ZooKeys 365: 71–82 (2013)
doi: 10.3897/zookeys.365.5725
www.zookeys.org

RESEARCH ARTICLE

A peer-reviewed open-access journal
ZooKeys
Launched to accelerate biodiversity research

The chloroplast DNA locus *psbZ-trnfM* as a potential barcode marker in *Phoenix* L. (Arecaceae)

Marco Ballardini[1], Antonio Mercuri[2], Claudio Littardi[1], Summar Abbas[3], Marie Couderc[4], Bertha Ludeña[4], Jean-Christophe Pintaud[4]

1 *Centro Studi e Ricerche per le Palme - Sanremo (CSRP), Corso F. Cavallotti 113, I-18038 Sanremo (IM), Italy* **2** *Consiglio per la Ricerca e la sperimentazione in Agricoltura - Unità di Ricerca per la Floricoltura e le Specie Ornamentali (CRA-FSO), Corso degli Inglesi 508, I-18038 Sanremo (IM), Italy* **3** *Institute of Horticultural Sciences, University of Agriculture, 38040 Faisalabad, Pakistan* **4** *UMR DIADE/DYNADIV, Institut de Recherche pour le Développement (IRD), 911 Av. Agropolis, F-34394 Montpellier Cedex 5, France*

Corresponding author: *Marco Ballardini* (m_ballardini@hotmail.com)

Academic editor: *K. Jordaens* | Received 1 June 2013 | Accepted 3 December 2013 | Published 30 December 2013

Citation: Ballardini M, Mercuri A, Littardi C, Abbas S, Couderc M, Ludeña B, Pintaud JC (2013) The chloroplast DNA locus *psbZ-trnfM* as a potential barcode marker in *Phoenix* L. (Arecaceae). In: Nagy ZT, Backeljau T, De Meyer M, Jordaens K (Eds) DNA barcoding: a practical tool for fundamental and applied biodiversity research. ZooKeys 365: 71–82. doi: 10.3897/zookeys.365.5725

Abstract

The genus *Phoenix* (*Arecaceae*) comprises 14 species distributed from Cape Verde Islands to SE Asia. It includes the economically important species *Phoenix dactylifera*. The paucity of differential morphological and anatomical useful characters, and interspecific hybridization, make identification of *Phoenix* species difficult. In this context, the development of reliable DNA markers for species and hybrid identification would be of great utility. Previous studies identified a 12 bp polymorphic chloroplast minisatellite in the *trnG*(GCC)-*trnfM*(CAU) spacer, and showed its potential for species identification in *Phoenix*. In this work, in order to develop an efficient DNA barcode marker for *Phoenix*, a longer cpDNA region (700 bp) comprising the mentioned minisatellite, and located between the *psbZ* and *trnfM(CAU)* genes, was sequenced. One hundred and thirty-six individuals, representing all *Phoenix* species except *P. andamanensis*, were analysed. The minisatellite showed 2-7 repetitions of the 12 bp motif, with 1-3 out of seven haplotypes per species. *Phoenix reclinata* and *P. canariensis* had species-specific haplotypes. Additional polymorphisms were found in the flanking regions of the minisatellite, including substitutions, indels and homopolymers. All this information allowed us to identify unambiguously eight out of the 13 species, and overall 80% of the individuals sampled. *Phoenix rupicola* and *P. theophrasti* had the same haplotype, and so had *P. atlantica*, *P. dactylifera*, and *P. sylvestris* (the "date palm complex" *sensu* Pintaud et al. 2013). For these species, additional molecular markers will be required for their unambiguous identification. The *psbZ-trnfM*(CAU) region therefore could be considered as a good basis for the establishment of a DNA barcoding system in *Phoenix*, and is potentially useful for the identification of the female parent in *Phoenix* hybrids.

Keywords
Chloroplast *psbZ-trnfM*(CAU) region, DNA barcode, minisatellite, palms

Introduction

Taxonomy and phylogeny of *Phoenix* L.

The genus *Phoenix* L. (*Arecaceae*) comprises 14 species (Govaerts and Dransfield 2005), distributed from the E Atlantic (Macaronesia), through Africa, the Mediterranean region, S Asia to islands in the Indian Ocean (Madagascar, Andaman) and the NW Pacific (Taïwan and N Philippines). *Phoenix* is morphologically and phylogenetically highly divergent from the other palm genera, and constitutes the monogeneric tribe *Phoeniceae* within the subfamily Coryphoideae (Asmussen et al. 2006, Dransfield et al. 2008). The position of *Phoenix* within the subfamily Coryphoideae has been confirmed by a generic-level phylogenetic analysis of the entire palm family (*Arecaceae*) that included plastid and nuclear DNA sequences, cpDNA RFLPs and morphological data (Baker et al. 2009).

