J. Bamboo and Rattan, Vol. 17, Nos. 1, pp. 1-10 (2018)

Limitations of the Recommended DNA Barcode Regions in Slow Evolving Plants: A Case Study of Rattans in the Western Ghats of India

Anoja Kurian[1], Suma Arun Dev[1], Sreekumar V. B[2], Muralidharan E. M[1*]
[1]*Forest Genetics and Biotechnology Division, Kerala Forest Research Institute, Peechi, Thrissur, Kerala, India*
[2]*Forest Ecology and Biodiversity Conservation Division, Kerala Forest Research Institute, Peechi, Thrissur, Kerala, India.*

ABSTRACT: Presence of environmental plasticity, homoplasies, look-alike species and species complexes, makes morphology based identification within genus Calamus very complicated. DNA barcoding is employed to enable accurate species identification within this genus in order to resolve these taxonomic complexities. We investigated the species discriminatory power of standard barcode loci (rbcL, matK and psbA-trnH) and their different combinations (rbcL+matK, matK+psbA-trnH, rbcL+psbA-trnH, rbcL+matK+psbA-trnH) using distance and similarity based analyses for 20 species of Calamus found in Western Ghats of India. In the present study standard DNA barcodes could not discriminate Calamus species. rbcL sequences did not show any nucleotide differences while barcoding gap exhibited by matK and psbA-trnH barcode regions and their combinations is not significant enough for successful species delimitation. The candidate DNA barcode regions adopted in the analysis failed to provide species- specific DNA barcodes, due to slow evolutionary rate in palms. This necessitates the need to explore new barcode regions other than plastid regions to delimit species boundaries in genus Calamus. Fast evolving regions like low copy nuclear regions can be exploited for their ability in species discrimination.

Keywords: Calamus, palms, lowcopynuclearregion, DNA barcoding, species discrimination

INTRODUCTION

Genus *Calamus,* commonly known as 'rattans', consists of spiny climbing palms, belonging to subfamily Calamoideae (Family: Arecaceae). They are commercially used in furniture industry, for matting, basketry and handicrafts. The massive genus *Calamus* (370 species) has been referred to as "protean" in Arecaceae (Uhl and Dransfield, 1987). They are predominantly seen in Asia and their distribution ranges from Indian subcontinent and south China southwards and east through Malaysia and Indonesia to Fiji, Vanuatu and tropical and subtropical regions of eastern Australia, represented by a single species in African continent (Uhl and Dransfield, 1987; Dransfield, 1992). In India, genus represented by 46 species, distributed in three phytogeographical regions *viz*. Peninsular India, Eastern Himalayas and the Andaman and Nicobar Islands (Renuka, 2001).

Even though palms have numerous observable characteristics such as leaf, stem, fruit and inflorescence structures, in depth studies in taxonomy are difficult due to unavailability of flowers and fruits most of the year (Uhl *et al*., 1995). Lack of sufficient herbarium specimens for comparison also creates an obstacle for accurate species identification (Sreekumar *et al*., 2006). Further, presence of environmental plasticity, homoplasies, look-alike species and species complexes makes morphology based identification within genus, very complicated. Modern molecular techniques like DNA barcoding are gaining prominence and expected to lend a hand in overcoming these difficulties as a supplementary tool to enable proper identification and classification of species within genus. An ideal DNA barcode could provide significant species discrimination and identification using short stretch of DNA, with high recovery rates and universality (Hebert *et al*., 2003; Chase *et al*., 2005; Kress *et al*., 2005). In

