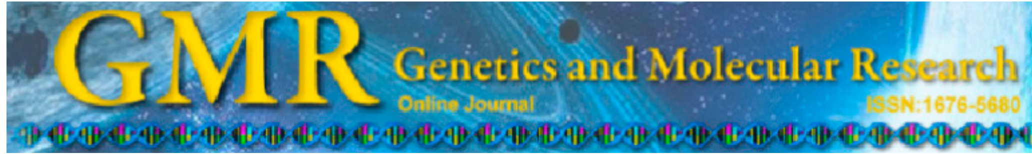

DNA barcoding based on plastid *mat*K and RNA polymerase for assessing the genetic identity of date (*Phoenix dactylifera* L.) cultivars

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. 13 (2): 3527-3536 (2014)
Received May 17, 2013
Accepted October 30, 2013
Published February 14, 2014
DOI http://dx.doi.org/10.4238/2014.February.14.2

ABSTRACT. The cultivated date palm is the most agriculturally important species of the Arecaceae family. The standard chloroplast DNA barcode for land plants recommended by the Consortium for the Barcode of Life plant working group needs to be evaluated for a wide range of plant species. Therefore, we assessed the potential of the *mat*K and *rpo*C1 markers for the authentication of date cultivars. There is not one universal method to authenticate date cultivars. In this study, 11 different date cultivars were sequenced and analyzed for *mat*K and *rpo*C1 genes by using bioinformatic tools to establish a cultivar-specific molecular monogram. The chloroplast *mat*K marker was more informative than the *rpo*C1 chloroplast DNA markers. Phylogenetic trees were constructed on the basis of the *mat*K and *rpo*C1 sequences, and the results suggested that *mat*K alone or in combination with *rpo*C1 can be used for determining the levels of genetic variation and for barcoding.

Key words: Arecaceae; Barcoding; *Phoenix dactylifera*; *mat*K; *rpo*C1

 ©FUNPEC-RP www.funpecrp.com.br

INTRODUCTION

Phoenix dactylifera L. (2n = 36) is a dioecious perennial crop belonging to the Arecaceae family and commonly known as date palm. It is widely cultivated in many countries extending from North Africa to the Middle East, including many states of the Arabian Gulf Cooperation Countries. Many date cultivars, owing to their diverse therapeutic properties, are utilized in traditional medicine (Bulpitt et al., 2007; Sghaier-Hammami et al., 2009). Date palms are also used as sources of food, farm income, and other products for local desert dwellers. All parts of the date tree are used for various industrial purposes. The extracts of date fruits have been reported to have biologically active antioxidants and antimutagenic activities (Biglari et al., 2008; Saafi et al., 2009).

The term DNA barcode was first coined by Hebert et al. (2003) and has gained worldwide attention in the scientific community (Chen et al., 2010). Currently, chloroplast DNA markers are used for the accurate identification and authentication of plant species (Chase et al., 2005; CBOL Plant Working Group, 2009; Devey et al., 2009; Fazekas et al., 2009; Chen et al., 2010; Al-Qurainy et al., 2011a). Variations in DNA sequences are very helpful in the development of unique markers, which can be used as a DNA barcode for that species. Many loci from the plastid genome, including *rbcL*, *rpoB*, *rpo*C1, *trnH-psbA* spacer, and *mat*K, have been tested for DNA barcoding of plants with different degrees of success (Kress et al., 2005; Lahaye et al., 2008; Hollingsworth et al., 2011). Thus far, no consensus sequence has been identified as a universal barcode in land plants. For the accurate and reproducible identification of species, analysis of more than one locus would be required (Kress et al, 2005; Kress and Erickson, 2007; Lahaye et al., 2008; Fazekas et al., 2008; CBOL Plant Working Group, 2009). Sequencing- and non-sequencing-based markers have been used in many medicinal and non-medicinal plant species for the detection of adulterants in the local herbal markets (Khan et al., 2011; Al-Qurainy et al., 2011b). CBOL Plant Working Group (2009) recommended two-locus combination of *mat*K and *rbcl* as the plant barcode. The morphological as well as biochemical markers used in the identification of plant species have many limitations due to their low reproducibility. Moreover, visually differentiating seeds of different species is difficult (Khan et al., 2011). Therefore, the main objective of this study is to assess the efficiency of 2 loci, *mat*K and *rpo*C1, as barcodes for the precise authentication of date cultivars.

MATERIAL AND METHODS

Plant material

Eleven date samples were collected from Al Ain city, United Arab Emirates. Plant material consisted of young leaves sampled from adult trees (Table 1). Leaf samples were individually placed in plastic pouches and transported to the laboratory and stored at -80°C until processing for DNA extraction.

