, the "Neighbour-Joining" and "Maximum Parsimony" (MP) trees were created. Instances that included gaps and mismatches were removed from the dataset via the Complete Deletion option. Calculated the consistency indexes (CI), retention indexes (RI), and homoplasy index (HI) [25]. The equation $R = [A * G * k1 + T * C * k2] / [(A + G) * (T + C)]$ was used to estimate the transition/transversion ratio t_i/t_v , wherein A, G, C, and T indicate the frequency distribution of the four nucleotides [23]. The analysis provides the number of substitutions per site for nucleotides observed across the sequences.

Indices of polymorphism and demographic history were determined through analysis of the aligned sequences using DnaSP software, 5.10.01, developed by Librado, and Rozas [26]. To quantify the genetic diversity among cultivars within the ITS sequences, the nucleotide diversity indices (Pi) [27] and the haplotype diversity indices (Hd) [28], along with their respective standard deviations (SD), were employed. The evaluated average (pairwise nucleotide differences) (K) for selective neutrality was evaluated by means of Tajima's D statistic [29] as well as Fu and Li's for D* and F* methods developed by [30]. However, to evaluate the demographic parameters using the distribution of pairwise sequence differences

(mismatch distribution) of site [31] frequency spectra (distribution of the allelic frequency at a site) of [29], in addition to population size change and population time, The $S_n(t)$ and value of $S_n(t)/a1$ for the anticipated number of segregating sites among 10 sequences were derived. $S_2(t)$, the expected number of sites separating two sequences, and $S_2(t)/a1$ represent the average number of pairwise differences. Thus, the expected number of time generations to be in equilibrium was estimated. The multi-variate PCA (principal component analysis) scatter chart plot has been included using DAMBE version 6 [22]. As well, the graphical display of the genetic relationship among the cultivars of date palms that were identified by the detection of haplotypes was carried out using the (NETWORK) programme 4.6.1.0, as described by. Bandelt, *et al.* [32].

3. Results and Discussions

3.1. Genetic diversity analysis

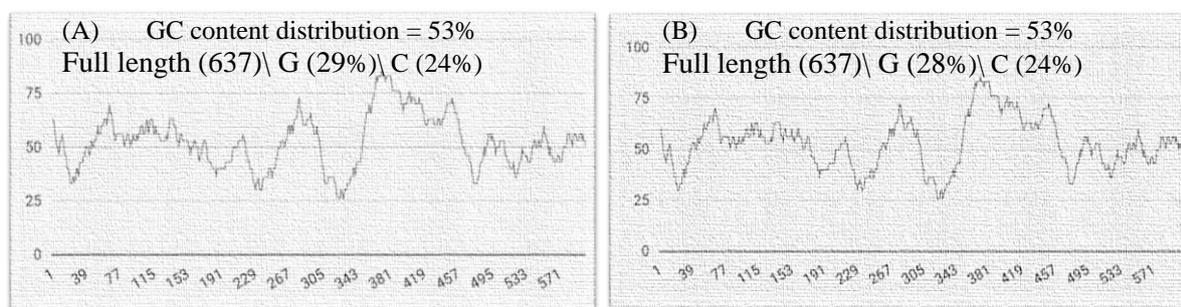
3.1.1. length of sequence and GC content

The amplified segment, with a stretch of approximately 600 bp, corresponds to the internal transcribed spacer 1 (ITS1) region. The BLAST results authenticate and confirm the symmetry of the sequences as ITS regions in the palm cultivars. The length estimation is not consistent with previous findings by Mainaa *et al.* [5] that the average length of the Tunisian date palm is 442.7 bp. While the length estimate is consistent with prior studies that have been reported for other species, including members of the Asteraceae family, where the entire ITS region's length discrepancy varied from 650 to 750 bp [33] and 600 bp for *Quercus* species by Bellarosa *et al.* [34], the ITS region generated from the species *Ficus carica*, on the other hand, has an average length of 697.5 bp, as discovered by Baraket *et al.* [6]. Kehie *et al.* conducted a comprehensive analysis of the complete ITS sequence from the Naga King Chilli, revealing an average length of 620 bp [35]. Similar investigations in different plant taxa, such as *Cucurbita pepo* at 187 bp and *Cucumis sativus* at 229 bp, have been reported [36]. As well as differences in wheat (597 to 605 bp) and barley (595 to 598 bp) length [37].

In fact, the range of GC content percentages on record in the ITS1 region was 53%–52%, with an average of 52.7 for all cultivars studied. While Cytosine remained consistent across all cultivars studied, it appears that the difference in GC content may be totally attributable to the fluctuation of the nitrogenous base Guanine. (see figs. 1, A, and B). Actually, it appears that the date palm's rDNA didn't undergo methylation, as in the Cucurbitaceae family, as reported by Hemleben, *et al.* [38]. It is noteworthy that the GC content obtained from this study may be compared to that determined in Tunisian date palms, which varied between 49% and 49.7% across an entire ITS region [5]. Consistent with various other plant species belonging to the Asteraceae family, the average GC content across the entire ITS region was found to be 51% [33]. Additionally, the higher level observed in *Ficus carica* L. with the GC content based on regions of ITS1, ITS2, and (5.8S genes) varied from 53.4 to 68, 52.3 to 67.7, and 46.7 to 56.6, respectively [6]. Notably, the depressed variation in the GC content of ITS1 indicates homogeneity among various cultivars, corroborating the high conservation grade of Iraqi date palms. However, adaptation to the environment makes it evident that ITS are subject to selective pressure and temperature.

