MOLECULAR SYSTEMATICS OF RATTANS OF SOUTH INDIA

Thesis submitted to

Cochin University of Science and Technology

in partial fulfilment of the

requirements for the award of the Degree of

Doctor of Philosophy

Under

Faculty of Science

by

ANOJA KURIAN

Reg. No: 4389

KSCSTE - KERALA FOREST RESEARCH INSTITUTE

THRISSUR-680653, KERALA, INDIA

November 2018

MOLECULAR SYSTEMATICS OF RATTANS OF SOUTH INDIA

Ph.D. Thesis under the Faculty of Science

Author

Anoja Kurian
Research Scholar
Department of Biotechnology
Forest Genetics and Biotechnology Division
Kerala Forest Research Institute
Thrissur -680653
Kerala, India
Email id: anojakurian@gmail.com

Supervising Guide

Dr. E.M. Muralidharan
Senior Principal Scientist
Department of Biotechnology
Forest Genetics and Biotechnology Division
Kerala Forest Research Institute
Thrissur -680653
Kerala, India
Email id: emmurali@kfri.org.in

Co-supervisor

Dr. Sreekumar V.B.
Scientist
Department of Botany
Forest Ecology and Biodiversity Conservation Division
Kerala Forest Research Institute
Thrissur -680653
Kerala, India
Email id: vbskumar@gmail.com

November 2018

KSCSTE-Kerala Forest Research Institute
(An Institution under Kerala State Council for Science, Technology and Environment)
Peechi - 680 653, Thrissur, Kerala, India

CERTIFICATE

This is to certify that the thesis entitled "**Molecular Systematics of Rattans of South India**" is a bona fide record of research work carried out by **Ms. Anoja Kurian** under our supervision in the Department of Biotechnology, Kerala Forest Research Institute, Thrissur. The results presented in this thesis or parts of it have not been presented for the award of any other degree or diploma or any other similar titles or recognition.

This is to certify that all the relevant corrections and modifications suggested by the audience during the pre-synopsis seminar and recommended by Doctoral Committee of the candidate have been incorporated in this thesis.

Dr. E. M. Muralidharan
(Supervisor)
Senior Principal Scientist
Department of Biotechnology
Forest Genetics and Biotechnology Division
Kerala Forest Research Institute

Dr. Sreekumar V.B.
(Co-supervisor)
Scientist
Department of Botany
Forest Ecology and Biodiversity Conservation Division
Kerala Forest Research Institute

Phone : 0487-2690100; Fax: 0487-2690111, e-mail : kfri@kfri.org, Website: www.kfri.res.in

DECLARATION

I hereby declare that the thesis entitled "**Molecular Systematics of Rattans of South India**" submitted to Cochin University of Science and Technology for the award of the degree of Doctor of Philosophy, is the authentic record of independent research work carried out by me under the supervision and guidance of Dr. E.M. Muralidharan and Dr. V.B. Sreekumar, Scientists, Kerala Forest Research Institute, and that no part thereof has previously formed the basis for the award of any degree or diploma or any other similar titles or recognition.

Thrissur
01-11-2018

Anoja Kurian

Dedicated To My
Parents, Husband and Daughter

ACKNOWLEDGEMENTS

I would like to express my deep reverence to God Almighty for his benevolence and for blessing me with a loving and supportive family and also for giving me the strength to complete this work.

With high esteem and gratification, I express my gratitude and indebtedness to Dr. E.M. Muralidharan, Senior Principal Scientist & Program Coordinator, Forest Genetics & Biotechnology Division, KFRI, my guide for his valuable guidance and suggestions throughout the course of this research work. I also express my heartfelt gratitude to Dr. V.B. Sreekumar, Scientist, Department of Botany, KFRI, my Co-guide, for his timely help and guidance. Their esteemed advices, immense knowledge and profound research experiences helped me a lot throughout the course of this study.

My sincere thanks are to Dr. Syam Viswananth, Director, Kerala Forest Research Institute (KFRI) as well as Dr. S. Pradeep Kumar, Dr. Bransdon S. Corrie, Dr. P.G. Latha, Dr. P.S. Easa and Dr. K. V. Sankaran, former Directors, KFRI for facilitating the successful completion of my doctoral work.

I also take this opportunity to thank Dr. T.V. Sajeev, Research Coordinator and Dr. E.A. Jayson, Dr. M. Balagopalan, Dr. K. Swarupanandan, former Research Coordinators, KFRI for their valuable suggestions and help rendered during this period. I convey my gratitude to Dr. Anita V., Nodal Officer and Dr. M.P. Sujatha, Dr. T.K. Dhamodaran, former Nodal Officer for their sensible advices to accomplish the officialities for the doctoral programme.

I thank the Kerala Forest Department, Govt. of Kerala and Karnataka Forest Department, Govt. of Karnataka for their permission to collect plant samples. The financial support received from the Kerala State Council for Science Technology and Environment (KSCSTE) is also acknowledged.

With great respect, I record my deep sense of gratitude to Dr. Suma Arun Dev, Scientist, Biotechnology department, KFRI for her valuable comments and suggestions to improvise my work.

I express my deep sense of gratitude to my lab members and beloved friends, Ms. Sijimol K., Ms. Prathibha P.S., Ms. Swathi Balakrishnan, Ms. Remya Unnikrishnan., Mr. Subin K. and Mr. Arun M.A. for their constant support and the warm companion ship extended to me. I thank each of you for all the fun we have had in the lab. They have stood by me through thick and thin, giving me the confidence that I needed. I also express my heartfelt gratitude to my friends, Mr. Sanil M.S., Mr. Saju Michael, Mr. Joyner T., Mr. Anoop B.S., Mr. Subin K..K., Ms. Sudha S., Ms. Rani Susan Thomas and Ms. Remya S.R. for providing

me with exceptional support and inspiration during the study. I also thank Mr. Sabik S., Mr. Arun Kumar and Mr. Nirmesh for extending their help for the field work of this study. The co-operation, help and lively environment extended to me by the fellow researchers of the Department of Biotechnology, KFRI are sincerely appreciated. A special thanks to all my friends especially Dr. Jinu Mathew Valayil, Dr. Suni Anie Mathew, Mr. Sajan Raju and Mr. Luigi Bianchi for their love and tremendous support.

I acknowledge library and LAN staff for their admirable services. I am indebted to all my friends, colleagues at KFRI and my school-college mates for their immense moral support and motivation to strive towards the goal.

I bow my head before my parents for their love, support, encouragement and above all, never ending faith in my dreams and aspirations. Specifically I want to appreciate them for taking care of my daughter while I was away. I cannot forget the contribution of my brother who has been a constant reservoir of support at all those times when I needed it the most. My husband has deep-rooted influence on me and without his constant wishes, love, encouragement and support it would not be possible for me to fulfil this project. My heart felt regard goes to my parents-in-law for their love and support. To my beloved daughter, my love for her are beyond words and I would like to express my thanks for being such a good girl, always cheering me up. I dedicate this piece of work to my family.

I sincerely thank everyone who has been a part of my life, my relatives, friends, teachers and well-wishers, as well as expressing my sincere apology if I could not mention any name personally. I humbly dedicate each and every success of mine, however small they may be, to each one of them…

Anoja Kurian

CONTENTS

		Page No.
	LIST OF TABLES	i
	LIST OF FIGURES	ii
	LIST OF APPENDICES	iv
	ABBREVIATIONS	v
	ABSTRACT	vi
1.	**INTRODUCTION**	1
1.1.	An Overview	1
1.2.	Palms	3
1.2.1.	Rattans	3
1.2.2.	Distribution of Rattans in India	5
1.3.	Genus *Calamus*	5
1.3.1.	*Calamus* of South India	7
1.3.2.	Species identification in the genus *Calamus*	9
1.4.	DNA Barcoding	10
1.5.	Phylogenetics in palms	11
1.5.1.	Phylogenetic classification of Palms	11
1.6.	Biogeography	12
1.6.1.	Continental drift	13
1.7.	Palm Biogeography	14
1.8.	**Objectives**	14
2.	**REVIEW OF LITERATURE**	16
2.1.	DNA BARCODING	16
2.1.1.	DNA barcodes	16
2.1.1.1.	Chloroplast Regions	17
2.1.1.2.	Nuclear Region	19
2.1.2.	Future of DNA Barcoding	20
2.1.3.	Barcoding in Palms	21
2.2.	PHYLOGENY	22

2.2.1. Molecular Phylogeny 23
2.2.1.1. Mitochondrial DNA 24
2.2.1.2. Chloroplast DNA 24
2.2.1.3. Nuclear DNA 25
2.2.1.4. Low copy nuclear regions 25
2.2.2. Palm Phylogeny 26
2.2.2.1. Molecular Phylogenetic Studies in Palms 27
2.3. BIOGEOGRAPHY 32
2.3.1. Biogeography of palms 32
2.3.2. Fossil evidence of palms in India 33
2.3.3. Formation of India 33
3. DNA BARCODING IN *CALAMUS* 36
3.1. INTRODUCTION 36
3.2. MATERIALS AND METHODS 38
3.2.1. Taxon Sampling 38
3.2.2. DNA extraction, PCR amplification and Sequencing 38
3.2.3. Sequence alignment 40
3.2.4. Data analysis 40
3.2.4. 1. Distance-based analysis 41
3.2.4.2. Similarity based analysis 41
3.2.4.3. Tree based analysis 41
3.3. RESULTS 42
3.3.1. Amplification and sequence analysis 42
3.3.2. Distance-based analysis 44
3.3.3. Similarity based analysis 45
3.3.4. Tree Based Analyses 46
3.4. DISCUSSION 48
4 **RESOLVING THE TAXONOMIC COMPLEXITIES IN *CALAMUS* OF SOUTH INDIA** 54
4.1. INTRODUCTION 54
4.1.1. Taxonomic revision of *C. metzianus, C.rotang* and *C. rivalis* 55

4.1.2. *Calamus gamblei* species complex 55
4.1.3. Analyzing morphologically similar species 56
4.2. MATERIALS AND METHODS 59
4.2.1. Sample collection 59
4.2.2. Data Analysis 59
4.3. RESULTS 60
4.3.1. Taxonomic revision of *C. metzianus*, *C. rotang* and *C. rivali*s 60
4.3.2. *Calamus gamblei* species complex 63
4.4. DISCUSSION 63
4.4.1. Taxonomic revision of *C. metzianus*, *C. rotang* and *C. rivali*s 63
4.4.2. *Calamus gamblei* species complex 65
4.4.3. Analyzing morphologically similar species 65
4.4.3.1. *C. hookerianus* and *C. pseudotenuis* 65
4.4.3.2. *C. delessertianus* and *C. dransfieldii* 66
5 **PHYLOGENETIC RELATIONSHIPS WITHIN THE GENUS *CALAMUS*** 67
5.1. INTRODUCTION 67
5.2. MATERIALS AND METHODS 70
5.2.1. Sampling 70
5.2.2. DNA extraction, PCR amplification and Sequencing 70
5.2.3. Phylogenetic analysis 71
5.3. RESULTS 72
5.4. DISCUSSION 75
6 BIOGEOGRAPHY AND PHYLOGENETIC RELATIONSHIP OF SUBFAMILY CALAMOIDEAE 80
6.1. INTRODUCTION 80
6.2. MATERIALS AND METHODS 83
6.2.1. Taxon Sampling 83
6.2.2. DNA extraction, PCR amplification and Sequencing 84
6.2.3. Phylogenetic analysis 85
6.2.4. Dating phylogenetic tree 85

6.2.5. Ancestral area reconstruction 86
6.3. RESULTS 87
6.3.1. Phylogenetic analysis 87
6.3.2. Dating phylogenetic tree 87
6.3.3. Ancestral area reconstruction 90
6.4. DISCUSSION 93
6.4.1. Rattan in India 96
6.4.2. *Calamus* in India 98
7 SUMMARY AND CONCLUSIONS 100
8 REFERENCES 103
PUBLICATIONS 134
APPENDICES

LIST OF TABLES

No.	Titles	Page No.
	Chapter 3	
3.1	Barcoding primers used and PCR reaction conditions	39
3.2	Sequence characteristics of eight DNA regions and their combinations	43
3.3	Parameters of interspecific divergence and intraspecific variation of eight barcode regions and their combination	44
3.4	Test of significance of interspecific divergence among loci in Wilcoxon-signed rank test	46
3.5	Identification success rates of eight barcode regions singly and their possible combinations using TaxonDNA under 'Best Match' and 'Best Close Match' criteria	47
	Chapter 4	
4.1	Summary statistics of four barcodes and their combinations	61
4.2	Unique nucleotide differences and Tajima's test (1993) statistics for barcode loci	63
	Chapter 5	
5.1	Species of India from Beccari's Classification	68
5.2	Primers for the phylogenetic gene regions and their PCR reaction conditions	72
5.3	Sequence characteristics and models of gene regions	75
	Chapter 6	
6.1	Phylogenetic classification of subfamily Calamoideae (Dransfield *et al*., 2005)	81
6.2	Phylogeny primers and PCR reaction conditions	84

LIST OF FIGURES

No.	Figure Legends	Page No.
	Chapter 1	
1.1	Plate 1	6
1.2	Morphology of *Calamus thwaitesii* in different geographical regions	8
	Chapter 3	
3.1	PCR amplified products of eight analysed DNA barcode regions	43
3.2	DNA barcoding gap for eight barcode loci and their possible combinations	45
3.3	Phylogenetic tree construction using low copy nuclear region RPB2, based on GTR +G model and AIC in MrBayes v.3.2.2.	49
3.4	Neighbor-joining tree (NJ) of the genus *Calamu*s using RPB2 basedon Kimura 2 Parameter model using MEGA v.6.0 with bootstrap percentage	50
	Chapter 4	
4.1	Fruit morphology in Calamus Gamblei species complex	57-58
4.2	Multiple Sequence Alignment showing transversions in the *Calamus gamblei* species complex	58
4.3	Multiple Sequence Alignment of *mat*K sequences of *C. rotang* and *C. metzianus*	62
4.4	Multiple Sequence Alignment showing unique insertions in *psbA-trnH* sequences of *C. rotang*	62
4.5	Dendrogram depicting genetic relationships based on *psbA-trnH matK* and RPB2sequences	62
	Chapter 5	
5.1	PCR amplified products of the selected barcode regions	73
5.2	Bayesian consensus tree using chloroplast (*rps16, atp*H- *atp*I, and *trn*L-*trn*F) and nuclear (RPB2 and ITS2) regions in the genus *Calamus*	74

Chapter 6

6.1	Bayesian consensus tree using RPB2 and *atp*H-*atp*I in the sub family Calamoideae	88
6.2	Chronogram based on Bayesian approach of the subfamily Calamoideae. Geological epoch is shown below the tree	89
6.3	Graphical output from Bayesian Binary MCMC analyses (exported from RASP) (B) Biogeographical regions: A (India and Sri Lanka), B (Eurasia - to Wallace's Line, including Andaman and Nicobar islands), C (Africa), D(Australia)and E (South America) (C) Colour keys to represents possible ancestral ranges at different nodes	91
6.4	Graphical output from DEC analyses (exported from RASP) (B) Biogeographical regions: A (India and Sri Lanka), B (Eurasia - to Wallace's Line, including Andaman and Nicobar islands), C (Africa), D(Australia)and E (South America) (C) Colour key to represents possible ancestral ranges at different nodes	92

APPENDICES

No.	Titles
I	List of samples collected for DNA barcoding and their locations
II	Barcode regions and Genbank accessions
III	Barcode regions of *C. metzianus* and *C. rotang* and their Genbank accessions
IV	Samples collected for phylogenetic analyses and their collected locations
V	Sequences retrieved from NCBI for study
VI	Phylogenetic markers and their Genbank accessions

ABBREVIATIONS

AIC	-	Akaike Information Criterion
BBM	-	Bayesian Binary MCMC Method
BCM	-	Best Close Match
BI	-	Bayesian Inference
BLAST	-	Basic Local Alignment Search Tool
BM	-	Best Match
CBOL	-	Consortium for the Barcode of Life
CTAB	-	Cetyl trimethyl ammonium bromide
DEC		Dispersal-Extinction-Cladogenesis
dNTPs	-	Deoxyribonucleotide triphosphate
ESS	-	Estimated Sample Size
HPD	-	Highest Posterior Density
JC	-	Jukes-Cantor
K2P	-	Kimura two Parameter
MCMC	-	Markov Chain Monte Carlo
ME	-	Minimum Evolution
ML	-	Maximum Likelihood
MP	-	Maximum Parsimony
MSA	-	Multiple Sequence Alignment
My	-	Millions of Years
Mya	-	Million Years Ago
NCBI	-	National Center for Biotechnology Information
NJ	-	Neighbor-joining
PCR	-	Polymerase Chain Reaction
PP	-	Posterior Probability
RAPD	-	Random Amplified Polymorphic DNAs
SNP	-	Single Nucleotide Polymorphism
SPSS	-	Statistical Package for Social Science
UPGMA	-	Unweighted Pair Group Method Using Arithmetic Averages

ABSTRACT

ABSTRACT

Rattans are spiny climbing palms of subfamily Calamoideae with solid stems and scaly fruits. *Calamus*, one among the 13 climbing genera constitute the largest rattan genus, with 520 species under subtribe Calaminae of subfamily Calamoideae. Even though palms have numerous observable morphological characteristics such as leaf, stem, fruit and inflorescence structures for identification, extensive studies in taxonomy were difficult due to the absence of flowers and fruits most of the year. Lack of sufficient herbarium specimens also hinders comparison of species'. Several taxonomic complexities like homoplasies look-alike species, environmental plasticity and species complexes, encountered within this group make species delimitation using morphological characters alone, complicated. In such circumstances, molecular markers have aided to resolve various taxonomic complexities in rattans and also established its potential as an important tool for rattan classification with high level of precision. With the advent of DNA barcoding techniques, an efficient molecular tool for identification at species level has become possible for all living organisms. Developing suitable barcodes for species in the genus *Calamus* will help to resolve some of the taxonomic complexities in this group of plants. In order to address the major taxonomical complexities, the pattern of relatedness between the species also need to be analysed. Understanding phylogenetic relationships and biogeographical distribution pattern is also essential in order to trace their evolutionary history and strengthen taxonomy.

DNA barcoding has been used in this study to understand the species boundaries of the 21 *Calamus* species of Western Ghats. Seven candidate barcode loci of the plastid genome (four coding regions *viz. matK*, *rbc*L, *rpo*B, *rpo*C and three intergenic spacers *viz. psb*A-*trn*H, *psb*K-*psb*I, *atp*F-*atp*H), a chloroplast region, *psb*Z-*trnf*M, two nuclear regions (ITS1 and ITS2) and two low copy nuclear gene regions (PRK and RPB2) were evaluated to identify discriminant DNA barcodes for genus *Calamus.* Different statistical methods such as distance-, tree-, and similarity- based analysis were adopted in this study. Considering all the regions studied, RPB2 nuclear barcode region appears to be an ideal DNA barcode for *Calamus* species identification as compared to other

candidate chloroplast barcode. The standard DNA barcodes could not discriminate the investigated *Calamus* species. The failure of standard barcoding regions and their combinations to delimit species- boundaries in the genus *Calamus* can be attributed to slow evolutionary rates of palm DNA. To reconcile the taxonomic complexities of the genus and palms in general, the use of low-copy nuclear gene, RPB2 which have the potential to resolve the taxonomic problems in the genus *Calamus,* is recommended from this study.

Using the recommended DNA barcode, low copy nuclear region RPB2, taxonomic complexities existing in *Calamus* species of South India were analysed. Species in the "*C. gamblei* species complex" which were morphologically similar were resolved using the RPB2 region. Based on the formation of a monophyletic clade in phylogenetic tree analysis in the molecular study as well as the absence of sufficient morphological differences between *C. dransfieldii* and *C. delesertianus*, they both are considered to be the same species. We consider the specimens treated as *C. delessertianus* from South India by Renuka (1999) to be morphological variation of *C. dransfieldii*. This study also suggest that *C. hookerianus* and *C. pseudotenuis* are closely related species based on the molecular data. Even though, there are also many morphological similarities, the merging of both can be suggested only after analysing more accessions from Karnataka and include specimens from Sri Lanka. The merging of *C. rivalis* with *C. metzianus* using morphology as well as molecular data is supported. The study also confirmed the misidentification of *C. metzianus* as *C. rotang* from the Western Ghats parts of Kerala.

Phylogenetic studies were carried out in order to determine the evolutionary process of the highly speciose genus *Calamus* with six gene regions *viz.* plastid (*rps16*, *atp*H- *atp*I, *psb*A-*trn*H and *trn*L-*trn*F) as well as nuclear regions (RPB2 and ITS2 nuclear) using Bayesian Inference. This helped to build up a better understanding of relationship between *Calamus* species in India. This study analysed the importance of both morphological as well as molecular characters for resolving the phylogenetic relationship between the *Calamus* species. A better classification of *Calamus* of Western Ghats was obtained, in which the genus *Calamus* was clustered into six different clades, which grouped based on morphological similarities.

To get a global picture on the phylogeny of subfamily Calamoideae, phylogeny of the subfamily was analysed using two gene regions- *atp*H-*atp*I and RPB2. This further aided in better understanding of the systematic position of the genus *Calamus*. Phylogenetic analysis, dating the phylogenetic tree and ancestral area reconstruction was performed using respective software. Phylogenetic tree resolved Calamoideae into 3 major clades each representing tribe Lepidocaryeae, Eugeissoneae, and Calameae. *Calamus* formed a monophyletic clade with *Daemonorops*, which was found nested within the genus *Calamus*, thereby supporting the merging of *Daemonorops* with *Calamus.* Molecular dating estimated the crown age of *Calamus* to be in Eocene. Supporting the previous studies, the origin of Calamoideae is found to be in Asia, followed by its dispersal to Africa, South America and Australia. Presence of Calameae in India is attributed to its dispersal from Asia. The biogeographical studies mainly focused on present day Indian *Calamus*, which was found to have diverged around 35 Mya (millions of years ago) in the Eocene. As per the present study they reached India through dispersal from Eurasian plate, after the collision of India with Eurasian plate in Eocene. Frequent dispersal events were evident in *Calamus* species of Southern India. Indian *Calamus* was found to be non-monophyletic in origin. Presence of early fossils of Calamoideae subfamily in South India, pointed to dispersal through the island arcs or transoceanic dispersal of the species' prior to collision followed by extinction due to climatic changes/ volcanic eruptions during early periods.

In spite of the complexities encountered in molecular studies of palms, due to slow evolutionary rate of palm DNA, an efficient DNA barcode to resolve the taxonomic complexities present in *Calamus* were identified. The phylogenetics and biogeographical studies helped in understanding the systematics of *Calamus* and dispersal of this genus into India.

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

1.1. An Overview

All biodiversity on earth is the result of evolution. On Earth, life began 4 billion years ago and it has been evolving since then. In the beginning, all living things on earth were single celled organisms, multicellular organism evolved, and later the diversity of life on earth increased day by day. The evolutionary theory of Darwin, propounded in his *"Origin of species"* in 1859 had a profound effect on its application in the classification of plant and animal groups. Systematics is the scientific study of the diversity of organisms (Mayr, 1969), including their classification, origin, evolution, diversity and distribution. This relies on morphological, genetic and evolutionary relationships among organisms and hence includes both taxonomy and phylogeny. Phylogenetics provides evolutionary relationship between diverse species by comparing species specific characters under the natural assumption that morphologically similar species are genetically closer. During earlier studies, phylogenetics only dealt with physical, or morphological features, however recent trends in phylogenetics targets both morphological as well as genetic data (DNA/protein sequence). Amalgamation of both morphological and genetic data could provide a better understanding of highly problematic taxa. Thus phylogenetic relationship relies on proper identification of species, which is further supported by molecular data. Similarly, taxonomy is one of the components of systematics for describing, naming, and the classification of organisms using morphological, behavioral, genetic, as well as biochemical observations into different species (Simpson, 1961; Mayr, 1969).

Biological species concept defines species as group of interbreeding natural populations that are reproductively isolated from other such groups (Mayr, 1942). Over years, different 'species concepts' have been established (Mayden, 1997; Wheeler and Meier, 2000), none of those could provide a completely satisfactory, definition to the "species concept". Different groups of biologists supported the different concepts (Mayden, 1997; de Queiroz, 1998; Harrison, *et al*., 1998), hence

the concept of species itself is confusing with respect to species delimitation. Even after listing twenty four species concepts by Mayden (1997), there are even more alternative definitions for species. Many of these concepts lead to different conclusions concerning numbers and species boundaries. Thus, the species concept problem is closely related to species delimitation. Organisms may appear to be alike but be different species; sometimes it may look different and yet be the same species; in some situations, many characteristics can vary within single species. Hence, determining species boundaries from observed data is often challenging.

Over the past 250 years, around 1.78 million species of plants, animals, and microorganisms have been named by taxonomists but it is estimated that more than 30 million species are yet to be identified (Singh, 2012). This intensifies the urge to address various issues pertaining to the species concept that requires major attention. However species identification using morphological characteristics alone may be quite challenging especially in the case of plants and animals showing extreme environmental plasticity, homoplasy etc.

Morphological data are limited and insufficient to distinguish two organisms in certain instances, while molecular characters provide more data for analysis and large databases of DNA sequences allow comparison across many organisms. Moreover, molecular data shows evolutionary change that can be easily modelled, usually providing an ample supply of character change typically producing data sets that can be analysed statistically. This also helps for the correct identification of species, there by resolving the taxonomic complexities prevailing in the group. Several molecular approaches had been attempted to resolve such problems in taxonomy and the technique of DNA barcoding showed promise as a supplementary tool to resolve the taxonomical complexities in several organisms.

Palms are one of the many plant groups which require taxonomical revisions owing to the fact that major phylogenetic studies have failed to resolve taxonomical complexities, mostly due to misidentification of the species. It could be attributed to the environmental plasticity which is one of the reasons why taxonomic relationships of palms still remains inadequately understood (Baker *et al*., 2011).

1.2. Palms

Palms are among the most important and characteristic components of forest ecosystems in terms of species diversity, abundance of individuals and impact on the environment (Couvreur *et al*., 2011). They belongs to an iconic family (Palmae/ Arecaeaea) of oldest monocotyledonous flowering plants having immense ecological as well as economic significance (oil palm, date palm, coconut, rattan etc.) and fortunately with rich fossil records available for study (Janssen and Bremer, 2004; Peters *et al*., 2004; Harley, 2006). They are widely distributed in the tropics and sub-tropics, principally in rain forest and in some dry habitats to a lesser extent (Dransfield *et al*., 2008). The current classification of Palmae consists of six subfamilies, includes 14 tribes and 38 subtribes (Dransfield and Uhl, 1986; Uhl and Dransfield, 1987). The initial description of species within this family was described by Dransfield and Uhl (1986), who placed 189 genera into six subfamilies and 13 tribes. Among six subfamilies, Calamoideae is the most diverse and the largest group comprising 13 genera and 600 species classified into three tribes (Lepidocaryeae, Eugeissoneae, Calameae) and eight subtribes (Uhl and Dransfield, 1987; Dransfield, 2000), concentrated solely in the old world tropics and subtropics of Asia and equatorial Africa (Dransfield, 1996). Although, Calamoideae contains tree palms and acaulescent palms, it is perhaps best known for its climbing members, rattans, which are highly diverse, particularly in South-East Asia.

1.2.1. Rattans

Rattans are spiny climbing palms of subfamily Calamoideae with solid stems and scaly fruits (Basu, 1985). The commercially important rattans belong to 13 genera (*Calamus, Calospatha, Ceratolobus, Daemonorops, Eremospatha, Korthalsia, Laccosperma, Myrialepis, Oncocalamus, Plectocomia, Plectocomiopsis, Pogonotium, Retispatha*) (Dransfield, 1992). Their distribution is exclusively restricted to tropical and subtropical regions since they are physiologically not able to undergo dormant periods and are hence susceptible to frosts (Tomlinson, 2006). Rattans are concentrated mainly in the Old World tropics. Rattans are not reported

from Madagascar. Rattan genera, *Laccosperma*, *Eremospatha* and *Oncocalamus*, are endemic to Africa. Rattans are distributed in Latin America, Mexico, Chile, Argentina, Caribbean, Africa and South-East Asian regions. A rich diversity of rattans is present in Malaysia, Indonesia, Philippines, China, Bangladesh, Sri Lanka, Myanmar and India (Dransfield *et al.*, 2008).

Even though some species are shrubby non-climbers, they are still included in rattans as their floral characters are similar to that of climbers. Sago palm, Bertam palm and tree palms such as *Raphia* (Raphia) and *Metroxylon* (Sago palm) and shrub palms such as *Salacca* (Salak) (Uhl and Dransfield, 1987), are also considered as rattans. Most rattans are lianas and climb by means of either a cirrus (an extension of the leaf rachis) or flagellum (a modified inflorescence), both of which are armed with recurved, grappling spines (Stiegel *et al*., 2011). The leaf sheaths, which are a tube-like part of the leaf, encircle the stem before the actual leaf stalk branches out. Leaves are pinnate, rarely bifid, sometimes with terminal cirrus (Renuka *et al.*, 2010), and a knee (a swelling of the leaf sheath at the junction of the petiole base) is present in most of the climbing species. In some species ochrea forms chambers, inhabited by ants, which is believed to protect species from climbers and other insects. Rattans have an efficient stem vasculature for conductive requirements (Tomlinson, 2006). The thickness of fibre cell wall, fibre percentage, distribution of tissues and cellular characteristics impart stiffness and determines mechanical behavior of rattans (Bhat and Thulasidas, 1989).

Rattans are ecologically and economically important in Asian rainforests (Gentry, 1991), due to their unique properties such as strength, lightness, durability, appearance, flexibility and economic value. They are one of the main non-wood forest products, provide livelihood potential for people living in rural areas and are widely used as a source material for handicraft as well as furniture industries (Dransfield and Manokaran, 1994; Mohan and Tandon, 1997; Sreekumar and Renuka, 2006). Over the past few decades, rattans are heavily exploited for commercial purposes. Due to indiscriminate extraction, over exploitation and

deforestation of canes from the forests, the populations of rattan species have been facing destruction in the past few decades.

1.2.1. Distribution of Rattans in India

India is one of the richest sources of rattans, comprising 60 species under four genera, *viz*. *Calamus, Daemonorops, Korthalsia* and *Plectocomia* distributed in three major phyto-geographical areas viz. Peninsular India, sub-Himalayas and Andaman and Nicobar Islands (Renuka, 2001). Genus *Salacca,* present in North eastern region of India belonging to Calameae tribe (Calamoideae), are considered as pseudorattans which belong to subtribe Salaccinae. *Calamus* and *Daemonorops* belong to subtribe Calaminae, while *Korthalsia* and *Plectocomia* belong to subtribes Korthalsiinae and Plectocomiinae respectively. All these three subtribes belong to the tribe Calameae (Dransfield *et al*., 2001).

Out of the four predominant genera of rattans found in India, *Calamus* is an important genus comprising more than fifty per cent of the total palms in India, includes commercially important species (Basu, 1985; Goel, 1992). In India, *Calamus* is distributed in all the three major phyto-geographical areas. The genus *Calamus* is widely distributed in evergreen, semi-evergreen and moist deciduous forests of the Western and the Eastern Ghats in the peninsular region (Renuka, 2001).

1.3. Genus *Calamus*

Calamus the largest rattan genus, is one among thirteen climbing genera with 520 species under subtribe Calaminae of subfamily Calamoideae (Uhl and Dransfield, 1987; Baker, 2015). The genus consists of spiny, erect or high climbing, pleonanthic rattans (multiple flowering) with solid stem found as clustering (clump-forming) or solitary (Fig.1.1). They are characterized by overlapping scales on spiny fruit which serves as pre-adaptation for climbing nature (Dransfield, 1992). Like all climbing palms they are charecterised by the means of either a cirrus or flagellum.

PLATE 1

1. ***Calamus gamblei***; A. Habit; B. Leaf sheath

2. ***Calamus nagbettai***; A. Habit; B. Young leaf

The distribution of this predominant Asian genus ranges from Indian subcontinent and south China southwards and east through Malaysia and Indonesia to Fiji, Vanuatu and tropical and subtropical parts of eastern Australia. A single species (*C. deerratus*) is found in Africa exhibiting high variability among populations (Uhl and Dransfield, 1987; Dransfield, 1992).

Pollination of the species are mainly by insects, birds, primates, other mammals, including bats and elephant (Lee *et al.*, 1995; Renuka, 1995; Bogh, 1996; Dransfield, 2001). The species identification has relied mainly on vegetative as well as floral characters. Since they are dioecious, the identification of sex is quite difficult, as they remain indistinguishable until flowering stage and the inflorescence in both sexes being similar. Therefore, it is very difficult to tackle sex-related issues in vegetative growth stages of the species without the availability of a sex-linked molecular marker. Flowering is annual and pleonanthic, produced continually and fruiting do not result in the death of the stem. Flowering of most of the Indian rattans generally starts between October-January and fruit maturation occurred during April-June.

1.3.1. *Calamus* of South India

Genus *Calamus* was proposed by Beccari (1908) who categorized species into 16 infra-generic groups. Assigning species to each group precisely was challenging due to lack of sufficient plant material for identification. Later, it was difficult to include species under any of these recognized groups, as they exhibit a mixture of defined characters of different groups. Another problem in systematics of this genus is high degree of intra-specific variability which often led to treatment of closely related species as different taxa. In this situation, phylogenetic re-grouping is required and molecular markers can act as a tool for accurate identification and classification based on molecular sequence data (Baker *et al.*, 2000a).

There are 21 species of *Calamus* reported from south India (*Calamus brandisii* Becc., *C. delessertianus* Becc., *C. dransfieldii* Renuka, *C. hookerianus* Becc., *C. shendrurunii* Anto, Renuka & Sreek., *C. neelagiricus* Renuka, *C. vimnalis* Willd.,

C. rotang L., *C. metzianus* Schltdl., *C. wightii* Griff., *C. pseudotenuis* Becc., *C. nagbettai* R.R. Fernandez & Dey, *C. karnatakensis* Renuka & Lakshmana &, *C. lakshmanae* Renuka, *C. prasinus* Lakshmana & Renuka, *C. stoloniferus* Renuka, *C. travancoricus* Bedd., *C. vattayila* Renuka, *C. thwaitesii* Becc., *C. lacciferus* Lakshmana & Renuka and *C. gamblei* Becc.) (Renuka *et al.*, 2010). Identification of *Calamus* species is mainly based on morphological characters, as well as floral characters, if available. So there exists confusion in accurate identification of species due to their indistinguishable characters such as in look-alike species, *C. lacciferous* and *C. gamblei* which showed morphological similarities. Environmental plasticity exhibited by *C. thwaitesii* possesses different stem armature and different spine arrangement in different environments, creates difficulties in their identification (Fig. 1.2).

Fig. 1.2. Morphology of *Calamus thwaitesii* in different geographical regions.

The existences of "species complexes" in the genus *Calamus* was reported by Sreekumar and Henderson (2014) and Atria *et al*. (2017). The "*C. gamblei* speciescomplex" comprising the species *C. gamblei, C. lacciferus, C. neelagiricus,*

C. prasinus, *C. dransfieldii*, *C. renukae* and *C. shendurunii* are distributed in the Western Ghats of peninsular India (Sreekumar and Henderson, 2014).

Large number of endemic rattan species are already threatened and in the verge of becoming endangered (Lakshmana, 1995; Lyngdoh *et al.*, 2005). Out of 600 reported species of rattan, the World Conservation Union (IUCN) Red List records 117 taxa as threatened to some degree. Of the 117 species, 21 are endangered, 38 are regarded as vulnerable, 28 as rare and 30 as indeterminate (Walter and Gillett, 1998). Of the reported 160 species in India, six are critically endangered and 27 are vulnerable (Renuka, 2001). *Calamus rheedii* formerly reported from south India could not be relocated after past 10 decades revealed that the species may be extinct or likely to be on the way of extinction (Lakshmana, 1995). *C. hugelianus*, earlier reported from three districts of Karnataka, becoming rare nowadays. *C. nagbettai* has been already included in Red Data Book. The species richness of *C. dransfieldii, C. karnatakensis, C. lakshmanae, C. prasinus, C. stoloniferous, C. travancoricus* and *C. vattayila* have been found decreased from the natural distribution zones. Most of the populations of *C. thwaitesii* of Western Ghats of India and Sri Lanka are fragmented as a result of continuous deforestation and overexploitation (Sreekumar and Renuka, 2006). Efforts to conserve valuable rattan resources by introducing sustainable management systems have not proved very successful, and conservation of the resource in reserves where rattan harvesting is strictly forbidden requires control and policing which have proved virtually impossible in all reserves in Southeast Asia (Dransfield, 2001). Therefore, there is an urgent need for the proper management and the effective conservation of rattan resources.

