Cultivar-level phylogeny using chloroplast DNA barcode *psbK-psbI* spacers for identification of Emirati date palm (*Phoenix dactylifera* L.) varieties

M.R. Enan[1,2] and A. Ahmed[1]

[1]Biology Department, College of Science, United Arab Emirates University, United Arab Emirates
[2]Agricultural Research Center,
Agricultural Genetic Engineering Research Institute, Giza, Egypt

Corresponding author: M.R. Enan
E-mail: mohamed.enan@uaeu.ac.ae

Genet. Mol. Res. 15 (3): gmr.15038470
Received January 20, 2016
Accepted March 11, 2016
Published August 5, 2016
DOI http://dx.doi.org/10.4238/gmr.15038470

ABSTRACT. The efficacy of genetic material for use as DNA barcodes is under constant evaluation and improvement as new barcodes offering better resolution and efficiency of amplification for specific species groups are identified. In this study, the chloroplast intergenic spacer *psbK-psbI* was evaluated for the first time as a DNA barcode for distinguishing date palm cultivars. Nucleotide sequences were aligned using MEGA 6.0 to calculate pairwise divergence among the cultivars. The analyzed data illustrated a considerable level of variability in the genetic pool of the selected cultivars (0.009). In fact, five haplotypes were detected among 30 cultivars examined, yielding a haplotype diversity of 0.685. An unweighted pair group method with arithmetic

mean phylogenetic tree was constructed and shows a well-defined relationship among date palm cultivar varieties. On the other hand, selective neutrality investigations using Tajima test and Fu and Li tests were negative, providing evidence that date palm has been undergoing rapid expansion and recent population growth. Thus, we suggest that the *psbK-psbI* spacer can be successfully used to construct reliable phylogenetic trees for *P. dactylifera*.

Key words: *Phoenix dactylifera psbK-psbI* intergenic spacer; Chloroplast DNA barcode

INTRODUCTION

DNA barcoding has become a progressively important tool for taxonomic studies and species delimitation, as well as for the discovery of new (cryptic) species (Hajibabaei HYPERLINK "http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3890672/"et al.HYPERLINK "http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3890672/", 2007). A consortium of scientists has suggested the two-locus combination of *rbcL* and *matK* genes as a universal plant barcode (CBOL Plant Working Group, 2009), while other authors (Chen HYPERLINK "http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3890672/"et alHYPERLINK http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3890672/.HYPERLINK "http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3890672/", 2010) have suggested the ITS2 region as a more effective nuclear DNA barcode. However, despite all efforts, no locus (alone or in combination), has been confirmed to be 100% efficient as a common DNA barcode for plants at the species level. The first DNA barcoding investigation in palms (Jeanson HYPERLINK "http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3890672/"et alHYPERLINK http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3890672/.HYPERLINK "http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3890672/", 2011) achieved 92% success in species discrimination by applying a combination of three cpDNA markers (the plastids matK and rbcL and the nuclear ITS2) to the tribe Caryoteae. Palmae (Arecaceae) is the third most significant plant family in terms of human use; it includes about 180 genera and 2000 species distributed across the tropics and subtropics (Naeem et al., 2014).

Date palm (*Phoenix dactylifera* L.) is one of the earliest-domesticated fruit trees, and has great socioeconomic significance and nutritional value. It is the major income-producing crop in arid and desert areas (Hodel and Johnson, 2007). There are nearly 5000 date palm cultivars around the world (Jaradat and Zaid, 2004). The economic importance of date palm is largely due to its nutritionally valuable fruit, which contains 44-80% carbohydrates, 0.2-0.5% fat, 2.3-5.6% protein, and 6-12% dietary fiber (Hadrami and Al-Kjhayri, 2012). Numerous medicinal properties have also been attributed to date palm, including treatment for intestinal ailments, colds, sore throat, toothaches, fever, gonorrhea, and cough (Vayalil, 2012).

