Systematics and biogeography of the pantropical genus *Manilkara* Adans. (Sapotaceae)

A thesis submitted for the degree of Doctor of Philosophy

Kate Armstrong

Institute of Evolutionary Biology
School of Biological Sciences
University of Edinburgh
&
Royal Botanic Garden Edinburgh

October 2010

Declaration:

I hereby declare that the work contained in this thesis is my own, unless otherwise acknowledged and cited. This thesis has not in whole or part been previously presented for any degree.

Kate Armstrong
Edinburgh
October 28, 2010

The image on the title page is of *Manilkara hexandra*, and was taken in Thailand by David Middleton (RBGE).

Abstract

Mechanisms for the generation of biodiversity in species-rich biomes such as rain forests remain unclear. Molecular phylogenies using DNA sequence data, calibrated with a temporal dimension offer a means of addressing this question, enabling the testing of different hypotheses on biogeographic histories and causes of diversification. *Manilkara* is a genus of trees in the Sapotaceae consisting of ~79 species distributed throughout the tropics (30 South and Central American, 35 African and 14 Southeast Asian). This species diversity in all major tropical regions of the globe makes it an ideal candidate for in-depth biogeographic studies.

Maximum parsimony and Bayesian analyses of nuclear (ITS) and chloroplast (*rpl32-trnL*, *rps16-trnK* and *trnS-trnFM*) sequences were used to reconstruct a species level phylogeny of *Manilkara* and related genera in the tribe Mimusopeae. *Manilkara*, as currently defined, is not monophyletic due to the placement of three Asian taxa (*M. fasciculata*, *M. dissecta* and *M. udoido*), which are more closely related to the Madagascan genera *Labourdonnaisia* and *Faucherea* than to *Manilkara s.s*. and need to be re-circumscribed in a new genus. *Letestua* is nested in *Manilkara* and the genera *Faucherea* and *Labourdonnaisia* are not monophyletic. Nuclear and chloroplast datasets were mostly congruent, however, three instances of hard incongruence were demonstrated, suggesting chloroplast capture events.

Bayesian analyses of ITS sequences using a relaxed molecular clock calibrated with fossils, focused on testing biogeographical hypotheses on the origin of *Manilkara*'s pantropical disjunct distribution and spatio-temporal diversification patterns on each continent. Mimusopeae, originated during the Eocene ~46-57 Ma and fossil evidence supports its existence in the boreotopical region of the northern hemisphere during this time. This suggests that the tribe may have evolved there and found refuge in Africa when Oligocene climatic cooling made higher latitudes uninhabitable for megathermal taxa. The subtribe Manilkarinae was resolved as ~42-36 Myo. These ages fall on the Eocene-Oligocene boundary and the crown node age coincides with the onset of Oligocene cooling and the closing of the boreotropical route. The genus *Manilkara* is estimated to have evolved ~36-33 Ma. The current distribution of the genus could not, therefore, have been the result of Gondwanan vicariance or migration through the boreotropics, but results instead support long distance dispersal as an important factor influencing the distribution of the group.

Resolution along the backbone of the phylogeny is weak and the area of origin is, therefore, difficult to determine. However, all sister taxa to *Manilkara* are African and this suggests that the most likely explanation is an African origin for the genus with subsequent inter-continental dispersal during the Miocene. *Manilkara* spread from Africa to the Neotropics and Asia via at least three separate long distance dispersal events. A single lineage dispersed to the Neotropics ~27-21 Ma and spread across the Isthmus of Panama before its closure. Another lineage dispersed to Southeast Asia ~30-25 Ma from mainland Africa and subsequently diversified throughout the region. A third dispersal from Madagascar to the Sahul Shelf, occurred ~31-16 Ma in the *M. fasciculata*/*dissecta*/*udoido* lineage.

In South America, diversification is consistent with both aridification and the rearrangement of drainage patterns in the Amazon basin as a result of Andean orogeny. The Atlantic coastal forest clade and the Amazonian clade of *Manilkara* split from one another ~14 Ma, at approximately the same time as the dry biomes of the Cerrado and Caatinga were forming between them. In Africa diversification coincides with Tertiary cycles of aridification and uplift of the east African plateaux. In Southeast Asia Wallace's Line did not affect the dispersal of *Manilkara*. Instead, the limiting factor was the appearance of land in New Guinea ~10 Ma, which coincides with the dispersal and establishment of new taxa east of Wallace's Line.

Spatio-temporal patterns of diversification in *Manilkara* were compared to those of 34 other wet tropical genera which have intercontinental disjunctions. Ages of disjunctions ranged from the Eocene to the Pliocene, indicating that compilation of the tropical rain forest biome is a dynamic process which has been occurring throughout the Tertiary. Recent migration via long distance dispersal is a significant phenomenon in biome construction. Geo-climatic events have also been shown to be important drivers of diversification in all continental regions.

Acknowledgements

Firstly, I would like to give my sincere thanks and gratitude to my supervisors James Richardson, Richard Milne and Graham Stone, who have been extremely supportive throughout my project and have always made the time to discuss my ideas and queries, to steer me in the right direction, and to read and comment on several drafts of my thesis.

This doctoral research was made possible by a scholarship from the Torrance Bequest at the University of Edinburgh, and for this support, I am immensely grateful.

My research has greatly benefited from discussions with many people, notably Vanessa Plana, who suggested this topic in the first place and who advised on all aspects of *Manilkara* taxonomy; David Harris, who was always willing to look at *Manilkara* specimens and give his taxonomic opinion; Laurent Gautier who advised me on Madagascan species of *Manilkara* and guided me in understanding the taxonomy of *Faucherea* and *Labourdonnaisia*; Terry Pennington who is the source of knowledge for all things Sapotaceous; and Josh Clayton, who sowed the seeds for this project in his M.Sc. thesis.

Thanks to Helen Hoy for arranging my specimen loans and to all the herbaria, which lent me their specimens and allowed me to sample them for DNA extraction: A, BR, BISH, E, FHO, G, K, L, MO, NY, P and WAG. Thanks also to Adele Smith for support in the herbarium and Martin Pullan for technical help with PADME.

I am grateful to Graham Hardy for providing invaluable assistance obtaining extensive interlibrary loans on obscure fossil references; Jonathan Wingerath for taking digital images of all the Mimusopeae fossils in the Smithsonian Institute collections and sending them to me for inspection; to Shusheng Hu for images of fossils from the Peabody Museum; to Terry Pennington for his opinion on the identification of these Sapotaceae macrofossils; to Madeline Harley for identifying fossil Sapotaceae pollen and for answering my numerous queries on the subject; and to Aaron Pan for helpful suggestions about fossil resources and for sharing information about his Ethiopian Sapotaceae leaf fossils, which are used as a calibration point in this thesis.

Carrying out fieldwork in Indonesia enriched my understanding of *Manilkara*'s ecology and Southeast Asian biogeography. I am very grateful to RISTEK for granting me permission to collect specimens in West Papua and Sulawesi and to the staff of the Herbarium Bogoriense, particularly the director Eko Baroto Walujo and Teguh Triono, for supporting my research there. Thanks also to Charlie Danny Heatubun and Rani Asmarayani for accompanying me on my fieldwork and making sure everything ran smoothly. Fieldwork would not have been possible without grants from the Carnegie Trust, the Systematics Association, the Davis Expedition Fund, and the Royal Geographical Society, whose support is very gratefully acknowledged. Thanks also to Kate Eden for mounting my specimens collected during this fieldwork.

Many people kindly provided silica-gel dried leaf samples, DNA isolations and sequence data, which have greatly enhanced this project. For this, thanks are due to Arne Anderberg, Jerome Chave, Laurent Gautier, David Harris, Yamama Naciri, Terry Pennington, Rachun Pooma and Adi Suprapto.

Venturing into the lab was a new experience as part of this Ph.D. and I have many people to thank for helping me make the transition from a field botanist to a gel jock. I am very grateful to Alex Clark, Michelle Hollingsworth, Jane Squirrell and Ruth McGregor for expert advice and general support in the lab. Thanks also to Vania Azevedo for *Manilkara* microsatellite primers and Damon Little for advice on amplifying problematic samples and providing the extra special Taq to do the trick. I also received a great amount of support from my fellow Ph.D. students, particularly Chantel Davies, who facilitated my first tentative steps

in the lab and Bhaskar Adhikari, and Daniel Thomas, who always had excellent suggestions of how to make protocols work well.

I am indebted to James Nicholls for teaching me the ins and outs of Bayesian analyses, how to use MrBayes, BEAST and Bayes factor testing and for generally being incredibly patient, helpful and up-beat throughout the process. Thanks also to Pablo Fuentes and the rest of the Stone lab for their support and for letting me monopolize their BigMac during endless MrBayes and BEAST runs. Thanks to Duncan Reddish for setting up space on the RBGE server for me to run even more analyses in tandem.

Toby Pennington and Matt Lavin both took time to help me with r8s and answer all my queries about the software; for that I am very grateful. While analyzing my data, both James Tosh and Laura Kelly helped me by answering numerous questions on the finer points of molecular analysis. Thanks also to Nikki Harrison, Tiina Sarkinen and Daniel Thomas for advice on ancestral area reconstruction in BEAST.

Throughout this process my fellow Ph.D. students have been firm friends, providing both moral support and technical advice: Bhaskar Adhikari, Mobina Shaukat Ali, Rhiannon Chrichton, Chantel Davies, Jane Droop, Emma Goodyer, Nikki Harrison, Sabina Knees, Tobias Marczweski, Sumudu Rubasinghe, Tiina Sarkinen, Daniel Thomas and Alex Twyford.

A special thanks goes to Alex Twyford, who read and made valuable comments on an earlier version of this manuscript and to Bhaskar Adhikari, who has been a great help in checking through my references and assisting me with the technicalities of assembling a thesis. I am also grateful to Caroline Muir, who took the time to show me the basics of InDesign.

Thanks to Tiina Sarkinen, Frans Mortengren, Andy Matthews and Susanna Holland for buoying me up when I felt deflated and providing lovely food, wonderful conversation and plenty of laughter during the writing of this thesis.

Finally, I am eternally grateful to my parents, who have been unwavering in their encouragement and support of all my botanical pursuits throughout the years.

TABLE OF CONTENTS

Declaration......i
Abstract......ii
Acknowledgements......iv

Chapter I – Introduction
1.1 Introduction......1
1.2 History of early biogeographical thought......1
1.3 Darwin, Hooker & Wallace – dispersalists versus extensionists......2
1.4 Vicariance versus dispersal biogeography......3
1.5 Current trends in biogeography......4
1.6 Why *Manilkara*?......5
1.7 Aims of the thesis and overview of chapter content......6

Chapter II – Taxonomic history of *Manilkara*
2.1 Introduction to classification in the Sapotaceae......8
2.2 Classification within tribe Mimusopeae......9
2.3 Classification within subtribe Manilkarinae......11
2.4 History of the names Sapotaceae and *Manilkara*......12
2.5 History of the classification of *Manilkara*......13
2.6 Recent advances in Sapotaceae classification......15
2.7 Taxonomic aims of the thesis......16
2.8 Appendix......19

Chapter III – Molecular and phylogenetic materials and methods
3.1 Choice of plant material for DNA extraction......20
  3.1.1 Acquisition of material and sequences......20
  3.1.2 Selection of species for inclusion in analyses......20
  3.1.3 Outgroups to *Manilkara*......21
  3.1.4 Material removal from herbarium specimens......21
  3.1.5 Sampling strategy......22
3.2 Choice of genomic regions......22
  3.2.1 Nuclear regions......22
  3.2.2 Chloroplast regions......24
  3.2.3 Microsatellite flanking regions......25
3.3 Lab protocols......26
  3.3.1 DNA extraction......26
  3.3.2 PCR amplification......26
  3.3.3 Nested PCR reactions for DNA extracted from herbarium material......27
  3.3.4 Purification and sequencing......27
3.4 Sequence editing and alignment......28
3.5 Regional coverage of species sequenced......28
3.6 Phylogenetic tools and methods......28
  3.6.1 Phylogeny reconstruction methods......28
  3.6.2 Maximum parsimony......29

3.6.3 Maximum likelihood....29
3.6.4 Bayesian inference....29
3.6.5 Methods chosen for use in this study....30
3.7 Appendices....31

Chapter IV – Phylogenetic support for *Manilkara*: Monophyly and generic delimitation in the tribe Mimusopeae and subtribe Manilkarinae
4.1 Introduction....40
4.2 Phylogenetic methods....40
4.2.1 Taxon selection and sequence data....40
4.2.2 Chloroplast gap-coding....41
4.2.3 Why nuclear and chloroplast regions were not combined....41
4.2.4 Parsimony analysis settings in PAUP*....42
4.2.5 Bayesian analysis....42
4.2.5.1 MrBayes settings....42
4.2.5.2 Partitioning and model selection....42
4.2.6 Interpretation of support values....43
4.3 Results....43
4.3.1 Model selection and Bayes factor tests....43
4.3.2 ITS phylogenies – overview of clade support....46
4.3.3 Chloroplast phylogenies – overview of clade support....49
4.3.4 Comparison of nuclear and chloroplast phylogenies....52
4.3.4.1 Congruence....52
4.3.4.2 Incongruence....53
4.3.4.3 Comparison of phylogenetic methods, genes and a summary of monophyly tests....54
4.4 Discussion....55
4.4.1 Delimitation of the tribe Mimusopeae and subtribes Mimusopinae, Manilkarinae and Glueminae....55
4.4.2 Delimitation of genera within the Manilkarinae....56
4.4.2.1 *Labramia*, *Labourdonnaisia* and *Faucherea*....56
4.4.2.2 *Manilkara* and *Letestua*....57
4.4.3 Implications for classification – summary of findings....58
4.4.4 Delimitation of species complexes within *Manilkara*....58
4.4.5 Hard incongruence between nuclear and chloroplast trees - evidence for chloroplast capture?....59
4.4.5.1 Clades U and R: chloroplast capture across the Indian Ocean?....60
4.4.5.2 Clade V: chloroplast capture across the Atlantic Ocean?....61
4.4.5.3 Other possible instances of chloroplast capture....62
4.4.5.4 Precedents for chloroplast capture across the geographic boundaries....62
4.5Appendices....63

Chapter V – Historical biogeography of tropical forests and their intercontinental disjunctions

5.1 Overview of regional biogeography and trends in angiosperm disjunctions in the Cenozoic....66
5.2 Origin of eudicots....67
5.3 Gondwanan vicariance....67
5.4 Regional history of the boreotropics....70
5.5 Overview of "interplate dispersal paths for megathermal angiosperms"....75
5.6 Long distance dispersal....76
5.7 Regional biogeographic histories....83
5.7.1 Regional history of South America....83
5.7.2 Regional history of Africa....88
5.7.3 Regional history of Asia....92
5.7.3.1 Southeast Asia....92
5.7.3.2 Indian subcontinent....96

Chapter VI – The Sapotaceae fossil record

6.1 Introduction....99
6.2 Overview of the Sapotaceae fossil record at family level....99
6.2.1 Cretaceous-Paleocene....99
6.2.2 Eocene....100
6.2.3 Oligocene....101
6.2.4 Miocene....102
6.2.5 Pleistocene....103
6.2.6 Summary of Sapotaceae fossil records....104
6.3 Fossils from the tribe Mimusopeae....104
6.3.1 Cretaceous-Paleocene and Eocene....104
6.3.2 Oligocene....105
6.3.3 Miocene and Pleistocene....106
6.3.4 Summary of Mimusopeae fossil records....107
6.3.5 Appendices....109

Chapter VII – Molecular dating and biogeographical analysis....123
7.1 Introduction....123
7.2 Dating phylogenies....123
7.2.1 Molecular clock theory....123
7.2.2 Approaches for temporal calibration of phylogenies....125
7.2.2.1 Use of fossils as calibration points....125
7.2.2.2 Secondary calibration methods in the absence of fossil data....126
7.2.2.3 Calibration using geological events....126
7.2.2.4 Applying rates from other studies....127
7.2.3 Using a relaxed molecular clock....127
7.3 Methodological approach....129
7.3.1 Introduction to methodological approaches used in this chapter....129

7.3.2 Choice of fossil calibration points..129
7.3.3 Dataset for dating analyses: taxa and outgroup selection.....................................132
7.3.4 Bayes factor tests for model selection & clock-likeness in MrBayes & BEAST...132
7.3.5 Molecular dating analyses..133
7.3.5.1 BEAST program and settings...133
7.3.5.2 r8s program settings...135
7.3.5.3 Testing the utility of different fossil calibration points in BEAST & r8s......135
7.3.5.4 Fossil cross-validation...136
7.3.6 Ancestral area reconstruction in BEAST...137
7.3.7 Hypothesis testing with area constraints in MrBayes...138
7.4 Results...139
7.4.1 Molecular dating results...139
7.4.1.1 Bayes factor tests for model selection..139
7.4.1.2 Testing the utility of different fossil calibration points in BEAST and r8s....140
7.4.1.3 Choice of fossil calibration scenario and analytical method.........................144
7.4.1.4 Relaxed uncorrelated lognormal clock versus a strict clock in BEAST.........146
7.4.1.5 Node ages..146
7.4.1.6 Estimated substitution rates..148
7.4.2 Ancestral area reconstruction results...148
7.4.3 Area constraints hypothesis testing results..151
7.5 Discussion..153
7.5.1 Origin and means by which *Manilkara* achieved its pantropical distribution........153
7.5.1.1 Evidence for origin in Africa...153
7.5.1.2 Splits in the phylogeny consistent with Gondwanan vicariance...................153
7.5.1.3 Splits in the phylogeny consistent with the boreotropics hypothesis............153
7.5.1.4 Splits in the phylogeny consistent with long distance dispersal....................155
7.5.1.4.1 Origin of *Manilkara* in Africa followed by long distance dispersal....155
7.5.1.4.2 Dispersal to Madagascar and the Mascarenes...................................157
7.5.1.4.3 Dispersal to South America - followed by dispersal to Central America & the Caribbean islands...157
7.5.1.4.4 Dispersal to Asia...158
7.5.1.4.5 Further dispersal events suggested by hard incongruence between the nuclear and chloroplast datasets...159
7.5.2 Regional diversification in *Manilkara*..159
7.5.2.1 Regional diversification patterns in the Neotropics......................................159
7.5.2.2 Regional diversification patterns in Africa and Madagascar.........................161
7.5.2.3 Regional diversification patterns in Southeast Asia......................................163
7.5.3 Comparison of mean estimated substitution rates across the entire phylogeny.....165
7.6 Conclusions...165
7.7 Appendices..167

Chapter VIII – Conclusions
8.1 Overview of findings....175
8.1.1 Systematics....175
8.1.2 Biogeography....176
8.2 Relevance of this study in relation to other biome level studies....178
8.2.1 Northern and Southern hemisphere meta-analyses....178
8.2.2 Neotropical meta-analyses....178
8.2.3 African meta-analyses....179
8.2.4 Southeast Asian and Australasian meta-analyses....179
8.3 Theories about the historical construction of tropical rain forest....180
8.4 The bigger picture – comparison with other tropical disjunct taxa....181
8.4.1 Methods....181
8.4.2 Results....181
8.4.3 Discussion – implications for tropical rain forest biome age....182
8.4.4 Suggestions for future research....184
8.5 Appendix....185

References....188

LIST OF FIGURES

Figure 1.1 Map depicting number of *Manilkara* species in each main continental region...... 6

Figure 2.1 Illustration of *Manilkara smithiana*..12

Figure 2.2 Illustration of *Manilkara* petal dissection...12

Figure 2.3 Sapotaceae family level phylogeny based on *ndhF* and morphology from Swenson & Anderberg 2005...17

Figure 2.4 Sapotoideae tribal level phylogeny based on *ndhF, trnC-petN, petN-psbM, psbM-trnD* and *trnH-psbA* from Smedmark *et al* 2006..18

Figure 3.1 Diagram of the ITS & ETS regions...23

Figure 3.2 Diagram of regions surveyed in the chloroplast genome.......................................25

Figure 4.1 One out of ten thousand most parsimonious trees reconstructed from the ITS dataset in PAUP*...46

Figure 4.2 Bayesian majority rule consensus tree reconstructed from the ITS dataset in MrBayes..47

Figure 4.3 One out of ten thousand most parsimonious trees reconstructed from the chloroplast dataset in PAUP*..49

Figure 4.4 Bayesian majority rule consensus tree reconstructed from the chloroplast dataset in MrBayes..50

Figure 5.1 Overview of tectonic plates and their direction of movement...............................67

Figure 5.2 Break-up of Gondwana showing vicariance and collision times between continental fragments...68

Figure 5.3 Geological area cladogram representing relationships between Gondwanan fragments and time of phylogenetic splits..69

Figure 5.4 Depiction of boreotropical migration pathways available during the Eocene ThermalMaximum..71

Figure 5.5 Hypothetical area cladogram depicting a boreotropical migration pattern............73

Figure 5.6 Graph of expected successful dispersal and establishment over distance & time 77

Figure 5.7 Modern ocean currents..79

Figure 5.8 Major vegetation types in tropical South America...83

Figure 5.9 Walvis Ridge – Rio Grande Rise Hotspot...85

Figure 5.10 The Proto Greater Antilles...86

Figure 5.11 Hypothetical area cladogram depicting ages of splits between lineages corresponding with paleogeographical phenomena in the Neotropics..................................87

Figure 5.12 Vegetation map of Africa...88

Figure 5.13 Alternative hypotheses of African rainforest origins...91

Figure 5.14 Hypothetical area cladogram depicting diversification in African rain forest taxa coinciding with cycles of aridification....92

Figure 5.15 Map of forest types in Malesia....93

Figure 5.16 Wallace's, Weber's and Lydekker's Lines....95

Figure 5.17 Area cladogram depicting hypothetical dispersal across Wallace's Line in accordance with two paleogeological scenarios....96

Figure 5.18 Area cladogram depicting a scenario for the migration of taxa from Gondwana to Southeast Asia via Indian rafting....98

Figure 6.1 Sapotaceae fossils recorded from the Cretaceous-Paleocene....100

Figure 6.2 Sapotaceae fossils recorded from the Eocene....101

Figure 6.3 Sapotaceae fossils recorded from the Oligocene....102

Figure 6.4 Sapotaceae fossils recorded from the Miocene....103

Figure 6.5 Sapotaceae fossils recorded from the Pleistocene....103

Figure 6.6 Mimusopeae fossils recorded from the Cretaceous-Paleocene....105

Figure 6.7 Mimusopeae fossils recorded from the Eocene....105

Figure 6.8 Mimusopeae fossils recorded from the Oligocene....106

Figure 6.9 Mimusopeae fossils recorded from the Miocene....107

Figure 6.10 Mimusopeae fossils recorded from the Pleistocene....107

Figure 7.1 Illustration of stem and crown nodes in a hypothetical phylogeny....125

Figure 7.2 Cladogram depicting placement of fossil taxa for testing calibration scenarios..131

Figure 7.3 Maps of regions coded for taxa in the BEAST ancestral area analysis....138

Figure 7.4 Age ranges reconstructed for the subtribe Manilkarinae using each of the fossil calibration scenarios and each of the molecular dating methods....141

Figure 7.5 Age ranges reconstructed for the *Faucherea*, *Labourdonnaisia* and small *Manilkara* clade using each of the fossil calibration scenarios and each of the molecular dating methods....142

Figure 7.6 Age ranges reconstructed for the Neotropical *Manilkara* clade using each of the fossil calibration scenarios and each of the molecular dating methods....143

Figure 7.7 Maximum clade credibility chronogram of the ITS dataset derived from an analysis in BEAST....147

Figure 7.8 Maximum clade credibility chronogram of the ITS dataset derived from a phylogeographic analysis in BEAST with ancestral areas reconstructed....149

Figure 7.9 Map indicating age and direction of long distance dispersal events in the tribe Mimusopeae....155

LIST OF TABLES

Table 2.1 Pennington's (1991) tribe Mimusopeae....9

Table 2.2 Diagnostic characters of genera in Pennington's subtribe Manilkarinae....10

Table 4.1 Summary of sequence data and settings....44

Table 4.2 Bayes factor comparison of different models & clock settings in the ITS dataset..44

Table 4.3 ITS dataset Bayes factor model testing part two....45

Table 4.4 Bayes factor comparison of different models and clock settings in the chloroplast dataset....45

Table 4.5 Summary of results of monophyly tests....54

Table 5.1 Examples of dated phylogenetic studies, which conform to a Gondwanan vicariance scenario....70

Table 5.2 Examples of dated phylogenetic studies, which conform to a megathermal boreotropical migration scenario during the Paleocene-Early Oligocene....74

Table 5.3 Examples of dated phylogenetic studies, which conform to a microthermal boreotropical migration scenario from the Mid-Oligocene through the Miocene....74

Table 5.4 Hypothesized rafting times across the Atlantic Ocean, Caribbean Sea and Southeast Indian Ocean....78

Table 5.5 Examples of dated phylogenetic studies, which exhibit long distance dispersal between the Neotropics and Africa....80

Table 5.6 Examples of dated phylogenetic studies, which conform to a scenario of recent long distance dispersal....81

Table 5.7 Examples of dated phylogenetic studies, which conform to a scenario of dispersal into Southeast Asia following rafting on the Indian subcontinent....98

Table 7.1 Taxon sets defined in BEAST....133

Table 7.2 Bayes factor comparison in MrBayes of different models and clock settings on the ITS dataset....139

Table 7.3 Summary of sequence data and settings in BEAST....140

Table 7.4 The molecular and fossil age for each calibration point....144

Table 7.5 Relaxed versus strict clock Bayes factor test....146

Table 7.6 Mean crown & stem node ages in BEAST for the profiled calibration scenarios.148

Table 7.7 Overall substitution rates calculated in BEAST for different fossil calibration scenarios....148

Table 7.8 Percentage likelihood of ancestral areas for profiled nodes....150

Table 7.9 Bayes factor test for area constraints....152

Table 7.10 Inferred long distance dispersal events in *Manilkara* and *Mimusops*..................155

Table 7.11 Summary of regional diversification events within *Manilkara* in relation to climate and geology...164

Table 7.12 ITS substitution rates for other angiosperm genera..165

LIST OF APPENDICES

Appendix 2.1 Illustration of *Manilkara subsericea*....19

Appendix 3.1 List of specimens from which DNA was extracted....31

Appendix 3.2 Species which were not sampled as part of this study....35

Appendix 3.3 ITS primers....35

Appendix 3.4 Chloroplast primers....36

Appendix 3.5 Chloroplast monomorphic loci....37

Appendix 3.6 ITS PCR recipe....37

Appendix 3.7 Chloroplast PCR recipe....37

Appendix 3.8 ITS PCR program A....37

Appendix 3.9 ITS PCR program B....38

Appendix 3.10 Chloroplast PCR program....38

Appendix 3.11 Monomorphic loci PCR program....38

Appendix 3.12 Annealing temperatures for monomorphic loci....38

Appendix 3.13 ExoSAP-IT protocol....39

Appendix 3.14 Exo-SAP-IT PCR purification protocol....39

Appendix 3.15 CEQ sequencing PCR reaction....39

Appendix 3.16 CEQ sequencing PCR protocol....39

Appendix 3.17 CEQ Sequencing PCR clean-up reaction....39

Appendix 3.18 BigDye recipe....39

Appendix 3.19 BigDye sequencing PCR protocol....39

Appendix 4.1 Specimen data for taxa included in the ITS and cpDNA analyses....63

Appendix 6.1 Fossil pollen records from the literature....109

Appendix 6.2 Fossil wood records from the literature....115

Appendix 6.3 Fossil leaf records from the literature....116

Appendix 6.4 Fossil fruit records from the literature....116

Appendix 6.5 British Museum London Clay specimens....117

Appendix 6.6 Smithsonian Institute fossil specimens....118

Appendix 6.7 Yale Peabody Museum fossil specimens....121

Appendix 6.8 University of California Berkeley Museum of Palaeontology fossil specimens 122

Appendix 7.1 Specimen data for taxa included in the different analyses in Chapter VII.....167

Appendix 7.2 Mean node age with 95% HPD age range reconstructed for profiled nodes in BEAST using all fossil calibration scenarios..........172

Appendix 7.3 Mean ages with confidence intervals reconstructed for profiled nodes in r8s using penalized likelihood and all fossil calibration scenarios..........173

Appendix 7.4 Mean ages with confidence intervals reconstructed for profiled nodes in r8s using nonparametric rate smoothing and all fossil calibration scenarios..........174

Appendix 8.1 Surveyed dated phylogenetic studies of genera with intercontinental disjunctions..........185

“Probably only a botanist living some ten or hundred thousands of years later (if botanists are still roaming then) may be more lucky (or not) when put face to face with this group of plants.”

H.J. Lam, 1941 – on the difficulties of revising *Manilkara* and the subtribe Manilkarinae

Chapter I – Introduction

1.1 Introduction

Biodiversity is unevenly distributed across the globe. It is most intensely concentrated in the tropics, particularly in wet tropical forests, which are the most species-rich biomes on the planet. Even within the tropics, there are significant differences between the numbers of taxa found in each of the continental regions. It is estimated that there are *c*. 27,000 species of flowering plants in tropical Africa (Lebrun 2001; Lebrun & Stork 2003), compared with *c*. 90,000 for South America (Thomas 1999) and *c*. 50,000 for Southeast Asia (Whitmore 1998). This uneven species diversity raises fundamental questions about the pattern and tempo of speciation and extinction on each continent and the processes, which have triggered these phenomena in one region rather than another.

In order to investigate current patterns of biodiversity and how tropical forests have been assembled through time, an understanding of historical biogeography and the mechanisms which influence the speciation, spread and extinction of organisms is necessary. The following sections outline the history of biogeography as a discipline and place the current study into context, highlighting the major advances and movements in biogeographical thought throughout the last ~300 years.

1.2 History of early biogeographical thought

Throughout history biologists have recognised that plants and animals are not evenly spread across the world and that similar environments on different continents can harbour very different biotas. The discipline of biogeography has always sought to document and understand spatial patterns of biodiversity and reconstruct the origin, dispersal and extinction of organisms. It utilizes the present distribution of taxa in combination with hypotheses on relationships between these taxa to infer historical connections or dispersal between the landmasses across which the taxa are distributed.

Biogeographic theory has evolved over time and one of its early influential contributors was George-Louis Leclerc, Compte de Buffon, who observed that environmentally similar but isolated regions have distinct assemblages of mammals and birds. From this he concluded that organisms improved or degenerated after dispersing from a centre of origin and that this change was driven by differences in climate. He further hypothesized that during migration, populations became separated and modified over time until tropical biota on each continent became gradually more and more dissimilar. This became the first principle of biogeography, known as “Buffon’s Law” (Buffon, 1761, 1776; Lomolino *et al* 2006).

Alexander von Humboldt, known as the father of phytogeography, expanded Buffon’s Law to include plants. He noticed that plant assemblages are strongly correlated with climate and that in addition to latitudinal gradients, floristic zonation could also be described along elevational zones or “floristic belts” ranging from the equatorial tropics to boreal arctic at the summit (von Humboldt 1805, 1808). Further elaborating upon Buffon’s Law, Augustin P. de

Candolle developed the concept of endemics and areas of endemism when he deduced that there were botanical regions, defined areas which support a certain number of indigenous species. He additionally believed that the few species which were found to be cosmopolitan were exceptions to the law of nature and their wide distributions were due to transport via water, wind, animals and humans (de Candolle 1820, 1855, Lomolino *et al* 2006). The geologist Charles Lyell put forth the idea that geological change was the result of the steady accumulation of minute changes over extremely long time spans and that geological remains from the past could be explained with reference to current, observable processes. Lyell documented evidence for sea level changes and the uplift and erosion of mountains and also found fossil evidence that many tropical life forms had formerly thrived in the now temperate regions of northern Europe (Lyell 1830, Lomolino *et al* 2006).

During this period Buffon's law was explained primarily in relation to climate and external factors, but it was generally acknowledged that external factors alone were an insufficient explanation. The causes of diversification in organisms had yet to be established.

1.3 Darwin, Hooker & Wallace – dispersalists versus extensionists

Charles Darwin, Joseph Dalton Hooker and Alfred Russell Wallace built upon the ideas of Buffon, Humboldt, de Candolle, Lyell and others. Darwin's revolutionary theory of evolution through natural selection provided the basis for understanding how lineages of organisms spread, adapt and change through time and space. He determined that the distribution and disjunction of species can result from long-distance dispersal, which in turn creates areas of endemism (Darwin 1859). The opposing force to dispersalists at the time, were extensionists – those who argued that long distance dispersal across great barriers (such as oceans) was too unlikely to explain disjunct distributions. They, instead, proposed that disjunctions were the result of the spread of organisms via now submerged ancient land bridges and continents.

Extension was championed by Joseph Dalton Hooker (1867, 1877). While he was a proponent of evolution through natural selection, he could not accept the universality of long-distance dispersal and instead suggested that tropical floras were remnants of a once continuous flora, which had been broken up by geological and climatic causes. He further hypothesized land bridges between South America and Africa as well as between Madagascar and India. Hooker developed and applied many of the principles which are now called vicariance biogeography.

Alfred Russell Wallace was also a key player in the development of the discipline of biogeography. He was the first person to analyze faunal regions based on the distribution of multiple groups of terrestrial animals and noticed the striking difference between the fauna of Borneo+Phillipines and that of Sulawesi+New Guinea, on either side of the Makassar Straits. Wallace recognised that this biological boundary (which came to be known as Wallace's Line) was a reflection of the different histories of the landmasses on either side of the divide and provided evidence for plate tectonics long before it was discovered by geologists (Wallace 1860, 1863, 1869, 1876).

While the wide-ranging land bridges envisaged by the extensionists have long been discredited by geology, the movement of landmasses on tectonic plates has come to the fore and the extensionist-dispersalist dispute has evolved into the vicariance versus dispersal debate.

1.4 Vicariance versus dispersal biogeography

Currently, there are two main methods used to explain intercontinental disjunctions: dispersal and vicariance. In dispersal biogeography, taxa disperse across pre-existing barriers such as oceans and mountain ranges, whereas vicariance biogeography is based on the principle that once continuous populations become broken up due to abiotic events such as plate tectonic movement, mountain building and climate change. Because dispersal is unfalsifiable and provides a potential explanation for nearly all distribution patterns, many methods of biogeographic analysis are based on the vicariance model. Vicariance hypotheses can be tested based on the concordance between phylogenetic relationships and distribution patterns of different taxa. Related taxa which exhibit the same distribution pattern are believed to share a common history and to have been influenced by the same isolating geological and climatological events (Lomolino & Heaney 2004, Lomolino *et al* 2006). Vicariance methodology can be broken down into two different approaches depending upon how distribution histories are reconstructed: panbiogeography and cladistic biogeography.

Panbiogeography, founded by Croizat (1952, 1958), is a mapping exercise which does not include phylogenetics. It aims to reconstruct the distribution history of a taxon by plotting the ranges of endemic species on a map and then drawing lines (or “tracks”) connecting the distributions of closely related taxa in different areas. Where tracks of unrelated taxa coincide, they become “generalized tracks,” which are believed to indicate the historical connections between once widespread continuous distributions that have subsequently become fragmented. Theoretically, the ways in which regional biota have developed over time and space could be reconstructed by plotting the generalized tracks on a map. However, this method is widely recognised as being flawed due to its lack of phylogenetic evidence and inability to test temporal hypotheses (Lomolino & Heaney 2004).

The alternative approach of cladistic (or vicariance) biogeography (Nelson 1969, Nelson 1974, Platnick & Nelson 1978, Nelson 1978, Nelson 1981) is based on Hennig’s (1966) principles of phylogenetic systematics and came to the fore just as Wegener’s (1912, 1966) plate tectonic theory began to gain acceptance. It combines cladograms of taxa and their regional distributions to create area cladograms, which represent hypotheses of historical relationships between areas. Repeated patterns of area relationships are believed to be due to the same underlying historical cause, such as the break-up of land masses. Cladistic biogeography assumes that if speciation events are the result of geographic isolation, then a phylogeny represents the relative timing of the separation of disjunct taxa. If these disjunctions are still present in the current distribution of a taxon, then the phylogeny can provide information on the historical relationships between geographic areas as inferred through taxon ancestor-descendant relationships. Therefore, if multiple unrelated taxa exhibit similar distributions and their area cladograms are congruent, then this may support a

vicariance hypothesis for a particular region. However, vicariance hypotheses can be falsified by incongruent area cladograms, disagreement with the fossil record, geology or climatic history. They can also be complicated by extinction, sympatric speciation and widespread, cosmopolitan taxa (Lomolino & Heaney 2004).

As an alternative to vicariance, Platnick & Nelson (1978) suggested that even with information from the fossil record, dispersal hypotheses are difficult to falsify because long-distance dispersal is both improbable and unpredictable. Each dispersal and colonization would be likely to represent an independent event for a single taxon rather than for multiple organisms simultaneously, making this scenario much less unlikely for an entire biota. They also reasoned that if long-distance, barrier-crossing dispersal were possible once, it should be possible multiple times, and repeated episodes of dispersal, colonization and extinction would invalidate the original assumption of allopatric speciation as well as complicate an area cladogram by not preserving the geographic history of speciation events.

1.5 Current trends in biogeography

Current methods in reconstructing historical biogeography follow Nelson & Platnick's lead with the cladistic biogeography approach, but tend to focus on the geographical history of a single taxon rather than the combined area histories of all the taxa in a biota. This is mainly due to a lack of phylogenetic data to study an entire biome, but that situation is changing with more and more studies emerging. Four major developments have helped to advance the discipline of biogeography in the last forty years: the validation of plate tectonic theory, the advent of molecular phylogenetics, the application of molecular clock theory and the use of fossils to incorporate time into phylogenies.

While the early trend in biogeography was in favour of long distance dispersal as championed by Darwin, following the popularization of plate tectonic theory in the 1970's, scientific opinion swung the other direction with vicariance gaining popularity and becoming the favoured explanation for intercontinental disjunctions. In this spirit of vicariance biogeography, pantropical taxa were frequently cited as being of Gondwanan origin, having attained their current distribution as a result of the splitting of the former supercontinent. Likewise, the disjunctions in many subtropical and temperate genera between North America and China were widely cited as being due to their former close proximity across the Bering and North Atlantic land bridges, which could have acted as a migration corridor. Although these theories were based on the congruence of distribution patterns and phylogenies, they could not be tested until the development of methods and software which incorporate time into phylogenies (i.e. r8s: Sanderson 2003, BEAST: Drummond *et al* 2006, and others). Dated phylogenies provide some evidence in support of Gondwanan vicariance, but this, as a causal factor, is not as ubiquitous as previously believed. Instead, there has been a counter-revolution in favour of dispersal as numerous studies demonstrate its importance in biome construction (Pennington & Dick 2004, Dick *et al* 2007, Renner 2004c). As such, molecular dating techniques and the use of fossil evidence are now seen as crucial in interpreting the history of a taxonomic lineage and the effects of dispersal and vicariance on current distribution patterns (Donoghue & Moore 2003).

