HAL
open science

In silico mining of microsatellites in coding sequences of the date palm (arecaceae) genome, characterization and transferability

Frédérique Aberlenc-Bertossi, Karina Castillo, Christine Tranchant-Dubreuil, Emira Cherif, Marco Ballardini, Sabira Abdoulkder, Muriel Gros-Balthazard, Nathalie Chabrillange, Sylvain Santoni, Antonio Mercuri, et al.

To cite this version:

Frédérique Aberlenc-Bertossi, Karina Castillo, Christine Tranchant-Dubreuil, Emira Cherif, Marco Ballardini, et al.. In silico mining of microsatellites in coding sequences of the date palm (arecaceae) genome, characterization and transferability. Applications in Plant Sciences, 2014, 2 (1), 10.3732/apps.1300058 . hal-01268884

HAL Id: hal-01268884

https://hal.science/hal-01268884v1

Submitted on 27 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

BioOne RESEARCH EVOLVED

In Silico Mining of Microsatellites in Coding Sequences of the Date Palm (Arecaceae) Genome, Characterization, and Transferability

Author(s): Frédérique Aberlenc-Bertossi, Karina Castillo, Christine Tranchant-Dubreuil, Emira Chérif , Marco Ballardini, Sabira Abdoulkader, Muriel Gros-Balthazard, Nathalie Chabrillange, Sylvain Santoni, Antonio Mercuri, and Jean-Christophe Pintaud

Source: Applications in Plant Sciences, 2(1)
Published By: Botanical Society of America
DOI: http://dx.doi.org/10.3732/apps.1300058
URL: http://www.bioone.org/doi/full/10.3732/apps.1300058

BioOne (www.bioone.org) is a nonprofit, online aggregation of core research in the biological, ecological, and environmental sciences. BioOne provides a sustainable online platform for over 170 journals and books published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Web site, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/page/terms_of_use.

Usage of BioOne content is strictly limited to personal, educational, and non-commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Applications in Plant Sciences

Applications in Plant Sciences 2014 2(1): 1300058

Primer Note

In silico mining of microsatellites in coding sequences of the date palm (Arecaceae) genome, characterization, and transferability[1]

Frédérique Aberlenc-Bertossi[2,8], Karina Castillo[2], Christine Tranchant-Dubreuil[2], Emira Chérif[2,3], Marco Ballardini[4], Sabira Abdoulkader[2,5], Muriel Gros-Balthazard[2,6], Nathalie Chabrillange[2], Sylvain Santoni[7], Antonio Mercuri[4], and Jean-Christophe Pintaud[2]

[2]IRD, UMR DIADE—BDP, DYNADIV, and EVODYN teams, 911 Av. Agropolis, BP 64501, 34394 Montpellier, Cedex 5, France; [3]Laboratoire de génétique moléculaire, immunologie et biotechnologie, Faculté des Sciences de Tunis, Campus Universitaire, 2092 El Manar, Tunis, Tunisia; [4]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; [5]ISV/CERD, route de l'Aéroport, BP 468, Djibouti; [6]Centre de Bio-Archéologie et d'Ecologie (UMR 5059 CNRS/Université Montpellier 2/EPHE/INRAP), Institut de Botanique, 163 Rue Auguste Broussonet, 34090 Montpellier, France; and [7]INRA, UMR AGAP, 2 Place Viala, 34060 Montpellier, Cedex 1, France

- *Premise of the study:* To complement existing sets of primarily dinucleotide microsatellite loci from noncoding sequences of date palm, we developed primers for tri- and hexanucleotide microsatellite loci identified within genes. Due to their conserved genomic locations, the primers should be useful in other palm taxa, and their utility was tested in seven other *Phoenix* species and in *Chamaerops*, *Livistona*, and *Hyphaene*.
- *Methods and Results:* Tandem repeat motifs of 3–6 bp were searched using a simple sequence repeat (SSR)–pipeline package in coding portions of the date palm draft genome sequence. Fifteen loci produced highly consistent amplification, intraspecific polymorphisms, and stepwise mutation patterns.
- *Conclusions:* These microsatellite loci showed sufficient levels of variability and transferability to make them useful for population genetic, selection signature, and interspecific gene flow studies in *Phoenix* and other Coryphoideae genera.