1.3.2. Species identification in the genus *Calamus*

Identification using morphological characters like inflorescence, leaf sheath, climbing organs, stem armature and fruit types played a leading role in rattan taxonomy. Even though they have numerous observable morphological characteristics for identification, in depth studies in taxonomy were found lacking (Uhl *et al*., 1995), due to their spiny nature and the absence of flowers and fruits

most of the year. Lack of sufficient herbarium specimens also found to be hurdle for species comparison (Sreekumar *et al.,* 2006). Several taxonomic complexities like homoplasies, presence of look-alike species, environmental plasticity and species complexes encountered within this group (Boer, 1968; Sreekumar and Henderson, 2014), make species identification using morphological characters alone, complicated.

A species may be referred to by many names because its distribution range encompasses a number of language groups. Often, one species can be given many names, reflecting the different uses of the plant or the various stages of development from juvenile to adult. Commonly, the blanket name, canes are given to a wide range of species (Shanley *et al*., 2002). However, it is essential to know the commercially important species and to identify them precisely. Thus a clear understanding of species boundaries is important for a stable taxonomic classification of rattan which can be further utilized for the effective conservation and sustainable management of the species. This information is critical for meaningful inventories of commercially important taxa and assessment of silvicultural potential of each species based on sound ecological knowledge (Dransfield, 2001). In this scenario, molecular tools can be introduced in order to resolve the existing taxonomic confusions in rattans, and for the stable rattan classification with high level of precision (Baker *et al.,* 2000b; Sreekumar *et al*., 2006).

1.4. DNA Barcoding

DNA barcoding is a molecular technique for the rapid, accurate and automatable species identification using standardized DNA sequences called DNA barcodes that enables cost effective species identification (Hebert *et al*., 2003; Vijayan and Tsou, 2010). In the case of plants, the Consortium for Barcode of Life (CBOL) has recommended the use of core barcode regions as *rbc*L and *mat*K as well as supplementary barcode regions (*psb*A-*trn*H, *rpo*B, *rpo*C, *psb*K-*psb*I and *atp*F-*atp*H) for species identification across the plant kingdom (CBOL, 2009). Even though with the advent of DNA barcoding many groups have undergone major taxonomical

revisions, very limited studies has been reported for the identification in palms using DNA barcoding (Jeanson *et al.,* 2011; Al-Qurainy *et al*., 2011; Yang *et al*., 2012; Ballardini *et al.*, 2013; Umapathy *et al*., 2015).

1.5. Phylogenetics in palms

Understanding phylogenetic relationships is essential to trace the evolutionary history of organisms. Though major studies on phylogeny has been reported, positioning of certain group in the tree of life are yet to be resolved. In palms, the systematic positioning of species is little understood due to unstable classifications. This requires immediate attention, as this could hamper comparative research on this plant group (Baker *et al*., 2011). All previous phylogenetic studies in palms indicated that morphological data alone are inadequate to derive phylogenetic relationships (Uhl *et al*., 1995; Baker *et al*., 1999a; Baker *et al.*, 1999b; Pintoud, 1999; Baker *et al*., 2000 a & b; Lewis *et al*., 2000; Henderson, 2002) and hence the use of molecular sequence data would generate information that would help for better understanding of the phylogeny. Molecular systematics in Palmae has gone through many barriers, due to slow evolutionary rate in DNA sequences The dearth of suitable markers has hampered molecular studies in palms, but in recent times has picked up pace by the introduction of new markers (Dransfield *et al*., 2008; Meerow, 2009; Barford, 2010; Baker and Couvreur, 2012; Roncal, 2013).

1.5.1. Phylogenetic classification of palms

Phylogenetic analyses using morphological and molecular data for Palmae supported the monophyly of Calamoideae (Uhl *et al*., 1995; Baker *et al*., 1999a; Asmussen *et al.*, 2000). But the interpretation of relationships within Calamoideae is more challenging due to the wide range of morphological diversity embraced by this subfamily. In the phylogenetic tree, endemic African rattan genera (*Laccosperma, Eremospatha* and *Oncocalamus*) forms a clade which is not closely related to predominantly Asian genera. Baker *et al.* (2000a), suggested that rattan habit has arisen on more than one occasion within Calamoideae. Despite the fact that significant morphological differences were present between African rattan

genera and Indian *Calamus*, African rattans were once considered to be a subgenera within the South East Asian genus *Calamus* (Mann and Wendland, 1864). Subsequently, they were raised to the rank of genus (Sunderland, 2012), and later grouped these three genera, into a single subtribe, Ancistrophyllinae based on molecular studies (Baker *et al*., 2000b; Dransfield *et al*., 2008). Previously, tribe Calameae, consisted of eight subtribes (Uhl and Dransfield, 1987), and later on the classification changed it to six subtribes (Calaminae, Plectocominae, Salaccinae, Korthalsinae, Metroxylinae and Pigafettinae). *Salacca* was considered in a new subtribe Salaccinae and *Korthalsia* in Korthalsinae. Similarly, tribe Lepidocaryae consist of two sub tribes (Uhl and Dransfield, 1987), and later African rattans were added to the tribe and Eugeissoneae formed a separate tribe. Subtribe Calaminae was found as two distinct clades with six genera, comprising both rattan and non-rattan genera (Baker *et al.*, 2000a). The presence of persistent tubular bracts which separated six genera of Calaminae, and later genera such as, *Calospatha*, *Ceratolobus*, *Daemonorops*, *Pogonotium* and *Retispatha* were subsumed in to a single genus *Calamus* (Dransfield *et al*., 2008; Baker, 2015). The preliminary studies on Malayan species showed that genus *Calamus* is paraphyletic (Kramadibrata, 1992), further confirmed by Baker *et al.* (2000b) using DNA sequence data. Thus the morphological heterogeneities and taxonomic complexities of some of the accepted groups, defined intuitively, suggest that the current classification may not fully reflect phylogenetic relationships among the genera. Therefore the phylogeny of the genus *Calamus* and its related genera needs further studies using molecular evidences for a stable taxonomic classification to emerge.

1.6. Biogeography

Species distribution and existence of a species in particular location, gives the information regarding their biogeographic origin as well as its evolution and dispersal derived from fossil data. It also requires the knowledge of their phylogenetic relationships and the changes in Earth's geography such as continental drift, mountain formation etc., that occurred when the organisms were evolving. Hence biogeography plays an important role in providing the evolutionary history

of a species. In order to understand biogeography it is vital to understand the continental drift.

1.6.1. Continental drift

Millions of years ago, dry land on Earth a single huge landmass, called Pangea covering nearly a third of the planet's surface in Paleozoic and early Mesozoic eras. This was surrounded by a global ocean called Panthalassa. This theory of continental drift, was first presented by Alfred Wegener in the German Geological Society in 1912 (Wegener, 1912 as cited by Demhardt, 2005). Later, approx. 180 million years ago (mya), Pangaea began to separate into two continents, southern Gondwana and northern Laurasia further divided by Tethys Sea. Gondwana land mass formed the present day South America, Africa, India, Madagascar, Australia, New Zealand, Antarctica and Laurasia comprised of today's North America, Europe and Asia, excluding Indian subcontinent.

Later on, Laurasia divided into the continents, after which it is named: Laurentia (now North America) and Eurasia (excluding Indian subcontinent). The break-up of Gondwana happened around 180 mya, during the Jurassic period. The landmass which formed today's South America and Africa split off from rest of Gondwana and began to move west. Later South America and Africa separated from each other about 140 mya. Rest of the Gondwana landmass split apart to form Antarctica, Australia, the Indian subcontinent, and island of Madagascar. Antarctica and Australia stuck together for long and separated only before about 50 million years. The plate containing Australia began to move northward and is still moving northward a few centimeters every year.

Due to this continental drift, large-scale horizontal movements of continents occurred, relative to one another and to the ocean basins during one or more episodes of geologic time and formed present world. The continental drift hypothesis will be better explained from plate tectonic movements. When a landmass breaks up, due to continental drift, geographical ranges of millions of species community are simultaneously divided (Crisci *et al*., 2009). Subsequently,

over millions of years, pairs of geographically separated taxa emerge from original species. Each pair of these sister taxa will have a divergence time from a common ancestor corresponding to the time of continental breakup (McIntyre *et al*., 2017).

1.7. Palm Biogeography

In palms, the distribution of the genus from its inferred ancestral area to its present distribution is still unclear. Several dispersal events have been happened in palms from their ancestral area, Laurasia to its present distribution in tropics where it has become a dominant element of the floras. Even though India is one of the richest sources of rattans, molecular studies reported so far regarding molecular dating as well as biogeographic origin of Indian species are limited.

1.8. Objectives

Due to morphological similarities, genus and species circumscriptions are difficult among *Calamus* and its relative genera. In order to resolve these issues, molecular tools which are independent of the environmental plasticity has to be adopted here. DNA barcoding is an efficient tool which delimit the species boundaries which uses the universal barcodes for the identification. In case of plants chloroplast regions are recommended over the other due to their evolutionary rate when compared with mitochondrial regions. Even though DNA barcodes in palms are from plastid regions, their reliability are still under doubt due to slow evolutionary rate of these regions in palms. Hence there is a need to evaluate the plastid regions and other regions here, inorder to find a suitable barcode for the identification of *Calamus* and to resolve the taxonomic complexities prevailing here, in South India.

Since a majority of the species are endemic, one is tempted to consider the species of South India to be monophyletic. This however needs further evaluation and confirmation. Biogeography also contributes in shaping the evolutionary history of species. The distribution pattern of genus *Calamus* in India thus need to be analysed. In order to find out evolutionary origin and biogeographic pattern of the genus in India, a global picture of subfamily Calamoideae is essential.

Based on the above background, the following objectives were therefore undertaken in this study:

1. **To generate DNA barcodes for *Calamus* species.**
2. **To study the phylogenetic relationship among the *Calamus* species of South India**

CHAPTER 2

REVIEW OF LITERATURE

2. REVIEW OF LITERATURE

2.1. DNA BARCODING

The use of DNA nucleotide sequence data has awakened a quickly growing interest in molecular systematic analyses which will provide accurate identification and classification of the genus. Ever since Hebert *et al*. (2003) suggested DNA barcoding for the identification of living organisms, this technique has been widely used for identification of plant species. 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; Kress *et al*., 2005; Chase *et al*., 2005).

In DNA barcoding, unique nucleotide sequence patterns of small DNA fragments (400–800 bp) are used as specific reference collections to identify specimens (Kress *et al*., 2005). The initial goal of DNA barcoding was to construct online libraries of sequence for all known species that can serve as a standard to which DNA barcodes of any identified or unidentified specimens can be matched (Ratnasingam and Hebert, 2007; Vijayan and Tsou, 2010). A DNA barcode should have important characteristic features such as universality, specificity on variation and easiness on employment. DNA barcoding has been widely applied in various fields of research such as conservation biology (Stoeckle, 2003), timber forensics and tracking (Asif and Cannon, 2005; Fuji, 2007), identification of adulterants (Dev *et al*., 2014) and in biosystematics (Gao *et al*., 2010; Pang *et al*., 2010).

2.1.1. DNA Barcodes

The mitochondrial gene *CO1* (mtDNA) has been found to be an efficient species identification tool in many animal groups (Hebert *et al*., 2003). However in plants mtDNA has low substitution rates and a rapidly changing gene content and structure, which makes this region unsuitable for barcoding (Wolfe *et al*., 1987). Consortium for the Barcode of Life Plant Working Group (CBOL) suggested the use of seven chloroplast genomic regions for species identification across plant

kingdom; four segments of coding genes (*mat*K, *rbc*L, *rpo*B, and *rpo*C1) and three non-coding spacers (*atp*F–*atp*H, *psb*A-*trn*H, and *psb*K–*psb*I).

2.1.1.1. Chloroplast Regions

Plastid coding and noncoding regions (*rbc*L, *mat*K and *psb*A- *trn*H intergenic spacer) are the commonly used regions for barcoding. *rbc*L encodes large subunit of rubilose-1, 5-bisphosphate carboxylase/oxygenase (RUBISCO), a critical photosynthetic enzyme (Zurawski *et al*., 1981). Even though CBOL suggested *rbc*L as a potential plant barcode, it is not efficient in species discrimination and most investigation groups are of the opinion that this can only be used in combination with other markers (Chase *et al*., 2007; Hollingsworth *et al*., 2009). *mat*K codes for maturase kinase involved in splicing of type-II introns from RNA transcripts (Neuhaus and Link, 1987). Considering its rapid evolutionary rate, it is recommended as a single barcode or in a combination barcode by some groups (Ford *et al*., 2009; Starr *et al*., 2009). Failure of PCR amplification was however shown in some species which is a limitation of *mat*K region (Wolfe *et al*., 1987). Because of high species discrimination, universality and ease of PCR amplification of *rbc*L and *mat*K, their combination was recommended by CBOL as standard two-locus barcode for plants.

Similarly, other chloroplast regions (*rpo*B, *rpo*C, *acc*D, *ycF5*, *ndh*J) are also suggested by certain working groups. *rpo*B, *rpo*C1, *rpo*C2 encodes RNA polymerase subunits and hence widely used in phylogenetic studies because of its high substitution rate (Tsumara *et al*., 1996; Guisinger *et al*., 2008). Even though they have high universality and yield high quality sequence, they were eliminated as single barcodes, because of their low species discrimination ability by CBOL-Plant Working Group. The *acc*D gene encoding β-carboxyl transferase subunit of acetyl-CoA carboxylase is present in all flowering plants, but is completely absent in grasses (Hajdukiewicz *et al*., 1997). This evolved five times faster than *rbc*L in plant taxa (Yasui and Ohnishi, 1998), and was proposed as barcode by Ford (2009). Royal Botanical Gardens, Kew proposed the short, single copy *Ycf5*, encoding a protein containing 313 amino acids (Yasui and Ohnishi, 1998), as a barcode region.

Although *Ycf5* is conserved across all land plants, this is hardly used due to poor universality and problems in aligning sequences (Wakasugi *et al*., 2001; Ford *et al*., 2009). NADH dehydrogenase is coded by *ndh*J (Nakazono and Hirai, 1993), which is tested for barcoding studies (Yasui and Ohnishi, 1998; Sass *et al*., 2007; Lahaye *et al*., 2008; Ford *et al*., 2009). The *ndh* gene found as a gene complex consisting of 11 genes, *ndh*A–*ndh*K, codes for subunits of a functional respiratory protein complex (Burrows *et al*., 1998). *ndh* is absent in certain gymnosperms *viz* Pinus (Wakasugi *et al*., 1994), Cuscuta (Haberhausen and Zetsche, 1994) and Gnetales, but present in other gymnosperms (Braukmann *et al.*, 2009); difficulties were reported in amplification of *ndh*J in orchids. Although Ford *et al*. (2009) suggested this as supplementary barcode due to its high species discriminatory ability in liverworts, pteridophytes and monocotyledons, this was not accepted by CBOL perhaps they were absent in some groups like conifers. Anyhow Lahaye *et al*. (2008) observed low discriminatory power for *ndhJ*.

Intergenic spacers *psb*A-*trn*H, *atp*F- *atp*H, *psb*K-*psb*I are also used as plant barcodes. Because of the high species discriminatory power, *psb*A-*trn*H was widely in plant DNA barcoding (Kress *et al*., 2005). Genome size of *psb*A-*trn*H varies from ~300 bp up to 1000 bp, which can sometimes lead to problems in bidirectional sequencing (CBOL, 2009; Chase *et al*., 2007). Combination of this region along with *rbc*L and with *mat*K were also reported (Chase *et al*., 2007; Newmaster *et al*., 2008; Erickson *et al*., 2008; van de Wiel *et al*., 2009). *atp*F- *atp*H encoding ATP subunits CFO I and CFO III respectively (Drager and Hallick, 1993), amplify easily but sequences alignment was difficult due to variation in length (Lahaye *et al*., 2008), and hence CBOL recommended this only as a supplementary loci. *psb*K and *psb*I genes which encodes for two low molecular mass polypeptides, K and I respectively of photosystem II (Meng *et al*., 1991). Unlike *atp*F-*atp*H, psbK-psbI showed easy sequence alignment and PCR amplification (Lahaye *et al*., 2008). However, due to inconsistency in getting unambiguous bidirectional sequences, it was only considered as supplementary barcode by CBOL, in spite of its high species discriminatory power (Lahaye *et al*., 2008). Even though *trn*L (UAA) – *trn*F (GAA) locus was not recommended by CBOL, this was suggested by Taberlet *et al*.

(2007) as a barcode. The secondary structure with alternation of conserved region, variable regions and usage in highly degraded tissues, was found as an advantage of *trn*L (UAA)–*trn*F (GAA) (Quandt *et al*., 2004; Taberlet *et al.*, 1991 & 2007).

As plastid DNA shows less variation, more than one region is selected to be barcodes in a tiered approach. Chase *et al*. (2007) proposed *rpo*C1, *rpo*B and *mat*K as well as *rpo*C1, *mat*K and *psb*A-*trn*H as viable markers for land plant barcoding. Instead of simply adding another locus to increase the number of region, a tiered method of using a first tier coding region common across the land plants and a second tier of non-coding regions provide resolution to the species level was also suggested (Newmaster *et al*., 2006). Multi-gene approaches are now common in many plant species.

2.1.1.2. Nuclear Region

Nuclear rDNA cistron is a multigene family encoding the nucleic acid core of ribosome, which evolved in a concerted manner such that higher level of overall sequence homogeneity exists among copies of rDNA within species that shows differences between species, giving it characters of an ideal barcode (Chase *et al.*, 2007). The internal transcribed spacer (ITS) has a greater discriminatory power over plastid regions at lower taxonomic levels. ITS1 and ITS2 regions can be separately amplified by anchoring primers in conserved region, facilitating easy amplification of this region from poor quality DNA (Kress *et al*., 2005). Recent hybridization, lineage sorting, recombination among copies, high mutation rate and pseudogene formation of cistrons are considered to be the reasons for intra individual variations in ITS (Rogers and Bendich, 1987; Alvarez and Wendel, 2003; Bailey *et al.*, 2003). nrITS was proposed as a potential candidate for barcoding in plants based on the presence of multiple copies in cells, availability of universal primers, high universality and good species discriminatory power. This was successfully used for the identification of several medicinal plants and their close relatives (Chen *et al.*, 2010).

2.1.2. Future of DNA Barcoding

Although universal DNA barcodes are not 100 percent efficient in species identifications across the tree of life, researchers have adopted DNA barcoding as tool in their respective fields (Joly *et al.,* 2014) and introduced many new barcode regions (Valentini *et al*., 2009; Ballardini *et al*., 2013; Techen *et al*., 2014; Joly *et al*., 2014; Paracchini *et al*., 2017).

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. Failure of plastid data to highlight species boundaries was reported in willows, wild potatoes, carex, bryophytes (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). As low-copy nuclear genes evolve up to five times faster than plastid genome and are found in rapidly diversifying lineages (Sang, 2002; Small *et al*., 2004; Norup *et al*., 2006), they are now being explored as additional barcodes to discriminate species. Low copy nuclear regions were also reported as DNA barcodes in *Clermontia* (Campanulaceae) and *Cyrtandra* (Gesneriaceae), where the plastid regions had a slow evolutionary rate (Pillon *et al*., 2013).

The use of complete plastid genome sequences as DNA barcodes has been suggested by Nock *et al.* (2011). Decreasing cost and increasing power of next generation sequencing technologies make complete plastome sequencing relatively easier. Plastid haplotypes frequently do not completely track species delimitations, so completely sequencing this genome will unambiguously help in barcoding. Next generation sequencing technologies will ultimately get to the point where obtaining vast amounts of sequence data from many individuals is feasible, which may tackle the problems of single copy genes by permitting fullest possible description of species boundaries using genetic data.

The availability of standardized universal primers makes shorter DNA sequences-mini-barcodes, a more suitable option for DNA barcode analysis through inexpensive and comprehensive large-scale species identification (Hajibabaei and McKenna, 2012). While for closely related species, whole-plastid-based barcodes, super barcodes offer great potential in species discrimination (Li *et al*., 2015). Continuing advances in sequencing technologies, may uphold whole genome sequencing and super barcodes as the most preferred choice for plant identification in near future.

2.1.3. Barcoding in Palms

Taxonomy of palms has improved in the past decades (Dransfield *et al.,* 2008), yet problems such as several species complexes remain, which is difficult to tackle by depending only on morphological features (eg. in *Calamus*) (Henderson and Martins, 2002; Henderson, 2011). A clear understanding of species delimitation is important for the stable taxonomic classification of rattans, the major group of palms, which can be utilised for their conservation and sustainable development. Lack of taxonomic knowledge is mainly due to difficulties in collection. However, it is essential to know the commercially important species of commercially important taxa (Dransfield, 2001). Molecular techniques were employed to sort existing taxonomic problems in rattans and can be considered as an important tool for rattan classification (Baker *et al.,* 2000b; Sreekumar *et al*., 2006).

DNA barcoding has been reported as an efficient supplementary tool for species identification in palms. The successful species identification using three markers – *mat*K, *rbc*L and nrITS2 in their combination was reported in tribe Caryoteae (subfamily Coryphoideae) (Jeanson *et al.,* 2011). Taxonomic confusions in their identification are existing in the genus *Phoenix* and using *psb*A-*trn*H as the DNA barcode, cultivars of *Phoenix dactylifera* had been successfully identified (Al-Qurainy *et al*., 2011). Later, *psb*Z-*trnf*M (CAU) region with a minisatellite was considered as an appropriate DNA barcode for species identification in Phoenix in order to resolve the hybridization issues (Ballardini *et al.,* 2013).

Even though *psb*A-*trn*H was suggested to be a good DNA barcode in *Calamus* by testing the Chinese specimens (Yang *et al*., 2012), due to the presence of indels (Kress *et al*., 2005) and high amount of intra specific variation, leading to difficulties in sequence alignment, indicates its inefficiency as a good DNA barcode. The inefficiency of single nuclear genic region for species discrimination was reported in *Daemonorops* (Umapathy *et al*., 2015). Even though barcoding studies using plastid regions has been reported in palms, the reliability of these barcode regions in rattans needs to be ascertained due to its slow evolutionary rate.

Plastid DNA seems to evolve slowly in palms than in other monocots (Gaut *et al*., 1992; 1996; Asmussen, 1999 a & b), where as in others, the rate of evolution seems unexpectedly high (Bayton, 2005). This slow rate in substitution rate makes hindrance to molecular studies to derive inter/intra specific relations (Jeanson *et al.*, 2011). Obstacles to the use of DNA sequences in studies of palm evolutionary biology have been attributed to low rates of substitution in palm chloroplast genomes (Gaut *et al*., 1992 & 1996), and incomplete concerted evolution resulting in intragenomic heterogeneity in nrDNA internal transcribed spacer sequence data (Baker *et al*., 2000a). Hence DNA barcoding in palms is expected to be offer a challenge.

2.2. PHYLOGENY

Phylogenetics is the branch of science that deals with the studies of evolutionary relationship among organisms (e.g. species, populations). The main objective of phylogeny is to construct the evolutionary history of a group of organisms based on either anatomy, morphology, cytology, ecology, biochemistry, genetics, distribution or molecular biology. The inference of phylogenies with computational methods has many important applications in medical and biological research, such as drug discovery and conservation biology. Phylogenetic trees have already witnessed applications in numerous practical domains, such as in conservation biology (Baker and Palumbi, 1994) (illegal whale hunting), epidemiology (Bush *et al*., 1999) (predictive evolution), forensics (Ou *et al*., 1992) (dental practice HIV transmission), gene function prediction (Chang and Donoghue, 2000) and drug

development (Halbur *et al.*, 1994). Other applications of phylogenies include multiple sequence alignment (Notredame *et al*., 2000; Edgar, 2004), protein structure prediction (Shindyalov *et al*., 1994), gene and protein function prediction (Eisen, 2003; La *et al.*, 2005), drug design (Searls, 2003) and commercial drug discovery (Bader *et al.*, 2001).

A phylogenetic tree depicts the evolutionary relationship of organisms. There are two main types of trees that can be found, rooted trees which have a single node from which all nodes are derived and unrooted trees, those that do not originate from one clear node. Phylogenetic tree construction methods are of two types, distance and character based. Unweighted pair group method using arithmetic averages (UPGMA) (Sneath and Sokal, 1973), Neighbor Joining (NJ) (Saitou and Nei, 1987), minimum evolution (ME method) (Sneath and Sokal, 1973) and Fitch and Margoliash (Fitch and Margoliash, 1967), are the commonly used methods based on distance matrix. The alternative to these methods are the character based methods such as maximum parsimony (MP) (Fitch, 1971), and maximum likelihood (ML) (Felsenstein, 1981), and Bayesian Inference (BI) methods which take a probabilistic approach to tree construction. Even though BI uses likelihood function like ML, BI is considered to be superior over others because it uses posterior probability, stochastic search algorithm (MCMC) and explicit models for sequence evolution.

2.2.1. Molecular Phylogeny

The advancement of DNA sequencing technologies has elevated phylogenetic analysis to utmost heights. Molecular data are an important source of information on phylogenetic relationships, since they are independent on morphological characters and environmental plasticity (Collin, 2003). Molecular data have had a profound impact on the field of plant systematics, and the application of DNA-sequence data to phylogenetic problems is now routine. In combination with explicit methods for phylogenetic analysis, molecular data have reshaped concepts of relationships and circumscriptions at all levels of taxonomic hierarchy (Qiu *et al.*, 1999; Soltis *et al*., 1998; Crawford, 2000).

Genomic DNA encompasses coding (conserved) as well as non-coding regions. Non coding regions are preferred in mutated sequence continues divergence till the mutated family gets a distinct identity. These non-coding regions, especially ribosomal DNA sequences hold great promise for phylogeny reconstruction and evolutionary directions. Usually three genomic regions (nuclear DNA, chloroplast DNA and mitochondrial DNA) are exploited by plant systematists to derive evolutionary relationships in plants (Caputo, 1997).

2.2.1.1. Mitochondrial DNA

Mitochondrial DNA (mt DNA) data has been used for resolving species-level phylogenies in animals because of its uniparental origin. They changed relatively fast in evolution and are useful at the short end of the evolutionary time scale to resolve changes that occurred less than 15 million years ago such as the separation of genera and species. More often, it is not the gene content of the mitochondria that is used, but the order of genes round the chromosomal ring (Brown *et al*., 1982). In plants, mt DNA are slowly evolving, with very low substitution rates and hence not found very suitable for phylogeny (Palmer, 1992). Therefore, the search for a plant molecular marker shifted to chloroplast and nuclear genomes with high substitution rates.

2.2.1.2. Chloroplast DNA

In early 90s, DNA based phylogenetic studies in plants were based on genes and spacers from plastid genome (Clegg, 1993; Olmstead and Palmer, 1994; Soltis *et al*., 1998). Chloroplast DNA became an important marker in tracing evolutionary pattern at taxonomic level, due to its high conservation, slow rate of sequence evolution, relatively short and constant size (Palmer, 1985). This region is generally used in systematic studies because they are structurally stable, haploid, non-recombinant and generally uniparentally inherited. Evolutionary processes, such as gene duplication and deletion, concerted evolution and pseudogene formation are not common in chloroplast genomes unlike in nuclear genes, which can dangerously distort the evolutionary history of DNA sequences relative to that of organisms

(Palmer, 1985). Chloroplast DNA molecules are highly conserved especially in closely related species (Olmstead and Palmer, 1994). Nucleotide substitutions of chloroplast DNA are low in land plants (Wolfe *et al*., 1987), which makes this suitable for studies in distantly related taxa. As chloroplast DNA have a medium rate of evolution when compared with that of the nuclear and mitochondrial regions, they (*rbc*L, *mat*K, *atp*B and *ndh*F) are used mainly at family level (Chase *et al*., 1993; Soltis *et al*., 1998). But non-coding regions (*psb*A-*trn*H, *trn*L-*trn*F) and introns (*rps*16, *rpL*16, *tr*nL, *trn*K and *rpo*C) are frequently used at lower taxonomic levels (Taberlet *et al*., 1991).

2.2.1.3. Nuclear DNA

Plant systematic community incorporated nuclear markers for phylogenetic reconstruction, recognizing the problems of using uniparentally inherited plastid and mitochondrial gene sequence for phylogeny (Rieseberg and Soltis, 1991). Nuclear DNA gained importance because of its high copy number and was well conserved enabling the use of same probes universally for amplification. In eukaryotes, rDNA, nuclear DNA coding for ribosomal RNA, have been classified into four major types (5SrDNA, 5.8srDNA, 18srDNA, and 28srDNA) and are highly conserved (Cullis *et al*., 1988). 18S, 5.8S, and 28S are separated by highly variable intergenic spacer regions (ITS1 and ITS2); each transcriptional unit is separated by intergenic spacer (IGS) (Mishler *et al*., 1994) and provide valuable data in phylogentic inferences. The genes coding for rRNA, are highly conserved because of their important structural role and mutations are unlikely to survive natural selection. It can therefore provide evidence about changes that occurred very early in animal evolution, such as the separation of classes within a phylum or even the origin of new phyla and they are also important for the phylogenetic analysis of distantly related species (Moritz *et al*., 1987).

2.2.1.4. Low copy nuclear regions

In previous instances, where plastid and nuclear barcode regions failed, low-copy nuclear regions have been successfully used in resolving phylogenetic relationships

especially at lower taxonomic levels, regardless of difficulties like paralogy, concerted evolution and intragenic polymorphism (Lewis and Doyle, 2001 & 2002; Carl and Doyle, 2001; Gunn, 2004; Roncal *et al.,* 2005; Bayton, 2005; Thomas *et al*., 2006; Loo *et al*., 2006; Norup *et al*., 2006). As these regions can evolve up to five times faster than plastid genome, it can resolve even recently and rapidly diversifying lineages (Sang, 2002; Small *et al*., 2004; Norup *et al*., 2006). The potential to accumulate datasets from multiple unlinked loci, independently evolving, and with biparental inheritance is of great advantage when adopting low copy nuclear regions (Randall *et al*., 2004).

2.2.2. Palm Phylogeny

Palm family (Palmae/ Arecaceae) comprises of approximately 2000 species under 191 genera, widespread across tropical and subtropical regions (Baker *et al*., 1999a; Govaerts and Dransfield, 2005). Palmae is classified as consists of six subfamilies, 14 tribes and 38 subtribes (Dransfield and Uhl, 1986; Uhl and Dransfield, 1987; Uhl and Dransfield, 1999). Among subfamilies, Calamoideae is the largest group comprises of 22 genera and 650 species classified into two tribes and eight subtribes (Uhl and Dransfield, 1987). Palms coming under commelinid clade of monocotyledons are monophyletic in origin (Asmussen *et al*., 2006; Chase *et al*., 2006). The initial description of species within this family was described by Dransfield and Uhl (1986), who placed 189 genera into six subfamilies and 13 tribes.

Phylogenetic analyses supported the monophyly of Calamoideae (Uhl *et al*., 1995; Baker *et al*., 1999a; Asmussen *et al.,* 2000). African rattans form a monophyletic group, as do all Asian Calamoideae except for Eugeissona. Calaminae is not monophyletic, six rattan genera forming a clade which is sister to a monophyletic Plectocomiinae. Eleiodoxa and *Salacca* form a clade in a grade of other genera with catkin-like rachillae, including the Metroxylinae, which appears not to be monophyletic and Pigafetta (Baker *et al.,* 2000a).

Phylogenetic analysis showed that genus *Calamus* had a paraphyletic origin; *Calamus* and four other genera, *Daemonorops, Retispatha, Ceratolobus* and *Pogonotium* formed a monophyletic group (Baker *et al*., 2000b). The main characters which separated six genera of Calaminae, are the presence of persistent tubular bracts, and later, the genus *Calospatha, Ceratolobus, Daemonorops, Pogonotium* and *Retispatha* were subsumed in *Calamus* (Dransfield *et al*., 2008; Baker, 2015).

2.2.2.1. Molecular Phylogenetic Studies in Palms

Phylogenetic relationship among palms were poorly understood (Dransfield, 2008), and this found to be an obstacle for comparative research on this important plant groups (Baker *et al*., 2011). Even though, palms have numerous observable characteristics such as leaf, stem, fruit and inflorescence structures, in-depth studies on morphological and anatomical features were found to be difficult. Homoplasy for some of the morphological characters often creates difficulties in resolving the relationship between families (Uhl *et al*., 1995). Even though, phylogenetic studies has been reported in palms (Pintoud, 1999; Henderson, 1999), these studies are insufficient in their interpretation, since morphological data alone were in adequate to resolve the phylogenetic relationships (Uhl *et al*., 1995; Baker *et al*., 1999b; Baker *et al*., 2000c; Lewis *et al*., 2000) and molecular data had to be introduced into palm phylogeny. Molecular phylogeny in palms has also gone through many barriers due to the slow rate of evolution in the DNA. Chloroplast DNA was observed to have 5 to 13 fold slow rate of evolution in palms (Wilson *et al*., 1990). Genes, *rbc*L has 5 times and *Adh* has 2.5 times slower evolutionary rate in palms compared to grasses (Gaut *et al*., 1992 & 1996). Later mitochondrial gene region was also found to have slower evolutionary rate (Eyre-Walker and Gaut, 1997). Baker *et al*. (2000a) analysed the phylogenetic relationship of 22 genera of subfamily Calamoideae using *rps16* intron, and his studies supported the early findings regarding the slower evolutionary rates of chloroplast regions. Substitution rates in *rbc*L in monocotyledons are inversely correlated with minimum generation time. The similar slow evolutionary rate in three regions- chloroplast, mitochondrial

and nuclear, indicates a correlation in synonymous sites evolution across genus (Eyre -Walker and Gaut, 1997). A single evolutionary process in pattern of palm diversification might be causing the slow rate of molecular evolution in palm DNA that affects all three genomes (Eyre -Walker and Gaut, 1997; Gaut *et al*., 1992). Obtaining well resolved phylogenetic trees at species level in Palmae challenging due to notoriously slow rate of molecular evolution, convergent evolution among some morphological characters, difficulties in outgroup choice and character polarisation (Hahn, 2002b). These factors remain a hindrance for using the well-known DNA regions in palm phylogeny, making their selection for phylogeny a difficult task.

Molecular phylogenetic studies in palms were very few in numbers in past (Baker and Cauvreur, 2012; Dransfield *et al.,* 2008), due to limitation of availability of, but such studies have been increasing in recent times (Meerow, 2009; Barford, 2010; Roncal, 2013). In the older days, Restriction Fragment Length Polymorphism (RFLP) data have been used for phylogenetic reconstruction in Palmae (Uhl *et al*., 1995). Later on, nuclear and plastid regions, such as *trn*L-*trn*F, *rps16* intron, *atp*B, *rbc*L, 5 S nr DNA, 5 S nr spacer have been included in palm phylogenetics (Barrow, 1998; Asmussen, 1999b; Baker *et al*., 2000 a & b; Asmussen and Chase, 2001; Hahn, 20002a).