*To whom correspondence should be addressed: emmurali@kfri.org

plants, Consortium of Barcode of Life (CBOL) has recommended the use of standard barcodes for species identification across plant kingdom *viz*., coding gene regions (*rbc*L and *mat*K) and a non-coding spacer region (*psb*A-*trn*H) (CBOL, 2009). Multiple loci approach has been introduced in plant DNA barcoding due to the failure of recommended single barcode regions for species discrimination (Newmaster *et al*., 2006). The successful species identification by using combination of three markers – *mat*K, *rbc*L and nrITS2 were reported in tribe Caryoteae of subfamily Coryphoideae (Jeanson *et al*., 2011). *psb*A-*trn*H was found to be potential barcode for identification of *Phoenix dactylifera* L. cultivars (Al-Qurainy *et al*., 2011). *psb*Z-*trnf*M (CAU) region has also found to be suitable barcode for species identification in this *Phoenix* (Ballardini *et al*., 2013). The use of *psb*A-*trn*H as single barcode was recommended in a preliminary assessment for DNA barcodes for 15 Chinese *Calamus* species (Yang *et al*., 2012). Even though barcoding studies using plastid regions has been reported in palms, reliability of these barcode regions in rattans is still under doubt due to slow evolutionary rate of these regions. In the present study, CBOL recommended standard barcoding loci (*rbc*L, *mat*K and *psb*A-*trn*H) and their combinations were evaluated to develop species specific DNA barcodes for accurate identification of species in taxonomically complex genus *Calamus*.

MATERIALS AND METHODS

Plant material and sampling

Multiple accessions of 20 species of *Calamus* (*C. brandissi, C. delessertianus, C. gamblei, C. hookerianus, C. karnatakensis, C. lacciferus, C. lakshmanae, C. metzianus, C. nagbettai, C. neelagiricus, C. pseudotenuis, , C. prasinus, C. rotang, C. shendrurunii, C. stoloniferous, C. thwaitesii, C. travancoricus, C. vattayila, C. viminalis* and *C. wightii*) were collected from their natural distributional zones in the Western Ghats, India. To capture the sequence variability in barcode regions within a putative species, at least three accessions of each species were collected. Voucher specimens were deposited in the KFRI Herbarium (KFRI, Kerala).

Table 1. Primer sequences of candidate DNA barcodes

Barcode region	Primer	Primer sequence 5'-3'	References
*rbc*L	1F	ATGTCACCACAAACAGAAAC	Fay *et al*.,
	724R	TCGCATGTACCTGCAGTAGC	1997
*mat*K	472F	CCCRTYCATCTGGAAATCTTGGTT	Yu *et al*.,
	1248R	GCTRTRATAATGAGAAAGATTTCTGC	2011
*psb*A-	*psb*A	GTWATGCAYG AACGTAATGCTC	Kress *et al*.,
*trn*H	*trn*H	CGCGCATGGTGGATTCACAATCC	2005

DNA extraction, polymerase chain reaction and sequencing

Total genomic DNA was extracted from fresh and silica gel dried leaf materials using modified Cetyl trimethyl ammonium bromide (CTAB) method (Doyle and Doyle, 1990) as well as using DNeasy Plant Mini Kit (Qiagen, Germany) according to manufacturer's protocol. Three standard barcode loci of plastid genome (two coding regions; *rbc*L, *mat*K and an intergenic spacer region, *psb*A-*trn*H) and their combinations (r*bc*L+*mat*K, *mat*K+*psb*A-*trn*H, *rbcL*+*psb*A-*trn*H, *rbc*L+*mat*K+*psb*A- *trn*H) were evaluated to develop discriminant DNA barcodes. Polymerase chain reaction (PCR) amplification was

performed in a PTC-100 thermocycler (BIO-RAD, India) using 2X Taq buffer (Genei, Bangalore) with 1.5 mM MgCl2, 200 mM dNTPs, 10 pm of each primer, 2U Taq DNA polymerase (Genei, Bangalore), 50-100 ng template DNA and distilled deionized water to give a final volume of 20µL. PCR reactions were performed in following temperature profile: initial denaturing step of 94°C for 5 min, followed by 30 cycles each of denaturing at 94°C for 30s, annealing at specific temperatures depending on primers (*rbc*L & *mat*K at 60°C and *psb*A-*trn*H at 59°C) for 1 min, extension at 72°C for 1 min, and a final extension at 72°C for 10 min. Primer information and optimal PCR conditions are provided in Table 1. Sanger's dideoxy sequencing was performed on purified PCR products (Chromous Biotech, Bangalore).

