Iraqi Journal of Biotechnology, 2025, Vol. 24, No. 1, 86-102

An Internal Transcribed Spacer1 (ITS1) Region to Assess Genetic Variety and Molecular Evolution for the Iraqi Date Palm

[1]Mohammed M. Hawash, [2]Laith M.J. Al-Shamma

[1, 2]Department of Biology, College of Science, University of Baghdad

Received: February 11, 2024 / Accepted: April 28, 2024 / Published: March 5, 2025

Abstract: In order to set up genetic relations between date palm cultivars and to illustrate their molecular evolution, the sequences of the internal transcribed spaser1 (ITS1) region were analysed in fifteen Iraqi date palm cultivars. Seven of them were propagated by tissue culture technique. The variability of GC content in that region was determined at an average of (52.7). Out of the 15 cultivars, the aligned Sequences permitted us to pinpoint 14 haplotypes. The resultant transition/transversion bias (R) is (0.8), demonstrating that transitions occur in this location less frequently than transversions. The aligned sequences had 637 characters, according to data from (ITS1), and the divergence values were found to vary from (0.0000%) to (0.0063%) with an average of (0.0023%). The results refer to recent demographic expansion for a population of Iraqi date palms that would return to equilibrium after 0.6N generations. The study also demonstrated that the Mir alhajj TC cultivar was found to be the most sensitive cultivar for tissue culture technique propagation. After doing a multi-variety PCA analysis among cultivars, it was discovered that (PC1 and PC2) have a variability of (0.9770 and 0.4218) and an Eigen value of (7.3222 and 6.3453), respectively.

Keywords: Palmevolution, mismatch distribution, sequence analysis, polymorphism, (*Phoenixmm dactylifera L.*).

Corresponding author: (Email: laith.alshamma@sc.uobaghdad.edu.iq).

Introduction

The date palm (Phoenix dactylifera L.) is a monocotyledonous plant that is perennial, diploid (2n = 2x = 36), and belongs to the tribe Phoeniceae of the family Palmae. The Greek words "phoenix," which mean "purple or red," and "dactylifera," meaning "finger," both are related to the fruit's colouring and shape (1). These plants that have been widely farmed in the Middle East and North Africa at latitudes 10° and 35° north and south of the equator since they were domesticated in Mesopotamia circa 3000 BC. As a result of being dioecious, the date palm displays a high level of genetic variety (2).

The date palm is mostly vegetatively propagated by branches, although it can also be done through tissue culture or through seeds. Although seed propagation creates new genotypes or forms of date palm, which is believed to be the main source of diversity in date palm, it is not the only source of these heterogeneity patterns.

Evaluating the genetic variety that is now accessible is one of the key objectives in plant genetic resource analysis. Date palm genetic diversity was investigated on many scales, including that between cultivars,

populations, or individual clones, as well as across various geographical locations. At the morphological, biochemical, physiological, or molecular levels, genetic diversity can be evaluated (3). By engaging in activities like agriculture, social interaction, artificial selection, and the spatiotemporal interchange and movement of germplasm, humans can also affect the genetic diversity of date palms. Among various oases and populations, the level and distribution of genetic diversity may differ. As a result of ecological, anthropogenic, geographic, and historical factors, according to reports date palm cultivars have a shared genetic background. As a result, correct distinction of the cultivars and individual plant assignments within each cultivar is a challenging undertaking, and errors are unavoidable. The genetic mingling of the date palms may also be to blame for this (4).

Although somaclonal variety in plants generated by tissue culture is occasionally acknowledged, this does not necessarily lead to the same plants. Date palms grown from tissue are frequently discovered to be anomalous, displaying common traits including broad leaflets, a slow rate of growth and development, and variegation (leaves).

Because the genus is dioecious, there can be a lot of interbreeding, which can result in interspecific hybrids. The natural variants and cultivars of the genus' species, according to Wrigley (5), should be considered to be a single species because of their physical similarity. Research has shown that molecular markers are particularly useful in establishing phylogenic linkages and cultivar characteristics among genotypes, and their usage has become essential in characterising these aspects.

Several molecular markers were used to evaluate the genetic diversity and genetic structure of Iraqi date palms throughout different geographic regions of Iraq. The bulk of these molecular markers were RAPDs and AFLP, in addition to microsatellite markers, ISSR, and SSR markers (6, 7). As reviewed by (8) and (9), employ palm biotechnology.

The aim of this study is to figure out genetic relationships and tissue culture impact at the molecular level among comparison to those propagated from offshoots, attempting to comprehend the evolutionary history and genetic fixed assets of the best cultivars of Iraqi date palms as a starting point for developing new cultivars.

Materials and methods

Date palm cultivars

For this study, 15 noteworthy date palm cultivars were chosen, of which 8 were reproduced by offshoots and 7 were reproduced using tissue culture tiqneque. The comparison of these cultivars' genetic make-ups amongst one another was the aim of this investigation.

These cultivars were chosen based on their economic viability, especially taking consumer preferences and fruit market demands into consideration. The original Arabic names of the date palm cultivars have changed due to changes in spelling and pronunciation. The most frequently used spellings were chosen for this investigation. In alphabetical order, the selected fruiting cultivars are (Barhi offs, Barhi tc, Khalas offs, khalas tc, $Majhool_1$ offs, $Majhool_1$ tc, $Majhool_2$ offs, Maktomi offs, Maktomi tc, Mir alhajj offs, Mir alhajj tc, Showaithy

offs, Showaithy tc, Um alduhan offs, and Um alduhan tc).

Plant materials

For each of the 15 chosen cultivars, fresh leaflet samples were gathered from sources. These samples were collected at five different places around Baghdad and Al Anbar. Five leaves from each of the 10 cultivars, with a range of around 15 cm, were picked at random. Only trees between the ages of 10 and 15 years that were healthy, established, and well-characterised were used as the source for these samples. In order to identify the cultivar to which these chosen trees belonged, professionals from the date palm stations of the Ministry of Agriculture, Baghdad University, and Janet Al-Nakheel Company Laboratory in Iraq gave assistance. In mid-April of the year 2023.

DNA extraction

Total genomic DNA was extracted from the entirety of modern leaflets that were purified from impurities and frozen under liquid nitrogen using the **FavorPrep™ Plant Genomic DNA Extraction Mini Kit** (sample size: 20 ~100 mg) from Korea.