Plastid regions of palms are conserved and more data are needed to construct phylogeny using plastids. In spite of its low evolutionary rate, plastids have been successfully used to resolve relationship at lower taxonomic levels (Asmussen *et al.,* 2000; Cuenca and Asmussen-Lange, 2007; Baker *et al.,* 2009). Monophyletic origin of Palmae is well supported by phylogenetic studies using chloroplast regions (Chase, 1993 & 1995). Relationship based on morphological and molecular data suggested that Coryphoideae, Calamoideae and Phytelephantoideae are monophyletic (Uhl *et al*., 1995). Asmussen (1999b) reported the monophyly of tribe *Genomeae* using *rps16* region as well as *rbc*L, *rps* 16 and *trn*L-*trn*F inter genic spacers (Asmussen and Chase, 2001). Chloroplast DNA introns and *rps16* data was

used to report Calyptrogyne, Calyptronome, Pholidostachys and Welfig, constituted a monophyletic group (Asmussen *et al.*, 1999a).

Faye *et al*. (2014) selected *atp*H-*atp*I, *psb*A-*trn*H, *psb*Z-*trn*fM and *rps*3-*rp*L from hundred primers for the phylogenetic reconstruction in monocots and palms. This study supported monophyly of subtribe Ancestrophyllinae, however the relation among genera, Laccospermae, Ermospathae, Oncocalamoideae remain unresolved. The *rps16* intron showed low levels of variation and this substantiated previous studies, where the conserved nature in chloroplast genome of palms (Baker *et al.*, 2000a) was found to be suitable for phylogenetic studies at family level rather than species level (Wilson, 1990; Gaut, 1996). *rps16* along with *trn*L-*trn*F reported to show a well resolved phylogeny in palms, led to change the systematic position of palm species, *Iriartea*, tribe Caryoteae and major clade comprising four subfamilies, Coryphoideae, Ceroxyloideae, Arecoideae and Phytelephantoideae (Asmussen *et al*., 2000). Even though Asmussen *et al*. (2000) reported poor performance of *trn*L-*trn*F region for phylogenetic utility among species or closely related genera in Palmae due to its conserved nature in palms, this spacer region revealed better relationship among higher taxonomic groups. Domenech *et al*. (2014) reported poor utility of plastid sequences in phylogenetic studies of Archontophoenicinae. Even at infrageneric level, very few studies have been reported using chloroplast regions to resolve phylogenetic relationship (Cuenca and Asmussen-Lange, 2007), most of the studies have reported the utility of plastid markers in combination with nuclear markers to resolve phylogenetic relationships (Dransfield *et al*., 2008; Roncal, 2013; Lewis and Doyle, 2001). Hence nuclear regions provide a good source of phylogenetic information at low taxonomic levels since they appear to evolve more rapidly than those from the plastid genome (Wolfe *et al.*, 1989; Eyre-Walker and Gaut, 1997; Gaut *et al*., 1998; Doyle and Doyle, 1999; Lewis *et al.*, 2000).

The nuclear internal transcribed spacer (nrITS) region shows higher level of polymorphism in Calamoideae due to concerted evolution which leads to higher level of homoplasy, suggesting that phylogenies derived from ITS data may not be reliable (Baker *et al*., 2000a). High levels of within-individual polymorphism were

identified in the ITS region, indicating that concerted evolution is not effectively homogenizing the ITS repeats. In the majority of cases, multiple clones from individuals resolved as monophyletic. However, when ITS combined with *rps16* intron, it supported well in phylogeny (Baker *et al.*, 2000). However these regions appear to have limited value in palm phylogenetics because of the lack of homogeneity among repeats within individual palm genomics (Baker *et al*., 2000a & b). 5SDNA was first used in palms by Barrow (1996) for the phylogenetic studies in Phoenix. The studies of Baker *et al.* (2000) showed the effect of this region in *Calamus* and related genera, indicating that this is a divergent region useful for discriminating lower taxonomic levels. With this region, the highly speciose genus *Calamus* was found to be non-monophyletic with all five remaining genera (*Calospatha, Retispatha, Daemonorops, Ceratolobus, and Pogonotium*) being embedded within it. The existence of intragenome polymorphism did not excessively interfere with phylogenetic reconstruction because, in majority of the studies, multiple clones obtained from individual species were resolved as monophyletic groups. This study indicated that 5S non-transcribed spacer is too divergent to be alignable across all taxa, although it has proved useful for lower taxonomic levels in subfamily (Baker *et al*., 2000b).

Since chloroplast and nuclear DNA have limitations in certain cases, low copy number nuclear genes provided robust evidence for relationships at intermediate and lower taxonomic level, which was failed to reveal by more slowly evolving plastid regions. At lower taxonomic levels, the rate of fast evolution of these genes can be advantageous. It is difficult to design PCR primers for poorly characterized low copy nuclear genes, amplification of this can be more problematic than other regions. The distinction of orthologous loci, complications related to concerted evolution or recombination among paralogous loci and intragenic polymorphism are challenges that must be overcome with low copy nuclear gene (Small *et al*., 2004). Despite these obstacles, low copy nuclear genes represent the largest source of molecular phylogenetic data in plants. Low copy nuclear genes like Malate Synthase (MS) are effective sources of data in palms, hold more phylogenetic information (Mort and Crawford, 2004). Lack of resolution among major lineages

within family using chloroplast and MS region may be due to rapid evolutionary radiation, without sufficient time for mutations to accumulate along internal branches (Lewis and Doyle, 2001). The phylogenetic utility of RPB2 was explored in most of the angiosperms (Denton *et al*., 1998; Oxelmann *et al*., 2004; Pfeil *et al*., 2004). Denton e*t al*. (1998) indicated the presence of single RPB2 orthologue, thereafter, Oxelmann *et al*. (2004) showed the presence of more than one RPB2 sequences in Zania, Nymphaeae, Magnolia, Maize and two asteroid groups. 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), have proven to be particularly useful as phylogenetic markers in palmae. PRK, gene encoding Calvin cycle enzyme phosphoribulokinase has been shown to be sufficiently variable to examine low level relationships within palms (Gunn *et al*., 2004; Lewis and Doyle, 2002; Lewis and Martinez, 2000; Roncal, 2005; Thomas *et al*., 2006). Non-coding intron RPB2 was also informative in exploring low level relationship in Arecoideae (Roncal, 2005). RPB2 is the low copy nuclear DNA from the gene encoding the second largest subunit of RNA polymerase II. PRK and RPB2 are highly informative and produce well resolved phylogeny in Arecoid palms (Baker *et al.,* 2011). Roncal *et al*. (2005) developed palm specific primers for intron 23 of RPB2 for phylogenetic studies of palms. Monophyly of subtribe Ptychospermatinae was supported with the help of RPB2 and PRK (Zona *et al*., 2011). RPB2 supported the resurrection and expanded circumscription of genus *Ponapea*, non monophyly of genera such as *Drymophlew*, *Ponapea* and *Velicha*. Their phylogenetic hypotheses suggested that Ptychospermeae diverged into six major clades with repeated radiation into Australia and Western Pacific (Zona *et al*., 2011). Molecular phylogenetic analysis of Arecaeae (Asmussen and Chase, 2001; Lewis and Doyle, 2002), confirmed the placement of Genomeae in subfamily Arecoideae supporting the previous classification (Uhl and Dransfield, 1987). The studies placed Genomeae within a clade that included the Arecoideae tribes such as Arecaeae, Cococeae, Podocococeae, Ceroxyloideae, Iriataceae and Hyphorbeae (Asmussen and Chase, 2001; Lewis and Doyle, 2002). Based on the literature, it is

revealed that the combination of nuclear, chloroplast and low copy nuclear regions will provide better resolution in phylogenetic analysis.

2.3. BIOGEOGRAPHY

2.3.1. Biogeography of palms

The known history of palms extends far back into Cretaceous period (Dransfield *et al*., 2008). Presence of unambiguous fossils in association with palms during the Turonian, revealed the existence of this family during Cretaceous period (89- 93.5 mya) (Kvacek and Herman, 2004). Molecular clock estimation of angiosperm phylogenies suggested the stem age for palm family to be ranging from 91 to 120 mya (Bremer, 2000). Hypotheses regarding palm origin in Gondwana (Moore, 1973; Baker and Dransfield, 2000) or Pangaea (Uhl and Dransfield, 1987), were rejected after the finding of Baker and Couvrier (2012), indicating the origin of palms in Lauraisa. They supported the previous study of Uhl and Dransfield suggesting that, earliest divergences among extant lineages of palms occurred in Laurasia (Uhl and Dransfield, 1987). The family members migrated out of Laurasia into Africa and South America, before the end of Cretaceous and thereafter, in Cenozoic extended their distributions to most of the tropical and subtropical regions of the world. Palms have achieved a pan-tropical/subtropical distribution, in spite of the fact that they are known to have relatively limited dispersal (Kissling *et al*., 2012). Calamoideae diverged from other palms at crown node of the family in Eurasia, expanding into Africa prior to its crown node divergence (80 mya). Later, in Campanian (77 mya), tribes Calameae and Lepidocaryeae diverged from each other, while former in Eurasia and the latter in Africa (Baker and Couvreur, 2012).

Fossil evidences revealed the existence of richer palm flora in Africa and India in past than at present. The low diversity of palms in Africa in contrast to Asia and America is due to Neogene aridification in Africa (Morley, 2000). Evidences of palm pollen were found from Africa during Campanian (83.5–70.6 mya) and

became abundant and more diverse during Maastrichtian, known as 'Late Cretaceous Palm Province' (70–65.5 mya) (Morley, 2000). Calamoideae has been recorded from the Maastrichtian in India (Herngreen *et al*., 1996).

2.3.2. Fossil evidence of palms in India

Earliest evidence of palm fossil in India was of a costa palmate palm leaves reported from Deccan Inter trappean sediments (Maastrichtian–Danian) and was considered as the oldest fossil record of costapalmate palm leaves from India and the Gondwana- derived continents (Srivastava *et al.,* 2014). The Maastrichtian–Danian sediments opens a new dispersal route other than dispersal event occurred from Indian Ocean into India (including Sri Lanka) during Miocene as suggested by Baker and Couvreur (2012). Subsequently, two palm leaf impressions described from Makum Coalfield, Assam had Late Oligocene age (28–23 mya) (Srivastava *et al.,* 2012). Fossil evidence for Eugeissona reported in the pre-collision Cenozoic of India (Morley, 1998). The fossil reports of palm stem woods from Deccan Intertrappean Beds and Lalgarh Formation, showed their affinities with Calamoid palms in India reported had Pliocene-Plastocene origin (Rao and Achutan, 1969; Bera and Banerjee, 1990). Fossil evidences suggested the presence of Nypa in early Pleocene in Meghalaya (Mehrothra, 2000). Well along, the records of the genus *Livistona* (Palmoxylon livistonoides) from Maastrichtian- Danian of Intertrappean beds as well as Eocene-Oligocene of western Himalaya and Miocene sediments of Australia were reported. Later, molecular dating confirmed *Livistona* to be a native of India which migrated to Australia through Asia after establishment of land connections during Neogene which explains the dispersal of palms in Australia (Crisp *et al.,* 2010).

2.3.3. Formation of India

About 200 mya, peninsular India was part of Gondwanan supercontinent (Mani, 1974; Briggs, 1989 & 2003; Hedges, 2003). From there, Indo-Madgascar plate

drifted away from East Africa, during Jurassic period, approximately 158-160 mya, followed by separation of peninsular India from Madagascar around 84-96 mya. The collision of Indian land mass with Eurasia resulted in the upliftment of Himalayan mountain range (Briggs, 2003), which acts as barrier for dispersal of species into and out of the Indian subcontinent. The Indian plate allegedly underwent a period of isolation for about 30- 40 million years (mya) after the separation from Madagascar, before colliding with the Eurasian plate around 40-50 mya. The fauna and flora of peninsular India is thus expected to be unique and highly endemic due to this long period of isolation (Keast, 1971).

However, a few studies also suggested that Indian landmass was not completely isolated during its northward journey based on the evidence of faunal links present (Briggs, 2003; Chatterjee and Scotese, 1999). Peninsular India was acting as a "biotic ferry", according to which the rafting Indian plate carried ancient Gondwanan forms to Asia (Mani, 1974; McKenna, 1995; Hedges, 2003; Briggs, 2003; Bossuyt *et al.,* 2006). According to this theory, existence of broad land bridge connections between drifting India and Africa or island arcs between India and Asia or India and Africa, allowed the exchange of flora and fauna (Briggs, 2003; Aliand and Aitchison, 2008; Chatterjee and Scotese, 2010). The limited fossil evidences showed that Indian plate had remained close to Africa and Madagascar even when it began to have contact with Eurasia (Briggs, 2003). An arrangement of tectonic plates during the end of Cretaceous was even depicted, in which the northern part of peninsular India was attached to the Eurasian plate and the southern part to Madagascar plate (Rage, 1996). Seychelles–Mascarene Plateau and nearby elevated sea-floor areas could have formed a passage between southern India and Madagascar in Late Cretaceous, for faunal elements to pass (Ali and Aitchison, 2008). This supports the assumption that many Asian and African elements dispersed into India.

An exchange of biota between Asia and peninsular India was suggested due to the collision of Indian and Eurasian plate (Mani, 1974). The dispersal of Gondwanan forms out of India and into Asia upon collision, regarded as "Out-of-India"

hypothesis and the dispersal of Asian elements into India and out of Asia so called "Out-of-Asia" and "Into-India" origin (Mani, 1974). The overall similarity between fauna of peninsular India and Southeast Asia has been reported in many studies, backing the dispersal between Asia and India (Mani, 1974; Blanford, 1901; Hora, 1949; Jayaram, 1949). In case of plants, the possibility of out of Asia and into India hypothesis has been reported by Lakhanpal (1970) and Bande (1992).

Molecular phylogenetic studies suggested that several taxa, such as primates, reptiles, amphibians and invertebrates, include clades in the Indian subcontinent have radiated *in situ* (Macey *et al*., 2000; Bossuyt and Milinkovitch, 2001; Raxworthy *et al*., 2002; Kohler and Glaubrecht, 2003; Bocxlaer *et al*., 2006; Datta-Roy *et al*., 2014). On the other hand, several groups clearly showed multiple origins via dispersals across Himalayas, as evidenced by speciose plant and animal genera widely distributed outside India (Datta-Roy *et al*., 2014). In plants, Yuan *et al*. (2003) have suggested multiple dispersals of Balsamaceae into India indicating the possibility of dispersal in plants.

CHAPTER 3

DNA BARCODING IN *CALAMUS*

3. DNA BARCODING IN *CALAMUS*

3.1. INTRODUCTION

The rattan genus *Calamus* Linnaeus is the largest of all palm genera, including 520 species distributed in the Asia-Pacific region, with one species in Africa (Baker, 2015). In India, *Calamus* is represented by 46 species which are distributed in three phytogeographical regions such as Peninsular India, Eastern Himalayas and Andaman and Nicobar Islands (Renuka, 2001). *Calamus* is predominantly distributed throughout the Western and Eastern Ghats of the peninsular region of the Indian subcontinent. The identification of this rattan is mainly based on morphological characters which includes both vegetative as well as floral characters such as leaf size, stem, fruit and inflorescences. The dependence on traditional taxonomy alone for species identification is difficult in the genus due to the unavailability of flowers and fruits throughout the year (Uhl *et al*., 1995). Several taxonomic complexities like homoplasies, look-alike species, environmental plasticity and species complexes existing within this group further complicated species identification using morphological characters (Boer, 1968; Sreekumar and Henderson, 2014; Atria *et al*., 2017).

The application of molecular technology had an immense potential to provide accurate identification and classification of species. DNA barcoding is a technique using small stretch of DNA sequences for accurate species identification (Hebert *et al*., 2003). DNA barcoding has wide application in various fields of research such as conservation biology (Stoeckle, 2003), timber forensics and tracking (Asif and Cannon, 2005; Fuji, 2007), identification of adulterants (Dev *et al*., 2014) and in biosystematics (Gao *et al*., 2010; Pang *et al*., 2010). Based on recoverability and sequence quality, Consortium for Barcode of Life (CBOL) Plant Working Group recommended core barcode regions, *rbc*L, *mat*K as well as supplementary barcode regions (*psb*A-*trn*H, *rpo*B, *rpo*C, *psb*K-*psb*I and *atp*F-*atp*H) for barcoding land plants either singly or in combinations (Chase *et al*., 2005; Kress *et al*., 2005; Newmaster *et al*., 2006; CBOL, 2009; Hollingsworth *et al*., 2009). The nuclear

Internal Transcribed Spacer (ITS) region has been reported as the most divergent barcode region in plants (Chase *et al*., 2005; Kress *et al*., 2005; China Plant BOL, 2011). Previous studies proposed new barcode regions in some plant taxa where the recommended DNA barcode regions failed to discriminate species (Valentini *et al*., 2009; Ballardini *et al*., 2013; Techen *et al*., 2014; Joly *et al*., 2014; Paracchini *et al*., 2017).

Since chloroplast regions have slow rate of evolution in palms (Wilson *et al*., 1990; Gaut *et al*., 1992), the use of recommended chloroplast barcode regions are inadequate for accurate species identification in Arecaecae. Nuclear ribosomal ITS2 have proven their utility for species identification in some species of palm (Jeanson *et al*., 2011), but the presence of multiple copies lead to lack of accurate species identification in Calamoideae (Baker *et al*., 2000a). Even though *psb*A-*trn*H has been recommended as barcode for 15 Chinese *Calamus* species (Yang *et al*., 2012), 58 per cent of species discrimination is an indicative of inefficiency of employed barcode. The chloroplast region *psb*Z-*trnf*M, has been recommended as potential DNA barcode region for species identification within genus *Phoenix* (Arecaceae) (Ballardini *et al*., 2013). Previous phylogenetic studies in palms also indicated the utility of low copy nuclear regions such as PRK (phosphoribulokinase), RPB2 (RNA polymerase II) and MS (malate synthase) to resolve taxonomic complexities at lower levels in palms (Lewis and Martinez, 2000; Carl and Doyle, 2001; Lewis and Doyle, 2002; Bayton, 2005; Roncal *et al*., 2005; Loo *et al*., 2006; Norup, 2006; Thomas *et al*., 2006; Zona *et al*., 2011).

In this context, present study aims to identify an ideal DNA barcode to address prevailing taxonomic complexities in the genus *Calamus.* The study analyses species discrimination ability of 12 barcode regions, chloroplast (*rbc*L, *mat*K, *psb*A-*trn*H, *rpo*C, *rpo*B, *psb*K-*psb*I, *atp*F-*atp*H, *psb*Z-*trnf*M) and nuclear regions (ITS1, ITS2, PRK and RPB2) for the identification of *Calamus* species They were used here either alone or in combination for species discrimination using distance-, tree-, and similarity- based methods.

3.2. MATERIALS AND METHODS

3.2.1. Taxon Sampling

Multiple accessions of 21 species of *Calamus* (*Calamus brandisii, C. hookerianus, C. delessertianus, C. dransfieldii, C. shendrurunii, C. neelagiricus, C. vimnalis, C. rotang, C. metzianus, C. wightii, C. pseudotenuis, C. nagbettai, C. karnatakensis, C. lakshmanae, C. prasinus, C. stoloniferus, C. travancoricus, C. vattayila, C. thwaitesii, C. lacciferus* and *C. gamblei*) (Renuka *et al.*, 2010), were collected from their natural distribution zones in the Western Ghats of India (Appendix I). Taking account of the effect of geographical sampling scale and intraspecific variability on DNA barcoding, multiple accessions were included. Voucher specimens were deposited in the Kerala Forest Research Institute (KFRI) Herbarium.

3.2.2. DNA extraction, PCR amplification and Sequencing

Total genomic DNA was extracted from fresh and silica gel dried leaf materials using modified Cetyltrimethyl ammonium bromide (CTAB) method (Doyle and Doyle, 1990), as well as using DNeasy Plant Mini Kit (Qiagen, Germany) according to the manufacturer's protocol. Seven standard barcode loci of plastid genome (four coding regions *viz. mat*K, *rbc*L, *rpo*B, *rpo*C, three intergenic spacers *viz. psb*A-*trn*H, *psb*K-*psb*I, *atp*F-*atp*H), chloroplast non-coding region, *psb*Z-*trnf*M, two high copy nuclear regions (ITS1and ITS2) and two low copy nuclear gene regions (PRK and RPB2) were evaluated.

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 $MgCl_2$, 200 μM dNTPs, 10 pmoles of each primer, and 2U Taq DNA polymerase (Genei, Bangalore), 50-100 ng template DNA and enough distilled deionized water to give a final volume of 20 μL. PCR reactions were performed in the following regime: initial denaturation at 94°C for 5 min, 30 cycles of denaturation at 94°C for 30 sec, primer annealing at specified temperatures and an extension at 72°C for 1 min followed by final extension at 72°C for 10 min. The

primer information and optimal PCR annealing temperatures are provided in Table 3.1. The amplified products were resolved in 2 per cent agarose gel and documented using a gel documentation system (Syngene, UK). The PCR products were purified using a Nucleospin Gel and PCR Clean-up kit (Macherey Nagel, USA) as per the manufacturer's protocol, and quantified using a spectrophotometer (Nanodrop, Thermo Scientific, USA). Sanger dideoxy sequencing was performed in both forward and reverse directions (Chromous, Banglore).

Table 3.1. Barcoding primers used and PCR reaction conditions

Barcode region	**Primer**	**Primer sequence 5'-3' (Reference)**	**Annealing temperature**
*rbc*L	1F 724R	ATGTCACCACAAACAGAAAC TCGCATGTACCTGCAGTAGC (Kress *et al*., 2005)	60°C - 40 sec.
*mat*K	472F 1248R	CCCRTYCATCTGGAAATCTTGGTT GCTRTRATAATGAGAAAGATTTCTGC (Yu *et al*., 2011a)	60°C - 40 sec.
*psb*A-*trn*H	*trn*H *psb*A	GTWATGCAYGAACGTAATGCTC CGCGCATGGTGGATTCACAATCC (Kress *et al*., 2005)	59°C - 50 sec.
*rpo*C	*rpo*C *rpo*C	GGCAAAGAGGGAAGATTTCG CCATAAGCATATCTTGAGTTGG (Sass *et al*., 2007)	58°C - 50 sec.
*rpo*B	*rpo*B *F* *rpo*B *R*	AAGTGCATTGTTGGAACTGG GATCCCAGCATCACAATTCC (Ford *et al*., 2009)	58°C - 40 sec.
*psb*K-*psbI*	*psb*K *psb*I	TTAGCCTTTGTTTGGCAAG AGAGTTTGAGAGTAAGCAT (Lahaye *et al*., 2008)	55°C - 50 sec.
*atp*F-*atp*H	*atp*F *atp*H	ACTCGCACACACTCCCTTTCC GCTTTTATGGAAGCTTTAACCAAT (Hollingsworth *et al*., 2009)	60°C - 50 sec.
*psb*Z-*trnf*M	*psbZ* *trnfM*	GGTACMTCATTATGGATTGG GCGGAGTAGAGCAGTTTGGT (Scarcelli *et al.,* 2011)	50-65°C- 1 min
ITS1	ITSF ITSR	TCCGTAGGTGAACCTGCGG GCTGCGTTCATCGATGC (White *et al*., 1990)	60°C - 40 sec.
ITS2	ITS3F IT4R	GCATCGATGAAGAACGCAGC TCCTCCGCTTATTGATATGC	60°C - 40 sec.

		(White *et al*., 1990)	
PRK	PRKF PRKR	GTGATATGGAAGAACGTGG ATTCCAGGGTATGAGCAGC (Lewis and Doyle, 2002)	50-60°C- 1 min
RPB2	RPB2F RPB2R	CAACTTATTGAGTGCATCATGG CCACGCATCTGATATCCAC (Ronçal *et al*., 2005)	58°C - 40 sec.

3.2.3. Sequence alignment

The raw sequences were edited manually using BioEdit (Hall, 1999), and aligned using default alignment parameters in CLUSTAL X (Jeanmougin *et al*., 1998). The refined sequences were confirmed via BLASTn (http://blast.ncbi.nlm.nih.gov/Blast.cgi) against nucleotide database and deposited in NCBI GenBank and details are appended (Appendix II).

3.2.4. Data analysis

Three different statistical analyses viz. distance based, similarity based and tree based methods were performed to evaluate species discrimination within genus. In Distance method, we computed intra- and interspecific pairwise distances based on the Kimura two-parameter (K2P) model (Srivathsan and Meier, 2012), and the best performing region was obtained from the analysed region based on the barcoding gap. We determined the closest match of a sequence by comparing it to all other sequences in the aligned data set by using 'best match' and 'best close match' in the similarity based method (Meier *et al.,* 2006). Similarity based method help us to confirm the previous finding, regarding the best performing region. Tree-Building method was used to generate a graphical representation of the results in a tree form, which is useful for determining the power of given markers to discriminate between closely related species (Zhang *et al*., 2013). From the phylogenetic tree constructed, it is able to find the unidentified samples and resolve the taxonomic issues.

3.2.4.1. Distance-based analysis

For pairwise genetic distance (PWG-distance) method, the genetic distance was determined by MEGA v.6.0.using Kimura 2-Parameter (K2P) model (Kimura, 1980), with complete deletion option (Tamura *et al*., 2012). The average interspecific distance, theta prime and minimum interspecific distances were used to represent interspecific divergence. The average intraspecific distance, theta and maximum intraspecific distance (coalescent depth) were calculated to evaluate intraspecific variation using K2P model (Meyer and Paulay, 2005; Chen *et al*., 2010) for the studied barcode regions. DNA barcoding gap was calculated as the differences between average intraspecific and average interspecific distances (Erickson *et al*., 2008). The significance of barcoding gap was assessed using Wilcoxon matched pairs signed rank test in SPSS v 17.0 (SPSS Inc, 2007), for selected barcoding loci and possible combination (RPB2, *mat*K + *psb*A-*trn*H, *psb*A-*trn*H and *mat*K) in *Calamus*.

3.2.4.2. Similarity based analysis

Two methods, Best Match (BM) and Best Close Match (BCM) were employed to determine species identification success in *Calamus* species using TaxonDNA / Species Identifier 1.7.7 based on uncorrected p-distances (Meier *et al*., 2006). For 'BCM' method, threshold similarity values were computed from pairwise summary, in order to define how similar a barcode match. The criteria for successful identification, ambiguous identification, incorrect identification, and no match were set according to previous studies (Meier *et al*., 2006).

3.2.4.3. Tree based analysis

Neighbor-joining tree (NJ) (Saitou and Nei, 1987), was also constructed for each dataset of studied barcode regions using MEGAv.6.0. adopting K2P model (Kimura, 1980; Tamura *et al*., 2013). Phylogenetic tree was constructed for individual data matrix as well as for the concatenated datasets by Maximum likelihood and Bayesian Inference using RAxML-HPC2 7.6.3 (Stamatakis, 2006) and Mr Bayes

3.2.2 (Ronquist and Huelsenbeck, 2003), respectively (http://www.phylo.org/). jModelTest 2.1.4 (Posada and Buckley, 2004), was used to select a best fit model of nucleotide substitution under Akaike Information Criterion (AIC) prior Maximum likelihood analysis and Bayesian analysis, ML trees were assessed by 1000 bootstrap replicates of heuristic searches (Stamatakis, 2006), Bayesian Markov Chain Monte Carlo (MCMC) algorithm was run 5,00,000 generations with one cold chain and 3 heated chains, starting from random trees and sampling trees every 100 generations, until getting standard deviation value below 0.01. The 50 per cent majority-rule consensus trees were constructed after the first 25 per cent of sampled trees were removed during burn in period. The posterior probability (PP) of each topological bipartition was calculated across the remaining trees in Bayesian Inference. The constructed final tree was visualized in Figtree v1.3.1 (Rambaut, 2009).

3.3. RESULTS

3.3.1. Amplification and sequence analysis

Basic sequence information of selected DNA barcode regions *viz.* *rbc*L, *mat*K, *psb*A-*trn*H, *rpo*C, *rpo*B, *psb*K-*psb*I, *atp*F-*atp*H and RPB2 and their combination was compared (Table 3.2). ITS1 and ITS2 gave multiple amplified products and failed to produce good quality chromatogram peaks. The *psb*Z-*trnf*M and PRK regions also failed to yield good amplification. The remaining barcoding regions exhibited good PCR amplification and sequencing rate (Fig. 3.1). Among the nine regions studied, RPB2 showed highest percentage of parsimony informative sites (11.6 %) followed by *psb*A-*trn*H (6.74 %), *mat*K+ *psb*A-*trn*H (6.2 %) and *mat*K (4.51 %). Parsimony informative sites in *rbc*L, *rpo*B, *rpo*C, *psb*K-*psb*I, *atp*F-*atp*H were negligible when compared to other barcode regions. RPB2 showed greatest percentage of nucleotide variation (11.6 %) based on number of variable sites. Indels were more prevalent in *psb*A-*trn*H when compared to other spacers sequence length varied from 500bp to 800bp.

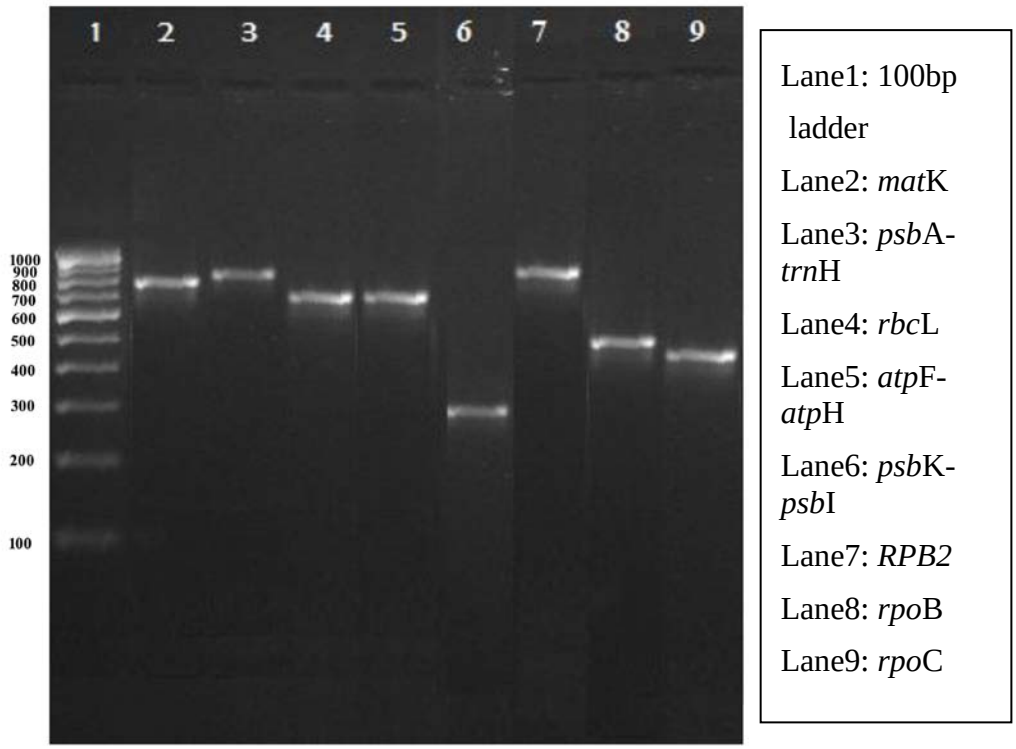

Fig. 3.1. PCR amplified products of eight analysed DNA barcode regions

Table 3.2. Sequence characteristics of eight DNA regions and their combination

Comparison	***rbc*L**	***mat*K**	***psb*A-*trn*H**	***rpo*B**	***rpo*C**	***psb*K-*psb*I**	***atp*F-*atp*H**	***mat*K + *psb*A-*trn*H**	**RPB2**
Sequence length (bp)	674	709	786	463	456	271	637	1499	732
Conserved region	666	659	675	456	453	271	620	1330	642
Variable region	7	50	87	7	3	4	16	142	90
Parsimony informative site	4	32	53	4	2	4	11	93	85
Singleton site	3	18	33	3	1	0	5	47	5

3.3.2. Distance-based analysis

Genetic divergence was estimated using six parameters in MEGA 6.0 (Table 3.3). Maximum average interspecific distance was observed in RPB2 (0.24), followed by *psb*A-*trn*H (0.23), the combination of *mat*K + *psb*A-*trn*H (0.19), *mat*K (0.17), *atp*F-*atp*H (0.004), *rpo*B (0.036), *psb*K-*psb*I (0.03), *rbc*L (0.028), and *rpo*C (0.02). Highest average intra specific distance was found in *mat*K + *psb*A-*trn*H combination (0.0119). *rpo*C, *rpo*B, *rbc*L and *psb*K-*psb*I showed lowest average intra specific distances.

Table 3.3. Parameters of interspecific divergence and intraspecific variation of eight barcode regions and their combination

Parameters	***rbc*L**	***mat*K**	***psb*A-*trn*H**	***rpo*B**	***rpo*C**	***psb*K-*psb*I**	***atp*F-*atp*H**	***mat*K + *psb*A-*trn*H**	**RPB2**
Average Inter specific	0.028±0.019	0.170±0.04	0.23±0.077	0.036±0.026	0.02±0.016	0.03±0.028	0.07±0.036	0.191±0.040	0.246±0.074
Average theta prime	0.0019±0.0014	0.009±0.003	0.017±0.005	0.002±0.001	0.0009±0.0009	0.002±0.001	0.004±0.002	0.012±0.003	0.013±0.004
Interspecific distance	0	0	0	0	0	0	0	0	0
Average intraspecific distance	0.0004±0.0004	0.009±0.002	0.0076±0.002	0.0004±0.0003	0.0004±0.0003	0.0004±0.0004	0.0011±0.0003	0.0119±0.002	0.0008±0.0008
Maximum intra specific (coalescent depth)	0.002±0.002	0.023±0.007	0.027±0.007	0.004±0.003	0.003±+0.002	0.002±0.002	0.015±0.002	0.088±0.001	0.010±0.003
Average theta	0.0004±0.0004	0.0008±0.00005	0.009±0.0027	0.0004±0.0004	0.0004±0.0003	0.0004±0.0004	0.0009±0.0003	0.0116±0.004	0.0009±0.0005
Barcoding gap	0.0276	0.161	0.2224	0.0356	0.0196	0.0296	0.0689	0.1791	0.2452

A significant barcoding gap (0.2452) was shown by RPB2, able to discriminate the selected *Calamus* species (Fig. 3.2). A less barcoding gap was observed in five barcode loci *rpo*C, *rpo*B, *rbc*L, *atp*F-*atp*H and *psb*K-*psb*I i.e. 0.0196, 0.0356, 0.0276, 0.0689 and 0.0296, respectively. Eventhough, *psb*A-*trn*H, *mat*K and their combination (*mat*K + *psb*A-*trn*H) displayed distinct barcode gaps (0.2224, 0.161 and 0.1791 respectively), but were lower than RPB2. Wilcoxon's signed rank test displayed by RPB2 barcode was significantly higher than that of other barcode regions (*mat*K + *psb*A-*trn*H, *psb*A-*trn*H, *mat*K) (Table 3.4).

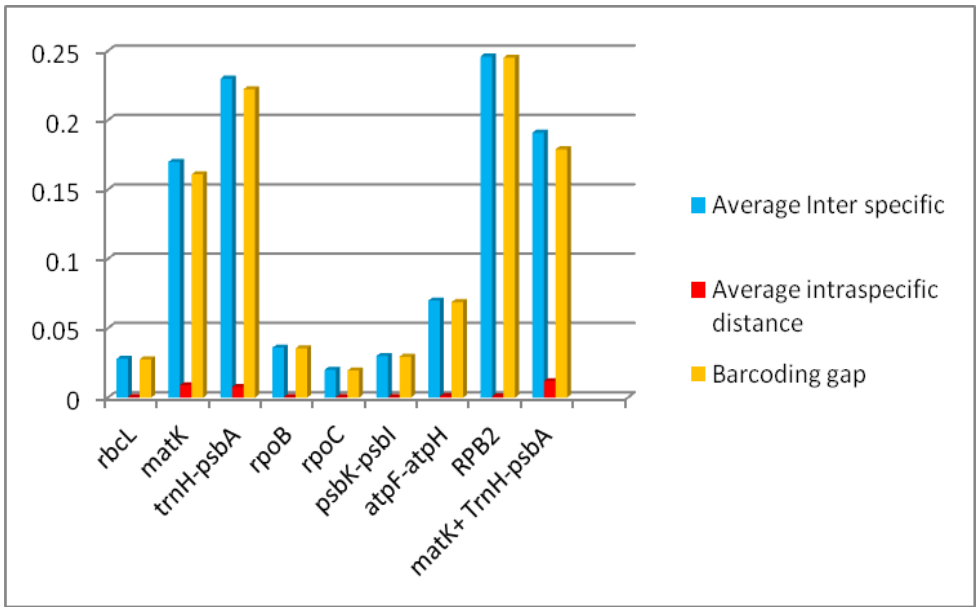

Fig. 3.2. DNA barcoding gap for eight barcode loci and best possible combinations

3.3.3. Similarity based analysis

Among the analysed barcode regions, RPB2 had highest successful identification rate (96.96%) while *rpo*C (0%) failed to show any successful identification under the 'BM' methods (Table 3.5). The results of the 'BCM' method were similar with that of the 'BM' method, for all the regions. The combination of *mat*K + *psb*A-*trn*H were able to discriminate about 62.7 %, while *psb*A-*trn*H and *mat*K as single locus had 44.6 % and 42.8 % species discrimination efficiencies respectively.