Data analysis

Sequences were initially edited in *BioEdit* (Hall, 1999) and aligned using multiple alignment parameters in *CLUSTAL X* (Jeanmougin *et al*., 1998). The generated sequences were confirmed via *BLASTn* (http://blast.ncbi.nlm.nih.gov/Blast.cgi) against online nucleotide database and were deposited in Gen-Bank. Basic sequence statistics including conserved sites, variable sites, singletons and transition/ transversion ratio were calculated using default parameters in *MEGA v.6.0.* The average interspecific distance, theta prime and minimum interspecific distance were used to calculate interspecific divergence. Average intraspecific distance, theta and maximum intraspecific distance (coalescent depth) were calculated to evaluate intraspecific variation using K2P model in *MEGA v.6.0* (Kimura, 1980; Tamura *et al*., 2013) for the analyzed barcodes. DNA barcoding gap was calculated to find discriminatory power of studied barcodes. Similarity based approach using Best Match (BM) and Best Close Match (BCM) criteria based on uncorrected p-distances in *TAXONDNA* was also calculated to determine closest match of a barcode sequence by comparing it to all other sequences in aligned data set (Meier *et al*., 2006).

RESULTS

The analyzed barcode regions were successfully amplified with 100 per cent PCR efficiency in multiple accessions of *Calamus* species. The edited sequences were deposited in GenBank and accession numbers are provided in Table 2. Basic sequence information was calculated for the DNA barcode regions and their combinations (Table 3). *rbc*L showed similar sequences in all the studied species without any nucleotide differences, showed its inefficiency to discriminate species, while *mat*K and *psb*A-*trn*H barcode regions showed nucleotide variations in multiple sequence alignment. *psb*A-*trn*H (6.74%) showed highest percentage of parsimony informative sites followed by *mat*K+*psb*A-*trn*H (5.6%) and *mat*K (4.51%). Parsimony informative sites in *rbc*L were negligible when compared to other barcoding regions. Indels were more prevalent in *psb*A-*trn*H and aligned sequence lengths varied from 500bp to 800bp.

Table 2. List of barcode regions and Gen-Bank accessions

Barcode region	Gen-Bank Accession
*rbc*L	MG 907310-MG 907369
*mat*K	MG 907370-MG 907438
*psb*A-*trn*H	MG 907439-MG907489
	MG 911722-MG911727

Table 3. Basic statistics of DNA barcode regions and their combinations in *Calamus*

Barcode region	Sequence length (bp)	Conserved region	Variable region	Parsimony informative site	Singleton sites
*rbc*L	674	666	7	4	3
*mat*K	709	659	50	32	18
*psb*A-*trn*H	786	675	87	53	33
*rbc*L+*mat*K	1, 383	1, 325	57	36	21
*rbc*L+*psb*A-*trn*H	1, 460	1, 341	94	57	36
*mat*K+*psb*A-*trn*H	1, 495	1, 334	144	85	39
*rbc*L+*mat*K+*psb*A-*trn*H	2, 169	2, 000	151	89	42

Genetic divergence was estimated using six parameters for interspecific and intraspecific distances were provided (average interspecific distance, theta prime, minimum interspecific distance, average intraspecific distance, theta and maximum intraspecific distance (coalescent depth) in *MEGA v*.6.0 (Table 4). The maximum interspecific distance was observed in *psb*A-*trn*H (0.23), followed by *mat*K+*psb*A- *trn*H (0.19) while *rbc*L possess lowest inter specific distances. In barcoding gap analysis, *psb*A-*trn*H exhibited highest barcoding gap (0.222) followed by *mat*K+*psb*A-*trn*H and *mat*K (0.178 and 0.161, respectively) and rbcl showed the least (0.027) (Figure 1). Other combinations showed lower values and the lowest barcoding gap was observed in *rbc*L region (0.027) (Figure 1). However, barcoding gap exhibited by *psb*A-*trn*H and *mat*K barcode regions and their combinations was not significant enough for successful species identification as evident from Wilcoxon's signed rank test. The rate of species identification success was calculated using similarity-based method with BM and BCM parameters in *TAXONDNA* (Table 5). The identification success using BCM method was similar to that of BM method in all the analysed barcode regions and their combinations. *psb*A-*trn*H and *mat*K showed 44.6 per cent and 42.8 per cent species discrimination efficiencies respectively, *rbc*L shows only 10 per cent correct identification under BM method in *TAXONDNA*. The three locus combination (*rbc*L+*mat*K+*psb*A-*trn*H) showed highest species discrimination efficiency (65.05 %) followed by *mat*K+*psb*A-*trn*H (62.7%), while *rbc*L+*mat*K, *rbcL*+*psb*A-*trn*H combinations failed to show even 50 per cent identification success.