A 100-base pair ladder was used as the DNA molecular weight marker in this experiment. The gel was then exposed to an electrical field for one hour at 70 volts and 65 amps. In 1X TBE buffer, the DNA was examined using a UV transilluminator.

PCR amplification of the (ITS1) region

A premix kit for polymerase chain reaction (PCR) amplification was used, which contained deoxyribonucleotides (DNTPs) with a concentration of 2.5 mM, reaction buffer with a concentration of 10X, gel loading buffer with a concentration of 1X, as well as I-Taq DNA Polymerase with an activity of 2.5U. The 25 µl total volume of the PCR amplification mixture used for the particular diagnostic gene reaction included 1.5 µl of DNA, 5 µl of PCR PreMix, and 1 µl of each primer with a concentration of 10 pmol. To get the final amount of 25 µl, 16.5 µl of distilled water was then added. The universal ITS5 (5′-ATGATAACTCGACGGACCGC-3′) and ITS2 (5′-TCTTCGAGCCCCCAACTTTC-3′) primers were used for PCR amplification. The thermal cycling conditions were as follows: a 5-minute initial denaturation phase at 95 °C, 35 cycles of denaturation at 95 °C for 45 seconds, annealing at 57 °C for 1 minute, extension at 72 °C for 1 minute, and finally a 5-minute final extension phase at 72 °C. Applied Biosystems' Gene Amp PCR system 9700, a thermal cycler, was used to carry out the amplification procedure. After that, 1.5% agarose gel electrophoresis was used to separate all of the PCR results, and they were all then coloured with red safe dye (Intron Korea).

DNA sequencing and alignment

Sequencing was performed for the gene encoding the ITS1 region of the nr DNA by taking the PCR products of the species included in the study in a volume of 25 µl with a pair of primers (forward + reverse) 10 pmol/µl for each gene after eppendrof pcr tube 0.2 µl, clean, sterile, tightly packed, and sent to Macrogen Biotechnology Company in South Korea for determination of nucleotide sequences, with Sanger dideoxy sequencing tiqneque. Thus, the Basic Local Alignment Search Tool (BLAST) was used to compare the acquired sequences to the GenBank database. Further, there was a high similarity with the standard gene sequence under ID (XR_005510437.1)

available at the National Centre for Biotechnology Information (NCBI).

Sequence alignment and analysis

All details of the resulting 15 sequences have been deposited into the GenBank database (accession numbers OQ911641.1–OQ911655.1) based on the ITS1 region of 18S nrDNA (Table 2). The nucleotide sequences were aligned using the DAMBE version 7.3.32 (10) and analysed with the MEGA version 11.0.13 software (11). The alignment had been manually validated. In order to establish the pairwise sequencing differences across cultivars in the ITS1 region, the "Maximum Composite Likelihood (MCL)" approach was used (11). In addition, the GC content of each sequence was determined online by Biologics Corp. (https://www.BiologicsCorp.com/tools/GCcontent/). The "neighbour-joining" (NJ) approach, with 1000 bootstrap replications were executed for constructing the "neighbour joining and maximum parsimony" trees. The transitions/transversions rate (ti/tv) was determined by applying the equation R = [A*G*k1 T*C*k2]/[(A+ G)*(T + C)], where A, G, C, and T stand for each frequency of the four nucleotides (11). Along with the consistency indexes (CI) and retention indexes, the scores for the homoplasy index (RI) have also been established. The analysis displays the number of substitutions for each site identified in the sequences. For measuring genetic diversity among cultivars, by evaluating the aligned sequences using the DnaSP programme version 5.10.01 (12), the polymorphism indices, demographic history, and haplotype indices (Hd), together with nucleotide substitution and diversity (Pi), were determined, with a standard deviation, across the (ITS1) sequences. The mean pairwise nucleotide difference (K) was estimated. Using Tajima's D and Fu and Li's D* and F* approaches, selective neutrality was also analysed. The investigation of the distribution of allelic frequency at a site, the mismatch distribution of pairwise sequencing differences, and population size change were also included in the evaluation of demographic variables. Network analysis using the (NETWORK) application version 4.6.1.0 was used to illustrate the genetic relationships among the date palm cultivars as evidenced by the haplotypes (13). The scatter chart plot generated by the multivariate PCA was also included using DAMBE version 6 (10).

Results and discussion

The amplified section has a length of approximately 600 base pairs, according to an internal transcribed spacer 1 (ITS1) region of nuclear riposomial DNA, as shown in (Figure 1).

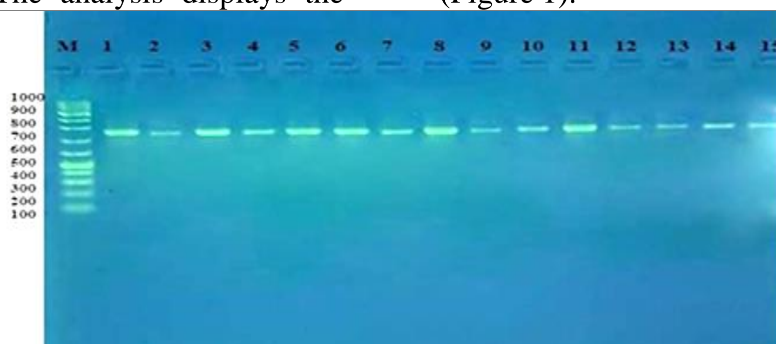

Figure (1): PCR product the band size. The product was electrophoresis on 1.5% agarose at 5 volt/cm2. 1x TBE buffer for 1:30 hours. M: DNA ladder (100).

A sequence based on ITS1 of the nr DNA sequences was constructed after a set of the best 15 Iraqi date palm cultivars were investigated. The sequence is 637 base pairs lengthy, and the nucleotide composition frequencies for A, T, C, and G are 23.88%,22.92%, 24.29%, and 28.91%, respectively. while the average size of the Tunisian date palm, according to a report by Mainaa et al. (14), is 442.7 bp, and the nucleotide proportions seem about equal, with A accounting for 24.76%, T for 25.67%, C for 27.52%, and G for 21.95%. However, similar outcomes have also been noted for other plant species, such as members of the Asteraceae family, where the overall length variation of the ITS region ranged from (650) to (750) bp. Average frequencies for adenine, thymine, cytosine, and guanine were (25%), (24%), (26%), and 25%, respectively, whereas the overall GC content and AT content were (51% and 49%, respectively) (15). Contrarily, it was discovered by Baraket et al. (16) that the ITS region originated from the Tunisian fig cultivars, which had an average length of 697.5 bp and a nucleotide makeup of 19.7% adenine, 18.6% thymine, 31.4% cytosine, and 30.2% guanine. whereas the nucleotide frequencies for A, T, C, and G were (18.85%), (17.56%), (33.95%), and (29.64%), respectively, in line with the evaluation of the entire ITS sequence in the Naga King Chilli by Kehie et al. (17), which reported that it had an average length of (620) bp. Additionally, the sequence length of the ITS region in angiosperms varied from 565 to 700 bp, while in species from the Coniferales, Cycadales, Ginkgoales, and Gnetales orders, it varied from 975 to 3125bp (18).