Although mechanisms which generate and maintain biodiversity in species-rich biomes remain unclear, various scenarios have been postulated for the evolution of tropical forest floras on a regional scale. These vary from the 'museum model' (Stebbins 1974), which suggests that a stable tropical climate allowed species to accumulate over time, to the 'engine model,' (Haffer 1969) which suggests that climates may have been unstable due to glacial cycles resulting in cooling, drying or changes in sea-level during which rainforest species may have withdrawn to small refugial pockets. On a local scale, dated molecular phylogenies have provided evidence for speciation that may have been due to recent climatic changes such as aridification or geological phenomena such as the uplift of mountain ranges in the Neotropics and Africa (e.g. Richardson *et al* 2001, Hughes & Eastwood 2006, Simon *et al* 2009, Plana *et al* 2004, Couvreur *et al* 2008).

Comparative analyses of multiple dated, species level phylogenies of biome-endemic lineages are beginning to piece together regional ecosystem histories in relation to geology, climate and ecology (e.g. Worldwide: Linder 2008; Southern Hemisphere: Sanmartin & Ronquist 2004; Northern Hemisphere: Xiang *et al* 1998, 2000, Wen 1999, Donoghue *et al* 2001, Donoghue & Smith 2004; Andes: Sarkinen 2010; Cerrado: Simon *et al* 2009; Neotropical dry forests: Pennington & Dick 2004, Pennington *et al* 2006; South African Cape flora: Linder 2003, Linder & Hardy 2004, Galley & Linder 2006, Warren & Hawkins 2006, Verboom 2009; Australia: Crisp *et al* 2004; Malesia: Richardson *et al* 2010). These studies show that while some biomes are composed of ancient lineages which have evolved gradually, a great deal of contemporary floristic diversity is relatively recent and can be attributed to radiation since the mid-Miocene climatic optimum.

Such an analysis has not yet been attempted across the wet tropics, but the potential for using a pantropical genus such as *Manilkara* as a proxy to determine patterns in the history of the assembly of tropical forests in the Neotropics, Africa and Southeast Asia will be tested in this thesis.

1.6 Why *Manilkara*?

Pantropically distributed taxa, such as *Manilkara*, are excellent models for studying the evolution of tropical forests and variation in the rates of diversification between species on different continents. While many angiosperm genera have disjunct distributions, comparatively few of these are pantropical with taxa spanning all major tropical regions of the world, and even fewer of these pantropical disjunct genera have been revised taxonomically or have fully resolved species level molecular phylogenies upon which to test biogeographical scenarios. This is partially due to the fact that many of them are very species-rich, making it difficult to achieve complete taxon sampling, i.e. *Psychotria* (Rubiaceae) ~1850 spp., *Begonia* (Begoniaceae) ~1400 spp., *Ficus* (Moraceae) ~850 spp., *Phyllanthus* (Phyllanthaceae) ~800 spp., *Diospyros* (Ebenaceae) ~550 spp., *Capparis* (Capparidaceae) ~250 spp., *Terminalia* (Combretaceae) ~200 spp., *Xanthoxylum* (Rutaceae) ~200 spp., *Garcinia* (Guttiferae) ~200 spp., *Strychnos* (Loganiaceae) ~190 spp., *Homalium* (Salicaceae) ~180 spp., *Xylopia* (Annonaceae) ~160 spp. (Mabberley 2008).

Manilkara is a genus of trees in the Sapotaceae consisting of ~79 species distributed throughout the tropics (30 in South and Central America, 35 in Africa and 14 in Southeast Asia). With relatively fewer species than the large pantropical genera mentioned above, a complete taxon sample is more achievable. Additionally, its even spread of species diversity across all major tropical regions of the globe makes *Manilkara* an ideal candidate for in-depth biogeographic investigation. Moreover, the Neotropical and African species have been recently revised (Pennington 1990, Plana unpublished manuscript) and numerous family-level molecular studies have been carried out (Anderberg & Swenson 2003, Swenson & Anderberg 2005, Smedmark *et al* 2006, Smedmark & Anderberg 2007) providing a strong phylogenetic framework within which to assess biogeographic hypotheses.

Various questions arise from such a distribution. Where and when did *Manilkara* originate and what factors have contributed to the pantropical distribution we see today? Has vicariance or dispersal played a more prominent role in creating this intercontinental disjunction? What does the timing of diversification on different continents tell us about the historical assembly of tropical forests in each region? Molecular phylogenies using DNA sequence data, calibrated with a temporal dimension, offer a means of testing different hypotheses on biogeographic histories and causes of diversification.

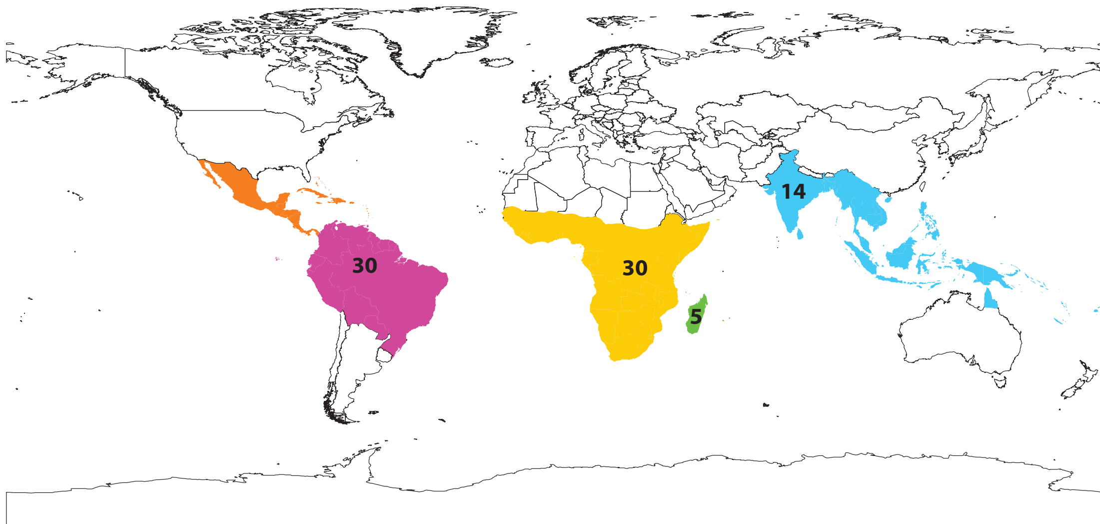


Figure 1.1 Map depicting number of *Manilkara* species in each main continental region: 30 in the Neotropics, 35 in Africa and Madagascar and 14 in Asia.

1.7 Aims of the thesis and overview of chapter content

The principal objective of this research is to investigate the factors, which have affected the historical assembly of tropical forests by studying the pattern and cadence of species diversification in *Manilkara* and then comparing this with the spatio-temporal patterns found in phylogenetic studies of other tropical angiosperms. As such, this study considers whether vicariance or dispersal has played a greater role in the assembly of tropical forest taxa and the timeframe in which these mechanisms have influenced cladogenesis with respect to historical geology and climate.

An understanding of taxonomic relationships at inter and intra-generic levels is fundamental to biogeographic investigations. Without this knowledge it is not possible to determine which biogeographic questions are appropriate to ask. Therefore, a primary aim of this thesis is to reconstruct the species level relationships in *Manilkara* as well as relationships between clades in the subtribe Manilkarinae and tribe Mimusopeae in order to test whether traditional morphology-based classifications are supported by molecular data and to determine whether the groups are monophyletic prior to biogeographic investigations.

To provide a background to the phylogenetic analyses presented in Chapter IV, the taxonomic placement of *Manilkara* within the family Sapotaceae, the tribe Mimusopeae and the subtribe Manilkarinae as well as morphological characters for generic and tribal delimitation are discussed in Chapter II. General molecular phylogenetic methods including DNA extraction, amplification and sequence generation are outlined in Chapter III. Relationships between the taxa are investigated in Chapter IV through the reconstruction of a species-level phylogeny of *Manilkara* and related genera using parsimony and Bayesian methods and both nuclear (ITS) and chloroplast (*rps16-trnK, rpl32-trnL, trnS-trnFM*) gene regions.

In order to interpret the geographic patterns reflected in taxon relationships, an understanding of the routes for migration between tropical regions is necessary, providing a spatial and temporal context for disjunctions and age ranges of clades. Chapter V presents a detailed overview of dispersal and vicariance scenarios for tropical intercontinental disjunctions and outlines the geological and climatological history of the Neotropics, Africa and Southeast Asia throughout the Tertiary in order to provide a framework for biogeographic hypothesis testing. Biogeographic patterns in other tropical angiosperm groups are also surveyed based on phylogenetic evidence. In Chapter VI Sapotaceae fossil history is reviewed in order to determine appropriate fossils for calibration of the phylogeny and to present further background information for hypothesis testing.

Chapter VII considers the biogeographic questions and hypotheses presented in Chapter V. The historical biogeography of *Manilkara* is investigated using different approaches to molecular phylogenetic dating and assessing the effect of various fossil calibration points. Additionally, the relative roles of vicariance and inter-continental dispersal are tested in relation to the ages of splits between lineages and the geological and climatological history on each continent. Ancestral areas are also reconstructed. Results are interpreted within a global biogeographic context and compared to dispersal-vicariance patterns found in other tropical taxa.

In the final Chapter, VIII, the main taxonomic and biogeographic findings of the research are discussed in relation to the original questions posed about the pattern and tempo of species diversification throughout the tropics and the historical assembly of tropical forests. Lastly, in order to synthesize what is currently known about the historical composition of this biome, a comparison is made between the ages of tropical forest genera with intercontinental disjunctions as evidenced by dated molecular phylogenies. Results are discussed in comparison with *Manilkara* and recommendations are made for future research.

Chapter II – Taxonomic history of *Manilkara*

2.1 Introduction to classification in the Sapotaceae

Sapotaceae is a family of predominantly tropical and subtropical trees and shrubs, which make up a significant component of lowland, wet forest in Africa, Asia and the Neotropics. Aside from its ecological prominence, the family is also economically important, being utilized for its timber, fruit and latex. Within the genus *Manilkara* alone, there are numerous economically important taxa. *M. bidentata*, *M. huberi*, *M. obovata* and *M. kauki* are just a few of the many species which are used commercially for timber and are known for their heavy, durable, rot-resistant wood. The latex of *Manilkara zapota* is tapped to form “chicle,” (the original chewing gum) and *M. bidentata* latex is the source of “balata,” formerly used in the manufacture of golf ball shells and machine belts. *M. zapota* fruit (“sapodilla”) is also widely cultivated, while other species, although not in cultivation, are all edible. Additionally, a related species within the tribe Mimusopeae, *Vitellaria paradoxa* (syn = *Butyrospermum parkii*), is known in the cosmetics industry as the source of shea butter (a fat pressed from the seed), and is an important component of many skin creams (Pennington 1991, Govaerts *et al* 2001, Mathews 2009).

According to recent phylogenetic analyses by the Angiosperm Phylogeny Group III (2009), the Sapotaceae is placed within the asterid order Ericales and is closely related to Ebenaceae, Styraceae and Sympolocaceae, although exact sister group relationships are still uncertain. Sapotaceae has always been regarded as a family in which generic delimitation is problematic, due to the large number of taxa with overlapping morphological variation. According to Pennington (1991), “characters unique to a genus are extremely rare in the Sapotaceae, so the use of single characters to define genera causes instability, depending which character is selected.” Because no single character is constant enough to define a group, stable groups must be delimited by suites of characters. Historically, this lack of discrete defining characters has led to confusion and a proliferation of differing classification systems based on broader versus narrower delimitation. This is evidenced by classification schemes, which range from 115 genera in 15 tribes and 4 subfamilies (Aubreville 1964) to 63 genera in six tribes and three subfamilies (Baehni 1965).

The most recent family-wide classification by Pennington (1991) includes 53 genera and more than 1,100 species. Pennington recognises five tribes in the Sapotaceae: Chrysophylleae, Isonandreae, Omphalocarpeae, Mimusopeae, and Sideroxyleae, with Mimusopeae subdivided into three subtribes: Mimusopinae, Manilkarinae and Gluemine. *Manilkara* is placed within Pennington’s subtribe Manilkarinae. Relationships within the genus *Manilkara*, its relationship to other genera within the subtribe Manilkarinae, and to members of Pennington’s tribe Mimusopeae will be investigated in Chapter IV. These groups are, therefore, briefly discussed further and the important characteristics of the subtribes are represented in the Table 2.1.

2.2 Classification within tribe Mimusopeae

The tribe Mimusopeae has the most complicated floral structure in the family Sapotaceae and is typically distinguished by the presence of tripartite corolla lobes (one median and two lateral lobes), petaloid staminodes, and a basal-basiventral seed scar. Pennington (1991) diagnoses the three subtribes of the Mimusopeae as follows:

subtribe 1. Mimusopinae: calyx of two whorls of four sepals, the outer whorl valvate; corolla lobes, stamens, staminodes, ovary loculi usually eight; staminodes usually hairy; fruit indehiscent; seed scar usually small and basal.

subtribe 2. Manilkarinae: calyx of two whorls of three sepals, the outer whorl valvate, corolla lobes, stamens usually six, less frequently 12-18; staminodes six, or absent, glabrous; fruit indehiscent; seed scar usually elongate, basi-ventral.

subtribe 3. Glueminae: calyx a single whorl of five imbricate or quincuncial sepals; corolla lobes, stamens, staminodes ovary loculi usually five; staminodes hairy or glabrous; fruit dehiscent or not, seed scar long, usually narrow, adaxial.

Table 2.1 Pennington's (1991) tribe Mimusopeae

Subtribe	Genera	Number of species
Mimusopinae	*Autranella*	1
	Baillonella	1
	Mimusops	47
	Tieghemella	2
	Vitellaria	1
	Vitellariopsis	5
Manilkarinae	*Faucherea*	11
	Labourdonnaisia	7
	Labramia	9
	Letestua	1
	Manilkara	79
	Northia	1
Glueminae	*Eberhardtia*	3
	Gluema	1
	Inhambanella	2
	Lecomtedoxa	5
	Neolemonniera	3

Table 2.2 Diagnostic characters of genera in Pennington's subtribe Manilkarinae

Genus	Calyx	Corolla	Stamens	Staminodes	Ovary	Seed scar	Endosperm
Manilkara **Adanson (1763)**	2 x 3 free	6(-9), usually divided to the base into three segments	6(-12)	(0-)6(-12) alternating with stamens	hairy or glabrous	basi-ventral usually narrowly elongate	copious
Labramia **A. de Candolle (1844)**	2 x 3 free	6(-8), divided to the base into three segments	6(-8)	6(-8), usually vestigial	glabrous	adaxial, broad, full length of the seed	copious
Faucherea **Lecomte (1920)**	2 x 3 free or slightly united	6(-11), entire/no appendages	6(-11)	6(-11), usually reduced	hairy	basi-ventral, broad, less than half as long as the seed	copious
Labourdonnaisia **Bojer (1837)**	2 x 3 free	(10-)12-18, entire or with few small lateral teeth	11-18 (-21)	absent or vestigial	hairy	basal or basiventral, broad, often strongly concave	copious
Northia **J.D. Hooker (1884)**	2 x 3 free	6, appendages vestigal or absent	6	absent or vestigial	hairy	adaxial, broad, covering about 1/3 of the seed surface	absent
Letestua **Lecomte (1920)**	2 x (2) 3 free	12-18, divided to the base into three segments	12-18	absent but occasionally some stamens lacking anthers	hairy	adaxial, narrowly elongate	copious

2.3 Classification within subtribe Manilkarinae

With its broad pantropical distribution, *Manilkara* is the largest and most widespread of the six genera in the subtribe Manilkarinae. It has traditionally been distinguished from other genera in the Sapotaceae by the following suite of characters: presence of two calyx whorls each consisting of three sepals, a corolla of six petals, each often divided into three segments, the median segment erect and clasping the opposite stamen, while lateral segments spread horizontally, staminodes variously shaped and alternating with stamens (Figs. 2.1 and 2.2, Appendix 2.1). The closely related genera *Labramia* (nine spp.), *Faucherea* (eleven spp.) and *Labourdonnaisia* (seven spp.) are all endemic to Madagascar and the Mascarenes. They only differ slightly from *Manilkara* in the development of the corolla lobe segments and staminodes, and in *Labourdonnaisia*, also in the higher number of corolla lobes and stamens. The monotypic genus *Letestua* is endemic to West Africa and is distinguished from all other genera in the Manilkarinae (except *Labourdonnaisia*) by its high number of corolla lobes, stamens and ovules as well as its long adaxial seed scar (in which it differs from *Labourdonnaisia*). According to Pennington (1991) "with up to three times the normal compliment of corolla lobes and each one divided to the base in three segments, *Letestua* presents the most complicated corolla structure of any Sapotaceae." Lastly, *Northia* is a monotypic genus, endemic to the Seychelles and differs from other members of the Manilkarinae in its lack of petal appendages and staminodes (although both are occasionally vestigial) and its seed characteristics including the shape of its seed scar and lack of endosperm. Table 2.2 presents a more detailed comparison of the genera.

Like other genera in the family, *Manilkara* too has had a convoluted taxonomic history as evidenced by the numerous segregate genera (*Achras*, *Chiclea*, *Manilkaropsis*, *Mopania*, *Nispero* and *Northiopsis*), which have now been sunk into synonymy by various authors (Govaerts *et al* 2001). This discrepancy in generic circumscription is due in large part to confusion over the varying degrees of development of the corolla lobe segments and staminodes, with some authors placing more weight than others on presence or absence of these characters. However, variation in corolla lobe dissection has apparently evolved independently in the Neotropics, Africa and Asia (Lam 1941). For example: *M. zapota* in Central America can range between undivided to fully divided corolla lobes (Fig. 2.2). The same trend is found in African *M. discolor* (Hemsley 1966), whereas Polynesian *M. hoshinoi* has an un-divided corolla. Therefore, placing all species with undivided corolla lobes in a separate genus (as did Baehni 1965) is likely to create an artificial group of unrelated species. Pennington (1991) adds that aside from the degree of development of the corolla lobe segments and staminodes, "*Manilkara* is a remarkably consistent pantropical genus" in terms of its morphology. To give some background to the taxonomic hypotheses presented in Chapter IV, a history of the classification of *Manilkara* is presented below.

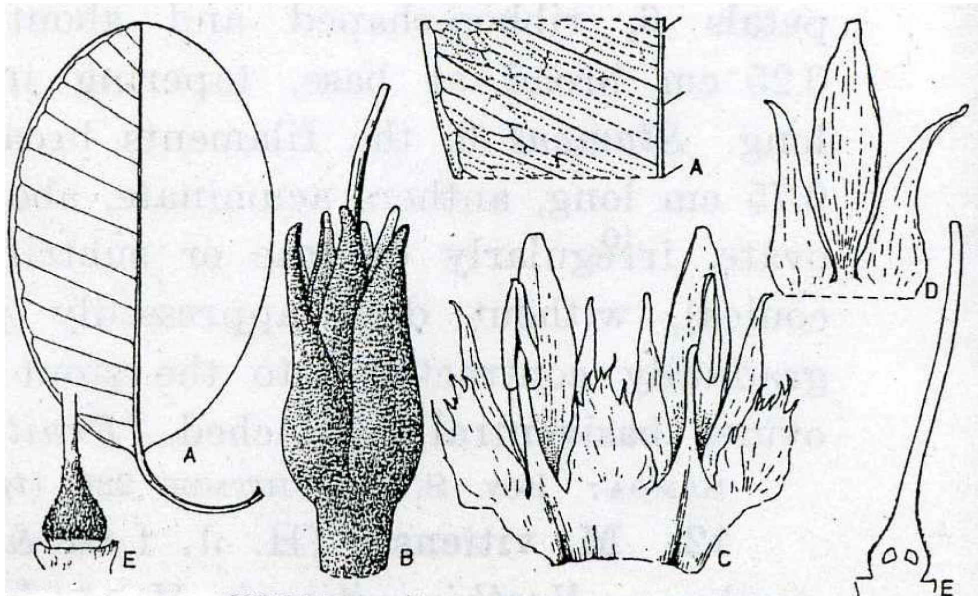


Figure 2.1 An illustration of the Fijian species *Manilkara smithiana*, detailing (a) leaf shape and venation, (b) a closed calyx with protruding stigma, (c) a dissected section of corolla with stamens opposite the petals and alternating with prominent lacinate staminodes, (d) a single petal with two appendages, and (e) the gynoecium. Excerpted from van Royen 1953.

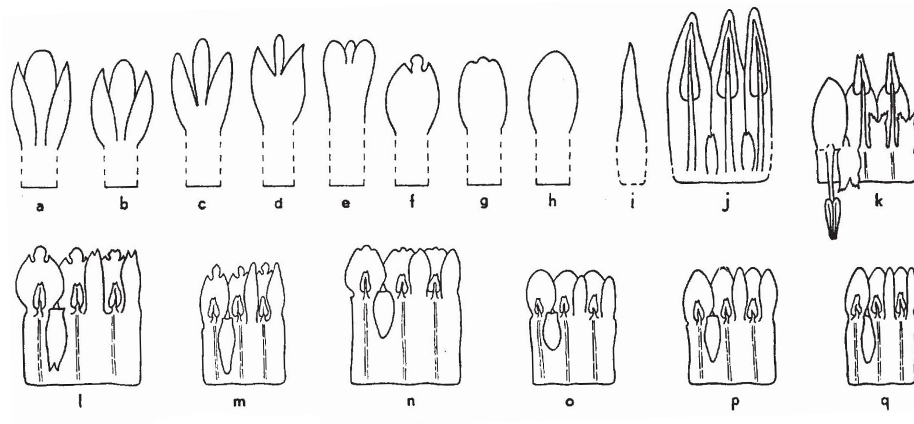


Figure 2.2 An illustration of petal dissection, or lack thereof, in *Manilkara*, excerpted from Lam 1941. Petal (a) exhibits the well-developed appendages typical of most *Manilkara* species, whereas few species have entire petals, as exhibited in (h) and (i). A range of intermediate variation also exists. (j) through (q) illustrate the floral architecture of *Manilkara*, with petals opposite the stamens and alternating with staminodes.

2.4 History of the names Sapotaceae and *Manilkara*

The family name Sapotaceae is derived from the genus *Sapota*, now a synonym of *Manilkara*. "Sapota" is the Spanish transliteration of the Nahuatl (Aztec) word "*tzapotl*" or "*tzapocuahuitl*" meaning "a soft edible fruit" in reference to the cultivated species *Manilkara zapota* and *Pouteria sapota* (de Riojas & de Poll 2007).

The name *Manyl-kara* was first published by van Rheede in Hortus Malabaricus (1683), a conspectus of the economic plants of Malabar (the present Indian state of Kerala) to describe the plant currently known as *Manilkara kauki* (Manilal 2003). The book refers to various

names for the plant, among them: *Manyl-kara* of the Malabaris, *Manil-gale* of the Brahmins, *Fruta Manilha* of the Portuguese and *Leo-bessen* or *Chineesche Pruymnen* of the Belgians. *Manilkara kauki* is not native to India and it is clear from the text that this is a cultivated plant brought to Malabar from China and the Philippines, used mainly for its edible fruits but also to treat ailments ranging from boils to "beriberi." Van Rheede indicates that "Manil" was adapted from the Portuguese "Manilhas Insulas" (for Manila, Philippine Islands) and "kara" refers to the Malayam for edible fruits.

In 1763, Adanson listed the name *Manilkara* in a table along with other proposed genera and their characters in his "Familles des Plantes." The name was later taken up by Dubard (1915), who formally described the genus and distinguished it from *Mimusops*. Then, in 1953, after considerable on-going debate over the classification of *Manilkara* versus *Achras*, van Royen chose to combine the two genera. Although *Achras* was the older formally published name, its floral morphology was not representative of the majority of species in the new genus, and Lam & van Royen (1953), therefore, proposed to conserve the name *Manilkara* Adans. (1763) against *Achras* L. (1753) – a decision which was accepted at the International Botanical Congress in Paris in 1954.

2.5 History of the classification of *Manilkara*

In his *Species Plantarum* (1753), Linnaeus created the genera *Mimusops* (mimo = ape, opsis = face, apparently named after the form of the corolla) for *Mimusops kauki* (van Rheede's *Manyl-kara*) and *Achras* (meaning "wild pear") for *Achras zapota* (the cultivated sapodilla or chicle tree).

Jussieu (1789) was the first to recognise these Linnean genera together as a distinct, homogeneous group, "Sapotae" or "les Sapotilles" with seven genera including both *Mimusops* and *Achras*. Then in 1844, DeCandolle subdivided this new family into six informal groups based on: 1) presence/absence of staminodes, 2) number of corolla lobes versus calyx lobes, 3) number and placement of stamens versus number of corolla lobes, 4) presence/absence of corolla lobe appendages. These characters are of high importance in contemporary classifications of the Sapotaceae (Pennington 1991). However, in DeCandolle's classification, there was a misunderstanding over the difference between corolla lobe appendages and staminodes. He mistakenly treated the poorly developed corolla lobe appendages in some genera as staminodes and the well-developed corolla lobe appendages in others as additional pairs of corolla lobes, resulting in a somewhat confused classification.

Species now recognised as *Manilkara* fall into two of DeCandolle's genera: *Sapota* and *Mimusops*. DeCandolle divided *Mimusops* into two sections: Quaternaria (flower parts in fours with eight fertile stamens) and Ternaria (flower parts in threes, with six fertile stamens). Quaternaria reflects the contemporary circumscription of *Mimusops* (aside from the inclusion of *M. kauki*) whereas Ternaria is predominantly comprised of species now placed in *Manilkara*. DeCandolle was the first to segregate species of *Mimusops* based on these important characters.

In their treatment of the Sapotaceae, Bentham and Hooker (1876) again broke the genera into six informal groups based predominantly on the number of floral parts in each whorl. Their recognition of the difference between a uniseriate and biseriate calyx was an important addition to the classification of sapotaceous genera. However, like DeCandolle before them, they also made mistakes interpreting some of the floral structures and treated the corolla lobe appendages of *Mimusops* as additional corolla lobes.

Hartog (1878) was the first to make formal divisions in the Sapotaceae based on the presence or absence of corolla lobe appendages and staminodes. However, because *Achras* lacks petal appendages but has staminodes, it was not placed with *Mimusops*, which does have petal appendages and staminodes.

Radlkofer (1888) based his classification of the genera on the presence or absence of staminodes and the number of fertile stamens, while tribal level divisions were based on the presence or absence of endosperm, simple or divided corolla lobes, presence of stipules, number of floral parts and single versus double calyx whorls. He was the first to place value on the presence of stipules as a generic level character. This classification resulted in some heterogeneous groups. Again *Achras* and *Mimusops* were placed in separate tribes.

Engler's two accounts of the Sapotaceae (1890-1891) both recognize two tribes distinguished by the presence or absence of corolla lobe appendages. Once again, *Mimusops* and *Achras* were placed in separate tribes based on the high weight given to this one character. Baillon's (1891) revision of the Sapotaceae improves upon previous classifications because it employs more characters for the higher level divisions and includes seed scar characters at the generic level for the first time. Yet *Mimusops* and *Sapota* (a synonym for *Achras*) are again placed in separate series.

Dubard (1912 & 1915) further clarified generic relationships within Sapotaceae, basing his higher level divisions primarily on androecium characters. While *Achras* was still kept separate from *Mimusops*, Dubard resurrected the name *Manilkara* and for the first time distinguished it as a genus separate from *Mimusops* recognizing the 3x2 (for *Manilkara*) versus 4x2 (for *Mimusops*) calyx whorls and the position of the seed scar as important characters for dividing the two genera.

Lam (1939) based his subdivisions of the Sapotaceae on Dubard's classification, but included calyx and corolla characters alongside Dubard's androecium characters at primary rank to distinguish his three subfamilies. While he recognized *Manilkara* and *Mimusops* in the same subfamily, he continued to keep *Achras* separate due to its lack of petal appendages. However, in his 1941 review of the Mimusopoideae (paying particular attention to the Asian species) Lam added the genus *Achras* to his tribe Manilkareae – a classification followed by contemporary authors. In doing so, he noted a very important point: "a most striking feature of the Manilkareae is the tendency of the reduction of both dorsal appendages and staminodes...both reductions are, generally speaking, independent of one another." He also indicated that it is clear that variations in the degree of division of the corolla lobes have occurred independently in America, Africa, Asia and the Far East. Lam's observation, that

the reduction series of these two characters were independent of one another, was a giant step forward and resulted in a more robust classification followed by subsequent authors. However, Lam concludes his 1941 revision of the Manilkareae with the sobering thought: "Probably only a botanist living some ten or hundred thousands of years later (if botanists are still roaming then) may be more lucky (or not) when put face to face with this group of plants."

Gilly (1943) discussed the generic limits of *Achras* and *Manilkara* stating: "Because of the trend of fusion of the exterior staminodes (petal appendages) and because vegetative, fruit, seed and perianth characters are fundamentally the same in both groups, I can see no real reason for maintaining the Sapodilla-Nispero (*Achras*) and the Balata (*Manilkara*) complex as separate genera." He, therefore, united the two genera under *Manilkara* (since there were many fewer species described in *Achras*, thus requiring fewer name changes). Van Royen (1953) reviewed Lam's 1941 classification in light of Gilly's conclusion and chose to formally sink *Achras* into *Manilkara* with a proposal to conserve the name.

Baehni's (1938) revision of the Sapotaceae differed from all previous classifications in that it placed primary importance on the position of the seed scar (basal or lateral). His weighting of seed scar characters above all else resulted in some anomalous groupings, including the odd placement of *Manilkara* and *Mimusops* in different subfamilies. In his subsequent (1965) inventory of genera he discusses characters both uniting and separating *Manilkara* and *Achras*, but although he appears in favour of uniting them (actually stating "*Manilkara*s with simple lobes are *Achras*."), continues to keep them separate on the basis of their seed scar characteristics. Baehni's, rejection of certain important characters (such as number of floral parts per whorl) and refusal to accept variability in generic characters resulted in a chaotic, artificial circumscription unlike that of any other author before or since.

Aubreville's (1964) assessment of the Sapotaceae followed Lam's (1939) classification including the tribes Mimusopinae and Manilkarinae. The result reflects the current classification of the family. Its only detraction is that many genera are narrowly defined and often based on single variable characters.

Pennington's (1991) comprehensive revision and synthesis of familial classification is the current standard reference for delimitation of genera within the family, alongside Govaerts *et al*'s (2001) checklist. His classification follows on from those of Lam and Aubreville and is based on a range of characters. In Pennington's system the subtribe Manilkarinae (including the genera *Faucherea*, *Labourdonnaisia*, *Labramia*, *Letestua*, *Manilkara* and *Northia*) is placed in the tribe Mimusopeae and is delimited by a calyx of two whorls of three sepals (the outer whorl valvate), corolla lobes and stamens usually six, less frequently 12-18, staminodes six or absent, glabrous; fruit indehiscent; seed scar usually elongate, basi-ventral.

2.6 Recent advances in Sapotaceae classification

Subsequent to Pennington's (1991) classification, Anderberg & Swenson (2003) carried out the first family-wide phylogenetic analysis based on sequences of the chloroplast region

*ndh*F (Fig. 2.3). In it, the monophyly of the tribe Mimusopeae was not supported, due in part to the low resolution provided by a single chloroplast marker of low variability. For the same reason, the monophyly of the subtribes Mimusopinae and Manilkarinae were not supported or refuted, but the subtribe Glueminae was clearly shown to be paraphyletic and for the most part, not closely related to the other two subtribes. In a further analysis of the same data with the addition of morphology, Swenson & Anderberg (2005) determined that the tribe Mimusopeae is best split into two groups, one with the subtribes Manilkarinae and Mimusopinae and the other a collection of genera with, as yet, uncertain affinity. They also reiterated that Pennington's Glueminae are not allied with Manilkarinae and Mimusopinae, which were this time both recovered as monophyletic, but with poor jacknife support. As with previous classifications, they found that no single unambiguous morphological character could diagnose the tribe. Swenson & Anderberg (2005) then suggested a new classification of the Sapotaceae proposing three subfamilies corresponding to the three main clades recovered in their phylogeny: Sarcospermatoideae, Sapotoideae and Chrysophylloideae. Their subfamily Sapotoideae includes their tribes Sapoteae and Sideroxyleae. Sapoteae is composed of Pennington's subtribes Mimusopinae and Manilkarinae (but not Glueminae) and one part of the tribe Isonandreae.

Following on from Swenson & Anderberg's (2005) familial reclassification, Smedmark *et al* (2006) investigated subfamilial relationships in the Sapotoideae with *ndh*F and the additional chloroplast regions: *trn*H-*psb*A, *trn*C-*trn*D, *trn*C-*psb*M, *psb*M-*trn*D (Fig. 2.4). In this analysis, Sapoteae *s.str.* (which includes the subtribes Mimusopinae and Manilkarinae) is strongly supported in the Bayesian analysis (posterior probability 1), but only poorly supported in the parsimony (bootstrap 50). There was weak support (pp 64, bs < 50) for the monophyly of subtribe Manilkarinae with the exclusion of the genus *Northia*, which had uncertain placement, sister to *Inhambanella*. *Manilkara* was also recovered as monophyletic (pp 1, bs 55) with the inclusion of the monotypic genus *Letestua*. Mimusopinae was resolved as paraphyletic, but *Mimusops* was shown to be monophyletic (pp 1, bs 100) with *Tieghemella* as its sister group and *Vitellariopsis* was shown to be monophyletic (pp 1, bs 97) with *Vitellaria* as its sister. As in previous analyses, Glueminae was paraphyletic and not a natural group, unrelated to the rest of the tribe Mimusopeae. Synapomorphies for Sapoteae *s.str*. were shown to be the presence of two calyx whorls, each with three (Manilkarinae) or four (Mimusopinae) sepals and with the central corolla lobe clasping the stamen. Other than the misplacement of the Glueminae and the genus *Northia*, Pennington's (1991) classification holds up remarkably well in light of the new molecular data.

2.7 Taxonomic aims of the thesis

The monophyly of the genus *Manilkara*, relationships within the subtribe Manilkarinae and the classification of Pennington's tribe Mimusopeae are investigated with further sampling and sequencing of the nuclear ribosomal ITS region and the chloroplast regions *rps16-trn*K, *rpl32-trn*L, *trnS-trn*FM region in Chapter IV.

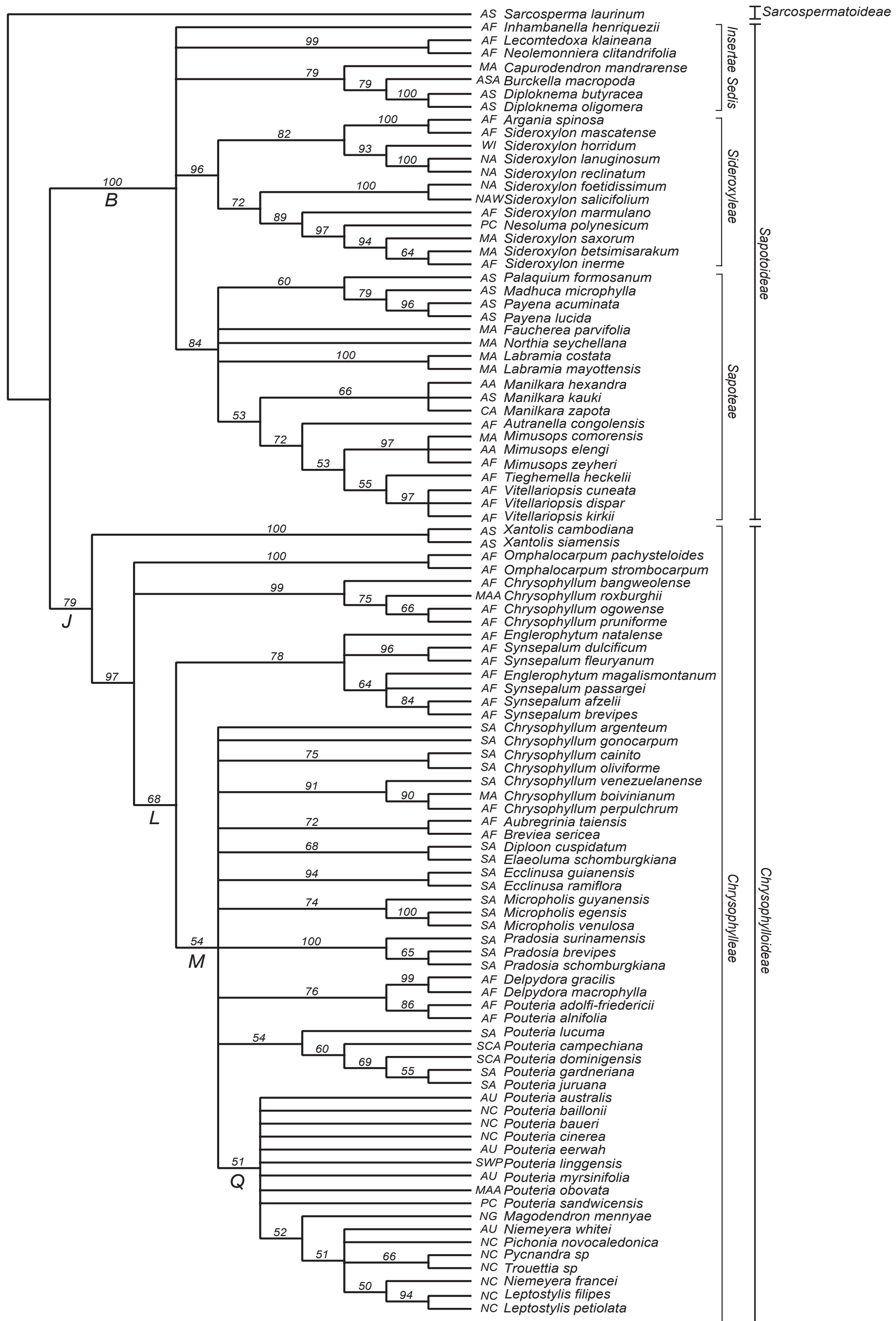


Figure 2.3 Sapotaceae family level phylogeny based on *ndhF* and morphology with a summary of proposed subfamilial classification presented in Swenson & Anderberg 2005. Subfamilies are listed in brackets to the right of the phylogeny. Jacknife support for groups is noted above branches. Distribution areas are listed to the left of taxon names. AF = Africa, AS = Asia, ASA = Asia to Australia, AU = Australia, CA = Central America (including Mexico), MA = Madagascar (including Mayotte Island, Seychelles), MAA = Madagascar to Australia, NA = North America, NAW = North America and West Indes, NC = New Caledonia, NG = New Guinea, PC = Pacific, SA = South America, SCA = South and Central America, SWP = South West Pacific, WI = West Indes.