Key words: Arecaceae; Coryphoideae; microsatellite/SSR mining; *Phoenix dactylifera*; transferability.

The date palm (*Phoenix dactylifera* L.) is a monocotyledon species belonging to the Arecaceae family, and is widely cultivated in North Africa, the Sahel (from the Atlantic to the Red Sea), the Middle East, and eastward to the Indus Valley. The date palm is well adapted to cultivation in arid and semiarid areas, and it has been introduced in warm and dry regions worldwide. Mainly grown for its fruits, the date palm represents an important ecological and socioeconomic resource.

Despite the increasing number of studies on date palm, there are still not enough molecular markers available for a number of applications. Most published microsatellite or simple sequence repeat (SSR) markers are dinucleotide loci from unknown noncoding regions of the genome, generally isolated from microsatellite-enriched DNA libraries (Billotte et al., 2004; Arabnezhad et al., 2012). The increasing amount of available genome sequence data offers new prospects for microsatellite marker development through in silico mining, a promising approach for date palm (Cherif et al., 2013), based on the recently published date palm genome sequence (Al-Dous et al., 2011) and expressed sequence tags (ESTs) (Zhao et al., 2012). Our aim was to develop new markers from coding sequences to ensure clear stepwise mutation patterns usable for genetic diversity, dating, and selection signature analyses, and also to facilitate transferability to other species.

METHODS AND RESULTS

In silico microsatellite mining and primer design were performed on the date palm genome draft sequence version 2 (Al-Dous et al., 2011), with the Perl script SSR_pipeline-v2.pl (Poncet et al., 2006), which incorporates three free software programs: Tandem Repeats Finder (Benson, 1999), Primer3 (Rozen and Skaletsky, 2000), and BLAST (Altschul et al., 1990). The multi-FASTA file of all 19,414 predicted genes (full and partial; PDK20.mRNA.fsa) and the multi-FASTA file with all scaffold sequences (PDK20.fsa) from version 2 of the date palm genome research program at Weill Cornell Medical College in Qatar were downloaded from http://qatar-weill.cornell.edu/research/datepalmGenome/download.html. The search identified 204 genes containing coding sequences with microsatellites, 150 of which were suitable for primer design, but only 103 had nonduplicated primer annealing sites. Among them, we retained loci having perfect trinucleotide motifs with six (excluding those without annotation) or more (with or without annotations) repeats, and hexanucleotide motifs with at least four repeats (with or without annotation).

Of the 47 primer pairs finally retained, 33 generated expected PCR amplification patterns in a preliminary test with eight *P. dactylifera* individuals (Table 1). The 33 loci were further tested on 16 individuals representing *P. dactylifera* (7),

[1]Manuscript received 8 July 2013; revision accepted 24 September 2013.

This work was partially funded by the Agence Universitaire de la Francophonie (AUF)–Projets méditerranéens de recherche scientifique inter-universitaire (MeRSI) project 6313PS001 "Ressources génétiques et moléculaires du palmier dattier" (2010-2013).

[8]Author for correspondence: frederique.aberlenc@ird.fr

doi:10.3732/apps.1300058

 © 2014 Aberlenc-Bertossi et al. Published by the Botanical Society of America. This work is licensed under a Creative Commons Attribution License (CC-BY-NC-SA).

TABLE 1. Characteristics of 33 microsatellite markers developed for *Phoenix* species. The putative annotation was done using the BLASTX program and the UniProtKB/Swiss-Prot protein database with an *E*-value cutoff of 10^{-5}.