The taxonomy, phylogeny and evolution of the genus itself have been assessed using morphological and molecular approaches. According to Barrow (1998), both morphological, and molecular data of the 5S intergenic spacer of the nuclear ribosomal 5S DNA unit supported the existence of two clades of closely related species. The first clade included *P. dactylifera*, *P. sylvestris*, *P. theophrasti* and *P. canariensis* -the so-called "date-palm complex"-, and *P. atlantica* (Pintaud et al. 2010). The second group comprised the sister species *P. paludosa* and *P. roebelenii*. However, Barrow's (1998) molecular analysis included only 11 out of the 13 species recognized at that time, since *P. atlantica* was left as an insufficiently known taxon. Its status as a valid species was confirmed later by Henderson et al. (2006). Using one plastid and 16 nuclear microsatellite markers, Pintaud et al. (2010) demonstrated that all members of the "date-palm complex" are distinct species. Moreover, their data suggested that *P. atlantica* and *P. dactylifera* were sister species. Unfortunately, *P. paludosa* and *P. andamanensis* were not included in their analyses. Combining sequence data of the chloroplast *psbZ-trnfM* and *rpl16-rps*3 loci, Pintaud et al. (2013) depicted five distinct phylogenetic lineages within *Phoenix* (*P. loureiroi-acaulis-pusilla*, *P. roebelenii-paludosa*, *P. caespitosa*, *P. reclinata*, and *P. rupicola-theophrasti-canariensis-dactylifera-atlantica-sylvestris*), and restricted the "date palm complex" to *P. dactylifera-atlantica-sylvestris*. This complex could be distinguished by the presence of a 3-repetitions haplotype of a 20 bp minisatellite motif at the *rpl*16-*rps*3 locus, that was absent in all other species. *Phoenix andamanensis* was the only taxon not included in their study.

The cultivated date palm *P. dactylifera* L. is the most important fruit crop in the Middle East and North African countries. This species was probably domesticated around 4,000 B.C. in the Mesopotamia-Arabic Gulf area (Nesbitt 1993, Zohary and Hopf 2000, Tengberg 2012) and is nowadays distributed worldwide.

Phoenix species are largely interfertile and many interspecific hybrids have been recognized or suspected (Greuter 1967, Wrigley 1995). The spread of the domesticated *Phoenix dactylifera* resulted in situations of sympatry with wild species, promoting interspecific gene flow, in particular with the endemic *P. canariensis* in the Canary Islands (González-Pérez et al. 2004), and possibly with *P. theophrasti* in Turkey (Boydak and Barrow 1995), *P. atlantica* in the Cape Verde Islands (Henderson et al. 2006), and *P. sylvestris* in NW India (Newton et al. 2013). Moreover, spontaneous and directed hybridization between species is an important aspect of *Phoenix* ornamental cultivation (Tournay 2009).

Added to the common hybridization process between *Phoenix* species, the paucity of systematically useful morphological and anatomical characters within the genus (Barrow 1998), makes it difficult to establish a comprehensive taxonomy of the genus *Phoenix*. Because of this confusing situation, a reliable DNA marker set (barcode) to discriminate among *Phoenix* species and hybrids would be extremely useful.

DNA barcoding

Hebert et al. (2003) introduced the concept of "DNA barcode" as a new approach to taxon recognition, assuming that a short standardised DNA sequence can distinguish individuals of a species because genetic differentiation between species exceeds that within species. Since then, DNA barcoding has become increasingly important as a tool in taxonomic studies and species delimitation, as well as in the discovery of new (cryptic) species (e.g. DeSalle et al. 2005, Hebert et al. 2004, Hebert and Gregory 2005, Savolainen et al. 2005, Hajibabaei et al. 2007). A consortium of scientists suggested the two-locus combination of *rbc*L + *mat*K plastid genes as the universal plant barcode (CBOL 2009), while other authors (Chen et al. 2010, Yao et al. 2010) proposed the ITS2 region as a more efficient barcode. The China Plant BOL Group (2011) highlighted the importance of both sampling multiple individuals and using markers with different modes of inheritance, and suggested to incorporate the ITS1/ITS2 region into the core barcode for seed plants.

However, despite all efforts, no locus (alone or in combination), has proven to be 100% efficient as universal DNA barcode in plants at the species level.

The first DNA barcoding analysis in palms (Jeanson et al. 2011) achieved a 92% success in species discrimination by applying a combination of three markers (the plastid *mat*K and *rbc*L, and the nuclear ITS2) to the tribe *Caryoteae* (subfamily *Coryphoideae*).