To ensure the conservation of high-quality varieties, tissue culturing and direct shoot regeneration methods are recommended (Khan and Bi Bi, 2012). Determination of genetic relationships among date palm varieties is of primary importance for the characterization of date palm germplasm, propagation programs, and conservation purposes (Haider et al.,

2012), but using biochemical markers (Gothwal et al., 2013) for genotype identification is difficult, and these markers can be altered by environment. Numerous molecular markers have been used to investigate genetic variety, including RAPD (Sedra et al., 1998; Mirbahar et al., 2014), ISSRs (Zehdi et al., 2002), SSRs (Yusuf et al., 2015), RAMPO (Rhouma et al., 2008), and AFLP (Khierallah et al., 2011). These nrDNA markers show significant polymorphism among date palm cultivars, but defining the cultivars remains problematic. However, cpDNA barcodes can also be used to assess phylogeny (Jamil et al., 2014). The authors (Enan and Ahamed, 2014) first attempted cpDNA analysis in Emirati date palm cultivars. Suitable molecular markers are needed to facilitate a deeper and sufficient awareness of the genetic diversity of date palm varieties.

Thus, the ultimate objective of this paper was to generate a phylogenetic relationship among a set of cultivars using the photosystem II *psbK-psbI* intergenic spacer.

MATERIAL AND METHODS

Plant material

Fresh young leaves of 30 different varieties (Table 1) of date palm (*P. dactylifera* L.) were collected from various areas of the United Arab Emirates. Three replicates were collected per cultivar and the samples were stored at -80°C.

Table 1. Date palm cultivars studied, accession numbers, and variation in the *psbK-psbI* region.

psbK-psbI						*psbK-psbI*					
Label	Ecotype	Accession No.	Length (bp)	GC (%)	AT (%)	Label	Ecotype	Accession No.	Length (bp)	GC (%)	AT (%)
K1	Zaghlul	KT593916	393	29.5	70.5	K16	Fard	KT764056	363	29.2	70.8
K2	Gashhabash	KT593917	366	29.0	71.0	K17	Jabiri	KT764057	396	29.5	70.5
K3	Khesab	KT593918	364	29.2	70.8	K18	Nagdi	KT764058	400	29.8	70.2
K4	Hatemy	KT593919	362	29.3	70.7	K19	Ashrasi	KT764059	364	29.1	70.9
K5	Anghal	KT764048	366	29	71	K20	Barhi	KT764060	366	29	71
K6	Lulu	KT764049	396	29.8	70.2	K21	BuMoaan	KT764061	336	29.4	70.6
K7	Degletnoor	KT764050	366	29	71	K22	Breem	KT764062	375	29.3	70.7
K8	AbuAzouq	KT764051	366	29	71	K23	Maktoom	KT764063	336	28.7	71.3
K9	Chichi	KT764052	392	29.3	70.7	K24	Diri	KT764064	367	28.6	71.4
K10	UmSala	KT898908	346	30.3	69.7	K25	Anwan	KT764065	367	28.9	71.1
K11	Khenezi	KT764053	366	28.7	71.3	K26	Khalas	KT764066	413	29.8	70.2
K12	AinBakr	KT898909	327	30.6	69.4	K27	Rabie	KT764067	401	29.6	70.4
K13	AbuKebal	KT898910	289	29.4	70.6	K28	Raziz	KT764068	367	28.6	71.4
K14	Dabbas	KT764054	377	29.2	70.8	K29	Azmy	KT764069	367	28.9	71.1
K15	Khadroui	KT764055	401	29.7	70.3	K30	Madhoun	KT764070	392	29.3	70.7
								Average	374 bp	29.3%	70.7%

DNA extraction

Total genomic DNA was extracted from fresh plant material using a DNeasy™ Plant Mini Kit (Qiagen, UK). For each sample, genomic DNA was extracted from 100 mg frozen leaf tissue, which was first ground using a BeadBlaster homogenizer (Benchmark Scientific, USA). Extracted DNA was quantified by means of a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific Inc., USA) and visualized on 1% agarose gel stained with ethidium bromide.