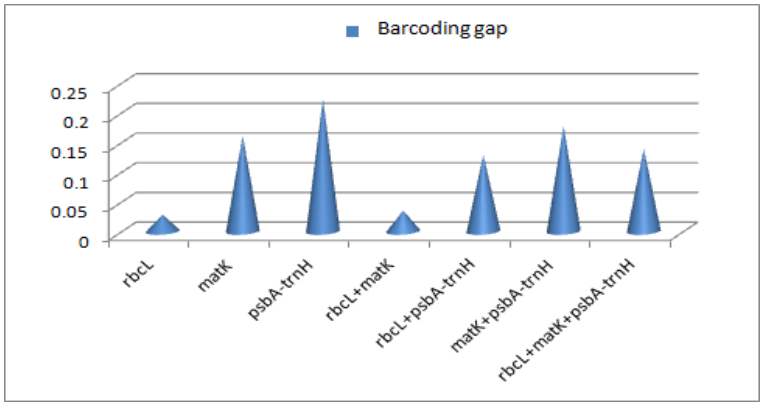

Figure 1. DNA barcoding gap for standard barcode loci and their combinations

Table 4. Genetic divergence parameters of three intergenic spacers and their combinations

Barcode region	Average Interspecific distance	Average theta prime	Inter-specific distance	Average intraspecific distance	Maximum intraspecific distance	Average theta
*rbc*L	0.028±0.019	0.0019±0.0014	0	0.0004±0.0004	0.002±0.002	0.0004±0.0004
*mat*K	0.170±0.04	0.009±0.003	0	0.009±0.002	0.023±0.007	0.0008±0.0005
*psb*A-*trn*H	0.23±0.077	0.017±0.005	0	0.0076±0.002	0.027±0.007	0.009±0.0027
*rbc*L+*mat*K	0.037±0.018	0.003±0.0012	0	0.002±0.0003	0.01±0.007	0.0006±0.0002
*rbc*L+*psb*A-trnH	0.13±0.030	0.008±0.002	0	0.001±0.0003	0.03±0.001	0.005±0.001
*mat*K+*psb*A-*trn*H	0.19±0.04	0.012±0.003	0	0.0119±0.002	0.088±0.001	0.0116±0.004
rbcL+*matK*+*psbA*-*trnH*	0.14±0.020	0.010±0.001	0	0.001±0.0003	0.01±0.001	0.004±0.001

Table 5. Identification success rates of barcode regions using *TAXONDNA* under 'best match' (BM) and 'best close match' (BCM) method

Barcode region	Best match (BM)			Best close Match		
	Correct Identification	Ambiguous Identification	Incorrect Identification	Correct Identification	Ambiguous Identification	Incorrect identification
*rbc*L-*mat*K	42.85%	34.28%	22.85%	42.85%	34.28%	22.85%
*psb*A-*trn*H	44.64%	32.14%	23.21%	44.64%	32.14%	23.21%
*rbc*L+*mat*K	45.05%	29.65%	25.3%	45.05%	29.65%	25.3%
*rbc*L+*psb*A- *trn*H	48.1%	41%	10.9%	48.1%	41%	10.9%
*mat*K+*psb*A-*trn*H	62.74%	3.92%	33.33%	62.74%	3.92%	33.33%
rbcL+*mat*K+*psb*A-*trn*H	65.05%	16.65%	18.35%	65.05%	16.65%	18.35%