Table 1: Iraqi date palm samples accessions, query number, number and type of substitution sequence, and percentage of identical

Cultivar	Type of substitution	Location	Nucleotide	Sequence ID with Submission	Sequence ID with compare	Identities
Majhool offs	Transversion	439	C\A	ID:OQ911644.1	ID: XR_005510437.1	99%
	Transversion	455	G\T			
Mir alhajj offs	Transition	562	C\T	ID:OQ911646.1	ID: XR_005510437.1	99%
	Transition	677	A\G			
Showaithy tc	Transversion	298	C\A	ID:OQ911649.1	ID: XR_005510437.1	99%
	Transition	337	A\G			
	Transversion	608	A\T			
Majhool tc	Transversion	735	T\G	ID:OQ911650.1	ID: XR_005510437.1	99%
	Transversion	784	C\A			
Barhi tc	-----	-----	-----	ID:OQ911651.1	ID: XR_005510437.1	100%
Mir alhajj tc	Transversion	304	A\T	ID:OQ911652.1	ID: XR_005510437.1	99%
	Transition	423	G\A			
	Transition	562	C\T			
Um alduhan tc	Transition	677	A\G	ID:OQ911653.1	ID: XR_005510437.1	99%
	Transversion	645	G\T			
	Transition	694	T\C			
Maktomi tc	Transversion	349	T\A	ID:OQ911654.1	ID: XR_005510437.1	99%
	Transition	350	G\A			
	Transversion	406	C\G			
Khalas tc	-----	-----	-----	ID:OQ911655.1	ID: XR_005510437.1	100%



Figures 1: show percentage and variation of GC distribution among rDNA sequences of Iraqi date palm cultivars. A=53% \ B=52%

3.1.2. Variability of Nucleotide composition and mutations

Purine bases had a transition/transversion ratio of $K1 = 1.809$, while pyrimidine bases had $K2 = 1.12$. For all bases, a total transition/transversion ratio (R) of 0.751 was found. According to Baraket *et al.* [6], the transition/transversion ratio R observed in the ITS1 area is equal to the value discovered after analysing the complete ITS region in *Ficus carica* of 0.7, and in addition to *Lolium* and *Fescue* complex accessions, R was 0.94 [39] but less than those found in wild barley (7.4) and wheat (6.90) [37]. As well, the Asteraceae family (ti/tv) was 1.43 [33] and *Capsicum* sp. was 3.746 [35]. The various substitutions found are listed in Table 2, which demonstrates that transitions rather than transversions are more common in Iraqi date palm RNA. The transitions G\A and C\T occur more frequently than A\G and T\C transitions, while transversions C\A and T\A are more recurrent than other transversions, as shown in Table 1. The outcome presented is discordant with the findings of [5] in the Tunisian date palm due to transitions occurring with greater frequency compared to

transversions, which exhibit a contrary trend. The observed genetic differences could have arisen from stress exposure of the callus tissue during cell culture conditions, or may have had an epigenetic basis, potentially leading to an elevated occurrence of uncommon nucleotide transversions between purine and pyrimidine bases. However, the Mir alhaji cultivar exhibits a high degree of sensitivity to chemical materials used during tissue culture propagation.

The diversity in the constitution of the nucleotide sequence alignment of the Internal Transcribed Spacer 1 (ITS1) restricted within a 637 (bp) character matrix has demonstrated the existence of 621 preserved sites in conjunction with 16 variable sites, composed of 14 singleton sites and 2 informative sites (parsimony) at positions 289 and 404. The ITS sequences have nucleotide frequencies that are adenine (A): 0.238 = 23.88%; thymine (T): 0.229 = 22.97%; cytosine (C): 0.242 = 24.25%; and guanine (G): 0.289 = 28.90%, while in the Asteraceae family are (25%), (24%), (26%), and (25%) for A, T, C, and G, respectively, with an average GC of 51%, as reported by Amar, *et al.* [33]. Whereas the nucleotide frequencies of Tunisian date palm cultivars were A of 0.247, T of 0.256, C of 0.276, and G of 0.219 [5]. On the other hand, [6] reported that the fundamental compositions in *Ficus carica* for A, T, C, and G were 19.7%, 18.6%, 31.4%, and 30.2%, respectively. Whereas [35] evaluated the nucleotide frequencies in "Naga King Chilli" and observed that the frequencies for adenine, thymine, cytosine, and guanine were respectively 18.85%, 17.56%, 33.95%, and 29.64%.

Table 2: nucleotide substitution rates inferred from the nrDNA ITS1 region

	A	T	C	G
A	-	6.59	6.95	14.24
T	6.85	-	8.56	8.29
C	6.85	8.11	-	8.29
G	11.76	6.59	6.95	-

NOTE:* Each entry in the provided data represents the probability of substitution (r) from one base (row) to another base (column). Transitional substitutions are denoted in bold, while transversal substitutions are shown in *italics*.