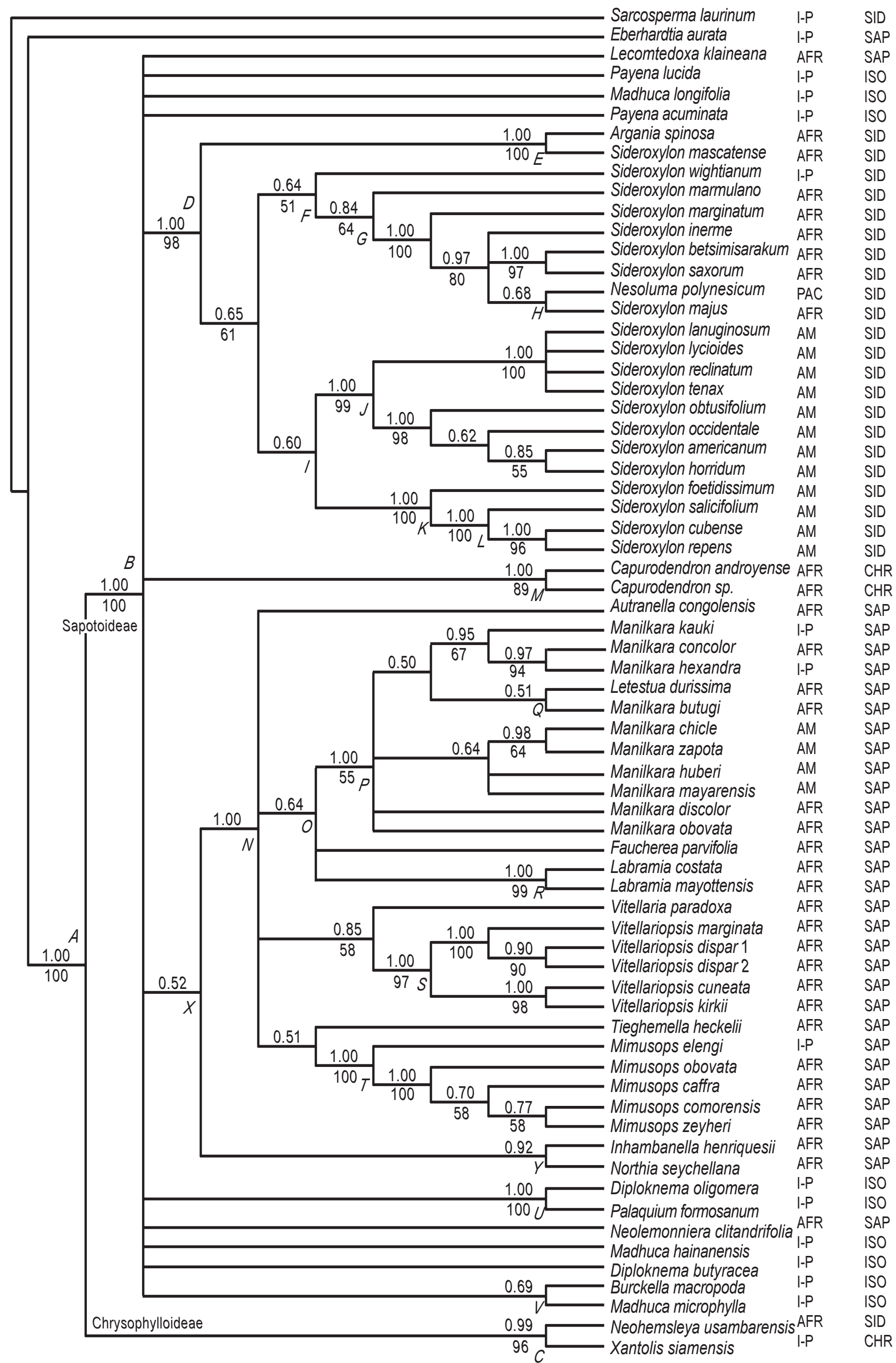
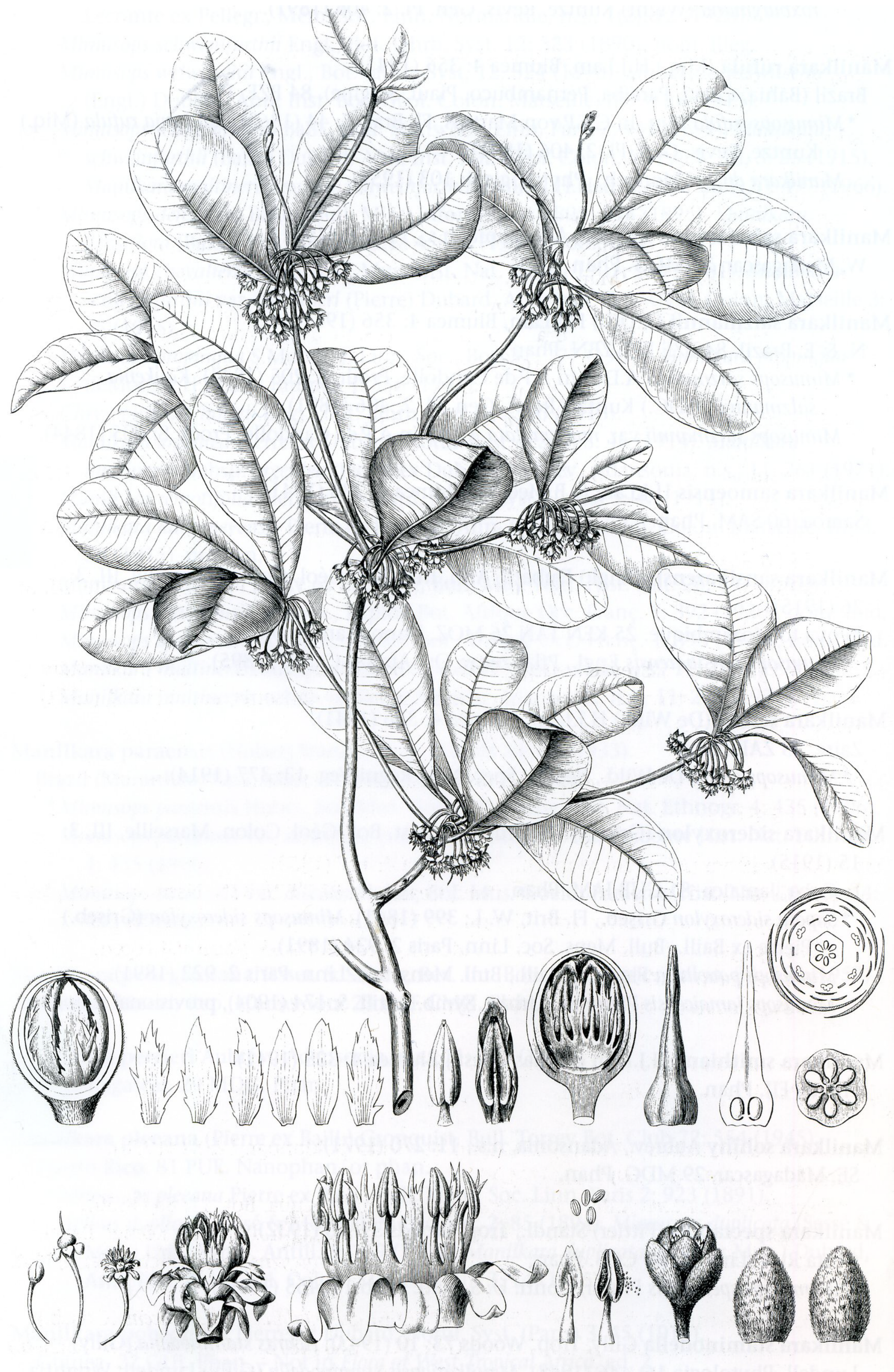


Figure 2.4 Sapotoideae tribal level phylogeny based on *ndhF, trnC-petN, petN-psbM, psbM-trnD*, and *trnH-psbA* from Smedmark *et al* 2006. Majority rule consensus tree with posterior probabilities of clades noted above branches.

2.8 Appendix



Appendix 2.1 An illustration of the Brazilian species *Manilkara subsericea*, showing the typical architecture (*Terminalia*-type branching) of *Manilkara*, and floral structure. Excerpted from Govaerts *et al* 1991, originally published in Flora Brasiliensis in 1863.

Chapter III – Molecular and phylogenetic materials and methods

3.1 Choice of plant material for DNA extraction

With approximately 79 species relatively evenly distributed across the tropics (30 South & Central American, 35 African and 14 Southeast Asian) *Manilkara* is an excellent model for studying historical biogeography. A well-sampled species-level phylogeny has the potential to answer questions about the origin of the genus and its diversification through time and space. The target of this study was, therefore, to sample as many *Manilkara* species as possible from across all three major tropical regions of the world (the Neotropics, Africa and Southeast Asia). Whereas the genus has a wide distribution, most individual species do not and acquiring fresh or silica gel-dried material of all 79 species proved to be extremely difficult. Consequently, the majority of DNA samples were obtained from herbarium specimens.

3.1.1 Acquisition of material and sequences

The following herbaria kindly granted permission to remove samples: E, K, FHO, P, L, WAG, G, BR, MO, NY, A and BISH. Additional material was provided by David Harris (E), Terry Pennington (K/FHO), Rachun Pooma (BKF), Laurent Gautier (G), Arne Anderberg (S) and Adi Suprapto (BO/Purwodadi) and also acquired during my own fieldwork on the Indonesian islands of West Papua and Sulawesi. In total, DNA was extracted from 104 samples, representing 69 out of the 79 (87%) species of *Manilkara* worldwide (Appendix 3.1). The remaining nine unsampled species are poorly known (often only from the type specimen) and were not possible to access (Appendix 3.2). Some *Manilkara* sequences were also acquired from previous M.Sc. research carried out at RBGE on the genus by Josh Clayton (2003), from research as part of the BRIDGE project in French Guiana by Jerome Chave and from Arne Anderberg and Jenny Smedmark at the Natural History Museum in Stockholm. Details of taxa sampled in each analysis are given in chapters IV and VII.

3.1.2 Selection of species for inclusion in analyses

Any molecular phylogenetic analysis is dependent upon a strong taxonomic background, especially concerning which species are real entities. Selection of species for the present study required consideration of differing opinions about whether certain taxa should be recognised. Eighty-two species of *Manilkara* are recognised in Govaerts *et al*'s (2001) Sapotaceae checklist. However, as the African taxa are still being revised, this list includes six taxa (*Manilkara adolf-friederici, Manilkara casteelsii, Manilkara longistyla, Manilkara microphylla, Manilkara seretii* & *Manilkara sylvestris*), which have been sunk into synonymy by Plana (unpublished) and are, therefore, not included in this study. Additionally, four species (*M. cuneifolia, M. lacera, M. multivervis, M. welwitschii*), which are sunk into synonymy in the Sapotaceae checklist, are considered for resurrection by Plana. These species are included in this study. Two newly described species (*M. lososiana* & *M. yangambensis*) are also included (Plana, unpublished). Twelve other species in Govaerts *et al* (2001) are poorly known and were not available for sampling (*Manilkara bolivarensis, M.*

dardanoi, M. doeringii, M. dukensis, M. fischeri, M. frondosa, M. ilidensis, M. kribensis, M. kurziana, M. nicholsonii, M. pobeguinii and *M. spectabilis*). See Appendix 3.2 for details.

3.1.3 Outgroups to *Manilkara*

Sister groups in the Tribe Mimusopeae (Pennington 1991) subfamily Sapotoideae (Swenson & Anderberg 2005) were also sampled including (sampled species/total number of species in genus according to Govaerts *et al* 2001) subtribe Manilkarinae: *Labramia* (4/9), *Faucherea* (4/11), *Labourdonnaisia* (3/7), *Letestua* (1/1), *Northia* (1/1) and subtribe Mimusopinae: *Mimusops* (7/47), *Autranella* (1/1), *Baillonella* (1/1), *Vitellaria* (1/1), *Vitellariopsis* (4/5), *Tieghemella* (1/2). Further outgroups were sampled in the tribes Isonandreae and Sideroxyleae of the subfamily Sapotoideae (Swenson & Anderberg 2005) in order to use fossils from these groups for molecular dating analyses and to test the monophyly of the tribe Mimusopeae. Genera in the subfamilies Chrysophylloideae and Sarcospermatoideae (Swenson & Anderberg 2005) were used to root the phylogenies. Outgroup sequences were predominantly acquired from colleagues at the Swedish Museum of Natural History and through M.Sc. research at RBGE on the Isonandreae by Azrul M. Bakar (2009). Sequences of *Mimusops* were also provided by Yamama Naciri in Geneva.

3.1.4 Material removal from herbarium specimens

Care was taken not to damage or disfigure specimens when removing material for DNA sampling. This was done by removing material from the fragment packet in the first instance or from under another leaf where possible. Recently collected and greener specimens were favoured over older specimens with brown leaves, the DNA of which is likely to be more degraded. The oldest DNA sample was taken from the holotype of *Manilkara samoensis*, which was collected in 1878, however the majority of sampled specimens were collected between the late 1960's and the present.

DNA naturally degrades over time in a specimen, but rapid desiccation using silica gel or air-drying significantly aids preservation. The Schweinfurth method of specimen collection (Schrenk 1888, Bridson & Foreman 2000), which is used widely in the Southeast Asian tropics, requires that the specimens be preserved in alcohol in the field before being dried upon return to a herbarium. This technique severely degrades DNA, making it extremely difficult to amplify large regions through PCR. Specimens collected with the Schweinfurth method are, therefore, not the best choice for use in phylogenetic studies. Nonetheless, in some instances very few accessions of particular Southeast Asian species (*M. hoshinoi, M. napali, M. kanosiensis,* and *M. celebica*) were available and despite the fact that they were collected in alcohol the specimens were sampled, though with limited PCR success.

3.1.5 Sampling strategy

Each available *Manilkara* species was sampled at least once. Some species with broad distributions or variable morphology were sampled multiple times in order to discern whether putative taxonomic relationships were valid. Examples of such species are *M. obovata*, *M. mabokeensis*, *M. kauki* and *M. bidentata*. Vanessa Plana advised on the sampling of African taxa and Terry Pennington advised on the Neotropical taxa as they have each revised the species on those continents respectively.

3.2 Choice of genomic regions

Previous studies in the Sapotaceae (Anderberg & Swenson 2003, Swenson & Anderberg 2005, Bartish *et al* 2005, Smedmark *et al* 2006, Smedmark & Anderberg 2007, Triono *et al* 2007, Swenson *et al* 2007 a & b, Swenson *et al* 2008 a & b) have made use of a variety of genes from both the nuclear (ITS, ETS, ChsI) and chloroplast regions (*ndhF, trnC–petN*, *petN–psbM*, *psbM–trnD*, *trnH*–psbA, *psB-psbH, rpl*20-*rps*12, *trnS-trnG, trnL-trnF, atpβ-rbcL*). Different gene regions are capable of resolving phylogenetic relationships at different levels in a phylogeny depending on their mutation rate (slow or fast), which is dependent on the degree of selection pressure (i.e. protein coding regions are under greater selection pressure than non-coding regions) (Soltis & Soltis 1998, Judd *et al* 2002, Small *et al* 2004). More rapidly evolving regions such as the nuclear ribosomal internal transcribed spacer (ITS) can be effective at resolving relationships between closely related species (Baldwin *et al* 1995, Hershkovitz & Lewis 1996, Feliner & Rossello 2007), whereas the more slowly evolving chloroplast regions can be useful for resolving deeper level nodes, including the backbone of a phylogeny (Soltis & Soltis 1998, Shaw *et al* 2005, Shaw *et al* 2007). Utilizing a selection of both nuclear and chloroplast regions to study the evolutionary history of a taxonomic group is also an important way to test for incongruence between datasets, to determine whether hybridization has occurred and potentially directionality of gene flow, and also to make certain that the reconstructed phylogeny is representative of the evolution of the organism rather than an individual gene (Soltis & Soltis 1998, Small *et al* 2004, Edwards 2009). As such phylogenies based on each gene region can be considered alternative phylogenetic hypotheses (Doyle 1992, Maddison 1997). Additionally, if gene trees are congruent, they can be combined to give stronger support for evolutionary relationships between species in an organismal tree.

3.2.1 Nuclear regions

Due to its ease of amplification with universal primers and its ability to resolve species level relationships, ITS (internal transcribed spacer region) has become the single most commonly used nuclear region in angiosperms (Baldwin *et al* 1995, Alvarez & Wendel 2003, Small *et al* 2004). The ITS region is a non-coding spacer, which is part of the nuclear ribosomal 18S-5.8S-26S cistron. The ITS 1 spacer is situated between the 18S and 5.8S and the ITS 2 spacer sits between the 5.8S and 26S (Baldwin *et al* 1995, Judd *et al* 2002, Poczai & Hyvonen 2009) (Fig. 3.1). Whilst non-coding ITS 1 & 2 can be highly variable, the 5.8S gene, believed to function in ribosome translocation and protein elongation (Elea & Nazar

1997), is more conserved. Hundreds to thousands of copies of ITS are present in the genome in multiple tandem arrays, which undergo homogenization through concerted evolution and may result in a single predominant sequence across all arrays (Small *et al* 2004). This high copy number makes ITS easier to amplify. In previous publications on the Sapotaceae (Smedmark & Anderberg 2007, Swenson *et al* 2008), ITS has been shown to have more informative characters than cpDNA markers of an equivalent length. Because of this and for the sake of compatibility with other Sapotaceae datasets, ITS was chosen as the primary nuclear region for this study. ITS primer sequences are given in Appendix 3.3.

The external transcribed spacer region (ETS), which is adjacent to 18S at the 5' end (Fig. 3.1), has also been used by other Sapotaceae researchers (U. Swenson pers. comm. 2008), but because it is linked to ITS, it would be less likely to give an independent estimate of the phylogeny (Baldwin & Markos 1998). Additionally, ETS amplification in Sapotaceae can be complicated due to the difficulty of finding suitable sites for the forward primer to bind (M. Myrenas, pers. comm. 2008). For these reasons, ETS was not used in this study.

Gene duplication events can result in incorrect organismal relationships if paralogous copies are compared. Care must, therefore, be taken to ensure that sequence data is derived from orthologous gene copies (Buckler *et al* 1997, Alvarez & Wendel 2003, Small *et al* 2004, Feliner & Rossello 2007, Poczai & Hyvonen 2009).

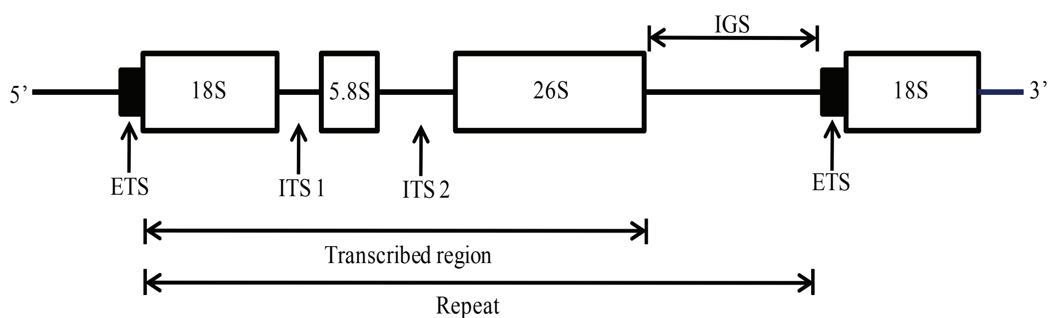


Figure 3.1. Diagram of the ITS and ETS regions adapted from Soltis & Soltis (1998). The non-coding internal transcribed spacer regions are denoted by ITS1 and ITS2. 18S and 26S are the genes which code for ribosomal RNA. The intergeneric spacer, which separates the repeating ITS region, is denoted by IGS, and the external transcribed spacer region (ETS) is located at the 5' end of the 18S region.

3.2.2. Chloroplast regions

Chloroplasts are abundant in plant cells, making their DNA easy to extract and amplify. Additionally, because they are usually uniparentally inherited (typically maternally in angiosperms) chloroplast genes are generally single copy, avoiding the potential problems of paralogy which are common in nuclear genes (Soltis & Soltis 1998, Small *et al* 1998, Small *et al* 2004). Yet, despite these positive characteristics, their conservative mutation rates mean that chloroplast regions are often not variable enough to resolve relationships between closely related species. Shaw *et al* (2007) found that the more variable noncoding chloroplast regions have rarely been employed in phylogenetic studies, whereas the most widely used regions are among the least variable. As a result, they reviewed a range of non-coding chloroplast regions, compared their mutation rates to discern potentially informative character (pic) values and suggested a list of the "fastest" chloroplast markers to test.

Additional considerations when choosing gene regions are ease of amplification and sequencing, and compatibility with other existing datasets. Research on the subfamily Sapotoideae (Smedmark *et al* 2006, Smedmark & Anderberg 2007) utilized five regions: *ndhF, trnC–petN*, *petN–psbM*, *psbM–trnD* and *trnH*–psbA. While they were informative at the subfamily level, they only showed moderate resolution between species in the genus *Sideroxylon*. In their study of chloroplast regions, Shaw *et al* (2007) did not review all of these markers, but the two they did study *psbM–trnD* and *trnH*–psbA were found to have low pic values, coming in 23rd and 24th respectively out of the 34 regions surveyed. The relative invariability of these regions suggests that they would not be the most useful markers for species level studies.

Jennifer Petersen, working on population level studies of *Chrysophyllum* (subfamily Chrysophylloideae), found the regions *trnS-trnFM*, *rps16*, *rpl16,* and *trnD-trnT* (Shaw *et al* 2007) to be useful (pers. comm. 2008). They were ranked by Shaw *et al* (2007) as 24th, 17th, 13th, and 10th respectively in terms of their pic values. According to recommendations in Shaw *et al* (2007), Smedmark *et al* (2006), Smedmark & Anderberg (2007) and by Peterson (pers. comm. 2008) the following twelve chloroplast regions were surveyed (bracketed number relates to pic value ranking out of 34 regions in Shaw *et al* 2007): *rpl32-trnL* (1st), *trnQ-rps16* (2nd), *trnV-ndhC* (3rd), *ndhF-rpl32* (4th), *psbD-trnT* (5th), *psbJ-petA* (6th), *rps16-trnK* (7th), *atpI-atpH* (8th), *trnD-trnT* (10th), *trnS-trnFM* (11th), *rps16* intron (17th), *trnH-psbA* (24th). Figure 3.2 illustrates the approximate locations and sizes of these regions in the chloroplast genome. Primers for these regions were tested on four species of *Manilkara* from the Neotropics, Africa and Asia in distinct lineages according to Clayton (2003). The regions *rpl32-trnL, rps16-trnK* and *trnS-trnFM* were found to be the easiest to amplify consistently and had high pic values as recorded by Shaw *et al* (2007). They were, therefore, utilized in this study. All tested chloroplast primer sequences are given in Appendix 3.4.

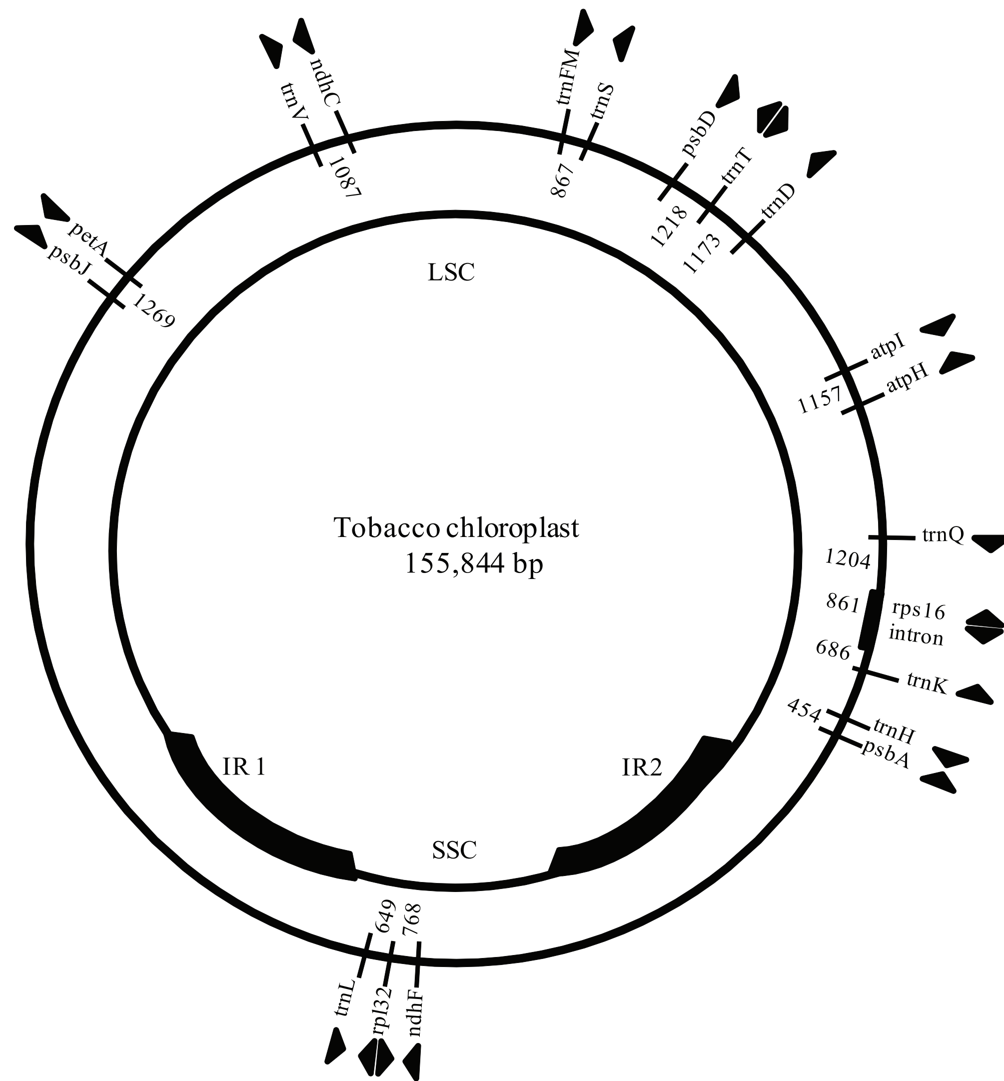


Figure 3.2. Diagram of the tobacco chloroplast genome with triangular arrows depicting approximate locations of regions surveyed in this study. Numbers between regions denote the number of base pairs in the relevant spacer region in the tobacco chloroplast genome. IR1 and 2: inverted repeats 1 and 2; LSC: large single copy region; SSC: small single copy region. Adapted from Shaw *et al* 2005 and Shaw *et al* 2007.

3.2.3 Microsatellite flanking regions

A new method for using microsatellite flanking regions in phylogeny reconstruction was proposed by Chatrou *et al* (2009). In a study on the genus *Annona*, they state that "flanking regions had a 3.5-10 fold higher substitution rate compared to two commonly used chloroplast markers (*rbcl* and *trnLF*), have no rate heterogeneity among nucleotide positions, evolve in a clock-like fashion and show no evidence of saturation," making them potentially useful for reconstructing species level relationships in angiosperm phylogenies. These are important points because saturation can be a problem in highly variable regions, where base changes have occurred on multiple occasions resulting in potentially high levels of homoplasy. Azevedo (2005) developed microsatellite markers for population level studies of *Manilkara huberi*. As a by-product of this study she designed primers for seven loci, which turned out to be monomorphic, i.e. with no allelic variation (MH05, MH09, MH10, MH14,

MH18, MH27, MH28), and which are potentially useful in phylogenetic studies. Primers developed specifically for *M. huberi* in this study (Appendix 3.5) were tested across a range of eight *Manilkara* species from each continent to determine their utility. MH09 in particular had some limited success amplifying across the sampled taxa (five out of eight species), whereas the other six primers were unsuccessful for species other than *M. huberi*. Due to the lack of amplification success, it was not effective to follow this line of investigation further.

3.3 Lab protocols

3.3.1 DNA extraction

Total DNA was extracted from herbarium specimens and silica gel-dried leaf samples. 20-30 mg of dried leaf material was placed in a 2ml Eppendorf tube with an angular tungsten grinding bead and submerged in liquid nitrogen. Samples were then agitated using a Mixer Mill at 20Hz for 1.5 minutes. This process of freezing and grinding was repeated up to six times until the leaf material was sufficiently pulverized.

CTAB (cetyltrimethylammonium bromide) (Doyle & Doyle 1990) and Qiagen Plant DNeasy Mini Kit DNA extraction methods were tested against one another with eight samples each (six herbarium specimens and two silica-gel). Neither method produced visible bands in the post-extraction gel for the herbarium specimen samples. However, silica gel-dried samples showed bands in the CTAB, but not the Qiagen extraction. It is common not to get bands in a gel before the first PCR when working with herbarium specimens and it was decided after discussion with Swedish colleagues at the Natural History Museum in Stockholm working on the Sapotaceae (J. Smedmark, A. Anderberg & U. Swenson) that the Qiagen method would produce a more pure extract, particularly from these degraded DNA samples. They determined in previous studies (Smedmark *et al* 2006, Smedmark & Anderberg 2007) that the Qiagen Plant DNeasy DNA extraction protocol was optimal and recommended its use. All total DNA extractions were then carried out using Qiagen's Plant DNeasy Mini Kit following the manufacturer's instructions.

3.3.2 PCR amplification

25 μl PCR reactions were set up using the recipes in Appendices 3.6 and 3.7. Betaine and BSA (Bovine Serum Albumin) were used as additives in most reactions to improve specificity and consistency respectively. Betaine reduces secondary structure and BSA is a stabilizing agent in enzymatic reactions. Amplification of DNA was carried out using the primers in Appendices 3.3, 3.4 and 3.5 and the PCR settings in Appendices 3.8, 3.9, 3.10 and 3.11 on a Tetrad2 BioRad DNA Engine. ITS PCRs using the primers ITS 1, 2g, 3p, 4, 5 and 8 followed the program "ITS PCR program A" in Appendix 3.8 and primers 18SF, 26SR, N18L, C26A, 5.8SN, 5.8SC followed the program "ITS PCR program B" in Appendix 3.9. All chloroplast regions were amplified using the *rpl16* program of Shaw *et al* (2005) outlined in Appendix 3.10, which is "slow and cold and has proven to be effective across a wide range of genomic and taxonomic regions." Monomorphic loci primers were amplified using the program in Appendix 3.11 and annealing temperature was varied for the different regions to determine the optimal PCR conditions (Appendix 3.12).

3.3.3 Nested PCR reactions for DNA extracted from herbarium material

Often extraction from herbarium specimens yields low quantities of degraded DNA and requires nested PCR in order to amplify DNA in quantities sufficient for sequencing. For ITS, nested PCR was accomplished by using ITS 5 and ITS 8, (the furthest external primers situated in the 18S and 26s regions respectively) in the first round of PCR. 1µl of the PCR product from this reaction was then used in the second (“nested”) PCR using the internal primers ITS 1 & ITS 4 and the same PCR program. In the majority of cases there were no bands visible in the gel following the first PCR, but after the second PCR, bands were apparent. Occasionally further internal primers, ITS 2g and ITS 3p, were used in place of ITS 1 and ITS 4 when amplification using the later primers was unsuccessful. ITS primers 18SF, 26SR, N18L, C26A, 5.8SN and 5.8SC were also tested, but were less effective than those outlined above and were, therefore, not used.

The nested PCR strategy was used in approximately 95% of the ITS reactions and 25% of chloroplast reactions. *Manilkara*-specific internal primers were designed (Appendix 3.4) for each of the chloroplast regions in the study (*rpl32-trnL, rps16-trnK* and *trnS-trnFM*) using Primer 3 (http://frodo.wi.mit.edu/primer3/input.htm) and Premier Biosoft’s Net Primer facility (http://www.premierbiosoft.com/company/news/NetPrimer.html).

This nested approach is highly effective at amplifying degraded DNA from herbarium specimens, but because it is very sensitive, it is also capable of amplifying small fragments of DNA in the lab other than the target group. To prevent contamination during PCR, filtered pipette tips were used and the barrels of communal pipettes were swabbed with alcohol as was the flow hood bench space.

3.3.4 Purification and sequencing

PCR purification was done initially with GFX microspin columns following the manufacturer’s instructions, but too much DNA was lost through this process, and the ExoSAP-IT (GE Healthcare) method was later adopted as it proved to be more efficient as judged by preparation time and sequencing quality. The ExoSAP-IT recipe and PCR protocol are outlined in Appendices 3.13 and 3.14.

Initially sequencing was done at the Royal Botanic Garden Edinburgh on a Beckman Coulter CEQ 8800 sequencer and later it was outsourced to the University of Edinburgh’s GenePool facility, which uses an ABI 3730 sequencer (see: http://genepool.bio.ed.ac.uk).

To prepare samples for sequencing on the Beckman Coulter CEQ 8800, sequencing reactions were carried out using the CEQ recipe (Appendix 3.15) and PCR program (Appendix 3.16) and then cleaned using the following protocol. Sequencing PCR reactions were diluted with 20µl dH_2O and transferred to tube with stop solution (see recipe in Appendix 3.17). 60µl of 100% ice cold ethanol was then added and mixed thoroughly by pipetting up and down. The mixture was then centrifuged at 13.000g for 15 minutes at 4°C and the supernatant was removed. 200µl of 70% ice cold ethanol was then added to the tube and centrifuged

at 13.000g for five minutes at 4°C. Again, 200µl of 70% ice cold ethanol was added to the tube and centrifuged at 13.000g for five minutes at 4°C. After removing the supernatant, the DNA pellet was vacuum dried for two to five minutes until no ethanol remained and then re-suspended in 35µl of Sample Loading Solution.

Samples sequenced through the GenePool service first underwent a sequencing PCR with BigDye following the PCR recipe and program in Appendices 3.18 and 3.19. Sequencing reactions were then carried out by the GenePool service.

3.4 Sequence editing and alignment

Forward and reverse sequences were assembled into contiguous sequences (contigs) and edited using the alignment software Sequencher ver. 4.7. A nucleotide BLAST (highly similar sequences megablast) search was carried out in GenBank (http://blast.ncbi.nlm.nih.gov/Blast.cgi) against all organisms in the NCBI database to determine whether contamination was an issue. Sequences found to be contaminated were excluded from analyses. Edited contigs were assembled and aligned by eye in MacClade ver. 4.08 (Maddison & Maddison 2008) or in BioEdit ver. 7.0.5 (Hall 2005). Variable nucleotide positions in the alignment were cross-checked against the original electropherograms in Sequencher to verify that base calls had been made correctly.

3.5 Regional coverage of species sequenced

Out of the 35 African *Manilkara* species, 26 were successfully sequenced for at least one region and 9 were unsampled, amounting to a 74% coverage. Of the 30 Neotropical species, 24 were successfully sequenced and six were unsampled, giving an 80% coverage, and out of the 14 Asian species, all were successfully sequenced for at least one region, giving 100% coverage. In total 64 out of the 79 *Manilkara* species worldwide were successfully sequenced for at least one region. This amounts to an 81% coverage of the genus.

3.6 Phylogenetic tools and methods

3.6.1 Phylogeny reconstruction methods

The goal of molecular phylogenetics is to convert DNA sequence data into a tree which represents the evolutionary relationships between taxa. There are many methods for doing this, which fall into two main categories. The first is a “distance” approach, in which sequence data is converted into a pair-wise distance matrix before being input into a tree building method which estimates the evolutionary distance between sequences (Page & Holmes 1998, Vandamme 2009, Holder & Lewis 2003). However, in the process of reducing sequence data to distances, the relationship between the character states and the tree is lost. Examples of programs which use the distance method along with a clustering algorithm are Neighbor Joining and UPGMA. Clustering algorithms are fast and often produce robust trees, but the outcome can depend upon the order in which the sequences were added to the analysis. Another significant limitation of clustering is that it does not allow for hypothesis

testing or a way of measuring fitness if two trees explain the data equally well (Page & Holmes 1998). The second approach is a “discrete” one, which considers each nucleotide site directly in the tree building process. Discrete methods retain information about which sites contribute to branch lengths enabling the reconstruction of ancestral character states and hypothesis testing. Maximum parsimony, maximum likelihood and Bayesian inference follow a discrete approach to phylogeny reconstruction and utilize an optimality criterion to choose amongst a set of trees by assigning each tree a score, which is a function of the relationship between the tree and the data. Optimality methods require a specific function that relates the data to the tree, for example a model of how sequences evolve. They also allow for the evaluation of tree fitness enabling the comparison of how competing evolutionary hypotheses fit the data (Page & Holmes 1998, Vandamme 2009).

3.6.2 Maximum parsimony

In maximum parsimony the optimality criterion follows the principle of parsimony, where the tree requiring the least number of evolutionary changes to explain the data is preferred. Trees generated through a parsimony approach are scored according to how many steps (evolutionary changes) are required to explain the distribution of each character. The main objection to parsimony is that it can be inconsistent under some conditions such as long branch attraction, where highly divergent taxa are falsely resolved as sister to one another as a result of convergent evolution of states, which becomes more likely with higher mutation rates (Page & Holmes 1998).

3.6.3 Maximum likelihood

Maximum likelihood finds the tree which is the most likely to have produced the observed data (i.e. the probability of the data given the hypothesis) (Felsenstein 1981). Likelihood methods differ from parsimony in that they can incorporate explicit models of sequence evolution and allow for evolutionary hypothesis testing (Holder & Lewis 2003, Page & Holmes 1998, Vandamme 2009). They also depend upon the quality of the specified model (such as parameters which account for the rate of evolution); an incorrect model can produce a biased result. Maximum likelihood is often considered to be a better method than parsimony because it is statistically consistent and enables modelling of evolutionary processes (Lewis 1998).

3.6.4 Bayesian inference

Bayesian inference uses the likelihood function and implements the same models of evolutionary change as maximum likelihood (in this instance, it calculates the probability of the hypothesis given the data). However, it differs in that it is based on Bayes’ theorem, which relates the posterior probability of a tree to the likelihood of the data and the prior probability of the tree with the specified evolutionary model (Huelsenbeck *et al* 2001). Through Metropolis Coupled Markov Chain Monte Carlo (MCMCMC) simulation a posterior probability distribution is generated as a result of how well the chosen model fits the data (Larget & Simon 1999, Holder & Lewis 2003, Ronquist *et al* 2009). The prior

distribution describes the probability of different trees given previous knowledge, whereas the posterior probability distribution describes the probability of trees considering the prior distribution, the model and the data (Archibald *et al* 2003). Unlike parsimony and likelihood methods, Bayesian inference does not produce a single (or set of) most likely trees. Rather, it produces a distribution of trees, sampled in proportion to their likelihood. Posterior probability values reflect the probability that a clade is "true" given the priors, model and data (Ronquist *et al* 2009). Commonly cited drawbacks to Bayesian inference are the need to set prior probabilities which can introduce bias and the evaluation of when sampling chains have converged (Archibald *et al* 2003).