PCR amplification and nucleotide sequencing

Three replicate PCR assays were performed to amplify the *psbK-psbI* intergenic spacer located within the large single copy sequence of date palm chloroplast DNA using universal *psbK-psbI* primers *psbK*-F: 5'-TTAGCCTTTGTTTGGCAAG-3' and *psbI*-R: 5'-AGAGTTTGAGAGTAAGCAT-3' (Lee et al., 2007). The PCRs were prepared in a total 25-µL volume containing the following: 12.5 µL Taq PCR Master Mix (Qiagen, UK) in a final concentration of 200 µM each deoxynucleotide and 1.5 mM $MgCl_2$, 1 µM each primer (Eurofins MWG Operon, Germany), and 2 µL (20 ng) genomic DNA, adjusted with DNase-free sterile water. PCR amplification was performed using a T100 thermal cycler (Bio-Rad, USA) at 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 60 s. A final elongation step at 72°C for 10 min was performed to make sure any remaining single-strand DNA was fully extended. Owing to the presence of double bands in the amplicons, the PCR products were separated on 1% agarose gel; the bands were excised and purified using a QIAEX II Gel Extraction Kit (Qiagen, UK). Cycle sequencing products were created using a BigDye Terminator v3.1 Kit (Applied Biosystems, USA), then analyzed on an ABI 310 Automated DNA Sequencer (Applied BioSystems, USA). The PCR amplification primers were used also as sequencing primers.

Data analysis

The *psbK-psbI* sequences of all date palm varieties were uploaded to NCBI, and the Basic Local Alignment Search Tool (BLASTn) was used to perform queries for comparison to already-reported sequences in GenBank. All sequences generated were deposited in GenBank for reference; their accession numbers are provided in Table 1. The sequencing data acquired for all date palm ecotypes of the *psbK-psbI* spacer was aligned using CLUSTALW. Pairwise distance, transitional/transversional substitutions, and the maximum likelihood substitution matrix (Tamura, 1992) were estimated with the MEGA 6.2 software (Tamura et al., 2013). Genetic variation among the date cultivars was estimated by calculating the number of haplotypes, haplotype diversity (Hd), and parsimony-informative sites using the DnaSP 5.10.01 software (Librado and Rozas, 2009). Phylogenetic trees were inferred by the maximum likelihood (ML), neighbor-joining tree (NJ), and unweighted pair group method with arithmetic mean (UPGMA) methods, which were performed using MEGA 6.2 (Tamura et al., 2013). The topologies of the phylogenetic trees were evaluated using the bootstrap resampling method with 1000 replicates. Codon positions included were 1st + 2nd + 3rd + noncoding. To test population expansion, Tajima's D (Tajima, 1989), and Fu and Li's (1993) neutrality tests were performed in order to examine the null hypothesis that sequences are evolving according to neutral expectations.

RESULTS

Sequence characteristics

Successful double-stranded DNA amplifications and sequencing were obtained for

all the cultivars studied, with a success rate of 100%. The sequence data from the *psbK-psbI* region was aligned and subjected to BLASTn using NCBI. The similarity index ratio was checked against the *P. dactylifera* chloroplast complete genome (Accession No. GU811709.2). Accession numbers for all sequences were obtained from GenBank and published under the accession numbers listed in Table 1. For the *psbK-psbI* spacer, DNA sequences varied from 346 bp in the *Umsalla* cultivar to 417 bp in the *Abukebal* cultivar, with an average length of 374 bp. In addition, the GC content varied from 28.6 to 30.3%, and the AT content from 69.7 to 71.4% (Table 1). The substitution matrix and rates were estimated using ML under Tamura's 3-parameter model. Sequence analysis showed obvious nucleotide polymorphism among date palm varieties. For *psbK-psbI*, the analysis involved 30 nucleotide sequences; there were a total of 347 positions in the final dataset. The T to G, A to C, G to C, and C to G transversion rate was 2.35; that of A to T, G to T, T to A, and C to A was 5.51. The A to G and C to T transition rate was 10.26; that of C to T, T to C, and A to G was 24.03. Frequency of nucleotide substitution was as follows: A = 35.04, T/U = 35.04, C = 14.96, and G = 14.96 (Table 2). Pairwise distance was calculated on the basis of the *psbK-psbI* region, and the values of genetic diversity ranged from 0.000 to 0.009, as shown in Table 3. These very low distance values show that all varieties were closely related to each other. Polymorphism among the dataset was noted. The Hd, variance of haplotype diversity, theta (per site) from Eta, and nucleotide diversity among all varieties were found to be 0.685, 0.00419, 0.0022, and 0.00243, respectively (Table 4).