Table 3.4. Test of significance of interspecific divergence among loci in Wilcoxon-signed rank test

W+	**W-**	**Relative ranks**		**n**	**P>**	**Result**
		W+	**W-**			
RPB2	*mat*K+*trn*H	7398	4229	153	0.001	RPB2> *matK*+*trnH*
RPB2	*psb*A- *trn*H	9763	4942	171	0.001	RPB2> *psb*A- *trn*H
RPB2	*mat*K	10266	4268	171	0.001	RPB2> *mat*K
matK+ *psb*A- *trn*H	*psb*A- *trn*H	8767	3014	153	0.001	*mat*K+*trn*H> *psb*A- *trn*H
*mat*K+ *psb*A- *trn*H	*mat*K	7922	3705	153	0.001	*mat*K+*trn*H> *mat*K
*psb*A- *trn*H	*mat*K	12480	2225	171	0.001	*psb*A- *trn*H >*mat*K

3.3.4.Tree based analyses

The discriminatory success of single or combined barcodes, were also determined by evaluating the percentage of each species being monophyletic using a phylogenetic tree. Phylogenetic tree constructed based on Maximum Likelihood, Bayesian Inference and Neighbour-Joining showed similar topologies. RPB2 gave only a robust phylogenetic tree with large number of monophyletic clades in the genus *Calamus* in tree based analyses. The remaining barcode loci failed to delimit species boundaries in phylogenetic tree.

Table 3.5. Identification success rates of eight barcode regions singly and their possible combinations using TaxonDNA under 'Best Match' and 'Best Close Match' criteria

Parameters Best Match (BM) & Best Close Match (BCM)	***rbc*L**	***mat*K**	***psb*A-*trn*H**	***rpo*B**	***rpo*C**	***psb*K-*psb*I**	***atp*F-*atp*H**	***mat*K+*trn*H-*psb*A**	**RPB2**
Correct identifications according to "BM":	6.25%	42.85%	44.64%	4.41%	0	8.33%	22.5%	62.74%	96.96%
Ambiguous according to "BM":	93.75%	34.28%	32.14%	95.58%	96.61%	90.0%	70.0%	3.92%	1.51%
Incorrect identifications according to "BM":	0.0%	22.85%	23.21%	0.0%	3.38%	1.66%	7.5%	33.33%	1.51%
Correct identifications according to "BCM":	6.25%	42.85%	44.64%	4.41%	0	8.33%	22.5%	62.74%	96.96%
Ambiguous according to "BCM":	93.75%	34.28%	32.14%	95.58%	96.61%	90.0%	70.0%	3.92%	1.51%
Incorrect identifications according to "BCM"	0.0%	22.85%	23.21%	0.0%	3.38%	1.66%	7.5%	33.33%	1.51%

*psb*A-*trn*H, *mat*K, *atp*F-*atp*H and *mat*K + *psb*A-*trn*H combination showed monophyletic clade which are species-specific but failed to give complete resolution. The analysed barcode regions, *rpo*C, *rpo*B, *rbc*L and *psb*K-*psb*I failed to resolve species in the phylogenetic analysis. RPB2 gave more resolved topology in which multiple accessions of all species formed well defined clusters with higher bootstrap support and good posterior probability value (1.00). Furthermore, BI and NJ trees formed a well resolved tree using RPB2 gene region, wherein the members of the genus *Calamus*, formed a single clade with majority of the species clustered into different subclades (Fig. 3.3 and Fig. 3.4 respectively). For the analysis we have taken the BI tree as its shows better phylogenetic relationship (Fig. 3.3).Out of the 21 species analysed, 15 species (*C. brandisii, C. shendrurunii, C. neelagiricus, C. vimnalis, C. rotang, C. wightii, C. nagbettai, C. karnatakensis, C. lakshmanae, C. prasinus, C. stoloniferus, C. travancoricus, C. vattayila, C. thwaitesii*, and *C. lacciferus*) displayed well supported distinct monophyletic clade in phylogenetic tree. *C. metzianus, C. gamblei* and *C. hookerianus* showed polytomy in the phylogenetic analysis. *C. pseudotenuis* clustered with *C. hookerianus, C. dransfieldii* and *C. delessertianus* whereas *C. thwaitesii* was found as standalone species with strongly supported bootstrap value (100 %) in all constructed phylogenetic trees.

3.4. DISCUSSION

The 12 studied barcoding loci viz. *rbc*L, *mat*K, *psb*A-*trn*H, *rpo*C, *rpo*B, *psb*K-*psb*I, *atp*F-*atp*H, *psb*Z-*trn*fM, ITS1, ITS2, PRK and RPB2 and their combinations had different success rates in discriminating the species owing to many factors, including mutation rate and inheritance mode. Previous studies reported low species discrimination power of *rbc*L, *mat*K, *rpo*B, *rpo*C, *atp*F-*atp*H and *psb*K-*psb*I (Feng *et al*., 2013; Han *et al*., 2016) as observed in present study. ITS1 and ITS2 failed to give good PCR amplification efficiency probably due to the presence of multiple copies in Calamoideae as reported by Baker *et al*. (2000a). *psb*A-*trn*H exhibited dramatic differences in sequence lengths due to presence of insertion/deletions among the analysed individuals of species, which led to the difficulties in sequence

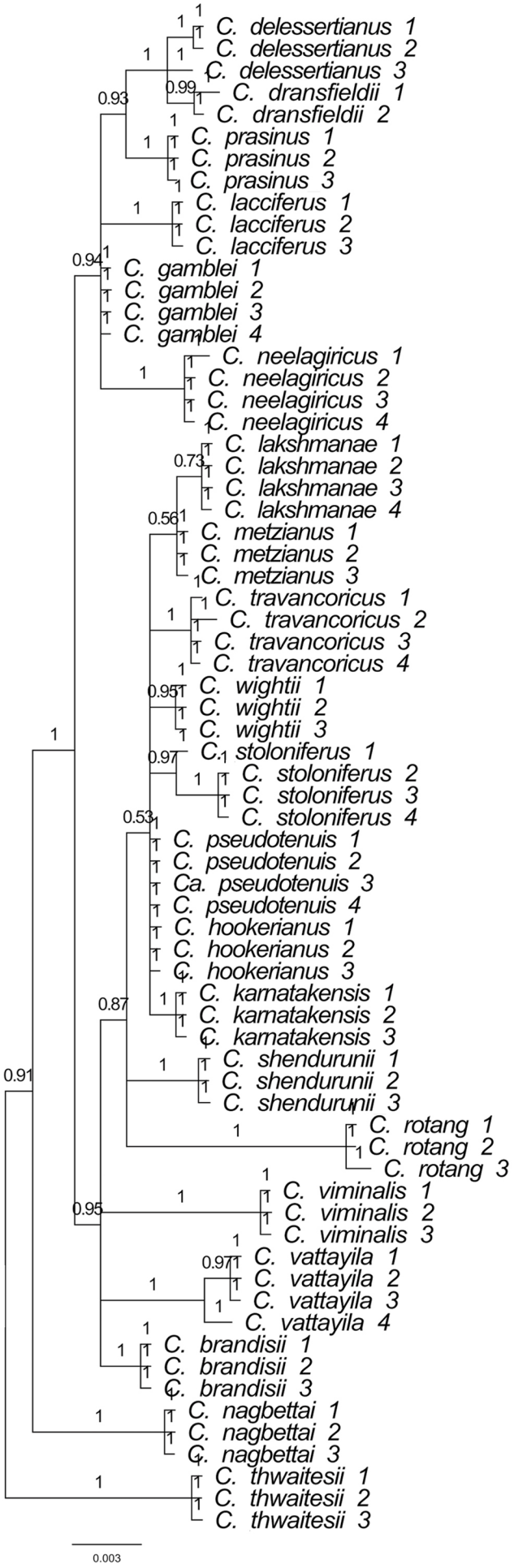

Fig. 3.3. Phylogenetic tree construction using low copy nuclear region RPB2, based on GTR +G model and AIC in MrBayes v.3.2.2.

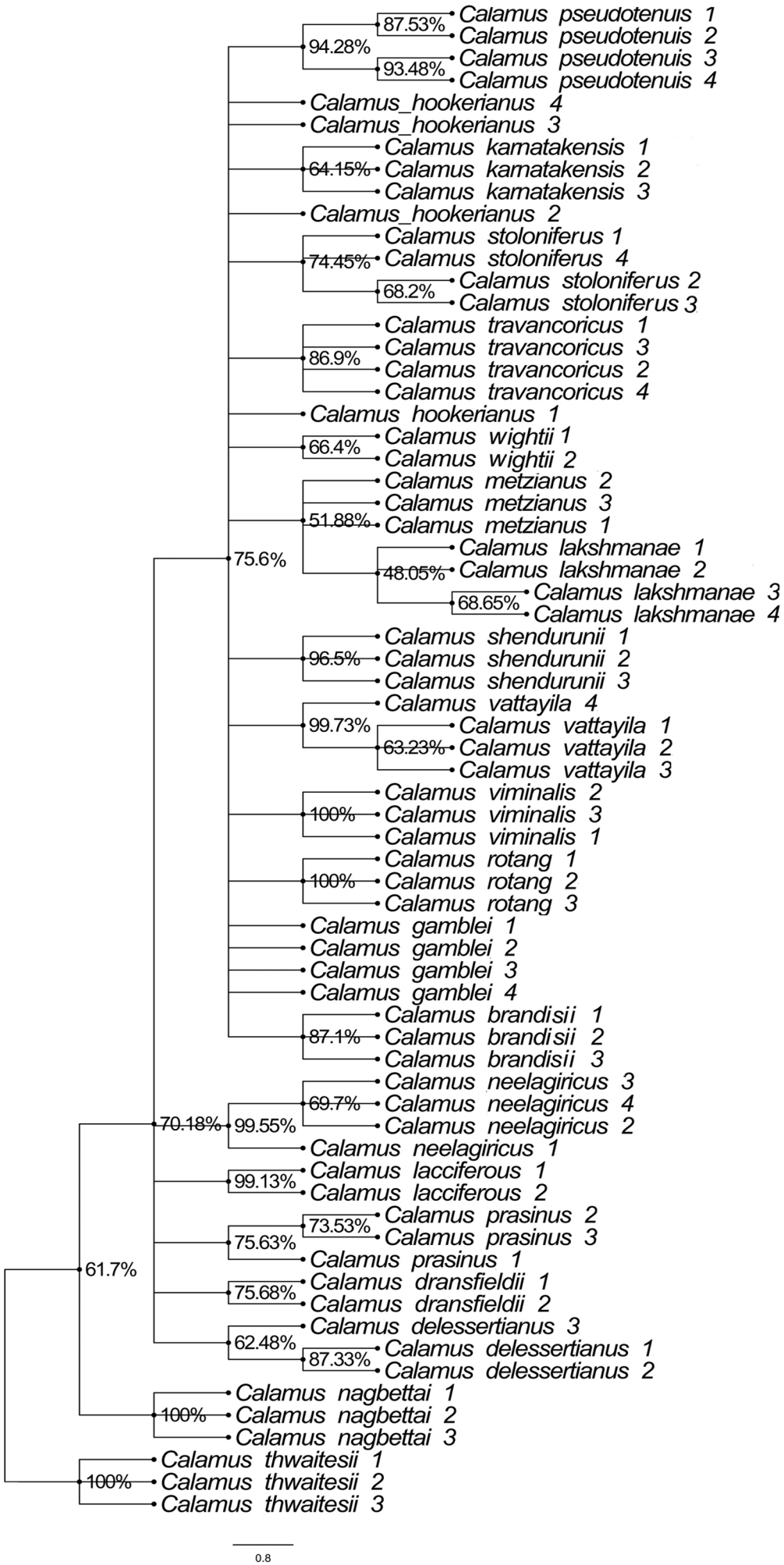

Fig. 3.4. Neighbor-joining tree (NJ) of the genus *Calamus* using RPB2 based on Kimura 2 Parameter model using MEGA v.6.0 with bootstrap percentage.

alignment making it unsuitable for species identification (Kress *et al*., 2005; CBOL, 2009). Barcoding gap exhibited by *mat*K and *psb*A-*trn*H was not successful in species delimitation due to high intraspecific differences. Considering three analyses *viz*. distance based, similarity based and tree based, RPB2 nuclear barcode region appears to be an ideal DNA barcode for *Calamus* species identification as compared to other candidate chloroplast barcode regions

To date, species identification success was reported only in a few plant taxa using the universal DNA barcoding regions. However, in *Calamus* along with some other taxa such as Alooideae (Daru *et al*., 2017), wild potatoes (Spooner *et al*., 2009), willows (Percy *et al*., 2014), Indian berberis (Roy *et al*., 2010), it poses a much greater challenge. Plastid data often failed to highlight species boundaries in sedges, willows, potatoes, etc. (Percy *et al*., 2008; Starr *et al*., 2009; Spooner *et al*., 2009; Hassel *et al*., 2013), due to widespread hybridization, introgression, or incomplete lineage sorting (Hollingsworth *et al*., 2011).

In palms, substitution rate estimates from restriction site variation in chloroplast DNA was 5 to 13-fold slower than for grasses (Wilson *et al*., 1990). The slow evolutionary rate of palm DNA hinders the use of plastid genes in palm molecular studies (Gaut *et al*., 1992& 1996; Wilson *et al*., 1990). Calamoideae possess lowest substitution rate (1.3×10^{-10}) in chloroplast region when compared with other palms (Wilson *et al*., 1990), clearly indicates the inefficiency of plastid primers to discriminate species. Even though, *psb*A-*trn*H (Al-Qurainy *et al*., 2011; Yang *et al*., 2012) and nuclear ITS2 (Jeanson *et al*., 2011), have proven their utility as barcodes in palms, they failed to succeed here. The nuclear ITS region, recommended for barcoding, reported to have incomplete concerted evolution and consequent intragenomic heterogeneity among copies in palms (Baker *et al*., 2000a). Though *psb*A-*trn*H was suggested to be a good DNA barcode in *Calamus* (Yang *et al*., 2012), yet the presence of indels (Kress *et al*., 2005) and high amount of intra specific variation, leading to difficulties in sequence alignment, foreshadows its inefficiency as a good DNA barcode. In *Daemonorops*, species discrimination cannot be derived alone with single plastid barcode region or nuclear marker

(Umapathy *et al*., 2015). Hence using recommended barcodes in *Calamus* for species identification becomes a tedious task. Intense debate and confusion have originated in plant community regarding the search for universal barcodes and many solutions are put forth from different corners in terms of combination of barcodes and introduction of new regions (Valentini *et al*., 2009; Ballardini *et al*., 2013; Techen *et al*., 2014; Joly *et al.*, 2014; Paracchini *et al*., 2017).

In previous instances, where plastid and nuclear barcode regions failed, low-copy nuclear regions have been successfully used in resolving phylogenetic relationships especially at lower taxonomic levels, regardless of the difficulties like paralogy, concerted evolution and intragenic polymorphism (Lewis and Doyle, 2001 &2002; Carl and Doyle, 2001; Gunn, 2004; Roncal *et al*., 2005; Bayton, 2005; Thomas *et al*., 2006; Norup *et al*., 2006). As these regions can evolve up to five times faster than the plastid genome, it can resolve even recently and rapidly diversifying lineages (Sang, 2002; Small *et al*., 2004; Norup *et al*., 2006). Low copy nuclear regions were used for DNA barcoding in *Clermontia* (Campanulaceae) and *Cyrtandra* (Gesneriaceae), where plastid genes had slow evolutionary rates (Pillon *et al*., 2013).

Low copy nuclear region, RPB2 was efficiently used in palm phylogenetic studies, to resolve relationships at taxonomic levels. This was reported to be more informative than chloroplast regions and has fewer molecular evolutionary problems than other nrDNA regions (Roncal *et al*., 2005). The efficiency of this low copy nuclear region was explored earlier in angiosperm systematics (Denton *et al*., 1998; Oxelman and Bremer, 2000; Oxelman *et al*., 2004; Pfeil *et al*., 2004). Even though paralogs of RPB2 were identified in core eudicots (Oxelman and Bremer, 2000; Oxelman *et al*., 2004), no paralogs were reported in palm tribe *Geonomeae* (Ronçal *et al*., 2005). Even though the utility of low copy nuclear regions like RPB2 as a barcode was restricted because of their low copy number (Kress *et al*., 2005). PCR amplification success rate was 100 per cent in the present study. Further, p-distance, similarity and tree based analyses also proved the utility of RPB2 as a possible DNA barcode to distinguish *Calamus*.

The development of suitable DNA barcodes for the accurate species identification in taxonomically complex genus *Calamus* was challenging due to the inefficiency of the recommended plastid regions. In the present study, the low copy nuclear region, RPB2 showed highest species resolution and was successful in unravelling the species complexities in the genus. RPB2 nuclear region can be recommended as the barcode region for resolving the taxonomic problems in the genus *Calamus* and a promising barcode region for palms in general.

CHAPTER 4

RESOLVING THE TAXONOMIC COMPLEXITIES IN *CALAMUS* OF SOUTH INDIA

4. RESOLVING THE TAXONOMIC COMPLEXITIES IN *CALAMUS* OF SOUTH INDIA

4.1. INTRODUCTION

Genus *Calamus* is the largest of Arecaceae, the only rattan genus distributed in peninsular part of India. Taxonomic complexities like homoplasies, look-alike species, environmental plasticity and species complexes, prevail within this group (Boer, 1968; Sreekumar and Henderson, 2014), which make species identification complicated. Certain morphological characters, showed a considerable amount of variation within individuals of *Calamus* genus, which was found to be unreliable for species identification (Dransfield, 1979; Renuka, 1994; Sreekumar *et al.,* 2006). The use of DNA nucleotide sequence data has awakened a quickly growing interest in molecular systematic analyses which will give a hand to provide accurate identification and classification of the species. Molecular tools like DNA barcoding can lend a helping hand to resolve the identification issues within this genus. DNA barcoding have broad scientific applications in the of species identification. It is of great utility where traditional methods are ambiguous, for the identification of cryptic species, resolving taxonomic complexities (Hajibabaei *et al*., 2007; Chen *et al.,* 2011), conservation (Shivji *et al.,* 2002). In forest forensics barcoding aids in timber identification, illegal felling (Asif and Cannon, 2005; Fuji, 2007) and resolving timber adulteration (Dev *et al.,* 2014).

The highly specious genus *Calamus* was reported to show high rate of morphological variation (Dransfield, 1979; Renuka, 1994). Unresolved confusions in delimiting certain species' as in the case of *C. metzianus, C. rotang* and *C. rivali*s and existence of *C. gamblei* species complex are among the few taxonomic complexities existing in the southern part of India which require more evaluation. Morphological affinities between certain species also make species identification more problematic. *C. hookerianus, C. pseudotenuis,* and *C. dransfieldii, C. delessertianus* are few examples for these, which has to be analysed.

4.1.1. Taxonomic revision of *C. metzianus, C.rotang* and *C. rivalis*

Calamus rotang L. was first reported to be found in the Coramandal coast, Madras, Nellore, Chingleput, and Kurnool at Courtallum (Beccari, 1908). A slender rattan collected from the coastal belt and sacred groves from Kerala was first identified as *Calamus rotang* L. (Renuka, 1992). But, later this was revealed to be a misidentification of *Calamus rivalis* (Renuka and Sreekumar, 2006), which was widely distributed in Sri Lanka also. Similarly, *C. rotang* reported from Asramam compound of Kollam dist. and some sacred groves of Kerala (Sreeja and Khaleel, 2010), and *C. metzianus*, a slender species described from Sri Lanka are having morphological confusions. In India, *C. metzianus* was originally collected from Canara district of Karnataka in 1853 and later Renuka and Bhat (1987), reported the same from Nilambur, Kerala. According to Beccari (1908), *C. metzianus* is the only continental form of *C. rivalis* distinguished from one another in the presence of characteristics such as having larger fruits with distinctly channelled fruit scales in the former when compared with the later species. Taxonomic reconsideration was attempted through the analysis of different Indian and Sri Lankan populations of *C. rivalis* and *C. metzianus* using RAPD markers and morphological characters and results the merging of *C. rivalis* and *C. metzianus* (Sreekumar *et al*., 2006). RAPD had limitation in species discrimination due to lack of reproducibility.

4.1.2. *Calamus gamblei* species complex

The existence of "species complexes" in the genus *Calamus* were reported (Sreekumar and Henderson, 2014; Atria *et al*., 2017). "*C. gamblei* species complex" comprising of *C. gamblei, C. lacciferus, C. neelagiricus, C. prasinus, C. dransfieldii, C. renukae,* and *C. shendurunii* distributed in the Western Ghats of peninsular India. *C. gamblei* species complex did not consistently differ from one another and the morphological characters used to distinguish them vary within as well as amongst species. All species in this complex share similar morphological characters such as the presence of long, parallel, apical pinnae joined at their bases, pinnae with long spines on the veins adaxially, recurved rachillae with pistillate

dyads arranged in alternate, non-opposite rows so that one side of the rachillae is without flowers, pistillate dyads, born on a distinct pedicel and yellow fruits with raised scales. Morphology of fruits of this complex are shown in Fig. 4.1. Species included in this complex are difficult to demarcate using morphology alone and hence molecular technique like DNA barcoding has to be assessed.

4.1.3. Analyzing morphologically similar species

C. delessertianus was first described by Beccari in 1908 from India, based on a portion of pistillate specimen collected by Roxburgh. Later, in 1914 Beccari, considered that *C. delessertianus* was conspecific with *C. tenuis*. In 1999, Renuka contradicted this, and stated that *C. delessertianus* (Roxburgh's specimen) and *C. tenuis* are different. She supported her findings by quoting the presence of *C. delessertianus* and absence of *C. tenuis* in South India. Later in 2014, Sreekumar and Henderson synonymised *C. delessertianus* (described by Beccari in 1908) with C. tenuis after observing the holotypes (G) and isotype (FI) images of *C. delessertianus* specimen. According to Renuka (1999) *C. dransfieldii* initially identified from Dhoni Hills, Palghat (Renuka, 1986), was different from *C. delessertianus* (described by Renuka, 1999), only by its pale green, sparingly spiny leaf sheath and pale green leaflets arranged more closer than *C. delesertianus*. *C. delesertianus* has dark green coloured heavily spiny leaf sheath and dark green leaflets.

Similarly, *C. hookerianus* and *C. pseudotenuis* are morphologically similar, in Beccari's classification (1908), *C. hookerianus* and *C. pseudotenuis* were included in a single group, the 'group V B', as both are scandant and slender palms, having similar very long partial inflorescence with spikelets on each sides. He has also commented on the "great similarities" between them. They have numerous, equidistant leaflets which are narrowly linear. Both have sub-globose ovate, fruit with scales in 18 series, yellowish base, finely erosely toothed margin and seeds have non-ruminate endosperm. These are difficult to distinguish through traditional taxonomy based identification.

Calamus shendurunii

Calamus lacciferus

Calamus prasinus

Calamus delessertianus

Fig. 4.1. Fruit morphology in *Calamus gamblei* species complex.

Calamus neelagiricus *Calamus gamblei*

Fig. 4.1. Fruit morphology in *Calamus gamblei* species complex cont.

Fig. 4.2. Multiple Sequence Alignment showing transversions in the *Calamus gamblei* species complex.

RPB2, low-copy nuclear region has been found to be efficient in species discrimination of the genus *Calamus* in our previous study. In the present study, species similarities between *C. metzianus*, *C.rivalis* and *C.rotang* were analysed using DNA barcoding region such as *rbc*L, *mat*K, *psb*A-*trn*H, RPB2 and their combinations. Using RPB2 as a DNA barcode, the *C. gamblei* species complex has to be evaluated. As well as, the species showing morphological affinities has to be analysed to confirm their identity.

4.2. MATERIALS AND METHODS

4.2.1. Sample collection

Samples of *C. rivalis* were collected from Cherthala (Alappuzha) and Asramam (Kollam). *C. metzianus* was collected from Honavar (Karnataka), Edayilekadu and Thekkumbad (Kannur), Pattakarimba (Nilambur, Malappuram) and Ponnakudam (Ernakulam). Multiple accessions of *C. rotang* were collected from Trichy (Tamil Nadu) and Warangal (Andhra Pradesh). Mulitple accessions of *C. gamblei, C. lacciferus, C. neelagiricus, C. prasinus, C. dransfieldii, C. shendurunii, C. delessertianus, C. hookerianus* and *C. pseudotenuis* and its multiples were collected from their natural distribution zones. Samples were collected from all populations and vouchers were deposited in Kerala Forest Research Institute (KFRI) herbarium. Procedures for DNA isolation, amplification using *rbc*L, *mat*K, *psb*A-*trn*H, RPB2 and their sequencing has been described in the previous chapter.

4.2.2. Data Analysis

In the analysis, Tajima's relative rate test for testing molecular clock hypothesis was performed to test the constancy of evolutionary rates among three species, *C. metzianus, C. rivalis* and *C. rotang,* keeping *C. rotang* (collected from type location Trichi) as an outgroup sequence (Tajima, 1983). P-value and X^2 Tajima's test static were calculated with one degree of freedom. Genetic distance analysis was conducted using Kimura 2-parameter (K2P) method in MEGAv.6.0 (Tamura *et al.*, 2013). All positions containing gaps and missing data were eliminated from dataset

adopted complete deletion option. Genetic distance analysis was conducted using Kimura 2-parameter method in MEGA v.6.0. Neighbor-joining (NJ) tree was constructed using Kimura 2 Parameter distance as model with 1000 bootstrap replicates using MEGA v.6.0. In all cases indels were treated as complete deletion. A species was discriminated when more than 50% of the sampled individuals fell in the same monophyletic group in the NJ tree. This relatively low threshold has been chosen to reflect minimum probability for which a correct identification would be more likely than a wrong identification.

Phylogenetic tree was also constructed based on Bayesian Inference using RPB2 for analysing *Calamus gamblei* species complex (Fig.3.2). Sequence analysis was carried out for the morphologically similar species using CLUSTAL X, as well as these species was analysed in phylogenetic tree (Fig. 4.2).

4.3. RESULTS

Sequence length and basic sequence statistics like conserved sites, variable sites, singletons and transition/transversion ratio, based on the CLUSTAL X alignment aswell as with alignment explorer in MEGAv.6.0, are provided in Table 4.1. Among four standard barcodes and its combinations, total number of aligned nucleotides, conserved sites, variable but uninformative and parsimony informative sites showed variation except for *rbc*L (Table 4.1). *rbc*L barcode region didn't have nucleotide variation among analysed sequences and hence their combination were excluded in this study. Generally, a very low percentage of parsimony informative characters (1 or 2%) were observed in all three analysed barcodes and their combinations.

4.3.1. Taxonomic revision of *C. metzianus, C. rotang* and *C. rivalis*

Unique nucleotide differences were found in *C. metzianus, C. rivalis* and *C.rotang* based on Tajima's test (Table 4.2). Sequence alignments for SNPs and unique SNPs were found with RPB2 in *C. rotang* were examined. Unique insertions of 6 bp and 20 bp were also found in *C. rotang* using *mat*K and *psb*A-*trn*H respectively (Fig. 4.3

and Fig. 4.4). Both transitions and transversions were found in *C. rotang* (samples collected from Trichi) using RPB2 and *mat*K. Transitions were more common than transversions, and hence transition/ transversion ratio was >1 in RPB2 and *mat*K. The concatenated sequences using three barcode regions were used to generate a dendrogram depicted genetic relationship among *C. metzianus*, *C. rivalis* and *C. rotang*. *C. metzianus* and *C. rivalis* which was collected from Kerala formed a monophyletic group (Fig. 4.5).

Table 4.1: Summary statistics of four barcodes and their combinations

Barcode regions	**Sequence length (bp)**	**Conserved region**	**Variable region**	**Parsimony informative site**	**Singleton sites**
*rbc*L	673	673	0	0	0
*mat*K	703	694	7	6	1
*psb*A-*trn*H	733	718	14	7	7
RPB2	725	710	15	15	0
*mat*K+ *psb*A-*trn*H	1436	1414	21	13	8
*mat*K + RPB2	1428	1406	22	21	1
*psb*A-*trn*H+ RPB2	1458	1428	29	22	7
*mat*K+*psb*A-*trn*H+ RPB2	2161	2124	36	28	8

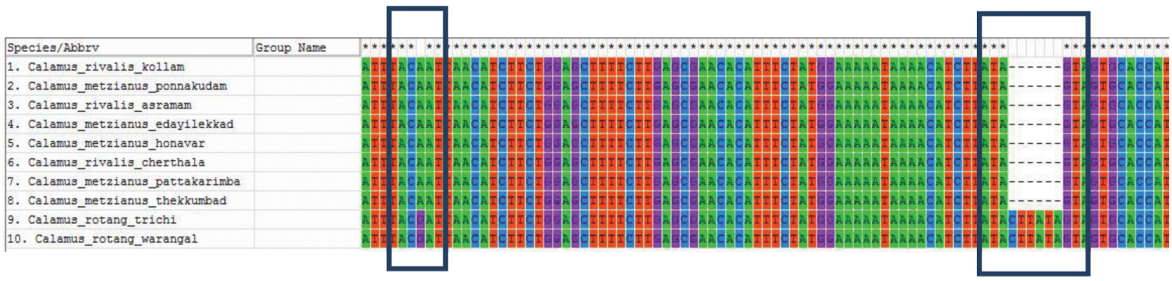

Fig. 4.3. Multiple sequence alignment of *mat*K sequences of *C. rotang* and *C. metzianus*

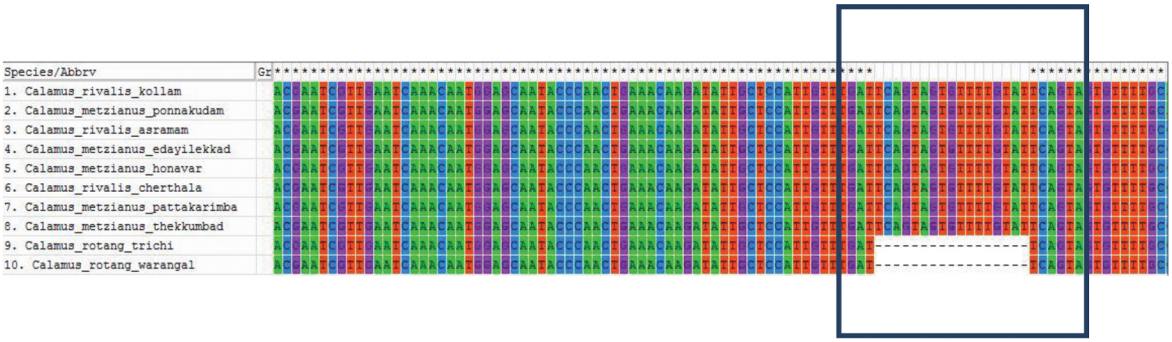

Fig.4.4. Multiple sequence alignment showing unique insertions in *psbA-trnH* sequences of *C. rotang*

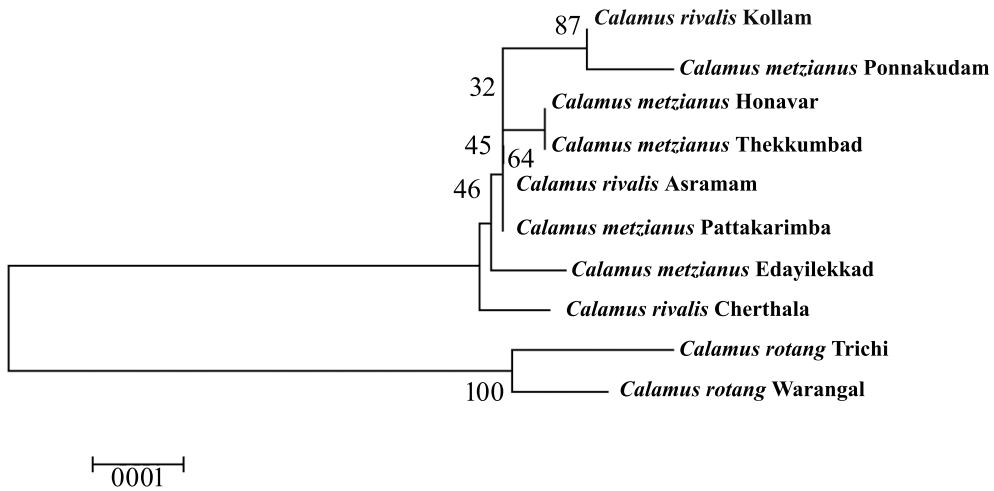

Fig. 4.5. Dendrogram depicting genetic relationships based on *trnH-psbA matK, RPB2* sequences

4.3.2. *Calamus gamblei* species complex

The present study revealed, six species namely *C. gamblei, C. lacciferus, C. neelagiricus, C. prasinus, C. dransfieldii* and *C. delessertianus* have a similar transversion and came in a single clade except *C. shendurunii* (Fig. 3.3). This confirmed close similarity among species of the complex except *C. shendurunii.*

Species, *C. dransfieldii* and *C. delessertianus* was found to be clustered together in the phylogram (Fig. 3.3) showing sequence similarities. *C. hookerianus* and *C. pseudotenuis* was found to be similar without any differences in RPB2 sequences. A common deletion of 300 bp in the *psb*A-*trn*H region for both *C. hookerianus* and *C. pseudotenuis* showed the similarity of both species.

Table 4.2. Unique nucleotide differences and Tajima's test (1993) statistics for barcode loci.

Measures	***rbc*L**	***mat*K**	***psb*A-*trn*H**	**RPB2**
Divergent sites in all four sequences	0	0	0	0
Unique differences in *C. metzianus*	0	0	1	0
Unique differences in *C. rivalis*	0	0	1	0
Unique differences in *C. rotang*	0	5	6	15
Tajima's test statistics between *C. metzianus, C. rivalis* and *C. rotang*	X^2=0.00 P=1.00 with 1 degree of freedom	X^2=0.00 P=1.00 with 1 degree of freedom	X^2=0.00 P=1.00 with 1 degree of freedom	X^2=0.00 P=1.00 with 1 degree of freedom

4.4 DISCUSSION

4.3.3. Taxonomic revision of *C. metzianus, C. rotang* and *C. rivalis*

Among the four tested barcoding regions, unique SNPs and indels were found in *mat*K, *psbA-trnH,* RPB2 which efficiently used to discriminate in *C. rotang. C.*

rivalis was found to be genetically similar with *C. metzianus,* in all the four analysed barcode regions. DNA barcoding regions failed to separate these two as different species which supported the previous study in which *C. rivalis* merged with *C. metzianus* using morphological characters and RAPD markers (Sreekumar *et al*., 2006). Species collected from coastal areas (Cherthala) and sacred groves (Ponnakudam, Edayilakkadu, Thekkumbad) of Kerala was earlier reported as *C. rotang* (Renuka, 1992; Sreeja and Khaleel, 2010), didn't show genetic sequence similarity with *C. rotang* collected from Mukkombu, Trichi and Warangal, while this had more affinity with *C. metzianus* without specific insertions in *mat*K and *psb*A-*trn*H barcode regions. The present study revealed the misidentification of species. In phylogenetic tree construction using three barcode regions and their combinations, samples collected from Ponnakudam, Edayilakkadu, Thekkumbad, Cherthala, Pattakarimba, Asramam, Kollam and Honavar found to be in a single clade in dendrogram using the three barcoding regions and their combination, which clearly depicts the close similarity of the species except species collected from Trichi and Warangal. Species from Trichi and Warangal is of *C. rotang.*

Narrow species concept was followed in the genus *Calamus* (Dransfield, 2000), and genera of Arecaeace (Trail, 1877; Dransfield and Beentje, 1995). The distinction of *C. metzianus* from *C. rivalis* is based on its larger fruit size and fruit scale channel in the middle and the rest of the morphological features are almost same for both species. But it was found that fruit size and fruit scale channel showed considerable amount of variation within individuals and are found to be undependable for species delimitation (Sreekumar *et al*., 2006). The four barcode adopted in this study failed to demarcate *C. rivalis* from *C. metzianus*, thus the merging of *C. rivalis* and *C. metzianus* were suggested. The slender rattan present in coastal areas and sacred groves of Kerala are *C. metzianus* and they are not *C. rotang.* C. *rotang* is genetically different from those rattans present in coastal areas and sacred groves of the Wastern Ghats, and they are not reported from Kerala till now. C. *rotang* is present in the Eastern Ghats region.