The ITS1 sequences of the nr DNA provided evidence of a genetic modification. The study of nine cultivars revealed the presence of eight haplotypes. The diversity of haplotypes (Hd) and nucleotides (p), along with their standard deviation, are shown in Table 3. Calculation was performed on the sequences, resulting in the determination of estimates of $0.972 \pm \text{SD } 0.064$ and $0.00611 \pm \text{SD } 0.00231$, respectively. The calculation of pairwise nucleotide differences (K) yielded a value of 3.889, indicating a relatively high level of polymorphism compared to findings from other researchers. However, the ITS region displayed a limited degree of genetic diversity. Conversely, *Ficus carica* at 35.34 and *Capiscum* sp. at 9.267 recorded noteworthy average pairwise nucleotide differences (K), suggesting that these specific species' ITS genomes have significant levels of polymorphism [6, 35]. However, it must be noted that the values reported by Mainaa, *et al.* [5] (0.686) were comparatively smaller.

The analysis of genetic distances utilising the "Maximum Likelihood Composite (MCL)" method resulted in a range of values from 0.0000 to 0.0062, with a mean value of 0.0027 (Fig. 2). Notably, no distance of (0.0000) was observed for the "Barhi tc" and "Khalas tc" cultivars, indicating a remarkable degree of similarity in their ITS sequences. Furthermore, a significant level of resemblance was found among the "Majhool offs", "Majhool tc", and "Um alduhan tc" cultivars, with a distance of 0.0031, whereas a distance of 0.0047 was recorded

between the "Showaithy tc", "Maktomi tc", and "Mir alhajj offs" cultivars. However, the "Mir alhajj tc" gene displayed the greatest distance, as evidenced by the higher value of 0.0063, indicating its marked dissimilarity from other cultivars.

Table 3: tests for neutrality and polymorphism based on the ITS1 region of Iraqi palm cultivars

<i>Phoenix dactylifera</i> sequences	The value
No. of sequences	9
Length of alignment	637 bp
monomorphic patterns	621
Variable characters	16
Singleton variable sites	14
parsimony	2
H	8
Variance of Hd	0.00409
Pi ± SD	0.0061 ± 0.0023
Hd ± SD	0.972 ± 0.064
Pi(JC)	0.00613
Theta (per site) from Eta	0.00924
K	3.889
R2 statistic	0.0768
"Sn(t)"	0.891
value of "Sn(t)"	0.328
"S2(t)"	0.259
average of "S2(t)/a1"	0.259
Tajima's D	-1.64543 (0.10 > P > 0.05)
Fu and Li's D	-1.71785 (P > 0.10)
Fu and Li's F	-1.90378 (P > 0.10)
Fu's Fs statistic	-3.471
PCA1	variance 0.9599% with an Eigen value 7.7099
PCA2	variance 0.9766% with an Eigen value 6.7500

Phylogenetic trees are constructed through the utilisation of both "Maximum Parsimony (MP)" and "Neighbour-Joining (NJ)" methodologies. The parsimony analysis uncovered 39 trees, resulting in the most parsimonious tree at 16 steps in length. The subsequent evaluation of this tree produced a "consistency index (CI)" value of 1.000, a "retention index (RI)" value of 1.000, and a "homoplasy index (HI)" of 0.000, thus signifying the non-existence of homoplastic characters within the analysed sequences.

3.2.2. Mismatch distribution

The curve of distribution of date palm mismatches over the entire dataset indicates that the population size of date palms has been slowly expanded. Moreover, Harpending's raggedness [31] and the Ramos-Onsins statistic [41] ($r = 0.1134$ and $R2 = 0.0768$) have supported these findings (Figs. 4, A and B). In addition, the negative worth of Fu's F_s does not show that the cultivars of date palm in Iraq have recently had a significant population expansion (Table 3). As depicted in Fig. 3, the mismatch distribution curves do not seem to align with the notion of a recent and rapid demographic expansion of these sequences. Instead, the ITS1 sequences appear to be indicative of an older evolutionary trajectory.

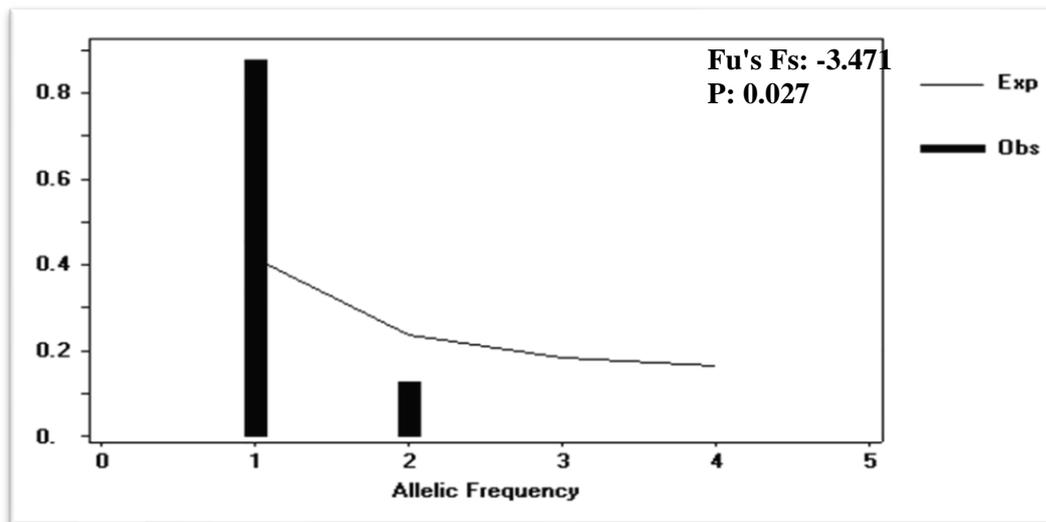


Figure 3: displays the frequency spectrum of rDNA sequences discovered in date palm at ITS1 sites. The distributions determined for neutrality and balance (mutation drift) appear in the spectrum as solid lines.