DISCUSSION

In spite of its general acceptance as a universal barcode, *rbc*L generated no species specific differences for discriminating *Calamus* species, supporting previous studies (Kress *et al*., 2005; Spooner, 2009; Han *et al*., 2016) pointed out its inefficiency as potential barcode. The recommended barcodes, *mat*K and *psb*A-*trn*H also failed to show significant species identification in present study. *psb*A-*trn*H recommended single barcode for 15 Chinese *Calamus* species with 58 per cent of species discrimination was not found as suitable barcode for 21 Indian *Calamus* species with 44.6 percent identification success in the present study as reported by Yang *et al*. (2012). *psb*A-*trn*H exhibited dramatic differences in sequence lengths among congeneric species, due to the presence of indels as reported (Kress *et al*., 2005) leading to difficulties in sequence alignment in the present study. The combinations of barcoding regions (*rbc*L+*mat*K, *mat*K+*psb*A-*trn*H, *rbc*L+*psb*A-*trn*H, *rbc*L+*mat*K+*psb*A-*trn*H) also failed to give sufficient barcoding gap for species identification in the genus and hardly any species-specific nucleotide changes could be observed. The species assignment in genus was performed by 'Bestmatch (BM)', 'Best Close Match (BCM)' of *TAXONDNA*. These statistical parameters have been utilized for species assignments to evaluate success rate of species identification. In BM criteria, *TAXONDNA* finds the closest barcode match for each query and in BCM, the program plots relative frequency of intraspecific distances to determine threshold value below 95 per cent of all intraspecific

distances (Meier *et al*., 2006). In genus *Calamus*, the tested loci were not successful with BM and BCM criteria, showed less than 70% identification rate. The combination of three barcode regions (*rbc*L+*mat*K+*psb*A-*trn*H) showed highest species identification rate (65%) in both BM and BCM, but failed to delimit species boundaries. Among analysed barcoding regions, *psbA-trnH* showed highest species identification rate while *rbcL* the least. Similarity based analysis revealed inefficiency of recommended barcode regions and their combinations for species identification.

Failure of plastid data to highlight species boundaries, as in the present study was also reported in other groups like willows, wild potatoes, carex, bryophytes, among others (Percy *et al*., 2008; Spooner, 2009; Starr *et al*., 2009; Hassel *et al*., 2013). This has been attributed to widespread hybridization, introgression, or incomplete lineage sorting (Hollingsworth *et al*., 2011). The failure of recommended plastid barcode regions can attribute to slow evolutionary rate of plastid genes in palms as evidenced in earlier studies (Wilson *et al.*, 1990; Gaut *et al*., 1996). Substitution rate estimates from restriction site variation in chloroplast DNA of palms were 5 to 13-fold slower than the estimates for grasses (Wilson *et al*., 1990). Later, this was confirmed by identifying substitution rates in *rbc*L in palms, which is 5 times slower than that of grasses (Gaut *et al*., 1996). In palms, Calamoideae have the lowest substitution rate of chloroplast DNA (1.3 x 10 -10) (Wilson *et al*., 1990) which clearly indicated the inefficiency of plastid primers to discriminate species in the subfamily. The nrITS region recommended in some studies (Chase *et al*., 2005; Kress *et al*., 2005) as most divergent gene region for plant barcoding has eliminated in this study due to the presence of multiple copies as indicated by Baker *et al*. (2000) in Calamoideae. Hence, the idea of using universal DNA barcodes in *Calamus* for identification become a challenge, as all recommended barcode regions are of plastid origin. The slow evolutionary rate of cpDNA along with incomplete concerted evolution of nrDNA (Wilson *et al*., 1990; Gaut *et al*., 1992; Hahn, 1999; Lewis *et al*., 2000), necessitate the search of additional markers for species identification in palms. As low-copy nuclear genes evolve up to five times faster than plastid genome, they are being explored recently to resolve slow evolving plants (Sang, 2002; Small *et al*., 2004; Norup *et al*., 2006). Low-copy nuclear regions such as PRK (phosphoribulokinase), RPB2 (RNA polymerase II) and MS (malate synthase) have been successfully used in resolving palm phylogenetic relationships especially at lower taxonomic levels, regardless of difficulties like paralogy, concerted evolution and intragenic polymorphism e.g. *Hyophorbe* (Lewis and Martinez, 2000), Areceae (Lewis and Doyle,2002),

Heterospathe and *Rhopaloblaste* (Norup, 2004), Cocoeae (Gunn, 2004), *Borasseae* (Bayton, 2005), *Geonomeae* (Roncal *et al*., 2005), *Chamaedorea* (Thomas *et al*., 2006), *Arecinae* (Loo *et al*., 2006), *Areceae (*Norup *et al*., 2006), *Ceroxyloideae* (Trénel *et al*., 2007), *Ptychospermatinae* (Zona *et al*., 2011). Single copy genes such as waxy (Mason-Gamer *et al*., 1998), *pistillata* (Bailey and Doyle, 1999) and chloroplast expressed glutamine synthetase (Emshwiller and Doyle, 1999) also have proven to be useful as phylogenetic markers in other families. Low copy nuclear regions (*Clerm2*, *Clerm4*, *Cyrt2* and *Cyrt4*) were reported as DNA barcodes in *Clermontia* (Campanulaceae) and *Cyrtandra* (Gesneriaceae), while plastid barcode had a slow evolutionary rate as suggested in Brassicaceae (Pillon *et al*., 2013; Sun *et al*., 2015).