3.6.5. Methods chosen for use in this study

After reviewing different phylogenetic methods it was decided to use a parsimony and a likelihood (Bayesian) approach and juxtapose these independent estimates to see if they agree on topology and branch support. Maximum parsimony analyses were carried out using PAUP* (Swofford 2003), Bayesian analyses were carried out using MrBayes 3.1 (Huelsenbeck & Ronquist 2001) and BEAST v.1.5.3 (Drummond & Rambaut 2007) was used to incorporate time into the analyses using a relaxed clock method in conjunction with fossil calibration points. BEAST was also used to reconstruct ancestral areas. Further details of specific methods used in these programs including settings and parameters are discussed in Chapters IV and VII.

3.7 Appendices

Appendix 3.1 List of specimens from which DNA was extracted as part of this project. Success with amplification and sequencing is noted in the final column. Not all sequenced accessions were used in the final analyses. Specimens included in analyses are noted in the relevant chapters.

Species	Edinburgh DNA accession number	Collector and number	Country of origin	Success with sequencing
Faucherea sp.	EDNA07-01933	A. Anderberg, A233	Madagascar	successfully sequenced ITS/cpDNA
Labourdonnaisia calophylloides	EDNA08-02271	R. Capuron, 28171SF	Reunion	successfully sequenced ITS/cpDNA
Labourdonnaisia madagascariensis	EDNA07-02212	R. Capuron, 27747SF	Madagascar	successfully sequenced ITS/cpDNA
Labourdonnaisia revoluta	EDNA07-02271	Lorence, DL1602	Mauritius	successfully sequenced ITS/cpDNA
Labramia costata	EDNA07-02272	Schatz & Gentry, 2094	Madagascar	successfully sequenced ITS/cpDNA
Labramia louvelii	EDNA07-01927	A. Anderberg, A245	Madagascar	successfully sequenced ITS/cpDNA
Manilkara bella	EDNA08-02267	Folli, 501	Brazil	successfully sequenced ITS/cpDNA
Manilkara bequaertii	EDNA07-02081	F. Breteler, 15348	Gabon	successfully sequenced ITS/cpDNA
Manilkara bidentata	EDNA06-05906	T. Pennington, 1203	Peru	successfully sequenced ITS/cpDNA
Manilkara boivinii	EDNA06-05905	L. Gautier, 3278	Madagascar	successfully sequenced ITS/cpDNA
Manilkara capuronii	EDNA07-02079	R. Capuron, 11.377SF	Madagascar	successfully sequenced ITS/cpDNA
Manilkara cavalcantei	EDNA07-02205	Vicentini *et al*, 527	Brazil	successfully sequenced ITS/cpDNA
Manilkara celebica	EDNA08-02339	Neth. Ind. For. Service, bb. 30152	Indonesia	successfully sequenced cpDNA only
Manilkara celebica	EDNA07-02055	M. van Gabel, 30 (bb.32.365)	Indonesia	poor quality sequence
Manilkara chicle	EDNA07-02200	Neill & Vincelli, 3309	Nicaragua	poor quality sequence
Manilkara concolor	EDNA07-1079 a,b,c	Nuvungu & Boane, 307	Mozambique	poor quality sequence
Manilkara cuneifolia	EDNA07-02264	G. McPherson, 16792	Gabon	successfully sequenced ITS/cpDNA
Manilkara dawei	EDNA07-01928	D.J. Harris, 7707	Central African Republic	successfully sequenced ITS/cpDNA
Manilkara decrescens	EDNA08-02268	Carvahlo & Lewis, 1113	Brazil	successfully sequenced ITS/cpDNA
Manilkara discolor	EDNA07-02268	J.D. & E.G. Chapman, 6689	Malawi	poor quality sequence
Manilkara dissecta	EDNA07-02060	W.A. Whistler, W3889	Samoa	poor quality sequence
Manilkara dissecta	EDNA07-02061	J.P. Wilson & Kajewski, 983	Vanuatu	poor quality sequence
Manilkara elata	EDNA08-02265	Jardin *et al*, 2277	Brazil	successfully sequenced ITS/cpDNA
Manilkara excels	EDNA07-02258 a,b	Duke, 13559	Brazil	poor quality sequence
Manilkara excise	EDNA08-02266	Harris, 8961	Jamaica	poor quality sequence
Manilkara fasciculata	EDNA08-02258	K. Armstrong, 353	Indonesia	successfully sequenced ITS/cpDNA
Manilkara fasciculata	EDNA08-02259	K. Armstrong, 354	Indonesia	successfully sequenced ITS only
Manilkara fasciculata	EDNA07-02089	Hildebrand, 225, bb.33.925	Indonesia	poor quality sequence
Manilkara fouilloyana	EDNA07-01082	Bouquet, 2431	Congo	poor quality sequence

Species	Edinburgh DNA accession number	Collector and number	Country of origin	Success with sequencing
Manilkara fouilloyana	EDNA07-02267	G. McPherson, 16173	Gabon	successfully sequenced ITS/cpDNA
Manilkara gonavensis	EDNA08-02264	Ekman, 8741	Haiti	successfully sequenced ITS/cpDNA
Manilkara hexandra	EDNA07-02051	R.B. Mazumdar, 10043	India	poor quality sequence
Manilkara hexandra	EDNA07-02053	P.L. Comanor, 868	Sri Lanka	successfully sequenced ITS/cpDNA
Manilkara hexandra	EDNA07-02090	van Beusekom & van Beusekom, 1954	Sri Lanka	successfully sequenced ITS only
Manilkara hoshinoi	EDNA07-02054	F.H. Damon, 217	Papua New Guinea	successfully sequenced ITS only
Manilkara hoshinoi	EDNA08-02340	M. Hoshino, 2138	Palau	successfully sequenced ITS/cpDNA
Manilkara huberi	EDNA07-01926	O. Poncy, OP1828	French Guiana	successfully sequenced ITS/cpDNA
Manilkara inundata	EDNA07-02093	Sothers & Saraiva, 22	Brazil	successfully sequenced ITS/cpDNA
Manilkara jamiqui	EDNA07-02201	Urquiola & Dressler, 529	Cuba	successfully sequenced ITS/cpDNA
Manilkara kanosiensis	EDNA07-02198	C.E. Carr, 11237	Papua New Guinea	successfully sequenced cpDNA only
Manilkara kanosiensis	EDNA08-02342	Neth. Ind. For. Service, bb. 24311	Papua New Guinea	poor quality sequence
Manilkara kauki	EDNA08-02260	K. Armstrong, 379	Indonesia	successfully sequenced ITS/cpDNA
Manilkara kauki	EDNA07-02062	P.J. Wester, 26	Philippines	poor quality sequence
Manilkara kauki	EDNA07-02063	Vetaga, 34	Fiji	poor quality sequence
Manilkara kauki	EDNA07-02064	S. Dewol & C. Phillipps, SAN 89959	Malaysia	poor quality sequence
Manilkara kauki	EDNA07-02065	A. Hoogerwerf, 134	Indonesia	poor quality sequence
Manilkara kauki	EDNA07-02066	P.I. Forster, PIF8924	Australia	poor quality sequence
Manilkara koechlinii	EDNA07-02270	Bouquet, 378	Congo	poor quality sequence
Manilkara lacera	EDNA07-01095 a,b	D.J. Harris, 8200A	Gabon	successfully sequenced ITS/cpDNA
Manilkara letestui	EDNA07-02080	J.J. Bos, 5604	Cameroon	successfully sequenced cpDNA only
Manilkara letouzeyi	EDNA08-02338	R. Letouzey, 4444	Cameroon	successfully sequenced ITS/cpDNA
Manilkara letouzeyi	EDNA07-01078 a,b	Dechamps, 177	Cameroon	poor quality sequence
Manilkara littoralis	EDNA07-02052	Maung Gale, 14654	Burma	successfully sequenced ITS/cpDNA
Manilkara longifolia	EDNA07-02092	Thomas *et al*, 8076	Brazil	successfully sequenced ITS/cpDNA
Manilkara lososiana	EDNA07-02088	D. Kenfack, 625	Cameroon	successfully sequenced ITS/cpDNA
Manilkara mabokeensis	EDNA07-01094 a,b	D.J. Harris, 7164	Central African Republic	successfully sequenced ITS/cpDNA
Manilkara mabokeensis	EDNA07-01929	D.J. Harris, 4324	Central African Republic	successfully sequenced ITS only
Manilkara maxima	EDNA07-02091	Sant'Ana *et al*, 670	Brazil	successfully sequenced ITS/cpDNA
Manilkara mayarensis	EDNA07-02210	Ekman, 15053	Cuba	successfully sequenced ITS/cpDNA

Species	Edinburgh DNA accession number	Collector and number	Country of origin	Success with sequencing
Manilkara mochisia	EDNA07-02269	Bidgood *et al*, 2286	Tanzania	successfully sequenced cpDNA only
Manilkara multifida	EDNA08-02269	Carvahlo & Guedes, 1903	Brazil	poor quality sequence
Manilkara napali	EDNA08-02341	F. Schram, B.W. 1636	Indonesia	successfully sequenced cpDNA only
Manilkara napali	EDNA07-02197	F. Schram, B.W. 2841	Indonesia	poor quality sequence
Manilkara obovata (butugi-type)	EDNA07-02262	Friis & Vollesen, 740	Sudan	successfully sequenced ITS/cpDNA
Manilkara obovata (multinervis-type)	EDNA07-02263	Schmidt *et al*, 3274	Ghana	successfully sequenced ITS/cpDNA
Manilkara obovata (obovata-type)	EDNA07-01930	D.J. Harris, 7759	Central African Republic	successfully sequenced ITS/cpDNA
Manilkara obovata (obovata-type)	EDNA08-02261	GAF Malanda, 7	Democratic Republic of Congo	successfully sequenced ITS/cpDNA
Manilkara obovata (obovata-type)	EDNA07-02261	Fanshawe, 2991	Rhodesia	poor quality sequence
Manilkara paraensis	EDNA07-02206	Zarucchi *et al*, 2526	Brazil	successfully sequenced ITS/cpDNA
Manilkara pelligriniana	EDNA07-02087	D.J. Harris & M. Fay, 1843	Cameroon	successfully sequenced ITS/cpDNA
Manilkara perrieri	EDNA07-02082	R. Capuron, 28132-SF	Madagascar	successfully sequenced ITS/cpDNA
Manilkara perrieri	EDNA07-02259	Randriamampionona, 248	Madagascar	poor quality sequence
Manilkara pleena	EDNA07-02209	A. Lioger & P. Lioger, 33453	Puerto Rico	successfully sequenced ITS/cpDNA
Manilkara pubicarpa	EDNA08-02262	Forest Dept. British Guyana, 5860	Guyana	successfully sequenced cpDNA only
Manilkara roxburghiana	EDNA07-02199 a,b	Matthew & Rajendren, 44790	India	successfully sequenced cpDNA only
Manilkara rufula	EDNA07-02208	G. Ignacio & A. Caurenio, 37	Brazil	successfully sequenced ITS/cpDNA
Manilkara sahafarensis	EDNA07-02085	R. Capuron, 20.965-SF	Madagascar	successfully sequenced ITS/cpDNA
Manilkara salzmanii	EDNA07-02207	Jardim *et al*, 2277	Brazil	successfully sequenced ITS/cpDNA
Manilkara samoensis	EDNA08-02263	S.J. Whilmee, 226	Samoa	successfully sequenced cpDNA only
Manilkara sansibarensis	EDNA07-01083 a,b,c	Abeid, 742	Tanzania	successfully sequenced ITS/cpDNA
Manilkara sideroxylon	EDNA07-02203	Ekman, 16173	Cuba	successfully sequenced ITS/cpDNA
Manilkara smithiana	EDNA07-02057	A.C. Smith, 1450	Fiji	successfully sequenced ITS/cpDNA
Manilkara sohihy	EDNA07-01081 a,b,c	Centre Technique Forestier, 26.591-SF	Madagascar	poor quality sequence
*Manilkara sp.*1	EDNA07-02260	P. Sita, 4107	Congo	successfully sequenced ITS only
Manilkara sp. 2	EDNA08-02256	Purwodadi Botanic Garden, P199701118/SO118	Indonesia	successfully sequenced ITS only
*Manilkara sp.*3	EDNA08-02257	K. Armstrong, 378	Indonesia, cultivated	successfully sequenced ITS only
Manilkara staminodella	EDNA07-02204	A. Anderberg *et al*, 50	Costa Rica	successfully sequenced ITS/cpDNA

Species	Edinburgh DNA accession number	Collector and number	Country of origin	Success with sequencing
Manilkara suarezensis	EDNA07-02259	Randriamampionona, 248	Madagascar	successfully sequenced ITS/cpDNA
Manilkara subsericea	EDNA07-02202	Hatschbach & Souza, 51302	Brazil	successfully sequenced ITS/cpDNA
Manilkara sulcata	EDNA07-02086	Frontier-Tanzania Coastal Forest Research Programme, 1045	Tanzania	successfully sequenced cpDNA only
Manilkara tampoloensis	EDNA07-01080 a,b,c	Services des Eaux et Forets de Madagascar, 5617	Madagascar	poor quality sequence
Manilkara triflora	EDNA08-02343	Fonseca *et al*, 2887	Brazil	successfully sequenced ITS/cpDNA
Manilkara triflora	EDNA07-02094	Daly *et al*, D788	Brazil	poor quality sequence
Manilkara udoido	EDNA07-02058	S. Slappy, LR26622	Palau	successfully sequenced ITS/cpDNA
Manilkara udoido	EDNA07-02059	Herbst, Stemmerman & Canfield, 9436	Palau	successfully sequenced ITS only
Manilkara valenzuelana	EDNA07-02211	A. Lioger & P. Lioger, 22980	Dominican Republic	successfully sequenced ITS/cpDNA
Manilkara vitiensis	EDNA08-02345	Smith, 1461	Fiji	successfully sequenced ITS/cpDNA
Manilkara vitiensis	EDNA07-02056	O. Degener & I. Degener, 32,216	Fiji	poor quality sequence
Manilkara welwitschii	EDNA08-02346	van den Houten *et al*, 25	Gabon	successfully sequenced ITS/cpDNA
Manilkara welwitschii	EDNA07-02265	F.J. Breteler & J.J.F.E. de Wilde, 655	Gabon	poor quality sequence
Manilkara yangambensis	EDNA07-02083	A. Leonard, 5824	Belgian Congo	poor quality sequence
Manilkara yangambensis	EDNA08-02344	C. Evrard, 1499	Belgian Congo	successfully sequenced cpDNA only
Manilkara zenkeri	EDNA07-02266	Letouzey, 13677	Cameroon	poor quality sequence
Manilkara zenkeri	EDNA07-02084	Doumenge, 526	Cameroon	successfully sequenced ITS/cpDNA

Appendix 3.2 Species listed in the Sapotaceae checklist (Govaerts *et al* 2001), which were not sampled as part of this study. In some cases this is due to the fact that a sample was unavailable, mainly because the species is only known from the type or the specimen has been lost. In other cases, this is because the species is sunk into synonymy with another species by Plana (unpublished) in her revision of the African species of *Manilkara*. These explanations are given for each species in the columns below.

Species	Distribution	Reason for exclusion
Manilkara adolf-friederici	Democratic Republic of Congo	= *M. welwitschii*
Manilkara bolivarensis	Venezuela	sample unavailable
Manilkara casteelsii	Democratic Republic of Congo	= *M. obovata*
Manilkara dardanoi	Brazil	sample unavailable
Manilkara doeringii	Togo	type lost - unknown
Manilkara dukensis	Cameroon	type lost - unknown
Manilkara fischeri	Tanzania	type lost - unknown
Manilkara frondosa	Angola	poorly known - related to *M. dawei*
Manilkara ilidensis	Cameroon	type lost, probably = *M. letestui*
Manilkara kribensis	Cameroon	type lost - unknown
Manilkara kurziana	Myanmar	no specimens found - unknown
Manilkara longistyla	Democratic Republic of Congo	= *M. dawei*
Manilkara microphylla	Republic of Congo-Gabon	= *M. welwitschii*
Manilkara nicholsonii	South Africa	new species only known from type
Manilkara pobeguinii	Guinea	type lost - unknown
Manilkara seretii	Democratic Republic of Congo	= *M. obovata*
Manilkara spectabilis	Costa Rica	sample unavailable
Manilkara sylvestris	Ivory Coast	= *M. obovata*

Appendix 3.3 ITS primers

Primer	Direction	Primer sequence (5'-3')	Author
ITS1	forward	GTAGGTGAACCTGCAGAAGGA	modified White *et al* 1990
ITS2g	reverse	GTGACACCCAGGCAGACGT	modified Moeller & Cronk 1997
ITS3p	forward	GCATCGATGAAGAACGTAGC	Moeller & Cronk 1997
ITS4	reverse	TCCTCCGCTTATTGATATGC	White *et al* 1990
ITS5p	forward	GGAAGGAGAAGTCGTAACAAG	Moeller & Cronk 1997
ITS8p	reverse	CACGCTTCTCCAGACTACA	Moeller & Cronk 1997
18SF	forward	GAACCTTATCGTTTAGAGGAAGG	Rydin *et al* 2004
26SR	reverse	CCGCCAGATTTTCACGCTGGGC	Rydin *et al* 2004
N18L	forward	AAGTCGTAACAAGGTTTCCGTAGGTG	Youngbae Suh - unpublished
C26A	reverse	TTTCTTTTCCTCCGCT	Youngbae Suh – unpublished
5.8SN	forward	ATCGAGTCTTTGAACGCA	Youngbae Suh - unpublished
5.8SC	reverse	TGCGTTCAAAGACTCGAT	Youngbae Suh - unpublished

Appendix 3.4 Chloroplast Primers

Primer name	Direction	Primer sequence (5'-3')	Author
atpI-atpH	forward	TATTTACAAGYGGTATTCAAGCT	Shaw *et al* 2007
atpI-atpH	reverse	CCAAYCCAGCAGCAATAAC	Shaw *et al* 2007
ndhF-rpl32	forward	CCAATATCCCTTYYTTTTCCAA	Shaw *et al* 2007
ndhF-rpl32	reverse	GAAAGGTATKATCCAYGMATATT	Shaw *et al* 2007
psbD-trnT	forward	CTCCGTARCCAGTCATCCATA	Shaw *et al* 2007
psbD-trnT	reverse	CCCTTTTAACTCAGTGGTAG	Shaw *et al* 2007
psbJ-petA	forward	ATAGGTACTGTARCYGGTATT	Shaw *et al* 2007
psbJ-petA	reverse	AACARTTYGARAAGGTTCAAT T	Shaw *et al* 2007
rpl32-trnL	forward	CTGCTTCCTAAGAGCAGCGT	Shaw *et al* 2007
rpl32-trnL	reverse	CAGTTCCAAAAAAACGTACTTC	Shaw *et al* 2007
rpl32-trnL-intF	forward	TCGTCGAGATTGAAGAGTCA	self-designed
rpl32-trnL-intR	reverse	TCTCTTTTGACCGGAAATTCA	self-designed
*rpl32_trnL*_int_2_F	forward	GGCGGCTGCTCAACTTAT	self-designed
*rpl32_trnL*_int_2_R	reverse	TCTCTTTTGACCGGAAATTCA	self-designed
rps16	forward	AAACGATGTGGTARAAAGCAAC	Shaw *et al* 2005
rps16	reverse	AACATCWATTGCAASGATTCGATA	Shaw *et al* 2005
rps16–trnK	forward	AAAGTGGGTTTTTATGATCC	Shaw *et al* 2007
rps16–trnK	reverse	TTAAAAGCCGAGTACTCTACC	Shaw *et al* 2007
rps16-trnK-intF	forward	TGTTCCTGCTATTCTATATTTCCTTG	self-designed
rps16-trnK-intR	reverse	GATGTGTAGATACAATCAGAATCAAAA	self-designed
*rps16_trnK*_int_2_F	forward	GGGTGCTCAACCTACAGAAA	self-designed
*rps16_trnK*_int_2_R	reverse	ACGAGGCAATCAAAACATTG	self-designed
trnD-trnT	forward	ACCAATTGAACTACAATCCC	Demesure *et al* 1995
trnD-trnT	reverse	CTACCACTGAGTTAAAAGGG	Demesure *et al* 1995
trnH-psbA	forward	ACTGCCTTGATCCACTTGGC	Hamilton 1999
trnH-psbA	reverse	CGAAGCTCCATCTACAAATGG	Hamilton 1999
trnQ-rps16	forward	GCGTGGCCAAGYGGTAAGGC	Shaw *et al* 2007
trnQ-rps16	reverse	GTTGCTTTYTACCACATCGTTT	Shaw *et al* 2007
trnS-trnFM	forward	GAGAGAGAGGGATTCGAACC	Demesure *et al* 1995
trnS-trnFM	reverse	CATAACCTTGAGGTCACGGG	Demesure *et al* 1995
*trnS-trnFM*_int.F	forward	ACTCAGCCATCTCTCCGAAA	self-designed
*trnS-trnFM*_int.R	reverse	TTTGGGGTGAGAGGAAAAGA	self-designed
*trnS-trnFM*_int_2_F	forward	AACCACTCAGCCATCTCTCC	self-designed
*trnS-trnFM*_int_2_R	reverse	GAACCCCTACACTATCACGG	self-designed
trnV-ndhC	forward	GTCTACGGTTCGARTCCGTA	Shaw *et al* 2007
trnV-ndhC	reverse	TATTATTAGAAATGYCCARAAAATATCATATTC	Shaw *et al* 2007

Appendix 3.5 Chloroplast monomorphic loci

Primer name	Direction	Primer sequence	Author
MH 05	forward	TCCGCTCAAGCTTATCAATG	Vania Azevedo - unpublished
MH 05	reverse	ACTAGGGACCGGAAAAGGAA	Vania Azevedo - unpublished
MH 09	forward	GGTTTTCCTGCGTACCTCCT	Vania Azevedo - unpublished
MH 09	reverse	GAGTGGGAGTGAGAGGCTGT	Vania Azevedo - unpublished
MH 10	forward	GTCGAGGAGGGCTTCTGTAA	Vania Azevedo - unpublished
MH 10	reverse	TGGAGTGAAGAAGAGGAGTTGTT	Vania Azevedo - unpublished
MH 14	forward	GACCCTCACTCAGGCTACGA	Vania Azevedo - unpublished
MH 14	reverse	ACTTACAGTGGGCGGATGAT	Vania Azevedo - unpublished
MH 18	forward	GCGCTAAGGGACTCTTCTTG	Vania Azevedo - unpublished
MH 18	reverse	ACCAAAGTCTCGTGGGGTAA	Vania Azevedo - unpublished
MH 27	forward	CTGGCAGTGCTGCTAAGTGA	Vania Azevedo - unpublished
MH 27	reverse	CAAGTCCGGCCATAATATAACA	Vania Azevedo - unpublished
MH 28	forward	CATTCATGTCGAGGATGCTG	Vania Azevedo - unpublished
MH 28	reverse	AACAAAAGCGCGCACAAA	Vania Azevedo - unpublished

Appendix 3.6 ITS PCR recipe, 25 µl reaction

Reagent	volume
ddH_2O	5.75 µl
dNTP (2mM)	2.5 µl
NH_4 (10x reaction buffer)	2.5 µl
$MgCl_2$ (25mM)	1.25 µl
Forward primer (10µM)	0.75 µl
Reverse primer (10µM)	0.75 µl
Betaine (5mM)	10 µl
Bovine Serum Albumin (BSA 0.4%)	0.25 µl
BioTaq DNA polymerase (Bioline UK) 5 units/µl	0.25 µl
DNA template	1 µl

Appendix 3.7 Chloroplast PCR recipe, 25 µl reaction

Reagent	volume
ddH_2O	15.25 µl
dNTP (2mM)	2.5 µl
NH_4 (10x reaction buffer)	2.5 µl
$MgCl_2$ (25mM)	1.25 µl
Forward primer (10µM)	0.75 µl
Reverse primer (10µM)	0.75 µl
Bovine Serum Albumin (BSA 0.4%)	0.8 µl
BioTaq DNA polymerase (Bioline UK) 5 units/µl	0.2 µl
DNA template	1 µl

Appendix 3.8 ITS PCR program A (for ITS 1, 2g, 3p, 4, 5, & 8)

1. Incubate at 95ºC for 5 minutes (denaturation)
2. Incubate at 95ºC for 30 seconds (denaturation)
3. Incubate at 50ºC for 30 seconds (annealing)
4. Incubate at 72ºC for 1 minute and 30 seconds (extension)
5. Cycle to step 2 for 34 more times
6. Incubate at 72ºC for 8 minutes (extension)
7. Incubate at 10ºC “forever” (finished)

The second, nested PCR uses the same program as above, but step 5 has only 29 cycles.

Appendix 3.9 ITS PCR program B (for ITS 18SF, 26SR, N18L, C26A, 5.8SN, 5.8SC)

1. Incubate at 95º C for 2 minutes (denaturation)
2. Incubate at 95º C for 30 seconds (denaturation)
3. Incubate at 56º C for 30 seconds (annealing)
4. Incubate at 72º C for 1 minute & 30 seconds (extension)
5. Cycle to step 2 for 3 more times
6. Incubate at 95º C for 30 seconds (denaturation)
7. Incubate at 54º C for 30 seconds (annealing)
8. Incubate at 72º C for 1 min. 30 seconds (extension)
9. Cycle to step 6 for 3 more times
10. Incubate at 95º C for 30 seconds (denaturation)
11. Incubate at 53º C for 30 seconds (annealing)
12. Incubate at 72º C for 1 minute & 30 seconds (extension)
13. Cycle to step 10 for 31 more times
14. Incubate at 72º C for 7 minute (extension)
15. Incubate at 10º C "forever" (finished)

Appendix 3.10 Chloroplast PCR program

Shaw's Rpl16 program for chloroplast regions:
1. Incubate at 80º C for 5 minutes (denaturation)
2. Incubate at 95º C for 1 minute (denaturation)
3. Incubate at 50º C for 1 minute (annealing)
4. Ramp to 65º C at 0.3º C per second
5. Incubate at 65º C for 4 minutes (extension)
6. Cycle to step 2 for 29 more times
7. Incubate at 65º C for 5 minutes (extension)
8. Incubate at 10ºC "forever" (finished)

Appendix 3.11 Monomorphic loci PCR program

1. 94º C for 5 min. (denaturation)
2. 94º C for 1 min. (denaturation)
3. 52º C for 1 min. (annealing)
4. 72º C for 1 min. (extension)
5. Cycle to step 2 for 29 more times
6. 72º C for 7 min. (extension)
7. 10º C "forever" (finished)

Appendix 3.12 Annealing temperatures (at step 3) for different monomorphic loci

Loci	T°C
Mh05	56
Mh09	56
Mh10	56
Mh14	52
Mh18	56
Mh27	60
Mh28	56

Appendix 3.13 ExoSAP-IT protocol, 7 µl reaction

Reagent	volume
PCR product	5 µl
ExoSAP-IT (GE Healthcare)	2 µl

Appendix 3.14 ExoSAP-IT PCR purification protocol

1. Incubate at 37º C for 15 minutes
2. Incubate at 80º C for 15 minutes
3. Incubate at 10º C "forever"

Appendix 3.15 CEQ sequencing PCR, 10 µl reaction

Reagent	volume
ddH_2O	1 µl
DCTS Quickstart mix	4 µl
Primer (10µM)	1 µl
DNA template	4 µl

Appendix 3.16 CEQ sequencing PCR protocol

1. Incubate at 96º C for 20 seconds (denaturation)
2. Incubate at 50º C for 20 seconds (annealing)
3. Incubate at 60º C for 4 minutes (extension)
4. Cycle to 1 for 34 times
5. Incubate at 4º C "forever" (finished)

Appendix 3.17 CEQ Sequencing PCR clean-up/stop solution, 5 µl reaction

Reagent	volume
Sigma H_2O	1.6 µl
NaOAc (3mM) pH5.2	2 µl
Glycogen	1 µl
EDTA (0.5 mM)	0.4 µl

Appendix 3.18 BigDye recipe, 10 µl reaction

Reagent	volume
ddH_2O	5.68 µl
BigDye (Applied Biosystems)	1 µl
Sequencing Buffer	2 µl
Primer (10µM)	0.32 µl
DNA template	1 µl

Appendix 3.19 BigDye sequencing PCR protocol

1. Incubate at 95º C for 30 seconds (denaturation)
2. Incubate at 50º C for 20 seconds (annealing)
3. Incubate at 60º C for 4 minutes (extension)
4. Cycle to 1 for 24 times
5. Incubate at 4º C "forever" (finished)

Chapter IV- Phylogenetic support for *Manilkara*: Monophyly and generic delimitation in the tribe Mimusopeae and subtribe Manilkarinae

4.1 Introduction

In Chapter II the current classification of the tribe Mimusopeae, subtribe Manilkarinae and the genus *Manilkara* was introduced. This chapter investigates the validity of Pennington's (1991) classification and Swenson & Anderberg's (2005) re-classification through phylogeny reconstruction of the nuclear ribosomal region ITS and the chloroplast regions *trnS-trnFM*, *rpl32-trnL* and *rps16-trnK*. Specifically, the monophyly of the tribe Mimusopeae, subtribes Glueminae, Mimusopinae and Manilkarinae, and the genera *Manilkara*, *Labramia*, *Faucherea* and *Labourdonnaisia* are tested with molecular data.

Different phylogeny reconstruction methods discussed in Chapter III (maximum parsimony and Bayesian inference) are compared, as are nuclear and chloroplast data to determine whether estimates from the different methods and genes agree with one another. If conflict exists between nuclear and chloroplast tree topologies, this may be due to chloroplast capture, which is the introgression of the chloroplast genome from one species into another following hybridization (Tsitrone *et al* 2003). Introgression of the nuclear genome is also a possible explanation. The implications of any incongruence in the data sets are discussed, as are suggestions for reclassification supported by molecular evidence.

4.2 Phylogenetic methods

4.2.1 Taxon selection and sequence data

Evolutionary relationships were reconstructed using nuclear (ITS) and chloroplast (*rpl32-trnL, rps16-trnK and trnS-trnFM*) sequences. Sequences with Edinburgh DNA (EDNA) numbers were generated as part of this study as per the methods outlined in Chapter II, whereas sequences designated AA were donated by Jenny Smedmark and Arne Anderberg at the Natural History Museum in Stockholm or were taken from previously published data (Swenson & Anderberg 2005, Smedmark *et al* 2006). *Mimusops* sequences designated as being from Geneva were contributed by Yamama Naciri and Laurent Gautier. In total, 111 accessions were included in the analysis (Appendix 4.1), and among these, all four sequences were available from 83 accessions. From the remaining 27 accessions, certain sequences either could not be amplified (EDNA material) or were not available (sequences donated from other groups). These accessions comprised 16 from which only ITS sequences were available; four from which ITS and only one of the three cpDNA sequences were available; three from which all chloroplast sequences were available, but ITS was not; and five from which two of the three chloroplast sequences were available, but ITS was not (Appendix 4.1). Therefore, the nuclear and chloroplast datasets comprised 101 and 95 accessions respectively. Phylogenetic trees were rooted using *Sarcosperma*, which has been shown in previous studies to be sister to the rest of the family (Anderberg & Swenson 2003).

4.2.2 Chloroplast gap-coding

Potentially informative indels in the chloroplast dataset were coded according to the simple indel coding method of Simmons & Ochoterena (2000). Twenty six indels were coded in *rpl32-trnL*, twenty one in *rps16-trnK* and thirteen in *trnS-trnFM*. Ambiguous alignment regions 113-118 and 380-459 in *rps16-trnK* were excluded. Indel events in ITS were so frequent that their coding as additional characters was deemed to be too ambiguous. Gaps were treated as missing data and all characters were equally weighted.

4.2.3 Why nuclear and chloroplast regions were not combined

It is a common misconception that the topology of gene trees is directly representative of species trees. Due to various biological factors such as gene duplication, horizontal gene transfer, introgression and lineage sorting, gene trees often differ substantially in topology from the species trees in which they are embedded (Maddison 1997, Page & Holmes 1998, Edwards 2009, Degnan & Rosenberg 2009). In order to reduce error in species tree reconstruction, it has been the trend in phylogenetics to generate multiple single gene data sets and concatenate them in a "total evidence" approach (Kluge 1989, 2004). However, some of the very characteristics which make this method popular (the fact that genes evolving at different rates and under different selection pressures can resolve nodes at different depths in a phylogeny) also make it problematic by potentially confounding topology as well as branch lengths and support values. According to Edwards (2009) "concatenation will, under many circumstances, be a worse approximation of the underlying diversity of gene trees than will approaches that allow for gene tree heterogeneity, because gene trees will always differ from one another subtly even when topologically congruent." Therefore, as a rule, datasets from different genes should not be combined if a) they can be shown to have incongruent topologies, and/or b) their DNA sequences have significantly different substitution rates.

Consequently, in this study the nrDNA and cpDNA datasets were analysed separately because hard (or well supported) incongruence was demonstrated between the topologies reconstructed from the two genomes, and because they have different rates of evolution. Additionally, in a combined analysis, the lack of informative sites in the chloroplast data compared to the nuclear data would translate into an over-representation of the ITS data ("swamping") and poor support values due to conflicting topologies. However, the different chloroplast regions (*rpl32-trnL, rps16-trnK* and *trnS-trnFM*) were combined in a single analysis because they are all part of a single, linked region with similar selection pressures, and substitution rates. The nrDNA and cpDNA datasets were analysed separately using two common phylogenetic methods: heuristic searches in maximum parsimony (Fitch 1971) and Bayesian inference (Yang & Rannala 1997) using Metropolis-coupled Markov chain Monte Carlo (MCMCMC) searches (Larget & Simon, 1999).

4.2.4 Parsimony analysis settings in PAUP*

Maximum parsimony analyses were carried out using PAUP* version 4.0b10 (Swofford 1993, 2003) in a two-step process. The first analysis involved a heuristic search of 10,000 replicates beginning with multiple starting trees, with random stepwise addition and branch swapping on best trees only, using tree-bisection-reconnection (TBR), MulTrees on and saving no more than 10 trees at each step with a score of 10. Characters were equally weighted and the character states were unordered. All trees from this analysis were saved and input in a second heuristic search using the same settings but saving no more than 10,000 trees with a score of 5. A strict consensus tree was generated from all trees in memory. Bootstrapping (Felsenstein 1985) was used to recover support values for clades. A full heuristic search was carried out with 10,000 replicates of simple stepwise sequence addition, swapping on best trees only, using TBR and MulTrees with 10 trees saved per replicate with a score of 10 and retaining groups with a frequency greater than 50%. Bootstrap values over 85% were considered well-supported, whereas 75-84% were moderately supported and 50-74% were considered poorly supported.

4.2.5 Bayesian analysis

4.2.5.1 MrBayes settings

Bayesian analyses were carried out using MrBayes 3.1 (Huelsenbeck & Ronquist 2001). Two independent runs of four MCMCMC chains each (three heated and one cold) were run with a temperature setting of 0.15 for 10,000,000 generations for the chloroplast and a temperature of 0.10 for 8,000,000 generations for the ITS dataset, which was found to provide sufficient mixing between chains and convergence between runs. Trees were sampled every 8,000 generations and a 10% burn-in was removed from the sampled set of trees (leaving a final sample of 1000 and 800 trees for chloroplast and ITS respectively), which were used to produce a majority rule consensus tree. Convergence of models was determined to have occurred when the standard deviation of split frequencies for two runs reached 0.01 (Ronquist *et al* 2005). This was backed-up by visual confirmation of parameter convergence of traces in Tracer v.1.5 (Rambaut & Drummond 2009). Clade support values are posterior probabilities (pp); pp values of 100-95 % indicate strong support, values of 94-90 % indicate moderate support and values between 89-55% indicate weak support for nodes. The output tree files were visualised in FigTree v.1.3.1

4.2.5.2 Partitioning and model selection

The suitability of different substitution models and partitioning strategies were assessed using Bayes factor comparisons between alternative approaches (Kass & Rafferty 1995). The appropriate model was selected through a process of model simplification, which incorporates phylogenetic uncertainty while comparing different models of sequence evolution. Multiple analyses were run with different models. All base substitution types were represented in each partition of the sequence data, so testing began with a GTR + I + G (general time reversible model, plus gamma distributed rate variation, plus a proportion of

invariant sites) for a single partition followed by more complex partitioning strategies and model simplification until the appropriate strategy was determined (Tables 4.2, 4.3 and 4.4). The most complex partitioning strategy in ITS had three separate partitions for ITS 1, 5.8S and ITS 2 respectively, while for the chloroplast regions, the most complex strategy included four partitions for *rps16-trnK*, *rpl32-trnL*, *trnS-trnFM* and coded gaps respectively. Natural logarithms (ln) of harmonic mean likelihoods (HML) were given in the .log file output in MrBayes and Bayes factors were determined using twice their difference (2ΔlnHML) for each model, each run. According to Kass & Rafferty (1995), a Bayes factor difference >10 is an appropriate measure of support for one model over another. However, given that there can be a difference of 15 points in log likelihood values between the two chains in a single analysis, a difference of >15 is instead used here to indicate decisive support of one model over another. The simplest partitioning strategy and substitution models supported in the Bayes factor tests for which posterior distributions were stationary and unimodal were selected for each region.

An assessment of the clock-likeness of the data was also made through Bayes factor comparison between an unconstrained non-clock, a uniform clock and a birth-death strict clock model in MrBayes for the molecular model selected. A birth-death strict clock model was supported over non-clock and uniform clock models for each region.

4.2.6 Interpretation of support values

Clade support comparisons between bootstrap and posterior probability values were interpreted following Taylor & Piel (2004), who found that parsimony bootstrap values were not significantly different from accuracy and may not represent underestimates, whereas posterior probabilities may overestimate values, but have low false-positive error rates (0% to 2.8%) at the highest values (.99%). Therefore, in comparison strong support was viewed as a combination of >85% for parsimony bootstrap values and >95% for posterior probabilities.

4.3 Results

4.3.1 Model selection and Bayes factor tests

The Bayes factor test revealed that for the ITS dataset three partitions (ITS1, 5.8S and ITS2) each with a GTR + I + G model, was the most appropriate (Table 4.3). Additionally, each chloroplast region (*rpl*32-*trn*L, *rps*16-*trn*K, *trn*S-*trn*FM, and coded gaps) was designated its own partition and a GTR + I model was found to be most suitable for each partition (Table 4.4). In instances where the Bayes factor difference between two competing models was not deemed significant, the simpler model with fewer parameters was preferred (i.e. as in Tables 4.2 and 4.4). A summary of data for each region including sequence length, number of informative sites, partitions, model choice and clock setting is presented in Table 4.1.

