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, Wiley, 2014, 2 (1), 10.3732/apps.1300058. hal-01268884

HAL Id: hal-01268884
https://hal.archives-ouvertes.fr/hal-01268884
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
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
IN SILEX MINING OF MICROSATellites IN CODING SEQUENCES OF THE DATE PALM (ARECACEAE) GENOME, CHARACTERIZATION, AND TRANSFERABILITY

FRÉDÉRIQUE ABERLENC-BERTOSSI, KARINA CASTILLO, CHRISTINE TRANCHANT-DUBEUIL, EMIRA CHÉRIF, MARCO BALLARDINI, SABIRA ABDOUNAKER, MURIEL GROS-BALTHAZARD, NATHALIE CHABRI LLANGE, SYLVAIN SANTONI, ANTONIO MERCURY, AND JEAN-CHRISTOPHE PINTAUD

1IRD, UMR DIADE—BDP, DYNADIV, and EVODYN teams, 911 Av. Agropolis, BP 64501, 34394 Montpellier, Cedex 5, France; 2Laboratoire de génétique moléculaire, immunologie et biotechnologie, Faculté des Sciences de Tunis, Campus Universitaire, 2092 El Manar, Tunis, Tunisia; 3Consiglio per la ricerca e la sperimentazione in agricoltura—Unità di Ricerca per la Floricoltura e le Specie Ornamentali (CRA-FSO), Corso degli Inglese 508, I-18038 Sanremo (IM), Italy; 4Consiglio per la ricerca e la sperimentazione in agricoltura—Unità di Ricerca per la Floricoltura e le Specie Ornamentali (CRA-FSO), Corso degli Inglese 508, I-18038 Sanremo (IM), Italy; 5ISV/CERD, route de l’Aéroport, BP 468, Djibouti; 6Centre de Bio-Archéologie et d’Ecologie (UMR 5059 CNRS/Université Montpellier 2/EPHE/INRAP), Institut de Botanique, 163 Rue Auguste Broussouet, 34090 Montpellier, France; and 7INRA, 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: Arecaeeae; Coryphoideae; microsatellite/SSR mining; Phoenix dactylifera; transferability.

The date palm (Phoenix dactylifera L.) is a monocotyledon species belonging to the Arecaeeae 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 characterization, and transferability.