Investigating the taxonomic status of *P. atlantica*, in comparison with its close relatives *P. dactylifera*, *P. canariensis* and *P. sylvestris*, Henderson et al. (2006) identified a polymorphic cpDNA minisatellite locus, situated within the *trnG*(GCC)-*trnfM*(CAU) intergenic spacer. Its structure was based on the 12 bp motif CTAACTACTATA repeated in tandem 2-6 times. Four haplotypes were observed: one specific of *P. canariensis*, one restricted to some individuals of *P. sylvestris*, and two shared between *P. dactylifera*, *P. atlantica* and *P. sylvestris*. Pintaud et al. (2010) studied this locus in 12

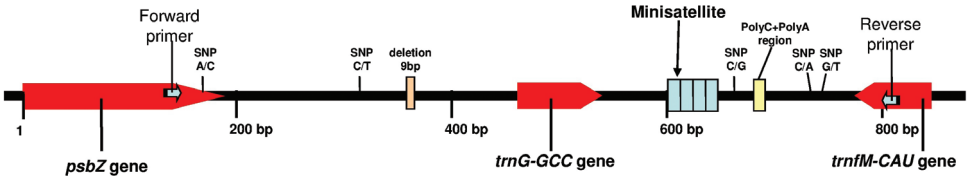

Figure 1. The sequenced cpDNA *psbZ-trnfM* region. The location of PCR primers used and polymorphisms found in this study are shown. DNA fragment length refers to the *P. dactylifera cv. Khalas* cpDNA sequence (Yang et al. 2010), characterised by a 4-repetitions minisatellite haplotype (NCBI Reference Sequence: NC_013991.2).

Phoenix species, identifying five haplotypes, whose pattern of variation was strongly associated with species. The maximum number of haplotypes per species was three (*P. roebelenii*). Yet, most of the haplotypes were shared between species, viz. the 3-repetitions haplotype was the most common haplotype within the genus, and was shared by eight out of the 12 species. *Phoenix canariensis* was the only taxon characterised by the 5-repetitions haplotype. Hence, despite the promising information obtained, the minisatellite alone did not allow to distinguish all *Phoenix* species.

Given the potential of the *trnG*(GCC)*-trnfM*(CAU) spacer for barcoding in *Phoenix*, we examined a wider cpDNA region, viz. a ~700 bp sequence *psbZ-trnfM*(CAU) (Figure 1), in search of an efficient DNA barcode locus for species delimitation and identification of female parents in hybrids in the genus *Phoenix*.

Methods

Taxon sampling

One hundred and thirty-six individuals, belonging to 13 *Phoenix* species, with emphasis on *P. dactylifera,* were analysed in this work (Appendix). *Phoenix andamanensis* was not included in the analysis due to a lack of material.

DNA sequencing

For each sample, genomic DNA was extracted from 40 mg of freeze-dried leaf tissue which was first grinded using a bead-mill homogenizer Tissuelyser (Qiagen, France). Extraction was performed using the DNeasy Plant Mini Kit protocol along with the QIAcube robotic workstation for DNA automated purification (Qiagen, France). Extracted DNA was quantified by means of a Nanodrop ND1000 spectrophotometer (Thermo Fisher Scientific Inc., USA) and visualized on 1% agarose gels stained with ethidium bromide.

The PCR amplification was carried out using the monocotyledoneous universal primers *psbZ*-IGS-F: GGTACMTCATTATGGATTGG, and *trnfM*-IGS-R: GCG-

GAGTAGAGCAGTTTGGT (Scarcelli et al. 2011). The amplified cpDNA fragment was approximately 700 bp long. PCR reactions were prepared in 25 µl of total volume, containing the following reagent concentrations: 5 ng/µl DNA template, 0.2 µM each of forward and reverse primers, 2X Failsafe PCR PreMix E (Epicentre Biotechnologies, Madison USA), 2.5 U/µl Failsafe Enzyme Mix (Epicentre Biotechnologies, USA), and DNase-free sterile water. PCR parameters were the following: an initial denaturation step at 94 °C for 3 min, then 35 cycles at 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 1 min, and a final elongation step at 72 °C for 10 min. PCR products were controlled on 1% agarose gels stained with ethidium bromide and then purified using Ampure Agencourt kit (Agencourt Bioscience Corporation, USA). Their quantification was done by means of a Nanodrop ND1000 spectrophotometer (Thermo Fisher Scientific Inc., USA). Cycle sequencing was carried out using the Big Dye Terminator v3.1 kit (Applied Biosystems, USA). Cycle sequencing products were purified using the CleanSeq Agencourt Kit (Agencourt Bioscience Corporation, USA) and were then analysed on an ABI 3130 automated DNA Sequencer (Applied BioSystems, USA).

Sequence alignment and identification success

The chromatograms obtained with the forward and reverse primers were combined and edited with SeqMan II 5.00 software (DNASTAR Inc., USA), to generate consensus sequences, which were aligned in BioEdit (Hall 1999), using the Clustal W algorithm. The obtained alignment was further improved manually with MESQUITE v2.75 (Maddison and Maddison 2007). The observed polymorphisms were positioned in reference to the complete chloroplast genome sequence of *P. dactylifera* cv. Khalas, available in GenBank (accession NC_013991.2).