Locus	Primer sequences (5′–3′)[a]	Repeat motif	Size range (bp)[b]	Scaffold ID	Start	Stop	Gene annotation	*E*-value	Organism
mPdIRD01	F: CTCGGAAGGGTATGGACAAA R: TTGCCTTCGACGTGGTAGTA	$(AAG)_3$	200	PDK_20s1306691	24393	24401	Putative pectinesterase/pectinesterase inhibitor 28	4.00E-87	*Arabidopsis thaliana*
mPdIRD03	F: CATTGATCCAACACCACCAC R: GCCAAAACCAGCTCTGGTAAC	$(CCT)_6$	192–198	PDK_20s1315791	3431	3448	Cysteine-rich receptor-like protein kinase 2	1.00E-166	*Arabidopsis thaliana*
mPdIRD04	F: TTGGTGGCCTTTCTCAGAGT R: TGGGATCAAAGTAGGGTTGG	$(AGC)_6$	255–261	PDK_20s13282911	9405	9422	*S*-adenosylmethionine synthase 10	6.00E-77	*Oryza sativa*
mPdIRD05	F: CTATCAGGATGGGGGTGATG R: ACCCATCTGCATAGCTCCAG	$(GAT)_6$	301–302	PDK_20s1366071	11666	11683	DEAD-box ATP-dependent RNA helicase ISE2, chloroplastic	3.00E-09	*Arabidopsis thaliana*
mPdIRD07	F: TGCAATACGATGGCAGAGTC R: CCTTGCAAGTTTTCCACACC	$(TGG)_6$	182–212	PDK_20s1387131	3737	3754	No hit	–	–
mPdIRD08	F: CTATTGGGTCCCTTGGTGAG R: TGACTGCTCGTCATCAGGTC	$(GAT)_6$	202	PDK_20s1402051	10945	10962	No hit	–	–
mPdIRD10	F: ATGCGTTCATCTCCCTTGAG R: GCTGCAAACATCATCCTCAC	$(CAG)_6$	194–214	PDK_20s1405881	31976	31993	No hit	–	–
mPdIRD11	F: GAGTTGGAGGCAAAACCAGA R: CCACAAAACCCTTGTCTTCC	$(GAT)_6$	309–317	PDK_20s1422271	4385	4402	Two-component response regulator-like APRR9	5.00E-18	*Arabidopsis thaliana*
mPdIRD13	F: GCGGAGACAGGAGATGGTAA R: CTTGACTGCTTCTGCTGCTG	$(CAC)_6$	198–227	PDK_20s1496731	12538	12555	Trihelix transcription factor GT-2	8.00E-62	*Arabidopsis thaliana*
mPdIRD14	F: GAGGGGTTCACGTTTGTGTC R: GCACCAAGCACAAGAGCAAT	$(GCG)_6$	163	PDK_20s1505351	9121	9138	Probable ascorbate-specific transmembrane electron transporter 1	1.00E-82	*Oryza sativa*
mPdIRD15	F: CCGAGTCTGGCGAAGTAAAC R: CTCCCCTTCCTCATCCTCTC	$(GAA)_6$	406–408	PDK_20s1507261	2378	2395	Eukaryotic translation initiation factor 2 subunit beta	1.00E-22	Wheat
mPdIRD16	F: CTGTCCGATCGAATTCTGC R: GGACATCTCTTTGCGGTCAT	$(CAG)_6$	197–214	PDK_20s1521921	7038	7055	Probable WRKY transcription factor 41	3.00E-47	*Arabidopsis thaliana*
mPdIRD17	F: GTGGGAGAAACCCGAAGAAT R: CTGCTGCCTCATCTGCATT	$(AGC)_6$	199–202	PDK_20s1549911	54838	54855	Flowering time control protein FCA	3.00E-38	*Arabidopsis thaliana*
mPdIRD20	F: TTGAATGGTCCCCTGTAGGT R: GTCCCAGCATGATTGCAGTA	$(AGT)_6$	341–373	PDK_20s1640771	6702	6719	Transcription factor bHLH62	7.00E-57	*Arabidopsis thaliana*
mPdIRD22	F: GGCTGTATGGGAAAGACCTG R: CCTGCTGCATATTCTTCGTG	$(GAA)_6$	231–271	PDK_20s1726541	2878	2895	Probable peptide/nitrate transporter At1g59740	4.00E-40	*Arabidopsis thaliana*
mPdIRD24	F: GCTCCTGCAGAACCTGAAAC R: GGACATCACCGTCCAATTCT	$(AAG)_6$	184	PDK_20s1762671	5194	5211	Probable nucleolar protein 5-1	2.00E-46	*Arabidopsis thaliana*
mPdIRD25	F: CACTGGAAATTCAGGGCCTA R: CCCAATTTCTCAGCCAAGAC	$(AGG)_6$	193–205	PDK_20s1831761	4692	4709	Heat stress transcription factor A-2c	8.00E-135	*Oryza sativa*
mPdIRD26	F: CCTCCAGTTCATGCTTCTCC R: GAGCAGACCCGACAGACAAT	$(ACC)_7$	189–192	PDK_20s130094114	13441	13461	Protein transport protein Sec24-like At3g07100	4.00E-99	*Arabidopsis thaliana*
mPdIRD28	F: GAAACGGTATCGGGATGATG R: TTAACGACGCCGTTTCCT	$(TGA)_7$	299–306	PDK_20s1327431	28753	28773	Nuclear cap-binding protein subunit 2	3.00E-82	*Arabidopsis thaliana*