Table 2. Maximum likelihood values of transitional (bold) and transversional (italics) substitutions of nucleotides based on the *psbK-psbI* spacer for 30 date palm varieties as calculated by MEGA 6.0.

	A	T/U	C	G
A	-	***5.51***	***2.35***	**10.26**
T/U	***5.51***	-	**10.26**	***2.35***
C	***5.51***	**24.03**	-	***2.35***
G	**24.03**	***5.51***	***2.35***	-

Phylogenetic analysis

The identification power of three tree-building methods was tested among the date palm varieties: NJ, ML, and UPGMA. In this study, no differences were found between the results of NJ-, ML- and UPGMA-tree based analysis, so the resolution power was calculated based on UPGMA analysis alone. A phylogenetic tree built based on the *psbK-psbI* sequences of the date palm varieties showed very little genetic distance (Figure 1), indicating close genetic similarity among them. This dendrogram supported the varieties' organization into two main clusters, as denoted by clusters I and II. Cluster I includes 22 varieties, which can be separated into three sub-clusters. Sub-cluster A comprises 15 varieties with a 40% bootstrap value. Sub-cluster B includes two varieties, *Umsalla* and *Fard*, which were allied with 64% bootstrap similarity. Sub-cluster C comprises 5 varieties, namely *Hatmey*, *Dabass*, *Khesab*, *Nagdi*, and *Ashrasi*, with a 41% bootstrap value. Cluster II includes eight varieties, *Abukebal*, *Jabiri*, *Khenezi*, *Maktoom*, *BuMaan*, *Chichi*, *Diri*, and *Raziz*, which were clustered with a 41% bootstrap value (Figure 1).

Table 3. Pairwise distance among *psbK-psbI* sequences from 30 date palm varieties, as calculated by MEGA 6.0.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
[K1]																													
[K2]	0.000																												
[K3]	0.003	0.003																											
[K4]	0.003	0.003	0.000																										
[K5]	0.000	0.000	0.003	0.003																									
[K6]	0.000	0.000	0.003	0.003	0.000																								
[K7]	0.000	0.000	0.003	0.003	0.000	0.000																							
[K8]	0.000	0.000	0.003	0.003	0.000	0.000	0.000																						
[K9]	0.003	0.003	0.009	0.009	0.003	0.003	0.003	0.003																					
[K10]	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.009																				
[K11]	0.003	0.003	0.009	0.009	0.003	0.003	0.003	0.003	0.000	0.009																			
[K12]	0.000	0.000	0.003	0.003	0.000	0.000	0.000	0.000	0.003	0.003	0.003																		
[K13]	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.006	0.003	0.003																	
[K14]	0.003	0.003	**0.000**	0.000	0.003	0.003	0.003	0.003	0.009	0.003	0.009	0.003	0.003																
[K15]	0.000	0.000	0.003	0.003	0.000	0.000	0.000	0.000	0.003	0.003	0.003	0.000	0.003	0.003															
[K16]	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.009	0.000	0.009	0.003	0.006	0.003	0.003														
[K17]	0.003	0.003	0.009	0.009	0.003	0.003	0.003	0.003	0.000	0.009	0.000	0.003	0.003	0.009	0.003	0.009													
[K18]	0.003	0.003	0.000	0.000	0.003	0.003	0.003	0.003	0.009	0.003	0.009	0.003	0.003	0.000	0.003	0.003	0.009												
[K19]	0.003	0.003	0.000	0.000	0.003	0.003	0.003	0.003	**0.009**	0.003	0.009	0.003	0.003	0.000	0.003	0.003	0.009	0.000											
[K20]	0.000	0.000	0.003	0.003	0.000	0.000	0.000	0.000	0.003	0.003	0.003	0.000	0.003	0.003	0.000	0.003	0.003	0.003											
[K21]	0.003	0.003	0.009	0.009	0.003	0.003	0.003	0.003	0.000	0.009	0.000	0.003	0.003	0.009	0.003	0.009	0.003	0.009	0.009	0.003									
[K22]	0.000	0.000	0.003	0.003	0.000	0.000	0.000	0.000	0.003	0.003	0.003	0.000	0.003	0.009	0.000	0.003	0.003	0.003	0.003	0.000	0.003								
[K23]	0.003	0.003	0.009	0.009	0.003	0.003	0.003	0.003	0.000	0.009	0.000	0.003	0.003	0.009	0.003	0.009	0.000	0.009	0.009	0.003	0.000	0.003							
[K24]	0.003	0.003	0.009	0.009	0.003	0.003	0.003	0.003	0.000	0.009	0.000	0.003	0.003	0.009	0.003	0.009	0.000	0.009	0.009	0.003	0.000	0.003	0.000						
[K25]	0.000	0.000	0.003	0.003	0.000	0.000	0.000	0.000	0.003	0.003	0.003	0.000	0.003	0.003	0.000	0.003	0.003	0.003	0.003	0.000	0.003	0.000	0.003	0.003					
[K26]	0.000	0.000	0.003	0.003	0.000	0.000	0.000	0.000	0.003	0.003	0.003	0.000	0.003	0.003	0.000	0.003	0.003	0.003	0.003	0.000	0.003	0.000	0.003	0.003	0.000				
[K27]	0.000	0.000	0.003	0.003	0.000	0.000	0.000	0.000	0.003	0.003	0.003	0.000	0.003	0.003	0.000	0.003	0.003	0.003	0.003	0.000	0.003	0.000	0.003	0.003	0.000	0.000			
[K28]	0.003	0.003	0.009	0.009	0.003	0.003	0.003	0.003	0.000	0.009	0.000	0.003	0.003	0.009	0.003	0.009	0.000	0.009	0.009	0.003	0.000	0.003	0.000	0.000	0.003	0.003	0.003		
[K29]	0.000	0.000	0.003	0.003	0.000	0.000	0.000	0.000	0.003	0.003	0.003	0.000	0.003	0.003	0.000	0.003	0.003	0.003	0.003	0.000	0.003	0.000	0.003	0.003	0.000	0.000	0.000	0.003	
[K30]	0.000	0.000	0.003	0.003	0.000	0.000	0.000	0.000	0.003	0.003	0.003	0.000	0.003	0.003	0.000	0.003	0.003	0.003	0.003	0.000	0.003	0.000	0.003	0.003	0.000	0.000	0.000	0.003	0.000