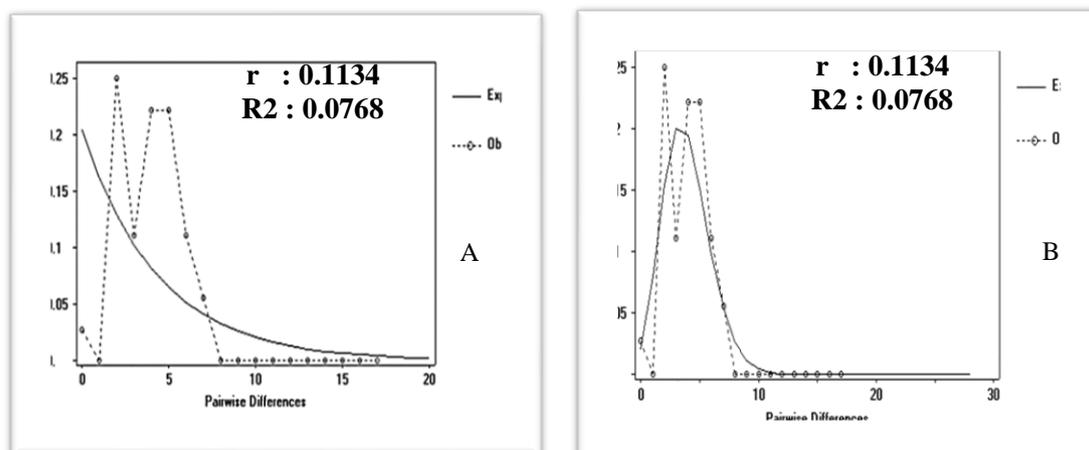


Figure 4: Mismatch distribution of palm population. (A) displaying the observed distribution expected values for Population Size Changes(PSC) with an initial theta of 0.000 and a final tau of 3.889. and (B) have a start theta of 0.000, a final theta of 1000, and an ultimate tau of 3.889.

3.2.3. Population Size Change of the date palm

Between the nine sequences, $S_n(t)$, the anticipated number of segregating sites was 0.891, with a $S_n(t)/\alpha l$ value of 0.328. $S_2(t)$, the anticipated average number of pairwise differences

among two sequences, was determined to be 0.259, and an average for $S_2(t)/a_1$ of 0.259. The nine cultivars' segregating sites are expected to be in equilibrium with those expected in a population after $0.6N$ generations. (Figure 5). In contrast to the conclusion by Kitavi [42],

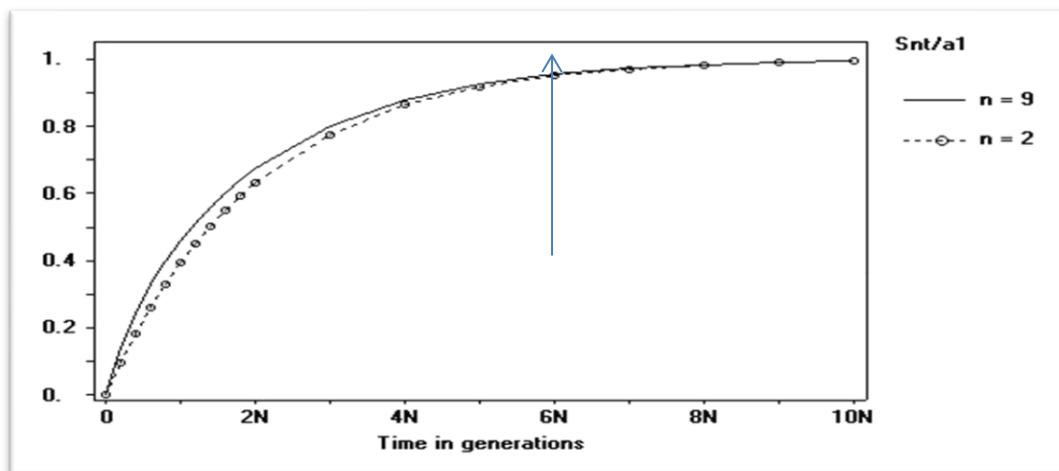


Figure 5: Mismatch distribution of palm population. showing the projected distribution and measured distribution of pairwise shifts for a population expansion model with start theta 0.000, final theta 1000, final tau 3.889, and time $0.6N$ generations.

when examining the African bananas in a community that stayed in balance up until the ninth generation, the number of segregating sites among the 89 varieties was greater than anticipated. The $S_n(t)$ value was 2.713, the $S_n(t)/a_1$ quotient was 0.536, the $S_2(t)$ value was assessed at 0.259, and the average $S_2(t)/a_1$ value was 0.25. Whereas Tajima, [43] demonstrated that the rate at which variation increases is relatively slow, particularly when $n = 2$. Specifically, it requires $1.4N$ generations for this figure to halve from its maximum value. Conversely, when $n = 100$, this process takes only $0.5N$ generations. Furthermore, the researcher suggested that larger sample sizes lead to the number of segregating sites increasing at a faster rate.