Table 1. Continued.

Locus	Primer sequences (5′–3′)[a]	Repeat motif	Size range (bp)[b]	Scaffold ID	Start	Stop	Gene annotation	*E*-value	Organism
mPdIRD29	F: GGCTCCACCATCATTGACA R: AACAGCATCGACTGCCTTCT	$(CCA)_7$	205–217	PDK_20s1359471	804	824	Putative pectinesterase 14	1.00E-34	*Arabidopsis thaliana*
mPdIRD30	F: GCAGATGGTTGAAAGCTCCT R: CCCCATTAACAGGATCAACG	$(TCA)_7$	218–224	PDK_20s1398581	15353	15373	No hit		
mPdIRD31	F: GCAGGTGGACTGCAAAATCT R: CTATTGGGGTGCTGATCCAT	$(CCA)_7$	343–372	PDK_20s1419261	29072	29092	Flowering time control protein FY	4.00E-76	*Arabidopsis thaliana*
mPdIRD32	F: AAGAAGACATTCCGGCTGGT R: GCGGGTGTGTGATATTGATG	$(ATC)_7$	148–163	PDK_20s1457341	3172	3192	Probable alpha-glucosidase Os06g0675700	0.0	*Oryza sativa*
mPdIRD33	F: GGAGCATACAGTGGGTTTGC R: CAGCCTGGGAATGAGGATAG	$(CAG)_7$	189–213	PDK_20s1569281	5206	5226	Putative clathrin assembly protein At4g25940	6.00E-133	*Arabidopsis thaliana*
mPdIRD35	F: CAGCCCCTTACTCAGACTGG R: CCCATAAGCTGATTGTGCTG	$(GCA)_7$	209	PDK_20s1690511	5056	5076	No hit		
mPdIRD36	F: GACACGTTGACGATGTGGAA R: CCATTGCTGTTGAGGAGGAG	$(TCA)_8$	162–177	PDK_20s1457341	3210	3233	Probable alpha-glucosidase Os06g0675700	0.0	*Oryza sativa*
mPdIRD37	F: TTTCCTGCTCGAAAGACACC R: CTTAGCCAGCCTCCACACTC	$(AGC)_9$	171–191	PDK_20s1521781	15593	15619	Hydroxyphenylpyruvate reductase	3.00E-71	*Solenostemon scutellarioides*
mPdIRD40	F: GAGAGATGCGTCAGGGAATC R: CCAGAATCTTCCAAGCAAGC	$(CCAGTG)_4$	175–211	PDK_20s1327401	16193	16216	No hit		
mPdIRD42	F: GAGGCAAAACTATGGGAAGC R: TTCACTGGAGCAAGGGTAGG	$(CCAGCA)_4$	82–86	PDK_20s1397171	13789	13812	Histone-lysine *N*-methyltransferase SUVR2	6.00E-04	*Arabidopsis thaliana*
mPdIRD43	F: GCAGCCATTGCTTACAGTGA R: TAAACTGCTGCCTTCCTTGG	$(AACCCT)_4$	202–208	PDK_20s1411101	2862	2885	Chaperone protein ClpB1	2.00E-05	*Arabidopsis thaliana*
mPdIRD44	F: CAGATCCGGGAGATGATGAA R: AGCAGGAGCAGCTGCATAA	$(TGGTGC)_4$	263	PDK_20s1467201	3121	3144	Two-component response regulator ARR2	2.00E-06	*Arabidopsis thaliana*
mPdIRD45	F: TAGCCTGTGCATGTTCGTTG R: AACAGCAGCTGATGGTGATG	$(AGCATC)_4$	197	PDK_20s1473281	13788	13811	No hit		
mPdIRD46	F: ATGGGTCCATTGGAGGAACT R: GACGGAGACCTTGACTGCTC	$(CAGGCA)_4$	173–197	PDK_20s1677871	3983	4006	Protein spotted leaf 11	0.0	*Oryza sativa*