CONCLUSIONS

The candidate DNA barcode regions recommended by CBOL and their combinations, failed to provide species-specific markers in the genus *Calamus* due to slow evolutionary rate of chloroplast DNA in palms. In situations where cpDNA and nrDNA are unable to generate reliable data, low-copy nuclear genes have a higher potential due to higher rate of evolution than organellar sequences. Hence, the use of recommended regions will add a new dimension to techniques available for species identification and for resolving taxonomic complexities. Therefore, we recommended the use of low-copy nuclear regions such as PRK, RPB2, MS etc. for the accurate identification of *Calmaus* species.

REFERENCES

Al-Qurainy, F., Khan, S., Al-Hemaid, F.M., Ali, A.A., Tarroum, M. and 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. The International Journal of Molecular Sciences,* 12: 6871–6880.

Bailey, C.D. and Doyle, J.J. (1999). Potential phylogenetic utility of the low-copy nuclear gene *pistillata* in dicotyledonous plants: comparison to nrDNA ITS and *trnL* intron in *Sphaero cardamum* and other Brassicaceae. *Molecular Phylogenetics and Evolution,* 13: 20–30.

Baker, W.J., Hedderson, T.A. and Dransfield, J. (2000). Molecular phylogenetics of subfamily Calamoideae (Palmae) based on nrDNA *ITS* and cpDNA *rps16*intron sequence data. *Molecular Phylogenetics and Evolution,* 14(2):195–217.

Ballardini, M., Mercuri, A., Littardi, C., Abbas, S., Couderc, M., Ludeña, B. and Pintaud, J.C. (2013). The chloroplast DNA locus *psbZ-trnf*M as a potential barcode marker in *Phoenix* L. (Arecaceae). *ZooKeys,* 365: 71–82.

Bayton, R.P. (2005). Borassus, L. and the Borassoid palms: systematics and evolution. PhD thesis, University of Reading, UK.

CBOL Plant Working Group (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences USA,* 106: 12794-12797.

Chase, M. W., Cowan, R.S., Hollingsworth, P.M., van den Berg, C., Madrinan, S., Petersen, G. Seberg, O., Jørgsensen, T., Cameron, K.M., Carine, M., Pedersen, N., Hedderson, T.A.J., Conrad, F., Salazar, G.A., Richardson, J.E., Hollingsworth, M.L., Barraclough, T.G., Kelly, L. and Wilkinson, M. (2005). A proposal for a standardized protocol to barcode all land plants. *Taxon.* 56: 295- 299.

Doyle, J.J. and Doyle, J.L. (1990). A rapid total DNA preparation procedure for fresh plant tissue. *Focus,* 12: 13-15.

Dransfield, J. (1992). The rattans of Sarawak. Kew, Royal Botanic Gardens and Sarawak Forest Department, UK.

Emshwiller, E. and Doyle, J.J. (1999). Chloroplast-Expressed Glutamine Synthetase (ncpGS): Potential Utility for Phylogenetic Studies with an Example from Oxalis (Oxalidaceae). *Molecular Phylogenetics and Evolution,* 12 (3): 310-319.

Fay, M.F., Swensen, S.M. and Chase, M.W. (1992). Taxonomic affinities of *Medusagyne oppositifolia* (Medusagynaceae). *Kew Bulletin,* 52: 111-120.

Gaut, B.S., Morton, B.R., Mc Caig, B.C. and Clegg, M.T. (1992). Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbc*L. *Proceedings of the National Academy of Sciences USA,* 93: 10274–10279.

Gunn, B.F. (2004). The phylogeny of the Cocoeae (Arecaceae) with emphasis on *Cocos nucifera*. *Annals of the Missouri Botanical Garden,* 91:505–522.