Genetic variation

Nucleotide composition variance and mutational events

GC content

In reality, the ITS1 region's GC content detection rate varied between 52% and 53%, with an average of 52.5%. In fact, the range of GC content documented for the ITS region as a whole was between 53% and 52%, with an average of 52.5 for all cultivars evaluated. While cytosine was constant across all cultivars under investigation, it appears that the variation in GC content may be entirely attributed to variations in the nitrogenous base guanine. See Figure (2). Contrary to Cucurbitaceae, as described by Hemleben et al. (19), the date palm's rDNA appears to have more genetic conservation and resistance against methylation. It is important to note that the GC content found in this study is comparable to the range of 49% to 49.7% that was found in Tunisian date palms over the complete ITS district (14). The average GC content over the whole ITS was 51% (15), which is comparable to other plant species, including those in the Asteraceae family. Additionally, the GC content based on regions of the ITS1, ITS2, and (5.8S genes) ranged from 53.4 to 68, 52.3 to 67.7, and 46.7 to 56.6, respectively, in *Ficus carica* L. (16). Notably, the homogeneity between cultivars is indicated by the minimal variability of the GC content of the

ITS1 region, supporting the high level of conservation seen in Iraqi date palms. However, it is clear from adaptation to the environment that ITS are sensitive to temperature and pressure.

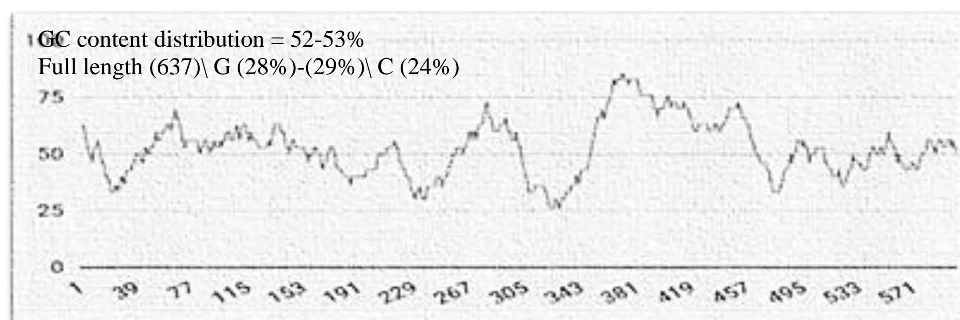

Figure (2): Show percentage and variation of GC distribution among rDNA sequences of Iraqi date palm cultivars, ranging from 52%, to 53%.

The transition/transversion (R)

The transition/transversion ratios for purine nucleotides were K_1 = 1.793 and for pyrimidine bases, K_2 = 1.357. A total transition/transversion ratio (R) of 0.8 was discovered for all bases. Tables (1) lists the different substitutions discovered and shows that transitions are less frequent in Iraqi date palm sequences than transversions. The transitions G\A and C\T are occur more frequently than A\G and T\C transitions. The majority of the observed transition/transversion ratio (R) of 0.8, based on the ITS1 region, exhibits similarities to the ratio revealed in the assessment of the entire ITS region of Tunisian Fig, a ratio of 0.7 noted by (16). It is interesting to note that this ratio is lower than the ratio of 4.375 found across an entire ITS region of Tunisian date palms, as documented by Mainaa et al. (14). Whereas the (ti/tv) ratio for the Asteraceae family was reported by Amar et al. (15) to be 1.43, and in another study by Kehie et al. (17), the ratio for *Capsicum* sp. proved to be 3.746. Additionally, Sharma et al. (20) states that the ratios for wheat and wild barley were 6.9 and 7.4, respectively. Noted, the ratio lowered due to lower transition events in comparison to transversion occurrences.

Table (1): Relative frequencies of nucleotide substitutions estimated in the ITS1 region of nrDNA, entire cultivars under study.

	A	T	C	G
A	-	*6.39*	*6.77*	**14.45**
T	*6.66*	-	**9.19**	*8.06*
C	*6.66*	**8.67**	-	*8.06*
G	**11.94**	*6.39*	*6.77*	-

NOTE1*:- 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 transversional substitutions are shown in italics.

Table (2): Iraqi date palm samples show the ID of the accession sequences, ID of the compared sequence, and type of substitution sequence, and percentage of identify.

Cultivar			Substitutes type	Nucleotide substitutes	ID of the accession sequences	ID of the compared sequence	Identities
Showaithy	Offs+tc	Offs alone	Tv	C\A	ID:OQ911641.1	ID:XR_005510437.1	99%
			Ti	A\G			
Maktomi			Tv	C\G	ID:OQ911642.1	ID:XR_005510437.1	99%
Barhi			Ti	A\G	ID:OQ911643.1	ID:XR_005510437.1	99%
Majhool 1			Tv	C\A	ID:OQ911644.1	ID:XR_005510437.1	99%
			Tv	G\T			
Mir alhajj			Ti	C\T	ID:OQ911646.1	ID:XR_005510437.1	99%
			Ti	A\G			
Khalas			Tv	A\C	ID:OQ911647.1	ID:XR_005510437.1	99%
			Tv	T\A			
			Tv	T\A			
Um alduhan			Ti	T\C	ID:OQ911648.1	ID:XR_005510437.1	99%
			Tv	G\T			
			Ti	T\C			
Majhool 2			Ti	G\A	ID:OQ911645.1	ID:XR_005510437.1	99%
Showaithy		Tc alone	Tv	C\A	ID:OQ911649.1	ID:XR_005510437.1	99%
			Ti	A\G			
			Tv	A\T			
Majhool			Tv	T\G	ID:OQ911650.1	ID:XR_005510437.1	99%
			Tv	C\A			
Barhi			-----	-----	ID:OQ911651.1	ID:XR_005510437.1	100%
Mir alhajj			Tv	A\T	ID:OQ911652.1	ID:XR_005510437.1	99%
			Ti	G\A			
			Ti	C\T			
			Ti	A\G			
Um alduhan			Tv	G\T	ID:OQ911653.1	ID:XR_005510437.1	99%
			Ti	T\C			
Maktomi			Tv	T\A	ID:OQ911654.1	ID:XR_005510437.1	99%
			Ti	G\A			
			Tv	C\G			
Khalas			-----	----	ID:OQ911655.1	ID: XR_005510437. 1	100%

Note2*: (Majhool 2 offs)* means that the cultivar was propagated by offshoot (offs), but the mother was propagated from tissue culture (tc).