3.2.4. Distribution of haplotypes according to ITS1 sequencing

The haplotype network, as depicted in Figure 6, which is according to the internal transcribed spacer1 sequences of the nr DNA, reveals a conspicuous evolution of Iraqi date palms from an inherited haplotype. The network of relationships showed interconnections, as reflected by the central positioning of the founder haplotype structure represented by the "Barhi tc" cultivar. This network structure underscores the genetic cohesion among these date palm varieties. Consequently, it can be inferred that the "Barhi tc" cultivar represents an ancestral for other cultivars.

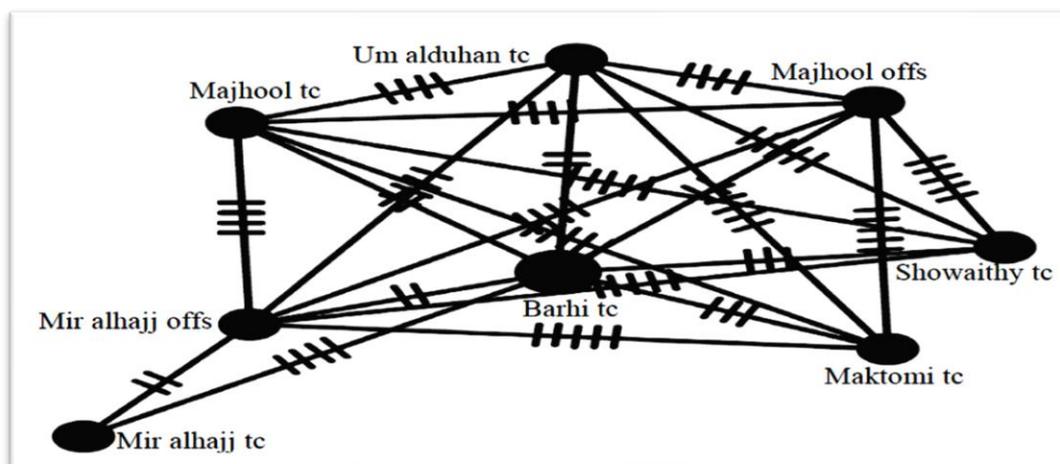


Figure 6: the ITS1 sequences' indicated network of haplotypes, shows interconnectedness among the cultivars. The hatch marks are proportional to the number of mutations.

The haplotype serves as the ancestor of the residual sequences that have undergone changes over the course of evolution. The substitutes for the Barhi tc cultivar were enumerated in the following manner: four with "Mir alhajj tc", three with "Maktomi tc" and "Showaithy tc" cultivars, and two with "Majhool tc", "Majhool offs", "Um alduhan tc", and "Mir alhajj tc". Additionally, the analysis also illustrates a network connecting the haplotype "Mir alhajj tc" cultivar with the "Mir alhajj offs" cultivar, with other branches indicating the occurrence of numerous substitutions. The five substitutions were taken between "Maktomi tc" and each of "Majhool tc", "Mir alhajj offs", and "Um alduhan tc", as well as between "Showaithy tc" and both of "Majhool tc", "Majhool offs", "Mir alhajj offs", and "Um alduhan offs". Moreover, the four substitutions comprised a rotation between "Um alduhan tc" and both of "Majhool tc", "Majhool offs", and "Mir alhajj offs", and between "Majhool offs" and both of "Majhool tc" and "Mir alhajj offs", and between "Mir alhajj tc" and "Mir alhajj offs". The hatch marks in Figure 6 correspond to the number of mutational occurrences, reflecting the antiquity of the date palm trees in Iraq.

3.2.5. principal component analysis (PCA)

The scatter plot obtained from the principal component analysis (PCA) of the first two components, PC1 and PC2, derived from the ITS1 region showed that PC1 accounted for a variance of 0.9599 among the cultivars and had an eigenvalue of 7.7099. In contrast, PC2 exhibited a variance of 0.9766 and an eigenvalue of 6.7500. Similar to the investigation of Moroccan date palms conducted by Ibrahimi, *et al.* [44], the analysis of the combined data from SSR and DAMD markers revealed that the first three axes were (7.3%), (5.75%), and (4.38%), respectively. While [45] utilised morphological markers to identify a variation of 31.86% with an Eigen value of 4.67 for PC1. Additionally, at an Eigen value of 4.63, PC2 exhibited a variance of 17.93%. As with the other analyses, no clustering based on collection location or propagation methods was observed. Similarly, groupings similar to those identified in the phylogenetic tree were apparent in the outcomes of the PCA scatter diagram. Although the Iraqi date palm varieties have been widely cultivated for a long time, the use of principal component analysis (PCA) has demonstrated that the cultivars possess diverse relationships among themselves. Furthermore, the dendrogram's topology and the cultivars' distribution in the PCA analysis have shown that continuous morphological alternation with more frequency than genetic diversity usually characterises the germplasm of Iraqi date palm (Fig. 7). This outcome is in line with the NJ phylogenetic tree in Figure 2 and the haplotype distribution in Figure 6.