Table 4.1 Summary of sequence data and settings

region	partition	total aligned length (including gaps) + indels coded	parsimony informative sites	variable sites	model	clock setting
nuclear	ITS 1	316	127	194	GTR + I + G	birth-death clock
nuclear	5.8S	165	14	25	GTR + I + G	birth-death clock
nuclear	ITS2	297	134	184	GTR + I + G	birth-death clock
chloroplast	*rpl*32-*trn*L	1130 + 26 indels	76	170	GTR + I	birth-death clock
chloroplast	*rps*16-*trn*K	1134 + 21 indels	60	132	GTR + I	birth-death clock
chloroplast	*trn*S-*trn*FM	999 + 13 indels	41	83	GTR + I	birth-death clock

Table 4.2 Bayes factor comparison of different models and clock settings in the ITS dataset. First partitioning strategy was tested, then model selection, then the appropriate clock model. A Bayes factor value of 2.22 shows support for the three partition model: ITS1 = GTR+I+G, 5.8S = GTR+I+G, ITS2 = GTR+I+G with a birth-death clock, however this value is not significant by the criteria set out and the simpler model of ITS1= GTR+I+G, 5.8S = GTR+I, ITS2 = GTR+I+G with a birth-death clock was, therefore, preferred.

Models per partition and clock strategy imposed	Ln HML	Bayes factor: 2ΔlnHML comparison of values shown between the chosen model and each alternative
1 partition (ITS1 + 5.8S + ITS2) GTR+I+G	-6144.36	248.64
2 partition ITS 1 & 2 – GTR+I+G 5.8s – GTR+I+G	-6036.35	281.26
2 partition ITS 1 & 2 – GTR+I+G 5.8s – GTR+I	-6048.29	305.13
3 partition ITS1 - GTR+I+G 5.8S - GTR+I+G ITS2 - GTR+I+G	-5915.56	39.68
3 partition ITS1 - GTR+I+G 5.8S - GTR+I ITS2 - GTR+I+G	-6007.95	224.46
3 partition ITS1 - GTR+I+G 5.8S - GTR+G ITS2 - GTR+I+G	-5999.17	206.9
3 partition ITS1 - GTR+I+G 5.8S - GTR+I+G ITS2 - GTR+I+G uniform clock	-5919.3	47.16
3 partition ITS1 - GTR+I+G 5.8S - GTR+I+G ITS2 - GTR+I+G birth-death clock	-5894.61	2.22
3 partition **ITS1 - GTR+I+G** **5.8S - GTR+I** **ITS2 - GTR+I+G** **birth-death clock**	**-5895.72**	**Chosen model**
3 partition ITS1 - GTR+I+G 5.8S - GTR+G ITS2 - GTR+I+G birth-death clock	-5901.7	11.96

Table 4.3 ITS dataset model testing part two. After the initial Bayes factor tests, the number of taxa in the dataset set was increased, so the best three models were re-tested to determine whether the optimal model had changed with the additional taxa. The most appropriate model was determined to be the three partition model: ITS1 = GTR+I+G, 5.8S = GTR+I+G, ITS2 = GTR+I+G with a birth-death clock, supported by a Bayes factor value of 12.72.

Models per partition and clock strategy imposed	**Ln HML**	**Bayes factor: 2ΔlnHML comparison of values shown between the chosen model and each alternative**
3 partition ITS1 - GTR+I+G 5.8S - GTR+I ITS2 - GTR+I+G birth-death clock	-7547.15	17.2
3 partition **ITS1 - GTR+I+G** **5.8S - GTR+I+G** **ITS2 - GTR+I+G** **birth-death clock**	**-7538.55**	**Chosen model**
3 partition ITS1 - GTR+I+G 5.8S - GTR+G ITS2 - GTR+I+G birth-death clock	-7544.91	12.72

Table 4.4 Bayes factor comparison of different models and clock settings in the chloroplast dataset. First the partitioning strategy was tested, then model selection, then the appropriate clock model. A Bayes factor value of 14.86 shows moderate support for the models: *rpl32-trnL* = GTR+I, *rps16-trnK* = GTR+I, *trnS-trnFM* = GTR+I+G, gap coding rates = equal and a birth-death clock. However, the simpler model with all three partitions using GTR+I was not significantly worse and was, therefore, chosen.

Models per partition and clock strategy imposed	**Ln HML**	**Bayes factor: 2ΔlnHML comparison of values shown between the chosen model and each alternative**
4 partitions *rpl32-trnL* = GTR+I+G *rps16-trnK* = GTR+I+G *trnS-trnFM* = GTR+I+G gap coding rates = + G	-7685.39	131.72
4 partitions *rpl32-trnL* = GTR+I *rps16-trnK* = GTR+I+G *trnS-trnFM* = GTR+I+G gap coding rates = equal	-7683	126.94
4 partitions *rpl32-trnL* = GTR+I *rps16-trnK* = GTR+I *trnS-trnFM* = GTR+I+G gap coding rates = equal	-7678.67	118.28
4 partitions *rpl32-trnL* = GTR+I *rps16-trnK* = GTR+I *trnS-trnFM* = GTR+I+G gap coding rates = equal uniform clock	-7672.81	106.56
4 partitions *rpl32-trnL* = GTR+I *rps16-trnK* = GTR+I *trnS-trnFM* = GTR+I+G gap coding rates = equal birth-death clock	-7612.1	14.86
4 partitions ***rpl32-trnL* = GTR+I** ***rps16-trnK* = GTR+I** ***trnS-trnFM* = GTR+I** **gap coding rates = equal** **birth-death clock**	**-7619.53**	**Chosen model**

4.3.2 ITS phylogenies - overview of clade support

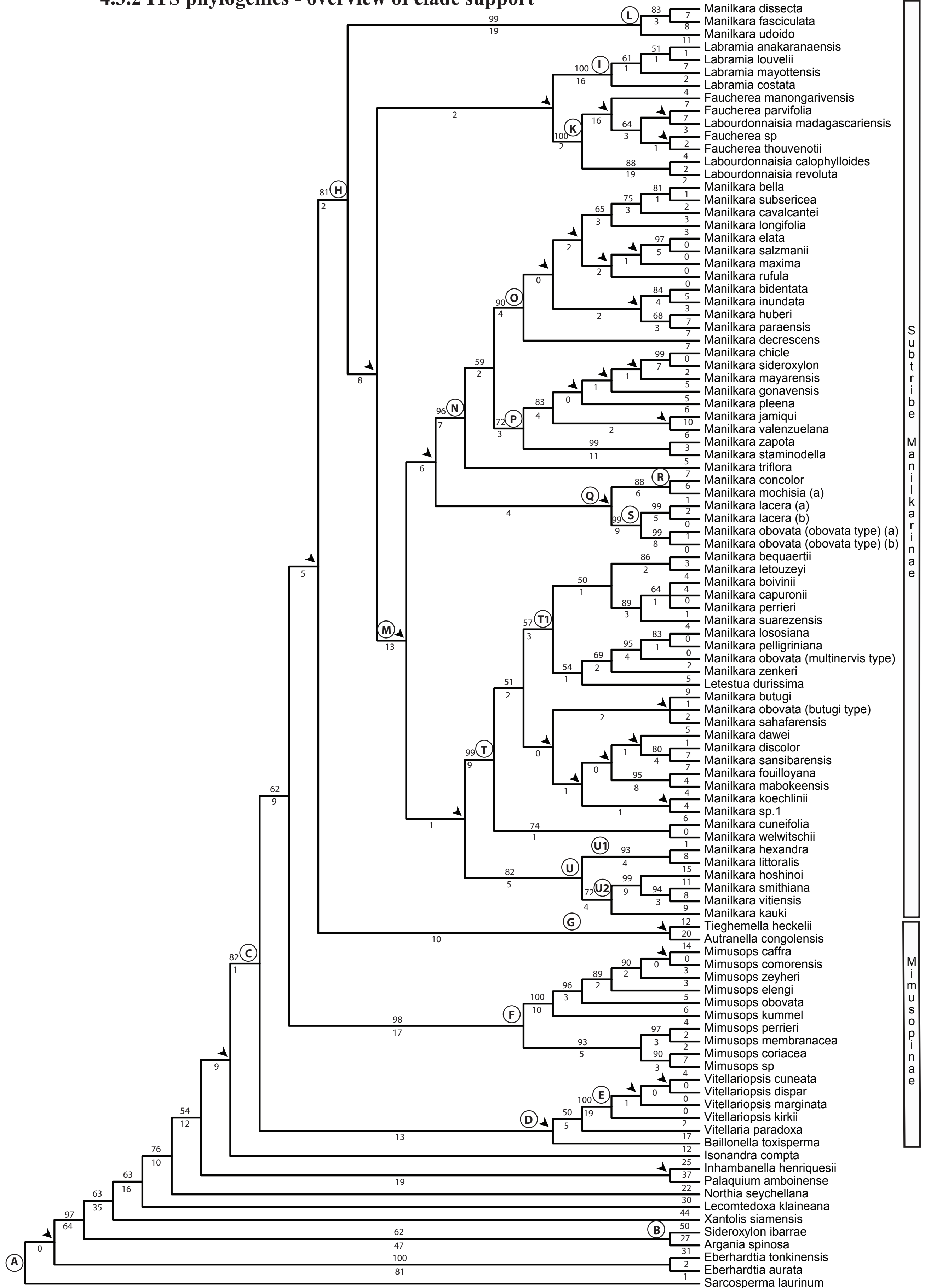


Figure 4.1 One out of ten thousand most parsimonious trees reconstructed from the ITS dataset in PAUP*. Where clades in this tree have bootstrap support >50%, bootstrap values are indicated above branches; branch lengths are indicated below. Arrows indicate branches which collapse in the strict consensus tree. The tree has a length of 1167 steps, with a consistency index of 0.527, a retention index of 0.754 and a rescaled consistency index of 0.397.

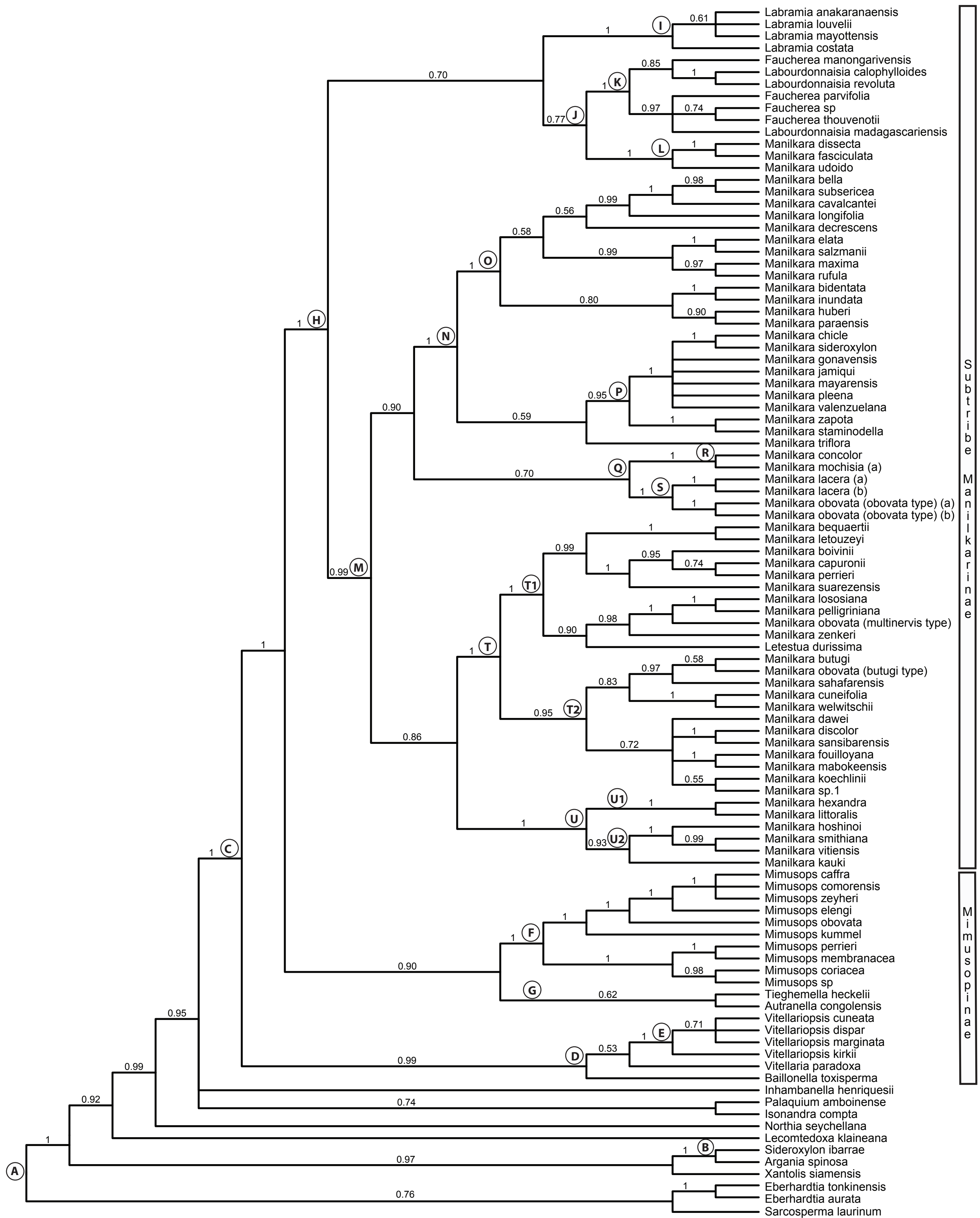


Figure 4.2 Bayesian majority rule consensus tree reconstructed from the ITS dataset in MrBayes. Posterior probability values are indicated above branches.

The parsimony and Bayesian phylogenies are both composed of a basal grade including the outgroup *Sarcosperma*, *Eberhardtia*, a grade of *Argania, Sideroxylon, Xantolis*, *Lecomtedoxa*, *Northia*, *Palaquium, Inhambanella* and *Isonandra* within which is nested a large clade of all other species (clade C). Clade C is strongly supported in the Bayesian analysis (pp 1) and moderately supported (bs 82) in the parsimony analysis. It comprises a basal clade of *Baillonella*, *Vitellaria* and *Vitellariopsis* (clade D), the strongly supported (bs 98/pp 1) clade F containing all *Mimusops* species examined, a weakly supported (bs <50/pp 62) *Autranella* + *Tieghemella* clade (clade G) and a clade containing the subtribe Manilkarinae (clade H).

Clade D is strongly supported (pp 0.99) in the Bayesian but weakly supported (bs <50) in the parsimony analysis. Within this clade the *Vitellariopsis* subclade (clade E) is also a strongly supported (bs 100/pp 1) monophyletic group. Relationships between clades F, G and H are not well-resolved, with a weakly supported (pp 0.90) sister relationship between clades F and G in the Bayesian analysis only. Clades D, F and G thus contain the genera *Vitellariopsis, Vitellaria*, *Baillonella, Mimusops*, *Tieghemella* and *Autranella*, which comprise the subtribe Mimusopinae in Pennington's (1991) classification. While these genera are shown to be closely related, they are not resolved as a monophyletic group, but instead form a grade, within which a monophyletic subtribe Manilkarinae (clade H) is nested.

The Manilkarinae clade (H) is moderately supported in the parsimony (bs 81), but strongly supported in the Bayesian analysis (pp 1). It is comprised of a *Labramia* subclade (clade I), a *Faucherea/ Labourdonnaisia* subclade (clade K), a clade of three Asian *Manilkara* species (*M. fasciculata*, *M. udoido* and *M. dissecta*; clade L), and the large clade M, which comprises all other *Manilkara* species examined, plus *Letestua durissima*. The first three clades all have 1.0 Bayesian support and 100% (clades I and K) or 99% (clade L) bootstrap support. However, relationships between these four clades were not well-supported, with a clade (J) comprising subclades K and L weakly supported (pp 0.77) in the Bayesian analysis only. Although *Labourdonnaisia* and *Faucherea* together form a strongly supported (bs 100/pp 1) monophyletic group (clade K), neither genus alone is monophyletic.

Clade M was strongly supported in the Bayesian analysis (0.99 pp), but unsupported in the parsimony analysis (bs <50). It comprised four continent-specific clades: a small African clade (Q), a large African-Madagascan clade (T), an Asian clade (U) and a Neotropical clade (N). The small African clade Q is poorly supported in both analyses (bs <50/pp 0.70), but its constituent subclades R (bs 88/pp 1) and S (bs 99/pp 1) have strong support. The larger of the two African clades (T), is strongly supported in both the parsimony and Bayesian analyses (bs 99/pp 1), as is the Asian clade (U) (bs 82/pp 1). The Neotropical clade (N) was also strongly supported (bs 96/pp 1). It comprises two subclades of South American (clade O) and Central American/Caribbean (clade P) species, plus *M. triflora*. Relationships between these three were not resolved.

4.3.3 Chloroplast phylogenies – overview of clade support

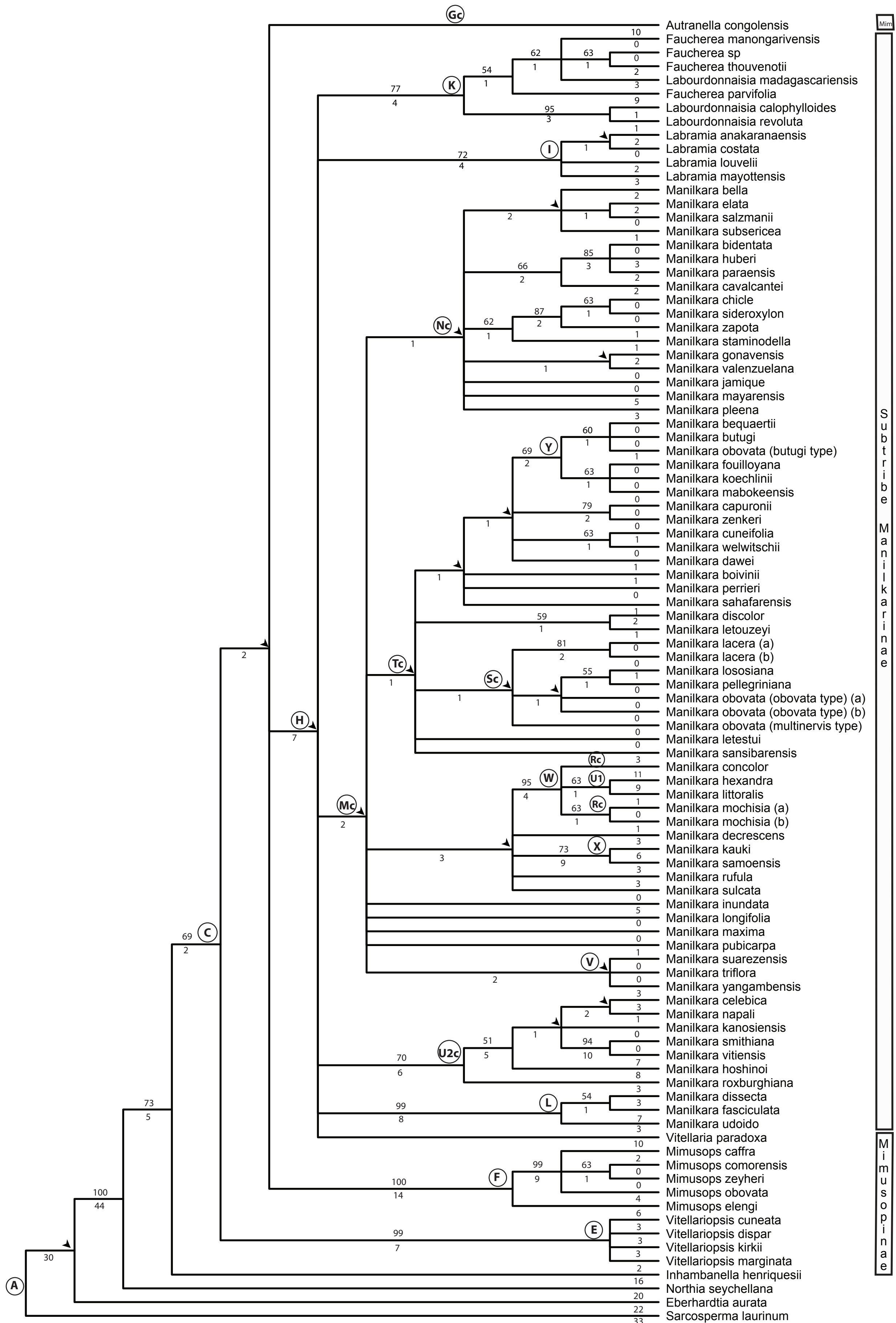


Figure 4.3 One out of ten thousand most parsimonious trees reconstructed from the chloroplast dataset recovered in PAUP*. Bootstrap values are indicated above branches and branch lengths are indicated below. Arrows indicate branches which collapse in a strict consensus. The tree has a length of 495 steps, with a consistency index of 0.822, a retention index of 0.843 and a rescaled consistency index of 0.693.

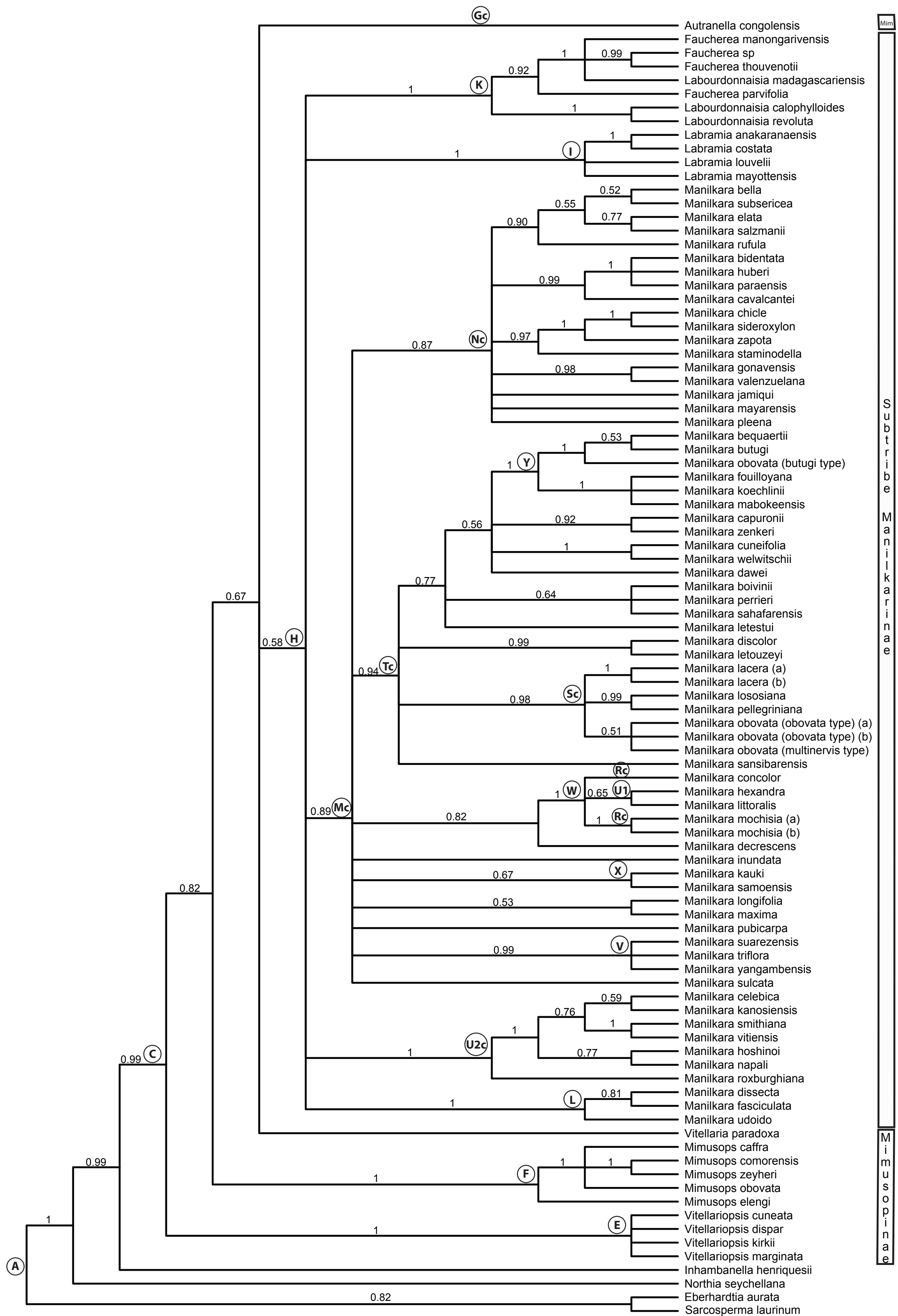


Figure 4.4 Bayesian majority rule consensus tree reconstructed from the chloroplast dataset in MrBayes. Posterior probability values are indicated above branches.

The two chloroplast phylogenies contain some clades (G_C, M_C, N_C, R_C, S_C, T_C, $U2_C$), which are similar to those found in the ITS phylogenies, but have more or less taxa. Therefore, to demonstrate the topological similarities between these clades in the different analyses, yet differentiate between them, those clades in the chloroplast tree which differ slightly are annotated with a $_C$. For instance, the Neotropical *Manilkara* clade is marked as clade N in the ITS analyses, but N_C in the chloroplast analyses.

Both the parsimony and Bayesian chloroplast phylogenies included the large clade C, with *Inhambanella, Northia*, *Eberhardtia* and *Sarcosperma* as its successive sister groups. In the Bayesian analysis, clade C comprised clade E (*Vitellariopsis*), clade F (*Mimusops*), *Autranella* (the only representative of the ITS clade G examined for cpDNA), *Vitellaria* and the large clade H (subtribe Manilkarinae). The parsimony analysis differed from the Bayesian analysis in that *Vitellaria* was resolved within clade H, although there was no bootstrap support for this clade. Therefore, clade D from the ITS analysis is not recovered in these chloroplast analyses. In both analyses, clade E was basal within clade C. This relationship was moderately supported in the parsimony analysis (bs <69), but strongly supported in the Bayesian analysis (pp 0.99). Clade H lacked support in either analysis (bs 58/pp <50), (both with and without *Vitellaria*) and is only worthy of mention here because it was strongly supported in the ITS analysis. As with the ITS analysis, Pennington's (1991) subtribe Mimusopinae was non-monophyletic and was instead resolved as a basal grade to the subtribe Manilkarinae (clade H excluding *Vitellaria*). Clades E and F were strongly supported in both analyses (bs 99/pp 1 and bs 100/pp 1 respectively).

In both analyses, clade H included five clades: clade I (*Labramia*), clade K (*Faucherea/ Labourdonnaisia*), clade L (*Manilkara fasciculata*, *M. udoido* and *M. dissecta*), clade U2 (seven Asian *Manilkara* species), and the large clade M_C, which contained all *Manilkara* species except those in clades L and $U2_C$. Clade M_C, therefore, differed from Clade M in the ITS analysis in that it did not contain clade $U2_C$. Relationships between these clades were not resolved. In the parsimony analysis *Vitellaria* was also placed in clade H, in an unresolved position. Clade L was strongly supported in both analyses (bs 99/pp 1), whereas Clades I (*Labramia*) and K (*Faucherea/Labourdonnaisia*) had strong Bayesian support (pp 1) but moderate bootstrap support of 72 and 77 respectively. Within clade K, neither *Faucherea* nor *Labourdonnaisia* were monophyletic. Clade $U2_C$ was more strongly supported in the Bayesian than the parsimony analysis (bs 70/pp 1), and contained *M. smithiana, M. vitiensis, M. hoshinoi, M. celebica, M. kanosiensis, M. napali,* and *M. roxburghiana*, the last four of which were not included in the ITS analysis. It was, otherwise, similar to one of the two subclades of the ITS analyses (clade U2), except that it did not include *M. kauki*. Clade U was, therefore, polyphyletic in the cpDNA analysis.

The main *Manilkara* clade (clade M_C) was poorly supported (bs <50/0.89 pp) in both the parsimony and Bayesian analyses. Within Clade M_C, only two large clades were resolved, clades T_C (African species) and N_C (Neotropical species), plus two small mixed area clades, V and W; the remaining seven species had positions that were either unresolved or poorly supported within Clade M_C. Clade W was well-supported (bs 95/pp 1), and is comprised

of the Asian clade U1 (*M. hexandra* and *M. littoralis*), plus all the members of the African clade R (*M. concolor* and *M. mochisia* a and b). Clade V was supported only in the Bayesian analysis (pp 0.99) and included three species: *Manilkara triflora, M. yangambensis,* and *M. suarezensis*. These three species differed both in their geographic distribution (Brazil, Congo and Madagascar respectively) and their placement in the ITS tree. One other small clade (X) within clade M_C contained *M. kauki* and *M. samoensis* with weak support (bs 73/pp 0.67). Relationships between these clades and the remaining species (*M. longifolia* and *M. maxima, M. inundata*, *M. pubicarpa* and *M. sulcata*) were either unresolved or very weakly supported.

Clade N_C was poorly supported (bs <50/pp 0.87). It contained most of the species from clade N of the ITS analysis, but was not identical because *M. triflora, M. longifolia* and *M. maxima* fell outside it. Hence clade N_C contained only Neotropical taxa, but unlike clade N from the ITS analysis, it did not contain all Neotropical taxa examined. In the Bayesian cpDNA analysis Clade N_C thus comprised *M. bella, M. bidentata, M. cavalcantei, M. chicle, M. elata, M. gonavensis, M. huberi, M. jamiqui, M. mayarensis, M. paraensis, M. pleena, M. rufula, M. salzmanii, M. sideroxylon, M. staminodella, M. subsericea, M. valenzuelana* and *M. zapota,*. However, in the parsimony cpDNA analysis, Clade N_C is different again, because *M. rufula* is placed elsewhere in an unresolved position.

Clade T_C received moderate support in the Bayesian analysis only (bs < 50/pp 0.94). Within it, two large subclades received strong Bayesian support. One was clade S_C (pp 0.98), which contained all the species from clade S from the ITS analysis (i.e. *M. lacera* a & b, and *M. obovata* - obovata type a & b) plus *Manilkara obovata* (multinervis-type) as well as *M. lososiana*, and *M. pellegriniana*. The second, clade Y, comprised *M. bequaertii, M. butugi, M. obovata*-butugi type*, M. fouilloyana, M. koechlinii,* and *M. mabokeensis*. This clade was better supported in the Bayesian than the parsimony analysis (bs 69/pp 1), and was not present in the ITS analysis. Among species not in these clades, three species pairs received strong support in at least one analysis; these were *M. capuronii* and *M. zenkeri* (bs 79/pp 0.92), *M. cuneifolia* and *M. welwitschii* (bs 63/pp 1), *M. discolor* and *M. letouzeyi* (bs 59/pp 0.99). Relationships among the remaining species of clade T_C, i.e. *M. dawei, M. boivinii, M. perrierii, M. sahafarensis, M. letestui, M. sansibarensis*, were either unresolved or poorly supported.

4.3.4 Comparison of nuclear and chloroplast phylogenies

4.3.4.1 Congruence

The nuclear and chloroplast trees are mostly congruent, reconstructing many of the same main clades. Leaving aside species that were present in one analysis only, the following clades were present in both analyses: clade C (subtribes Mimusopinae & Manilkarinae), clade E (*Vitellariopsis*), clade F (*Mimusops*), clade G (*Autranella/Tieghemella*), clade I (*Labramia*), clade K (*Faucherea/Labourdonnaisia*), clade L (*Manilkara dissecta*, *M. fasciculata* and *M. udoido*), and clade U1 (*Manilkara hexandra* and *M. littoralis*). Clade H (Manilkarinae) is almost constant across all analyses, except for the cpDNA parsimony analysis, which also contained *Vitellaria*, though with weak support. Clades M, N and T

from the ITS analyses all appear with slight alterations in the cpDNA analyses: clade M_C (cpDNA) differed from clade M (ITS) in that it did not include clade U2; clade N_C (cpDNA) differed from clade N (ITS) in that *M. decrescens, M. rufula* (parsimony only), *M. inundata, M. longifolia*, *M. maxima M. triflora* and *M. pubicarpa* are not in clade N_C (although *M. pubicarpa* was not available for ITS); clade T_C (cpDNA) differed from clade T (ITS) in that it includes the species of ITS clade S for cpDNA only.

4.3.4.2 Incongruence

There are a number of examples of soft incongruence (i.e. poorly supported conflicting positions) between taxa in the nuclear and chloroplast analyses within the subtribe Manilkarinae and the genus *Manilkara*. As these are unsupported, they will not be discussed here in depth. However, there are also a few examples of hard incongruence (well supported conflicting positions), which are detailed below.

In the ITS analysis, all Asian species other than the three in clade L form a monophyletic clade, U. This clade comprised two subclades, U1 and U2. However, in the chloroplast analysis, clade U1 and $U2_C$ appeared in different parts of the tree: clade $U2_C$ fell outside clade M_C, whereas clade U1 (*M. hexandra* and *M. littoralis*) fell within clade W inside clade M_C. In addition, clade U2 contained *M. kauki* whereas clade $U2_C$ did not; in the cpDNA analysis *M. kauki* was weakly supported as sister to *M. samoensis* (which was not in the ITS analysis), and the position of this pair within clade M_C was unresolved. Therefore, Asian species appeared to constitute four lineages in the cpDNA analysis, but just two in the ITS analysis.

The two African species *M. mochisia* and *M. concolor* are closely associated in both analyses, however only in the ITS analysis do they form a monophyletic clade (R). In the cpDNA analysis, *M. mochisia* and *M. concolor* are resolved in a strongly supported clade (W) with the Asian clade U1 (*M. hexandra* and *M. littoralis*) which, in the ITS tree are resolved within a clade of other Asian taxa (U). Both W and R fall within clade M/M_C; however the relationships of clade W are unresolved for cpDNA, whereas for ITS, clade R is weakly linked to clade S and these together are sister to clade N with moderate (pp 0.90) Bayesian support.

Another hard incongruence between the nuclear and chloroplast trees is in the position of the three taxa: *Manilkara yangambensis*, *M. triflora* and *M. suarezensis*, which in the cpDNA analysis only formed the monophyletic clade V. In the ITS analysis, however, this clade was not present: the Brazilian *M. triflora* was basal to clade N, whereas the Madagascan *M. suarezensis* was resolved within the main African clade (T). The Congolese species *M. yangambensis* was not included in the ITS analysis.

Other instances of hard incongruence between the ITS and chloroplast analyses also exist within the main African *Manilkara* clade (T/T_C). In subclade Y, a sister relationship between *M. bequaertii*, *M. butugi* and *M. obovata* (butugi type) is moderately to strongly supported (bs 60/pp 1) in the chloroplast analysis, but not in the ITS analysis, where *M. bequaertii* is resolved in a separate clade from *M. butugi* and *M. obovata* (butugi type).

M. capuronii and *M. zenkeri* are resolved as sister to one another with moderate support (bs 79/pp .92) in the chloroplast analyses, but in separate subclades of clade T1 in the ITS analysis.

M. letouzeyi and *M. discolor* (bs 59/pp .99) are also moderately to strongly supported as sister to one another in the chloroplast analysis, but resolved in separate subclades (T1 and T2 respectively) in the ITS analysis.

M. lacera, *M. lososiana*, *M. pellegriniana*, *M. obovata* (obovata type) and *M. obovata* (multinervis type) are resolved as a strongly supported monophyletic group in the Bayesian chloroplast analysis (pp .98) but poorly supported in the chloroplast parsimony analysis (bs > 50). Within the ITS analysis, these taxa are resolved in two separate, strongly supported but distantly related clades: *M. lososiana*, *M. pellegriniana* and *M. obovata* (multinervis type) in clade T1 (bs 95/pp 1) and *M. lacera* and *M. obovata* (obovata type) in clade S (bs 99/pp 1).

Furthermore, within the South American clade (O), *M. cavalcantei*, an Amazonian species, is resolved within a clade of Atlantic coastal forest species comprising *M. bella*, *M. subsericea*, *M. longifolia*, *M. elata*, *M. salzmanii*, *M. maxima* and *M. rufula* (bs >50/pp 0.58). In the chloroplast analysis, *M. cavalcantei* is instead moderately to strongly supported (bs 66/pp .99) as belonging to a clade of other Amazonian species: *M. bidentata*, *M. huberi* and *M. paraensis*.

4.3.4.3 Comparison of phylogenetic methods, genes and a summary of monophyly tests

Overall, tree topologies recovered for an individual region using parsimony and Bayesian methods were in agreement. Parsimony bootstrap values and Bayesian posterior probabilities were also generally consistent in their level of support for clades.

The consensus on taxonomic classification between the different phylogenetic methods and gene regions is that Pennington's (1991) tribe Mimusopeae is polyphyletic, as is the subtribe Glueminae. Subtribe Mimusopinae is paraphyletic and Manilkarinae is monophyletic. Based on the classification in Govaerts *et al*'s (2001) Sapotaceae checklist, *Manilkara*, *Faucherea* and *Labourdonnaisia* are all paraphyletic and *Labramia* is monophyletic. A synopsis of these results is given in Table 4.5. The implications of these molecular results for classification in relation to morphology are discussed in more detail in the sections below.

Table 4.5 Summary of results of monophyly tests on taxa described in Pennington's (1991) classification

Taxon /region & analysis	nrDNA Bayesian	nrDNA parsimony	cpDNA Bayesian	cpDNA parsimony
Mimusopeae	polyphyletic	polyphyletic	untested	untested
Glueminae	polyphyletic	polyphyletic	untested	untested
Mimusopinae	paraphyletic	paraphyletic	paraphyletic	paraphyletic
Manilkarinae	monophyletic	monophyletic	monophyletic	monophyletic
Manilkara	paraphyletic	paraphyletic	paraphyletic	paraphyletic
Labramia	monophyletic	monophyletic	monophyletic	monophyletic
Labourdonnaisia	paraphyletic	paraphyletic	paraphyletic	paraphyletic
Faucherea	paraphyletic	paraphyletic	paraphyletic	paraphyletic

4.4 Discussion

4.4.1 Delimitation of the tribe Mimusopeae and subtribes Mimusopinae, Manilkarinae and Glueminae

The tribe Mimusopeae (composed of the three subtribes: Glueminae, Mimusopinae and Manilkarinae) as circumscribed by Pennington (1991) is resolved as polyphyletic in the ITS analyses due to the placement of the genera *Eberhardtia*, *Lecomtedoxa* and *Inhambanella*, which are classified in Pennington's subtribe Glueminae and *Northia*, which is in Pennington's subtribe Manilkarinae. However, with the exclusion of *Northia* and subtribe Glueminae, the monophyly of the remaining taxa in tribe Mimusopeae (clade C) was strongly supported in both the ITS analyses (bs 82/pp 1).