[a] Annealing temperature for all primers is 60°C.
[b] Size ranges were compiled from all amplification experiments conducted on seven *Phoenix* species.

P. reclinata Jacq. (2), *P. roebelenii* O'Brien (2), *P. rupicola* T. Anderson (2), *P. theophrasti* Greuter (2), and the interspecific hybrid *P. canariensis* × *P. sylvestris* (Table 2). Among these loci, 15 showed consistent amplification and promising polymorphisms across the sample and were further investigated in a variable number of individuals (80–1000) of the aforementioned species, including population samplings of *P. dactylifera* and *P. reclinata*. The transferability of 10 loci was also evaluated in *Chamaerops humilis* L., resulting in 100% positive amplification, with eight polymorphic loci displaying two to 12 alleles among seven to 51 individuals (Table 3). Moreover, the amplification of one *Hyphaene thebaica* Mart. individual and one *Livistona carinensis* (Chiov.) J. Dransf. & N.W. Uhl individual was tested for five loci, with both species giving positive amplification results in three loci (mPdIRD25, mPdIRD31, and mPdIRD33).

DNA from these individuals was extracted from freeze-dried or silica-dried leaf tissue. Samples were reduced into a fine powder using either an IKA A10 analytical grinder (IKA-Werke, Staufen, Germany) or a QIAGEN TissueLyser and QIAGEN DNeasy Plant Mini, Maxi, or 96-well kits (QIAGEN, Courtaboeuf, France). PCR reactions were performed in a thermocycler (Biometra GmbH, Göttingen, Germany, or Eppendorf AG, Hamburg, Germany) in a total reaction mixture of 25 µL, containing: 10 ng of total genomic DNA, 1× PCR buffer, 2 mM $MgCl_2$, 200 µM dNTP, 0.5 U of *Taq* DNA polymerase, 0.4 pmol of the forward primer labeled with a 5′ M13 tail, 2 pmol of the reverse primer, and 2 pmol of the fluorochrome-marked M13 tail, plus sterile water to reach the final volume. The fluorochromes used were either 6-FAM, HEX, or TAMRA. The PCR parameters were as follows: denaturation for 2 min at 94°C; followed by six cycles at 94°C for 45 s, 60°C for 1 min, and 72°C for 1 min; then 30 cycles at 94°C for 45 s, 55°C for 1 min, and 72°C for 1.5 min; then 10 cycles at 94°C for 45 min, 53°C for 1 min, 72°C for 1.5 min; and a final elongation step at 72°C for 10 min.