Table 4. Sequence polymorphism among the date palm varieties based on the *psbK-psbI* spacer.

No. of haplotypes	Haplotype diversity (Hd)	Transition/ transversion bias (R)	Variance of haplotype diversity	Theta (per site) from Eta	Nucleotide diversity (Pi)	Parsimony-informative sites	Tajima's D	Fu and Li's D	Fu and Li's F
5	0.685	2.86	0.00419	0.0022	0.00243	2	-0.68535 (P > 0.1)	-0.34824 (P > 0.1)	-0.54036 (P > 0.1)

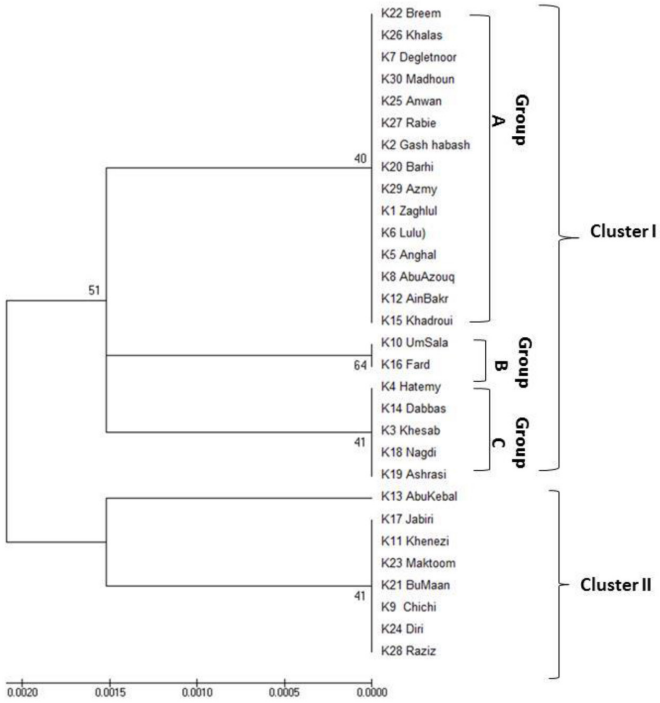

Figure 1. Phylogenetic tree of date palm varieties constructed on the basis of *psbK-psbI* spacer sequences using the UPGMA method. Branch length was calculated by Tamura 3-parameter method. Bootstrap (1000 replications) analysis was performed to establish the confidence level of the topology of the consensus tree.