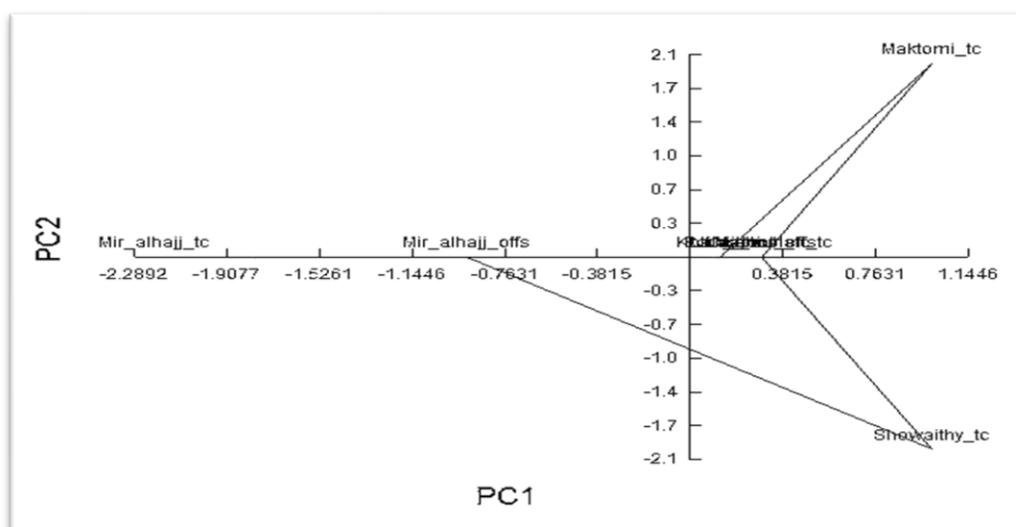


Figure 7: Analysing the 9 date palm cultivars based on ITS1 region using PCA

The results of our study indicate the existence of a common genetic basis across different date palm genotypes, albeit with some molecular variations. This study's results are close to aligning with previous papers analysing the Tunisian date palm [5] and Moroccan date palm [44]. The ITS investigated in this particular study exhibited lower levels of polymorphism patterns in comparison to other researchers that have employed different molecular tools on Iraqi date palms, as demonstrated by [11, 15, 46] utilizing "RAPD" markers, [13, 47-48] ALFP markers, and [13-15] using microsatellite. Other studies have focused on Iraqi date palm characteristics [49-50].

Conclusion

All the conclusions derived from our inquiry indicate that all the examined cultivars are descended from a shared progenitor, despite the presence of specific genetic differences. Especially when it concerns cultivars propagated through tissue culture, there is evidence that contradicts assertions suggesting that tissue culture will give genetically identical individuals. However, the findings suggest that tissue culture might result in additional genetic variety, which would then manifest itself in the phenotypic and productive traits of palm plants. Notably, the ITS technique, especially with plants created through tissue culture, is a useful tool with easy and rapid DNA markers for early detection of genetic variants. Moreover, it is crucial to gather molecular information regarding every cultivar cultivated in Iraq using multiple genetic markers. Therefore, a thorough understanding of the extent and distribution of genetic diversity is crucial for assessing the genetic predisposition of a cultivar. While the long-life cycle of the date palm poses a significant challenge for genetic studies and conventional breeding efforts, a more robust assessment of the existing genetic diversity is still needed. The slow lifecycle makes genetics research difficult, emphasizing the need for more thorough characterization of available genetic variation. In order to gain a comprehensive understanding of Iraqi palm genetics, the genetic sequences of these cultivars were recorded in the NCBI Gene Bank.

Acknowledgements

This work was supported by the Department of Biology Graduate Studies, College of Science, University of Baghdad. We would like to thank Dr. Ayyad W. Alshahwany, for offering aid in acquiring specimens and for contributing labour and direction.