Hahn, W.J. (1999). Molecular systematic studies of the Palmae. *Memoirs of the New York Botanical Garden,* 83:47–70.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis: Department of Microbiology, North Carolina State University.

Han Wei, Y., Duan, D., Ma, X.F., Jia, Y., Liu, Z.L., Zhao, G.F. and Li, Z.H. (2016). Efficient Identification of the Forest Tree Species in Aceraceae Using DNA Barcodes. *Frontiers in Plant Science,* 7: 1707.

Hassel, K., Segreto, R. and Ekrem, T. (2013). Restricted variation in plant barcoding markers limits identification in closely related bryophyte species. *Molecular Ecology Resources,* 13: 1047–1057.

Hebert, P.D.N., Cywinska, A., Ball, S.L. and deWaard, J.R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London Biological Sciences,* 270: 313–321.

Hollingsworth, P.M., Graham, S.W. and Little, D.M. (2011). Choosing and Using a Plant DNA Barcode. *PLoSONE,* 6(5): e19254.

Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G. and Gibson, T.J. (1998). Multiple sequence alignment with Clustal X. *Trends in Biochemical* Sci*ences,* 23: 403–405.

Jeanson, M.L., Labat, J.N. and Little, D.P. (2011). DNA barcoding: a new tool for palm taxonomists? *Annals of Botany,* 108: 1445–1451.

Kimura, M.A. (1980). Simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution,* 16: 111-120.

Kress, W.J., Wurdack, K.J., Zimmer, E.A., Weigt, L.A. and Janzen, D.H. (2005). Use of DNA barcodes

to identify flowering plants. *Proceedings of the National Academy of Sciences USA*, 102 (23): 8369-8374.

Lewis, C.E. and Doyle, J.J. (2002). Phylogenetic analysis of tribe Areceae (Arecaceae) using two low-copy nuclear genes. *Plant Systematics and Evolution*, 236: 1–17.

Lewis, C.E., Baker, W.J. and Asmussen, C.B. (2000). DNA and palm evolution. *Palms*, 44(1): 19–24.

Lewis, C.E. and Martinez, N. (2000). Identity of the *Hyophorbe* palms at the botanical garden of Cienfuegos, Cuba. *Palms,* 44(2): 93–97.

Loo, A.H.B., Dransfield, J., Chase, M.W. and Baker, W.J. (2006). Low-copy nuclear DNA, phylogeny and the evolution of dichogamy in the betel nut palms and their relatives (Arecinae; Arecaceae). *Molecular Phylogenetics and Evolution,* 39: 598–618.

Mason-Gamer, R.J., Weil, C. and Kellogg, E.A. (1998). Granule-bound starch synthase: structure, function, and phylogenetic utility. *Molecular Biology and Evolution,* 15: 1658–1673.

Meier, R., Shiyang, K., Vaidya, G. and Ng, P.K. (2006). DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Systematic Biology,* 55: 715–728.

Newmaster, S.G., Fazekas, A. and Ragupathy, S. (2006). DNA barcoding in the land plants: evaluation of *rbcL* in a multigene tiered approach. *Canadian Journal of Botany*, 84:335–341.

Norup, M.V. (2004). A molecular systematic study of *Heterospathe* and *Rhopaloblaste* (Arecaceae, Areceae). Master's Thesis, University of Aarhus.

Norup, M.V., Dransfield, J., Chase, M.W., Barfod, A.S., Fernando, E.S. and Baker, W.J. (2006). Homoplasious character combinations and generic delimitations: a case study from the Indo-Pacific arecoid palms (Arecaceae: Areceae). *American Journal of Botany,* 93: 1065–1080.

Percy, D.M., George, W.A., Quentin, C.C., Aron, J.F., Prasad, R.K., Kevin, S.B., Brian, C.H., Steven, G.N., Spencer, C.H.B. and Sean, W.G. (2014). Understanding the spectacular failure of DNA barcoding in willows (*Salix*): Does this result from a trans-specific selective sweep? *Molecular Ecology,* 19: 4737-56.