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 Skaltsky, 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), P. bchafti (7), P. canariensis (3), P. falconeri (3), P. marginata (3), and P. pachypoda (2) (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.
| Locus   | Primer sequences (5′–3′) | Repeat motif | Size range (bp) | Scaffold ID | Start | Stop | Gene annotation                                      | E-value | Organism             |
|---------|--------------------------|--------------|----------------|-------------|-------|------|---------------------------------------------------|---------|----------------------|
| mPdIRD01 | F: CTCGGAAGGGTATGGAACAA  | (AAG)$_3$    | 200            | PDK_20s1306691 | 24393 | 24401 | Putative pectinesterase/pectinesterase inhibitor 28 | 4.00E-87 | *Arabidopsis thaliana* |
|         | R: TTGCTTGGGCTGATGAGTA   |              |                |             |       |      |                                                   |         |                      |
| mPdIRD03 | F: CATTAGTCAACACACACCAC  | (CCT)$_6$    | 192–198        | PDK_20s1315791 | 3431  | 3448 | Cysteine-rich receptor-like protein kinase 2      | 1.00E-166 | *Arabidopsis thaliana* |
|         | R: GCAAACACAGCTCTGGTACAC |              |                |             | 9405  | 9422 |                                                   |         |                      |
| mPdIRD04 | F: TCATTAGTCAACATGGTTGG  | (GAT)$_6$    | 301–302        | PDK_20s1366071 | 11666 | 11683 | DEAD-box ATP-dependent RNA helicase ISE2, chloroplastic | 3.00E-09 | *Arabidopsis thaliana* |
|         | R: ACCATCCATGAGCTCCAG    |              |                |             | 3737  | 3754 |                                                   |         | *Oryza sativa*        |
| mPdIRD05 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 202            | PDK_20s1402051 | 10945 | 10962 | No hit –                                            |         |                      |
|         | R: TGACTGCTCGTCATCAGGGG  |              |                |             | 194-214 |     |                                                   |         |                      |
| mPdIRD06 | F: ATGCGTTCATCTCCCTTGAG | (CAG)$_6$    | 184            | PDK_20s1405881 | 31976 | 31993 |                                                   | 4.00E-82 | *Wheat*              |
|         | R: CCTGCAAAACATCATCTCCAC |              |                |             |       |      |                                                   |         |                      |
| mPdIRD07 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 319–321        | PDK_20s1387131 | 3737  | 3754 | No hit –                                            |         |                      |
|         | R: CTTGCAAGTCTCTCCAACAC |              |                |             | 194-214 |     |                                                   |         |                      |
| mPdIRD08 | F: TGACTGCTCGTCATCAGGGG  | (AAG)$_6$    | 202            | PDK_20s1402051 | 10945 | 10962 | No hit –                                            |         |                      |
|         | R: ACCATCCATGAGCTCCAG    |              |                |             | 194-214 |     |                                                   |         |                      |
| mPdIRD10 | F: ATGCGTTCATCTCCCTTGAG | (CAG)$_6$    | 184            | PDK_20s1387131 | 3737  | 3754 | No hit –                                            |         |                      |
|         | R: CCTGCAAAACATCATCTCCAC |              |                |             |       |      |                                                   |         |                      |
| mPdIRD11 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 309–317        | PDK_20s1422721 | 4385  | 4402 | Two-component response regulator-like APRR9       | 5.00E-18 | *Arabidopsis thaliana* |
|         | R: CACAAACACCTCTGGTCC    |              |                |             |       |      |                                                   |         |                      |
| mPdIRD13 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 309–317        | PDK_20s1496731 | 12538 | 12555 | Trihelix transcription factor GT-2               | 8.00E-62 | *Arabidopsis thaliana* |
|         | R: CACAAACACCTCTGGTCC    |              |                |             |       |      |                                                   |         |                      |
| mPdIRD14 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 198–227        | PDK_20s1503531 | 9121  | 9138 | Probable ascorbate-specific transmembrane electron transporter 1 | 1.00E-82 | *Oryza sativa*        |
|         | R: CACAAACACCTCTGGTCC    |              |                |             |       |      |                                                   |         |                      |
| mPdIRD15 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 309–317        | PDK_20s1507261 | 2378  | 2395 | Eukaryotic translation initiation factor 2 subunit beta | 1.00E-22 | *Wheat*              |
|         | R: CACAAACACCTCTGGTCC    |              |                |             |       |      |                                                   |         |                      |
| mPdIRD16 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 198–227        | PDK_20s1521921 | 7038  | 7055 | Probable WRY1 transcription factor 41             | 3.00E-47 | *Arabidopsis thaliana* |
|         | R: CACAAACACCTCTGGTCC    |              |                |             |       |      |                                                   |         |                      |
| mPdIRD17 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 198–227        | PDK_20s1549911 | 54838 | 54855 | Flowering time control protein FCA              | 3.00E-38 | *Arabidopsis thaliana* |
|         | R: CACAAACACCTCTGGTCC    |              |                |             |       |      |                                                   |         |                      |
| mPdIRD20 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 309–317        | PDK_20s1640771 | 6702  | 6719 | Transcription factor bHLH62                      | 7.00E-57 | *Arabidopsis thaliana* |
|         | R: CACAAACACCTCTGGTCC    |              |                |             |       |      |                                                   |         |                      |
| mPdIRD22 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 309–317        | PDK_20s172551 | 2878  | 2895 | Probable peptide/nitrate transporter At1g59740 | 4.00E-40 | *Arabidopsis thaliana* |
|         | R: CACAAACACCTCTGGTCC    |              |                |             |       |      |                                                   |         |                      |
| mPdIRD24 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 309–317        | PDK_20s176271 | 5194  | 5211 | Probable nucleolar protein 5-1                  | 2.00E-46 | *Arabidopsis thaliana* |
|         | R: CACAAACACCTCTGGTCC    |              |                |             |       |      |                                                   |         |                      |
| mPdIRD25 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 309–317        | PDK_20s1813761 | 4692  | 4709 | Heat stress transcription factor A-2c             | 8.00E-135| *Oryza sativa*        |
|         | R: CACAAACACCTCTGGTCC    |              |                |             |       |      |                                                   |         |                      |
| mPdIRD26 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 309–317        | PDK_20s13004114 | 13441 | 13461 | Protein transport protein Sec24-like At3g07100 | 4.00E-99 | *Arabidopsis thaliana* |
|         | R: CACAAACACCTCTGGTCC    |              |                |             |       |      |                                                   |         |                      |
| mPdIRD28 | F: CATTAGTCAACATGGTTGG  | (GAT)$_6$    | 309–317        | PDK_20s1327431 | 28753 | 28773 | Nuclear cap-binding protein subunit 2           | 3.00E-82 | *Arabidopsis thaliana* |
|         | R: CACAAACACCTCTGGTCC    |              |                |             |       |      |                                                   |         |                      |
TABLE 1. Continued.