To assess the potential of the *psbZ-trnfM* region as a barcode for accurate species identification, we evaluated the proportion of correct identifications using TaxonDNA (Meier et al. 2006). The Best Match and Best Close Match tests were run for species with > 1 individual and with nearly complete sequences, which resulted in a reduced dataset of 11 species (excluding *P. acaulis* and *P. atlantica*) and 121 individuals. Because of this constraint, the two species represented by only one individual were analysed by direct comparison of their sequences. Moreover, direct sequence comparison included not only nucleotide substitutions as in the TaxonDNA analysis, but also indels, minisatellites and homopolymers.

Results

The amplification of the plastid target region *psbZ-trnfM*(CAU) was successful for all samples, and the sequencing with both primers was achieved for 123 individuals, while a single read (forward or reverse) was retrieved for the other 13 individuals, whose sequences were approximately 20% shorter.

Table 1. Distribution of observed polymorphisms in the region *psbZ-trnfM*(CAU)

Substitutions[a]					9 bp deletion[a]	Minisatellite[c]	Homo-polymer[a]	Species[b]
36607	36754	37099	37183	37190	36795–36803	37050–37098	37128–37139	
Haplotypes recorded in a single species[d] (80.1% total sampling)								
C	T	G	A	T	absent	**5M1+1M2**(4)	7 C + 5 A	*P. canariensis* (7)
C	T	**C**	A	T	absent	**2M1+5bp+1M2**(6)	**6 C + 5 A**	*P. reclinata* (4)
C	T	G	A	T	absent	**1M1+2M2**(7)	7 C + 5 A	*P. reclinata* (6)
C	T	G	**C**	T	absent	6M1+1M2(5)	7 C + 5 A	*P. caespitosa* (2)
C	C	G	A	**G**	absent	2M1+1M2(1)	7 C + 5 A	*P. loureiroi* (1)
A	C	G	A	T	absent	3M1+1M2(2)	6 C + 6 A	*P. loureiroi* (1)
A	C	G	A	T	absent	2M1+1M2(1)	7 C + 5 A	*P. acaulis* (1)
C	C	G	A	T	absent	2M1+1M2(1)	6 C + 6 A	*P. pusilla* (2)
C	T	G	A	T	present	2M1+1M2(1)	6 C + 6 A	*P. paludosa* (2)
C	T	G	A	T	present	4M1+1M2(3)	**5 C + 7 A**	*P. roebelenii* (3)
C	T	G	A	T	present	3M1+1M2(2)	**5 C + 7 A**	*P. roebelenii* (1)
C	T	G	A	T	absent	4M1+1M2(3)	7 C + 5 A	*P. dactylifera* (78)
C	T	G	A	T	absent	2M1+1M2(1)	**8 C + 5 A**	*P. sylvestris* (1)
Haplotypes shared by two species (5.1%)								
C	T	G	A	T	absent	6M1+1M2(5)	7 C + 5 A	*P. rupicola* (3)
								P. theophrasti (4)
Haplotypes shared by three species (14.8%)								
C	T	G	A	T	absent	3M1+1M2(2)	7 C + 5 A	*P. atlantica* (1)
								P. dactylifera (16)
								P. sylvestris (3)

[a] Position in the complete chloroplast genome of *Phoenix dactylifera* 'Khalas' accession NC_013991.2.

[b] Number of individuals analysed for each species in parentheses (total sampling of 136 specimens).

[c] Number of repetitions of the 12 bp minisatellite units, including number of units of motif 1 (M1) and motif 2 (M2) as represented in Figure 2.

[d] Species-specific mutations in bold.

(1–7) Minisatellites haplotypes as reported in Figure 2 (1 to 7).

The analysis of the intra- and interspecific variation within the sequenced region by direct observation of the sequence alignment showed four mutation types that contributed to the separation of *Phoenix* species: single nucleotide polymorphisms (SNPs), indels, length variation at the 12 bp minisatellite locus, and in homopolymers, allowing in total to identify unambiguously eight out of the 13 species (Table 1).

The minisatellite located in the *trnG*(GCC)-*trnfM*(CAU) intergenic spacer showed seven haplotypes. Most haplotypes corresponded to a Variable Number Tandem Repeat (VNTR) stepwise mutational pattern of 12 bp units. These units corresponded to two motifs: CTAACTACTATA (motif 1) and GTAGTTAGTATA (motif 2), which form between themselves a pattern of 12 bp inverted repeats shifted with respect to the boundaries of the mutational units (Figure 2). One haplotype, found in four out of ten *P. reclinata* individuals, departed from this pattern, with two complete units of motif 1 plus an incomplete third unit with a 7 bp-deletion (CTAACTA) (haplotype 6; Figure 2). These four specimens were further characterized by a SNP (C instead of G)