DNA extraction

Leaf samples were immersed in liquid nitrogen and crushed using sterile mortar and pestle to obtain a fine powder. DNeasy plant mini kit (Qiagen, Germany) was used for DNA

isolation. The quality of the extracted DNA was determined using gel electrophoresis and Nanodrop 2000 spectrophotometer (Thermo Scientific, USA).

Table 1. Sources of samples, voucher information, and GenBank database accession numbers of DNA sequences of taxa used in the present study.

Cultivars	Abbreviation	*mat*K			*rpo*C1	
		Locality	Accession No.	G+C (%)	Accession No.	G+C (%)
AbuMaan	PD1	Al Ain, UAE	KC437393	33.8	KC793979	42.1
Ngal	PD2	Al Ain, UAE	KC771273	34.5	-	-
Fard	PD4	Al Ain, UAE	KC437392	34.5	KC793975	42.2
Khenezi	PD5	Al Ain, UAE	KC437387	34.5	KC793976	42.2
Khalas	PD6	Al Ain, UAE	KC437389	34.5	KC793980	46.2
Khalasuae	PD7	Al Ain, UAE	KC771274	30.4	KC793978	42.5
Gashzabad	PD8	Al Ain, UAE	KC437394	33.8	KC793972	42.2
Hilali	PD9	Al Ain, UAE	KC437390	34.5	KC793974	42.2
Nmishi	PD10	Al Ain, UAE	KC437388	34.5	KC793973	42.2
Barhi	PD11	Al Ain, UAE	KC771272	34.8	-	-
Jaberi	PD12	Al Ain, UAE	KC437391	34.5	KC793977	42.2

(-) = sequencing of PCR products failed.

Amplification of the *mat*K and *rpo*C1 genes

A total volume of 25 µL of PCR mixture contained the following: 12.5 µL Taq PCR Master Mix (Qiagen), yielding a final concentration of 200 µM of each deoxynucleotide and 1.5 mM $MgCl_2$, 1 µM of each primer (**Table S1**; Eurofins MWG Operon, Germany), 2 µL (50 ng) genomic DNA, and the rest was adjusted with sterile distilled water. PCR amplification was performed using a T100 thermal cycler (BioRad, USA) as follows: 95°C for 1 min, followed by 35 cycles of 95°C for 30 s, 50°C for 30 s, and 72°C for 60 s, followed by an elongation step at 72°C for 5 min. All PCR conditions were used as described previously (Sass et al., 2007; Yu et al., 2011). The universal primers for the 2 loci are listed in **Table S1**.

Agarose gel electrophoresis

A long (14 x 11 cm) 1.5% agarose gel in 1X TBE buffer containing 0.5 µg/mL ethidium bromide was used for electrophoresis of the PCR products. Gel images were obtained using gel document (Major Bioscience, Taiwan) UV transilluminator imaging system. The sizes of the PCR products resulting from the primer pairs of the specific barcoding gene were determined using a 100-bp ladder (Qiagen) and Un-Scan-It gel version 6.1 (Silk Scientific Inc., USA).

DNA sequencing and data analysis

The sequencing reaction was performed using a BigDye Terminator cycle sequencing kit (Perkin-Elmer, Applied Biosystems) according to manufacturer instructions for the ABI PRISM 310 DNA Analyzer (Perkin-Elmer, Applied Biosystems). All sequences generated in the present study were deposited in GenBank for reference; their accession numbers are provided in Table 1. The Basic Local Alignment SearchTool (BLAST) was used to detect ho-

mologous sequences to those obtained for date cultivars. When the sequences were confirmed to be *mat*K and *rpoC*1, phylogenetic trees were inferred with the maximum likelihood (ML), neighbor-joining tree (NJ), and UPGMA methods. The topologies of the phylogenetic trees were evaluated using the bootstrap resampling method with 1000 replicates. Codon positions included were 1st + 2nd + 3rd + noncoding. Pairwise distance, transitional/transversional substitutions, and phylogenetic analyses were conducted using MEGA5.0 (Tamura et al., 2004, 2007). Genetic variation among date cultivars was estimated by calculating the number of polymorphic sites and mutations, haplotype diversity, and nucleotide diversity by using the DnaSP software (Librado and Rozas, 2009).