The composition of the nucleotide sequence alignment within the Internal Transcribed Spacer 1 (ITS1) showcases remarkable diversity. from a character matrix of 637 base pairs (bp), reveals the presence of 615 conserved sites alongside 22 sites that are variations (Table 3). Among these variable sites, 15 are singletons at Site positions: 31, 76, 77, 109, 150, 166, 182, 287, 321, 335, 356, 462, 511, 619, and 620, while seven sites at positions: 25, 64, 133, 289, 372, 404, and 421, provide informative sites, which are considered parsimony informative. For A, T, C, and G, respectively, the frequencies of nucleotides revealed in the ITS1 sequences were 23.88%, 22.92%, 24.29%, and 28.91%.

Genetic relations of ITS1 sequences

The analysis of the distance matrix revealed that the genetic distance ranged from (0.0000) to (0.0063), with an average of (0.0023). No distance

(0.0000) has been shown between the cultivars of Barhi and Khalas. Whereas Mir alhajj offs cultivars were revealed to have a genetic distance of (0.0008). Um alduhan tc, Majhool tc, Majhool 1 offs, and Showaithy offs have a distance of (0.0031), but Barhi offs, Majhool 2 offs, and Maktomi offs have a degree of similarity with a distance of (0.0016). The distance among Khalas offs, Showaithy tc, Maktomi tc, and Um alduhan offs is 0.0047, indicating a significant similarity in their ITS1 sequences. Important to keep in mind is that the Mir alhajj tc cultivar indicated a greater degree of genetic distance (0.0063) see (Figure 3). For constructing phylogenetic trees, the Maximum Parsimony (MP) and Neighbour-Joining (NJ) techniques were used. The evaluation identified a maximum parsimony of 43 trees, allowing for the identification of the most parsimonious tree with 25 steps. evidenced by the consistency index of (0.880), retention index of (0.571), and homoplasy index of (0.120) that were obtained. It provides evidence that the analysed sequences exhibited homoplasic characters. The "neighbour-joining dendrogram" grouped Iraqi understudy date palm cultivars into two major groups, as depicted in (Figure 3). The first cluster is composed of two cultivars, Mir alhajj tc and Mir alhajj offs, despite a significant difference in distance between them. The second cluster is divided into two sub-groups. The first sub-group is represented by Showaithy tc, Showaithy offs, Maktomi tc, and Maktomi offs cultivars, which show a significant divergence in distance in dependence on propagation technique and a remarkable similarity between the two cultivars, Majhool 2 offs and Barhi offs. The second sub-group includes all of the surviving cultivars as Barhi tc and Khalas tc cultivars with significant similarity, while Um alduhan tc, Um alduhan offs, Majhool tc, and Khalas offs cultivars demonstrate the presence of a significant amount of variability that characterises all cultivars under study.

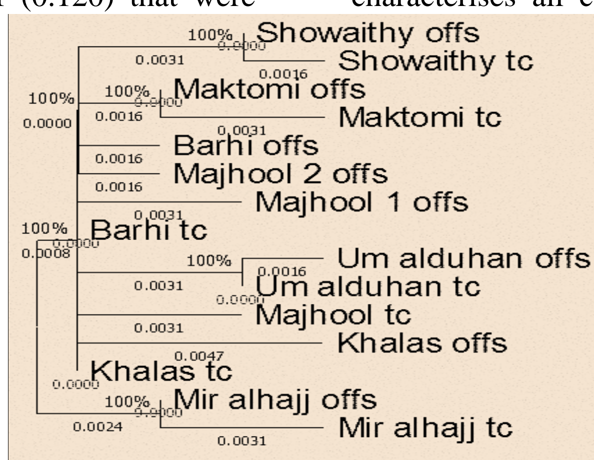

Figure (3): Neighbor Joining tree based on the ITS1 region of the nr DNA of date palm cultivars.

Genetic evolution

Tests for selective neutrality

Indeed, to assess if the observed diversity patterns in the nuclear ribosomal DNA sequences of Iraqi date palm cultivars significantly deviated from a neutral equilibrium hypothesis, we applied two commonly used tests to check for the selective neutrality of the observed mutations, employing both

Tajima's and Fu and Li's methods, in order to test the null hypothesis. According to our results, Tajima D in the ITS1 region revealed statistically significant negative values (D = -1.84929*) (P < 0.05). On the other hand, the statistical analyses employing Fu and Li on the full sample exhibited negative but not significant results (D* = -1.68225 (P>0.10); F* = -1.98806 (P>0.10) as in (Table 3). Indicators that the disequilibrium between mutation and gene drift with a limited number of rare mutations as detected within the sequences under study may be drawn from both of the two tests, especially the Tajma test, support the selective neutrality hypothesis. The Fu statistic is evaluated to give further information about the reason for a deviation from neutrality. This specific statistic is more useful for assessing population evolution and recent population expansion, as well as for detecting deviations from neutrality. Reviewing (Table 3 and Figure 4) reveals that the Fu's Fs parameter's values, which are shown as (-11.021), are high and highly negative. According to calculations, the R2 index has low values (0.0590) in relation to the internal transcribed spacer 1 (ITS1) of the nuclear ribosomal DNA (nr DNA). In contrast to the values reported by Mainaa et al. (14). for Tunisian date palm cultivars (Pi = 0.00155, Hd = 0.552), the levels of nucleotide and haplotype diversities observed in the Iraqi date palm (Pi = 0.0058 ± 0.0007, and Hd = 0.990 ± 0.028) appear to be relatively high (Table 3). Although the levels of haplotype diversity are comparable to those previously reported, it appears that these levels of nucleotide diversity are relatively low in comparison to other species, such as *Ficus carica*, as reported by Baraket et al. (16) with a haplotype diversity (Hd) of 0.996 and a nucleotide diversity (Pi) of 0.072). Furthermore, in the *Capsicum* sp. Kehie *et al.* (17) noted a haplotype diversity (Hd) of 1 and a nucleotide diversity (Pi) of 0.01499. Nuclear ribosomal DNA's internal transcribed spacer 1 (ITS1) sequence analysis indicates a relatively low degree of genetic diversity. The studied sequences reveal observed haplotypes in 14 out of 15 cultivars. It has minimal genetic diversity in this region, as evidenced by the average pairwise nucleotide difference (K) of 3.733 see (Table 3). The value is greater than the value determined by Mainaa et al. (14) in the Tunisian date palm (K = 0.686), despite the fact that on the other hand, the values of K reported by Baraket et al. (16) and Kehie et al. (17) in *Ficus carica* species (35.34) and *Capiscum* sp. (9.267), respectively, show a considerable level of variability over an entire ITS sequence for those specific species. Indeed, the deviation from selective neutrality can be attributed to the rare mutations discovered in the singleton sequences under study, according to the Tajima as well Fu and Li tests used to ascertain the cause of this divergence. The Fu statistic, which is well-known for its high efficacy in detecting departures from neutrality as well as its capacity to test for population growth and recent expansion, is computed. The distribution of mismatches in the *P. dactylifera* dataset as a whole indicates that the date palm population continues to expand, although slowly. Along with the statistical values of the Fu F's in the