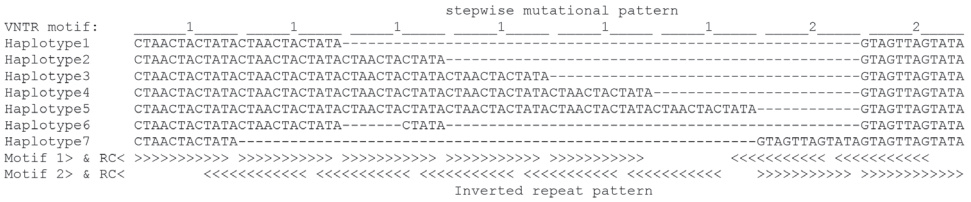

Figure 2. Structure and variation of the minisatellite in the *trnG-trnfM* intergenic spacer. The repeats of the two mutational motifs (1 and 2) are indicated above the sequence alignment of the 7 haplotypes recorded. The pattern of inverted repeats generated by the two motifs and their reverse complements (RC) is shown below the alignment. See Table 1 for haplotype distribution among species.

at position 37099. The other six *P. reclinata* samples were also unique in having two repeats of motif 2 instead of one as in the six other haplotypes (haplotype 7; Figure 2). *Phoenix canariensis* was characterised by a private haplotype with five repeats of motif 1 (haplotype 4; Figure 2). The maximum number of haplotypes per species was three; in that case, one or two of them were shared by different species (Table1).

Additional deletions and SNPs were detected in some of the analysed species, both upstream and downstream of the minisatellite (Table 1). *Phoenix roebelenii* and *P. paludosa* shared a 9 bp-deletion (GTACTTTAC, upstream to the minisatellite, in position 36795–36803), *P. acaulis, P. loureiroi*, and *P. pusilla* shared a SNP at position 36754 (C instead of T), while *P. caespitosa* had a SNP at position 37183 (C instead of A). One of the three samples of *P. loureiroi* showed a SNP in position 37190 (G instead of T), and another sample shared a SNP with the single specimen of *P. acaulis* (A instead of C in position 36607). Some more differences among species and/or individuals were found in an homopolymer region (C_n + A_n), located at position 37128–37139, downstream to the minisatellite (Table 1).

Phoenix theophrasti and *P. rupicola* shared the 6 repetitions of motif 1 minisatellite haplotype (haplotype 5; Figure 2) and could not be distinguished from each other (*P. caespitosa* had also the 6 repetitions haplotype but was further differentiated by a species-specific SNP). Within the "date palm-complex" (Pintaud et al. 2013), a 3 repetitions of motif 1 haplotype was shared by some specimens of *P. atlantica*, *P. dactylifera,* and *P. sylvestris* (haplotype 2; Figure 2), while a majority of individuals of *P. dactylifera* were uniquely identified by a 4 repetitions of motif 1 haplotype (haplotype 3; Figure 2).

The TaxonDNA pairwise comparison analysis of the 121 samples retained resulted in 115 sequences with a closest match at 0%. There were 18 allospecific matches at 0% (15.65%). At the individual level, the Best Match test, and the Best Close Match test with a threshold of 3%, resulted in 82.64% correct identifications, 14.87% ambiguous identifications and 2.47% incorrect identifications. At the species level, however, only *P. caespitosa* could be unambiguously identified, since it was the only species in the sampling with an autapomorphic SNP.

The haplotype sequences used, and the new ones obtained during this study, are deposited in GenBank under accessions: JF745571, EU043486, EU043484, EU043485, JX970915–970936.

Discussion

In this study, we tested the usefulness of the *psbZ-trnfM*(CAU) region as a barcode locus in *Phoenix*. The successful amplification and sequencing of this marker within all of the analysed species confirms its value in terms of universality. Moreover, its high performance should allow the acquisition of barcode information even with partially degraded DNA samples.

TaxonDNA unambiguously identified a single species, *P. caespitosa*, due to the scarcity of SNPs, most of them shared by two or more species, or on the contrary restricted to a subset of individuals within species. Therefore, it is important to take into account the other polymorphisms (indels, minisatellites and homopolymers) which usually represent half or more of the mutations in non-coding chloroplast DNA (Scarcelli et al. 2011). However, at the individual level, the Best Match and Best Close Match tests resulted in more than 80% correct identifications, which is indicative of the barcoding potential of the marker studied.

The 9 bp-deletion, shared by *P. roebelenii* and *P. paludosa*, supports Barrow's conclusions (1998), as well as Pintaud et al.'s (2013), regarding the close relationship between these two taxa.

Regarding the 12 bp minisatellite, our results revealed much more complexity than previously reported (Pintaud et al. 2010). This could be explained by the increased sampling of the present study, and also by differences in methodology, i.e. sequencing versus genotyping. In particular, the genotyping data of Pintaud et al. (2010) did not detect the 7 bp-deletion found within the minisatellite of some *P. reclinata* samples, and were also misled by the size homoplasy between haplotype 1 and 7 (Figure 2). We therefore recommend that sequence data should be obtained before performing any study based on genotyping, in order to have a solid basis to interpret genotyping data.