RESULTS AND DISCUSSION

The main objective of this study was to amplify and characterize two loci *rpo*C1 and *mat*K from the chloroplast genome to assess their suitability for the resolution of date cultivars. Electrophoretic analysis of PCR products based on partially amplified *mat*K and *rpo*C1 genes resulted in a single amplified DNA band. In a previous study, the *rpo*C1 locus was not successfully amplified from the date palm cultivars (Al-Quariny et al., 2011a). The amplification of *mat*K was not problematic, as suggested by Kress and Erickson (2007).

The phylogenetic relationships among the date palm cultivars have been evaluated in different countries such as Tunisia, USA (California), and Morocco by using various conventional molecular markers such as ISSR, AFLP, RAPD (Sedra et al., 1998; Al-Khalifah and Askari, 2003; Elshibli and Korpelainen, 2008), and microsatellites (Zehdi et al., 2004; Elshibli and Korpelainen, 2008). These markers showed high polymorphism among the date palm cultivars, but were ineffective in characterizing them. Genetically, the date palm is highly diverse due to existence of large number of cultivars distributed across different habitats (Khan et al., 2011). The chloroplast genome has been used successfully for the identification of various cultivars of date palm in Saudi Arabia (Al-Qurainy et al., 2011a). In this study, 3 phylogenetic methods were applied using one barcode locus or in combinations to evaluate the recovery of cultivars. When all the sequences for a given locus were considered, *mat*K could form cultivar-specific clusters. A single cultivar, *Khalasuae* (United Arab Emirates) was recovered as a single genotype in the first cluster. Figures 1, 2, and **Figure S1** show phylogenetic trees constructed by using UPGMA, ML, and NJ, respectively; the second cluster consisted of *Gashzabad* and *Abu Maan*. All the remaining cultivars were included in the third cluster that had 2 sub-clusters: the first sub-cluster included *khalas* (Saudia Arabia), *Nmishi*, *Hillali*, *Fardi*, *Jaberi*, *Ngal*, and *Barhi*, and the second sub-cluster included *Khenezi*. The cultivar-specific clusters in *mat*K trees resolved with high bootstrap confidence levels (95-100%) by using UPGMA and NJ, while, with ML, the bootstrap levels were 76-100%. In *rpo*C1 trees constructed by using UPGMA, ML, and NJ, as shown in Figures 3, 4, and **Figure S2**), *khalas* (Saudia Arabia) and *Abu Maan* were placed in the first cluster, whereas the second cluster included *Gashzabad*, *Nmishi*, *Khenezi*, *Jaberi*, *Hillali*, *Fardi*, and *Ngal*; *Khalasuae* was included in a separate sub-cluster. Bootstrap confidence levels were 88-100% with *rpo*C1 barcode when all 3 phylogenetic methods were used. ML showed the same tree topology as UPGMA and NJ with single locus analysis for the recovery of date cultivars. It can be assumed that, if different methods yield the same tree, then a robust estimation will become possible (Hosseini et al., 2012). The *mat*K locus provided better cultivar recovery compared to *rpo*C1. We compared the performance of combined *mat*K +

*rpo*C1 barcodes by using unweighted pair group (Figure 5), ML (Figure 6), and NJ (**Figure S3**) methods; the results indicated that there was no marked increase in the recovery of cultivars. In a recent study, the *rpo*C1 marker exhibited the lowest level of resolution among the evaluated regions (*mat*K > *atpF*-*atpH* > *rbcL* > *trnH*-*psbA* > *rpo*C1; Burgess et al., 2011). Phylogenetic methods were applied in a recently conducted study of barcoding species by using each barcode locus alone or in combination with others to evaluate species recovery (Roy et al., 2010). The *mat*K region has been used to construct the phylogeny of legumes and species belonging to the Fabaceae family (Wojciechowski et al., 2004; Gao et al., 2011). In a previous study, *rpo*C1 and *psbA trnH* GUG *ndhF* were used to construct the phylogeny of the genus *Lathyrus* (Asmussen and Liston, 1998).