| Locus   | Primer sequences (5′–3′) | Repeat motif | Size range (bp) | Scaffold ID | Start  | Stop   | Gene annotation                      | E-value   | Organism                      |
|---------|--------------------------|--------------|-----------------|-------------|---------|--------|--------------------------------------|-----------|-------------------------------|
| mPdIRD29 | F: GGCTGGACCACCATGGACA   | (CCA)₇       | 205–217         | PDK_20s1359471 | 804     | 824    | Putative pectinesterase 14           | 1.00E-34  | *Arabidopsis thaliana*       |
|         | R: AACAGCATGAGCTGCTTCT    |              |                 |             |         |        |                                      |           |                              |
| mPdIRD30 | F: GCAAGTGCTGAAAGCTCCT   | (TCA)₇       | 218–224         | PDK_20s1398581 | 15353   | 15373  | No hit                               | 4.00E-76  | *Arabidopsis thaliana*       |
|         | R: CCCATTAAACAGGATCACGG   |              |                 |             |         |        |                                      |           |                              |
| mPdIRD31 | F: GCAAGTGCTGAAAGCTCCT   | (CCA)₇       | 343–372         | PDK_20s1419261 | 29072   | 29092  | Flowering time control protein FY   | 0.0       | *Oryza sativa*                |
|         | R: CTATTGGGCTGCTGATCCT    |              |                 |             |         |        |                                      |           |                              |
| mPdIRD32 | F: AGAGAAGATTTTGGGCCTGT  | (ATC)₇       | 148–163         | PDK_20s1457341 | 3172    | 3192   | Probable alpha-glucosidase Os06g0675700 | 0.0       | *Oryza sativa*                |
|         | R: GGAGGTGTGATGATTGATG    |              |                 |             |         |        |                                      |           |                              |
| mPdIRD33 | F: GGAGCATACAGTGGGGTTGC  | (CAG)₇       | 189–213         | PDK_20s1569281 | 5206    | 5226   | Putative clathrin assembly protein At4g25940 | 6.00E-133 | *Arabidopsis thaliana*       |
|         | R: CAGCTGGGAAAGAGGATAGG   |              |                 |             |         |        |                                      |           |                              |
| mPdIRD35 | F: CAGGGGTTACTCGAGATCG    | (GCA)₇       | 209              | PDK_20s1690511 | 5056    | 5076   | No hit                               | 0.0       | *Oryza sativa*                |
|         | R: CCCATAAGGCTGATTGCTG    |              |                 |             |         |        |                                      |           |                              |
| mPdIRD36 | F: GACATGTGGACGACGAAAGA  | (TCA₉)       | 162–177         | PDK_20s1457341 | 3210    | 3233   | Probable alpha-glucosidase Os06g0675700 | 0.0       | *Solenostemon scutellarioides*|
|         | R: CCATTGCTGATTGAGGAGG    |              |                 |             |         |        |                                      |           |                              |
| mPdIRD37 | F: TTTCTGCTGAGAAGACACC   | (AGC)₉       | 171–191         | PDK_20s1521781 | 15593   | 15619  | Hydroxyphenylpyruvate reductase      | 3.00E-71  | *Solenostemon scutellarioides*|
|         | R: CTTAGCCAGCCTCCACACTC   |              |                 |             |         |        |                                      |           |                              |
| mPdIRD40 | F: GAGGAAGTGGCTAGGGAGATC | (CCAGTG)₄    | 175–211         | PDK_20s1327401 | 16193   | 16216  | No hit                               | 0.0       | *Oryza sativa*                |
|         | R: CCAGAATCTCCTCAAGGACGC  |              |                 |             |         |        |                                      |           |                              |
| mPdIRD42 | F: GAGGGAAAATCTAGGAGGAGAC| (CCAGCA)₄    | 82–86           | PDK_20s1397171 | 13789   | 13812  | Histone-lysine N-methyltransferase SUVR2 | 6.00E-04  | *Arabidopsis thaliana*       |
|         | R: TTACCTGGAGCCAGGTTAGG   |              |                 |             |         |        |                                      |           |                              |
| mPdIRD43 | F: GCAAGCTATTGCTAGAGAAGA| (AACCCT)₄    | 202–208         | PDK_20s1411101 | 2862    | 2885   | Chaperone protein ClpB1              | 2.00E-05  | *Arabidopsis thaliana*       |
|         | R: TAAACTGCTCCTCTTTTG     |              |                 |             |         |        |                                      |           |                              |
| mPdIRD44 | F: CAGCTGCGAGAATGAGAAA   | (TGGTGCG)₄   | 263              | PDK_20s1467201 | 3121    | 3144   | Two-component response regulator ARR2 | 2.00E-06  | *Arabidopsis thaliana*       |
|         | R: AGCAAGCGACCTGCAAAGAT  |              |                 |             |         |        |                                      |           |                              |
| mPdIRD45 | F: TAGGCCTGCTGAGTTGCTT   | (AGCATC)₄    | 197              | PDK_20s1473281 | 13788   | 13811  | No hit                               | 0.0       | *Oryza sativa*                |
|         | R: AACAGCAGCTGATGGTATG    |              |                 |             |         |        |                                      |           |                              |
| mPdIRD46 | F: ATGGCTCTAGGATGGAGACT  | (CAGGCA)₄    | 173–197         | PDK_20s1677871 | 3983    | 4006   | Protein spotted leaf 11             | 0.0       | *Oryza sativa*                |
|         | R: GACCGGAGCTTAGTACTGCTC  |              |                 |             |         |        |                                      |           |                              |