In total, considering all mutation types, our results allowed us to efficiently identify eight out of 13 species. This indicates that the locus *psbZ-trnfM*(CAU) has some potential to yield DNA barcodes that can be used for species identification within the genus *Phoenix*. This locus could also be useful to identify the female parent in many interspecific crosses, such as *P. dactylifera* × *P. canariensis*. Hybrids involving *P. canariensis* as female parents are particularly easy to track because this species is monomorphic with a private haplotype at the locus studied. Hybrids between these two species are a concern for the genetic integrity of native populations of *P. canariensis* in the Canary Islands (González-Pérez et al. 2004). Such hybrids are also very common in ornamental plantings, for which they represent a valuable horticultural resource.

Nevertheless, in order to increase resolution, other DNA regions should be examined, in search of characters allowing the identification of all taxa. Given their proven utility in palms, the *psbA-trnH* locus (Al-Qurainy et al. 2011) and/or the ribosomal ITS2 (Jeanson et al. 2011) could be investigated in combination with *psbZ-trnfM* for this purpose. Special attention should be paid to the species group sharing haplotype 2 (Figure 2): *P. atlantica*, *P. dactylifera* and *P. sylvestris*. This group is composed of very closely related species, so difficulty in DNA barcoding for these species is expected.

On the other hand, in some cases, the morphological divergence is not associated to sequence divergence in the *psbZ-trnfM* region. For example, *P. rupicola* and *P. theophrasti* share the same haplotype despite considerable morphological differentiation and geographical isolation, the former being restricted to the E Himalayan, while the latter is an Aegean endemic. These two species possibly share plesiomorphic SNP states and may show convergence in the minisatellite haplotype. In contrast, *P. dactylifera* and *P. theophrasti* are phenotypically very similar, but can easily be distinguished at the *psbZ-trnfM*(CAU) region. The relation between morphological divergence and molecular divergence at the *psbZ-trnfM*(CAU) region among the *Phoenix* species needs to be addressed with a larger sampling within species as recommended by the China Plant BOL Group (2011).

Acknowledgements

We wish to thank the following persons and institutions for their valuable help during the various phases of our study: Rita Bregliano†, Paolo Curir, Laura De Benedetti, Federica Nicoletti (CRA-FSO, Sanremo, Italy); Frédérique Aberlenc-Bertossi, Natalie Chabrillange, Nora Scarcelli (IRD Montpellier, France); Muriel Latreille, Sylvain Santoni (AMM at INRA SupAgro, Montpellier, France); Patrizia Martini (IRF, Sanremo, Italy); Mauro Roggero (Cooperativa "Il Cammino", Sanremo, Italy); Giancarlo Pignatta ("U Risveiu Burdigotu" Association, Bordighera, Italy); the Natta family (Bordighera, Italy); Robert Castellana (CRP, Nice, France), Salwa Zehdi (University of Tunis, Tunisia). We also thank the three anonymous reviewers and the scientific editor for their useful comments and suggestions.

References

Al-Qurainy F, Khan S, Al-Hemaid FM, Ajmal Ali A, Tarroum M, Ashraf M (2011) Assessing molecular signature for some potential date (*Phoenix dactylifera* L.) Cultivars from Saudi Arabia, Based on Chloroplast DNA Sequences *rpoB* and *psbA-trnH.* International Journal of Molecular Sciences 12: 6871–6880. doi: 10.3390/ijms12106871

Asmussen CB, Dransfield J, Deickmann V, Barfod AS, Pintaud J-C, Baker W (2006) A new subfamily classification of the palm family (Arecaceae): evidence form plastid DNA phylogeny. Botanical Journal of the Linnean Society 151: 15–38. http://pure.au.dk/portal/files/43885556/Asmussen_et_al_2006.pdf, doi: 10.1111/j.1095-8339.2006.00521.x

Baker WJ, Savolainen V, Asmussen-Lange CB, Chase M, Dransfield J, Forest F, Harley M, Uhl N, Wilkinson M (2009) Complete Generic-level Phylogenetic Analyses of Palms (Arecaceae) with Comparisons of Supertree and Supermatrix Approaches. Systematic Biology 58: 240–256. doi: 10.1093/sysbio/syp021

Barrow S (1998) A monograph of *Phoenix* L. (Palmae: Coryphoideae). Kew Bulletin 53: 513–575. doi: 10.2307/4110478

Boydak M, Barrow S (1995) A new locality for *Phoenix* in Turkey: Gölköy-Bödrum. Principes 39: 117–122.