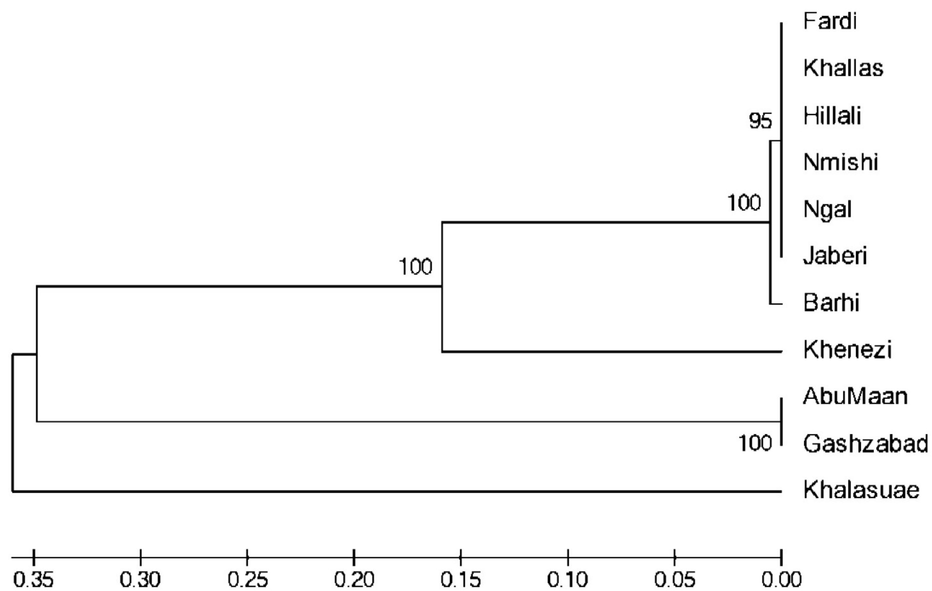

Figure 1. UPGMA tree based on Kimura 2-parameter of the *mat*K. The branch support was assessed with 1000 replicates. The support values above 50% were shown.

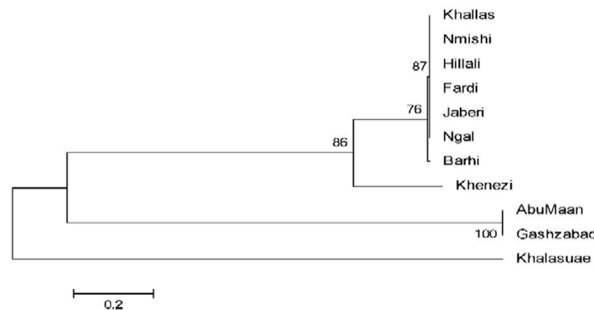

Figure 2. Maximum-likelihood tree constructed from nucleotide sequences of the *mat*K gene. Bootstrap values of the 50% majority rule are indicated and 1000 bootstrap replicates were used.

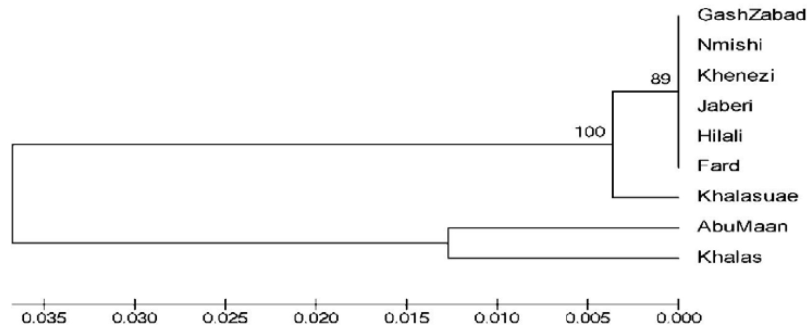

Figure 3. UPGMA tree based on Kimura 2-parameter of the *rpo*C1. The branch support was assessed with 1000 replicates. The support values above 50% were shown.

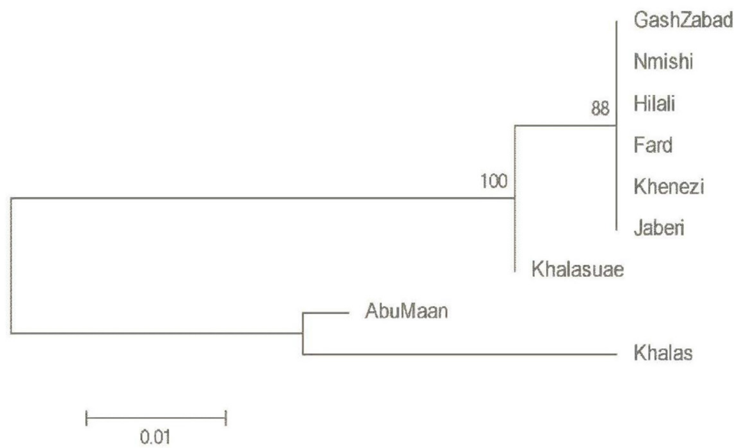

Figure 4. Maximum-likelihood tree constructed from nucleotide sequences of the *rpo*C1 gene. Bootstrap values of the 50% majority rule are indicated and 1000 bootstrap replicates were used.