entire sample of (-11.021), and the negative and significant D-Tajima -1.84929* (P<0.05) likewise provides clear proof of the population's expansion. For all sequences in the dataset, they were (F*: -1.98806 (P > 0.10); D*: -1.68225 (P > 0.10). These results point to changes in the population structure of the date palm cultivars under study. The computation of the R2 index, which reveals a low value of 0.0590, and the raggedness statistic of (r) 0.0307, which is less than 0.05, for the internal transcribed spacer1 (ITS1) region, serves as additional confirmation of our findings. The mismatched distribution is considered an indicator of expansion time. The data point to a recent demographic expansion in Iraq's date palms. The Tajima parameter D's high negative values appear to support this. These outcomes are in contrast to previous research on Tunisian date palm cultivars, which also evaluated the full ITS region (14).

Table (3): Tests for neutrality and sequence polymorphism performed on nr DNA.

Sequences analysis	The value	Sequences analysis	The value
No. of sequences	15	**K**	**3.733**
Length of Alignment	637 bp	**R2 statistic**	**0.0590**
Monomorphic characters	615	**S15(t)**	**1.220**
No. of Variable	22	**Value of S15(t)/a1**	**0.375**
Singleton variable sites	15	**S2(t)**	**0.259**
Parsimony informative	7	**average of S2(t)/a1**	**0.259**
No. of haplotypes (H)	14	**Tajima's D**	**-1.84929* (P < 0.05)**
Variance of haplotype diversity	0.00079	**Fu and Li's D***	**-1.68225 (P > 0.10)**
Pi ± SD	0.0058 ± 0.0007	**Fu and Li's F***	**-1.98806 (P > 0.10)**
Hd ± SD	0.990 ± 0.028	**Fu's Fs statistic**	**-11.021**
Pi(JC)	0.00589	**PCA1 variance 0.9770% Eigen value 7.3222**	
Theta (per site) from Eta	0.01062	**PCA2 variance 0.4218% Eigen value 6.3453**	

Fu's Fs: -11.021
P: 0.000

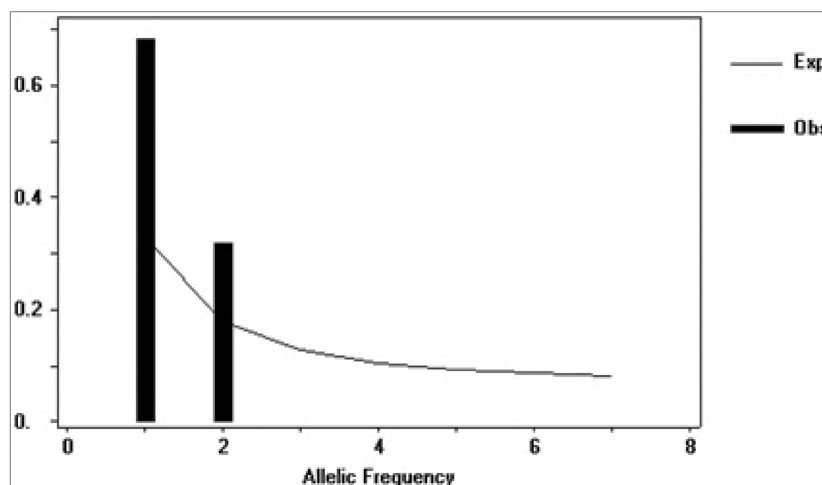

Figure (4): Displays a frequency spectrum of Iraqi date palm rDNA sequences detected in the ITS1 region. The distributions seen under balance and neutrality (mutation drift) are indicated by the solid lines drawn in the spectrum.

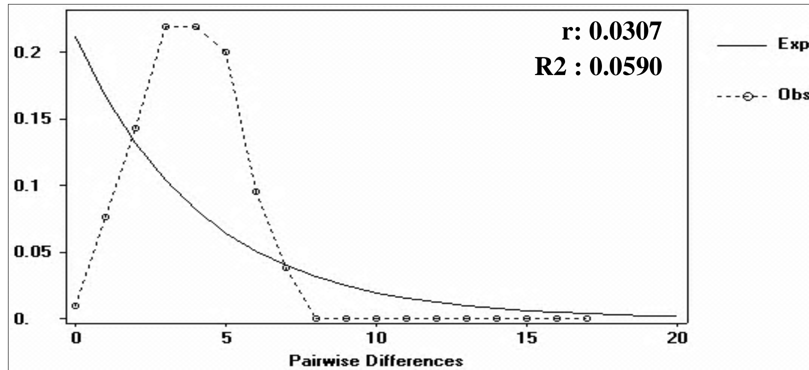

Figure (5): Mismatch distribution of palm population Displaying the observed distribution Expected Values for Population Size Changes with Initial Theta (0.000) and Final Tau (3.733).