Neutrality test

A commonly used approach for detecting selection is a neutrality test based on allelic frequencies. Selective neutrality tests were negative and insignificant: Tajima's D = -0.68535 (P > 0.1); Fu and Li's D = -0.34824 (P > 0.1); and Fu and Li's F = -0.54036 (P > 0.1) (Table 4).

DISCUSSION

Nuclear DNA markers RAPD, ISSR, and RAMPO reveal high levels of polymorphism, so the effective characterization of date palm at the cultivar level remains problematic (Zehdi et al., 2002; Rhouma et al., 2008; Haider et al., 2012). Non-coding regions of the chloroplast have been used for lower-level taxonomic studies under the assumption that non-coding regions should be under fewer functional constraints than coding regions and should provide greater levels of variation for phylogenetic analysis. Among the first regions to be exploited were

the *trnL-trnF* region (Taberlet et al., 1991) and the *atpB-rbcL* intergenic spacer (Manen and Natali, 1995). In the present study, the chloroplast *psbK-psbI* intergenic spacer was evaluated for the first time as a DNA barcode for distinguishing date palm cultivars. The maximum likelihood substitution matrix using the MEGA file shows the probability of substitution from one base to another. These changes include the substitution of a pyrimidine by a purine or a purine by a pyrimidine (transversion) and the substitution of a purine by a purine or a pyrimidine by a pyrimidine (transition). The *psbK-psbI* spacer sequences had more As and Ts relative to Cs and Gs, which is typical of noncoding DNA (Table 4). Base content may explain the occurrence of a relatively high proportion of transversions. In past substitution studies, it was found that when AT content was high, transversions occurred with higher frequency than in a high GC context (Bakker et al., 2000). In the present study, the relatively high AT values in the *psbK-psbI* sequences of date palm cultivars are consistent with those obtained by Akhtar et al. (2014). On the other hand, the number of parsimony-informative sites reached a value of 2, with five haplotypes among the 30 cultivars analyzed. This result is in contrast to the higher level of genetic variation (10 informative sites) found in 31 Tunisian date palm varieties using the *trnL-trnF* region (Rhouma et al., 2014). Pairwise genetic diversity among date palm cultivars using the *psbK-psbI* spacer was close to the values found for the *trnL-trnF s*pacer in fig (Baraket et al., 2009) and the *trnD-trnT* spacer in *Cunninghamia* (Lu et al., 2001). The relatively low level of polymorphism in cpDNA may be associated with high selection pressure by farmers to maintain pure stock or the restricted distribution of date palm in a specific area. In fact, the ecotypes are grouped into two main clusters, suggesting a narrow genetic basis among the cultivars studied despite their phenotypic distinctiveness. This concurs with the molecular phylogeny of eight Saudi date palm cultivars as determined by their cpDNA *rpoB* coding and *psbA-trnH* non-coding regions (Al-Qurainy et al. 2011). A negative Tajima's D test (Tajima, 1989) signifies an excess of low-frequency polymorphisms relative to expectation (Eswaran et al., 2005), indicating population size expansion and/or purifying selection. Tajima's test is based on the fact that under the neutral model, estimates of the number of segregating/polymorphic sites and of the average number of nucleotide differences are correlated. If the value of D is too large or too small, the neutral 'null' hypothesis is rejected. The high genetic similarity values in the present study suggest that date palm has been under high selection pressure. Tajima's and Fu and Li's tests rejected the neutrality hypothesis in the observed sample, indicating that there has been a recent and rapid expansion in the tree's evolutionary history.

In conclusion, the data showed that Emirati date palm cultivars are characterized by their genetic diversity. Despite low genetic variation levels in *Phoenix dactylifera* cultivars, the *psbK-psbI* barcode was not only the most reliable alternative for the identification of the cultivars but it was also useful for explaining the phylogenetic relationships at this taxonomic level.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

I would like to thank Khadija A. Alawanat, Ohood Almehrezi, Khadeijah Al-Zahmi, Maryam Qambar, Dima Al-Jawarneh, and Hessa Khasouni from the Biology Department, College of Science for providing plant materials.