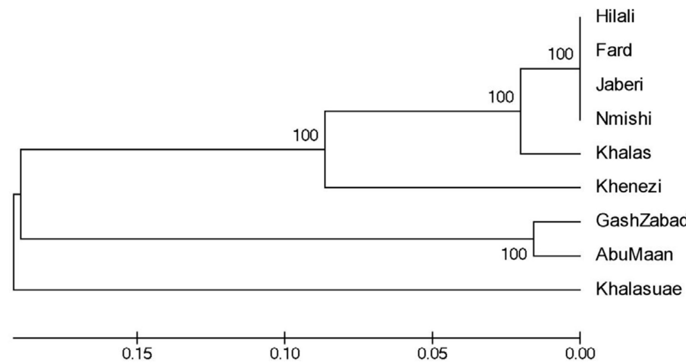

Figure 5. UPGMA tree based on Kimura 2-parameter using combined chloroplast loci (*mat*K + *rpo*C1). The branch support was assessed with 1000 replicates. The support values above 50% were shown.

©FUNPEC-RP www.funpecrp.com.br

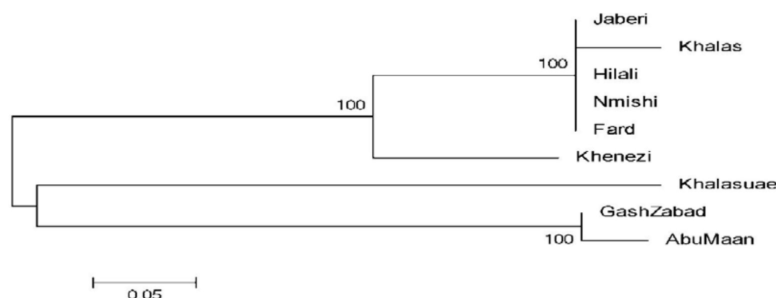

Figure 6. Maximum-likelihood tree constructed from nucleotide sequences using combined chloroplast loci (*mat*K + *rpo*C1). Bootstrap values of the 50% majority rule are indicated and 1000 bootstrap replicates were used.

Traditionally, cultivars and species were identified and characterized on the basis of morphological and physiological traits, which are sometimes difficult to discriminate (Kadkhodaei et al., 2010). DNA sequence-based identification is a more accurate method and has been used in many studies (Heinrich, 2008; Liu et al., 2011). The sequences generated in the present study, namely *mat*K and *rpo*C1, were deposited in GenBank. All sequences had few variations in the percent of guanine plus cytosine content (%GC) compared to that in the sequences of *mat*K and *rpo*C1 (Table 1). In the case of *mat*K, the nucleotide composition was biased toward the guanine and cytosine content with frequencies of 30.4 to 34.8%, respectively. In the case of *rpo*C1, the %GC content was 42.2 to 46.2% (Table 1). Similarly, Khan et al. (2012) found that the base composition of the *rpo*C1 gene sequence of *Ochradenus* species was 42-43%. The generated sequences of *mat*K for date cultivars were compared with the generated sequences of *rpo*C1, and significant results were found in terms of sequence divergence. In this study, no insertions/deletions were found (InDels) either in *mat*K or in *rpo*C1 (**Figure S4**). Similarly, no InDels were detected in the *rpo*C1 gene sequences of angiosperms (Samigullin et al., 1999). Internal spacers of the *rpo*C1 gene have been reported to have 6 base deletions and 11 base insertions (Khan et al., 2012). Lack of InDels in *rpo*C1 is attributed to its coding property, while the internal spacer of *rpo*C1 located in the non-coding region. The genetic distances among the 11 cultivars revealed by *mat*K sequences ranged from 0.00 to 0.722 (Table 2). The overall average distance (0.375) was higher than that found with *rpo*C1 (0.0305; Table 3) or when the combination of *mat*K + *rpo*C1 was used (0.241; Table 4). Our data are in agreement with those of Rhouma et al. (2008), where the genetic distance exhibited values ranging from 0.10-0.76 with mean of 0.34 in 40 date cultivars studied using random amplified microsatellite polymorphism markers. In the current study, for *mat*K, the genetic distance (0.375) among date cultivars was more than 10 times that of the *rpo*C1 distance (0.0305). The corresponding parsimony informative characters (2 variants) were 94, 17, and 112 with *mat*K, *rpo*C1, and *mat*K + *rpo*C1, respectively, and the total number of mutations were 419, 24, and 441, respectively, and the numbers of variable sites were 278, 24, and 302, respectively (Table 5). The overall transition to transversion ratio was 0.54, 0.51, and 0.48 with *mat*K, *rpo*C1, and combined *mat*K + *rpo*C1 (Table 5). When the overall outputs of pairwise distance and tree analyses were compared, the latter strategy resulted in better resolution of cultivars. In our study, *rpo*C1 was considered to possess less cultivar-discriminating power than *mat*K, possibly due to its minimal sequence variation (Hollingsworth et al., 2011).