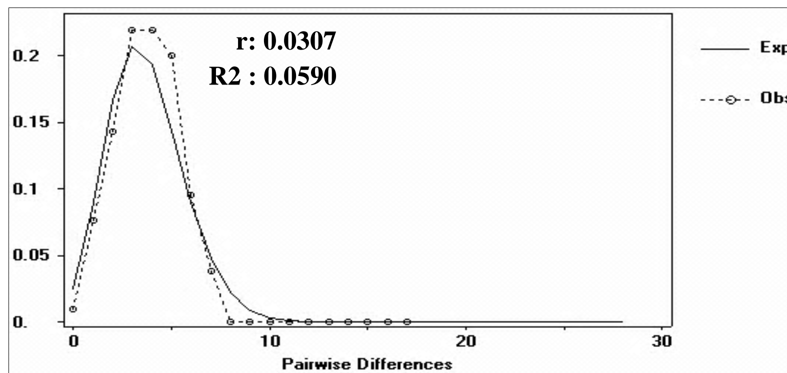

Figure (6): Mismatch distribution of date palm population. Displaying the observed distribution Expected Values for Population Size Changes at the Initial Theta (0.000), Final Theta (1000), and Final Tau (3.733).

Population size Change of palm

Within the 15 sequences, the expected number of segregating sites is Sn(t), estimated to be at (1.220), with a Sn(t)/al value of (0.375). The expected number of separating sites among two sequences, S2(t), or the average number of pairwise differences, has been determined to be (0.259). The ratio of S2(t)/a1 is calculated to be (0.259). After (0.6N) generations, it is anticipated that the total number of segregating sites among the 15 cultivars will be in equilibrium with the anticipated number, bringing the population back to equilibrium (Figure 7).

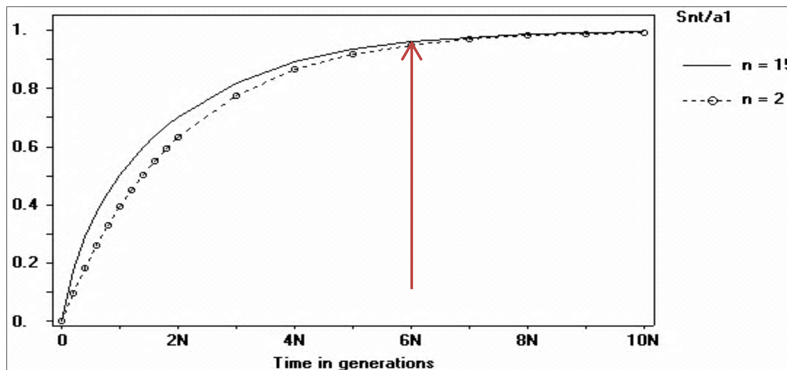

Figure (7): Mismatch in the population's distribution of date palms displaying the distribution of pairwise variations as seen and as predicted employing a population expansion equation at initiate (0.000), final (1000), final tau (3.733), and time (0.6N) generations.

The anticipated number of segregating sites in the 15 sequences was estimated at Sn(t) to be equal to 1.220, with a Sn(t)/al value of 0.375. The estimated number of segregating sites in two sequences of S2(t), or the average number of pairwise differences, was determined to be 0.259. S2(t)/a1 was likewise stated to have a ratio of 0.259. As a result, it is anticipated that after 0.6N generations, the total number of segregating sites among the 15 cultivars will be in equilibrium with the expected number in the population (Figure 7). Unlike Kitavi (21), which examined African bananas, the anticipated value of Sn(t) was 2.713 and the quotient of Sn(t)/al was 0.536. The predicted number of S2(t) was estimated to be 0.259, and 0.25 was the average number of S2(t)/a1. The segregating sites in the 89 cultivars were more numerous than anticipated in a population that was in equilibrium up until the ninth generation. While Tajima (22) demonstrated that, particularly when n = 2, the pace of variance expansion is frequently slow, Specifically, it takes 1.4N generations for this number to fall by half from its maximum value. On the other side, when n = 100, this approach only needs 0.5N generations. The researcher also showed that bigger sample sizes result in the creation of more segregating sites more quickly.

Haplotypes distribution based on ITS1 sequences

According to the ITS1 ribosomal DNA sequences, the haplotype network illustrated in (Figure 8) reveals an evident evolution of Iraqi date palms from an ancestral haplotype, which emphasises this fact by exhibiting a network that is radial-shaped and located the founder "Barhi tc" cultivar at the centre, as well as the founder haplotype connected with all other cultivars. As an outcome, it is implied that the "Barhi tc" cultivar acts as an ancestral for other cultivars. Notably, the majority of cultivars that the "Mir alhajj tc", "Maktomi tc", and "Showaithy tc" cultivars propagated from tissue culture are composed of other cultivars that were propagated from offshoots as ancestral to them, proving that tissue culture has not always produced individuals that are genetically identical. as well Since barely three centuries have passed since the invention of the tissue culture technique, it is possible that the influence of stress on the callus' development under chemical materials is the cause of these mutations. But regarding "Barhi tc" and "Um alduhan" cultivars as ancestrally to other cultivars propagated from offshoots. Thus, that may be either genetically identical or the impact of selective natural selection, recombination, or mutation reverse as a result of simultaneous evolution.

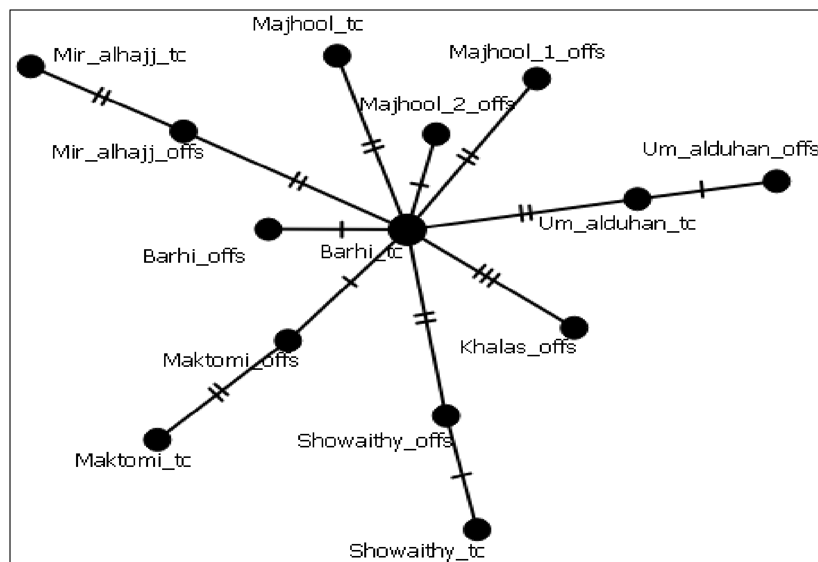

Figure (8): The network of haplotypes estimated from (ITS1) sequences reveals relationships among the 15 cultivars. The hatch marks are proportional to the number of mutations.