Subtribe Glueminae is composed of *Eberhardtia*, *Inhambanella*, *Lecomtedoxa, Northia*, *Gluema* and *Neolemonniera*, and is distinguished from the Mimusopinae and Manilkarinae by five-merous flowers and imbricate sepals in a single row, whereas the other subtribes have eight or six-merous flowers respectively, with valvate sepals in two rows (Pennington 1991, Anderberg & Smedmark 2003). Glueminae are well-sampled here in the ITS analysis, except for the missing genera *Gluema* and *Neolemonniera*. Regardless of the exclusion of these two taxa, it is evident that the morphological classification of this subtribe does not stand up to molecular scrutiny. Instead, the included genera formed part of a grade of lineages sister to clade C in the ITS analysis. *Eberhardtia* is resolved in a basal position sister to *Sarcosperma*, whereas *Lecomtedoxa* and *Northia* are resolved independently in separate lineages in-between the Sideroxyleae (represented by *Sideroxylon* and *Argania*) and Isonandreae (represented by *Isonandra* and *Palaquium*). Finally, *Inhambanella* is resolved independently as sister to the Isonandreae and basal to the Mimusopinae-Manilkarinae clade (C). Therefore, subtribe Glueminae is polyphyletic in the ITS analyses. Due to taxon sampling the polyphyly of subtribe Glueminae was not evident in the chloroplast analyses, but it has been demonstrated in previous chloroplast studies (Anderberg & Swenson 2003, Swenson & Anderberg 2005, Smedmark *et al* 2006). As such, Glueminae is clearly not a natural group and this finding corroborates that of Anderberg & Swenson (2003), who have formally transferred subtribe Glueminae out of the tribe Mimusopeae (= Sapoteae in their classification) and broken it up (Swenson & Anderberg 2005). Placement within the Sapotaceae and tribal association still remain to be resolved for some genera of the former Glueminae, particularly *Gluema* and *Neolemonniera*.

Subtribe Mimusopinae, which is comprised of taxa with 8-merous calyces, contains the genera *Vitellaria, Baillonella, Vitellariopsis, Tieghemella, Autranella* and *Mimusops*, and was not monophyletic in the present analysis. Instead it formed a grade, comprising three to four lineages subtending clade H, composed of the monophyletic subtribe Manilkarinae. In the ITS analysis, Mimusopinae subdivided into clades D (*Vitellaria, Baillonella, Vitellariopsis*), F (*Mimusops*) and G (*Tieghemella, Autranella*). The cpDNA analysis differed only in that Mimusopinae comprised four lineages rather than three because *Vitellaria* was not in clade D; however, the chloroplast dataset did not include the taxa *Baillonella* and *Tieghemella*. The paraphyly of subtribe Mimusopinae, comprising the three basal lineages within clade C,

suggests that the 8-merous calyx of the Mimusopinae is the ancestral state from which the 6-merous Manilkarinae calyx is derived.

With the exclusion of the genus *Northia*, which is placed in a basal grade in-between the Isonandreae and Sideroxyleae, the monophyly of subtribe Manilkarinae (clade H) is moderately to strongly supported in ITS (bs 81/pp 1), but poorly supported in the chloroplast analyses (bs < 50/pp 0.58).

Based on a family-level analysis of the Sapotaceae (including only 19 taxa classified in Pennington's tribe Mimusopeae) using the chloroplast gene *ndhF* and morphology, Swenson & Anderberg (2005) found the opposite pattern, i.e. that subtribe Mimusopinae was monophyletic and was, instead, subtended by a non-monophyletic subtribe Manilkarinae. However, support for this topology was poor, and when analysed solely with *ndhF* and without morphology, it was unsupported. Smedmark *et al*'s (2006) subsequent study of the subfamily Sapotoideae (including 31 taxa classified in Pennington's tribe Mimusopeae) with the chloroplast regions *ndhF, trnC-petN, petN-psbM, psbM-trnD* and *trnH-psbA* found that subtribes Mimusopinae and Manilkarinae were both monophyletic (except for the placement of *Autranella*) and sister to one another. The present study agrees more closely with these results, but differs in that subtribe Mimusopinae is clearly paraphyletic.

4.4.2 Delimitation of genera within the Manilkarinae

Although the phylogenetic delimitation of the subtribe Manilkarinae (the genera: *Manilkara*, *Labramia*, *Faucherea*, *Labourdonnaisia* and *Letestua*) is in agreement with the morphology-based classification, there is poor resolution along the backbone of the subtribe, particularly in the chloroplast trees. Therefore, the relationships between the main Manilkarinae clades remain ambiguous.

4.4.2.1 *Labramia*, *Labourdonnaisia* and *Faucherea*

Labramia, *Labourdonnaisia* and *Faucherea* are all small, morphologically similar genera, which are restricted to Madagascar and the Mascarenes. The monophyly of *Labramia* (clade I) was strongly supported in all analyses, as was the monophyly of clade K, which comprised both *Faucherea* and *Labourdonnaisia*. Between them, these genera contain two distinct lineages, but whether clades I and K are sister to each other, making the whole group monophyletic, is unclear. Such a relationship is weakly contradicted by the ITS Bayesian analysis but in the ITS parsimony and cpDNA analyses it is neither supported nor contradicted.

Within clade K, neither *Faucherea* nor *Labourdonnaisia* are monophyletic, therefore, their circumscription needs to be reconsidered. One possible solution would be to lump the two genera, as together they do form a strongly supported monophyletic group (clade K). However, *Labourdonnaisia* is distinct from *Faucherea* in having double the number of corolla lobes, stamens and ovary loculi.

It is also worth noting that there is incongruence between the ITS and cpDNA trees in the placement of *Faucherea manongarivensis*. In the ITS analysis, this species is resolved in a clade with *Labourdonnaisia calophylloides* and *L. revoluta*, whereas in the chloroplast tree, it is resolved in a clade with the rest of the *Faucherea* species and *L. madagascariensis,* pointing to a possible hybridization event. Geographical structure is also evident in the topology of the chloroplast phylogeny, with the Malagasy species of *Faucherea* and *Labourdonnaisia* (*L. calophylloides* & *L. revoluta*) resolved together in a weakly supported clade, while the Mascarene *Labourdonnaisia* species come out together in a strongly supported sister clade. However, in the ITS analyses, Malagasy and Mascarene species fall in a mixed clade. A broader sampling of all taxa in these genera should help to untangle possible hybridization events and resolve the conflict in classification, as reticulate evolution is a possible explanation for the failure of morphology to correspond to clade membership.

4.4.2.2 *Manilkara* and *Letestua*

As currently circumscribed *Manilkara* is not monophyletic, with three Asian species (*M. fasciculata*, *M. udoido* & *M. dissecta*) resolved in a separate clade (L) from all other *Manilkara* species in both the ITS and chloroplast analyses. The relationships of this clade are unclear; a sister relationship with clade K is very weakly supported in the Bayesian ITS analysis only.

In the ITS tree, all remaining species of *Manilkara* form a monophyletic clade (M), together with *Letestua*, although this topology is only strongly supported in the Bayesian analysis. In the cpDNA analysis, this clade is divided into two: clade $U2_C$ (*M. celebica, M. kanosiensis, M. smithiana, M. vitiensis, M. hoshinoi, M. napali,* and *M. roxburghiana*) falls outside of clade M_C, which contains all other species from ITS clade M. Therefore, *Manilkara* is biphyletic for ITS, but possibly triphyletic for cpDNA. However, the relative positions of clades M, I, K, and L (and for cpDNA only, $U2_C$), are unclear in any of these analyses. The monotypic genus *Letestua* was originally distinguished from *Manilkara* in having double the number of corolla lobes, stamens and ovary loculi, but lacking in staminodes. Here it was examined for ITS only, and based on this, it was nested within the African clade T of *Manilkara*. Smedmark *et al* (2006), using the chloroplast regions *ndhF, trnC-petN, petN-psbM, psbM-trnD* and *trnH-psbA*, also found it to be nested within *Manilkara*. A caveat in Smedmark *et al* (2006), noted that the sampled *Letestua* specimen was sterile and so its identification could not be confirmed. As this ITS sequence was from the same specimen, caution is advised and its placement is still somewhat questionable. *Letestua* is treated as belonging within *Manilkara* here, but fertile samples would be necessary to confirm its placement.

Within the large clade (M) of *Manilkara* there is strong geographical structure. In the ITS trees, *Manilkara* is comprised of five clades, of which clade U is Asian, clade N is Neotropical and clades T, R and S are all African. Clade T is the largest African clade and also includes Madagascan species. This geographic structure is also reflected in the chloroplast trees, although not as clearly because the positions of several species are not

resolved. These and other topological differences might reflect a lack of resolution. The biggest difference is that clade U is broken up, with one subclade ($U2_C$) falling outside clade $M_{C,}$ as noted above. This means that both *Manilkara* and Asian *Manilkara* contain one more distinct lineage for cpDNA than they do for ITS. Biogeography will be discussed in greater depth in Chapter VII.

4.4.3 Implications for classification – summary of findings

Based on molecular data, the tribe Mimusopeae as circumscribed by Pennington (1991) is not monophyletic. It should be re-circumscribed to exclude the subtribe Glueminae and the genus *Northia*, and should include the genera: *Manilkara*, *Faucherea*, *Labourdonnaisia*, *Labramia*, *Letestua*, *Mimusops*, *Baillonella*, *Tieghemella*, *Autranella*, *Vitellaria* and *Vitellariopsis.* This re-classification is roughly equivalent to Swenson & Anderberg's (2005) classification of their tribe Sapoteae, but differs significantly in that it excludes the genera *Northia*, *Madhuca*, *Payena*, and *Palaquium*. The revised classification presented here is in agreement with the findings of Smedmark *et al* (2006).

As the subtribe Mimusopinae has been shown not to be monophyletic there are two options for revising the taxonomy: a) abolish the subtribes Manilkarinae and Mimusopinae and only recognise the group at the tribal level of Sapoteae, or b) look for morphological characters which distinguish the subclades of the Mimusopinae enough to support their segregation as subtribes.

Within the Manilkarinae, the genera *Labourdonnaisia* and *Faucherea* were found not to be monophyletic. Broader sampling of all the taxa in these two small genera as well as an in-depth morphological study should be carried out before deciding whether to re-circumscribe the genera or lump them together.

Manilkara will also need to be re-circumscribed. A new genus should be created for the three Asian species *M. fasciculata, M. dissecta* and *M. udoido*, which is distinct from *Manilkara sensu stricto* in its coriaceous leaves with very tight venation; a character also found in *Faucherea*, *Labourdonnaisia* and *Labramia*. Additionally, if the results of this study and Smedmark *et al* (2006) can be confirmed using additional specimens, then *Letestua* should be sunk into *Manilkara*.

4.4.4 Delimitation of species complexes within *Manilkara*

The delimitation of some African *Manilkara* species is problematic due to a lack of clear diagnostic characters, with a number of taxa appearing to grade into one another. This is particularly true of the *Manilkara obovata* complex, which according to Plana (unpublished) "is highly variable with a number of modestly identifiable races, and which is almost certainly an example of either extensive hybridization or genetic plasticity and phenetic congruence among probably unrelated groups." A further possible explanation is that the defining characters of these races are all plesiomorphic, with none having evolved easily identified apomorphic traits. In Plana's taxonomic treatment, the *M. obovata* complex

includes: *M. obovata*, *M. mabokeensis*, *M. lacera*, *M. butugi* and *M. multinervis*. However, *M. fouilloyana, M. ferruginea M. sansibarensis,* and *M. welwitschii* are also described as being similar to *M. obovata* and easily confused when sterile.

Plana goes on to state, “As is apparent from its long and changeable synonymy, this highly variable species has been taxonomically difficult to evaluate. The *Manilkara obovata* tag has functioned as a lasting taxonomic trash bin and without more elaborate tools, such as the use of molecular markers, will unfortunately remain as such. Nevertheless, among the morass of eclectic specimens are distinct, sometimes large, populations which should not be ignored and lost in synonymy. In these groups there are well characterised cores of specimens with diagnostic characters, which unfortunately become blurred as a growing number of similar enough yet different specimens are introduced. These core populations do not merit taxonomic ranking for the simple reason that it would be impossible to define their morphological borders.”

The hypothesis that these taxa (*M. obovata*, *M. mabokeensis*, *M. lacera*, *M. butugi*, & *M. multinervis*) are part of a closely-related species complex was tested in the nuclear and chloroplast analyses presented in this chapter. Resolution between the species in both analyses suggests that they are all phylogenetically distinct entities. Analysis of additional genes may point towards hybridisation between lineages as a cause for morphological homogeneity, but this would need to be tested further. A better understanding of the taxonomic relationships of species in this group would require far more extensive sampling of material with multiple accessions per population to examine in detail how genetic, morphological and geographic differences correspond to one another.

4.4.5 Hard incongruence between nuclear and chloroplast trees – evidence for chloroplast capture?

Chloroplast capture, has frequently been cited as a potential cause for discrepancies between nuclear and chloroplast tree topologies (i.e. *Nothofagus* Acosta & Premoli 2010; *Thuja* Peng & Wang 2008; *Paeonia* Zhang *et al* 2009, Sang *et al* 1997; *Chrysophyllum* Swenson *et al* 2008; *Nesoluma* Smedmark & Anderberg 2007; *Gossypium* Wendel *et al* 1995; *Heuchera* Soltis *et al* 1991, Soltis & Kuzoff 1995; *Boykinia* Soltis *et al* 1996; *Mitella* Okuyama *et al* 2005; *Penstemon* Wolfe & Elisens 1995). It has also been implicated in incongruent chloroplast and morphology-based studies (i.e. *Rhododendron* subgenus *Hymenanthes* Milne *et al* 2010; *Macaranga* Banfer *et al* 2006; *Metrosideros* Percy *et al* 2008). According to Tsitrone *et al* (2003), conditions thought to be necessary for chloroplast capture may not be rare.

In addition to chloroplast capture, the introgression of nuclear regions, such as ITS, is also a potential explanation for gene tree incongruence as demonstrated in *Senecio* (Comes & Abbott 1999). Dengduangboripant *et al* (2007) show that although multicopy, ITS is inherited as a single homogenized unit. It is, therefore, possible for the F2 generation to segregate for the different parental copies (i.e. ITS capture can occur within two

generations). However, Petit & Excoffier (2009) suggest counter-intuitively that the more intraspecific gene flow that occurs, the less interspecific gene flow is expected and that markers associated with the most dispersing sex (i.e. biparentally inherited nrDNA in pollen) should better delimit species than markers associated with the less dispersing sex (i.e. maternally inherited cpDNA in seed). Therefore, for ITS, high within-species gene flow is expected along with low introgression. Conversely, chloroplast markers tend to show more inter-species gene flow and introgression. Additionally, when seed dispersal is particularly limited, cpDNA variation is more influenced by geography than taxonomy (Petit & Excoffier 2009). Given this scenario, chloroplast capture may be more likely than ITS capture.

4.4.5.1 Clades U and R: chloroplast capture across the Indian Ocean?

In this study, there is a hard incongruence between the topology of the nuclear and chloroplast trees, because the Asian species *M. hexandra* (Sri Lanka) and *M. littoralis* (Myanmar) (clade U1) are placed with other Asian species (clade U2) by ITS data, but with two African species *M. mochisia* (Zambia) and *M. concolor* (South Africa) (clade R) according to cpDNA. This suggests hybridization of taxa across the Indian Ocean possibly resulting in chloroplast capture. This putative relationship may be supported by morphology and ecology. *M. mochisia* and *M. concolor* can be distinguished from other African species in having relatively small glabrous leaves (which in *M. mochisia* are arranged in rosettes on short swollen shoots) and growing in East African dry, open woodland. Although *M. hexandra* is widespread from India to Thailand and can be morphologically variable, Sri Lankan populations also grow in dry, open woodland and exhibit similarly small, glabrous leaves, which are sometimes presented on short shoots like *M. mochisia*. However, this possible relationship would require further study.

If the chloroplast capture scenario is correct for these species, then hypotheses that would explain the incongruence include:

(i) dispersal of a clade R species from Africa to Asia, after which this species hybridises with an Asian lineage (ancestor of clade U1), donating its cpDNA and to this species, giving rise to *M. hexandra* and *M. littoralis,* but otherwise going extinct.

(ii) dispersal of a clade U1 species from Asia to Africa, after which this species hybridises with an African (clade R) lineage, capturing the chloroplasts of that lineage and giving rise to *M. mochisia* and *M. concolor*

Based on the phylogeny alone, the cpDNA lineage of clade U1 is probably monophyletic, whereas that of clade R may not be. Hence the African species appear paraphyletic with respect to the Asian species, and on this basis, hypothesis (i) seems the more likely of the two. Additionally, ancestral area reconstruction in Chapter VII suggests that *Manilkara* originated in Africa and then dispersed to Asia, a scenario which corroborates this hypothesis, because it demonstrates that dispersal has taken place in that direction previously.

4.4.5.2 Clade V: chloroplast capture across the Atlantic Ocean?

Intercontinental chloroplast capture may also be implicated in the case of clade V, which is resolved in the chloroplast analyses but not in the ITS analyses and is composed of *M. suarezensis* (Madagascar), *M. triflora* (Brazil) and *M. yangambensis* (Congo). The ITS analysis did not include *M. yangambensis,* but placed *M. triflora* with other Neotropical species in clade N, and *M. suarezensis* with other Madagascan species within a larger clade of African species (clade T). Therefore, ITS resolved at least two of the clade V species with species from the same landmass, but cpDNA did not, and resolved them together instead. Clade V is strongly supported (pp 0.99) in the Bayesian chloroplast analysis, but poorly supported (bs <50) in the parsimony analysis. However, the difference in strength of support between these analyses may be an artefact of sampling because *M. suarezensis* is only represented by *trnS-trnFM* in the chloroplast analysis, while *M. triflora* and *M. yangambensis* are represented by all three plastid regions.

Assuming that the correct species level relationships are resolved, clade V presents a case of long distance dispersal and chloroplast capture more remarkable than the clade R/U1 scenario, because it involves species from three landmasses, and hence two dispersal events. With this in mind, five hypotheses are possible:

(i) Africa to the Neotropics and Africa to Madagascar: Dispersal of the ancestor of *M. yangambensis* and *M. suarezensis* from Africa to the Neotropics, after which it hybridized with a member of the Neotropical lineage (clade N), donating its cpDNA to this species and giving rise to the ancestor of *M. triflora*. A second dispersal from Africa to Madagascar, without chloroplast capture, is also necessary to account for the Madagascan distribution of *M. suarezensis*.

(ii) Madagascar to Africa and subsequently Africa to the Neotropics: Dispersal of the ancestor of *M. yangambensis* and *M. suarezensis* from Madagascar to Africa, giving rise to *M. yangambensis*, which then dispersed to the Neotropics and hybridized with a member of the Neotropical lineage (clade N), donating its cpDNA to this species and giving rise to the ancestor of *M. triflora*.

(iii) Neotropics to Africa and subsequently Africa to Madagascar: Dispersal of the ancestor of *M. triflora* from the Neotropics to Africa, after which it hybridised with an African lineage (clade T), donating its cpDNA to this species, and giving rise to the ancestor of *M. yangambensis* in Congo, and *M. suarezensis,* which evolved after dispersal from Africa to Madagascar.

(iv) Madagascar to the Neotropics and subsequently Neotropics to Africa: Dispersal of the ancestor of *M. suarezensis* from Madagascar to the Neotropics, after which it hybridized with a member of the Neotropical lineage (clade N) donating its cpDNA to this species and giving rise to the ancestor of *M. triflora*, which then dispersed to Africa, giving rise to *M. yangambensis*.

(v) Neotropics to Madagascar and subsequently Madagascar to Africa: Dispersal of the ancestor of *M. triflora* from the Neotropics to Madagascar, after which it hybridized with a member of the Madagascan lineage (clade T) donating its cpDNA to this species and giving rise to the ancestor of *M. suarezensis*, which then dispersed to Africa, giving rise to *M. yangambensis*.

Out of these five scenarios, hypothesis (i) seems the most likely explanation, given that in Chapter VII Africa is reconstructed as the ancestral area for *Manilkara* with Oligo-Miocene dispersals from Africa to the Neotropics and to Madagascar. However, the fact that *M. suarezensis* and *M. triflora* have identical branch lengths and, thus, share the same number of substitutions is also consistent with hypotheses (iv) and (v). Better resolution and the inclusion of *M. yangambensis* in the ITS analysis would give a clearer picture.

4.4.5.3 Other possible instances of chloroplast capture

In the Bayesian cpDNA analysis, *Faucherea manongarviensis* is resolved within a clade of *Faucherea sp., F. thouvenotii*, *F. parvifolia* and *Labourdonnasia madagscariensis* (all of which are Malagasy species), whereas the Bayesian ITS analysis resolves the same clade excluding *F. manongarviensis*, which is instead grouped with the other two *Labourdonnasia* (Mascarene species), albeit with modest support. This could represent an instance of chloroplast capture between the Malagasy *F. manongarivensis* and the ancestor of the Mascarene *Labourdonnaisia calophylloides* and *L. revoluta.* Hybridisation between these species might explain why morphology does not fit with clade membership for the *Faucherea*/*Labourdonnasia* group.

4.4.5.4 Precedents for chloroplast capture across geographic boundaries

Although hybridization and chloroplast capture across long distances such as ocean barriers may seem unlikely, it has been suggested previously in numerous groups, including two taxa in the Sapotaceae. The species *Chrysophyllum cuneifolium* is inferred to have originated from an intercontinental hybridization event where the chloroplast is South American and the nuclear genome is African (Swenson *et al* 2008). Likewise, the Pacific genus *Nesoluma* is hypothesized to have arisen as a result of intercontinental hybridisation in the boreotropical region during the Eocene (Smedmark & Anderberg 2007). *Nesoluma* presents the opposite pattern to *Chrysophyllum*, where the chloroplast is African and the nuclear genome is Neotropical. Hybridization between New and Old World lineages has also been demonstrated in the pantropical genus *Gossypium* (Malvaceae) (Wendel *et al* 1995) and intercontinental chloroplast capture is hypothesized to have also occurred in the genus *Thuja* (Cupressaceae) where the eastern Asian species, *T. koraiensis*, may have obtained its chloroplast genome from the eastern North American species *T. occidentalis* (Peng & Wang 2008).

Investigating the exact mechanisms for these examples of chloroplast and nuclear topological incongruence would require more in-depth study involving cloning and the sequencing of additional nuclear and chloroplast regions. Intercontinental dispersal and biogeography are discussed in more depth in Chapters V and VII.

4.5 Appendices

Appendix 4.1 Specimen data for taxa included in the ITS and cpDNA analyses. A ✓ indicates that sequence data for this accession was included in the analyses. A – indicates that sequence data for this accession was unavailable.

Species	DNA accession number	Collector's number	Country of origin	ITS	*rpl*32-*trn*L	*rps*16-*trn*K	*trn*S-*trn*FM
Argania spinosa	AA	Nordenstam 9325	Morocco	✓	-	-	-
Autranella congolensis	AA	Bokdam 4401	Congo	✓	✓	✓	✓
Baillonella toxisperma	EDNA09_01453	Bourobou *s.n.*	Gabon	✓	-	-	-
Eberhardtia aurata	AA	G. Hao 534 Cultivated: S. China Bot. Gard.	Vietnam – SE China	✓	✓	✓	✓
Eberhardtia tonkinensis	Gen Bank AF456258	Yang,S.-X. unpublished	Yunnan, China	✓	-	-	-
Faucherea manongarivensis	EDNA06_05896	L. Gautier *et al* 3910	Madagascar	✓	✓	✓	✓
Faucherea parvifolia	EDNA06_05889	L. Gautier 163	Madagascar	✓	✓	✓	✓
Faucherea sp.	EDNA07_01933	A. Anderberg 233	Madagascar	✓	✓	✓	✓
Faucherea thouvenotii	EDNA06_05897	L. Gautier 3938	Madagascar	✓	✓	✓	✓
Inhambanella henriquesii	AA	de Winter & Vahrmeijer 8536	South Africa	✓	✓	✓	✓
Isonandra compta	AA	Emanuelsson 3039	Sri Lanka	✓	-	-	-
Labourdonnaisia callophylloides	EDNA08_02271	R. Capuron 28171SF	Reunion	✓	✓	✓	✓
Labourdonnaisia madagascariensis	EDNA07_02212	R. Capuron 27747SF	Madagascar	✓	✓	✓	✓
Labourdonnaisia revoluta	EDNA07_02271	Lorence 1602	Mauritius	✓	✓	✓	✓
Labramia anakaranaensis	EDNA06_05884	L. Gautier 4037	Madagascar	✓	✓	✓	✓
Labramia costata	EDNA07_02272	G. Schatz & A. Gentry 2094	Madagascar	✓	✓	✓	✓
Labramia louvelii	EDNA07_01927	A. Anderberg 245	Madagascar	✓	✓	✓	✓
Labramia mayottensis	AA	Labat *et al* 3309	Mayotte, Comores	✓	✓	✓	✓
Lecomtedoxa klaineana	AA	Veldhuizen 1509 Cultivated: Holland	Cameroon – Gabon	✓	-	-	-
Letestua durissima	AA	Normand *s.n.*	Congo	✓	-	-	-
Manilkara bella	EDNA08_02267	Folli 501	Brazil	✓	✓	✓	✓
Manilkara bequaertii	EDNA07_02081	F. Breteler 15348	Gabon	✓	✓	✓	✓
Manilkara bidentata	EDNA06_05887	T. Pennington 1203	Peru	✓	✓	✓	✓
Manilkara boivinii	EDNA06_05905	L. Gautier 3278	Madagascar	✓	✓	✓	✓
Manilkara butugi	EDNA06_05901	D.R. Chaffey 1252	Ethiopia	✓	✓	✓	✓
Manilkara capuronii	EDNA07_02079	R. Capuron 11.377SF	Madagascar	✓	✓	✓	✓
Manilkara cavalcantei	EDNA07_02205	Vicentini *et al* 527	Brazil	✓	✓	✓	✓
Manilkara celebica	EDNA08_02339	Neth. Ind. For. Service bb 30152	Sulawesi, Indonesia	-	✓	✓	-
Manilkara chicle	AA	Castillo *et al* 2083	Guatemala	✓	✓	✓	✓
Manilkara concolor	AA	Swenson & Karis 635	South Africa	✓	✓	✓	✓
Manilkara cuneifolia	EDNA07_02264	G. McPherson 16792	Gabon	✓	✓	✓	✓
Manilkara dawei	EDNA07_01928	D.J. Harris 7707	Central African Republic	✓	✓	✓	✓
Manilkara decrescens	EDNA08_02268	J.D. & E.G. Chapman 6689	Malawi	✓	✓	✓	✓

Species	DNA accession number	Collector's number	Country of origin	ITS	*rpl*32-*trn*L	*rps*16-*trn*K	*trn*S-*trn*FM
Manilkara discolor	EDNA06_05892	K. Vollesen 2460	Tanzania	✓	✓	✓	✓
Manilkara dissecta	EDNA06_05883	M. Gardner TNCA 4012	New Caledonia	✓	✓	✓	✓
Manilkara elata	EDNA08_02265	Jardin *et al* 2277	Brazil	✓	✓	✓	✓
Manilkara fasciculata	EDNA08_02258	K. Armstrong 353	West Papua, Indonesia	✓	✓	✓	✓
Manilkara fouilloyana	EDNA07_02267	G. McPherson 16173	Gabon	✓	✓	✓	✓
Manilkara gonavensis	EDNA08_02264	Ekman 8741	Haiti	✓	✓	✓	✓
Manilkara hexandra	EDNA07_02053	P.L. Comanor 868	Sri Lanka	✓	✓	✓	✓
Manilkara hoshinoi	EDNA08_02340	M. Hoshino 2138	Pulau	✓	✓	✓	✓
Manilkara huberi	EDNA07_01926	O. Poncy 1828	French Guiana	✓	✓	✓	✓
Manilkara inundata	EDNA07_02093	Sothers & Saraiva 22	Brazil	✓	✓	✓	✓
Manilkara jamiqui	EDNA07_02201	Urquiola & Dressler 529	Cuba	✓	✓	✓	✓
Manilkara kanosiensis	EDNA08_02342	Neth. Ind. For. Service bb. 24311	Papua New Guinea	-	✓	✓	-
Manilkara kauki	EDNA08_02260	K. Armstrong 379	Bali, Indonesia	✓	✓	✓	✓
Manilkara koechlinii	EDNA06_05893	J. Casier 443	Democratic Republic of Congo	✓	✓	✓	✓
Manilkara lacera (a)	EDNA07_01095b	D.J. Harris 8200A	Gabon	✓	✓	✓	✓
Manilkara lacera (b)	EDNA06_05894	X.M. van de Burgt 40	Gabon	✓	✓	✓	✓
Manilkara letestui	EDNA07_02080	J.J. Bos 5604	Cameroon	✓	-	-	✓
Manilkara letouzeyi	EDNA08_02338	R. Letouzey 4444	Cameroon	✓	✓	✓	✓
Manilkara littoralis	EDNA07_02052	Maung Gale 14654	Myanmar	✓	✓	✓	✓
Manilkara longifolia	EDNA07_02092	Thomas *et al* 8076	Brazil	✓	✓	✓	✓
Manilkara lososiana	EDNA07_02088	D. Kenfack 625	Cameroon	✓	✓	✓	✓
Manilkara mabokeensis	EDNA07_01094b	D.J. Harris 7164	Central African Republic	✓	✓	✓	✓
Manilkara maxima	EDNA07_02091	Sant'Ana *et al* 670	Brazil	✓	✓	✓	✓
Manilkara mayarensis	AA	Ekman 9971	Cuba	✓	✓	✓	✓
Manilkara mochisia (a)	EDNA06_05888	L. Gautier 4171	Zambia	✓	✓	✓	✓
Manilkara mochisia (b)	EDNA07_02269	Bidgood *et al* 2286	Tanzania	-	✓	✓	✓
Manilkara napali	EDNA08_02341	F. Schram BW 1636	West Papua, Indonesia	-	-	✓	✓
Manilkara obovata (butugi-type)	EDNA07_02262	Friis & Vollesen 740	Sudan	✓	✓	✓	✓
Manilkara obovata (multinervis-type)	EDNA07_02263	Schmidt *et al* 3274	Ghana	✓	✓	✓	✓
Manilkara obovata (obovata-type) (a)	EDNA07_01930	D.J. Harris 7759	Central African Republic	✓	✓	✓	✓
Manilkara obovata (obovata-type) (b)	EDNA08_02261	GAF Malanda 7	Democratic Republic of Congo	✓	✓	✓	✓
Manilkara paraensis	EDNA07_02206	Zarucchi *et al* 2526	Brazil	✓	✓	✓	✓
Manilkara pellegriniana	EDNA07_02087	D.J. Harris & M. Fay 1843	Cameroon	✓	✓	✓	✓
Manilkara perrieri	EDNA07_02082	R. Capuron 28132-SF	Madagascar	✓	✓	✓	✓
Manilkara pleena	EDNA07_02209	A. Lioger & P. Lioger 33453	Puerto Rico	✓	✓	✓	✓
Manilkara pubicarpa	EDNA08_02262	Forest Dept. British Guyana 5860	Guyana	-	✓	✓	✓
Manilkara roxburghiana	EDNA07_02199	Matthew & Rajendren 44790	India	-	✓	-	✓
Manilkara rufula	EDNA07_02208	G. Ignacio & A. Caurenio 37	Brazil	✓	✓	✓	✓
Manilkara sahafarensis	EDNA07_02085	R. Capuron 20.965-SF	Madagascar	✓	✓	✓	✓

Species	DNA accession number	Collector's number	Country of origin	ITS	*rpl*32-*trn*L	*rps*16-*trn*K	*trn*S-*trn*FM
Manilkara salzmanii	EDNA07_02207	Jardim *et al* 2277	Brazil	✓	✓	✓	✓
Manilkara samoensis	EDNA08_02263	S.J. Whilmee 226	Samoa	-	✓	-	✓
Manilkara sansibarensis	EDNA07_01083c	Abeid 272	Tanzania	✓	✓	✓	✓
Manilkara sideroxylon	EDNA07_02203	Ekman 16173	Cuba	✓	✓	✓	✓
Manilkara smithiana	EDNA07_02057	A.C. Smith 1450	Fiji	✓	✓	✓	✓
Manilkara sp. 1	EDNA07-02260	P. Sita 4107	Congo	✓	-	-	-
Manilkara staminodella	EDNA07_02204	Anderberg *et al* 50	Costa Rica	✓	✓	✓	✓
Manilkara suarezensis	EDNA07_02259	Randriamampionona 248	Madagascar	✓	-	-	✓
Manilkara subsericea	EDNA07_02202	Hatschbach & Souza 51302	Brazil	✓	✓	✓	✓
Manilkara sulcata	EDNA07_02086	Frontier-Tanzania Coastal Forest Research Programme 1045	Tanzania	✓	-	-	✓
Manilkara triflora	EDNA08_02343	Fonseca *et al* 2887	Brazil	✓	✓	✓	✓
Manilkara udoido	EDNA07_02058	S. Slappy LR26622	Palau	✓	✓	✓	✓
Manilkara valenzuelana	EDNA07_02211	A. Lioger & P. Lioger 22980	Dominican Republic	✓	✓	✓	✓
Manilkara vitiensis	EDNA08_02345	Smith 1461	Fiji	✓	✓	✓	✓
Manilkara welwitschii	EDNA06_05891	J.J.F.E. de Wilde & R.W. de Wilde-Bakhuizen 11385	Gabon	✓	✓	✓	✓
Manilkara yangambensis	EDNA08_02344	C. Evrard 1499	Democratic Republic of Congo	-	✓	✓	✓
Manilkara zapota	EDNA06_05886	J. Clayton 12	Trinidad	✓	✓	✓	✓
Manilkara zenkeri	EDNA07_02084	Doumenge 526	Cameroon	✓	-	✓	-
Mimusops caffra	AA	Swenson & Karis 636	South Africa	✓	✓	✓	✓
Mimusops comorensis	AA	Pignal & Ginguette 1065	Comoros Islands	✓	✓	✓	✓
Mimusops coriacea	Geneva	Bernadi 11891	Madagascar	✓	-	-	-
Mimusops elengi	AA	Chantaranothai 2305	Thailand	✓	✓	✓	✓
Mimusops kummel	Geneva	Kayambo 4996	Tanzania	✓	-	-	-
Mimusops membranacea	Geneva	Randrianaivo 126	Madagascar	✓	-	-	-
Mimusops obovata	AA	Swenson & Karis 633	South Africa	✓	✓	✓	✓
Mimusops perrieri	Geneva	S.F. 18297	Madagascar	✓	-	-	-
Mimusops sp. (voalala complex)	Geneva	Randrianaivo 583	Madagascar	✓	-	-	-
Mimusops zeyheri	AA	Dahlstrand 6386	South Africa	✓	✓	✓	✓
Northia seychellana	AA	L. Chong-Seng *s.n.*	Seychelles	✓	✓	✓	✓
Palaquium amboinense	AA	Iuijesundara *s.n.*	Sri Lanka	✓	-	-	-
Sarcosperma laurinum	AA	Saunders 2000	Hong Kong	✓	✓	✓	✓
Sideroxylon ibarre	AA	Lundell 19752	Guatemala	✓	-	-	-
Vitellaria paradoxa	AA	Neumann 1512	Benin	✓	✓	✓	✓
Tieghemella heckelii	AA	Jongkind 3936	Ghana	✓	-	-	-
Vitellariopsis cuneata	AA	Thomas 3662	Tanzania	✓	✓	✓	✓
Vitellariopsis dispar	AA	Pentz 2	South Africa	✓	✓	✓	✓
Vitellariopsis kirkii	AA	Robertson 4085	Kenya	✓	✓	✓	✓
Vitellariopsis marginata	AA	Chase 1122	South Africa	✓	✓	✓	✓
Xantolis siamensis	AA	Smitairi 1	Thailand	✓	-	-	-

Chapter V – Historical biogeography of tropical forests and their intercontinental disjunctions

5.1 Overview of regional biogeography and trends in angiosperm disjunctions in the Cenozoic

The processes which generate intercontinental disjunctions can be complex to decipher. Without knowledge of a temporal component, Gondwanan break-up, boreotropical migration and long distance dispersal can exhibit similar patterns. Often evidence of older vicariance events is overlain with more recent patterns of dispersal and radiation, which can confound attempts to reconstruct the historical biogeography of a taxon (e.g. Lavin *et al* 2004).

Biogeographic thought has undergone dramatic paradigm shifts over the past forty years. Prior to the validation of plate tectonic theory in the late 1960's, dispersal was championed as the biogeographer's main method of explaining disjunctions. However, once it became accepted that the earth's crust was not static and continental plates (Fig. 5.1) were in constant motion, vicariance via plate tectonic movement became the favoured scenario, while long distance dispersal was deemed to be a rare occurrence. The main objection to dispersal has been that it is unfalsifiable and therefore unscientific, yet if vicariance hypotheses are falsified based on age estimates, then by default, dispersal becomes a plausible explanation (de Queiroz, 2005). As such, timing is an important factor in determining which dispersal/vicariance pathways were available when. The reconstruction of biogeographic history has been attempted through studies of the fossil record and through cladistic morphological approaches. However, the fossil record is fragmentary and, therefore, cannot be relied upon to give the complete picture, and cladistic morphological approaches cannot be relied upon to accurately reconstruct the true phylogeny. Dated molecular phylogenies are a comparatively new tool, which have contributed greatly to historical biogeographic studies by allowing us to distinguish between competing biogeographic hypotheses (Pennington *et al* 2004). Different hypotheses imply different phylogenetic tree topologies with lineage splits occurring at nodes at particular ages (Donoghue & Moore 2003).

Plant disjunctions (particularly pantropical) have often been attributed to Gondwanan break-up (Raven & Axelrod 1974), but current studies have shown that many tropical groups are of more recent origin. Dated phylogenies indicate that while many family level disjunctions are the result of Gondwanan breakup, degradation of the boreotropical flora or other deep-time vicariance events, splits between genera have commonly been found to coincide with more recent long distance dispersal. Furthermore, species level disjunctions can sometimes be anthropogenic (Renner 2004).

Disjunctions caused by West Gondwanan break-up would be reflected in major phylogenetic splits between ~110 and 70 Ma. For boreotropical migration, divergence times between ~65-45 Ma would be expected. While long-distance dispersal could have occurred at any point in time, it is the only viable scenario for tropical disjunctions younger than ~33 Ma. In Chapter VI, these and other biogeographic hypotheses concerning the age of *Manilkara* and its origin will be tested. To provide some background, the history of these vicariance

and dispersal pathways are explained in more detail below. An overview of the geological and climatological history of the three main tropical continental regions (Neotropics, Africa and Southeast Asia) is also presented here so that a correlation can be made between phylogenetic patterns and the historical processes which may have shaped them.