Table 2. Pairwise distance among the date palm cultivar revealed by *mat*K.

	Barhi	Ngal	Jaberi	Fardi	Hilali	Khenezi	Khalas	Abu Maan	Gashzabad	Nmishi	Khalasuae
Barhi											
Ngal	0.01014										
Jaberi	0.01014	0.00000									
Fard	0.01014	0.00000	0.00000								
Hillali	0.01014	0.00000	0.00000	0.00000							
Khenezi	0.31757	0.31757	0.31757	0.31757	0.31757						
Khallas	0.01014	0.00000	0.00000	0.00000	0.00000	0.31757					
Abu Maan	0.69932	0.69595	0.69595	0.69595	0.69595	0.70270	0.69595				
Gashzabad	0.69932	0.69595	0.69595	0.69595	0.69595	0.70270	0.69595	0.00000			
Nmishi	0.01014	0.00000	0.00000	0.00000	0.00000	0.31757	0.00000	0.69595	0.69595		
Khalasuae	0.72297	0.71959	0.71959	0.71959	0.71959	0.71284	0.71959	0.72297	0.72297	0.71959	

Overall average distance = 0.375.

Table 3. Pairwise distance among the date palm cultivar revealed by *rpo*C1**.**

	Gashzabad	Nmishi	Hilali	Fard	Khenezi	Jaberi	Khalasuae	Abu Maan	Khalas
Gashzabad									
Nmishi	0.000								
Hilali	0.000	0.000							
Fard	0.000	0.000	0.000						
Khenezi	0.000	0.000	0.000	0.000					
Jaberi	0.000	0.000	0.000	0.000	0.000				
Khalasuae	0.007	0.007	0.007	0.007	0.007	0.007			
Abu Maan	0.069	0.069	0.069	0.069	0.069	0.069	0.061		
Khalas	0.090	0.090	0.090	0.090	0.090	0.090	0.081	0.026	0.000

Overall average distance = 0.0305.

Table 4. Pairwise distance among the date palm cultivar revealed by combined *mat*K and *rpo*C1 loci.

	Gashzabad	Nmishi	Hilali	Fard	Khenezi	Jaberi	Khalasuae	Abu Maan	Khalas
Gashzabad									
Nmishi	0.36077								
Hilali	0.36077	0.00000							
Fard	0.36077	0.00000	0.00000						
Khenezi	0.36427	0.16462	0.16462	0.16462					
Jaberi	0.36077	0.00000	0.00000	0.00000	0.16462				
Khalasuae	0.37828	0.37653	0.37653	0.37653	0.37303	0.37653			
Abu Maan	0.03152	0.39229	0.39229	0.39229	0.39580	0.39229	0.40280		
Khalas	0.40105	0.04028	0.04028	0.04028	0.20490	0.04028	0.40981	0.37303	

Overall average = 0.241.

Table 5. Summary characteristic of the two chloroplast markers evaluated in this study.

Marker	Aligned sequence length	Variable sites (S)	Percentage parsimony informative sites (two variants)	Nucleotide diversity (per site)	Total number of mutation (Eta)	Theta (per site) from Eta	Theta (per site) from S	Transition/ transversion bias	Haplotype (gene) diversity
*mat*K	296	278	31.76	0.37525	419	0.48329	0.82427	0.54	0.709
*rpo*C1	275	24	6.18	0.03051	24	0.03211	0.03411	0.51	0.583
*mat*K + *rpo*C1	571	302	19.61	0.24090	441	0.2841	0.30198	0.48	0.833

The *mat*K locus showed more polymorphic sites than the *rpo*C1 locus; hence, it was more informative and proved to be very effective in differencing the date palm cultivars. The *mat*K locus has been shown to provide high level of species recovery in several plant DNA barcoding studies on different floristic or biodiversity hotspots (Kress and Erickson, 2007; Lahaye