4.4.2 *Calamus gamblei* species complex

The type specimens and collected samples of six species (*C. gamblei, C. lacciferus, C. neelagiricus, C. prasinus, C. dransfieldii* and *C. shendurunii*) from the Western Ghats did not show consistent nucleotide differences. The morphological characters used to distinguish them varied within and amongst species and hence considered as *C. gamblei* complex (Sreekumar and Henderson, 2014). Many characters (like structure of pinnae and rachillae and fruits) are found similar in this species complex. Since, our results suggest that the specimens considered as *C. delessertianus* are similar to *C. dransfieldii* and as this clustered together with *C. delessertianus* in the phylogenetic tree, we consider the specimens treated as *C. delessertianus* in the *C. gamblei* complex. The present sudy revealed that, six species of the complex have a similar transversion in same position in RPB2 region, indicating the close affinity of these species in the complex except *C. shendurunii*. In spite of their similarities, each species formed specific monophyletic clades in the phylogenetic tree indicating that they are different species which are closely related (Fig. 3.3). It was hence possible to precisely discriminate these species using the RPB2 region, thus resolving the complex.

4.4.3. Analyzing morphologically similar species

4.4.3.1. *C. hookerianus* and *C. pseudotenuis*

Specimen examination and field observations unveiled that morphological character states of these two species, *C. hookerianus* and *C. pseudotenuis* overlapped with each other. Based on molecular analysis using tested barcode regions, *C. hookerianus* and *C. pseudotenuis* was found to be closely related to each other without any nucleotide differences. Hence *C. pseudotenuis* and *C. hookerianus was* recommended as closely related species based on similarity of RPB2 sequences and particularly the presence of a shared 300 bp indel in *psb*A-*trn*H spacer region besides the similarities in morphology. Further analysis of more specimens collected from Sri Lanka and Karnataka will be required for suggesting the merging of the species'.

4.4.3.2. *C. delessertianus* and *C. dransfieldii*

C. dransfieldii and *C. delessertianus* were morphologically similar in many characters. *C. delessertianus* has dark green coloured leaf sheath and leaflets while pale green for *C. dransfieldii.* Since we consider this colour difference to have least taxonomic significance, we cannot consider these to be different species'. No significant nucleotide differences were found in the analysed barcoding regions (*rbc*L, *mat*K, *psb*A-*trn*H, RPB2) of both the species and the differences in morphology were insignificant, to consider it as two different species. Hence, our results suggest that the specimens hitherto considered as *C. delessertianus* (Renuka, 1999) are similar to *C. dransfieldii* and to be considered as mere morphological variation of the latter species.

CHAPTER 5

PHYLOGENETIC RELATIONSHIPS WITHIN THE GENUS *CALAMUS*

5. PHYLOGENETIC RELATIONSHIPS WITHIN THE GENUS *CALAMUS*

5.1. INTRODUCTION

Genus *Calamus* L., commonly known as 'rattans', are spiny climbing palms, belonging to the subfamily Calamoideae (Arecaeae). Due to the extreme morphological heterogeneity, they are often regarded as the ''protean'' (Uhl and Dransfield, 1987). In India, *Calamus* is the only genus of rattans distributed in all the three major phyto-geographical areas *viz*. Peninsular India, sub-Himalayas and Andaman and Nicobar Islands and are the only rattan genus predominantly distributed in evergreen, semi-evergreen and moist deciduous forests of the Western and the Eastern Ghats of peninsular India (Renuka, 2001). Indian *Calamus*, comprise of more than fifty per cent of the total palms in India (Basu, 1985).

Rheede's *Hortus Malabaricus* (1693) gives the first authentic reference on taxonomy of Indian rattans. In 1832, four new species of rattans were reported from eastern part of India by Roxburgh. Griffith in *The Palms of British East India* (1844-1845) described 11 new species of rattans and in the Hooker's *Flora of British India* (1872 - 97), Beccari's work on Indian rattans were published describing nine species of *Calamus.* Later in 1908, Beccari's classical work in the monograph on *Calamus*, reported 16 groups, of which nine groups (I, II, V-IX, XIV, XV) where of Indian *Calamus* (Table 5.1) and included 11 species from Southern parts of India. In 1931, eleven species were reported by Gamble and Fischer. The detailed description of rattans of the Western Ghats and the Andaman and Nicobar islands were provided in the taxonomic manuals with the addition of several new species of Indian rattans (Renuka, 1992 and 1995).

In Beccari's monograph of *Calamus* (1908), this genus is classified in to 16 infra-generic groups based on morphology (Table 5.1). The first groups: I-X, consists of species with leaves, which are never cirriferous. The second class include, leaves which are shortly cirriferous with diminutive leaflets at summit or very exceptionally not cirriferous, and they belongs to groups XI-XIII. Groups XIV-XV

include leaves which are distinctively cirriferous. Anomalous types are included in group XVI.

Even though Beccari (1908) categorised species into 16 groups, it was difficult to assign the constant and precise characters to each group due to lack of enough plant materials for identification. Also, it was not possible to include more species under any of the recognized groups, as they exhibited a mixture of defined characters of different groups. Another problem prevailing in the systematics of this genus is their high degree of intra-specific variability often led to treat closely related species as different. Due to this variability, it was difficult for the genus and species circumscriptions within and between *Calamus* and its related genera (Sreekumar, 2005).

Table 5.1. Species of India from Beccari's Classification (1908)

Groups	**Species**
Group 1 B	*C. flagellum* Griff.
Group 2 A	*C. longisetus* Griff., *C. thwaitesii* Becc.
Group 5 B	*C. viminalis* Willd., *C. metzianus* Schlecht, *C. rivalis* Thw. *ex* Trim. *C. pseudotenuis* Becc., *C. hookerianus* Becc., *C. rotang* L., *C. brandisii* Becc. *C. delessertianus* Becc. *C. tetradactylus*, *C. tenuis* Roxb. and *C. floribundus* Griff.
Group 6	*C. guruba* Buch-Ham.
Group 7	*C. travancoricus* Bedd.
Group 8	*C. huegelianus* Mart. and *C. gamblei* Becc.
Group 9	*C. gracilis* Roxb.
Group 14	*C. andamanicus* Kurz.
Group 15A	*C. palustris* Griff. and *C. latifolius* Roxb.
Group 16B	*C. khasianus* Becc. and *C. nambariensis*

Species reported after Beccari

1. *C. vattayila* Renuka
2. *C. prasinus* Lakshmana & Renuka
3. *C. dranfieldii* Renuka
4. *C. lacciferus* Lakshmana & Renuka
5. *C. shendurunii* Anto, Renuka & Sreekumar

6. *C. nagbettai* Fernandez & Dey

7. *C. lakshmanae* Renuka

8. C. karnatakensis Renuka & Lakshmana

9. C. stoloniferus Renuka

10. C. *neelagiricus* Renuka

Very few studies have attempted to understand phylogenetic relationship within genus *Calamus* (Kramadibrata, 1992; Baker *et al.* 2000b; Sreekumar, 2005; Senthilkumar, 2015). Preliminary studies on phylogeny of Malayan species revealed that, genus *Calamus* is paraphyletic (Kramadibrata, 1992), which was further confirmed by Baker *et al.* (2000b) using DNA sequence data. Based on ITS and 5S nrDNA analysis, genus *Daemonorops* was found nested within *Calamus* (Asmussen *et al*., 2006) which was separated from *Calamus* based on floral characters. Later on, *Daemonorops* was merged within *Calamus* by Baker (2015). Earlier, genus *Calamus* was identified as paraphyletic, four other genera i.e. *Daemonorops*, *Calospatha*, *Retispatha*, *Ceratolobus* and *Pogonotium* including the genus *Calamus* found in single clade, based on *ITS* and *rps 16* sequences (Baker *et al.*, 2000b; Dransfield *et al*., 2008). Later on, these five genera of Calaminae was merged with *Calamus* and hence considered to be monophyletic (Baker, 2015). *Calamus* is the only genus of rattans present in Western Ghats of India and hosts many endemic species of *Calamus*.

The phylogenetic analysis of its members of Western Ghats using morphology does not support as a separate geographical entity because North Eastern species as well as species from Andaman and Nicobar islands nested together with it (Sreekumar, 2005). This study was only based on the morphological characters alone, and the most vital characters used in identification were recognized as highly homoplasious in phylogenetic trees based on low-density, world-wide taxon sampling of the genus (Baker *et al*., 2000a & b; Baker and Dransfield, 2008; Dransfield *et al*., 2008). Hence phylogenetic studies based on morphology are limited by taxonomic complexities in their interpretation (Baker *et al*., 2000a; Lewis *et al*., 2000; Uhl,

1995). Molecular data plays an important role here, in order to resolve these taxonomic complexities.

Globally, very few studies have been carried out to understand the phylogeny of *Calamus*. In the Indian sub-continent the studies are still in its initial stage and more studies should be undertaken to envisage phylogenetic relationship of *Calamus* of India and related genera of three different phytogeographical regions.

The present study attempted to derive phylogenetic relationships of the genus *Calamus* of Western Ghats and its related species of sub-Himalayas and Andaman and Nicobar Islands. The study will help to find out whether genus *Calamus* of the Western Ghats have a monophyletic or paraphyletic origin. Based on the previous works discussed above, six gene regions were used in this study, plastid (*rps16, atp*H- *atp*I, *psb*A-*trn*H and *trn*L-*trn*F) as well as nuclear regions (RPB2 and ITS2) to conduct the phylogenetic relationship of *Calamus*.

5.2. MATERIALS AND METHODS

5.2.1. Sampling

A total of 31 species from genus *Calamus* were collected from three geographic zones of India, as ingroup species for the present study. This study also included representative species from other genera of rattans in India *viz*. *Kortahlsia cheb* (*Korthalsia*), *Salacca ramosiana* (*Salacca*), *Plectocomia himalayana* (*Plectocomia*), *Daemonorops rarispinosa*, *Daemonorops manii* (*Daemonorops*). Leaf samples were collected from their natural distributional zones of the Western Ghats, Andaman and Nicobar Islands and North-eastern states of India. Species list and their collected location are appended (Appendix IV).The leaf tissues were stored in silica gel for further uses. Voucher specimens were deposited in herbarium of KFRI, Peechi.

5.2.2. DNA extraction, PCR amplification 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 by using DNeasy Plant Mini Kit (Qiagen, Germany) according to the manufacturer's protocol. Six gene regions, plastid (*rps16*, *atp*H-*atp*I, *psb*A-*trn*H and *trn*L-*trn*F) as well as nuclear regions (RPB2 and ITS2) were included in this study.

Polymerase chain reaction (PCR) amplification of targeted DNA regions was performed in a PTC-100 thermocycler (BIO-RAD, India) using 2X Taq buffer (Genei, Bangalore) with 1.5 mM $MgCl_2$, 200 μM dNTPs, 10 pm of each primer, and 2U Taq DNA polymerase (Genei, Bangalore), 50-100 ng template DNA and enough distilled deionized water to give a final volume of 20 μL. PCR reactions were performed in following regime: initial denaturation at 94°C for 5 min, 30 cycles of denaturation at 94°C for 30 sec, primer annealing at specified temperatures and an extension at 72°C for 1 min followed by final extension at 72°C for 10 min. Primer information and optimal PCR annealing temperatures are provided in Table 5.2. As the standard primers for ITS2 region failed to get amplification in Calamus species, new primers were designed for the study. The amplified products were resolved in 2 per cent agarose gel and documented using a gel documentation system (Syngene, UK). PCR products were purified using Nucleospin Gel and PCR Clean-up kit (Macherey Nagel, USA) as per the manufacturer's protocol, and quantified using a spectrophotometer (Nanodrop, Thermo Scientific, USA) and then were sequenced using Sanger dideoxy method in both the forward and reverse directions at Chromous, Banglore). DNA sequences were initially aligned using default pairwise and multiple alignment parameters in Clustal X (Jeanmougin *et al.*, 1998) and then rechecked and adjusted manually as necessary using MEGA v.6.0 (Tamura *et al*., 2013). Gaps were positioned to minimize nucleotide mismatches and treated as missing data in phylogenetic analysis.

5.2.3. Phylogenetic analysis

Phylogenetic analysis for the individual data matrix and combined data sets were conducted by Bayesian Inference (BI) using programme Mr Bayes 3.2.2 (Ronquist and Huelsenbeck, 2003) respectively. The jModelTest 2.1.4 (Posada and Buckley, 2004) was used to select a best fit model of nucleotide substitution. Bayesian

Markov Chain Monte Carlo (MCMC) algorithm was run 2,000,000 generations with four chains each and sampling trees every 100 generations. Effective sample size (ESS) was checked using Tracer v.1.5 (Rambaut and Drummond, 2009) and first 25 per cent of sampled trees was discarded as burn-in. The ESS value of >200 for the two combined runs and their overlapping marginal densities for two runs were examined and final trees were visualized in Figtree v.1.3.1 (Rambaut, 2009).

Table 5.2. Primers for the phylogenetic gene regions and their PCR reaction conditions

Phylogenetic gene regions	**Primer**	**Primer sequence 5'-3' (Reference)**	**Annealing temperature**
*trn*L-*trn*F	*Trn*L *trn*F	CGAAATCGGTAGACGCTAG ATTTGAACTGGTGACACGAG (Taberlet *et al*., 1991)	50°C -45 sec
*atp*H- *atp*I	*atp*H *atp*I	CCAAYCCAGCAGCAATAAC TATTTACAAGGTATTCAAGCT (Shaw *et al*., 2007)	60°C -50 sec
rps16	*rps16* F *rps16*R	GTGGTAGAAAGCAACGTGCGACTT TCGGGATCGAACATCAATTGCAAC (Oxelman *et al*., 1997)	60°C -45 sec
RPB2	RPB2F RPB2R	CAACTTATTGAGTGCATCATGG CCACGCATCTGATATCCAC (Rونçal *et al*., 2005)	58 °C -40 sec.
ITS2	5 F 5R	GCCGAGGGCACGCCTGCCT GATTCTCAAGCTGGGCTAT	56.5°C -35 sec
*psb*A- *trn*H	*trn*H *psb*A	GTWATGCAYGAACGTAATGCTC CGCGCATGGTGGATTCACAATCC (Kress *et al*., 2005)	59 °C -50 sec.

5.3. RESULTS

The data matrix used for multi-gene analysis was obtained from *rps16*, *atp*H-*atp*I, *trn*L-*trn*F, RPB2 and ITS2 sequences. *psb*A-*trn*H was excluded for analysis due to its high intra-specific variation (Fig. 5.1.). Generated sequences were deposited in GenBank public domain (https://www.ncbi.nlm.nih.gov/genbank) and accession numbers are appended here (Appendix VI). The aligned matrix was 3,894 characters long of which 3,272 were conserved, 483 were variable but parsimony-

uninformative and 161 were parsimony informative. Models selected for each region as the best-fit models according to Akaike Information Criterion (AIC) differences were given in Table 5.3.

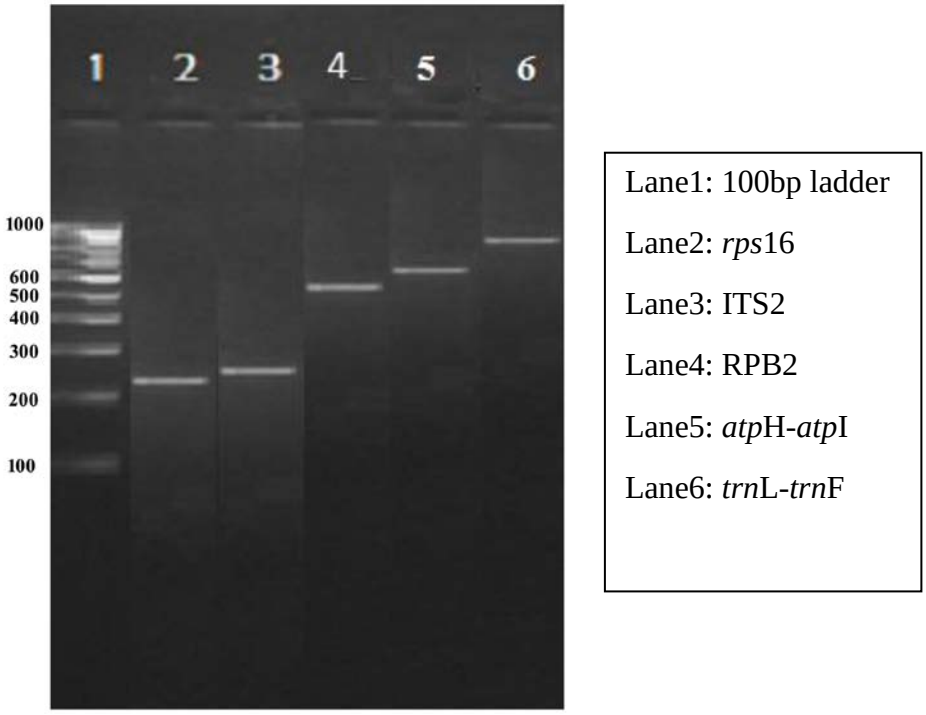

Fig. 5.1. PCR amplified products of the selected barcode regions

Phylogenetic tree construction using BI with plastid as well as nuclear regions separately showed incongruent topology. Phylogenetic analysis of concatenated nucleotide characters produced a BI tree with good posterior probability support (Fig. 5. 2). In the analysis, concatenated cpDNA and nuclear data set was partitioned so that each DNA region was analyzed using its own best-fit model.

Korthalsia, *Salacca* and *Plectocomia* stood as well separated outgroups in the phylogenetic tree, while *Daemonorops* was found clustered within *Calamus* clade. The constructed dendrogram consisted of eight major clusters in which *C. thwaitesii*, *C. flagellum* and *C. longisetus* clustered together to form first clade (PP=1.00) and second clade comprises of *C. nagbettai* and *C. andamanicus* (PP=1.00). Clade III for *Daemonorops* was found in between *Calamus* species (PP=1.00). Five species of *C. gamblei* complex clustered together in clade VI, except few including *C. gamblei*.

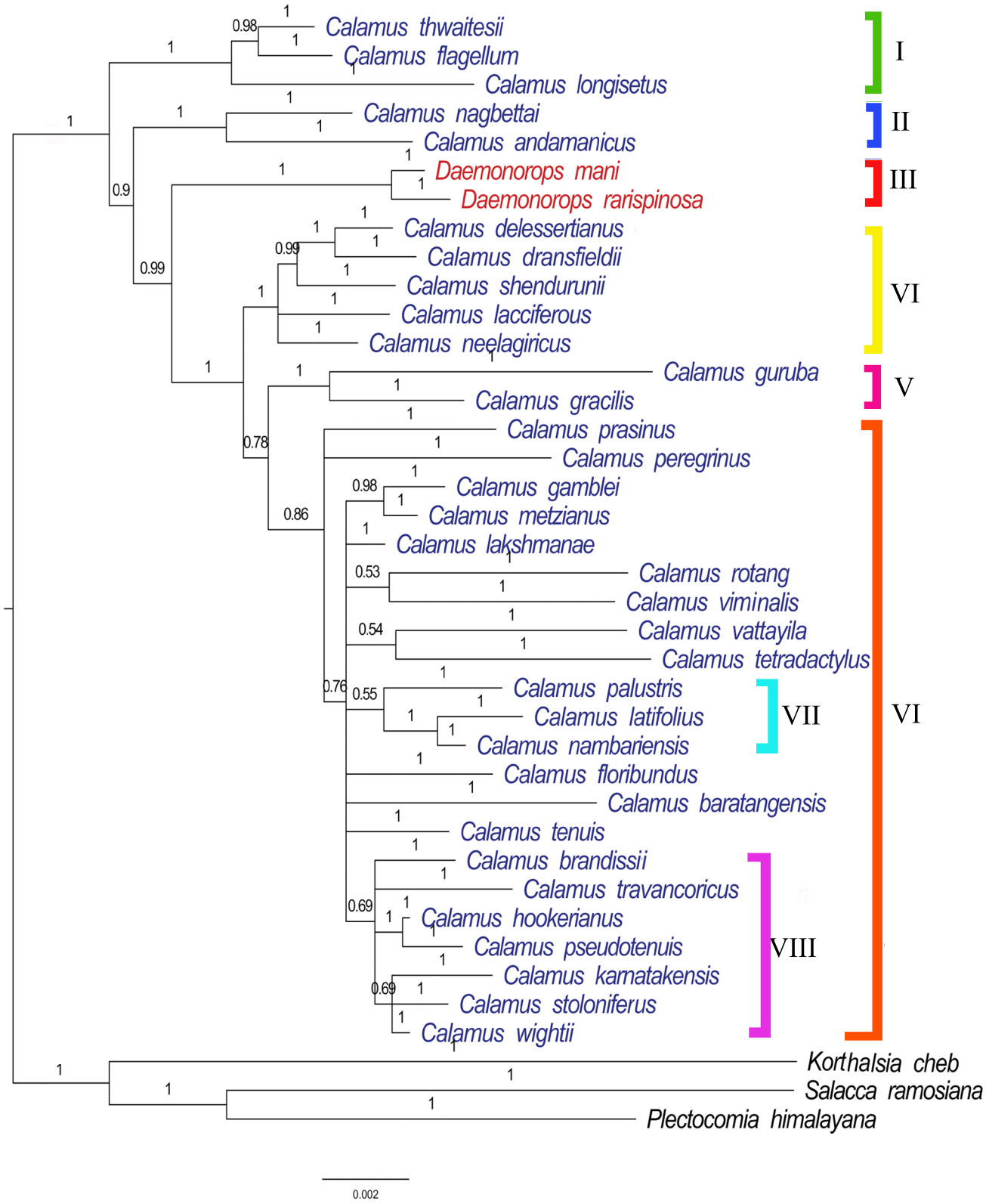

Fig. 5.2. Bayesian consensus tree using chloroplast (*rps16*, *atp*H- *atp*I, and *trn*L-*trn*F) and nuclear (RPB2 and ITS2) regions in the genus *Calamus*.

Even though the relationship between some *Calamus* species remains ambiguous, most of them are clearly distinct in this phylogenetic analysis.

Table 5.3. Sequence characteristics and models of gene regions

Gene regions	**Sequence length**	**Conserved sites**	**Variant sites**	**Parsimony Informative**	**Singleton sites**	**Models Selected**
rps16	256	245	6	2	0	F81
*atp*H-*atp*I	760	598	84	20	57	GTR + I+ G
*trn*L-*trn*F	816	765	43	8	65	HKY + G
RPB2	752	540	190	61	129	GTR + G
ITS2	351	246	98	40	56	GTR + I+ G

5.4. DISCUSSION

Species of *Calamus* are morphologically diverse, especially with regards to their growth habits, stem texture, climbing organs (cirrus and flagellum), leaflets arrangements and inflorescence architectures. Delimitation of Himalayan species that are close with those of South East Asia are often problematic (Henderson and Henderson, 2007). In Beccari's (1908) classification of *Calamus* based on the morphological similarities of the species, he mentioned that this classification was far from satisfactory because of lack of enough materials representing both vegetative and reproductive parts. Hence there can be differences from his classification.

In Beccari's classification (1908), *C. longisetus* and *C. thwaitesii* were grouped together and, similarly in this study too, they were found to be together in a clade (Fig 5.2). They form the first subclade in the genus *Calamus* along with *C.*

flagellum. *C. thwaitesii* (Western Ghats) was closely related to C. *flagellum* of North Eastern states which was further linked to *C. longisetus* found in Andaman Islands. They were flagelliferous, clustering, large diameter rattans. They were ecirrate and profusely armed with spines. Their fruits also have similar ovoid structure. Because of their morphological similarities they were grouped together in morphological classification and Sreekumar (2005) study, also group three together based on morphologicy. In spite of their different geographical isolation, they were grouped together in this study based on molecular analysis, as well as showed morphological similarities. The endemic species of Western Ghats, *C. nagbettai* also showed close affinities with *C. andamanicus* of Andaman Islands based on the presence of cirrate leaf (Clade II). The molecular analysis grouped them together as well and they formed the second clade of this genus. These large diameter rattans were similar with ovoid fruits with spines in vertical rows, brown with dark brown margin and have similar cup-shaped involucre. Genus *Daemonorops* formed the third clade which was found nested with in the genus *Calamus*, thence supporting the view (Baker, 2015) that it is to be merged with *Calamus*. *Daemonorops* was similar to *Calamus* in many morphological characters. Floral characters were however different like, the partial inflorescences which were embedded within the large primary spathe and secondary spathe. Molecular analysis in our study also supports the merging of *Daemonorops* with *Calamus* (Fig. 5.2).

The fourth cluster consisted of five species *viz. C. gamblei, C. dransfieldi, C. delessertianus, C. lacciferus and C. shendurunii*. They belong to the "*C. gamblei* species complex" distributed in the Western Ghats of peninsular India (Sreekumar and Henderson, 2014). The species here did not consistently differ from one another and the morphological characters used to distinguish them vary within as well as amongst species. All species in this complex share similar morphological characters such as the presence of long, parallel, apical pinnae joined at their bases, pinnae with long spines on the adaxial side of veins recurved rachillae with pistillate dyads arranged in alternate, non-opposite rows so that one side of the rachillae is without flowers, pistillate dyads, born on a distinct pedicel and yellow fruits with raised scales. Other species *viz. C. gamblei* and *C. prasinus* of this complex stood out from

this cluster, in a separate clade. Beccari had placed *C. gamblei* in the Group VIII and later Renuka (2010) placed other rattans of the present *C. gamblei* species complex here**.** But in the present study *C. gamblei* is found to be more related with *C. metzianus*. This has to be re-evaluated using other gene regions. Likewise morphologically dissimilar *C. vattayila* and *C. tetradactylus* which are allied together (PP=0.54) in this study, has to be re-evaluated.

The Indo- South East Asian species *C. gracilis* and *C. guruba* were found together in a cluster in clade V (PP = 1). Morphologically they show similarities with the presence of flagella, prominent knee and ocrea. They are clustering medium diameter rattans, present in the North eastern states of India and Bangladesh, while *C. guruba* is also distributed in Myanmar and Thailand as well. On the other hand, in Beccari's classification, *C. gracilis* and *C. guruba* were grouped in two different groups, group VI and IX. *C. peregrinus* a native of South East Asia is found together with *C. prasinus*, which is endemic to Western Ghats. Both are solitary, high climbing rattans with flagella and a conspicuous knee. Both have globose fruits.

Beccari treated *C. latifolius* and *C. nambariensis* in the two different groups of group XV, based on their morphological affinities. They are clustering, medium diameter, non- flagelliferous rattans, with cirrate leaves, triangular spines and globous fruits. In this study, they remained together as a single clade with PP=1 (clade VII). They were distributed in the Indo- South East Asian countries. *C. palustris* which was treated in the same group with *C. latifolius* by Beccari (1908), remained sister to *C. latifolius* and *C. nambariensis*. *C. palustris* is also a clustering medium diameter rattans with cirrus which are present in south East Asian countries and Andaman Nicobar Islands, while *C. latifolius* and *C. nambariensis* are present in North eastern parts of India.

Seven species of Western Ghats *viz*. *C. hookerianus, C. pseudotenuis, C. stoloniferus, C. karnatakensis, C. wightii, C. brandisii* and *C. travancoricus* formed a single clade (Clade VIII). This clade also consists of the newly reported species by Renuka, *C. karnatakensis* and *C. stoloniferus*. *C. karnatakensis, C. stoloniferus* and *C. pseudotenius,* which are characterized by cluster forming habit with homogenous

endosperm. They are clustering, flagelliferous, medium diameter rattans. Knee is present in all the seven species, while ocrea is present in all except *C. brandisii* and *C. travancoricus*. Beccari placed *C. viminalis, C. rotang, C. metzianus, C. pseudotenuis, C. hookerianus, C. delessertianus* and *C. brandisii* in the fifth group. But in this study, members of this group showed different clustering pattern in the dendrogram. Even though they were in different clades, *C. travancoricus*. *C. brandisii* and *C. travancoricus* were found allied together in this study supporting the conclusion of Sreekumar (2015). It is hard to conclude on the phylogenetic relationships between some of the *Calamus* species (*C. tenuis, C. floribundus, C. baratangensis* etc.) sampled here which are not resolved, pointing to the need for more sequence data to further clarify their relationship.

The pseudo-rattan genus *Salacca* was found closely together with *Plectocomia*, than with any other genus with PP=1. They together with *Korthalsia* form an out group to the *Calamus* clade. The knee, swelling of the leaf sheath at the junction of the petiole base (Tomlinson, 1962; Dransfield, 1978) is absent in *Korthalsia* and *Plectocomia* (Renuka *et al*., 2010).

In this study, some of the clades of *Calamus* remain ambiguous, with low posterior probability, requiring more gene regions in order to resolve the problems. Of the five gene regions were included in this study, *rps16* showed the least intraspecific variation and we recommend it only for analysing taxonomy at the higher level. While other regions like RPB2, *atp*H- *atp*I, ITS2 and *trn*L-*trn*F were successful at lower level phylogenetic studies. Because of the very high intra specific variations, *psb*A-*trn*H was eliminated from this study.

All the species of Western Ghats were flagelliferus, except, *C. nagbettai* which was cirrate. The cirrate character is more predominant in South East Asia which has a very high rattan diversity. Senthilkumar (2015) considered cirrate nature to be an ancestral trait of rattans. But in our studies the early diverged clade of *C. thwaitesii*, *C. longisetus* and *C. flagellum* were flagelliferus, indicating that cirrus need not be an ancestral trait. The clade of the cirrate species, *C. latifolius, C. palustris* and *C. nambariensis,* which was in the midst of flagellate rattans in the phylogenetic tree,

is considered to be recently diverged and therefore support the view that cirrate nature need not be the ancestral character. This character could have arisen in more than one occasion in this group. This study suggests that cirrate trait in Indian rattans may have appeared because of convergent evolution. The cirrate nature appears to be a nominal variation in the leaf morphology unlike the flagellum, lends support to this view.

Calamus species from North East India, Andaman and Nicobar Islands and South East Asia were seen nested within the species of Western Ghats in the phylogenetic tree. This makes the *Calamus* species of Western Ghats non- monophyletic. This may have resulted from the convergent evolution of characters or habit etc. Non-monophyly can also result from the recent dispersal of *Calamus* species from other regions in to Western Ghats.

This study analysed the importance of both morphological and as well as molecular characters for resolving the phylogenetic relationship between the *Calamus* species. A better classification of *Calamus* of Western Ghats (Including other Indian species) were obtained, in which the genus *Calamus* was grouped into six different groups. Morphological similarities between them were analysed and found that most of the clades in molecular analysis also showed morphological affinities. The present study concludes that, based on the five gene regions used, the genus *Calamus* is having a monophyletic origin. *Daemonorops* were found nested within the *Calamus* supporting the merging of this genus with *Calamus*. Some of the species yet remain to be unresolved and hence, we suggest that additional data is required in order to get a robust phylogeny of *Calamus.* Thorough sampling throughout India is required in order to get a clear picture of the phylogenetic relationship within Indian *Calamus.*

CHAPTER 6

BIOGEOGRAPHY AND PHYLOGENETIC RELATIONSHIP OF SUBFAMILY CALAMOIDEAE

6. BIOGEOGRAPHY AND PHYLOGENETIC RELATIONSHIP OF SUBFAMILY CALAMOIDEAE

6.1. INTRODUCTION

Calamoideae is the largest subfamily within Palmae (Arecaceae) comprises 22 genera and 650 species. They are characterized by the presence of reflexed, overlapping scales on pericarp, which gives the fruit its unique and distinctive appearance. Calamoideae constitutes tree palms, acaulescent palms and climbing palms (Dransfield, 2000).

Calamoideae is classified into three tribes, namely Eugeissoneae, Lepidocaryeae and Calameae and further into nine subtribes (Uhl and Dransfield, 1987; Dransfield, 2000) (Table 6.1). The members of Calamoideae were concentrated exclusively in the Old World tropics while four genera such as *Mauritia, Mauritiella, Lepidocaryum* and *Raphia* are present in the New World Tropics like Latin America, Mexico, Chile, Argentina and Caribbean. Three genera, *Laccosperma*, *Eremospatha* and *Oncocalamus* coming under subtribe Ancistrophyllinae are endemic to Africa (Sunderland, 2012), While the members of Calameae tribe were confined to South and South East Asia except few which found route to Australia. Tribe Eugeissoneae was present in South East Asia only.

Even though monophyly of subfamily Calamoideae is strongly supported by phylogenetic analyses using morphological and molecular data (Uhl *et al.*, 1995; Baker *et al.*, 1999b; Asmussen *et al.*, 2000), relationships within this subfamily are difficult to interpret due to the wide spectrum of existing morphological diversity (Baker *et al*., 2000a). In spite of the significant morphological differences present between African rattan genera (*Eremospatha*, *Laccosperma*, *Oncocalamus*) and genus *Calamus*, African rattans were considered to be subgenera within the South East Asian genus *Calamus* (Mann and Wendland, 1864). Subsequently, they were raised to the rank of genus (Sunderland, 2012), and later grouped into a single subtribe, Ancistrophyllinae (Baker *et al*., 2000a; Dransfield *et al*., 2008).

Table 6.1. Phylogenetic classification of subfamily Calamoideae (Dransfield *et al*., 2005)

TRIBE	SUBTRIBE	GENUS
Eugeissoneae		*Eugeissona*
Lepidocaryeae	Ancistrophyllinae	*Eremospatha* *Laccosperma* *Oncocalamus*
	Raphiinae	*Raphia*
	Mauritiinae	*Mauritia* *Mauritiella* *Lepidocaryum*
Calameae	Salaccinae	*Eleidoxa* *Salacca*
	Pigafettinae	*Pigafetta*
	Metroxylinae	*Metroxylon*
	Korthalsiinae	*Korthalsia*
	Plectocomiinae	*Myrialepis* *Plectocomiopsis* *Plectocomia*
	Calaminae	*Calamus* *Daemonorops* *Calospatha* *Ceratolobus* *Pogonotium* *Retispatha*

Similarly, several taxonomic confusions prevail in subtribe Calaminae, which comprises of six genera. The phylogenetic analysis revealed that the Calaminae is not monophyletic (Baker *et al*., 2000b). They also suspected the generic delimitation of several genera such as *Calospatha, Ceratolobus*, *Daemonorops*, *Pogonotium* and

Retispatha, and later, everything was merged with *Calamus* based on many studies (Baker and Dransfield, 2008; Baker, 2015). The presence of persistent tubular bracts is the main character which separated *Calamus* from related genera (Baker and Dransfield, 2008; Baker, 2015). Hence it is clear that, the morphology based phylogenetic studies are limited by taxonomic complexities in their interpretation since morphological data alone were not sufficient to resolve the relationships between Calamoideae (Uhl, 1995; Lewis *et al*., 2000; Baker *et al*., 2000b). Hence there is need for a molecular data to understand the evolutionary relationships of the species with related genera. Baker *et al*. (2000 a), analysed the phylogenetic relationship of 22 genera of subfamily Calamoideae using the combination of *rps16* intron and nrITS, and obtained a well-supported phylogenetic tree. However these regions appear to have limited value in palm phylogenetics because of the lack of homogeneity among repeats within individual palm genomics (Baker *et al*., 2000 a & b).