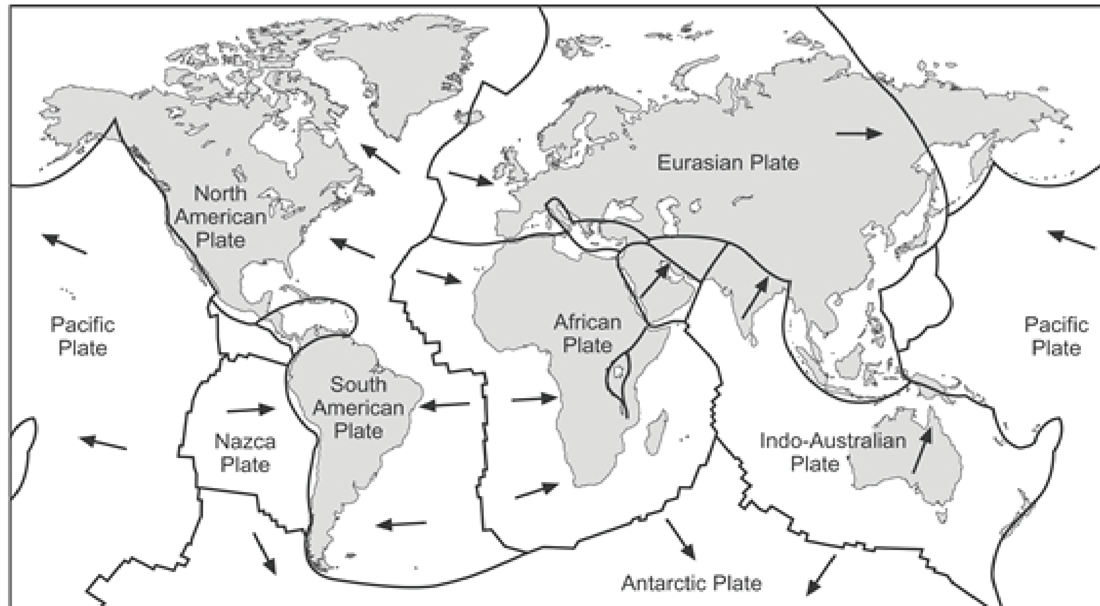


Figure 5.1 Overview of tectonic plates and their direction of movement from: http://www.esta-uk.net/jesei/platerid/plates.htm

5.2 Origin of eudicots

As evidenced by the occurrence of triaperturate pollen, early eudicots began to diversify during the Cretaceous (113-108 Ma) in West Gondwana (Gabon and Brazil) (Hickey & Doyle 1977, Doyle *et al* 1977) in a very warm, subhumid climate. They later (108-96 Ma) dispersed poleward into subtropical (but still megathermal - frost intolerant) climate zones and by 96-90 Ma they dominated the vegetation of most regions with major centers of radiation in the northern mid latitudes (Boreotropical province), southern mid latitudes (Gondwanan megathermal province) and the equatorial region (Morley 2000, 2003). Subsequent vicariance and dispersal within and between these regions occurred from this time onwards.

5.3 Gondwanan vicariance

The history of Gondwanan break-up is important in biogeography because it coincides with a major phase of angiosperm evolution and radiation ~130-90 Ma, (Crane *et al.*, 1995), marked by a wide range of extant families and genera first appearing in the fossil record (Tiffney, 1985a). Gondwana was a southern hemisphere supercontinent composed of what is now South America, Africa, Antarctica, India, Madagascar, New Guinea, Australia, New Zealand and New Caledonia. Vicariance between its component blocks has been cited to explain tropical intercontinental disjunctions, particularly between the southern continental extremities of South America, South Africa, Australia and New Zealand (Upchurch 2008, McLoughlin 2001). This "classic" Gondwanan distribution pattern is exhibited by *Nothofagus*, *Gunnera*, Proteaceae, Myrtaceae and Winteraceae (Table 5.1).

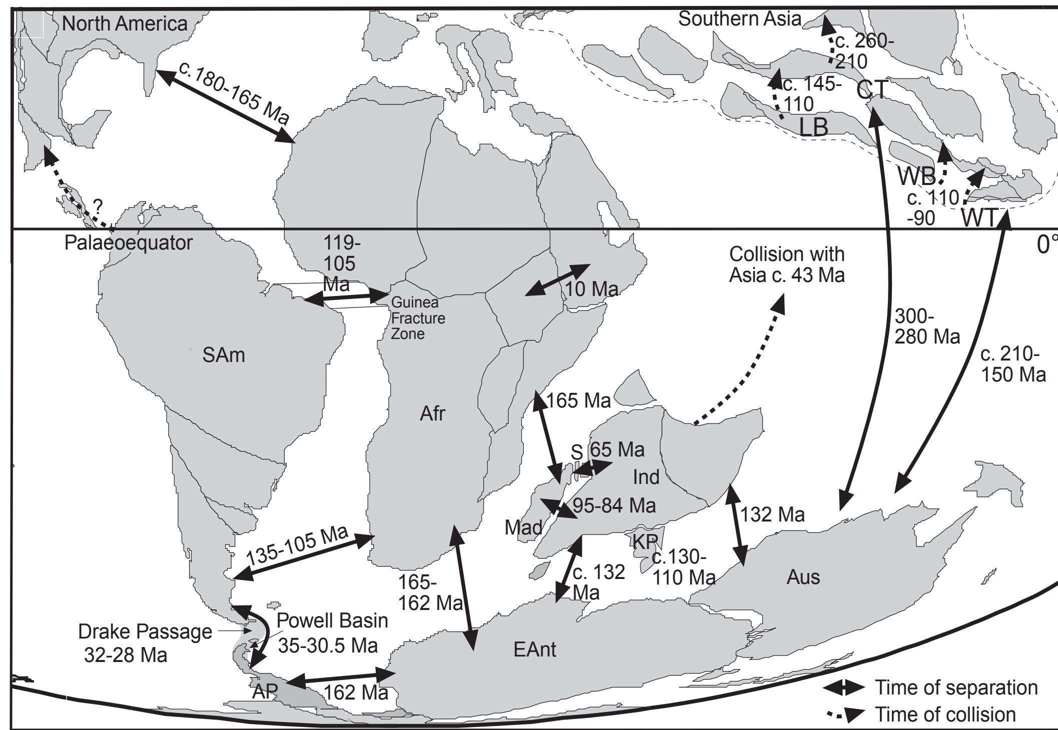


Figure 5.2 Break-up of Gondwana showing vicariance and collision times between continental fragments. Excerpted from McLoughlin 2001.

The initial fragmentation of Gondwana began during the Early Jurassic ~180 Ma, followed by intra-continental rifting and sea floor spreading ~165 Ma, which initiated the separation of Africa and Antarctica in the south and Madagascar in the east (Fig. 5.2). Madagascar then drifted to the southeast and its position with respect to Africa has not changed significantly since (McLoughlin 2001, Sanmartin 2002). Further sea-floor spreading in the South Atlantic about 135–130 Ma began to divide Africa and South America, although connections may have been maintained until ~119-105 Ma between Brazil and equatorial Africa as well as southern Africa and the Falklands Plateau. Volcanic islands and mid-Atlantic ridges are hypothesized to have facilitated dispersal between Africa and South America until ~95 Ma (Morley 2000). Following the fragmentation of West Gondwana, Africa was isolated until its collision with Eurasia beginning ~63 Ma (Sanmartin, 2002), which closed the Tethys sea by 14 Ma. Approximately 95-84 Ma the Indian Ocean began to open, rifting the Seychelles/India away from Madagascar/Africa. During their northward migration, the Seychelles became fixed with respect to Africa as India drifted towards Laurasia, colliding with the Asian plate ~43 Ma (Sanmartin 2002, McLoughlin 2001) (Fig. 5.2). Although a recent study suggests that the actual collision between India and Laurasia was more recent, beginning at ~35 Ma (Ali & Aitchison 2008).

Rifting between east Antarctica and Australia commenced ~96 Ma, but the continents remained in contact via Tasmania until ~64 Ma (McLoughlin 2001). However, it is likely that stepping stone dispersal was possible until ~52 Ma when the South Tasman Sea began to form between Australia and Antarctica. Fully marine conditions inhibited dispersal by the Late Eocene ~35 Ma. Paleocene ~60 Ma spreading in the Tasman Sea created the New Zealand-New Caledonia split from Australia and from this time through the mid-Oligocene

New Zealand and New Caledonia were submerged. Australia began to collide with the Philippine Plate during the Oligocene ~25 Ma (Hall, 1998) and is still migrating towards continental Southeast Asia (the Sunda shelf) today.

Australia and South America remained in contact across Antarctica until at least the Early Eocene (Sanmartin 2002) and, with its warm temperate climate, this region may have acted as an important migration corridor for Gondwanan angiosperms such as *Nothofagus* (Nothofagaceae) as well as *Ilex* (Aquifoliaceae), *Gunnera* (Gunneraceae), *Ascarina* (Chloranthaceae) and members of the Proteaceae, Winteraceae and Myrtaceae as evidenced by fossil pollen (Dettmann 1989). Tectonic movements had a profound effect on oceanic circulation, which in turn altered global climatic patterns, and by the late Eocene ~46 Ma a cooling climate would have limited potential overland migration. By ~30 Ma western Antarctica and southern South America had separated with the opening of the Drake Passage and the south circumpolar current was established, which enabled cold water to circulate in the southern hemisphere and initiated the onset of glaciation in Antarctica. Ice sheets became more extensive during the Pliocene following the closure of the Isthmus of Panama and the narrowing of the Indonesian flow-through as the Australian plate converged on Southeast Asia (Sanmartin 2002).

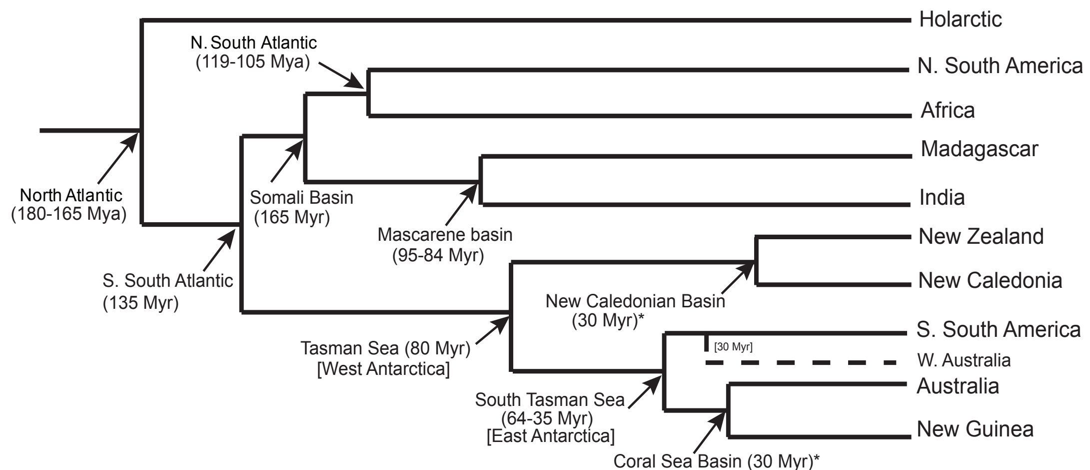


Figure 5.3 Geological area cladogram adapted from Sanmartin & Ronquist 2004 representing relationships between Gondwanan fragments and time of phylogenetic splits based on paleogeographic data. Vicariance is assumed to be at primary fragmentation. Asterisks (*) mark nodes, which are dated at 70-60 Ma in alternative reconstructions.

There is some debate (Upchurch 2008) over the pattern, which a Gondwanan disjunction would exhibit depending upon the preferred break-up scenario, i.e. whether Africa or East Gondwana (Antarctica, India, Madagascar and Australia) split from South America first. This argument, however, would only be relevant for taxa which are significantly older than ~80 Ma. The chronogram in Figure 5.3 depicts times of lineage splits which would be concordant with Gondwanan vicariance. Dated phylogenies of taxa, which have been shown to conform to a Gondwanan break-up pattern are presented in Table 5.1.

Table 5.1 Examples of (predominantly dated) phylogenetic studies of taxa which conform to a Gondwanan vicariance scenario. LDD = long distance dispersal.

Family/Taxon	Inferred Biogeographical History	Reference
Restionaceae	Gondwanan vicariance plus recent LDD	Linder *et al* 2003
Proteaceae	Gondwanan vicariance and trans-oceanic dispersal	Barker *et al* 2007
Nothofagaceae	Gondwanan vicariance followed by recent LDD between Australia & New Zealand	Knapp *et al* 2005, Cook & Crisp 2005
Winteraceae	Gondwanan vicariance	Doyle 2000
Myrtaceae eucalypt & melaleuca groups	Gondwanan vicariance and subsequent LDD	Ladiges *et al* 2003
Hernandiaceae	West Gondwanan vicariance (122 Ma split between the predominantly African–Madagascan–Malesian lineage of *Hazomalania*, *Hernandia* and *Illigera*, and an African–Neotropical lineage comprising *Gyrocarpus* and *Sparattanthelium*) followed by LDD	Michalak *et al* 2010
Monimiaceae	East Gondwanan vicariance followed by Oligo-Miocene LDD: trans-Pacific dispersal from Australasia to South America in *Mollinedia* (28–16 Ma), plus over-water dispersal in other clades between Australia, New Caledonia, New Zealand & across the Indian Ocean to Madagascar (20-29 Ma)	Renner *et al* 2010
Atherospermataceae	Gondwanan origin & diversification 100-140 Ma followed by LDD	Renner *et al* 2000
Annonaceae	Cretaceous Gondwanan origin & vicariance between South America and Africa	Richardson *et al* 2004
Sapotaceae *Chrysophyllum*	Hypothesized Gondwanan migration between South America and Australasia via Antarctica ~ 60-65 Ma	Bartish *et al* in press
Asteraceae Vernonieae	Origin in Gondwana/South Africa/Madagascar, followed by LDD	Keeley *et al* 2007
Gunneraceae *Gunnera*	Hypothesized Gondwanan vicariance backed up by Antarctic fossils - but phylogeny not dated	Wanntorp *et al* 2003

5.4 Regional history of the boreotropics

From the Early Paleocene (65 Ma) global temperatures began to increase and by the Early Eocene (55 Ma) they underwent a ~5-10° C rise over a period of ~10-20 thousand years (Wing *et al* 2005). This dramatic climatic warming event, the Paleocene-Eocene Thermal Maximum (PETM), enabled a megathermal flora to flourish in the northern hemisphere up to 45°-60° N latitude, across what is now North America, Greenland and Eurasia, with the main limiting factor being inadequate winter daylight (Tiffney 1985a, Tiffney 1985b, Lavin & Luckow 1993, Manchester & Tiffney 2001, Morley 2001, Harrington *et al* 2004). The presence of fossils of modern-day tropical taxa occurring in the northern latitudes, inspired Wolfe (1975) to coin the term "boreotropics" to describe this paleoregion. Dispersal pathways were available through the North Atlantic Land Bridge, which connected North America with Eurasia via Greenland (between ~50-25 Ma) and also the Bering Land Bridge to a lesser extent (on and off until ~5 Ma). The timing of dispersal events is indicated by the appearance of similar pollen formations in both regions (Tiffney 1985b, Manchester 1999), including that of Sapotaceae (Morley 2000, Tiffney & Manchester 2001). Some modern relatives of this flora can now be found in the Neotropics and Asia as well as Africa to a much lesser extent.

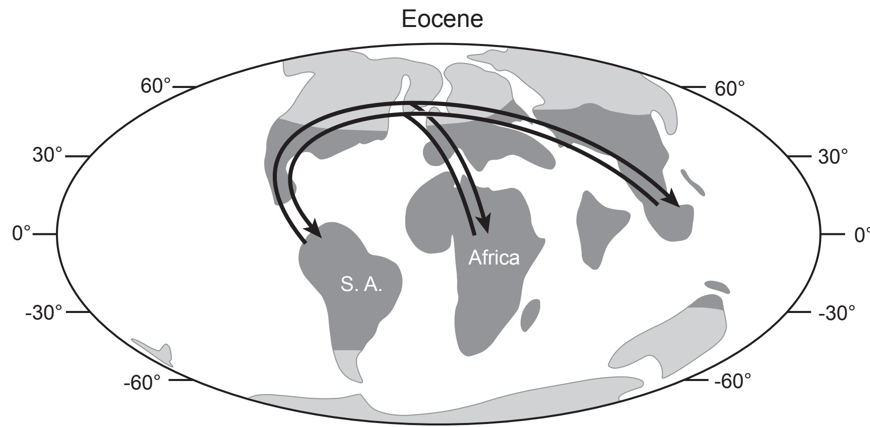


Figure 5.4 Depiction of boreotropical migration pathways available during the Eocene Thermal Maximum. Excerpted from Donoghue (2008).

At its greatest extent the boreotropical flora extended from Western North America across Europe to East Asia (Fig. 5.4), although it was probably not homogenous. Fossil assemblages suggest that although the localities share common elements, the boreotropical flora was regionalized, with different formations preserving different levels of diversity and was originally divided into Western North America/Asia and Eastern North America/Europe components by the Cannonball sea and the Turgai straits respectively (Tiffney & Manchester 2001, Milne 2006). In a comparison of thermophilic Middle Eocene (44 Ma) floras, Collinson & Hooker (2003) demonstrate that the Clarno flora of western North America shares 24% of genera with European floras (20% of genera and 10% species with the London Clay flora alone), clearly indicating an interchange between these regions prior to the Middle Eocene. Many of these important fossil assemblages trace the outline of the former Tethys Seaway: the Clarno nut flora of Oregon (Scott 1954; Manchester 1981, 1994, Wheeler & Manchester 2002), the Brandon Lignite of Vermont (Perkins 1905, Traverse 1953, 1955, 1994; Tiffney 1994), the London Clay flora of England (Reid & Chandler 1933), the Geiseltal flora of Germany (Mai 1976, Mai 1995), the Haselbach Flora of Germany (Mai & Walther 1978), the Burgas Flora of Bulgaria (Palamarev 1973), as well as formations in Bohemia (Knobloch & Konzalova 1998), Egypt (Chandler 1954), Kazakhstan (Makulbekov 1987, Shilin 2000) and Pakistan (Fredericksen 1994, Vimal 1952). According to Tiffney (1985a) the island chains of the Tethys may have been sources for the evolution of new taxa through allopatric speciation; spanning from eastern Asia to Europe and west as far as the Caribbean. These islands may have provided a route for the migration of taxa which evolved along the seaway. Additionally, modern refugia rich in boreotropical elements such as Central America, Southeastern North America, the Caucasus, the Himalayas and Southeast Asia are located on the margins of the paleo-Tethys.

The PETM warming trend culminated at ~52 Ma (Zachos 2001) and the climatic cooling which occurred after this time precipitated a major vegetational shift, whereby megathermal taxa became restricted to refugial sites at lower latitudes, while taxa suited to seasonal climates became more widespread (Tiffney & Manchester 2001). Global climates gradually warmed again for a period during the Miocene leading to the expansion of megathermal lineages which had retreated to refugial pockets, suggesting that many taxa had only gone

regionally extinct during the Oligocene. However, this warm phase was short-lived and cooler than the Eocene thermal maximum. From the mid-Miocene to the Pleistocene, global climates continued to cool, leading to the modernization of the northern hemisphere flora, with the spread of deciduous trees and herbs, while megathermal elements went extinct or retreated to equatorial regions (Tiffney & Manchester 2001).

Many of these megathermal taxa retreated southwards, finding refuge in East & Southeast Asia, Central America/Caribbean and to a much lesser extent in Africa. Morley (2001) suggests that the relative representation of boreotropical elements in Asia, Africa and the Neotropics reflects the different opportunities for southward dispersal to the tropical regions during the Late Tertiary. The European boreotropical flora would have encountered numerous barriers to its southern migration including the uplift of the Alps, the Mediterranean Sea and the Sahara desert, explaining why there is a paucity of boreotropical elements found in African rain forests today (Morley 2001) – only eight genera are represented in the contemporary African flora according to Tiffney (1985b). Boreotropical elements in the Americas probably took refuge along the southern margin of the North American plate, but until the Isthmus of Panama was formed during the late Miocene, would not have had a direct land connection to equatorial latitudes in South America. However, Cody *et al* (2010) indicate plant dispersal between South and Central America prior to the formation of the land bridge. Twenty two genera with boreotropical affinities are represented in the present-day Neotropical flora (Morley 2001). East and Southeast Asia harbour the largest number of boreotropical genera (34 according to Tiffney1985b & Morley 2001) and have long been recognised as important refugial areas for megathermal angiosperms. The continuous land connection between the northern latitudes and the equatorial zone in Asia has enabled megathermal taxa to migrate towards the tropics unhindered (Tiffney 1985a, b).

Tiffney (1985a) suggests that post-Eocene extinctions were probably more widespread in the North American boreotropics than in East Asia, because the lower topography of North America affords a more limited range of habitats than in East Asia. Additionally, the north-south orientation of the Appalachian mountains and the Mississippi river valley create a funnel for arctic air towards the Caribbean, whereas, in Asia the mountain ranges act as a barrier to cold air masses, yet have sufficient gaps for megathermal angiosperms to migrate southwards. Finally, although the two regions have similar mean monthly temperatures, the absolute minimum temperatures are lower in North America.

Based on dated molecular phylogenies and fossil evidence, many taxa which are now restricted to the tropics are believed to have inhabited the northern hemisphere during the PETM, i.e. Sideroxyleae in the Sapotaceae (Smedmark & Anderberg 2007), Meliaceae (Muellner *et al* 2010), Malpighiaceae (Davis *et al* 2002 a, b, 2004), Leguminosae (Lavin *et al* 2001), Annonaceae (Richardson *et al* 2004), Simaroubaceae (Clayton *et al* 2009) and Lauraceae (Chanderbali *et al* 2001). Details of these taxa and others are listed in Table 5.2. However, Collinson & Hooker (2003) caution that although the taxa are similar, the boreotropical flora does not equate to modern tropical forest, the make-up of which is different.

It is also important to note that while strictly megathermal (tropical, frost-intolerant) taxa would have only been able to exploit the boreotropical migration pathway from ~65-45 Ma, subtropical (mesothermal) and temperate (microthermal) taxa continued to thrive in the northern hemisphere well after the decline of the PETM when temperatures began to cool (Zachos *et al* 2002). Therefore, rather than there being an exact migration cut-off point following the PETM, there is more of a gradual progression towards cooler and drier climates which limited the distribution of tropical taxa, while still enabling the migration of subtropical taxa. Those microthermal taxa, which persisted in the northern hemisphere well into the Oligocene and Miocene (some until 10-5 Ma) are commonly referred to as being components of the "Tertiary relict flora" (rather than the boreotropical flora) with disjunct distributions between Eastern North America and East Asia (Wen 1999, Xiang *et al* 2000, Milne & Abbott 2002, Milne 2006). Some examples of taxa are listed in Table 5.3. While the Tertiary relict flora exhibits an important and widely studied distribution pattern, it is not considered strictly tropical and so is not discussed further here.

In terms of hypothesis testing, divergence times of ~65-33 Ma between lineages on different continents would be consistent with the boreotropical hypothesis and cladograms would be expected to exhibit Neotropical, African and/or tropical Asian taxa being derived from a northern hemisphere lineage (if still extant) as depicted in Figure 5.5.

Figure 5.5 One potential boreotropical phylogenetic pattern showing Neotropical, African and Southeast Asian taxa derived from tropical, northern hemisphere ancestors. Excerpted from Pennington & Dick (2004), modified from Lavin and Luckow (1993).

Table 5.2 Examples of (predominantly dated) phylogenetic studies of megathermal taxa, which conform to a scenario of intercontinental boreotropical migration during the Paleocene-Early Oligocene. LDD = long distance dispersal.

Family/Taxon	Inferred biogeographical history	Reference
Lauraceae *Cinnamomum* & *Persea* groups & Laureae	Gondwanan origin followed by boreotropical migration	Chanderbali *et al* 2001
Leguminosae *Poitea* & *Pictetia*	Boreotropical relicts (currently in the Greater Antilles)	Lavin *et al* 2001
Leguminosae *Dichrostachys* group & Robineae	Hypothesized boreotropical migration – phylogeny not dated	Lavin & Luckow 1993
Sapotaceae Sideroxyleae	Diversification 65-35 Ma followed by boreotropical migration	Smedmark & Anderberg 2007
Malpighiaceae Acridocarpoid clade	Origin in South America ~68 Ma with repeated migration of several clades into North America and subsequent boreotropical migration to Africa and Asia during the Eocene. Migration from Africa to Madagascar 50-35 Ma and further LDD from Madagascar to New Caledonia ~15-8 Ma	Davis *et al* 2002a, Davis *et al* 2002b, Davis *et al* 2004
Simaroubaceae	Cretaceous origin followed by boreotropical migration and subsequent LDD between Africa & Asia, across the Atlantic & around the Pacific & Indian Ocean basins	Clayton *et al* 2009
Moraceae	Cretaceous origin with migration via multiple land routes including through the boreotropics	Zerega *et al* 2005
Rubiaceae *Gaertnera*	Hypothesized boreotropical migration to Africa during the early Tertiary	Malcomber 2002
Meliaceae - ancestor of *Cedrella* & *Toona*	The ancestor of Neotropical *Cedrella* and Asian *Toona* was present 46-50 Ma in London Clay flora. Phylogenetic splits between the genera coincide with hypothesized boreotropical migration.	Muellner *et al* 2010
Annonaceae	Cretaceous Gondwanan origin & subsequent boreotropical migration	Richardson *et al* 2004
Annonaceae *Guatteria*	Origin in Africa followed by boreotropical migration during the Eocene and subsequent dispersal into South America during the Miocene	Erkens *et al* 2009
Magnoliaceae *Magnolia*	Cretaceous origin followed by boreotropical migration. Tropical disjuncts are Eocene and temperate disjuncts are Oligocene corresponding with cooling and break-up of boreotropics	Azuma *et al* 2001

Table 5.3 Examples of (predominantly dated) phylogenetic studies of microthermal taxa which conform to a scenario of intercontinental boreotropical migration from the Mid-Oligocene through the Miocene. BLB = Bering Land Bridge, NALB = North Atlantic Land Bridge.

Family/Taxon	Inferred Biogeographical History	Reference
Berberidaceae *Berberis/Mahonia*	Origin in North America and migration to Eurasia 34-26 Ma followed by dispersal to South America ~25 Ma	Adhikari 2010
Styracaceae *Styrax*	Eurasian origin and migration/dispersal through the boreotropics to the Americas	Fritsch *et al* 2001
Cornaceae *Cornus*	Northern hemisphere origin and migration through the boreotropics with dispersal to Africa and South America	Xiang *et al* 2006
Araceae *Arisaema*	Northern hemisphere origin and migration through the boreotropics with Oligo-Miocene dispersal to Africa	Renner 2004
Ericaceae *Rhododendron* subsection Pontica	American-Eurasian disjunction hypothesized to result from two migrations across the Bering Land Bridge during the Tertiary	Milne & Abbott 2002, Milne 2004
Symplocaceae *Symplocos*	Northern hemisphere origin and migration through the boreotropics	Wang *et al* 2004
Anacardiaceae *Toxicodendron*	Hypothesized Miocene boreotropical migration in three separate lineages: two temperate disjunctions across the BLB (13-7 Ma) and one tropical disjunction across the NALB (20 Ma)	Nie *et al* 2009
Hamamelidaceae *Hamamelis*	Eocene (51 Ma) northern hemisphere origin, East Asia-East American disjunction estimated to be late Miocene (7 Ma) with migration across the BLB	Xie *et al* 2010
Leguminosae *Cercis*	Migration across the NALB or dispersal between 32 & 6 Ma	Davis *et al* 2002
Altingiaceae *Liquidambar*	Possible migration across both the BLB & NALB with northern hemisphere fossils ranging from the Paleocene to the Miocene	Ickert-Bond & Wen 2006
Ephedraceae *Ephedra*	Migration from Asia to North America across the BLB at ~ 30 Ma followed by dispersal from North America to South America ~ 25 Ma	Ickert-Bond *et al* 2009

5.5 Overview of "interplate dispersal paths for megathermal angiosperms" (from Morley 2000 & 2001)

Since the early Tertiary (~65 Ma), when global climates were warmer and megathermal angiosperms covered a much greater area, various pathways for their dispersal and migration have been available at different times within a changing mosaic of plate tectonic movement and climate oscillations. Availability of dispersal routes was dependent upon a variety of factors including: connectivity/proximity of land, temperature, moisture, elevation and light levels. Taking all of this into consideration, Morley (2000) distinguished nine potential dispersal routes for megathermal angiosperms, which he further divided into two categories: post-Gondwana break-up routes and routes which have formed since the Middle Eocene following plate collision.

Dispersal routes in the first category were available between 60-49 Ma when warm climates globally encouraged the dispersal and radiation of megathermal angiosperms. This period began with the break-up of Gondwana in the late Cretaceous - early Tertiary and culminated during the early Eocene thermal maximum. During this period six main dispersal/migration pathways may have been available:

- a transatlantic path between Europe and North America (boreotropical route)
- a route from Europe to Africa (Late Cretaceous-Early Tertiary trans-Tethyan dispersal)
- a land bridge between North & South America (Greater Antilles - Aves Ridge island arcs)
- a trans-Atlantic path between Africa & South America (Walvis Ridge/Rio Grande Rise)
- routes between Africa & India (via dispersal from Madagascar)
- a land bridge between South America & East Gondwana (Antarctica/Australasia)

The Bering Land Bridge, which also existed throughout this time, may have been located too far north (75°) for the dispersal of megathermal taxa, although it was an important migration route for several microthermal Tertiary relict taxa (Manchester 1999, Milne & Abbott 2002, Milne 2006).

The second category of dispersal routes, which formed between the middle Eocene (~45 Ma) and the present are primarily attributed to tectonic plate collision. During this period global climates began to cool, which inhibited the dispersal of megathermal angiosperms outside the tropics (Morley 2000):

- the collision of the Indian plate with Asia
- the complex collision of the Australian plate with the Philippine Arc and Asian plates
- the formation of the Panamanian Isthmus

The fossil record suggests that dispersal between tectonic plates is unlikely to happen without appropriate land connections and climatic similarity. Successful dispersals are more likely to occur within the same latitudinal and climatic zones than between zones and Morley (2000) believes that this may be the reason why dispersals between North America & Europe during the early Tertiary and from India to Southeast Asia during the Middle Eocene were particularly successful. However, dispersals between tectonic plates can also continue long after the time of initial separation, as evidenced by fossil pollen – e.g. dispersals across the

Atlantic between South America and Africa continued after 100 Ma (Morley 2000). Specifics of connectivity between diverging plates are not well known and volcanic mantle plumes and islands left in their wake may also have provided an opportunity for filter dispersal following continental breakup. The biogeographic histories of these dispersal routes particular to a specific region are presented in the sections relating to each area below.

5.6 Long distance dispersal

Dated molecular phylogenies indicate that long distance dispersal is much more common than formerly believed (e.g. Melastomataceae Renner *et al* 2001; Renner 2004a, b; *Exacum* Yuan *et al* 2005; *Cyrtandra* Cronk *et al* 2005; Cucurbitaceae Schaefer *et al* 2008; *Pseuduvaria* Su & Saunders 2009; Aglaieae Muellner *et al* 2008; Tables 5.5 & 5.6) and is an important factor in determining the make-up of modern tropical floras. Continental islands such as Madagascar and New Zealand were traditionally believed to be composed of relicts from former continental biotas (de Quieroz 2005). However, based on phylogenetic evidence, Yoder & Nowak (2006) surmise that "Madagascar is an island primarily comprised of neoendemics that are the descendents of Cenozoic waif dispersers." This hypothesis is corroborated by Ali & Huber (2010) through palaeogeographic reconstructions and palaeo-oceanographic modelling of strong Palaeogene ocean currents from East Africa to Madagascar, which would support sweepstakes dispersal. Winkworth *et al* (2002) come to a similar conclusion about the flora of New Zealand, stating that numerous successful dispersal events have occurred since the late Tertiary and many Southern Hemisphere plant distributions have arisen only within the last 10 million years. It has also been shown that recent colonizers make up a significant component of continental biota. Pennington & Dick (2004) suggest that transoceanic immigrant lineages comprise ca. 20% of species of a tree community in Ecuador. This implies that modern floras are much more dynamic than has been previously accepted and contain elements of recent origin.

Transport of disseminules via wind and ocean currents, on rafting islands of vegetation expunged from tropical rivers, and on the feet or in the gut contents of birds are all mechanisms which have been proposed to explain inter-continental disjunctions. Although long distance dispersal followed by the successful establishment of a seedling may be a rare occurrence, it is nonetheless plausible due to the immense time scales involved. One viable dispersal event every few million years makes long distance dispersal a significantly more likely scenario. As Nathan *et al* (2008) point out "In studying the ecology and evolution of processes such as dispersal, we usually focus our attention on the prevailing events, assuming that rare events are unimportant. Yet frequency and importance are not necessarily positively correlated. More recent studies have shown that long jumps available through rare long distance dispersal events are much more influential than the numerous small steps available through local dispersal in determining the spread of invasive species or range expansion of native species after climatic range shifts" (Fig. 5.6). In the case of Hawaii, it has been estimated that successful dispersals have taken place once every 98,000 years for plants, once every 68,000 years for insects and less than once in 1Ma for birds to account for the archipelago's current diversity (Price & Clague 2002). As such, long distance dispersal

has been recognised as being fundamental to the generation of biodiversity on oceanic islands (Cowie & Holland 2006, Baldwin & Wagner 2010).

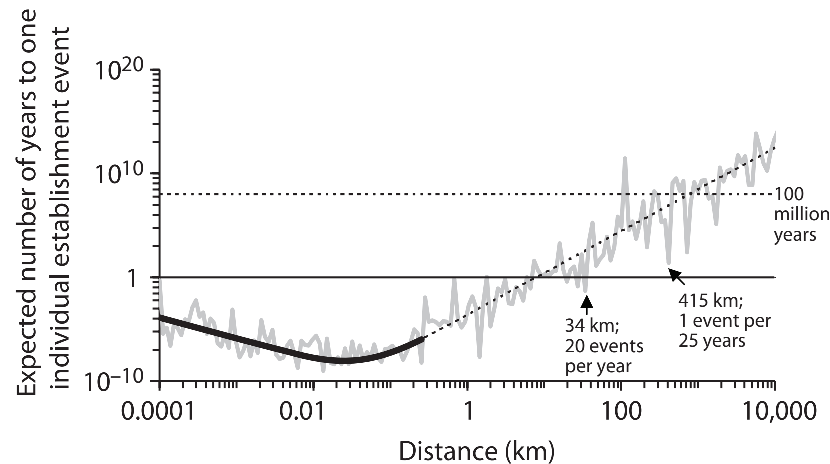


Figure 5.6 Graph of expected successful dispersal and establishment over distance and time. The expected time for a single effective dispersal event to occur is longer than one million years beyond 250 km. Nevertheless, an effective long distance dispersal event 415 km from the source, expected to occur once in almost 10 million years under the mean trend, may occur once in 25 years as a result of processes or events that "break the rules". Excerpted from Nathan *et al* 2006.

Extreme climatic events such as hurricanes and tornadoes have also been hypothesized to play a potentially important role in the dispersal of plant disseminules (Visher 1925, Nathan *et al* 2006, 2008). Interestingly, these same climatic events which may most effectively transport plant disseminules out to sea also have an impact on land, clearing territory for the establishment of new seedlings. In a post-hurricane study of *Manilkara bidentata* forest in Puerto Rico (You & Petty 1991) it was found that trees had a high survival rate, they rapidly adjusted to post-hurricane conditions and that there was increased seedling recruitment. They, therefore, deduced that such climatic events "play an important role in releasing suppressed seedling growth in *Manilkara* populations and that hurricanes may contribute to the abundance of *Manilkara* trees in the Luquillo Experimental Forest."

Givnish & Renner (2004) point out that "wind and ocean surface currents are not randomly distributed in space and time and should leave an evolutionary trace provided that they have been stable long enough to override lineage-specific differences in dispersal and establishment capability." Munoz *et al* (2004) demonstrated just such a pattern, showing that wind currents in the Southern hemisphere were the main mechanism in determining distributions for bryophytes and lichens (but not pteridophytes). This has also been shown to be the case in the Pacific for the wind-dispersed seeds of *Metrosideros* (Wright 2000 & 2001). While this may be a viable method for spore-bearing plants and those angiosperms with fine seed, few woody tropical angiosperms have seed which is small and light enough to be transported great distances by wind.

In the case of most tropical tree diaspores, transport via ocean currents or in rafting mats of vegetation is a much more feasible scenario than dispersal by wind. Darwin (1855, 1857, 1859) was an early proponent of this dispersal method and he carried out numerous

experiments demonstrating the survival of seeds which had been floating in sea water and also reported sightings of mats of vegetation floating in the ocean. The North and South Equatorial current systems act like a conveyor belt between the tropical regions of the world and can transport large objects such as a raft of vegetation between continents within less than three weeks (Houle 1998), e.g. between the Congo delta and eastern Brazil. While the direction of ocean currents is likely to have changed significantly during the Tertiary due to the closure of the Tethys and the Isthmus of Panama, it is possible that a narrower Atlantic Ocean would have facilitated even faster transport times between Africa and South America in particular. Table 5.5 lists fifteen taxa which exhibit this pattern of dispersal between the Neotropics and Africa.

Houle (1998) investigated the amount of time it would take an island of vegetation to raft across three bodies of water: the Atlantic between Congo & Brazil, the Caribbean Sea and the Southeast Indian Ocean from the northern plate of Australia to Sundaland at different times during the Paleogene. (Times are represented in Table 5.4.) He noted that the likelihood of transoceanic migrations was not equal for the three regions due to differences in paleocurrents and winds, which were running westerly in the Paleogene Atlantic Ocean and would have strongly favoured westerly migrations from Africa to the Neotropics. However, in the Southeast Indian Ocean, paleowinds were blowing in a northerly direction, while the currents were flowing to the south. The situation was reversed in the Caribbean, with paleowinds blowing south-westerly and paleocurrents flowing northerly. The longest time projected is just under a month rafting in the Southeast Indian Ocean, which is likely to be just within the time frame of viability for some tropical seeds.

Table 5.4 Hypothesized rafting times across the Atlantic Ocean, Caribbean Sea and Southeast Indian Ocean during different periods throughout the Tertiary according to Houle (1998).

Ocean System	rafting time 50 Ma	rafting time 40 Ma	rafting time 30 Ma
Atlantic Ocean	5.2-7.7 days	7.3-10.8 days	10-14.7 days
Caribbean Sea	11.2–18.2 days	10.2–16.6 days	9.3–15.1 days
SE Indian Ocean	24.2–25.6 days	18.4–19.5 days	11.5–12.2 days

Various studies on driftwood and other items have confirmed the ability of oceanic currents to transport objects extremely long distances. Hnatiuk & Rudall (1985) analysed driftwood genera (including *Mimusops*, Sapotaceae) found on Aldabra Atoll in the Indian Ocean and found that 83% of logs were Southeast Asian, 73% were Indian and 70% were African (some genera were found in more than one area). The marginally larger Southeast Asian component was attributed to rafting on the strong South Equatorial current (Fig. 5.7), which flows from Southeast Asia to Africa. A similar study done on driftwood, which washed ashore in Hawaii found that the vast majority of the wood had rafted from logging industries in the northeastern coast of North America (Oregon/British Columbia), whereas a few had come from the Philippines or Indonesia, one from Japan and one log from Central America (Strong & Skolmen 1963). In another survey, Barber *et al* (1959) found that *Nothofagus* and conifer driftwood on Macquarie island (in the Southern Ocean south of New Zealand) had floated from South and North America respectively. Likewise, Spennemann's (1997) study on drifted objects found on the Marshall Islands confirmed that dispersal had occurred from both east (North America & Hawaii) and west (Japan, Southeast Asia & Polynesia).