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 polymorphism 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 an Eppendorf (Hamburg, Germany) in a total reaction mixture of 25 μL, containing: 10 ng of total genomic DNA, 1× PCR buffer, 2 mM MgCl₂, 200 μM dNTP, 0.5 μL of Taq DNA polymerase, 0.4 pmol of the forward primer labeled with a 5' fluorochrome-marked M13 tail, plus sterile water to reach the final volume. The PCR mixture was denatured for 2 min at 94°C; followed by six 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 s, 55°C for 1 min, and 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 heterozygosity, Wright’s fixation index [Fₜₐₜ] 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.⁸

| Locus      | Pdac (7) | Prec (2) | Proc (2) | Prup (2) | Pthe (2) | Phyb (1) | All (16) | SM | Locus comment                           |
|------------|----------|----------|----------|----------|----------|----------|----------|----|-----------------------------------------|
| mPdIRD01   | M        | M        | M        | M        | M        | M        | M        | —  | 100% amplification, monomorphic         |
| mPdIRD03   | M        | M        | M        | M        | Failed   | M        | Failed   | P  | 3                                        |
| mPdIRD04   | M        | M        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD05   | M        | M        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD07   | M        | M        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD08   | M        | Failed   | Failed   | Failed   | Failed   | M        | Failed   | M  | 3                                        |
| mPdIRD10   | P        | P        | M        | Failed   | Failed   | Failed   | M        | P  | 3                                        |
| mPdIRD11   | P        | P        | P        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD13   | P        | P        | P        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD14   | M        | Failed   | Failed   | Failed   | Failed   | M        | Failed   | M  | Partial amplification, monomorphic      |
| mPdIRD15   | M        | M        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD16   | P        | M        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD17   | M        | M        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD20   | M        | P        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD22   | M        | M        | P        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD24   | M        | M        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD25   | P        | P        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD26   | P        | M        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD28   | P        | M        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD29   | P        | M        | Failed   | Failed   | Failed   | M        | Failed   | M  | Partial amplification, monomorphic      |
| mPdIRD30   | P        | Failed   | Failed   | Failed   | Failed   | M        | Failed   | M  | Partial amplification, monomorphic      |
| mPdIRD31   | P        | M        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD32   | M        | M        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD33   | P        | P        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD35   | M        | M        | M        | M        | M        | M        | M        | —  | 100% amplification, monomorphic         |
| mPdIRD36   | M        | M        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD37   | M        | M        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD40   | P        | P        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD42   | P        | P        | M        | M        | M        | M        | M        | P  | 3                                        |
| mPdIRD43   | P        | Failed   | Failed   | Failed   | Failed   | M        | Failed   | M  | Partial amplification, monomorphic      |
| mPdIRD44   | P        | Failed   | Failed   | Failed   | Failed   | M        | Failed   | M  | Partial amplification, monomorphic      |
| mPdIRD45   | P        | M        | M        | M        | M        | M        | M        | 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 roebelenii; Proc = Phoenix rupicola; Pthe = Phoenix theophrasti; Phyb = Phoenix canariensis × Phoenix sylvestris; SM = stepwise mutation pattern.