Biogeographical distribution of this subfamily remain complex and need to be explored. Biogeographically events above genus level in palms were inferred by comparing ancestral area reconstruction (Baker and Couvreur, 2012). The dispersal of different genera from its inferred ancestral area to its present distribution is still vague. The rich species diversity coupled with a wide distribution range of the genus *Calamus* in tropical rainforests leads to the need for addressing biogeographical questions and to investigate phenomena such as rapid radiations, shifts in diversification rates, character evolution and the evolution of key innovations, but also poses major taxonomic challenges. The biogeographical studies in Indian subcontinent gain its importance as it has a unique geological past (Briggs, 2003). Molecular studies undertaken regarding the dating of *Calamus* and the present distribution of the species and its radiation into India are limited.

The present study mainly focussed on Indian *Calamus*, as Peninsular India has a great importance in biogeography. It was a part of Gondwanan supercontinent, about 200 million years ago (mya) (Mani, 1974; Briggs, 1989 and 2003; Hedges, 2003). From Gondwana, the Indo-Madgascar plate drifted away from East Africa, during

Jurassic period (158-160 mya), followed by the separation of peninsular India from Madagascar (84-96 mya) (Briggs, 2003). Previous studies reported that, after the separation from Madagascar, Indian plate allegedly underwent a period of isolation for about 30- 40 my before the collision (40-50 mya), allowing the highly diverse and endemic biota to develop, before colliding with the Eurasian plate around 40-50 mya. During that time period India was considered to act as a biotic ferry, which exchanged biota from Africa into Asia and from Asia into India (Out-of-India and Out-of-Asia theories). In contrast to 'Biotic ferry theory', in which broad land bridge connections existed between drifting India and Africa or island arcs between India and Asia or India and Africa, allowing the exchange of biota, several models are also proposed (Chatterjee and Scotese, 1999; Briggs 2003).

The systematic positioning of Indian rattan in Calamoideae was also need to be analysed. The origin and distribution of *Calamus* in India are not completely addressed till now. In a study of climbing palms of India, Senthilkumar (2015), reported the dispersal of *Calamus* from South East Asia to India during late Oligocene, which failed to address the reasons for the presence of fossil records prior to this time period. Hence, in addition to phylogenetics, the biogeographic distribution pattern of these groups found lacking. In order to provide a better understanding in phylogenetics as well as biogeography of rattan resources of India, molecular studies using combination of both chloroplast and nuclear sequences are undertaken here.

6.2. MATERIALS AND METHODS

6.2.1. Taxon Sampling

Sampling consists of 35 Calamus species, single representative species from other four genera of rattans of India, (*Korthalsia*, *Salacca*, *Plectocomia* and Daemonorops) were collected from three geographic zones of India (Appendix IV). Seventeen species from subtribe Ancistrophyllinae and single representative species from subtribes, Calaminae (*Calamus aruensis* Becc.), Metroxylinae (*Metroxylon salomonense* (Warb.) Becc.), Mauritiinae (*Mauritiella armata* (Mart.) Burret,

Mauritia flexuosa L.f.; Lepidocaryum tenue Mart.), Raphiinae (Raphia palma-pinus (Gaertn.) Hutch, Raphia hookeri G. Mann and H. Wendl.) and tribe Eugeissoneae (*Eugeissona tristis* Griff.) were retrieved from NCBI (Faye *et al*., 2014). Thirty six species of *Calamus* from different geographical zones were included in this study. *Ceroxylon quindiuense* was chosen as outgroup. The sequences retrieved from NCBI public domain were appended here (Appendix V).

6.2.2. DNA extraction, PCR amplification 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). Polymerase Chain Reaction (PCR) amplification of targeted DNA regions was performed in a PTC-100 thermocycler (BIO-RAD, India) using 2X Taq buffer (Genei, Bangalore) with 1.5 mM $MgCl_2$, 200 μM dNTPs, 10 pm of each primer, and 2U Taq DNA polymerase (Genei, Bangalore), 50-100 ng template DNA and enough distilled deionized water to give a final volume of 20 μL. Primer information and standardised PCR conditions were mentioned in Table 6.2. The amplified products were resolved in 2 per cent agarose gel and documented using a gel documentation system (Syngene, UK). PCR products were purified using a Nucleospin Gel and PCR Clean-up kit (Macherey Nagel, USA) and quantified using a spectrophotometer (Nanodrop, Thermo Scientific, USA). The purified PCR products were sequenced in both forward and reverse directions using Sanger dideoxy chemistry (Chromous, Banglore).

Table 6.2. Phylogeny primers and PCR reaction conditions

Phylogenetc gene region	**Primers**	**Primer sequence 5'-3' (Reference)**	**Annealing temperature**
*atp*H- *atp*I	*atp*H *atp*I	CCAAYCCAGCAGCAATAAC TATTTACAAGGTATTCAAGCT (Shaw *et al*., 2007)	60°C - 50 sec
RPB2	RPB2F RPB2R	CAACTTATTGAGTGCATCATGG CCACGCATCTGATATCCAC (Ronçal *et al*., 2005)	58°C - 40 sec.

6.2.3. Phylogenetic analysis

Raw data chromatogram were initially aligned using default pairwise option and multiple alignment parameter in Clustal X (Jeanmougin *et al.,* 1998) and adjusted manually using MEGA v.6.0 (Tamura *et al*., 2013). Gaps were positioned to minimize nucleotide mismatches and treated as missing data in further analysis. The trees were rooted using out group *Ceroxylon quindiuense.* Phylogenetic analysis for individual data matrix and concatenated sequences were conducted by Bayesian Inference using Mr Bayes v.3.2.2 (Ronquist and Huelsenbeck, 2003). jModelTest v.2.1.4 (Posada and Buckley, 2004) was used to select a best fit model of nucleotide substitution prior to analysis under Akaike Information Criterion. Bayesian Markov Chain Monte Carlo (MCMC) algorithm was run 2,000,000 generations with one cold chain and one heated chain, starting from random trees and sampling trees every 100 generations. After the run, 25% trees were considered as burn-in and discarded. The convergence of molecular evolutionary parameters was examined using Tracer v.1.5 (Rambaut and Drummond, 2009) with ESS (effective sample size) of >200 for two runs together and examined the overlapping marginal densities for both runs. The constructed final tree was visualized in Figtree v.1.3.1 (Rambaut, 2009).

6.2.4. Dating phylogenetic tree

Divergence time was estimated using BEAST v.2.5.0 under relaxed molecular clock with uncorrelated lognormal distribution (Drummond and Rambaut, 2007). Bayesian approach was employed to estimate the divergence times of subfamily Calamoideae. Data sets were partitioned and determined the best substitution model as GTR based on AIC jModelTest v.2.1.4 (Posada and Buckley, 2004). Tree prior speciation Yule model was used based on ML inferred ultrametric starting tree in RPB2 and *atp*H-*atp*I sequences. Two calibration points were used for Calamoideae crown group, 76.6 mya (Lepidocaryeae and Calameae) and 46 mya (Calaminae) based on previous study (Baker and Couvreur, 2012). Monte Carlo Markov Chain was run for 10^9 generations and sampled every 1,000 generations. Effective Sample Size (ESS) was

examined in Tracer v.1.5 (Rambaut and Drummond, 2009) to ensure value exceeded 200 and tree files from two runs were combined using Log Combiner v. 2.4.4. Twenty percent of trees were removed as burn-in and resulting trees were summarized with Tree Annotator v. 2.4.4 and was visualized in Figtree v.1.3.1 (Rambaut, 2009).

6.2.5. Ancestral area reconstruction

The reconstruction of ancestral areas in a phylogeny is essential to understand the biogeographic diversification history of a lineage thereby enabling to infer the original location and dispersal routes of the organisms. For the reconstruction of ancestral areas, Calamoideae was divided into five regions, based on biogeographical distribution of studied species. Five areas includes India and Sri Lanka (A), Eurasia - to Wallace's Line, including Andaman and Nicobar islands (B), Africa (C), Australia (D) and South America (E). Each genera was assigned to one or more geographical areas based on previous information on genus distributions (Dransfield *et al*., 2008). Statistical Dispersal–Vicariance Analysis (S-DIVA) (Yan *et al*., 2010), Bayesian Binary MCMC Method (BBM) (Ronquist and Huelsenback, 2003) and Dispersal-Extinction-Cladogenesis (DEC) implemented in RASP v.3.1 (Yu *et al*., 2011b) were employed to construct ancestral area of Calamoideae. Analyses were based on the MCC tree from BEAST analyses. The frequencies of an ancestral range at a node in ancestral reconstructions are averaged over all trees in both methods (Yan *et al*., 2010). S-DIVA could not include in further analysis due to the presence of polytomies. The maximum number of areas was kept as 4 in BBM method. BBM analysis was conducted with MCMC chains were run simultaneously for 50, 00,000 generations and the state was sampled every 100 generations. Fixed JC +G (Jukes-Cantor+ Gamma) was used as model for BBM analysis.

Dispersal-Extinction-Cladogenesis (DEC) analyses is based on DEC models which require information about a single ultrametric-dated phylogeny and distributional information of extant species, Dispersal cost was set, where probability values of dispersion between areas were assigned from 0.0001 to 1.0, taking into account the

geographic closeness between them. DEC analysis was run on all of the 21, 101 trees from MCMC output.

6.3. RESULTS

Multiple sequence alignment (MSA) was performed in the anlaysed sequences using MEGA v.6.0. The concatenated sequences of *atp*H - *atp*I and RPB2 region in 63 species had a total size of 1514 bp with species specific indels were present in some species. All the generated sequences were deposited in GenBank (https://www.ncbi.nlm.nih.gov/genbank) and the details are appended (Appendix VI).

6.3.1. Phylogenetic analysis

Phylogentic tree using Bayesian Inference with good posterior probability (PP) values were obtained using concatenated *atp*H-*atp*I and RPB2 (Fig. 6.1). Subfamily Calamoideae were found as three clades representing three tribes, Lepidocaryeae, Eugeissoneae and Calameae (PP=1.00). Tribe Lepidocaryeae was further separated into three clades, I, II and III. African rattan subtribe Ancistrophyllinae comprising three genera, *Laccosperma*, *Eremospatha* and *Oncocalamus* formed a monophyletic clade I (PP=0.98). Genera *Mauritia, Mauritiella, Lepidocaryum* formed clade II while *Raphia* formed as clade III. *Eugeissona tristris*, taken as a representative for tribe Eugeissoneae showed sister relationship with Calameae and Lepidocaryeae (PP=1.00). Tribe Calameae formed as a monophyletic clade with five subtribes Calaminae, Korthalsinae, Salaccinae, Plectocominae and Metroxilinae. Salaccinae formed a sister clade with Plectocominae (PP=1.00) while genus *Daemonorops* found nested within genus *Calamus* (PP=1.00). African clade Ancistrophyllinae formed as distant clade with Asian genus *Calamus*.

6.3.2. Dating phylogenetic tree

The chronogram constructed based on concatenated sequences of RPB2 and *atp*H-*atp*I region estimated the divergence time of Calamoideae as 87 mya during

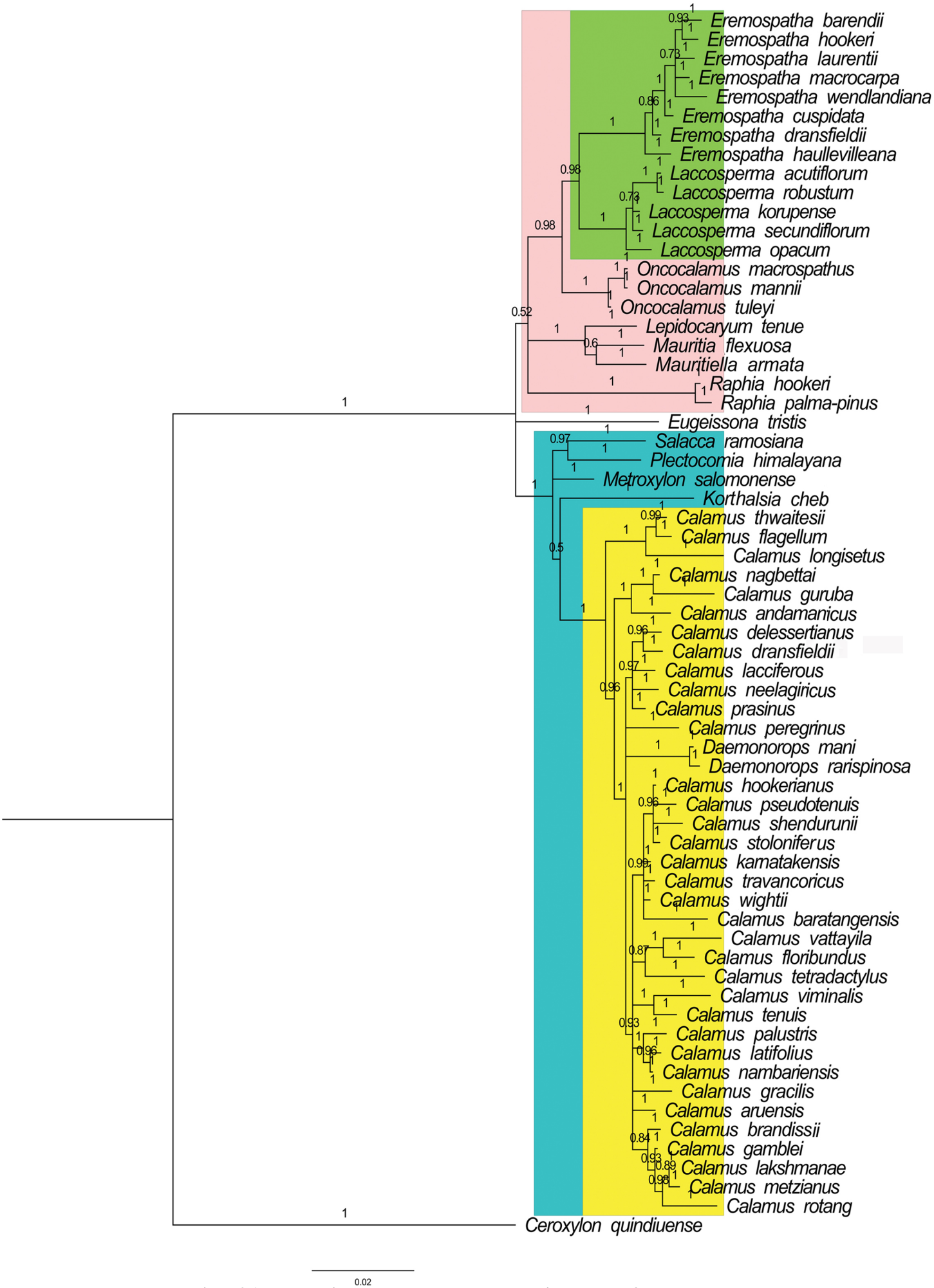

Fig. 6.1. Bayesian consensus tree using RPB2 and *atp*H-*atp*I in the sub family Calamoideae.

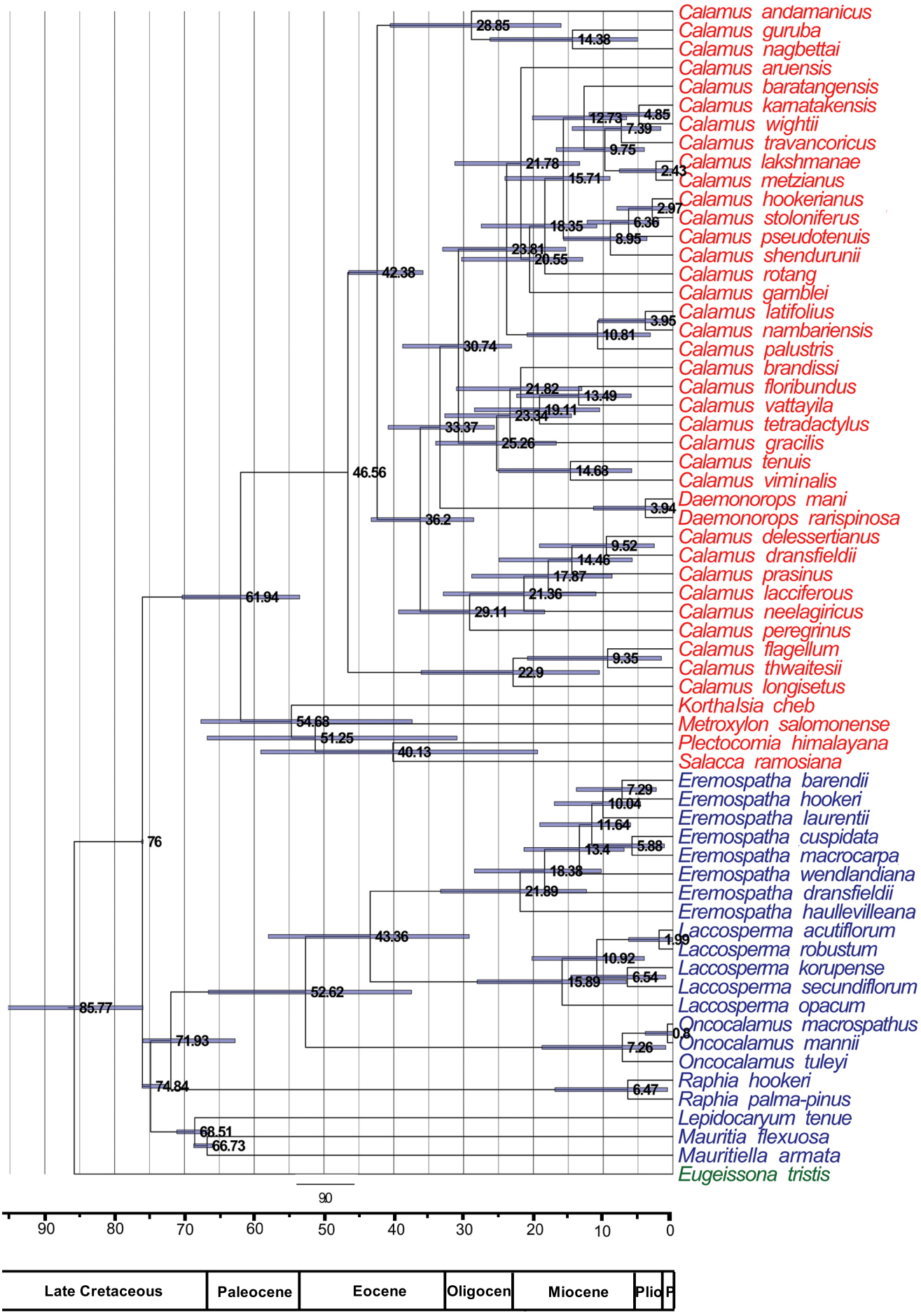

Fig. 6.2. Chronogram based on Bayesian approach of the subfamily Calamoideae. Geological epoch is shown below the tree.

Cretaceous period (Fig. 6.2). The diversification of Eugeissoneae clade from remaining groups of Calamoideae was estimated as 85.7 mya (95% HPD 75.9-111.7 mya). A split happened between tribes Lepidocaryeae and Calameae at 76 mya.

Diversification of subtribe Lepidocaryae clade happened during the late upper Cretaceous period 74.8 mya (95% HPD 72.5-76 mya). Divergence time of Calameae was estimated during Paleocene - 61.9 mya (95% HPD 53.4-70.3 mya) and Daemonorops diverged from the genus *Calamus* in Oligocene - 30.7 mya (95% HPD 25.6-40.8 mya). Within genus *Calamus*, *C. thwaitesii, C. longisetus* and *C. flagellum* were the oldest diverged species (46.5 mya) and diversification within this clade occurred around 22.9 mya. The most recently evolved species were *C. metzianus* and *C. lakshmanae* (2.4 mya).

6.3.3. Ancestral area reconstruction

Ancestral area reconstruction using BBM and DEC identified Asia as the origin of distribution of Calamoideae (Node 125, PP=1.00). This method suggests a complex biogeographical history in which dispersal and vicariance had a vital role in shaping the current distribution pattern of Calamoideae. BBM postulates 41 dispersals and five vicariance (Fig. 6.3) while DEC analysis showed five vicariance and 30 dispersal events to explain the current distribution of Calamoideae (Fig 6.4). Majority of the dispersal were in Calameae clade. Both BBM and DEC showed almost similar results. The basal node 125 suggested that the ancestors of Calamoideae originated in B (Asia excluding India) (Fig. 6.3, Fig 6.4.). Two dispersal events and a vicariance at node 24 splits into African and Asian lineages of Calamoideae. At node 24, Calamoideae splits into two clades, one formed as Asian Calameae and another leads to African Lepidocareae. Node 123 represented by Lepidocaryae suggested its early dispersal from Asia of to Africa. Two dispersal events and a vicariance evident at this node, resulting in African and South American lineages. Node 103 represents members of tribe Calameae and its possible ancestor range is B (Asia), suggested a dispersal event. Node 101 revealed a dispersal and vicariance in which one descendant (*Metroxylon* sps.) dispersed into Australia and another lineage underwent *in situ* diversification in Indo - Southeast

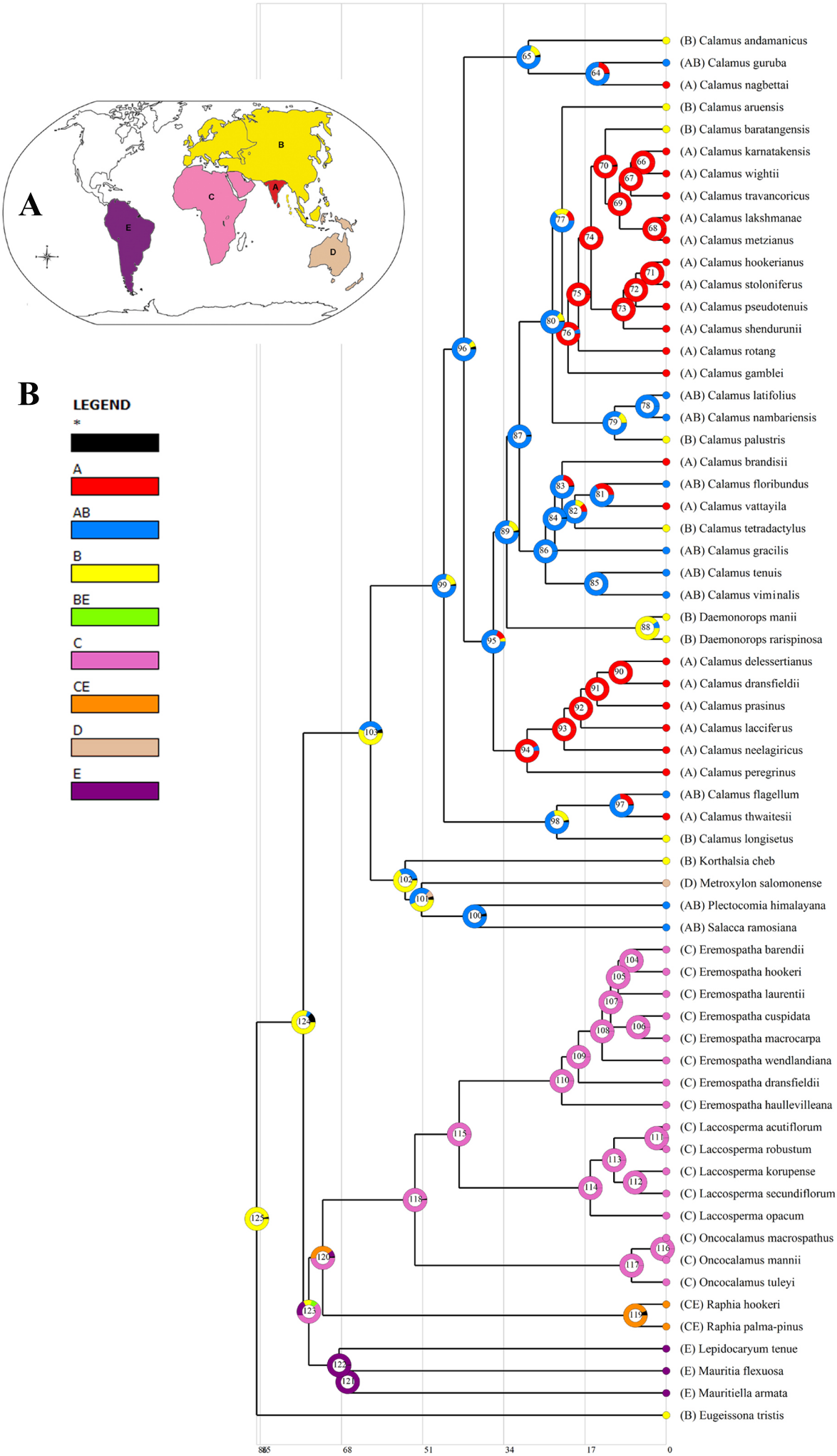

Fig. 6.3. Graphical output from Bayesian Binary MCMC analyses (exported from RASP) [A] Biogeographical regions: A (India and Sri Lanka), B (Eurasia - to Wallace's Line, including Andaman and Nicobar islands), C (Africa), D(Australia) and E (South America) [B] Colour keys to represents possible ancestral ranges at different nodes, * (Undefined ancestral ranges).

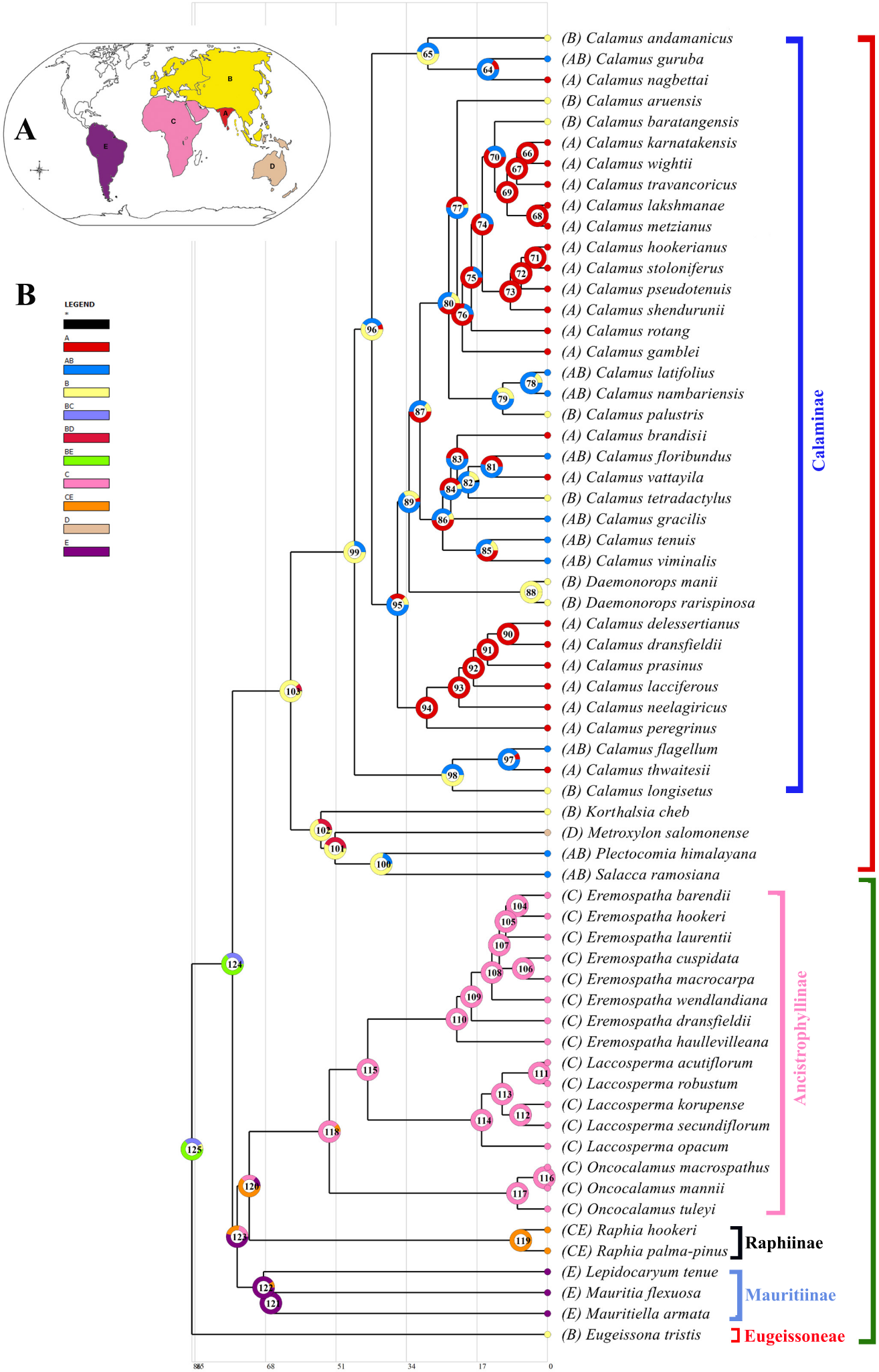

Fig. 6.4. Graphical output from DEC analyses (exported from RASP) [A] Biogeographical regions: A (India and Sri Lanka), B (Eurasia - to Wallace's Line, including Andaman and Nicobar islands), C (Africa), D(Australia) and E (South America) [B] Colour key to represents possible ancestral ranges at different nodes, * (Undefined ancestral ranges).

Asia. Numerus dispersal events were happened in this lineages represented as node 89 consists of genus *Calamus* of the Western Ghats, North Eastern Himalayas, Andaman and Nicobar islands.

6.4. DISCUSSION

Phylogenetic tree reconstruction of Calamoideae subfamily using *atp*H-*atp*I and RPB2, clearly separated three tribes into three monophyletic clades as supported by the traditional morphology based classification represented as single tribe (Dransfield *et al*., 2005). *Ceroxylon quindiuense* of subfamily Ceroxyloideae formed a separate outgroup from Calamoideae. Clade I consists of tribe Lepidocaryeae comprising three subtribes, Ancistrophyllinae, Raphiinae and Mauritiinae. Subtribes, Ancistrophyllinae and Raphiinae were considered earlier under tribe Calameae (Uhl and Dransfield, 1987) but later raised into subtribes. The African rattan Ancistrophyllinae further divided into three genera, *Eremospatha, Lepidocaryum* and *Oncocalamus*, this results are in agreement with the classification proposed by Dransfield *et al*., 2005. *Ermospatha* is the largest genus of Ancistrophyllinae comprises of eleven species (Sunderland, 2012), out of which eight species were included in the present study. Genus *Ermospatha* formed a single monophyletic clade supported the study of Faye *et al*. (2014). Morphologically, *Ermospatha* is distinct from other genera in the subtribe by the absence of spines on leaf sheath, lowermost smaller leaflets, often swept back across basal portion of leaf, inflorescences lacking conspicuous bracts and hermaphrodite flowers with very short anthers inserted on top of filaments (Sunderland, 2012). Genus *Laccosperma* had a sister relationship with *Ermospatha* in the phylogenetic analysis. *Laccosperma* and *Oncocalamus* had monophyletic origin. Out of four species reported by Sunderland (2012), three were taken here for the analysis.

Tribe Mauritiinae constitute of three genera *viz*. *Lepidocaryum*, *Mauritia* and *Mauritiella* formed as monophyletic clade. *Eugeissona tristris* comes in Clade II as a separate tribe was considered to be in Ancistrophyllinae of Calameae (Uhl and Dransfield, 1987) and later on Dransfield *et al*. (2005) raised this genera into a new tribe Eugeissoneae.

Clade III comprises of Asian tribe, Calameae which includes five subtribes (Korthalsinae, Calaminae, Metroxylinae, Salaccinae and Plectocominae). Largest genus *Calamus* of subtribe Calaminae were the major part of this study. *Plectocomia,* and pseudorattan, *Salacca* showed a sister relationship here. Genus *Daemonorops* was found nested within genus *Calamus* which supports the need for merging of *Daemonorops* with *Calamus.* There by genus *Calamus* forming a monophyletic clade and majority of the clades in the phylogenetic tree were strongly supported by posterior probability at least 0.9. A proper resolution between some of the species is lacking in our study due to the lack of samples. We also suggest the use of more sequence data to resolve the relationships within genera.

Ancistrophyllinae being a genera of Calameae was treated as a part of Asian *Calamus* found to be in different clade in this study and closely related to tribe Lepidocaryeae as reported by previous study (Uhl and Dransfield, 1987). Even though India was a part of Gondwana, African rattans doesn't belong to Indian clade of rattans.

The origin of palms was reported in Laurasia about 91–120 mya and the subfamily Calamoideae diverged from other palms during Cretaceous period (80 mya) (Uhl and Dransfield, 1987; Friis *et al*., 2004; Baker and Couvrier, 2012). Tribe Eugeissoneae was estimated to be the first clade diverged from remaining members of the subfamily in the Cretaceous period (85 mya) (95% HPD 75.9-111.7 mya), supporting the previous study (Baker and Couvrier, 2012). During Campanian (76 mya), tribes Calameae and Lepidocaryeae diverged from each other. Mauritiinae was the first clade to be diverged from Lepidocaryeae (74 mya, 95% HPD 72-76 mya). Divergence within Mauritiinae occurred in Maastrichtian (Upper Cretaceous) (68.5 mya, 95% HPD 66.3-71 mya) likewise divergence between Raphiinae and Ancistrophyllinae happened around same period (71.9 mya, 95% HPD 62.7-75.9 mya). Divergence in African rattan subtribe, Ancistrophyllinae occurred in Eocene (52.6 mya, 95% HPD 37-66.5 mya) and *Oncocalamus* was the first to diverge in this subtribe. Later *Laccosperma* and *Ermospatha* diverged from each other in Eocene, 43.3 mya (95% HPD 29-57.9 mya).

Divergence of Asian tribe Calameae occurred in Paleocene (61.9 mya) (95% HPD 53.4-70 mya) where as Korthalsinae (54 mya) and Metroxylinae (51 mya) diverged in Eocene. Plectocominae and Salaccineae diverged from each other during Eocene (40 mya) found to be sister to Korthalsinae and Metroxylinae. *Daemonorops*, found nested within genus *Calamus* diverged in Oligocene (33.3 mya). This study suggested merging of *Daemonorops* within the genus *Calamus* and further confirmed its monophyletic origin and was estimated to be diverged in Eocene.

Node 125 represent the subfamily Calamoideae estimated to be diverged from other palms at crown node in Eurasia (85 mya) expanding their distribution into South East Asia, Africa, South America and Australia (Fig. 6.3). The favoured ancestral range at node 124 is B (Asia 100 %), and this node also suggests a vicariance and a dispersal events between Africa and Asia. Tribes Calameae and Lepidocaryeae diverged from each other at this node (node 124) (Fig. 6.3), with the distribution of former in Eurasia and latter in Africa, supporting the hypothesis of Baker and Couvrier (2011). Asian tribe Calameae diverged from Lepidocaryeae around 76 mya in Eurasia and dispersed to South and South East Asia expanding to Australia, while Lepidocaryeae diverged in Eurasia and dispersed into Africa and South America. Node 123 represents the divergence of Ancistrophyllinae and Raphiinae from Mauritiinae of tribe Lepidocaryae at 74 mya, suggests both vicariance and a dispersal event from Africa (C) and CE (Africa + South America). During that time period, the dispersal events might have occurred through the inter-connected African and South American land masses with island arcs (Pennington and Dick, 2004). Both BBM and DEC analysis (Fig. 6.3, Fig. 6.4), suggested C (Africa) as a possible ancestral range at node 118 with a 100 % marginal probability for Ancistrophyllinae and they still remain endemic to Africa, while Raphiinae dispersed to South America as well with 76% marginal probability (node). Node 122 represents Mauritiinae which dispersed from Africa into South America and at present they are present only in South American land mass and got isolated in South America (Baker and Couvrier, 2012). Two possible ancestral ranges were suggested for node 122, E (South America) and CE (Africa + South America) and percentage of their occurrence are 70 and 30 %, respectively.