The PCR products were processed on an ABI 3130XL Genetic Analyzer (Applied Biosystems, Foster City, California, USA). Allele size scoring was performed with respect to a noncommercial ladder using GeneMapper version 3.7 software (Applied Biosystems).

Genetic analyses (number of alleles, observed and expected heterozygosities, Wright's fixation index [F_{IS}] and its significance calculated using the permutation test) were conducted with GENETIX version 4.05 software (Belkhir et al., 2004).

Each of the 15 loci tested were polymorphic in at least one *Phoenix* species (Tables 2 and 3). The loci mPdIRD25, mPdIRD30, mPdIRD31, mPdIRD33, and mPdIRD40 were particularly suitable in *P. dactylifera* with three to eight alleles, having a clear stepwise mutation pattern in accordance with the microsatellite motif (tri- or hexanucleotide), and showing little to moderate heterozygosity deficit. The loci mPdIRD13, mPdIRD25, mPdIRD31, and mPdIRD33 were useful in *Chamaerops humilis* with three to 12 alleles, confirming good intergeneric transferability. In addition, mPdIRD25, mPdIRD31, and mPdIRD33 were amplified in *Livistona carinensis* and *Hyphaene thebaica*.

Table 2. Test of functionality of the 33 loci across the *Phoenix* genus.[a]

Locus	Pdac (7)	Prec (2)	Proe (2)	Prup (2)	Pthe (2)	Phyb (1)	All (16)	SM[b]	Locus comment
mPdIRD01	M	M	M	M	M	M	M	—	100% amplification, monomorphic
mPdIRD03	P	M	M	M	Failed	Failed	P	3	Partial amplification, intra- or interspecific polymorphism
mPdIRD04	M	M	M	M	M	M	P	3	100% amplification, interspecific polymorphism
mPdIRD05	M	M	M	M	M	M	P	No	100% amplification, interspecific polymorphism
mPdIRD07	M	M	M	M	P	M	P	3	Partial amplification, intra- or interspecific polymorphism
mPdIRD08	M	Failed	Failed	Failed	Failed	Failed	M	—	Partial amplification, monomorphic
mPdIRD10	P	P	M	Failed	M	Failed	P	3	Partial amplification, intra- or interspecific polymorphism
mPdIRD11	P	P	P	M	M	M	P	3	100% amplification, intra- or interspecific polymorphism
mPdIRD13	P	P	P	M	P	M	P	3	100% amplification, intra- or interspecific polymorphism
mPdIRD14	M	Failed	Failed	Failed	Failed	Failed	M	—	Partial amplification, monomorphic
mPdIRD15	M	M	M	M	M	P	P	No	100% amplification, interspecific polymorphism
mPdIRD16	P	M	M	M	M	M	P	3	100% amplification, intra- or interspecific polymorphism
mPdIRD17	M	M	M	M	M	P	P	3	100% amplification, interspecific polymorphism
mPdIRD20	M	P	M	M	P	M	P	3	100% amplification, intra- or interspecific polymorphism
mPdIRD22	M	M	M	P	M	M	P	3	100% amplification, intra- or interspecific polymorphism
mPdIRD24	M	M	M	M	M	M	M	—	100% amplification, monomorphic
mPdIRD25	P	P	M	M	M	M	P	3	100% amplification, intra- or interspecific polymorphism
mPdIRD26	P	M	M	M	M	M	P	3	Partial amplification, intra- or interspecific polymorphism
mPdIRD28	P	M	P	M	M	M	P	No	100% amplification, intra- or interspecific polymorphism
mPdIRD29	P	P	M	Failed	Failed	Failed	P	3	Partial amplification, intra- or interspecific polymorphism
mPdIRD30	P	P	Failed	Failed	M	Failed	P	3	Partial amplification, intra- or interspecific polymorphism
mPdIRD31	P	P	M	M	M	M	P	3	100% amplification, intra- or interspecific polymorphism
mPdIRD32	M	P	M	M	M	P	P	3	100% amplification, intra- or interspecific polymorphism
mPdIRD33	P	P	M	M	M	M	P	3	100% amplification, intra- or interspecific polymorphism
mPdIRD35	M	M	M	M	M	M	M	—	100% amplification, monomorphic
mPdIRD36	M	P	M	M	M	P	P	3	100% amplification, intra- or interspecific polymorphism
mPdIRD37	P	P	P	P	M	P	P	3	Partial amplification, intra- or interspecific polymorphism
mPdIRD40	P	P	M	M	P	P	P	3	100% amplification, intra- or interspecific polymorphism
mPdIRD42	P	Failed	Failed	Failed	Failed	Failed	P	No	Partial amplification, intra- or interspecific polymorphism
mPdIRD43	P	M	M	M	M	M	P	6	100% amplification, intra- or interspecific polymorphism
mPdIRD44	P	Failed	Failed	Failed	Failed	Failed	P	No	Partial amplification, intra- or interspecific polymorphism
mPdIRD45	M	Failed	Failed	Failed	Failed	Failed	M	—	Partial amplification, monomorphic
mPdIRD46	P	P	P	P	P	P	P	6	100% amplification, intra- and interspecific polymorphism