REFERENCES

Akhtar W, Rasheed A, Shinwari ZK, Naqvi SMS, et al. (2014). Genetic characterization of different Pakistani date palm varieties. *Pak. J. Bot.* 46: 2095-2100.

Al-Qurainy F, Khan S, Al-Hemaid FM, Ali MA, et al. (2011). Assessing molecular signature for some potential date (*Phoenix dactylifera* L.) cultivars from Saudi Arabia, based on chloroplast DNA sequences *rpoB* and *psbA-trnH. Int. J. Mol. Sci.* 12: 6871-6880. http://dx.doi.org/10.3390/ijms12106871

Bakker FT, Culham A, Gomez-Martinez R, Carvalho J, et al. (2000). Patterns of nucleotide substitution in angiosperm cpDNA trnL (UAA)-trnF (GAA) regions. *Mol. Biol. Evol.* 17: 1146-1155. http://dx.doi.org/10.1093/oxfordjournals.molbev.a026397

Baraket G, Olfa S, Chatti K, Messaoud M, et al. (2009). Chloroplast DNA analysis in Tunisian fig cultivars (*Ficus carica* L.): Sequence variations of the *trnL-trnF* intergenic spacer. *Biochem. Syst. Ecol.* 36: 828-835. http://dx.doi.org/10.1016/j.bse.2008.09.005

CBOL Plant Working Group (2009). A DNA barcode for land plants. *Proc. Natl. Acad. Sci. USA* 106: 12794-12797.

Chen S, Yao H, Han J, Liu C, et al. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS One* 5: e8613. http://dx.doi.org/10.1371/journal.pone.0008613

Enan MR and Ahamed A (2014). DNA barcoding based on plastid *matK* and RNA polymerase for assessing the genetic identity of date (*Phoenix dactylifera* L.) cultivars. *Genet. Mol. Res.* 13: 3527-3536. http://dx.doi.org/10.4238/2014.February.14.2

Eswaran V, Harpending H and Rogers AR (2005). Genomics refutes an exclusively African origin of humans. *J. Hum. Evol.* 49: 1-18. http://dx.doi.org/10.1016/j.jhevol.2005.02.006

Fu YX and Li WH (1993). Statistical tests of neutrality of mutations. *Genetics* 133: 693-709.

Gothwal RK, Bhargava R, Yadav PK, Meghwal RR, et al. (2013). Evolutionary relationship study in date palm cultivars using morphological and biochemical parameters. *The Bioscan* 8: 1251-1254.

Hadrami E and Al-Kjhayri M (2012). Socioeconomic and traditional importance of date palm. *Emirates J. Food Agric* 24: 371-385.

Haider N, Nabulsi I and Ali M (2012). Phylogenetic relationships among date palm (*Phoenix dactylifera* L.) cultivars in Syria using RAPD and ISSR markers. *J. Plant Biol. Res.*: 12-24.

Hajibabaei M, Singer GAC, Hebert PDN and Hickey DA (2007). DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends Genet.* 23: 167-172. http://dx.doi.org/10.1016/j.tig.2007.02.001

Hodel DR and Johnson DV (2007). Imported and American varieties of dates (*Phoenix dactylifera* L.) in the United States. UC ANR publication 3498. University of California, Oakland, California, 112.

Jamil I, Qamarunnisa S, Azhar A, Shinwari ZK, et al. (2014). Subfamilial relationships within Solanaceae as inferred from *atpβ-rbcL* intergenic spacer. *Pak. J. Bot.* 46: 585-590.

Jaradat AA and Zaid A (2004). Quality traits of date palm fruits in a center of origin and center of diversity. *Food Agr. Environ* 2: 208-217.

Jeanson ML, Labat J-N and Little DP (2011). DNA barcoding: a new tool for palm taxonomists? *Ann. Bot. (Lond.)* 108: 1445-1451. http://dx.doi.org/10.1093/aob/mcr158

Khan S and Bi Bi T (2012). Direct shoot regeneration system for date palm (*Phoenix dactylifera* L.) cv. Dhakki as a means of micropropagation. *Pak. J. Bot.* 44: 1965-1971.