 ©FUNPEC-RP www.funpecrp.com.br

et al., 2008; CBOL, 2009; Chen et al., 2010) as in the present study. The genetic divergence evaluated among the date palm cultivars might be due to the dispersal of off-shoots, pollen grains, and seeds (Al-Qurainy et al., 2011a). Off-shoots and pollen grains are extensively distributed among farmers within a village, province, or country, while seed dispersal occurs by other means, such as via travelers and traders, across geographic borders (Elshibli and Korpelainen, 2008). The other reasons for sequence variability among the date cultivars could be due to the different mating frequencies, mutation rate, gene flow patterns, long-term evolution history, and human activities; these factors affect the genetic variation patterns among plant populations. Furthermore, environmental factors might be one of the reasons for variability (Nybom and Bartish, 2000). Haplotype (gene) diversity among date cultivars was 0.709, 0.583, and 0.833 with *mat*K, *rpo*C1, and combined *mat*K + *rpo*C1, respectively (Table 5). We found that the topology of phylogenetic trees constructed on the basis of *mat*K, *rpo*C1, and *mat*K + *rpo*C1 sequences were coherent with the number of haplotypes detected. The presence of more than one haplotype (**Table S2**) in the cultivars of *Phoenix dactylifera* from UAE might be explained as follows: a) the introduced cultivars had different origins; b) it originated from only one source, either through different introduction events or from a single event containing more than one haplotype. The local cultivar *Khalasuae* was separated as a single chloroplast haplotype. This might be explained by a unique introduction or multiple introductions of the same haplotype. Our results showed that the *Khalasuae* haplotype was not similar to *khalas* from Suadia Arabia. In conclusion, applying both nucleotide distance and phylogeny-based approaches, we found that *mat*K locus, either singly or in combination, could discriminate date cultivars.

ACKNOWLEDGMENTS

We thank Dr. Khaled Amiri, Head of the Biology Department at United Arab Emirates University for his help and support.

Supplementary material

REFERENCES

Al-Khalifah NS and Askari E (2003). Molecular phylogeny of date palm (*Phoenix dactylifera* L.) cultivars from Saudi Arabia by DNA fingerprinting. *Theor. Appl. Genet.* 107: 1266-1270.

Al-Qurainy F, Khan S, Al-Hemaid FM, Ali MA, et al. (2011a). 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.

Al-Qurainy F, Khan S, Tarroum M, Al-Hemaid FM, et al. (2011b). Molecular authentication of the medicinal herb *Ruta graveolens* (Rutaceae) and an adulterant using nuclear and chloroplast DNA markers. *Genet. Mol. Res.* 10: 2806-2816.

Asmussen C and Liston A (1998). Chloroplast DNA characters, phylogeny, and classification of *Lathyrus* (Fabaceae). *Am. J. Bot.* 85: 387.

Biglari F, AlKarkhi AFM and Easa AM (2008). Antioxidant activity and phenolic content of various date palm (*Phoenix dactylifera*) fruits from Iran. *Food Chem.* 107: 1636-1641.

Bulpitt CJ, Li Y, Bulpitt PF and Wang J (2007). The use of orchids in Chinese medicine. *J. R. Soc. Med.* 100: 558-563.

Burgess KS, Fazekas AJ, Kesanakurti PR, Graham SW, et al. (2011). Discriminating plant species in a local temperate flora using the *rbcL+ matK* DNA barcode. *Meth. Ecol. Evol.* 2: 333-340.

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

Chase MW, Salamin N, Wilkinson M, Dunwell JM, et al. (2005). Land plants and DNA barcodes: short-term and long-term goals. *Philos. Trans. R. Soc. Lond B. Biol. Sci.* 360: 1889-1895.

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.

Devey DS, Chase MW and Clarkson JJ (2009). A stuttering start to plant DNA barcoding: microsatellites present a