CBOL Plant Working Group (2009) A DNA barcode for land plants. Proceedings of the National Academy of Sciences of the USA 106: 12794–12797. doi: 10.1073/pnas.0905845106

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

China Plant BOL Group (2011) Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proceedings of the National Academy of Sciences of the USA 108: 19641–19646. doi: 10.1073/pnas.1104551108

DeSalle R, Egan M, Siddall M (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding. Philosophical Transactions of the Royal Society B 360: 1905–1916. doi: 10.1098/rstb.2005.1722

Dransfield J, Uhl NW, Asmussen CB, Baker WJ, Harley MM, Lewis CE (2008) Genera palmarum, the evolution and classification of palms. Royal Botanic Gardens, Kew, U.K.

González-Pérez MA, Caujapé-Castells J, Sosa PA (2004) Molecular evidence of hybridisation between the endemic *Phoenix canariensis* and the widespread *P. dactylifera* with Random Amplified Polymorphic DNA (RAPD) markers. Plant Systematics and Evolution 247: 165–175. doi: 10.1007/s00606-004-0166-7

Govaerts R, Dransfield J (2005) World checklist of palms. Royal Botanic Gardens, Kew, U.K.

Greuter W (1967) Beiträge zur Flora der Südägäis 8-9. Bauhinia 3: 243–254.

Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends in Genetics 23: 167–172. doi: 10.1016/j.tig.2007.02.001

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95–98.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proceedings of the Royal Society of London B 270: 313–321. doi: 10.1098/rspb.2002.2218

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences of the USA 101: 14812–14817. doi: 10.1073/pnas.0406166101

Hebert PDN, Gregory TR (2005) The promise of DNA barcoding for taxonomy. Systematic Biology 54: 852–859. doi: 10.1080/10635150500354886

Henderson SA, Billotte N, Pintaud J-C (2006) Genetic isolation of Cape Verde Island *Phoenix atlantica* (Arecaceae) revealed by microsatellite markers. Conservation Genetics 7: 213–223. doi: 10.1007/s10592-006-9128-7

Jeanson ML, Labat J-N, Little DP (2011) DNA barcoding: a new tool for palm taxonomists? Annals of Botany 108: 1445–1451. doi: 10.1093/aob/mcr158

Maddison W, Maddison D (2007) Mesquite 2. http://mesquiteproject.org

Meier R, Kwong S, Vaidya G, Ng PK (2006) DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Systematic Biology 55: 715–728. doi: 10.1080/10635150600969864

Nesbitt M (1993) Archaeobotanical evidence for early Dilmun diet at Saar, Bahrain. Arabian Archaeology Epigraphy 4: 20–47. doi: 10.1111/j.1600-0471.1993.tb00041.x

Newton C, Gros-Balthazard M, Ivorra S, Paradis L, Pintaud J-C, Terral JF (2013) *Phoenix dactylifera* and *P. sylvestris* in Northwestern India: a glimpse of their complex relationships. Palms 57: 37–50.

Pintaud J-C, Zehdi S, Couvreur T, Barrow S, Henderson S, Aberlenc-Bertossi F, Tregear J, Billotte N (2010) Species delimitation in the genus *Phoenix* (Arecaceae) based on SSR markers, with emphasis on the identity of the Date palm (*Phoenix dactylifera*). In: Seberg O, Petersen G, Barfod A, Davis J (Eds) Taxonomy of *Phoenix*. Diversity, phylogeny, and evolution in the Monocotyledons. Aarhus University Press, Denmark, 267–286.

Pintaud J-C, Ludeña B, Aberlenc-Bertossi F, Zehdi S, Gros-Balthazard M, Ivorra S, Terral J-F, Newton C, Tengberg M, Abdoulkader S, Daher A, Nabil M, Saro Hernández I, González-Pérez MA, Sosa P, Santoni S, Moussouni S, Si-Dehbi F, Bouguedoura N (2013) Biogeography of the Date Palm (*Phoenix dactylifera* L., *Arecaceae*): Insights on the Origin and on the Structure of Modern Diversity. In: Bouguedoura N, Bennaceur M, Pintaud J-C (Eds) Proceedings of the I International Symposium on Date Palm. Acta Horticulturae 994: 19–38. http://www.actahort.org/books/994/994_1.htm

Savolainen V, Cowan RS, Vogler AP, Roderick GK, Lane R (2005) Towards writing the encyclopaedia of life: an introduction to DNA barcoding. Philosophical Transactions of the Royal Society B 360: 1805–1811. doi: 10.1098/rstb.2005.1730

Scarcelli N, Barnaud A, Eiserhardt W, Treier UA, Seveno M, d'Anfray A, Vigouroux Y, Pintaud J-C (2011) A Set of 100 chloroplast primer pairs to study population genetics and phylogeny in Monocotyledons. PLoS ONE 6: e19954. doi: 10.1371/journal.pone.0019954

Tengberg M (2012) Beginnings and early history of date palm garden cultivation in the Middle East. Journal of Arid Environments 86: 139–147. doi: 10.1016/j.jaridenv.2011.11.022

Tournay F (2009) The Nabonnand family and palms. Palms 53: 119–123.