The minimum network was established by attaching those sequences to a radial shape in accordance with the number of mutational changes. in order to establish the genetic relationships among haplotypes in *P. dactylifera*. Employing nrDNA (ITS) sequences, as well, displays the genetic network of the identified haplotypes. The founder haplotype Barhi tc. represents, through which all other cultivars are linked under study. It was noted that the haplotypes emerged as "Mir alhajj tc" from "Mir alhajj offs" across two mutants, "Showaithy tc" from "Showaithy offs" across one mutant, and "Maktomi tc" from "Maktomi offs" across two mutants, but not compatible with the "Um alduhan offs" made up of "Um alduhan tc" in one mutant. Events imply that particular sequences have diverged lately, which may be due to issues occurring for the callus during tissue culture conditions. as well as asynchronous evolution that may play an important role through gene flow, genetic drift, and unequal crossing over, in addition to reverse mutation, which led to more polymorphism and then genetic diversity. The phylogenetic tree (Figure 3), and a network of haplotypes (Figure 8), are showed the same outcome.

Principal component analysis

The scatter plot obtained after performing a multiple-variate principal component analysis on the first two elements (PC1 and PC2) from the (ITS1) sequences indicated that PC1 possessed a variance of 0.9770% with respect to the cultivars and an Eigen value of 7.3222, whereas PC2 had a variance of 0.4218% and an Eigen value of 6.3453 (Figure 9).

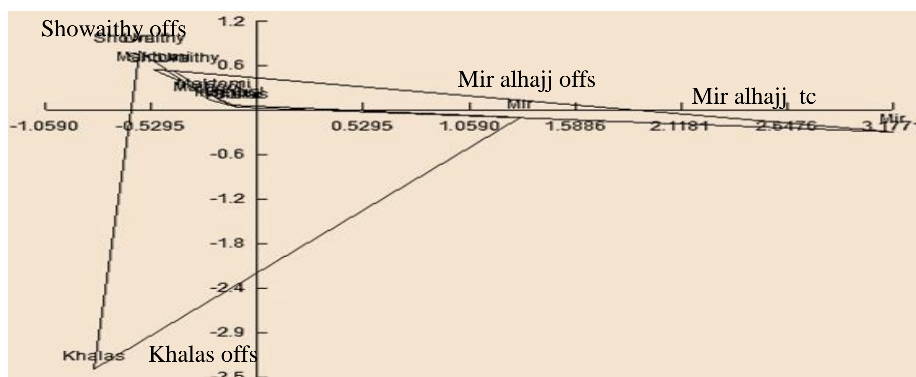

Figure (9): A scatter plot diagram for (principle component analysis), demonstrates the relative distribution of 15 date palm cultivars based on the (ITS1) region.

A study of population genetic variation using principal coordinates (PCA) cluster analysis tends to be more sensitive to relationships between individuals. For this reason, this multivariate technique was employed to supplement the cluster analysis information. Principal component analysis revealed two main clusters; the first and second Eigen vectors account for 6.66% and 6.51% of the variation, respectively. This analysis agrees with the population structure and dendrogram result in that. As with the other analyses, no clustering based on collection location or propagation methods was observed. much similar to the analysis on Moroccan date palms performed by Ibrahimi et al. (23). The clustering of the Iraqi cultivars with the Tata population was underlined by the PCA test plot. The examination of the combined data from SSR and DAMD markers showed that the first three axes were (7.3%), (5.75%), and (4.38%), respectively. On the other hand, Khierallah and Azhar (24) described that morphological markers were employed to identify a variation of 31.86% for PC1 with an Eigen value of 4.67. Furthermore, PC2 showed a variance of 17.0% and an Eigen value of 4.63. The results of the PCA's scatter diagram showed similarities between clusters and those in the phylogenetic tree. Despite the fact that Iraqi date palm cultivars have been cultivated for a very long time, PCA analysis has shown that these cultivars have a variety of associations with one another. Additionally, the topology of the phylogenetic tree and the dispersion of the cultivars in the PCA analysis showed that the germplasm of the Iraqi date palm is often characterised by continual morphological modification rather than genetic diversity. In the entire data set, the average ITS1 region percentage of genetic diversity of date palm cultivars under investigation was revealed to be 6%. The results from the Iraqi date palm were in perfect agreement with the research findings published by Mainaa *et al.* (14). When compared to other plant species, specifically *Daphnia pulex, Pinus sylvestris, Quercus petrea, Ficus carica, Juniperus osteosperma, and Capsicum* sp., it was found that the level of genetic variation found in both Tunisian and Iraqi date palm cultivars was relatively less significant. Therefore, it may be the coordinated evolution that led to minimising the variety among gene copies of the different cultivars under

study. Genetic distance measurement is based on the idea that mutations happen separately across the whole genome and that they are exponentially spaced apart over time. The timing of the subsequent mutation is therefore believed to be independent of the timing of earlier mutations. The relationships between each pair of cultivars are marked by a number of these substitutions, starting with a hypothetical common ancestor. The smallest of these genetic distances represents the degree of similarity with the comparative cultivar. The date palm cultivars were successfully grouped using the dendrogram for neighbour-joining, which is based on the internal transcribed spacer 1 (ITS1) region. It indicates that the varieties of date palm found in Iraq have an identical genetic history. This supports the idea that Iraq has an unique of a population, like the situation with Tunisian date palms (14). These results are consistent with other studies that looked at the Iraqi date palm using various molecular markers (6; 7; 25 and 26). On the other hand, since some of the population had more altered genetic patterns, it appears that this is completely different with individuals that are reproduced in tissue culture, disproving the prevailing notion that individuals propagated by tissue culture are genetically identical.