Note: M = monomorphic; P = polymorphic; Pdac = *Phoenix dactylifera*; Prec = *Phoenix reclinata*; Proe = *Phoenix roebelenii*; Prup = *Phoenix rupicola*; Pthe = *Phoenix theophrasti*; Phyb = *Phoenix canariensis* × *Phoenix sylvestris*; SM = stepwise mutation pattern.

[a] Species abbreviations are presented with the number of samples tested in parentheses. Herbarium voucher information: Pdac = *dac1*: cultivated, Kew, United Kingdom, MWC 1395 (K); *dac2*: cultivated, Elche, Spain, cv. 'Zahidi', MWC 1800/Barrow 77 (K); *dac3*: cultivated, Kew, MWC 1891 (K); *dac4*: cultivated, Kew, MWC 1398/Kew 1987-3379 (K); *dac5*: cultivated, Kew, MWC 1164 (K); *dac6*: feral, Gran Canaria, Pintaud 636 (G); *dac7*: cultivated Faisalabad, Pakistan, cv. 'Khadrawy', Pintaud 648 (G); Prec = *rec1*: Djibouti, Pintaud 642 (G); *rec2*: Zimbabwe, MWC 1874/Wilkin 724 (K); Proe = *roe1*: cultivated, Thailand, MWC 1161/Barrow 26 (K); *roe2*: cultivated, United Kingdom, MWC 1400/Kew 1987-530; Prup = *rup1*: cultivated, United Kingdom (from India), Pintaud 586 (G); *rup2*: Samchi, Bhutan, MWC 1162/Grierson and Long 3414 (K); Pthe = *the1*: cultivated, Sanremo, Italy, Pintaud 646 (G); Phyb = cultivated, Sanremo, Italy, no. 91005.

[b] In cases where stepwise mutation occurs, the number of base pairs of the repeat unit is given.

Table 3. Polymorphism characterization for 15 loci in *Phoenix* and 10 loci in *Chamaerops*.