Pillon, Y., Johansen, J., Sakishima, T., Chamala, S., Barbazuk, W.B., Roalson, E.H., Price, D.K. and Stacy, E.A. (2013). Potential use of low-copy nuclear genes in DNA barcoding: a comparison with plastid genes in two Hawaiian plant radiations. *BMC Evolutionary Biology*, 13: 35.

Renuka, C. (2001). Palms of India: status, threats and conservation strategies. In: Uma Shaanker R., Ganeshaiah K.N., Bawa K.S. (eds), Forest Genetic Resources: Status, Threats and Conservation Strategies. Oxford & IBH Publishing Co. Pvt. Ltd., Calcutta, pp. 197-209.

Roncal, J., Francisco-Ortega, J., Asmussen, C.B. and Lewis, C.E. (2005). Molecular phylogenetics of tribe Geonomeae (Arecaceae) using nuclear DNA sequences of phosphoribulokinase and RNA polymerase II. *Systematic Botany,* 30: 275–283.

Sang, T. (2002). Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical Reviews in Biochemistry and Molecular Biology,* 37:121–147.

Small, R., Cronn, R. and Wendel, J. (2004). L.A.S Johnson Review No.2. Use of nuclear genes for

phylogeny reconstruction in plants. *Australian Systematic Botany,* 17:145–170.

Spooner, D.M. (2009). DNA barcoding will frequently fail in complicated groups: An example in wild potatoes. *American Journal of Botany,* 96 (6): 1177–1189.

Sreekumar, V.B., Renuka, C., Suma, T.B. and Balasundaran, M. (2006). Taxonomic reconsideration of *Calamus rivalis* Thw. ex Trim. and *C. metzianus* Schlecht (Arecaceae) through morphometric and molecular analyses. *Botanical Studies,* 47: 443-452.

Starr, J.R., Naczi, R.F.C. and Chouinard, B.N. (2009). Plant DNA barcodes and species resolution in sedges (*Carex*, Cyperaceae). *Molecular Ecology Resources,* 9 (1): 151–163.

Sun, X.Q., Qu, Y.Q., Yao, H., Zhang, Y.M., Yan, Q.Q. and Hang, Y.Y. (2015). Applying DNA barcodes for identification of economically important species in Brassicaceae. *Genetics and Molecular Research,*14(4): 15050-61.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution,* 28: 2731–2739.

Thomas, M.M., Garwood, N.C., Baker, W.J., Henderson, S.A., Russell, S.J., Hodel, D.R. and Bateman, R.M. (2006). Molecular phylogeny of the palm genus Chamaedorea, based on low copy nuclear genes PRK and RPB2. *Molecular Phylogenetics and Evolution,* 38: 398–415.

Trénel, P., Gustafsson, M., Baker, W.J., Asmussen, C.B., Dransfield, J. and Borchsenius, F. (2007). Mid-tertiary dispersal, not Gondwanan vicariance explains distribution patterns in the wax palm subfamily (Ceroxyloideae: Arecaceae). *Molecular Phylogenetics and Evolution*, 45: 272–288.

Uhl, N.W. and Dransfield, J. (1987). Genera Palmarum. Allen Press, Lawrence, Kansas, USA.

Uhl, N.W., Dransfield, J., Davis, J.I., Luckow, M.A., Hansen, K.S. and Doyle, J.J. (1995). Phylogenetic relationships among palms: Cladistic analyses of morphological and chloroplast DNA restriction site variation. In: Rudall P.J., Cribb P.J., Cutler D.F., Humphries C.J. (eds), Monocotyledons: Systematics and Evolution 2. pp 623–661.

Wilson, M.A., Gaut, B. and Clegg, M.T. (1990). Chloroplast DNA evolves slowly in the palm family (Arecaceae). *Molecular Biology and Evolution,* 7: 303–314.

Yang, H.Q., Dong, Y.R., Gu., Z.J., Liang, N. and Yang, J.B. (2012). A preliminary assesment of *matK*, *rbcL* and *trnH-psbA* as DNA barcodes for *Calamus* (Arecaceae) species in China with a note on ITS. *Annales of Botanici Fennici*, 49: 319-330.

Yu, J., Xue, J.H. and Zhou, S.L. (2011). New universal matK primers for DNA barcoding angiosperms. *Journal of Systematics and Evolution,* 49 (3): 176-181.