Conclusion

Two conclusions may be drawn from this. First, it is possible that the existence of frequent common alleles that have been identified in the ancestral population is mostly caused by mixing among different cultivars. This implies that these alleles either indicate homoplasy or shared ancestral polymorphism through a common ancestral parent during hybridization. Second, if the two populations originated through hybridization from the same parents, but the individuals that were propagated from tissue culture had higher mutation rates as a result of exposure of the callus to chemical stress. Therefore, over the course of successive generations, these differences accumulated, and populations began to diverge due to the actions of the selective nature or evolutionary factors that were expressed in each group, resulting in some genetic diversity between the two populations. However, coordinated evolution may play a main role in minimising the variety among gene copies of different Iraqi date palm cultivars through various processes such as recombination, gene conservation, and equal crossing over. in addition to further processes such as natural selection or reversed mutation, isolation, genetic drift, gene flow restriction, or factors impacting the reduced diversity of the nrDNA spacer.

References

1. Linne, (1734) Cited in Keaney, T. H. (1906). Date verities and date cultures in Tunis. Washington, U.S.D.A. Bureau of Plant Industry, Bulletin no. 92
2. Munier, P. (1981). Origine de la culture de palmier dattier et sa propagation en Afrique. Fruits 36: 437-450..
3. Jaradat, A. A. (2014). Synthesis and assessment of date palm genetic diversity studies. Emirates Journal of Food and Agriculture. 26 (11): 934-952.
4. Saboori, S.; Noormohammadi, Z.; Sheidai M. and Marashi, S.S. (2021). Date Palm (*Phoenix dactylifera* L.) Cultivar Relationships Based on Chloroplast Genotyping. Iranian Journal of Science and Technology. Transaction A, Science 45(10): 833–840.
5. Wrigley, G. (1995). Date palm, *Phoenix dactylifera*. In: Smartt, J. and Simmonds,

N.W. (ed.), Evolution of crop plants, 2nd ed. Longman, London, 399-403.

6. Khierallah, H.S.; Al-Sammarraie, S.K. and Mohammed, H.I. (2014). Molecular characterization of some Iraqi date palm cultivars using RAPD and ISSR markers. Asian Journal of Scientific Research. 4(9): 490–503.
7. Hamwieh, A.; Farah, J.; Moussally, S.; Al-Sham'aa, K.; Almer, K.; Khierallah, H., *et al*. (2010). Development of 1000 microsatellite markers across the date palm (*Phoenix dactylifera L.*) genome. Acta Horticulturae 882: 269-277.
8. Ibrahim, K.M. (2018). Integrating date palm biotechnology with community, a review. Iraqi Journal of Biotechnology, 17(2): 1-12
9. Mohammed, A.K.; Abdulhassan A.A. and Al-Meshhdany W.Y. (2017). Biosorption of cromium ions from aqueous solutions by using date palm fibers. Iraqi Journal of Biotechnology,16(4): 8-14
10. Xia, X. (2017). DAMBE6: New tools for microbial genomics, phylogenetics and molecular evolution. Journal of Heredity 108(4): 431-437.
11. Tamura, K.; Stecher, G. and Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis version 11. Molecular Biology and Evolution 38(7): 3022-3027
12. Librado, P. and Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 25(11): 1451-1452.
13. Bandelet, H.J.; Forster, P. and Röhl, A. (1999). Median joining network for inferring intraspecific phylogenies. Molecular Biology and Evolution. 16 (1): 37-48.
14. Mainaa, N.; Baraketa, G.; Salhi-Hannachia A. and Sakkaa, H. (2019). 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. 247(37): 373–379.
15. Amar, M.H., Hassan, A.H.M. and El Sherbeny, E.A.M., (2012). Assessment of genetic diversity in some wild plants of Asteraceae family by ribosomal DNA sequence. Egyptian Journal of Genetics and Cytology. 41(2): 195–208.
16. Baraket, G.; Ben Abdelkrim, A.; Mars, M. and Salhi-Hannachi, A. (2013). Genetic diversity and molecular evolution of the internal transcribed spacer (ITSs) of nuclear ribosomal DNA in the Tunisian fig cultivars (*Ficus carica* L.; Moracea). Biochemical Systematics and Ecology. 48(7): 20-33.
17. Kehie, M.; Kumaria, S.; Sangeeta Devi, K. and Tandon, P. (2016). Genetic diversity and molecular evolution of Naga King Chili inferred from internal transcribed spacer sequence of nuclear ribosomal DNA. Meta Gene 2(7): 56–63.
18. Liston, A.; Robinson, W.A.; Oliphant, J.M. and Alvarez-Buylla, E.R. (1996). Length variation in the nuclear ribosomal DNA internal transcribed spacer region of non-flowering seed plants. Systematic Botany. 21(2), 109–120.
19. Hemleben, V.; Leweke, B.; Roth, A. and Stadler, I. (1982). Organization of highly repetitive satellite DNA of two Cucurbitaceae species (*Cucumis melo* and *Cucumis sativus*). Nucleic Acids Research. 10(2): 631-644.
20. Sharma, S.; Rustgi, S.; Balyan, H.S. and Gupta, P.K. (2002). Internal transcribed spacer (ITS) sequences of ribosomal DNA of wild barley and their comparison with ITS sequences in common wheat. Barley Genetic Newsletter. 32: 38–45.
21. Kitavi, M. (1989). Genetic diversity, evolutionary history and epigenetic analysis of East African highland bananas. Ph. D. Thesis. National University of Ireland, Galway. 291 p.
22. Tajima, F. (2009). The Effect of Change in Population Size on DNA Polymorphism, Genetics Society of America . Department of Biology, Kyushu University, Fukuoka 812, Japan.
23. Ibrahimi, M.; Brhadda, N.; Ziri, R.; Fokar, M.; Iraqi, D.; Gaboun, F., *et al*., (2023). Analysis of genetic diversity and population structure of Moroccan date palm (*Phoenix dactylifera L.*) using SSR and DAMD molecular markers, Journal of Genetic Engineering and Biotechnology. 21(1): 66.
24. Khierallah, H.S.M. and Azhar, H.D. (2016). Study of genetic diversity of Iraqi date palms using some morphological markers. International Journal of Current Microbiology and Applied Sciences, 5(3): 317–327.

25. Kareem, M.A.H.; Al-Saadi, A.H. and Naji, H.F. (2018). Genetic diversity of Iraqi date palm (*Phoenix dactylifera L.*) by using RAPD technique. Journal of University of Babylon, Pure and Applied Sciences, 26(1): 114–131.
26. Kareem, M.A.; Naji, H.F. and Al-Saadi, A.H. (2016). Determination of genetic diversity of Iraqi date palm (*Phoenix dactylifera L.*) by using ISSR technique. Euphrates Journal of Agriculture Science, 8(3): 56- 68.