Reference

- [1] Linne, Cited in T. H. Keaney, "Date varieties and date cultures in Tunisia," Washington, U.S.D.A.: "Bureau of Plant Industry". Bulletin no. 92, 1734, 1906.
- [2] M. Tengberg, "Beginnings and early history of date palm garden cultivation in the Middle East," *Journal of Arid Environments*, vol. 86, pp: 139-147, 2012.
- [3] C.P. Khare, "Indian Medicinal Plants—An Illustrated Dictionary". First Indian Reprint, Springer (India) Pvt. Ltd., New Delhi, 717-718, 2007.
- [4] A. El Hadrami, F. Daayf, and I. El Hadrami, "Date palm genetics and breeding," In: S.M. Jain, J.M. Al-Khayri and D.V. Johnson (Eds.), pp: 479-512, 2011. "Date Palm Biotechnology," Springer, Dordrecht.
- [5] N. Mainaa, G. Baraketa, A. Salhi-Hannachia and H.B. Sakkaa, "Sequence analysis and molecular evolution of Tunisian date palm cultivars (*Phoenix dactylifera* L.) based on the internal transcribed spacers (ITSs) region of the nuclear ribosomal DNA," *Scientia Horticulturae*, vol. 247, pp: 373–379, 2019.
- [6] G. Baraket, A. Ben Abdelkrim, M. Mars, and A. Salhi-Hannachi, "Genetic diversity and molecular evolution of the internal transcribed spacer (ITSs) of nuclear ribosomal DNA in the Tunisian fig cultivars (*Ficus carica* L.; Moraceae)," *Biochem. Syst. Ecol.*, vol. 48, pp: 20–33, 2013.
- [7] W. Ahmed, T. Feyissa, K. Tesfaye, and S. Farrakh, "Genetic diversity and population structure of date palms (*Phoenix dactylifera* L.) in Ethiopia using microsatellite markers". *Journal of Genetic Engineering and Biotechnology*, 19(1): 64, 2021.
- [8] N. Haider, " Phylogenetic relationships among date palm (*Phoenix dactylifera* L.) cultivars". Medicinal and aromatic plants of the world, eBooks Library - UNESCO-©Encyclopedia of Life Support Systems (EOLSS), 44p; <https://www.eolss.net/Sample-Chapters/C03/E6-79a-23.pdf>.
- [9] A. Jaradat, "Date palm genetics and breeding." In: S.M. Jain, J.M. Al-Khayri and D.V. Johnson (Eds.), pp: 19–71, 2011. "Date Palm Biotechnology," Springer, Dordrecht. J. M. Al-Khayri, S.M. Jain, and D.V. Johnson, "Date palm genetic resources and utilization," In: Afr. and Ame., J.M. Al-Khayri, S.M. Jain and D.V. Johnson, (eds) Springer, Dordrecht, 2015.
- [10] J. M. Al-Khayri, S.M. Jain, and D.V. Johnson, "Date palm genetic resources and utilization," In: *Afr. and Ame.*, J.M. Al-Khayri, S.M. Jain and D.V. Johnson, (eds) Springer, Dordrecht, 2015.
- [11] J.M.S. Jubrael, "Genetic characterization for some date palm cultivars in Iraq using RAPD markers," *IPA J. for Agri. Rese.*, vol. 11 no. 1, pp: 138–148, 2001. (in Arabic).
- [12] A.A. Al-Khateeb, G.R. Abdalla and H.M. Dinar, et al., "Auxin: cytokinin reactions in the in vitro micropropagation of date palm (*Phoenix dactylifera* L.)," *Egypt J. Appl. Sci.* vol. 17, no. 10, pp: 409–415, 2001.
- [13] H.S.M. Khierallah, S.M. Bader, M. Baum, and A. Hamwiah, "Genetic diversity of Iraqi date palms revealed by microsatellite polymorphism," *J. Am. Soc. Hort. Sci.*, vol. 136 no. 4, pp: 282–287, 2011b.
- [14] A. Hamwiah, J. Farah, and S. Moussally, et al., "Development of 1000 microsatellite markers across the date palm (*Phoenix dactylifera* L.) genome," *Acta Hort.* vol. 882, pp: 269–277, 2010.
- [15] H.S.M. Khierallah, S.K.I. Al-Sammaraie, H.I. Mohammed, "Molecular characterization of some Iraqi date palm cultivars using RAPD and ISSR markers," *J. Asia Sci. Res.*, vol. 4 no. 9, pp: 490–503, 2014.
- [16] B.G. Baldwin, M.J. Sanderson, M.J. Porter, M.F. Wojciechowski, C.S. Campbell, and M.J. Donoghue, "The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny," *Ann. Missouri Bot. Gard.* vol.82, pp: 247–277, 1995.
- [17] Ainouche, and Bayer, "On the origins of the tetraploid *Bromus* species (section *Bromus*, Poaceae): insights from internal transcribed spacer sequences of nuclear ribosomal DNA," *Genome*, vol. 40, no. 5, pp: 730–743, 1997.
- [18] N. Arnheim, "Concerted evolution of multi-gene families," In: M. Nei, and R.K. Koehn (eds.) "Evolution of genes and proteins," Sinauer, Sunderland, pp: 38-61, 1983.
- [19] J. Sambrook, E. F. Fritsch and T. Maniatis, "Molecular Cloning: A Laboratory Manual," in 2nd ed. Plainview, N.Y.: Cold Spring Harbor Laboratory Press, 1989.