	Phoenix all/Pdac/Prec		*Phoenix dactylifera*			*Chamaerops humilis*	
Locus	*N*	*A*	H_o	H_e	F_{IS}	*N*	*A*
mPdIRD11	18/9/2	2/2/2	—	—	—	—	—
mPdIRD13	700/560/25	10/2/4	—	—	—	7	4
mPdIRD16	100/87/2	3/2/1	—	—	—	7	2
mPdIRD20	100/87/2	5/1/2	—	—	—	7	5
mPdIRD22	100/87/2	5/1/1	—	—	—	—	—
mPdIRD25	300/108/60	5/4/2	0.29	0.42	0.31*	51	3
mPdIRD28	184/108/15	9/4/3	0.06	0.44	0.85*	—	—
mPdIRD30	83/28/15	4/3/2	0.11	0.10	–0.04	—	—
mPdIRD31	850/573/85	12/4/4	0.19	0.20	0.03	51	3
mPdIRD32	186/108/15	6/1/4	—	—	—	51	2
mPdIRD33	1000/618/85	12/4/8	0.19	0.23	0.16*	51	12
mPdIRD36	186/108/15	5/1/3	—	—	—	51	1
mPdIRD40	1000/645/85	11/8/6	0.47	0.53	0.11*	51	2
mPdIRD43	100/87/2	2/2/1	—	—	—	7	1
mPdIRD46	80/32/5	6/3/3	—	—	—	—	—

Note: *A* = number of alleles; F_{IS} = fixation index for inbreeding within populations; H_e = expected heterozygosity; H_o = observed heterozygosity; *N* = number of individuals tested; Pdac = *Phoenix dactylifera*; *Phoenix* all = all individuals of seven *Phoenix* species; Prec = *Phoenix reclinata*.

* Significant departure from Hardy–Weinberg equilibrium.

CONCLUSIONS

The loci described here are a useful addition to previously published microsatellite markers for palms. Their interspecific allelic differentiation makes them particularly suitable for hybrid and gene flow analysis within *Phoenix*. The most polymorphic loci can be added to other SSR loci to create marker sets for genetic diversity analysis in *P. dactylifera* and other species. Their transferability within the Coryphoideae subfamily will facilitate the study of species with limited molecular resources, such as *Chamaerops humilis*.

LITERATURE CITED

Al-Dous, E. K., B. George, M. E. Al-Mahmoud, M. Y. Al-Jaber, H. Wang, Y. M. Salameh, E. K. Al-Azwani, et al. 2011. *De novo* genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). *Nature Biotechnology* 29: 521–527.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410.

Arabnezhad, H., M. Bahar, H. R. Mohammadi, and M. Latifian. 2012. Development, characterization and use of microsatellite markers for germplasm analysis in date palm (*Phoenix dactylifera* L.). *Scientia Horticulturae* 134: 150–156.

Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme. 2004. GENETIX 4.05, logiciel sous Windows™ pour la génétique des populations. Laboratoire Génome, Populations, Interactions, Université de Montpellier II, Montpellier, France.

Benson, G. 1999. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Research* 27: 573–580.

Billotte, N., N. Marseillac, P. Brottier, J.-L. Noyer, J.-P. Jacquemoud-Collet, C. Moreau, T. Couvreur, et al. 2004. Nuclear microsatellite markers for the date palm (*Phoenix dactylifera* L.): Characterization and utility across the genus *Phoenix* and in other palm genera. *Molecular Ecology Notes* 4: 256–258.

Cherif, E., S. Zehdi, K. Castillo, N. Chabrillange, S. Abdoulkader, J.-C. Pintaud, S. Santoni, A. Salhi-Hannachi, et al. 2013. Male-specific DNA markers provide genetic evidence of an XY chromosome system, a recombination arrest and allow the tracing of paternal lineages in date palm. *New Phytologist* 197: 409–415.

Poncet, V., M. Rondeau, C. Tranchant, A. Cayrel, S. Hamon, A. de Kochko, and P. Hamon. 2006. SSR mining in coffee tree EST databases: Potential use of EST–SSRs as markers for the *Coffea* genus. *Molecular Genetics and Genomics* 276: 436–449.

Rozen, S., and H. J. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. *In* S. Misener and S. A. Krawetz [eds.], Methods in molecular biology, vol. 132: Bioinformatics methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

Zhao, Y., R. Williams, C. S. Prakash, and G. He. 2012. Identification and characterization of gene-based SSR markers in date palm (*Phoenix dactylifera* L.). *BMC Plant Biology* 12: 237.