Khierallah H, Bader S, Baum M and Hamwieh A (2011). Assessment of genetic diversity for some Iraqi date palms (*Phoenix dactylifera* L.) using amplified fragment length polymorphisms (AFLP) markers. *Afr. J. Biotechnol.* 10: 9570-9576.

Lee HL, Yi DK, Kim JS and Kim KJ (2007). Development of plant DNA barcoding markers from the variable noncoding regions of chloroplast genome. The second international barcode of life conference, Taipei, Taiwan.

Librado P and Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452. http://dx.doi.org/10.1093/bioinformatics/btp187

Lu SY, Peng CI, Cheng YP, Hong KH, et al. (2001). Chloroplast DNA phylogeography of *Cunninghamia konishii* (Cupressaceae), an endemic conifer of Taiwan. *Genome* 44: 797-807. http://dx.doi.org/10.1139/gen-44-5-797

Manen JF and Natali A (1995). Comparison of the evolution of ribulose-1, 5-biphosphate carboxylase (rbcL) and atpB-rbcL noncoding spacer sequences in a recent plant group, the tribe Rubieae (Rubiaceae). *J. Mol. Evol.* 41: 920-927. http://dx.doi.org/10.1007/BF00173172

Mirbahar AA, Markhand GS, Khan S and Abul-soad AA (2014). Molecular characterization of some Pakistani date palm (*Phoenix dactylifera* L.) cultivars by RAPD markers. *Pak. J. Bot.* 46: 619-625.

Naeem A, Khan AA, Cheema HMN, Khan IA, et al. (2014). DNA barcoding for species identification in the Palmae family. *Genet. Mol. Res.* 13: 10341-10348. http://dx.doi.org/10.4238/2014.December.4.29

Rhouma S, Dakhlaoui-dkhil S, Salem AOM, Azouzi SZ, et al. (2008). Genetic diversity and phylogenic relationships in date palms (*Phoenix dactylifera* L.) as assessed by random amplified microsatellite polymorphism markers (RAMPO). *Sci. Hortic. (Amsterdam)* 117: 53-57. http://dx.doi.org/10.1016/j.scienta.2008.03.013

Rhouma S, Sarra C, Zehdi SA, Khaled C, et al. (2014). Molecular polymorphism and phylogenetic relationships within Tunisian date palm (*Phoenix dactylifera* L.): Evidence of non-coding *trn*L-*trn*F regions of chloroplast DNA. *Sci. Hortic. (Amsterdam)* 170: 32-38. http://dx.doi.org/10.1016/j.scienta.2014.02.027

Sedra MH, Lashermes P, Trouslot P, Combes M, et al. (1998). Identification and genetic diversity analysis of date palm (*Phoenix dactylifera* L.) varieties of Morocco using RAPD markers. *Euphytica* 103: 75-82. http://dx.doi.org/10.1023/A:1018377827903

Taberlet P, Gielly L, Pautou G and Bouvet J (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol. Biol.* 17: 1105-1109. http://dx.doi.org/10.1007/BF00037152

Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595.

Tamura K (1992). Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. *Mol. Biol. Evol.* 9: 678-687.

Tamura K, Stecher G, Peterson D, Filipski A, et al. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30: 2725-2729. http://dx.doi.org/10.1093/molbev/mst197

Vayalil PK (2012). Date fruits (Phoenix dactylifera Linn): an emerging medicinal food. *Crit. Rev. Food Sci. Nutr.* 52: 249-271. http://dx.doi.org/10.1080/10408398.2010.499824

Yao H, Song J, Liu C, Luo K, et al. (2010). Use of ITS2 region as the universal DNA barcode for plants and animals. *PLoS One* 5: e13102. http://dx.doi.org/10.1371/journal.pone.0013102

Yusuf, AO, Culham A., Aljuhani W, Ataga1 CD, et al. (2015). Genetic Diversity of Nigerian Date Palm (*Phoenix dactylifera*) Germplasm based on Microsatellite Markers. *J. Biosci. Biotechnol.* 7: 121-132.

Zehdi SM, Trifi A and Ould Salem M (2002). Survey of inter simple sequence repeat polymorphisms in Tunisian date palms (*Phoenix dactylifera* L.). *J. Genet. Breed.* 56: 77-83.