previously overlooked problem in non-coding plastid regions. *Taxon* 58: 7-15.
Elshibli S and Korpelainen H (2008). Microsatellite markers reveal high genetic diversity in date palm (*Phoenix dactylifera* L.) germplasm from Sudan. *Genetica* 134: 251-260.
Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, et al. (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS One* 3: e2802.
Fazekas AJ, Kesanakurti PR, Burgess KS, Percy DM, et al. (2009). Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? *Mol. Ecol. Resour.* 9 (Suppl 1): 130-139.
Gao T, Sun Z, Yao H, Song J, et al. (2011). Identification of Fabaceae plants using the DNA barcode matK. *Planta Med.* 77: 92-94.
Hebert PD, Cywinska A, Ball SL and deWaard JR (2003). Biological identifications through DNA barcodes. *Proc. Biol. Sci.* 270: 313-321.
Heinrich M (2008). Ethnopharmacy and natural product research-multidisciplinary opportunities for research in the metabolomic age. *Phytochem. Lett.* 1: 1-5.
Hollingsworth PM, Graham SW and Little DP (2011). Choosing and using a plant DNA barcode. *PLoS One* 6: e19254.
Hosseini S, Go R, Dadkhah K and Nuruddin AA (2012). Studies on maturase k sequences and systematic classification of *bulbophyllum* in Peninsular Malaysia. *Pak. J. Bot.* 44: 2047.
Kadkhodaei S, Elahy M, Nekouei MK, Imani A, et al. (2010). A panel of cultivate specific marker based on polymorphisms at microsatellite markers for Iranian cultivated Almonds (*Prunus dulcis*). *Aust. J. Crop Sci.* 4: 730-736.
Khan S, Mirza KJ and Abdin MZ (2011). DNA fingerprinting for the authentication of *Ruta graveolens*. *Afr. J. Biotechnol.* 10: 8709-8715.
Khan S, Al-Qurainy F and Nadeem M (2012). Biotechnological approaches for conservation and improvement of rare and endangered plants of Saudi Arabia. *Saudi J. Biol. Sci.* 19: 1-11.
Kress WJ and Erickson DL (2007). A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS One* 2: e508.
Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, et al. (2005). Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci. U. S. A.* 102: 8369-8374.
Lahaye R, van der Bank M, Bogarin D, Warner J, et al. (2008). DNA barcoding the floras of biodiversity hotspots. *Proc. Natl. Acad. Sci. U. S. A.* 105: 2923-2928.
Librado P and Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.
Liu C, Gu Z, Yang W, Yang L, et al. (2011). Advances of biological taxonomy and species identification in medicinal plant species by DNA barcodes. *J. Am. Sci.* 7: 147-151.
Nybom H and Bartish IV (2000). Effects of life history traits and sampling strategies on genetic diversity estimates obtained with RAPD markers in plants. *Perspect Plant Ecol.* 3: 93-114.
Rhouma S, Dakhlaoui-Dkhil S, OuldMohamed Salem A and Zehdi Azouzi S (2008). Genetic diversity and phylogenic relationships in date-palms (*Phoenixdactylifera* L) as assessed by random amplified microsatellite polymorphism markers (RAMPOs). *Sci. Hort.* 117: 53-57.
Roy S, Tyagi A, Shukla V, Kumar A, et al. (2010). Universal plant DNA barcode loci may not work in complex groups: a case study with Indian berberis species. *PLoS One* 5: e13674.
Saafi EB, El-Arem A, Issaoui M and Hammami M (2009). Phenolic content and antioxidant activity of four date palm (*Phoenix dactylifera* L) fruit varieties grown in Tunisia. *Int. J. Food Sci.* 44: 2314-2319.
Samigullin TK, Martin WF, Troitsky AV and Antonov AS (1999). Molecular data from the chloroplast rpoC1 gene suggest a deep and distinct dichotomy of contemporary spermatophytes into two monophyla: gymnosperms (including Gnetales) and angiosperms. *J. Mol. Evol.* 49: 310-315.
Sass C, Little DP, Stevenson DW and Specht CD (2007). DNA barcoding in the cycadales: testing the potential of proposed barcoding markers for species identification of cycads. *PLoS One* 2: e1154.
Sedra MH, Lashermes P, Trouslot P and Combes M (1998). Identification and genetic diversity analysis of date palm (*Phoenix dactylifera* L.) varieties from Morocco using RAPD markers. *Euphytica* 103: 75-82.
Sghaier-Hammami B, Valledor L, Drira N and Jorrin-Novo JV (2009). Proteomic analysis of the development and germination of date palm (*Phoenix dactylifera* L.) zygotic embryos. *Proteomics* 9: 2543-2554.
Tamura K, Nei M and Kumar S (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. U. S. A.* 101: 11030-11035.
Tamura K, Dudley J, Nei M and Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596-1599.
Wojciechowski MF, Lavin M and Sanderson MJ (2004). A phylogeny of legumes (Leguminosae) based on analysis of the plastid *matK* gene resolves many well-supported subclades within the family. *Am. J. Bot.* 91: 1846-1862.
Yu J, Xue JH and Zhou SL (2011). New universal matK primers for DNA barcoding angiosperms. *J. Syst. Evol.* 49: 176-181.
Zehdi S, Trifi M, Billotte N, Marrakchi M, et al. (2004). Genetic diversity of Tunisian date palms (*Phoenixdactylifera* L.) revealed by nuclear microsatellite polymprohism. *Hereditas* 141: 278-287.

 ©FUNPEC-RP www.funpecrp.com.br