Node 103 represents members of subtribe Calameae originated in Eurasia and dispersed into South East Asia, Australia and India. Korthalsinae diverged from other groups in Asia and widely distributed in South East Asia at present (node 102, 85%). Salaccinae and Plectocominae dispersed from South East Asia into North eastern parts of India (node 100 with 70% marginal probability). While Metroxylinae dispersed to Australia from Eurasia (node 103), however records of this was found from India during Paleogene, which indicated the early occurrence of the taxa in India. Fossils of *Korthalsia* as well as *Salacca* found from southern parts of India during Paleogene, revealed their distribution in Southern parts of India (Ramanujam, 2004). The species are found to be extinct from their distributional due to the post Miocene deterioration of climate resulting in very much depleted and diminished precipitation (Ramanujam, 2004).

The fossils records of *Dicolpopollis* spp. in Eocene-Oligocene rocks of North America, Europe, Middle East, South East Asia and India showed affinities with extant *Calamus* confirmed the early distribution of the genus all over the world. The records of *Dicolpopollis* spp. was widely reported from Early Tertiary of Europe. However, early records of this were reported from the Maastrichtian of Somalia. These records, indicate the North African/Eurasian origin of Calaminae. While, this study supports that the ancestors of Calaminae originated in B (node 99) (Asia excluding India, 67%), followed by a dispersal event to the Indian region indicated by node 95 (87%). The existence of early fossil records also supports this findings. *Calamus* got diverged from other genera of Calameae 61.9 mya in Eurasia (node 99). Divergence of the genus *Calamus* in the crown node was occurred at 46 mya underwent various vicariance and dispersal events

6.4.1. Rattans in India

The earliest evidence of palm fossil in India was of a costapalmate palm leaves reported from Deccan Intertrappean sediments during Maastrichtian–Danian (72-66 mya), considered as oldest fossil record of costa-palmate palm leaves from India and the Gondwana- derived continents (Srivastava *et al*., 2014). This revealed the occurrence of palms in India before the collision of India with Eurasia. Fossils of

Dicolpopollis spp. has been widely recorded from the Eocene (56-33) of northern India. *Quilonipollenites* (Rao and Ramanujam, 1978) showed affinity with some extant species of *Eugeissona*, originally described from Neogene Quilon Beds of Kerala, India as well as from Neyveli lignites (Phadtare and Kulkarni, 1984), considered to be from early to middle Eocene (Saxena, 1992; Morley, 2000). Fossil evidence reported from South America, Africa and India in Cretaceous/Tertiary boundary also revealed the early existence of the species in Gondwanan land mass (Ediger *et al*., 1990). The hypothesis suggested that Eugeissoneae diverged from other Calamoideae in Gondwana and reached Sunda in Tertiary, through rafting via India (Baker and Dransfield, 2000). Plant dispersals might have occurred through humid, tropical corridor existed between India and Sunda region in middle Eocene during its collision with Eurasia (Morley, 1998; 2003) Dispersal events might have occurred during Cretaceous period through the island arcs connectivity existed between India and Gondwana (Ali and Aitchison, 2008). At present *Eugeissona* is absent in India, this might be the result of climate change occurred in Neogene and Quaternary (Morley, 2000). Fossil evidence of subtribe Calaminae from Somalia in Late Cretaceous period (Schrank, 1994) support the possibilities for dispersal of Calaminae into India from Africa during Cretaceous through island arcs or oceanic dispersal. The possibilities for dispersal from Asia also, cannot be neglected because of the presence of abundant fossil evidences of ancestors of *Calamus* from South East Asia (Ediger *et al*., 1990). Hence there are possibilities for the ancestors of rattans, reached India through island arcs or through oceanic dispersal from either Africa or Asia before collision. These ancestors of rattans might have got completely wiped out from India due to post Miocene deterioration of climate or by Neogene and Quaternary climate change and later on, the extant rattans dispersed into India after the collision. (Morley, 2000; Ramanujan 2004). Conti *et al*. (2002)'s hypothesis suggested that the absence of certain species in peninsular India might be due to the massive extinction of India's Gondwanan elements caused by Late Cretaceous/ Early Tertiary climatic changes/ volcanic eruptions. This might be the reason for the absence of other rattans from South India except the genus *Calamus*.

6.4.2. *Calamus* in India

Present day Indian *Calamus* was found to have diverged around 35 mya in Eocene, while India was already collided with the Eurasian plate (Eocene). Hence the present study revealed the distribution of genus *Calamus* in India occurred through various dispersal events from Eurasian plate, which enlightened the 'out-of-Asia and into-India-origin' hypothesis (Blanford, 1901; Hora, 1949; Jayaram, 1949; Mani, 1974; Lakhanpal, 1970; Bande, 1992). The absence of *Calamus* species in Madagascar supports the notion that the genus was not carried before separation of Indo-Madagascar, from Gondwana land mass to Eurasia through India.

First divergence in the genus *Calamus* occurred in Eurasia at 46.6 Mya whereas *C. flagellum*, *C. longisetus* and *C. thwaitesii* diverged from remaining species of the genus. The clade representing genus *Calamus* dispersed to two different geographical zones, Asia and India. In the clade, two species, *C. flagellum* and *C. thwaitesii* were present in two different phytogeographical zones, North Eastern Himalayas and Western Ghats and their ancestors are assumed to reach India from South East Asia into North East India before 20 Mya. Later on these species are dispersed throughout Peninsular India including Sri Lanka around 10 mya (node 98) (Fig. 6.4).

Similarly in the second clade of the genus, three species such as *C. nagbettai*, *C. andamanicus* and *C. guruba* were found in three different phytogeographical zones of two plates (India and Euraisa), diverged from main clade around 42 mya. *C. nagbettai* from Western Ghats was unique from other species of Western Ghats but showed similarity with species of South Eastern Asia by the presence of cirrate leaf. Flagelliferous *C. guruba* was found nested with cirrate species *C. nagbettai*, *C. andamanicus*, which indicates the existence of convergent evolution. *C. nagbettai* and *C. guruba* diverged from ancestors of *C. andamanicus* in Indo – Asia. They might have reached India through South East Asia and *C. nagbettai* further got dispersed into Western Ghats region (node 65, Fig. 6.4). *C. guruba* and *C. flagellum* are present both in North Eastern India as well as in South East Asia through dispersal events as evident in the analysis.

The ancestral area of clade I (node 80) (Fig. 6.4) are AB (India + Asia) and B (Asia excluding India) with a frequency of 74 % and 26 % respectively. This clade suggests a dispersal from South East Asia into Western Ghats through North east India with a distinct vicariance between Peninsular India and Andaman/ Nicobar Islands. Node 70 represents the independent transoceanic dispersals to Andaman and Nicobar Islands from the source area of Peninsular India. Similarly, a vicariance is found at node 77 representing the dispersal of *Calamus baratangesis* to South East Asia from Peninsular India. A complex biogeography, with various dispersal events in each node have been found in shaping the current distribution pattern of clade II (node 86) (Fig. 6.4). Species from South East Asia, Andaman and Nicobar Islands, North eastern India and Western Ghats are clustered together in this node (node 86) (Fig. 6.4). This might be due to numerous dispersal events occurring in these geographical zones very rapidly (nodes 81-86) (Fig. 6.4). Clade III (node 94) (Fig. 6.4) consisted of species from India, which included majority of the species from reported *C. gamblei* complex of the Western Ghats. *C. peregrinus* from North eastern states of India was found clustered along with the species of the Western Ghats. The present study revealed the non-monophyletic origin of *Calamus* of the Western Ghats. Non-monophyly might be attributed due to multiple dispersal events happened within this clades. Majority of the species of Western Ghats might have reached through the dispersal from North Eastern Himalayas, however dispersal to Andamans and from South Eastern Asia is also evident in the present study (node 80) (Fig. 6.4).

The present study provided evolutionary relationships and biogeographic histories of subfamily Calamoideae. The study strongly supported the monophyly of the subfamily Calamoideae as well as its three tribes, Eugeissoneae, Lepidocaryeae and Calameae. Genus *Daemonorops* were found nested within the genus *Calamus*, thence genus *Calamus* forming a monophyletic clade. This study also suggests that Genus *Calamus* might have reached India through dispersal from Eurasian plate, after the collision of India with Eurasia during Eocene (35 mya). *Calamus* species of Southern India showed non- monophyly due to the frequent dispersals between other regions.

CHAPTER 7

SUMMARY & CONCLUSIONS

7. SUMMARY AND CONCLUSIONS

DNA barcoding has been adopted to delimit species boundaries of taxonomically complex genus *Calamus* of the Western Ghats. The study analysed species discrimination efficiency of 12 barcode regions, chloroplast (*rbc*L, *mat*K, *psb*A-*trn*H, *rpo*C, *rpo*B, *psb*K-*psb*I, *atp*F-*atp*H, *psb*Z-*trnf*M) and nuclear regions (ITS1, ITS2, PRK and RPB2) for the species identification. Three different statistical analyses *viz.* distance based, similarity based and tree based methods were used to evaluate species discrimination of the recommended barcodes. Among the analysed barcodes, ITS1and ITS2 gave multiple amplified products due to the presence of multiple copies in Calamoideae. *psb*Z-*trnf*M and PRK regions also failed to yield good amplification. The recommended DNA barcode regions adopted in the study failed to provide species-specific DNA barcodes, due to slow evolutionary rate in palms. This necessitated the need to explore new barcode regions other than plastid regions to delimit species boundaries in genus *Calamus*. Hence, fast evolving regions like low copy nuclear regions were exploited for their ability in species discrimination.The three statistical methods highlighted the discriminatory potential of the low copy nuclear region, RPB2 when compared to other candidate chloroplast barcode for the identification of *Calamus* species. From this study, we recommends RPB2 low copy nuclear region as the barcode region for resolving the taxonomic problems in the genus *Calamus* and a promising barcode region for palms in general.

Low copy nuclear region, RPB2 showed highest species resolution and was successful in unravelling species complexities in the genus. The species comprised in '*C. gamblei* species complex, *viz. C. gamblei, C. lacciferus, C. neelagiricus, C. prasinus, C. dransfieldii, C. renukae* and *C. shendurunii* shared similar morphological characters These morphological characters did not consistently differ from one another, the characters used to distinguish them vary within as well as amongst species. RPB2 resolved the species complex existing in the genus and confirmed the close association of all the studied species except for *C. shendurunii.*

The study also confirms that, the so called C. *delessertianus* and *C. dransfieldii* are same, based on the barcoding analysis as well as based on the morphological similarities. We consider the specimens treated as *C. delessertianus* from South India by Renuka (1999) to be morphological variation of *C. dransfieldii*. We also suggest that *C. hookerianus* and *C. pseudotenuis* are closely related species based on the molecular data. Even though, there are also many morphological similarities, the merging of both can be suggested only after analysing more accessions from Karnataka and include specimens from Sri Lanka. The study also supports the merging of *C. rivalis* with *C. metzianus* using morphology as well as molecular data. The misidentification of *C. metzianus* as *C. rotang* by several researchers, from the Western Ghats of Kerala was observed in the study. Thus the present study substantiated the potential of DNA barcode, RPB2 to differentiate *Calamus* species with taxonomic complexities.

Calamus species from three geographic zones of India and representatives of four genera of Calameae tribe (*Korthalsia* sp., *Salacca sp., Plectocomia sp., Daemonorops* sp.) were selected.Chloroplast and nuclear DNA markers (*psb*A-*trn*H, *rps16*, ITS2, *trn*L-*trn*F, *atp*H-*atp*I and RPB2) were evaluated to analyse the phylogenetic relationship within genus *Calamus* and its related genera using Bayesian inference. This study suggests that cirrate trait might have appeared in Indian rattans due to the convergent evolution. Both morphological as well as molecular characters were analysed to understand the phylogenetic relationship between *Calamus* species and related genera of Calamoideae. Genus *Calamus* of Western Ghats, grouped into six major lineages, showed morphological similarities with remaining genera. But some of the species are yet to be resolved with the use of additional data in order to get a more robust phylogeny of *Calamus* are suggested. More sampling has to be carried out for obtaining a fully resolved phylogenetic tree of Indian *Calamus.*

Studies on the evolutionary relationships and biogeographic histories of subfamily Calamoideae provide valuable insights into geological and evolutionary processes underlying their current distribution patterns. Phylogeny of Calamoideae was

analysed using two gene regions- *atp*H-*atp*I and RPB2, to understand systematic position of genus *Calamus*. Different methods are adopted to trace out the phylogenetic origin of the species using biogeography and molecular dating in Calamoideae. For divergence time estimation, GTR+G model was used based on Akaike Information Criterion (AIC) in uncorrelated lognormal relaxed clock model. Phylogenetic tree resolved Calamoideae into three major clades each representing tribes namely Lepidocaryeae, Eugeissoneae, and Calameae. Even though both India and Africa were Gondwanan in origin, rattans present in the studied biogeographic regions were not closely associated. Genus *Daemonorops* present were found nested within genus *Calamus,* which there by supported the merging of the genus with *Calamus.* The chronogram constructed for the subfamily Calamoideae has been estimated to be diverged from other palms at crown node in Cretaceous period (85 Mya) in Eurasia. Eugeissoneae tribe was estimated to be the first clade to diverge from rest of Calamoideae and later, tribes Calameae and Lepidocaryeae diverged from each other about 76 Mya, former in Eurasia and the latter in Africa. A major subtribe Calameae originated in Eurasia and underwent various dispersal events to South East Asia, Australia and India.

The present study focused on estimation of divergence time of Indian *Calamus*, which was found to have diverged around 35 Mya in Eocene.They might have reached India through dispersal events from Eurasian plate, after the collision of India with Eurasia. *Calamus* species of Southern India showed non- monophyly due to the frequent dispersals between the three phytogeographical regions. Similarly *Calamus* of India also showed non-monophyly. Before the collision of India, fossil records of Calamoideae were found from Southern parts of India, this might have reached here either through island arcs or transoceanic dispersal from Africa or Asia and extinction occurred in post Miocene deterioration of climate or by Neogene and Quaternary climate change. A true picture of the exact biogeographic history of the subcontinent will emerge from evidence accrued from multiple sources and in the meanwhile the findings of this study of palms and rattans in particular is expected to contribute meaningfully to the discourse.

CHAPTER 8

REFERENCES

8. REFERENCES

Ali, J.R. and Aitchison, J.C. 2008. Gondwana to Asia: Plate tectonics, paleogeography and biological connectivity of the Indian sub-continent from the middle Jurassic through latest Eocene (166±35 Ma). Earth-Science Reviews, 88: 145-166.

Al-Qurainy, F., Khan, S., Al-Hemaid, F.M., Ajmal Ali, 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 *rpo*B and *psb*A-*trn*H. International Journal of Molecular Sciences, 12: 6871-6880.

Alvarez, I. and Wendel, J.F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. Molecular Phylogenetics and Evolution, 29: 417-434.

Asif, M.J. and Cannon, C.H. 2005. DNA extraction from processed wood: a case study for the identification of an endangered timber species (*Gonystylus bancanus*). Plant Molecular Biology Reporter, 23: 185-192.

Asmussen, C.B. 1999a. Relationships of the tribe Geonomae (Arecaceae) based on plastid *rps16* DNA sequences. Acta Botanica Venezuelica, 22: 65-76.

Asmussen, C.B. 1999b. Towards a chloroplast DNA phylogeny of the tribe Geonomeae (Palmae). In: Henderson A, Borchsenius F. (eds.), Evolution, variation and classification of palms. Bronx, The New York Botanical Garden Press, pp. 121-129.

Asmussen, C.B., Baker, W.J. and Dransfield, J. 2000. Phylogeny of the palm family (Arecaceae) based on *rps16* intron and *trn*L–*trn*F plastid DNA sequences. In: Systematics and Evolution of the Monocots—Proceedings of the Second International Conference on the Comparative Biology of the Monocots, Sydney, Wilson, K.L. and Morrison, D.A. (eds.), Vol. 1, CSIRO, Melbourne. pp. 525-537.

Asmussen, C.B. and Chase, M.W. 2001. Coding and noncoding plastid DNA in palm systematics. American Journal of Botany, 88: 1103 - 1117.

Asmussen, C.B., Dransfield, J., Deickmann, V., Barfod, A.S., Pintaud, J.C. and Baker, W.J. 2006. A new subfamily classification of the palm family (Arecaceae): evidence from plastid DNA phylogeny. Botanical Journal of the Linnean Society, 151: 15-38.

Atria, M., Mil, H., Baker, W.J., Dransfield, J. and Welzen, P.C. 2017. Morphometric analysis of rattan *Calamus javensis* complex (Arecaceae: Calamoideae). Systematic Botany, 42(3): 494-506.

Bader, D.A., Moret, B.M.E. and Vawter, L. 2001. Industrial applications of high-performance computing for phylogeny reconstruction. Proceedings of SPIE ITCom, 4528: 159-168.

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 Sphaerocardamum and other Brassicaceae. Molecular Phylogenetics and Evolution, 13: 20-30.

Bailey, C.D., Carr, T.G., Harris, S.A. and Hughes, C.E. 2003. Characterization of angiosperm nrDNA polymorphism, paralogy and pseudogenes. Molecular Phylogenetics and Evolution, 29: 435-455.

Baker, W.J. 2015. A revised delimitation of the rattan genus *Calamus* (Arecaceae). Phytotaxa, 197 (2): 139-152.

Baker, C.S. and Palumbi, S.R. 1994. Which whales are hunted? A molecular genetic approach to whaling. Science, 265: 1538-1539.

Baker, W.J., Asmussen, C.B., Barrow, S.C., Dransfield, J. and Hedderson, T.A. 1999a. A phylogenetic study of the palm family (Palmae) based on chloroplast DNA sequences from the *trn*L–*trn*F region. Plant Systematics and Evolution, 219: 111-126.

Baker, W.J., Dransfield, J., Harley, M.M. and Bruneau, A. 1999b. Morphology and cladistic analysis of subfamily Calamoideae (Palmae). In: Evolution and Classification of Palms, A. Henderson and F. Borchsenius, (eds.), Memoirs of the New York Botanical Garden, 83: 307-324.

Baker, W.J., Hedderson, T.A. and Dransfield, J. 2000a. Molecular phylogenetics of subfamily Calamoideae (Palmae) based on nrDNA ITS and cpDNA *rps16*

intron sequence data. Molecular Phylogenetics and Evolution, 14 (2): 195-217.

Baker, W.J., Hedderson, T.A. and Dransfield, J. 2000b. Molecular phylogenetics of *Calamus* (Palmae) and related rattan genera based on 5S nrDNA spacer sequence data. Molecular Phylogenetics and Evolution, 14 (2): 218-231.

Baker, W.J. and Dransfield, J. 2000. Towards a biogeographic explanation of the calamoid palms. Monocots: systematics and evolution, K.L. Wilson and D.A. Morrison (eds.), CSIRO, Melbourne. pp. 545-553.

Baker, W.J. and Dransfield, J. 2008. *Calospatha* subsumed in *Calamus* (Arecaceae: Calamoideae). Kew Bulletin, 63(1): 161-162.

Baker, W.J., Savolainen, V., Asmussen-Lange, C.B., Chase, M.W., Dransfield, J., Forest, F., Harley, M.M., Uhl, N.W. and Wilkinson, M. 2009. Complete generic-level phylogenetic analyses of palms (Arecaceae) with comparisons of supertree and supermatrix approaches. Systematic Biology, 58: 240-256.

Baker, W.J., Norup, M.V., Clarkson, J.J., Couvreur, T.L., Dowe, J.L., Lewis, C.E., Pintaud, J.C., Savolainen, V., Wilmot, T. and Chase, M.W. 2011. Phylogenetic relationships among arecoid palms (Arecaceae: Arecoideae). Annals of Botany, 108 (8): 1417-32.

Baker, W.J. and Couvreur, T.L.P. 2012. Global biogeography and diversification of palms sheds light on the evolution of tropical lineages. I: Historical Biogeography. Journal of Biogeography, 40: 274-285.

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

Bande, M.B. 1992. The Palaeogene vegetation of peninsular India (megafossil evidence). Palaeobotanist, 40: 275-284.

Barfod, A.S., Trenel, P. and Borchsenius, F. 2010. Drivers of diversification in the vegetable ivory palms (Arecaceae: Ceroxyloideae, Phytelepheae) – Vicariance or adaptive shifts in niche traits? In Diversity, phylogeny and evolution of Monocotyledons, O. Seberg, A. T. Peterson, A. Barfod, and A. M. Davis (eds.), Aarhus University Press. Pp. 225-243.

Barrow, S. 1996. A Monograph of *Phoenix L.* (Palmae: Coryphoideae). PhD Thesis, University of Reading, UK.

Barrow, S.C. 1998. A monograph of *Phoenix* L. (Palmae: Coryphoideae). Kew Bulletin, 53: 513-575.

Basu, S.K. 1985. The present status of rattan palms in India – An overview. In: Proceedings of the rattan seminar, Edited by Wang and Manokaran, Kuala Lumpur, Malaysia, pp. 77-94.

Bayton, R.P. 2005. *Borassus* L. and the Borassoid palms: systematics and evolution. Ph.D. Thesis, University of Reading, UK.

Beccari, O. 1908. Asiatic palms – Lepidocaryeae. Part 1. The species *Calamus*. Annals of Royal Botanical Garden Calcutta, 11: 1- 518.

Beccari, O. 1914. Asiatic palms – Lepidocaryeae. The species *Calamus*. Supplement to Part 1. Annals of Royal Botanical Garden Calcutta, 11 (Appendix): plates 1- 83.

Bera, S. and Banerjee, M. 1990. A new species of Palmoxylon and accretionary structures in the petrified woods from lateritic sediment in the western part of Bengal basin, India. Indian Journal of Earth Sciences, 17:78-79.

Bhat, K.M and Thulasidas, P.K. 1989. *Calamus metzianus* Schlecht. Why these rattans break? RIC Bulletin, 8 (1/4): 4-5.

Blanford, W.T. 1901. The distribution of vertebrate animals in India, Ceylon, and Burma; Philosophical Transactions of the Royal Society B. Biological Sciences, 194: 335-436.

Bocxlaer, I.V., Roelants, K., Biju, S.D., Nagaraju, J. and Bossuyt, F. 2006. Late Cretaceous vicariance in Gondwanan amphibians. PlosOne 74: 1-6.

Boer, W.J. 1968. The Genomoid palms. Verh. Kon. Ned. Akad. Wetensch, Afd. Natuurk. Tweede Sect. ser. 58: 1- 202.

Bogh, A. 1996. The reproductive phenology and pollination biology if four *Calamus* (Arecaceae) species in Thailand. Principes, 40: 5-15.

Bossuyt, F. and Milinkovitch, M. C. 2001. Amphibians as indicators of EarlyTertiary “out-of-India”dispersal of vertebrates. Science 292:92-95.

Bossuyt, F., Brown, R.M., Hillis, D.M., Cannatella, D.C. and Milinkovitch, M.C. 2006. Phylogeny and biogeography of a cosmopolitan frog radiation: Late Cretaceous diversification resulted in continent-scale endemism in the family Ranidae. Systematic Biology, 55: 579-594

Braukmann, T.W.A., Kuzmina, M. and Stefanović, S. 2009. Loss of all plastid ndh genes in Gnetales and conifers: Extent and evolutionary significance for the seed plant phylogeny. Current Genetics, 55: 323-337.

Bremer, K. 2000. Early Cretaceous lineages of monocot flowering plants. Proceedings of the National Academy of Sciences USA, 97: 4707-4711.

Briggs, J.C. 1989. The historic biogeography of India: Isolation or contact? Systematic Zoololgy, 3: 322-332.

Briggs, J.C. 2003. The biogeographic and tectonic history of India. Journal of Biogeography, 30: 381-388.

Brown, W.M., Prager, E.M., Wang, A. and Wilson, A.C. 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. Journal of Molecular Evolution, 18: 225-239

Burrows, P.A., Sazanov, L.A., Svab, Z., Maliga, P. and Nixon, P.J. 1998. Identification of a functional respiratory complex in chloroplasts through analysis of tobacco mutants containing disrupted plastid *ndh* genes. The EMBO Journal, 17: 868-876.

Bush, R.M., Bender, C.A., Subbarao, K., Cox, N.J. and Fitch, W.M. 1999. Predicting the evolution of human infuenza. Science, 286 (5446): 1921-1925.

Caputo, C. 1997. DNA phylogeny in plants: History and new perspectives. Lagascalia, 19 (1-2): 331-344.

Carl, E.L. and Doyle, J.J. 2001. Phylogenetic Utility of the Nuclear Gene Malate Synthase in the Palm Family (Arecaceae). Molecular Phylogenetics and Evolution, 19 (3): 409-420.

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

Chang, B.S. and Donoghue, M.J. 2000. Recreating ancestral proteins. Trends in Ecology and Evolution, 15: 109-114.

Chase, M.W., Soltisn, D.E., Olmstead, R.G., Morgan, D., Les, D.H., Mishler, B.D., Duvall, M.R., Price, R.A., Hills, H.G., Qiu, Y.L., Kron, K.A., Rettig, J.H., Conti, E., Palmer, J.D., Manhart, J.R., Sytsma, K.J., Michaels, H.J., Kress, W.J., Karol, K.G., Clark, W.D., Hedron, M., Gaut, B.S., Jansen, R.K., Kim, K.J., Wimpee, C. E, Smith, J. E, Furnier, G.R., Strauss, S.H., Xiang, Q.Y., Plunkett, G.M., Soltis, R.S., Swensen, S.M., Williams, S.E., Gadek, R.A., Quinn, C.J., Eguiarte, L.E., Golenberg, E., Learn, G.H. Jr., Graham, S.W., Barrett, S.C.H., Dayanandan, S. and Albert, V.A. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. Annals of the Missouri Botanical Garden, 80: 528-580.

Chase, M.W., Stevenson, D.W., Wilkin, R. and Rudall, R.J. 1995. Monocot systematics: a combined analysis. In: Rudall, R.J., Cribb, R.J., Cutler, D.E, Humphries, C.J. (eds.), Monocotyledons: systematics and evolution. Royal Botanic Gardens, Kew, pp. 685-730.

Chase, M.W., Salamin, N., Wilkinson, M., Dunwell, J.M., Kesanakurthi, R.P., Haider, N. and Savolainen, V. 2005. Land plants and DNA barcodes: short-term and long-term goals. Philosophical Transactions of the Royal Society B: Biological Sciences, 360: 1889-1895.

Chase M.W., Fay M.F., Devey D.S., Maurin O., Ronsted N., Davies T.J., Pillon Y., Peterson G., Seberg O., Tamura M.N., Asmussen C.B., Hilu K., Borsch T., Davis J.I., Stevenson D.W., Pires J.C., Givnish, T.J., Sytsma K.J., McPherson M.A., Graham S.W. and Rai H.S. 2006. Multigene analyses of monocot relationships: a summary. Aliso, 22: 63-75.

Chase, M.W., Cowan, R.S., Hollingsworth, P.M., Berg, C.V.D., Madriñán, 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. 2007. A proposal for a standardized protocol to barcode all land plants. Taxon, 56 (2): 295-299.

Chatterjee, S. and Scotese, C. 1999. The breakup of Gondwana and the evolution and biogeography of the Indian Plate. Proceedings of the National Academy of Sciences, 3: 397-425.

Chatterjee, S. and Scotese, C. 2010. The wandering Indian plate and its changing biogeography during the Late Cretaceous-Early Tertiary Period. In: Reitner J, Trauth MH, StuÈwe K, Yuen D. (eds.), Lecture Notes in Earth Sciences. New Aspects of Mesozoic Biodiversity. Springer-Verlag Berlin Heidelberg. pp. 105±126.

Chen, J., Li, Q., Kong, L., Yu, H. 2011. How DNA Barcodes Complement Taxonomy and Explore Species Diversity: The Case Study of a Poorly Understood Marine Fauna. PLoS ONE, 6(6): e21326.

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., Luo, K., Li, Y., Li, X., Jia, X., Lin, Y. and Leon, C. 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS ONE, 5: e8613.

China Plant BOL Group – Li, D.Z., Gao, L.M., Li, H.T., Wang, H., Ge, X.J., Liu, J.Q., Chen, Z.D., Zhou, S.L., Chen, S.L., Yang, J.B., Fu, C.X., Zeng, C.X., Yan, H.F., Zhu, Y.J., Sun, Y.S., Chen, S.Y., Zhao, L., Wang, K., Yang, T. and Duan, G.W. 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 USA, 108: 19641-19646.

Clegg, M.T. 1993. Chloroplast gene sequences and the study of plant evolution. Proceedings of the National Academy of Sciences, USA, 90: 363–367.

Collin, R. 2003. Phylogenetic relationships among calyptraeid gastropods and their implications for the biogeography of marine speciation. Systematic Biology, 52 (5): 618-640

Conti, E., Eriksson, T., Schonenberger, J., Systsma, K.J., Baum, D.A., 2002. Early Tertiary out-of-India dispersal of Crypteroniaceae: evidence from phylogeny and molecular dating. Evolution, 56: 1931-1942.

Couvreur, T.L.P., Forest, F. and Baker, W.J. 2011. Origin and global diversification patterns of tropical rain forests: inferences from a complete genus-level phylogeny of palms. BMC Biology, 9: 44.

Crawford, D.J. 2000. Plant macromolecular systematics in the past 50 years: one view. Taxon, 49: 479-50 l.

Crisci, J., Katinas, L., Posadas, P. and Crisci, J.V. 2009. Historical Biogeography: An Introduction. Harvard University Press.

Crisp, M.D., Isagi, Y., Kato, Y., Cook, L.G. and Bowman, D.M.S. 2010. Livistona palms in Australia: ancient relics or opportunistic immigrants?. Molecular Phylogenetics and Evolution, 54:512-523.

Cuenca, A. and Asmussen-Lange, C.B. 2007. Phylogeny of the palm tribe Chamaedoreeae (Arecaceae) based on plastid DNA sequences. Systematic Botany, 32: 250-263.

Cullis, C.A., Creisen, G.P., Gorman, S.W. and Teasdale, R.D. 1988. The 25S, 18s, and 5S ribosomal RNA genes from *Pinus radiata* D. Don, In: Cheliak, W.M. and Yapa, A.C. (eds.), Molecular genetics of forest trees, Petawa Natl For Inst Rep, Chalk River, Ontario, Canada 24-40.

Daru, B.H., van der Bank, M., Bello, A. and Yessoufou, K. 2017. Testing the reliability of standard and complementary DNA barcodes for the monocot subfamily Alooideae from South Africa. Genome, 60: 337-347.

Datta-Roy, A., Singh, M. and Karanth, K.P. 2014. Phylogeny of endemic skinks of the genus Lygosoma (Squamata: Scincidae) from India suggests an in situ radiation. Journal of Genetics, 93:1-16.

De Queiroz, K., Howard, D.J. and Berlocher, S.H. 1998. The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations, Endless forms: Species and speciation, New York Oxford University. pp. 57-75

Denton, A.L., Mc Conaughy, B.L. and Hall, B.D. 1998. Usefulness of RNA polymerase II coding sequences for estimation of green plant phylogeny. Molecular Biology and Evolution, 15: 1082-1085.

Dev, S.A., Muralidharan, E.M., Sujanapal, P. and Balasundaran, M. 2014. Identification of market adulterants in East Indian sandalwood using DNA barcoding. Annals of Forest Sciences, 71 (6): 517-522.

Domenech, B., Asmussen-Lange, C.B., Baker, W.J., Alapetite, E., Pintaud, J.C. and Nadot, S. 2014. A phylogenetic analysis of palm subtribe Archontophoenicinae (Arecaceae) based on 14 DNA regions. Botanical Journal of the Linnean Society, 175: 469-481.

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

Doyle, J.J., and Doyle, J.L. 1999. Nuclear protein-coding genes in phylogeny reconstruction and homology assessment: Some examples from Leguminosae. In: Molecular Systematics and Plant Evolution, P. M. Hollingsworth, R. M. Bateman, and R. J. Gornall (eds.), Taylor & Francis, London. pp. 229-254.

Drager, R.G. and Hallick, R.B. 1993. A novel *Euglena gracilis* chloroplast operon encoding 4 ATP synthase subunits and 2 ribosomal-proteins contains 17 Introns. Current Genetics, 23: 271- 280.

Dransfield, J. 1996. The rattan taxonomy and ecology. In: Rao, A.N., Rao, R.V. (eds), Proceedings of training courses cum workshop, rattan: taxonomy, ecology, silviculture, conservation, genetic improvement and biotechnology. IPGRI-APO and INBAR, Sarawak.

Dransfield, J. 1978. Growth forms of rain forest palms. In: P.B. Tomlinson and M.H. Zimmermann (eds.), Tropical trees as living systems. Cambridge University Press, Cambridge, UK. pp. 247- 268

Dransfield, J. 1979. A Manual on Rattans of Malay Peninsula. Malayan Forest Records No. 29. Forestry Dept., Peninsular Malaysia. 270 p.

Dransfield, J. and Uhl, N.W. 1986. An outline of a new classification of palms. Principes, 30: 3-11.

Dransfield, J. 1992. The taxonomy of rattans. In: Razali, W.M., Dransfield, J. and Manokaran, N. (eds.), A Guide to the Cultivation of Rattan. Forest Research Institute: Forest Record No. 35. Kuala Lumpur. Malaysia.

Dransfield, J. and Manokaran, N. 1994. Rattans. Plant Resources of South-East Asia No. 6. Indonesia, PROSEA Foundation, Bogor.

Dransfield, J. and Beentje, H.J. 1995. The Palms of Madagascar. Kew: Royal Botanic Gardens, and International Palm Society, Norwich.

Dransfield, J. 2000. General introduction to rattan - the biological background to exploitation and the history of rattan research. In: Dransfield, J., Florentino, O.T., Manokaran, N. (eds.), Rattan - Current Research Issues and Prospects for Conservation and Sustainable Development. Non Wood Forest Products 14. 5-7 December, FAO, Rome, pp. 23-34.

Dransfield, J. 2001. Gaps in taxonomic knowledge and confusion in nomenclature are challenges to rattan research and development. Taxonomy, biology and ecology of rattan. An international journal of forestry and forest industries. Unasylva, 52: 205

Dransfield, J., Uhl, N.W., Asmussen, C.B., Baker, W.J., Harley, M.M., and Lewis, C.E. 2005. A new phylogenetic classification of the palm family, Arecaceae. Kew Bulletin, 60: 559-569.

Dransfield, J., Uhl, N.W., Asmussen, C.B., Baker, W.J., Harley, M.M. and Lewis, C.E. 2008. Genera Palmarum the evolution and classification of palms, 2nd edn. Royal Botanic Gardens, Kew, Richmond, UK.

Drummond, A.J. and Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology, 7: 214-221.

Edgar, R.C. 2004. Muscle: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32(5): 1792- 1797.

Ediger, V.S., Bati, Z. and Alisan, C. 1990. Paleopalynology and paleoecology of *Calamus*-like disulcate pollen grains. Review of Palaeobotany and Palynology, 62: 97-105.

Eisen, J.A. 2003. Phylogenomics: intersection of evolution and genomics. Science, 300: 1706-1707.

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.

Erickson, D.L., Spouge, J., Resch, A., Weigt, L.A. and Kress, J.W. 2008. DNA barcoding in land plants: developing standards to quantify and maximize success. Taxon, 57: 1304-1316.

Eyre-Walker, A. and Gaut, B.S. 1997. Correlated rates of synonymous site evolution across plant genomes. Molecular Biology and Evolution, 14: 455-460.

Faye, A., Pintaud, J.C., Baker, W.J., Sonke, B. and Couvreur, T.L.P. 2014. A plastid phylogeny of the African rattans (Ancistrophyllinae, Arecaceae). Systematic Botany, 39(4): 1099-1107.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. Journal of Molecular Evolution, 17 (6): 368-376.

Feng, S., Jiang, Y., Wang, S., Jiang, M., Chen, Z., Ying, Q., Wang, H. 2013. Molecular identification of *Dendrobium* Species (*Orchidaceae*) based on the DNA barcode *ITS2* region and its application for phylogenetic study. International Journal of Molecular Sciences, 16: 21975-21988.