- [20] T.J. White, T. Bruns, S.J. Lee, and J.W. Taylor “Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and applications,” California USA: *Academic Press., San Diego*, vol.18 no. 1, pp: 315-322, 1990.
- [21] K. Tamura, G. Stecher, and S. Kumar, “MEGA11: Molecular Evolutionary Genetics Analysis version 11,” *Molecular Biology and Evolution*, vol. 38, pp: 3022-3027, 2021.
- [22] X. Xia, “DAMBE6: New tools for microbial genomics, phylogenetics and molecular evolution,” *Journal of Heredity*, vol. 108, pp: 431-437, 2017.
- [23] K. Tamura, M. Nei, and S. Kumar, “Prospects for inferring very large phylogenies by using the neighbor-joining method,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, pp: 11030–11035, 2004.
- [24] N. Saitou, and M. Nei, “The neighbor-joining method: a new method for reconstructing phylogenetic trees,” *Mol. Biol. Evol.*, vol. 4, pp: 406–425, 1987.
- [25] A.G. Kluge, and J.S. Farris, “Quantitative phylogenetics and the evolution of anurans,” *Systematic Zoology*, vol. 18, pp: 1–32, 1969.
- [26] P. Librado, and J. Rozas, “DnaSP v5: A software for comprehensive analysis of DNA polymorphism data,” *Bioinformatics*, vol. 25, , pp: 1451-1452, 2009.
- [27] T.H. Jukes, and C.R. Cantor, “Evolution of protein molecules,” In: H.N. Munroled, (Ed.), “Mammalian Protein Metabolism”. Academy Press, New York, pp: 31–132, 1969.
- [28] M. Nei, and F. Tajima, “Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data,” *Genetics*, vol. 105, pp: 207–217, 1983.
- [29] F. Tajima, “Statistical method for testing the neutral mutation hypothesis by DNA polymorphism,” *Genetics*, vol. 123, pp: 585–595, 1989.
- [30] Y.X. Fu, and W.H. Li, “Statistical tests of neutrality of mutations,” *Genetics*, vol. 133, pp: 693–709, 1993.
- [31] A.R. Rogers, and H. Harpending, “Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.*, vol. 9, pp: 552–569, 1992.
- [32] H.J. Bandelt, P. Forster, and A. Rohl, “Median-Joining Networks for Inferring Intraspecific Phylogenies,” *Molecular Biology and Evolution*, vol. 16, pp: 37-48, 1999.
- [33] M.H. Amar, A.H.M. Hassan, and E.A.M. El Sherbeny, “Assessment of genetic diversity in some wild plants of Asteraceae family by ribosomal DNA sequence,” *Egypt. J. Genet. Cytol.*, vol. 41, pp: 195–208, 2012.
- [34] R. Bellarosa, M.C. Simeone, A. Papini, and B. Schirone, “Utility of ITS sequence data for phylogenetic reconstruction of Italian *Quercus* spp,” *Mol. Phylogenet. Evol.*, vol. 34, pp: 355–370, 2005.
- [35] M. Kehie, S. Kumaria, K., Sangeeta Devi, and P. Tandon, “Genetic diversity and molecular evolution of Naga King Chili inferred from internal transcribed spacer sequence of nuclear ribosomal DNA,” *Meta Gene*, vol. 7, pp: 56–63, 2016.
- [36] R.A. Torres, M. Ganal, and V. Hemleben, “GC balance in the internal transcribed spacers ITS1 and ITS2 of nuclear ribosomal RNA genes,” *Journal of Molecular Evolution*, vol. 30, pp: 170–181, 1990.
- [37] S. Sharma, S. Rustgi, H.S. Balyan, and P.K. Gupta, “Internal transcribed spacer (ITS) sequences of ribosomal DNA of wild barley and their comparison with ITS sequences in common wheat,” *Barley Genet. Newslett.* pp: 32–45, 2002.
- [38] V. Hemleben, B. Leweke, A. Roth, and I. Stadler, “Organization of highly repetitive satellite DNA of two Cucurbitaceae species (*Cucumis melo* and *Cucumis sativus*),” *Nucleic Acids Research* vol.10, 631-644, 1982.
- [39] G. Charmet, C. Ravel, and F. Balfourier “Phylogenetic analysis in the *Festuca–Lolium* complex using molecular markers and ITS rDNA,” *Theor Appl Genet.* vol. 94, pp:1038–1046, 1997.
- [40] Y.X. Fu, “Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection,” *Genetics*, vol. 147, pp: 915–925, 1997.
- [41] S.E. Ramos-Onsins, and J. Rosas, “Statistical properties of new neutrality tests against population growth,” *Mol. Biol. Evol.*, vol. 19, pp: 2092–2100, 2002.
- [42] M. Kitavi, “Genetic diversity, evolutionary history and epigenetic analysis of East African highland bananas,” Ph. D. Thesis. National University of Ireland, Galway, p: 291, 1989.
- [43] F. Tajima, “The Effect of Change in Population Size on DNA Polymorphism,” Department of Biology, Kyushu University, Fukuoka 812, Japan. *Genetics Society of America*, 2009.

- [44] M. Ibrahimi, N. Brhadda, R. Ziri, M. Fokar, D. Iraqi, F. Gaboun, M. Labhilili, A. Habach, R. Meziani, J. Elfadile, R. Rabha Abdelwahd, and G. Diria, Analysis of genetic diversity and population structure of Moroccan date palm (*Phoenix dactylifera* L.) using SSR and DAMD molecular markers, *J. Genet. Eng. Biotechnol*, vol. 21 no. 1, p: 66, 2023.
- [45] H.S.M. Khierallah, and H.D. Azhar, "Study of genetic diversity of Iraqi date palms using some morphological markers," *Int. J. Curr. Microbiol App. Sci.*, vol. 5 no. 3, pp: 317–327, 2016.
- [46] M.S. Alansari, A.K.A. Al-kazaz, and H.S.M. Khierallah, "Assessment of Genetic Distance Among Some Iraqi Date Palm Cultivars (*Phoenix Dactylifera* L.) Using Randomly Amplified Polymorphic DNA". *Iraqi Journal of Science*, vol 55, no.4B, pp:1833-1843, 2014.
- [47] H.S.M. Khierallah, S.M. Bader, M. Baum, and M.A. Al-Khfaji, "Assessment of AFLP variations in date palm in vitro plantlets derived from inflorescence explant". The 4th Symp., date palm, KSA, 2008 Proc. The 4th Symp., date palm, KSA, pp: 552–564, 2007.
- [48] H.S.M. Khierallah, S.M. Bader, , M. Baum and A. Hamwieh, "Assessment of genetic diversity for some Iraqi date palms (*Phoenix dactylifera* L.) using AFLP markers". *Afr. J. Biotech.*, vol. 10 no. 47, pp: 9670–9676, 2011a.
- [49] A.A.S. Habib, and H.J. Attiya, "Evaluation the activity of Ajwa of date palm fruits extract as promoter rooting for stem cuttings of cassia (*Cassia surattensis* Berm)". *Iraqi Journal of Science*, vol. 57, no.3A, pp:1664-1670, 2016.
- [50] Y.H. Al-Juraisy, "Modulatory Effects of Date Palm (*Phoenix dactylifera* L.) Pits Extract Against Ethinylestradiol Induced Genotoxic Damage in Cultured Human Lymphocytes". *Iraqi Journal of Science*, vol. 57, no.1C, pp: 576-582, 2016.