-species abbreviations are presented with the number of samples tested in parentheses. Herbarium voucher information: Pdac = dacr1: cultivated, Kew, United Kingdom, MWC 1395 (K); dacr2: 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, Piantada 636 (G); dac7: cultivated Faisalabad, Pakistan, cv. ‘Khadrawy’, Piantada 648 (G); Prec = rec1: Djibouti, Piantada 642 (G); rec2: Zimbabwe, MWC 1874/Wilkin 724 (K); Proc = roe1: cultivated, Thailand, MWC 1161/Barrow 26 (K); roe2: cultivated, United Kingdom, MWC 1400/Kew 1987-530; Pthe = the1: cultivated, United Kingdom (from India), Piantada 586 (G); rup2: Samchi, Bhutan, MWC 1162/Grierson and Long 3414 (K); Pthe = the1: cultivated, Sanremo, Italy, Piantada 646 (G); Pthe = cultivated, Sanremo, Italy, no. 91005.

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

http://www.bioone.org/loi/apps

Aberlenc-Bertossi et al.—Phoenix dactylifera microsatellites

Applications in Plant Sciences 2014 2 (1): 1300058 doi:10.3732/apps.1300058
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**

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

| Locus       | N    | A    | F<sub>IS</sub> | H<sub>E</sub> | H<sub>o</sub> | Phoenix dactylifera | Chamaerops humilis |
|-------------|------|------|--------------|-------------|-------------|---------------------|-------------------|
| mPdIRD11    | 18/92| 2/2  | —            | —           | —           | —                   | —                 |
| mPdIRD13    | 700/560/25 | 10/2/4 | — | — | — | 7 | 4 |
| mPdIRD16    | 100/87/2 | 3/2/1 | 0.29 | 0.42 | 0.31* | 51 | 3 |
| mPdIRD20    | 100/87/2 | 5/1/2 | 0.06 | 0.44 | 0.85* | — | — |
| mPdIRD22    | 100/87/2 | 5/1/1 | 0.11 | 0.10 | —0.04 | — | — |
| mPdIRD25    | 300/108/60 | 5/4/2 | — | — | — | 7 | 5 |
| mPdIRD28    | 184/108/15 | 9/4/3 | 0.19 | 0.20 | 0.03 | 51 | 3 |
| mPdIRD30    | 83/28/15 | 4/3/2 | 0.19 | 0.23 | 0.16* | 51 | 3 |
| mPdIRD32    | 186/108/15 | 6/1/4 | — | — | — | 51 | 2 |
| mPdIRD33    | 1000/618/85 | 12/4/8 | — | — | — | 51 | 2 |
| mPdIRD36    | 186/108/15 | 5/1/3 | 0.19 | 0.23 | 0.16* | 51 | 12 |
| mPdIRD40    | 1000/645/85 | 11/8/6 | 0.19 | 0.23 | 0.16* | 51 | 1 |
| mPdIRD43    | 100/87/2 | 2/2/1 | 0.47 | 0.53 | 0.11* | 51 | 2 |
| mPdIRD46    | 80/32/5 | 6/3/3 | — | — | — | 51 | 1 |

* Significant departure from Hardy–Weinberg equilibrium.

**Note:** A = number of alleles; F<sub>IS</sub> = fixation index for inbreeding within populations; H<sub>E</sub> = expected heterozygosity; H<sub>o</sub> = observed heterozygosity; N = number of individuals tested; Pdac = Phoenix dactylifera; Phoenix all = all individuals of seven Phoenix species; Prec = Phoenix reclinata.

**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**

**Aberlenc-Bertossi et al.—Phoenix dactylifera microsatellites**

http://www.bioone.org/loi/apps