Fitch, W. 1971. Toward defining the course of evolution: minimum change for a specified tree topology. Systematic Zoology, 20: 406-416.

Fitch,W. M. and Margoliash, E. 1967. Construction of phylogenetic trees. Science, 155 (760): 279-284.

Ford, C.S., Ayres, K.L., Toomey, N., Haider, N., Stahl, J.V.A., Kelly, L.J., Wikström, N., Hollingsworth, P.M., Duff, R.J., Hoot, S.B., Cowan, R.S., ChaseM.W. and Wilkinson, M.J. 2009. Selection of candidate coding DNA barcoding regions for use on land plants. Botanical Journal of the Linnean Society, 159: 1-11.

Friis, E.M., Pedersen, K.R. and Crane, P.R. 2004. Araceae from the Early Cretaceous of Portugal: evidence on the emergence of monocotyledons. Proceedings of the National Academy of Sciences USA, 101: 16565-16570.

Fuji, T. 2007. Outline of the research project “Methods to identify wood species and origin of timber of Southeast Asia”. Proc. of the International Symposium on

Development of improved methods to identify *Shorea* species wood and its origin, September 25-26, 2007, University of Tokyo, Tokyo and IUFRO,19.

Gao, T., Yao, H., Song, J.Y., Liu, C., Zhu, Y., Ma, X., Pang, X., Xu, H., Chen, S. 2010. Identification of medicinal plants in the family Fabaceae using a potential DNA barcode ITS2. Journal of Ethnopharmacology, 130: 116-121.

Gaut, B.S., Muse, S.V., Clark, W.D. and Clegg, M.T. 1992. Relative rates of nucleotide substitution at the *rbcL* locus of the monocotyledonous plants. Journal of Molecular Evolution, 35: 292-303.

Gaut B.S, Morton B.R., McCaig, B.C., Clegg, M.T. 1996. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbcL*. Proceedings of the National Academy of Sciences, USA, 93: 10274-10279.

Gaut, B.S. 1998. Molecular clocks and nucleotide substitution rates in higher plants. Evolutionary Biology, 30: 93-120.

Gentry, A.H. 1991. The distribution and evolution of climbing plants. In: Putz F.E, Mooney H.A (eds.), The biology of vines. Cambridge University Press, Cambridge, pp. 3-49.

Goel, C.L. 1992. Tips on cultivation and harvesting techniques of canes (rattans) in India. In: ratttan management and utilization. Proceedings rattan-India Seminar, pp.174-179.

Govaerts, R. and Dransfield, J. 2005. World Checklist of Palms. Royal Botanic Gardens, Kew, pp. 223

Guisinger, M.M., Kuehl, J.V., Boore, J.L. and Jansen, R.K. 2008. Genome-wide analyses of Geraniaceae plastid DNA reveal unprecedented patterns of increased nucleotide substitutions. Proceedings of the National Academy of Sciences USA, 105: 18424-18429.

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

Haberhausen, G. and Zetsche, K. 1994. Functional loss of all ndh genes in an otherwise relatively unaltered plastid genome of the holo parasitic flowering plant Cuscuta reflexa. Plant Molecular Biology, 24: 217-222.

Hahn, W. 1999. Molecular systematic studies of the Palmae. Memoirs of the New York Botanical Garden, 83: 47-70.

Hahn, W.J. 2002a. A molecular phylogenetic study of the Palmae (Arecaceae) based on *atp*B, *rbc*L, and 18S nrDNA sequences. Systematic Biology, 51: 92-112.

Hahn,W.J. 2002b. A phylogenetic analysis of the Arecoid line of palms based on plastid DNA sequence data. Molecular Phylogenetics and Evolution, 23: 189-204.

Hajdukiewicz, P.T.J., Allison, L.A. and Maliga, P. 1997. The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. The EMBO Journal, 16: 4041-4048.

Hajibabaei, M., Singer, G.A., Hebert, P.D.N., Hickey, D.A. 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends in Genetics, 23 (4): 167-172.

Hajibabaei, M. and McKenna, C. 2012. DNA mini-barcodes. Methods in Molecular Biology, 858: 339-53.

Halbur, P., Lum, M.A., Meng, X., Morozow, I. and Paul, P.S. 1994. New porcine reproductive and respiratory syndrome virus DNA and proteins encoded by open-ended frames of an Iowa strain of the virus are used in vaccines against PRRSV in pigs. Patent-Filing WO9606619-A1.

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

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

Harley, M.M. 2006. A summary of fossil records for Arecaceae, Botanical Journal of the Linnean Society, 151: 39-67.

Harrison, R.G., Howard D.J. and Berlocher S.H. 1998. Linking evolutionary pattern and process, Endless forms: Species and speciation, New York Oxford University Press, pp. 19-31.

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 B: Biological Sciences, 270: 313-321.

Hedges, S.B. 2003. The coelacanth of frogs. Nature (London), 425: 669-670.

Henderson, A., Martins, R. 2002. Classification of specimens in the *Geonoma stricta* (Palmae) complex: the problem of leaf size and shape. Brittonia, 54: 202-212.

Henderson, A. 2011. A revision of *Geonoma* (Arecaceae). Phytotaxa, 17: 1-27.

Henderson, A. and Henderson, F. 2007. New species of *Calamus* from Laos and Myanmar. Taiwania, 52(2): 152-158.

Henderson, A. 2002. Evolution and ecology of palms. New York: The New York Botanical Garden Press.

Herngreen, G.F.W., Kedves, M., Rovinina, L.V., Smirnova, S.B. 1996. Cretaceous palynofloral provinces: a review. In: Jansonius J, Mcgregor D C, Palynology: principles and applications. Dallas: American Association of Stratigraphic Palynologists Foundation. pp. 1157-1188.

Hollingsworth, M.L., Clark, A.A., Forrest, L.L., Richardson, J., Pennington, R.T., Long, D.G., Cowan,R., Chase, M.W., Gaudeul, M. and Hollingsworth, P.M. 2009. Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. Molecular Ecology Resources, 9: 439-457.

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

Hora, S.L. 1949. Satpura hypothesis of the distribution of Malayan fauna and flora of peninsular India. Proceedings of the National Academy of Sciences, India, 15: 309-314.

Janssen, T. and Bremer, K. 2004. The age of major monocot groups inferred from 800 + rbcL sequences. Botanical Journal of the Linnean Society, 146: 385-398.

Jayaram, K.C. 1949. Distribution of Lizards of peninsular India with Malayan affinities. Proceedings of the National Academy of Sciences, India 15:403-409.

Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G. and Gibson, T.J. 1998. Multiple sequence alignment with Clustal X. Trends in Biochemical Sciences, 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.

Joly, S., Davies, T.J., Archambault, A., Bruneau, A., Derry, A., Kembel, S.W., Peres-Neto, P., Vamosi, J., and Wheeler, T.A. 2014. Ecology in the age of DNA barcoding: the resource, the promise and the challenges ahead. Molecular Ecology Resources, 14: 221-232.

Keast, A. 1971. Continental drift and the evolution of the biota on Southern Continents; Quarterly Review of Biology, 46: 335-378.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution, 16: 111-120.

Kissling, W.D., Eiserhardt, W.L., Baker, W.J., Borchsenius, F., Couvreur, T.L.P., Balslev, H. and Svenning, J.C. 2012. Cenozoic imprints on the phylogenetic structure of palm species assemblages worldwide. Proceedings of the National Academy of Sciences USA, 109: 7379-7384.

Kohler, F. and Glaubrecht, M. 2003. Morphology, reproductive biology and molecular genetics of ovoviviparous freshwater gastropods (Cerithioidea, Pachychilidae) from the Philippines, with description of a new genus *Jagora*. Zoologica Scripta, 33: 33-59.

Kramadibrata, P. 1992. A revision of the genus Calamus (Palmae) Section Macropodus sensu Furtado. Ph.D. thesis, University of Reading, UK.

Kress, J.W., 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: 8369-8374.

Kvacek, J. and Herman, A.B. 2004. Monocotyledons from the early Campanian (Cretaceous) of Grünbach, lower Austria. Review of Palaeobotany and Palynology, 128:323-353.

La, D., Sutch, B. and Livesay, D.R. 2005. Predicting protein functional sites with phylogenetic motifs. Proteins*:* Structure, Function and Bioinformatics, 58(2): 309-320.

Lahaye, R., Savolainen, V., Duthoit, S., Maurin, O. and van der Bank, M. 2008. A test of *psb*K–*psb*I and *atp*F–*atp*H as potential plant DNA barcodes using the flora of the Kruger National Park as a model system (South Africa). Nature Precedings. http://hdl.handle.net/10101/npre.2008.1896.1

Lakhanpal, R.N. 1970. Tertiary Floras of India and their bearing on the historical geology of the region. Taxon, 19: 675-694.

Lakshmana, A.C. 1995. Rattans of South India. Evergreen Publishers, Bangalore, India. pp. 1-179.

Lee, Y.F., Jong, K., Swaine, V.K. Chey, V.K. and Chung, A.Y.C. 1995. Pollination in the rattans *Calamus subinermis* and *Calamus caesius.* Sandakania, 6: 15-39.

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

Lewis**,** C.E., Baker, W. and Asmussen, C. B. 2000. DNA and palm evolution. Palms, 44:19-24.

Lewis, C.E. and Doyle, J.J. 2001. Phylogenetic utility of the nuclear gene malate synthase in the palm family (Arecaceae). Molecular Phylogenetics and Evolution, 19: 409-420.

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.

Li, X., Yang, Y., Henry, R.J., Rossetto, M., Wang, Y., Chen, S. 2015. Plant DNA barcoding: from gene to genome. Biological Reviews of Cambridge Philosophical Society 90(1):157-66.

Loo, A.H.B., Dransfield, J., Chase, M.W., 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.

Lyngdoh, N., Santosh, S.H., Ramesha, B.T., Rao, N.M., Ravikanth, G., Narayani, B., Ganeshaiah, K.N. and Shaanker, U.R. 2005. Rattan species richness and population genetic structure of *Calamus flagellum* in North-Eastern Himalaya, India. Journal of Bamboo and Rattan, 4: 293-307.

Macey, J.R., Schulte, J.A., Larson, A., Ananjeva, N.B., Wang, Y.Z., Pethiyagoda, R., Rastegar-Pouyani, N. and Papenfuss, T.J. 2000. Evaluating trans-tethys migration: An example using acrodont lizard phylogenetics. Systematic Biology, 49:233-256.

Mani, M.S. 1974. Biogeography of Peninsula; in Ecology and biogeography in India (The Hague: Dr W. Junk Publishers) pp. 614-646.

Mann, G. and Wendland, H. 1864. On the palm of western tropical Africa. Transactions of the Linnean *Society* of London, 24: 421-439.

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.

Mayden, R.L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. In: Claridge, M.A., Dawah, H.A., Wilson, M.R. (eds.), Species: the Units of Diversity. Chapman & Hall, London, pp. 381-424.

Mayr, E. 1942. Systematics and the Origin of Species (Columbia Univ. Press, New York).

Mayr, E. 1969. Principles of systematic zoology. McGraw-Hill, New York, NY. 428 p.

McIntyre, S.R.N., Lineweaver, C.H., Groves, C.P. and Chopra, A. 2017. Global biogeography since Pangaea**.** Proceedings of the Royal Society B: Biological Sciences, 284 (1856): 20170716

McKenna, M.C. 1995. The Mobile Indian Raft: A reply to Rage and Jaeger. Systematic Biology, 44: 265-271.

Meerow, A.W., Noblick, L., Borrone, J.W., Couvreur, T.L.P., Mauro- Herrera, M., Hahn, W. J., Kuhn, D.N., Nakamura, K., Oleas, N.H. and Schnell, R.J. 2009. Phylogenetic analysis of seven WRKY genes across the Palm subtribe Attaleinae (Areceaceae) identifies Syagrus as sister group of the Coconut. PLoS ONE, 4: e7353.

Mehrotra, R.C. 2000. Study of plant megafossils from the Tura Formation of Nangwalbibra, Garo Hills, Meghalaya, India. Palaeobotanist, 49 (2): 225-237.

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.

Meng, B.Y., Wakasugi, T. and Sugiura, M. 1991. Two promoters within the *psbK–psbl–trnG* gene-cluster in tobacco chloroplast DNA. Current Genetics, 20: 259-264.

Meyer, C.P. and Paulay, G. 2005. DNA Barcoding: Error Rates Based on Comprehensive Sampling. PLoS Biology, 3(12): e422.

Mishler, B.D., Lewis, L.A., Buchheim, M.A., Renzaglia, R.S., Garbary, D.J., Deleiche, C.F., Zechman, F.W., Kantz, T.S. and Capman R.L.1994. Phylogenetic relationships of the “green algae” and “bryophytes”. Annals of the Missouri Botanical Garden, 81:451-483.

Mohan, R.H.Y. and Tandon, R. 1997. Bamboo and rattans: from riches to rags. Proceedings of the Indian National Academy of Sciences, 63: 245-267.

Moore, H.E. 1973. Palms in the tropical forest ecosystems of Africa and South America. Tropical forest ecosystems in African and South America: a comparative review, B.J. Meggers, E.S. Ayensu and W.D. Duckworth (eds), Smithsonian Institute Press, Washington, DC, pp. 63-88.

Moritz, C., Dowling, T.E. and Brown, W.M. 1987. Evolution of Animal mitochondrial DNA-revelence for population biology and systematic. Annual Review of Ecology, Evolution, and Systematics, 18: 269-292.

Morley, R.J. 1998**.** Palynological evidence for Tertiary plant dispersals in the SE Asian region in relation to plate tectonics and climate. In: Hall, R. and Holloway J.D. (eds.), Biogeography and geological evolution of SE Asia. Backhuys Publishers, Leiden, The Netherlands. pp. 211-234.

Morley, R.J. 2000. Origin and Evolution of Tropical Rain Forests. Wiley, Chichester. pp. 362.

Morley, R.J. 2003. Inter plate dispersal paths for megathermal angiosperms. Perspectives in Plant Ecology, Evolution and Systematics, 6: 5-20.

Mort, M. and Crawford, D. 2004. The Continuing Search: Low-Copy Nuclear Sequences for Lower-Level Plant Molecular Phylogenetic Studies. Taxon, 53: 257.

Nakazono, M. and Hirai, A. 1993. Identification of the entire set of transferred chloroplast DNA sequences in the mitochondrial genome of rice. Molecular and General Genetics, 236: 341-346.

Neuhaus, H. and Link, G. 1987. The chloroplast *tRNA* (UUU) gene from mustard (Sinapsis alba) contains a class II intron potentially coding for a maturase-related polypeptide. Current Genetics, 11: 251-257.

Newmaster, S.G., Fazekas, A.J., Steeves, R.A.D. and Janovec, J. 2008. Testing candidate plant barcode regions in the Myristicaceae. Molecular Ecology Notes, 8: 480-490.

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

Nock, C.J., Waters, D.L.E., Edwards, M.A., Bowen, S.G. and Rice, N. 2011. Chloroplast genome sequences from total DNA for plant identification. Plant Biotechnology Journal, 9: 328-333.

Norup, M.V. 2006. A molecular systematic study of *Heterospathe* and *Rhopaloblaste* (Arecaceae, Areceae). Master's Thesis, University of Aarhus.

Notredame, C., Higgins, D. and Heringa, J. 2000. T-coffee: A novel method for multiple sequence alignments. Journal of Molecular Biology, 302: 205- 217.

Olmstead, R.G., Palmer, J.D. 1994. Chloroplast DNA systematics: A review of methods and data analysis. American Journal of Botany, 81: 1205-1224.

Ou, C.Y., Ciesielski, C.A., Myers, G., Bandea, C.I., Luo, C.C., Korber, B.T., Mullins, J.I., Schochetman, G., Berkelman, R.L. and Economou, A.N. 1992. Molecular epidemiology of HIV transmission in a dental practice. Science, 256(5060): 1165-1171.

Oxelman, B., Liden, M. and Berglund, D. 1997. Chloroplast rps16 intron phylogeny of the tribe Sileneae (Caryophyllaceae). Plant Systematics and Evolution, 206: 393-410.

Oxelman, B. and Bremer, B. 2000. Discovery of paralogous nuclear gene sequences coding for the second-largest subunit of RNA polymerase II (RPB2) and their phylogenetic utility in *Gentianales* of the asterids. Molecular Biology and Evolution, 17: 1131-1145.

Oxelman, B., Yoshikawa, N., Mcconaughy, B.L., Luo, J., Denton, A.L. and Hall, B.D. 2004. *RPB2* gene phylogeny in flowering plants, with particular emphasis on asterids. Molecular Phylogenetics and Evolution, 32: 462-479.

Palmer, 1992. Mitochondrial DNA in plant systematic: application and limitations In: Soltis, P.S., Soltis, D.E. and Doyle, J.J. (eds.), Molecular systematic of plants. London.

Palmer, J.D. 1985. Comparative organization of chloroplast genomes. Annual Review of Genetics, 19: 325–354.

Pang, X., Song, J., Zhu, Y., Xu, H., Huang, L. and Chen, S. 2010. Applying plant DNA barcodes for *Rosaceae* species identification. Cladistics, 27: 165-170.

Paracchini, V., Petrillo, M., Lievens, A., Puertas Gallardo, A., Martinsohn, J.T., Hofherr, J., Maquet, A., Silva, A.P.B., Kagkli, D.M., Querci, M., Patak, A. and Angers-Loustau A. 2017. Novel nuclear barcode regions for the identification of flatfish species. Food Control, 79: 297-308.

Pennington, R.T. and Dick, C.W. 2004. The role of immigrants in the assembly of the South American rainforest tree flora. Philosophical Transactions of the Royal Society B: Biological Sciences, 359, 1611-1622.

Percy, D.M., Argus, G.W., Cronk, Q.C., Fazekas, A.J., Kesanakurti, P.R., Burgess, K.S., Husband, B.C., Newmaster, S.G., Barrett, S.C. and Graham, S.W. 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.

Percy, D.M., Garver, A.M., Wagner, W.L., James, H.F., Cunningham, C.W., Miller, S.E., and Fleischer, R.C. 2008. Progressive island colonization and ancient origin of Hawaiian Metrosideros (*Myrtaceae*). Proceedings of the Royal Society *of* London. *Series* B, Biological Sciences, 275: 1479-1490.

Peters, H.A., Pauw, A., Silman, M.R., Terborgh, J.W. 2004. Falling palm fronds structure Amazonian rainforest sapling communities. Proceedings of the Royal Society B: Biological Sciences, 271: S367-S369.

Pfeil, B.E., Brubaker, C.L., Craven, L.A. and Crisp, M.D., 2004. Paralogy and orthology in the Malvaceae RPB2 gene family: investigation of gene duplication in hibiscus. Molecular Biology and Evolution, 21: 1428-1437.

Phadtare, N.R. and Kulkarni, A.R. 1984. Affinity of the genus Q*uilonipollenites* with the Malaysian palm *Eugeissona* Griffith. Pollen et Spores, 26: 217-226.

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.

Pintaud, J.C. 1999. A cladistic analysis of Archontophoenicinae (Palmae, Areceae), based on morphological and anatomical characters. Memoirs of the New York Botanical Garden, 83: 279 - 284.

Posada, D. and Buckley, T.R. 2004. Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. Systematic Biology, 53: 793-808.

Qiu, Y.L., Lee, J., Bernasconi-Quadroni, F., Soltis, D.E., Soltis, P.S., Zonis, M., Zimmer, E.A., Chen, Z., Sauolainen, V. and Chase, M.W. 1999. The earliest angiosperms: evidence from mitochondrial, plastid, and nuclear genomes. Nature, 402: 404-407.

Quandt, D., Müller, K., Stech, M., Frahm, J. P., Frey, W., Hilu, K. W. and Borsch, T. 2004. Molecular evolution of the chloroplast *trnL–F* region in land plants. Monographs in systematic botany from the Missouri Botanical Garden, 98: 13-37.

Rage, J.C. 1996. Le peuplement animal de Madagascar: unecomposante venue de Laurasie est-elle envisageable? In: Biogéographie de Madagascar, W R Lourenco (ed), Paris: ORSTOM, pp. 27-35.

Ramanujam, C.O.K. 2004. Palms through ages in southern India: A reconnaissance. Palaeobotanist, 53: 1-4.

Rambaut A. 2009. FigTree v1.3.1. Available from: http://tree.bio.ed.ac.uk/software/figtree

Rambaut, A. and Drummond, A.J., 2009. TRACER v1.5. http://tree.bio.ed.ac.uk/software/tracer/

Randall, D.C., Fleck, N.L., Shneerson, J.M. and File, S.E. 2004. The cognitive-enhancing properties of modafinil are limited in non-sleep-deprived middle-aged volunteers. Pharmacology Biochemistry and Behavior, 77: 547 - 55.

Rao, A.R. and Achuthan, V. 1969. Two new pieces of petrified palm woods from the Deccan Intertrappean beds of Mohgaonkalan, (M.P), India. In : Santapau H, Ghosh AK, Chanda S, Roy, S.K. and Chaudhari, S.K. (eds.), J Sen Memorial Volume, Ekta Press, Calcutta, pp. 319-329.

Rao, K.P. and Ramanujam, C.G.K. 1978. Palynology of the Neogene Quilon Beds of Kerala State in South India I. Spores of pteridophytes and pollen of monocotyledons. Palaeobotanist, 25: 397-427.

Ratnasingham, S. and Hebert, P.D.N. 2007. BOLD: The Barcode of Life Data System (www.barcodinglife.org). Molecular Ecology Notes, 7: 355-364.

Raxworthy, C. J., Forstner, M. R. J. and Nussbaum, R. A. 2002. Chameleon radiation by oceanic dispersal. Nature, 415:784-787.

Renuka, C. 1986. A new species of *Calamus* (Palmae) from India. Kew Bulletin, 42(2): 433-435.

Renuka, C. 1992. Rattans of the Western Ghats: a taxonomic manual. Kerala Forest Research Institute, Peechi, India. pp. 61.

Renuka, C. 1994. Genetic diversity and conservation of rattans. In: V. Ramanatha Rao and A. N. Rao (eds.), Biodiversity, Genetic Resources and Conservation working Group. Proceedings of the First NBAR workshop. *7* - 9 November, pp. 39 – 44.

Renuka, C. 1995. Reproductive biology of rattans. Report of an Expert Consultation held at Los Banos. In*:* Williams, J.T., Rao, R. I.V., Rao, A.N. (eds.), Genetic Enhancement of Bamboo and Rattan, Philippines, India, pp. 135-139.

Renuka, C. 1999. Notes on the identity of *Calamus delessertianus* Becc. Rheedea. 9(1): 81-84.

Renuka, C. 2001. Palms of India: status, threats and conservation strategies. In: Shaanker, U.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.

Renuka, C. and Bhat, K.M. 1987. On the occurrence of *Calamus metzianus* Schlecht in Kerala. Current Science, 56: *3*13- 314.

Renuka, C., Bhat, K.V. and Pandalai, R.C. 2010. Rattans of India. Kerala Forest Research Institute, Kerala, India. pp 1- 339.

Renuka, C., Sreekumar, V.B. 2006. *Calamus rivalis* Thw. ex Trim. (Arecaceae): A New Record For India. Journal of Economic and Taxonomic Botany, 30 (2):277-279

Rieseberg, L.H. and Soltis, D.E. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evolutionary Trends in Plant, 5: 65-84.

Rogers, S.O. and Bendich, A.J., 1987. Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. Plant Molecular Biology, 6: 339-345.

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.

Roncal, J., Kahn, F., Millan, B., Couvreur, T.L.P. and Pintaud, J.C. 2013. Cenozoic colonization and diversification patterns of tropical American palms: evidence from Astrocaryum (Arecaceae). Botanical Journal of the Linnean Society, 171: 120-139

Ronquist, F. and Huelsenbeck, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 19(12): 1572-1574.

Roy, S., Tyagi, A., Shukla, V., Kumar, A., Singh, U.M., Chaudhary, B.L., Datt, B., Bag, S.K., Singh, P.K., Nair, N.K., Husain, T. and Tuli, R. 2010. Universal plant DNA barcode loci may not work in complex groups: a case study with Indian berberis species. PLoS ONE, 5(10): e13674.

Saitou, N. and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution, 4(4): 406-425.

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

Sass, C., Little, D.P., Stevenson, D.W. and Specht, C.D. 2007. DNA Barcoding in the *Cycadales*: Testing the Potential of Proposed Barcoding Markers for Species Identification of *Cycads*. PLoS ONE, 2(11): e1154.

Saxena, R.K. 1992. Neyveli lignites and associated sediments – their palynology, palaeoecology, correlation and age. The Palaeobotanist, 40: 345-353.

Scarcelli, N., Barnaud, A., Eiserhardt, W., Treier, U.A., Seveno, M., d'Anfray, A., Vigouroux, Y. and Pintaud, J.C. 2011. A Set of 100 chloroplast primer pairs to study population genetics and phylogeny in Monocotyledons. PLoS ONE, 6(5): e19954.

Schrank, E. 1994. Palynology of the Yesomma Formation in Northern Somalia: a study of pollen spores and associated phytoplankton from the late Cretaceous Palmae Province. Palaeontographica Abteilung B, 231: 63-112.

Searls, D.B. 2003. Pharmacophylogenomics: gene, evolution, and drug targets. Nature Reviews Drug Discovery, 2: 613-623.

Senthilkumar, U. 2015. Phylogenetics and phylogeography of climbing palms of India (Tribe Calameae, Arecaceae). Doctoral thesis, Madras Christian College, India.

Shanley, P., Pierce, A.R., Laird, S.A., Guillen, S.A. 2002. Tapping the Green Market: Certification and Management of Non-Timber Forest Products. London: Earthscan

Shaw, J., Lickey, E.B., Schilling, E.E. and Small, R.L. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. American Journal of Botany, 94: 275 - 288.

Shindyalov, I.N., Kolchanov, N.A. and Sander, C. 1994. Can three dimensional contacts in protein structures be predicted by analysis of correlated mutations? Protein Engineering, Design and Selection, 7(3): 349-358.

Shivji, M., Clarke, C., Pank, M. 2002. Genetic identification of pelagic shark body parts for conservation and trade monitoring. Conservation Biology, 16: 1036-1047.

Simpson, G.G. 1961. Principles of animal taxonomy. Columbia University Press, New York. p. 247.

Singh, J.S. 2012. Biodiversity: An overview The Proceedings of the National Academy of Sciences, India, Section B*:* Biological Sciences, 82(2): 239.

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.

Sneath, P.H.A. and Sokal, R.R. 1973. "Numerical Taxonomy," in Numerical Taxonomy San Francisco, CA: Freeman, pp. 230-234.

Soltis, D.E., Soltis, P.S. and Doyle, J.J. 1998. Molecular systematics of plants II: DNA sequencing. Kluwer Academic, Boston.

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

SPSS Inc. Released 2007. SPSS for Windows, Version 17.0. Chicago, SPSS Inc

Sreeja, P. and Khaleel, K.M. 2010. Status of Mangroves in Thekkumbad, Kannur, Kerala, Journal of Experimental Science, 1 (8): 01-02.

Sreekumar, V.B. 2005. Systematics and phylogeny of the genus *Calamus* L. (Arecaceae) in the Western Ghats. Doctoral thesis, University of Calicut, India.

Sreekumar, V.B. and Henderson, A. 2014. Nomenclatural notes on Indian *Calamus* (Arecaceae). Phytotaxa, 166(2):145-149

Sreekumar, V.B. and Renuka, C. 2006. Assessment of genetic diversity in *Calamus thwaitesii* BECC. (Arecaceae) using RAPD markers. Biochemical Systematics and Ecology, 34: 397-405.

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.

Srivastava, G., Mehrotra, R.C. and Bauer, H. 2012. Palm leaves from the Late Oligocene sediments of Makum Coalfield, Assam, India. Journal of Earth System Sciences, 121(3): 747-754.

Srivastava, R., Srivastava, G., Dilcher, D.L. 2014. Coryphoid Palm Leaf Fossils from the Maastrichtian–Danian of Central India with Remarks on Phytogeography of the Coryphoideae (Arecaceae). PLoS ONE, 9(11): e111738.

Srivathsan, A. and Meier, R. 2012. On the inappropriate use of Kimura-2-parameter (K2P) divergences in the DNA-barcoding literature. Cladistics, 28:190-194.

Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics, 22(21): 2688-2690.

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.

Stiegel, S., Kessler, M., Getto, D., Thonhofer, J. and Siebert, S.F. 2011. Elevational patterns of species richness and density of rattan palms (Arecaceae:

Calamoideae) in Central Sulawesi, Indonesia. Biodiversity Conservation, 20: 1987-2005.

Stoeckle, M. 2003. Taxonomy, DNA, and the Bar Code of Life. BioScience, 53(9): 2-3.

Sunderland, T.C.H. 2012. A taxonomic revision of the rattans of Africa (Arecaceae: Calamoideae). Phytotaxa, 51: 1-7

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C. and Willerslev, E. 2007. Power and limitations of the chloroplast *trnL* (UAA) intron for plant DNA barcoding. Nucleic Acids Research, 35: e14.

Taberlet, P., Gielly, L., Pautou, G. and Bouvet, J. 1991. Universal primers for amplification of three non-coding regions of the chloroplast DNA. Plant Molecular Biology, 17: 1105-1109.

Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics, 105: 437-460.

Tamura, K., Battistuzzi, F.U., Billing-Ross, P., Murillo, O., Filipski, A. and Kumar, S. 2012. Estimating divergence times in large molecular phylogenies. Proceedings of the National Academy of Sciences USA, 109: 19333-19338.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution, 30: 2725-2729.

Techen, N., Parveen, I., Pan, Z. and Khan, I.A. 2014. DNA barcoding of medicinal plant material for identification. Current Opinion in Biotechnology, 25: 103-110.

Thomas, M.M., Garwood, N.C., Baker, W.J., Henderson, S.A., Russell, S.J., Hodel, D.R., Beteman, R.M. 2006. Molecular phylogeny of the palm genus *Chamaedorea*, based on the low-copy nuclear genes PRK and RPB2. Molecular Phylogenetics and Evolution, 38(2): 398-415.

Tomlinson, P.B. 1962. The leaf bases in palms its morphology and mechanical biology. Journal of the Arnold Arboretum, 93: 23-45.

Tomlinson, P.B. 2006. The uniqueness of palms. Botanical Journal of Linnean Society, 151: 5-14.

Trail, J. 1877. Descriptions of new species and varieties of palms collected in the valley of the Amazon in north Brazil. Journal of Botany, 14: 323- 333.

Trenel, P., Gustafsson, M.H.G., Baker, W.J., Asmussen-Lange, 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.

Tsumura, Y., Kawahara, T., Wickneswari, R. and Yoshimura, K. 1996. Molecular phylogeny of Dipterocarpaceae in Southeast Asia using RFLP of PCR-amplified chloroplast genes. Theoretical and Applied Genetics, 93: 22-29.

Uhl, N.W. and Dransfield, J. 1987. Genera Palmarum: a classification of palms based on the work of H. E. Moore, Jr. International Palm Society and L. H. Bailey Hortorium, Lawrence, Kansas, USA, Allen Press. pp. 1- 610.

Uhl, N.W. and Dransfield. J. 1999. Genera Palmarum: after ten years. In: Henderson A., Borchsenius E (eds.), Evolution and classification of palms. Memoirs of the New York Botanical Garden, 83: 245-253.

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. and Humphries, C. J. (eds.), Monocotyledons: Systematics and Evolution, 2: 623-661.

Umapathy, S., Duvuru, N., Munivenkatappa, S., Uma Shaanker, R. and Gudasalamani, R. 2015. Species delimitation in congenerics of Genus Daemonorops from India using DNA barcodes. Communications in Plant Science, 5 (1-2): 1-8.

Valentini, A., Miquel, C., Nawaz, M.A., Bellemain, E., Pompanom, F., Gielly, L., Cruaud, C., Nascetti, G., Wincker, P., Swenson, J.E. and Taberlet, P. 2009. New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the *trn*L approach. Molecular Ecology Resources, 9: 51-60.

Van de Wiel, C.C.M., van der Schoot, J., van Valkenburg, J.L., Duistermaat, C.H. and Smulders, M.J.M. 2009. DNA barcoding discriminates the noxious invasive plant species, floating pennywort *(Hydrocotyle ranunculoides* L.f.), from non-invasive relatives. Molecular Ecology Resources, 9: 1086-1091.

Vijayan, K. and Tsou, C. 2010. DNA barcoding in plants: taxonomy in a new perspective. Current Science, 99: 1530-1541.

Wakasugi, T., Tsudzuki, J., Ito, S., Nakashima, K., Sudzuki, T. and Sugiura, M. 1994. Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. Proceedings of the National Academy of Sciences USA, 91: 9794- 9798.

Wakasugi, T., Tsudzuki, T. and Sugiura, M. 2001. The genomic of land plant chloroplasts: gene content and alteration of genomic information by RNA editing. Photosynthesis Research, 70: 107- 118.

Walter, K.S. and Gillett, H.J. 1998. IUCN Red list of threatened plants. IUCN, Gland and Cambridge.

Wegener, A. 1912. Die Herausbildung der Grossformen der Erdinde (Kontinente und Ozeane), auf geophysikalischer Grundlage, Petermanns Geographische Mitteilungen, 63:185-195 as cited in Demhardt, I. J. 2005. Alfred Wegener's hypothesis on continental drift and its discussion in Petermanns Geographische mitteilungen, Polarforschung, 75 (1): 29 - 35.

Wheeler, Q.D. and Meier, R. 2000. Species concepts and phylogenetic theory: a debate. Columbia University Press, New York, NY. 230 p.

White, T.J., Bruns, T., Lee, S. and Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J., (eds.), PCR Protocols: A Guide to Methods and Applications. Academic Press: San Diego, USA, pp. 315-322.

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.

Wolfe, K.H., Sharp, P.M., and Li, W.H. 1989. Rates of synonymous substitution in plant nuclear genes. Journal of Molecular Evolution, 29: 208–211.

Wolfe, K.H., Li, W.H. and Sharp, P.M. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. Proceedings of the National Academy of Sciences, USA, pp. 84: 9054–9058.

Yan, Y., Harris, A.J. and Xingjin, H. 2010. S-DIVA (Statistical Dispersal-Vicariance Analysis): a tool for inferring biogeographic histories. Molecular Phylogenetics and Evolution, 56: 848- 850.

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. Annals of Botanici Fennici, 49: 319- 330.

Yasui, Y. and Ohnishi, O. 1998. Inter-specific relationships in *Fagopyrum* (Polygonaceae) revealed by the nucleotide sequences of the *rbcL* and *accD* genes and their intergenic region. American Journal of Botany, 85: 1134-1142.

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

Yu, Y., Harris, A.J. and He, X. 2011b. RASP (Reconstruct Ancestral State in Phylogenies) 2.0 beta. Available at: <http://mnh.scu.edu.cn/soft/blog/ RASP>.

Yuan, Y.M., Wohlhauser, S., Mo¨ller, M., Chassot, P., Mansion, G., Grant, J., Kupfer, P. and J. Klackenberg. 2003. Monophyly and relationships of the tribe Exaceae (Gentianaceae) inferred from nuclear ribosomal and chloroplast DNA sequences. Molecular Phylogenetics and Evolution, 28: 500-517.

Zhang, J., Kapli, P., Pavlidis, P. and Stamatakis, A. 2013. A general species delimitation method with applications to phylogenetic placements. Bioinformatics, 29: 2869-2876.

Zona, S., Francisco-Ortega, J., Jestrow, B., Baker, W.J. and Lewis, C.E. 2011. Molecular phylogenetics of the palm sub-tribe Ptychospermatinae (Arecaceae). American Journal of Botany, 98(10): 1716-1726.

Zurawski, G., Perrot, B., Bottomley, W. and Whitefield, P.R. 1981. The structure of the gene for the large subunit of ribulose-1, 5- bisphosphate carboxylase from spinach chloroplast DNA. Nucleic Acids Research, 9: 3251-3270.