Wrigley G (1995) Date palm (*Phoenix dactylifera* L). In: Smartt J, Simmonds NW (Eds) The Evolution of Crop Plants. Longman, Essex, 399–403.

Yang M, Zhang X, Liu G, Yin Y, Chen K, Yun Q, Zhao D, Al-Mssallem I, Yu J (2010) The complete chloroplast genome sequence of Date Palm (*Phoenix dactylifera* L.). PLoS ONE 5: e12762. doi: 10.1371/journal.pone.0012762

Yao H, Song J, Liu C, Luo K, Han J, Li Y, pang X, Xu H, Zhu Y, Xiao P, Chen S (2010) Use of ITS2 region as the universal DNA barcode for plants and animals. PLoS ONE 5: e13102. doi: 10.1371/journal.pone.0013102

Zohary D, Hopf M (2000) Domestication of plants in the Old World. The origin and spread of cultivated plants in the West Asia, Europe and the Nile Valley. 3rd Ed. Oxford University Press, New York.

Appendix

List of samples. DNA bank reference: IRD = Institut de Recherche pour le Développement, 911 Av. Agropolis, F-34394 Montpellier Cedex 5, France. Tissue bank reference: CRA-FSO = Consiglio per la Ricerca e la sperimentazione in Agricoltura - Unità di Ricerca per la Floricoltura e le Specie Ornamentali, Corso degli Inglesi 508, I-18038 Sanremo (IM), Italy.

Phoenix acaulis* Roxb**: MWC5559, Kew ,UK (IRD). ***Phoenix atlantica* A. Chev.**: SH25, Cape Verde (IRD). ***Phoenix caespitosa* Chiov.**: MWC1195, MWC1802, Kew, UK (IRD). ***Phoenix canariensis* Chabaud**: 93.100, 93.101, 93.103, 93.107, Sanremo, Italy (IRD, CRA-FSO); MWC1396, Kew, UK (IRD); JCP169, JCP210, Canary Isl., Spain (IRD). ***Phoenix dactylifera* L.**: 93.003, 93.004, 93.005, 93.025, 93.027, 93.030, 93.037, 93.043, 93.045, 93.047, 93.048, 93.049, 93.052, 93.054, 93.055, 93.056, 93.059, 93.060, 93.061, 93.065, 93.066, 93.067, 93.070, 93.071, 93.072, 93.073, 93.076, 93.077, 93.080, 93.085, 90.002, 90.003, 90.004, 90.005, 90.006, 90.007, 90.008, 90.009, 90.010, 90.011, 90.012, 90.013, 90.014, 90.015, 90.025, 90.026, 90.027, 90.028, 90.029, 91.005, Sanremo, Italy (IRD, CRA-FSO); 00.01, 00.02, 00.03, 00.04, 00.05, 00.06, 00.07, 00.08, 00.09, 00.10, 00.11, 00.13, 00.14, 00.83, 00.85, 00.88, 46.02, 46.04, 46.05, 46.06, 46.08, 46.09, 46.14, 46.15, 46.16, 46.17, 46.18, 46.19, 46.20, 46.21, 46.23, JCP413, JCP414, JCP415, JCP416, JCP417, Bordighera, Italy (IRD, CRA-FSO); DAT077-365, DAT079-366, Oman (IRD); JCP260, Murcia, Spain; JCP426, Elche, Spain; SZ1, SZ2, SZ5, SZ10, Tunisia (IRD). ***Phoenix loureiroi* Kunth *var. loureiroi: JCP409, Montgomery Botanical Garden, Miami, USA (IRD); MWC1187, Kew, UK (IRD). ***Phoenix paludosa* Roxb.**: MWC1190, MWC1877, Kew, UK (IRD). ***Phoenix pusilla* Gaertn.**: JCP213_5, Sri Lanka (IRD); MWC1806, Kew, UK (IRD). ***Phoenix reclinata* Jacq.**: ECH3-A, ECH4-A, ECH5-B, 91.001, 91.007, 91.008, 91.009, 91.033, 92.003, Sanremo, Italy (IRD, CRA-FSO); MWC1397, Kew, UK (IRD). ***Phoenix roebelenii* O'Brien**: ECH1-A, ECH2-A, Sanremo, Italy (IRD, CRA-FSO); MWC1400, MWC1805, Kew, UK (IRD). ***Phoenix rupicola* T. Anderson**: ECH6-A, ECH8-A, Sanremo, Italy (IRD, CRA-FSO); MWC1399, Kew, UK (IRD). ***Phoenix sylvestris* (L.) Roxb.**: DAT057-345, Elche, Spain (IRD); JCP214, The Palm Center, UK, (IRD); JCP405-388, Thuret, France (IRD); MWC1876, Kew, UK (IRD). ***Phoenix theophrasti* Greuter**: ECH7-A, ECH9-A, Sanremo, Italy (IRD, CRA-FSO); JCP215, The Palm Center, UK (IRD); MWC1163, Kew, UK.