Seeds of tropical taxa are also commonly found on foreign beaches having drifted from distant sources. Although Gunn *et al* (1976) suggest that less than 1% of tropical spermatophytes are viable after ocean wayfaring for more than a month, numerous other studies show that a range of taxa have successfully dispersed long distances this way. In a study on Fijian drift disseminules, Smith (1990) recorded 73 species in 36 families and noted that some non natives germinated (i.e. *Annona, Chrysobalanus* and *Xanthium*). Likewise, Green (1999) logged 63 species in 29 families in the Christmas Island drift seed flora, the majority of which were non-natives hypothesized to have arrived from Java and other nearby Indonesian islands. Some of these seeds displayed a high degree of viability on arrival (e.g. *Dioclea hexandra*, *Erythrina fusca* and *Mucuna gigantea*). Costin (1965) reported the germination of a *Caesalpina bonduc* seed collected on Macquarie Island, which he hypothesized had drifted 12,000 miles from Central America. Similarly, in New Zealand, Mason (1961) noted that of the non-native tropical seeds, which washed ashore, most were leguminous and seven out of twenty two *Entada* seeds germinated successfully.

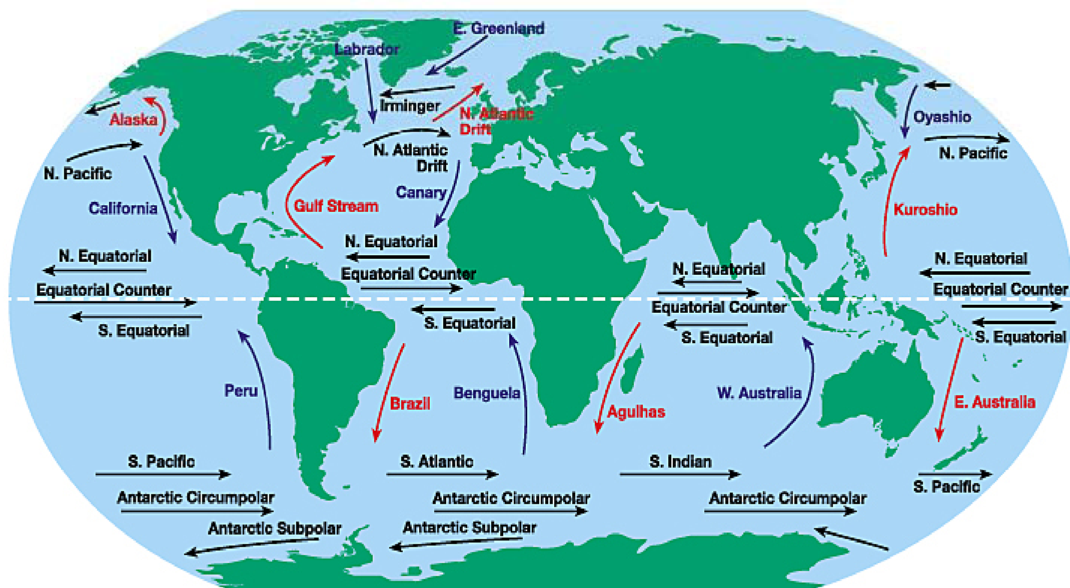


Figure 5.7 Modern ocean currents. From: https://fretzreview.wikispaces.com/file/view/Surface_currents.jpg/30705280/Surface_currents.jpg

Transoceanic dispersal of seeds by birds has also been suggested, but this is unlikely due to the fact that birds migrate north-south, not east-west and they void the contents of their guts frequently (Proctor 1968, Fukui 2003). Additionally, while it is possible for small seeds to be transported in soil on the feet of a bird, it is not likely for most tropical trees, which tend to have large seeds. However, chance events do happen and birds, along with their attached seeds, can get blown off course by tropical storms. Some have sporadically reached distant landmasses, as in the case of Tristan da Cunha which hosts five endemic land bird species (Renner 2004c) and North American birds which are regularly blown across the Atlantic to Shetland (Geoff Harper pers. comm. 2002). More recently human-mediated transport has created intercontinental disjunctions in the range of species. This has occurred unintentionally in the ballast of ships, in textiles and packing materials and more recently through air transport, as well as intentionally with the movement of economic plants around the world.

Dispersal may occur at any point in time and is, therefore, hard to test against specific vicariance events. However, since the Oligocene there have been no direct migration pathways available between the tropical regions of the world (Neotropics, Africa & Southeast Asia-Pacific), nor have there been any major vicariance events (i.e. tectonic plate rifting) which could explain geographical splits in lineages at this time. Therefore, post Eocene (~33 Ma to present) geographical splits in tropical taxa can best be explained through dispersal.

There are numerous instances in which dated phylogenies indicate relatively recent intercontinental dispersal in plants. i.e. across the Pacific Ocean: Sytsma *et al* 2008, Smith *et al* 2008, Michalak *et al* 2010; across the Indian Ocean: Baum *et al* 1998, Yuan *et al* 2005, Davis *et al* 2002, 2004, Renner *et al* 2001, Zhang *et al* 2007, Kuluju *et al* 2007, Li *et al* 2008, Michalak *et al* 2010, Bouetard *et al* 2010; and across the Atlantic Ocean: Lavin 2000, Skema 2003, Weeks *et al* 2007, Dick *et al* 2003, Sarkinen *et al* 2007, Renner 2004c. See Tables 5.5 and 5.6.

Table 5.5 Recent long distance dispersal between the Neotropics and Africa, examples from dated phylogenetic studies. LDD = long distance dispersal.

Family	Taxon	Direction	Age	Reference
Rapateaceae	*Maschalocephalus*	LDD from South America to Africa (Wind)	7 Ma	Givnish 2004
Bromeliaceae	*Pitcairnia*	LDD from South America to Africa (Wind)	12 Ma	Givnish 2004
Melastomataceae	Melastomeae	LDD from South America to Africa (Wind)	11 Ma	Renner *et al* 2001
Malvaceae	*Ceiba pentandra*	LDD from South America to Africa (Wind)	recent	Dick *et al* 2007
Leguminosae	*Macherium*	LDD from South America to Africa (Water?)	recent	Lavin 2000
Leguminosae	*Andira*	LDD from South America to Africa (Water)	recent	Skema 2003
Hernandiaceae	*Hernandia*	LDD from South America to Africa (Water?)	3 Ma	Michalak *et al* 2010
Annonaceae	*Annona*	LDD from South America to Africa (Water?)	16-13 Ma	Richardson *et al* 2004
Clusiaceae	*Symphonia globulifera*	LDD from Africa to South America (Water)	~17 Ma	Dick *et al* 2003
Zingiberaceae	*Renealmia*	LDD from Africa to South America (Water)	15.8-2.7 Ma	Sarkinen *et al* 2007
Rhizophoraceae	*Cassipourea* & *Rhizophora*	LDD from Africa to South America (Water)	recent	in Renner 2004c Schwarzbach pers. comm.
Arecaceae	*Raphia taedigera*	LDD from Africa to South America (Water)	recent	in Renner 2004c W. Baker pers. comm.
Arecaceae	*Elaeis oleifera*	LDD from Africa to South America (Water)	recent	in Renner 2004c W. Hahn pers. comm.
Malvaceae	*Gossypium*	LDD from Africa to South America (Wind?)	5-10 Ma	Cronn *et al* 2004, Wendel *et al* 1995
Burseraceae	*Commiphora*	LDD from Africa to South America (Water?)	Miocene	Weeks *et al* 2007

Table 5.6 Examples from dated phylogenetic studies which conform to a scenario of long distance dispersal. LDD = long distance dispersal.

Family/ Taxon	Inferred biogeographical history	Reference
Atherospermataceae	Gondwanan origin & diversification 100-140 Ma followed by LDD to New Zealand & New Caledonia 50-30 Ma	Renner *et al* 2000
Asteraceae Vernonieae	Origin in Gondwana/South Africa/Madagascar, followed by LDD with at least 2 trans-Atlantic dispersals between Old and New world	Keeley *et al* 2007
Myrtaceae	Possible LDD between Australasia, South America, Mediterranean & Africa	Sytsma *et al* 2004
Vochisiaceae	Possible Oligocene LDD from South America to Africa	Sytsma *et al* 2004
Simaroubaceae	Boreotropical migration and LDD between Africa & Asia, across the Atlantic & around the Pacific & Indian Ocean basins	Clayton *et al* 2009
Begoniaceae *Begonia*	Cretaceous origin, Eocene diversification in Africa and subsequent LDD ~20 Ma to Asia & Neotropics	Copestake *et al* 2009
Gesneriaceae *Cytandra*	Recent LDD from Southeast Asia into the Pacific	Cronk *et al* 2005
Ehretiaceae Cordiaceae Heliotropiaceae	Cretaceous origin followed by intercontinental migration and LDD	Gottschling *et al* 2004
Gentianaceae *Exacum*	Post-Eocene origin followed by LDD from Madagascar across the Indian Ocean Basin to India & Sri Lanka –followed by range expansion in Socotra-Arabia and Southeast Asia.	Yuan *et al* 2005
Cucurbitaceae	Cretaceous origin in Asia & 43 successful LDD events to Africa, Australia and the Americas in the past 60 Ma = 7 LDDs every 10 Ma. e.g. *Curcumis* LDD from Africa to Asia <10 Ma	Schaefer *et al* 2008, Renner *et al* 2007
Annonaceae *Pseuduvaria*	Origin in Sundaland ~8 Ma with subsequent dispersal east to New Guinea/Australia as well as back-dispersal to Sundaland	Su & Saunders 2009
Ericaceae *Rhododendron* section Vireya	Origin ~27 Ma with dispersal across Wallace's line ~12 Ma origin in mainland Southeast Asia with subsequent dispersal across the archipelago	Twyford & Richardson in prep., Brown 2006
Meliaceae Aglaieae	Late Eocene origin, Mio-Pliocene dispersal from Sundaland to India and eastwards throughout Southeast Asia to New Guinea with further Pliocene dispersals into the Pacific	Muellner *et al* 2008
Sapotaceae Chrysophylleae	Three+ Miocene LDD events from Australia to New Caledonia	Bartish *et al* 2005
Zingiberaceae *Etlingera*	Origin 18 Ma in Sumatra with eastward dispersal and new distinct species evolving after another 8 Ma as well as back dispersal	A. Poulsen pers. comm. 2010
Bombacaceae *Adansonia*	Miocene origin followed by LDD between Africa and Australia 2-15 Ma	Baum *et al* 1998
Proteaceae	Gondwanan origin and vicariance followed by Eocene LDD between Africa-Australia, Africa-South America & New Zealand –Australia	Barker *et al* 2007
Malpighiaceae Acridocarpoid clade	Origin in South America ~68 Ma with repeated migration of several clades into North America and subsequent boreotropical migration to Africa and Asia during the Eocene. Migration from Africa to Madagascar 50-35 Ma and further LDD from Madagascar to New Caledonia ~15-8 Ma	Davis *et al* 2002, Davis *et al* 2004a, Davis *et al* 2004b
Melastomataceae	Paleocene/Eocene diversification in the boreotropics followed by Oligocene migration to South America and subsequent LDD to Africa 14-12 Ma. From Africa LDD to Madagascar, India & Indochina. Plus, multiple Mio-Pliocene LDD of Asian Sonerilleae & Dissochaeteae to Madagascar & Africa	Renner *et al* 2001, Renner 2004a, Renner 2004b
Leguminosae *Atelia*	Recent LDD from Central to South America – three independent colonization events	Ireland *et al* 2010

Annonaceae	Cretaceous Gondwanan origin & vicariance between South America/Africa, subsequent boreotropical migration and LDD of 2 clades from Southeast Asia to Africa ~10-14 Ma as well as LDD from South America to Africa ~14-16 Ma, Southeast Asia to Australia ~19 Ma, e.g. *Uvaria* LDD from Asia to Africa 12-15 Ma	Richardson *et al* 2004
Rhamnaceae	Possible boreotropical migration of some groups and extensive LDD	Richardson *et al* 2004
Burseraceae	Laurasian origin ~60 Ma with boreotropical migration and subsequent LDD	Weeks *et al* 2005
Verbenaceae *Vitex*	At least one LDD event between Africa and South America 9.3 Ma	Cabral 2008
Clusiaceae Asian *Garcinia*	Dispersal from west to east Malesia during the Miocene ~10 Ma	Saleh 2006
Rubiaceae *Tricalysia, Ixora, Empogona*	*Tricalysia*: one Plio-Pleistocene dispersal from East Africa to Madagascar. *Ixora*: one dispersal between Africa & Madagascar 8 Ma, a second dispersal from Africa to Madagascar 2.7 Ma. *Empogona*: Recent dispersal of *E. ovalifolia* from Africa to Madagascar.	Tosh 2009
Santalaceae *Santalum*	Multiple LDD events from Australia throughout the Pacific during the late Miocene-Pliocene	Harbaugh & Baldwin 2007
Goodeniaceae *Scaevola*	At least six LDD events from Australia throughout the Pacific (undated)	Howarth *et al* 2003
Phyllanthaceae *Bridelia*	LDD from Asia to Africa once or twice and from Southeast Asia to Australia twice during the late Miocene	Li *et al* 2009
Rubiaceae *Gaertnera*	LDD from Africa to Asia during the late Miocene 5-6 Ma	Malcomber 2002
Euphorbiaceae *Macaranga* & *Mallotus*	LDD from Asia to Africa <27 Ma	Kulju *et al* 2007
Hernandiaceae *Hernandia* & *Illigera*	Oligocene-Miocene LDD: *Hernandia* exhibits both trans-Pacific and trans-Atlantic dispersal events, *Illigera* includes two trans-Indian Ocean dispersals.	Michalak *et al* 2010
Ebenaceae *Diospyros*	Multiple LDD events inferred from Southeast Asia to New Caledonia	Duangjai *et al* 2009
Piperaceae *Piper* & *Peperomia*	Multiple intercontinental LDD events between the Neotropics, Africa and Asia-Pacific	Smith *et al* 2008
Rubiaceae Knoxieae & Vanguerieae	Numerous dispersals from Africa and Asia to Madagascar as well as from Madagascar to the Comoros, Mascarenes and Seychelles	Wikström *et al* 2010
Rutaceae *Melicope*	Multiple dispersal events throughout the Pacific	Harbugh *et al* 2009
Proteaceae *Macadamia*	Inferred long-distance dispersal out of Australia	Mast 2008
Arecaceae Chamaedoreeae	Hypothesized multiple LDD events, two independent dispersals between North and South America in the mid Eocene and Miocene prior to the closing of the Isthmus of Panama and also LDD from the ancestral range to the Mascarenes. Alternatively, boreotropical migration could also be supported.	Cuenca *et al* 2008
Orchidaceae *Vanilla*	Three separate hypothesized LDD events during the Oligocene and Miocene from Africa to Asia, Africa to the Caribbean, and Africa to Indian Ocean Islands.	Bouetard *et al* 2010
Anisophylleaceae *Anisophyllea*	Split between Neotropical and Old World lineages of *Anisophyllea* ~23 Ma, Africa-Asian dispersal 22 Ma	Zhang *et al* 2007

5.7 Regional biogeographic histories

In relation to the vicariance and dispersal scenarios at different time scales presented above, an outline of regional biogeography in the Neotropics, Africa and Asia is presented below in order to give a regional context to the hypotheses which will be tested in Chapter VII.

5.7.1 Regional history of South America

In South America, the main lowland rain forest blocks are located in the Amazon basin, along the Brazilian Atlantic coast, and in the Chocó and the Magdalena valley of Colombia. Dry forest and savanna are also found in pockets throughout the region, most notably in the Caatinga and Cerrado respectively. These two dry biomes separate the Atlantic coastal forest from the Amazon, while the Andes separate the Chocó and the Magdalena valley from the Amazon basin (Fig.5.8) (Daly & Mitchell 2000).

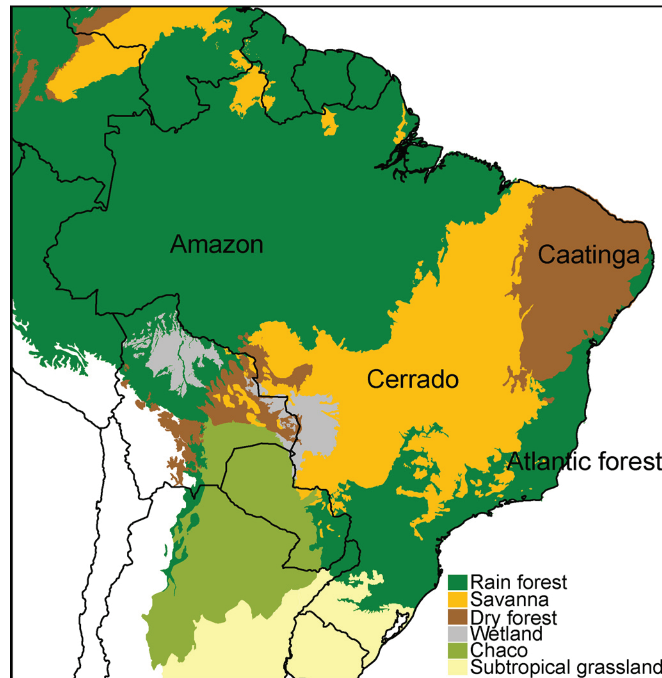


Figure 5.8 Major vegetation types in tropical South America, showing the wet Amazon and Atlantic Coastal forest biomes separated by the dry biomes of the Cerrado and Caatinga. Excerpted from Simon *et al* (2009).

During the Cretaceous, South America formed part of the supercontinent Gondwana. The opening of the south Atlantic ~105-119 Ma separated South America from Africa and cut it off from the other Gondwanan landmasses, except for its connection to West Antarctica. This southern link between Tierra del Fuego and Antarctica was severed by the late Oligocene ~28 Ma, opening up the Drake Passage (McLoughlin 2002). Throughout most of the Tertiary, South America remained isolated until the closure of the Isthmus of Panama (~3 Ma) resulted in a terrestrial connection with North America (Coates & Obando 1996).

According to pollen records, the Terminal Cretaceous extinction event severely reduced angiosperm diversity in South America (Morley 2000). Subsequently, tropical rainforests are believed to have only appeared during the Palaeocene (Wing *et al* 2009), but not significantly developed or diversified until the Eocene (Burnham & Johnson 2004). Jaramillo (2002) suggests that during the Eocene the tropical South American flora reached levels of diversity comparable with modern tropical rain forests and Hooghiemstra *et al* (2002) add that plant diversity may even have been higher than at present. During the Oligocene and Miocene the tropical lowland flora resembled contemporary lowland neotropical vegetation (van der Hammen 1991)

From the Cretaceous until the Late Oligocene, the paleo-Orinoco, which flowed northwest towards the Caribbean, was the main drainage system for Amazonia. However, during the Early Miocene (~25 Ma) the collision of the South American and the Nazca/East Pacific plates caused tectonic changes in the Amazon basin and thrust up the Eastern Cordillera of the Andes (Parra *et al* 2009), significantly altering the drainage patterns in the central shield areas of the continent (Hoorn *et al* 2010; Mora *et al* 2010). This drastic rearrangement of topography caused the separation of Amazon basin rain forest and that of the Chocó and the Magdalena Valley and localized high rainfall in the western Amazon (van der Hammen 1992, Hooghiemstra *et al* 2002, Sepulchre *et al* 2010). As a result of this tectonic activity, western Amazonia (just on the eastern side of the Andes) gradually became submerged creating a system of wetlands and lakes approximately one million square kilometres wide known as Lake Pebas. Western Amazonia remained flooded throughout the Mid to Late Miocene (~17-7 Ma) and was connected to the Caribbean by a marine incursion in the north near Lake Maracaibo. A second phase of rapid Andean orogeny during the Late Miocene (11-7 Ma) is hypothesized to have also uplifted the Guiana Shield, which blocked the Caribbean outlet of the paleo-Orinoco causing the re-arrangement of the Amazonian drainage system to the east and the eventual drying out of the Lake Pebas wetland system ~ 7 Ma (Hoorn *et al* 2010; Mora *et al* 2010, Antonelli *et al* 2008).

The exact age and sequence of Andean orogeny is debated and there is evidence that different sections of the range uplifted from south to north and west to east at different times (Parra *et al* 2009). Recent studies also indicate that the elevation of the Andes was relatively stable for tens of millions of years punctuated by rapid (1 to 4 million years) changes of 1.5 kilometers or more (Garzione *et al* 2008). A significant section is hypothesized to have reached only half its current height by 10 Ma, with the Eastern Cordillera of the Colombian Andes having achieved no more than 40% of its modern elevation by 4 Ma, suggesting considerable recent uplift (Gregory-Wodzicki, 2000, Graham 2009).

Aside from the Andean orogeny, the other major geologic event which greatly impacted the biogeography of South America was the formation of the Isthmus of Panama; a complex process spanning the past 15 Ma. The formation of this land bridge during the late Neogene ~3.5 Ma (Coates & Obando 1996, Pennington & Dick 2004) radically transformed ocean circulation and global climatic patterns and affected the ecology and geography of organisms by opening a direct terrestrial migration route between North and South America (Coates & Obando 1996, Iturralde-Vinent & MacPhee 1999).

Oscillations in temperature and rainfall during the Pleistocene ice-ages (2.7-2.2 Ma) are believed to have affected rain forest diversity in lowland areas of South America, and may have caused an exchange of floristic elements between different altitudinal belts, stimulating diversification in the periphery of the area (Hooghiemstra *et al* 2002), with the successive expansion and contraction of populations acting as a "species pump" (Morley 2000). Pleistocene refugia for tropical rain forest taxa have been proposed in three areas: in an arc along the lower flanks of the Andes, along the Atlantic seaboard, and in hill areas of the Guyana shield (Haffer 1969, Morley 2000). However, Pleistocene fossil pollen data presented by Colinvaux (2000) refute this theory and instead suggest that Amazon rain forests have been stable since the beginning of the Pleistocene, and that this stability has contributed to their current diversity.

South America was more or less biogeographically isolated for ~100 Ma, from the time of its split with Africa during the Cretaceous ~105 Ma (Goldblatt 1993) until the closure of the Isthmus of Panama ~3.5 Ma (Pennington & Dick 2004, Coates & Obando 1996). Although it was not directly connected to other landmasses throughout this time, dispersal may still have been possible via various hypothesized land bridges and stepping stones, as outlined below:

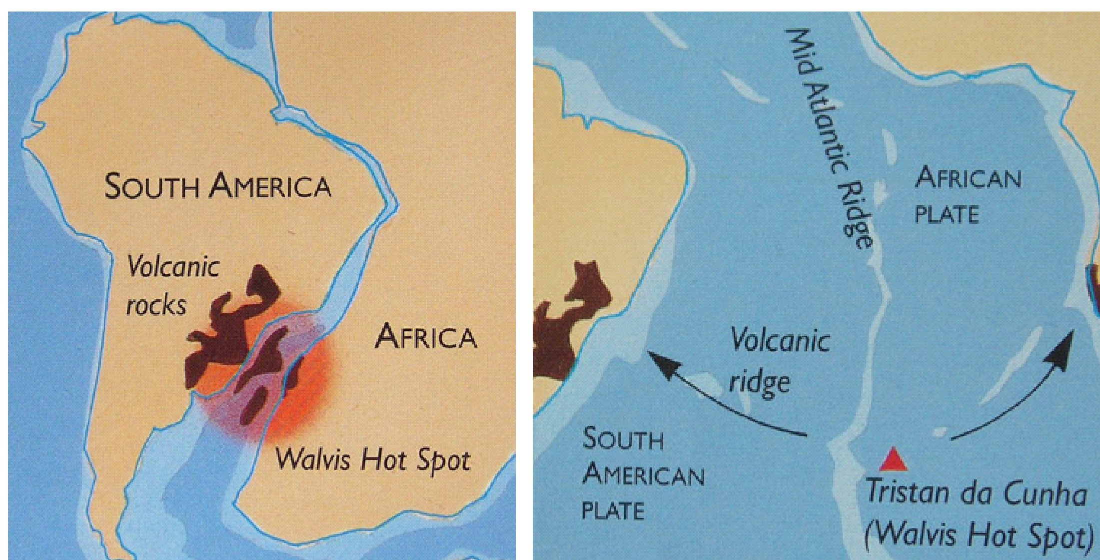


Figure 5.9 Walvis Ridge – Rio Grande Rise Hotspot depicted during the Cretaceous and in modern times. Excerpted from http://www.tristandc.com/earthvolcano.php

a) Walvis Ridge-Rio Grande Rise & Sierra Leone-Ceará Rises: dispersal from Africa to South America from the Cretaceous until possibly the Oligocene, through island chains between the Walvis Ridge and the Rio Grande Rise as well as the Sierra Leone and Ceará Rises. Dispersal may have been facilitated by lowered sea levels ~88 Ma (Morley 2000, Morley 2003). However, Pennington & Dick (2004) point out that the Ocean Drilling Program found no evidence of dry land in Late Cretaceous-Early Tertiary sediments along the Walvis ridge (Fig. 5.9)

b) Proto Antilles and GAARlandia (Greater Antilles –Aves Ridge island arcs): dispersal from North America to South America during the Middle-Late Eocene may have been possible through the islands of the proto-Antilles, which were submerged on the Caribbean

plate and became uplifted as it moved between North and South America and made contact with the Bahamas plate in the Middle Eocene. There is also suggestion that the Antilles in combination with the currently submerged Aves Ridge (GAARlandia) may have acted as a dispersal corridor during the Eocene-Oligocene boundary (35-33 Ma) (Iturralde-Vinent & MacPhee 1999, Pennington & Dick 2003) (Fig. 5.10).

c) Southern Gondwanan connection: dispersal between South America-Antarctica-Australia was possible from the late Cretaceous until the Eocene, although it was unlikely to be a significant dispersal pathway for megathermal angiosperms due to its near-polar position (Morley 2003, Pennington & Dick 2003) (Fig. 5.2).

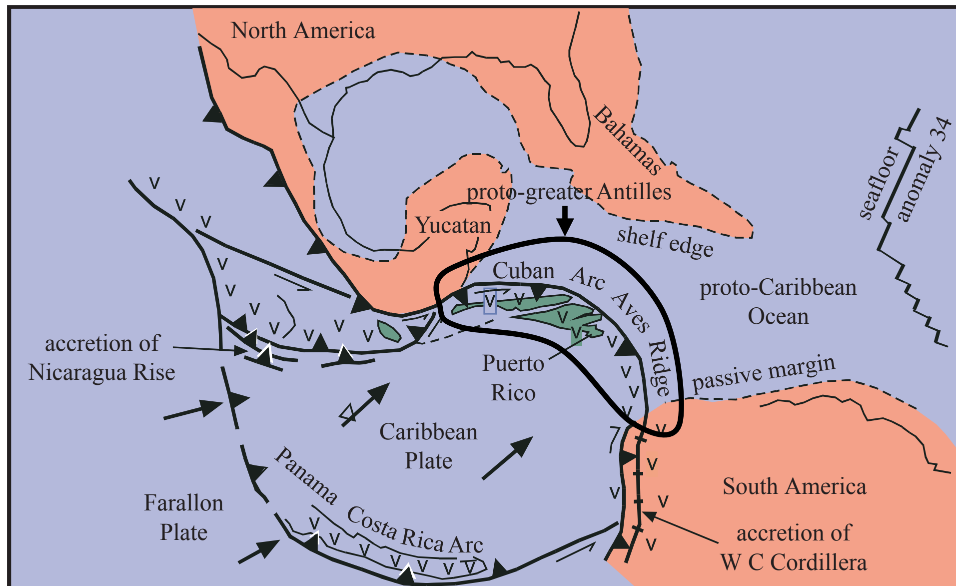


Figure 5.10 The Proto Greater Antilles as depicted by Pennington & Dick (2004)

Andean angiosperm radiations are beginning to be widely studied, including: *Lupinus*, Leguminosae (Hughes & Eastwood 2006), *Astragalus*, Leguminosae (Scherson *et al* 2008), *Amicia*, Leguminosae (Sarkinen 2010), Valerianaceae (Bell & Donoghue 2005, Moore & Donoghue 2007), *Halenia*, Gentianaceae (von Hagen & Kadereit 2003), *Gentianella*, Gentianaceae (von Hagen & Kadereit 2001), Iochrominae, Solanaceae (Smith & Baum 2006), *Espeletia*, Asteraceae (Rauscher 2002), Paranepheliinae, Asteraceae (Soejima *et al* 2008), as well as some transitional sub-Andean-Amazonian taxa including *Inga*, Leguminosae (Richardson *et al* 2001) and Rubiaceae (Antonelli *et al* 2008). These studies, primarily of sub-alpine and alpine species, suggest that a significant proportion of diversification in the Andes is relatively recent and rapid (since ~2 Ma), reflecting the uplift history of the mountain range. Research into the origins of Neotropical dry forest has also recently received attention (Pennington *et al* 2000, Pennington *et al* 2004, Mayle 2004, Lavin 2006, Becerra 2005, Dick & Wright 2005, Pennington *et al* 2010). These studies indicate dry forest biome ages ranging from the Mid-Miocene to the Pliocene.

In contrast, very few phylogenetic studies have focussed on Amazonian plant taxa (exceptions are Sarkinen *et al* 2007 and Antonelli *et al* 2008) and there is little to no information about the historical formation of other biomes in the region such as the

Atlantic Coastal forest of Brazil, the Caatinga and the Cerrado. However, one recent phylogenetic study (Simon *et al* 2009) indicates that dry-adapted Cerrado Leguminosae and Melastomataceae lineages diversified from the Late Miocene to the Pliocene (from 9.8 to 0.4 Ma) suggesting that the Cerrado biome has been in place since at least this time. Additionally, another study of *Coursetia* (Leguminosae) (Lavin 2006) reveals that species which inhabit the dry forest of the Brazilian Caatinga are 5-10 Ma.

Based on the geological history of the region and on these studies, phylogenetic splits reflecting speciation in lowland restricted lineages driven by Andean orogeny would be expected from ~15 Ma. The dry biomes of the Cerrado and Caatinga are both hypothesized to be of Mid Miocene age (~10 Ma). They separate two important centers of Neotropical rain forest endemism, the Atlantic Coastal Forest and Amazonia (Fig. 5.8). If a once-continuous ancestral distribution of rain forest species were to be split by a drying climate and the creation of these dry biomes, this would be expected to result in lineage splits between the Atlantic Coastal Forest and Amazonian forest blocks at around 10 Ma. The discontinuous nature of these two rain forest blocks is also likely to be related to the reorganization of drainage patterns in the Amazon basin following Andean uplift, with the formation of the modern Amazon delta in the northeast between these two regions 11-7 Ma. Likewise, if the paleo-wetland system of Lake Pebas had an effect on the diversification of lowland Amazonian taxa, this might be mirrored in phylogenetic splits between taxa either side of this system ~17 Ma (Antonelli *et al* 2008). If the creation of the Panama Isthmus was pivotal in allowing taxa to cross from South to Central America or vice versa, lineage splits at ~3 Ma would be predicted, while if the division between Central and South American lineages were due to dispersal, the split would be expected prior to 3 Ma. These scenarios are represented in the hypothetical chronogram in Figure 5.11. In addition, if Pleistocene climatic changes acted as a species pump, numerous splits would be expected between species that date to this epoch.

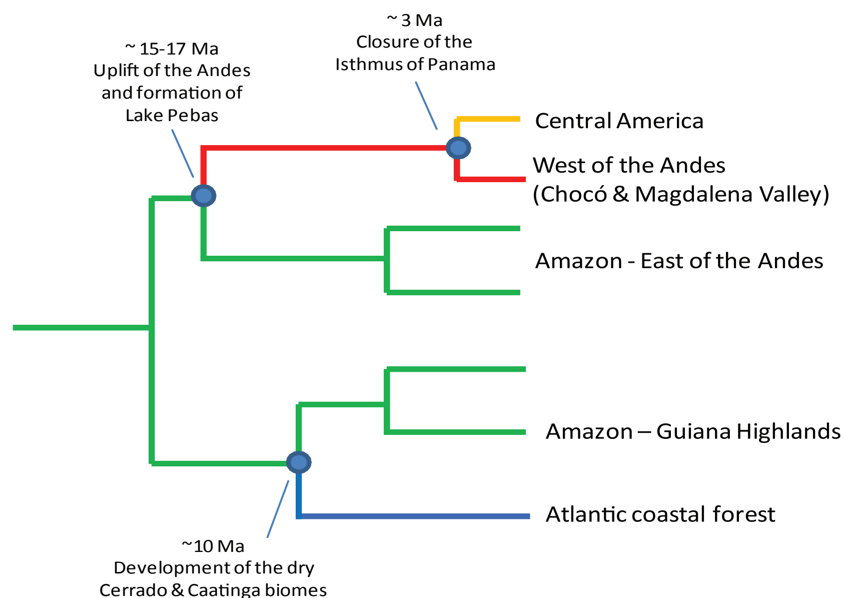


Figure 5.11 Hypothetical area cladogram depicting ages of splits between lineages corresponding with paleogeographical phenomena including Andean orogeny, aridification and the closure of the Isthmus of Panama.

5.7.2 Regional history of Africa

In Africa, rain forests are restricted to the west and central regions of the continent (Guineo-Congolia), and to small, isolated pockets of coastal and montane forests in the east (White 1983), while much of the remainder of the continent is covered in dry savanna or desert (Fig. 5.12).

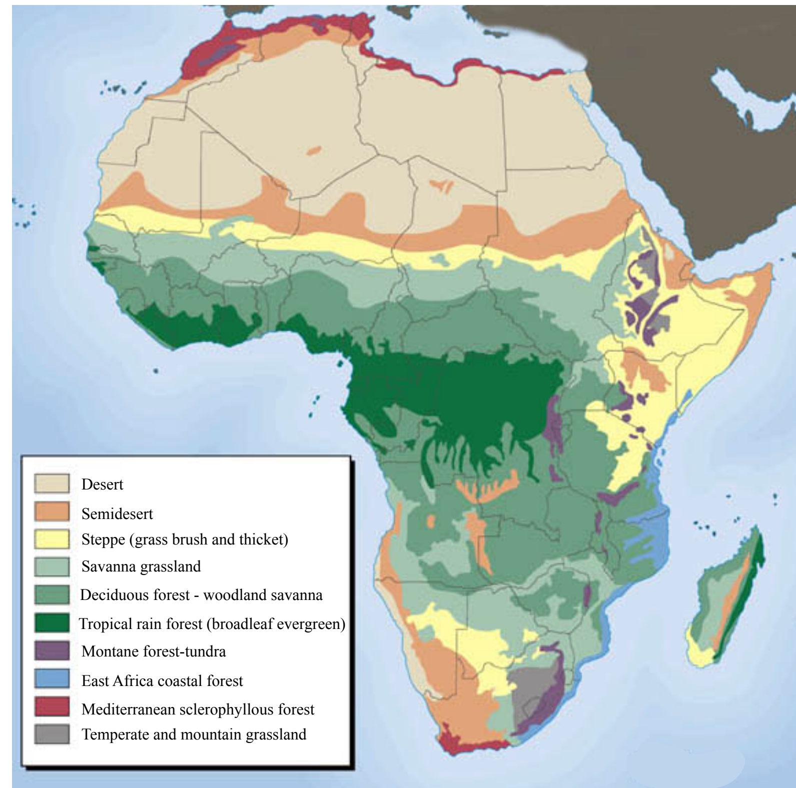


Figure 5.12 Vegetation map of Africa, adapted from: http://exploringafrica.matrix.msu.edu/

During the Cretaceous, Africa formed part of the Gondwanan supercontinent situated between South America and India-Antarctica-Australia. Africa broke away from Antarctica-Australia ca.165-162 Ma and separated from South America ~105 Ma, (McLoughlin 2002) but remained adjacent to Madagascar and the Indian subcontinent until 84-66 Ma (Morley 2000). By the Late Palaeocene (60-54 Ma.), Africa was isolated and slowly drifting northwards towards Eurasia. During this time, it was situated approximately 15-17 degrees south of its present latitude and had a stable climate. North Africa was covered in rainforest from the Gulf of Guinea to the Tethys, savanna bordered the rainforest to the south, and south and east Africa were covered in subtropical forest (Plana 2004, Lovett & Wasser 1993).

Extensive extinctions occurred throughout Africa during the Cretaceous/Tertiary boundary (65.5 Ma), with the number of pollen taxa falling by 50% during this period, and Morley (2000) suggests that this figure may apply to the entire African flora. However, from the Late Palaeocene (~60 Ma) until the end of the Eocene (~35 Ma), the West African flora continued to diversify unhindered. This is evidenced by the appearance and radiation of taxa indigenous to Africa and to a lesser extent dispersals from other areas (Morley 2000). There is also clear evidence of a continent-wide latitudinal temperature gradient and zonation of vegetation throughout the Palaeocene and Eocene. Palaeocene pollen studies suggest that

the Sahara region had already formed a major biogeographical divide at this time, either due to its upland topography or to arid climates. Microfossil assemblages from the Palaeocene/Eocene boundary (~57 Ma) also suggest a significant change in lowland vegetation, with the coexistence of a mosaic of dry and humid forest with savannah woodland throughout most of the Congo basin (Coetzee 1993).

By the Late Eocene (39-36 Ma.) latitudinal zonation had become less distinct, global climates were cooling and closed rainforests had become extensive at equatorial latitudes. The equatorial African flora appears to have changed during this period of cooling and contraction, with pollen of many extant tree taxa (particularly legumes) suddenly appearing in the fossil record (Morley 2000). Yet, while tropical forests contracted near the equator, the terminal Eocene "big chill" resulted in relatively few extinctions at equatorial latitudes in Africa. In Egypt pollen of tropical tree taxa becomes much more common (Kedeves 1971) and Morley (2000) suggests that this element of the flora may be an immigrant dispersing from Europe.

The first evidence for drier climates at lower latitudes comes during the late Oligocene 30-25 Ma with an increase in grass pollen coinciding with a fall in sea level. Morley (2000) has suggested that these grasses most likely formed the under storey of forests rather than open savannas. At the same time, pollen assemblages from West Africa are characterized by the appearance of many typically African rainforest trees and climbers

Climatic and tectonic events during the Miocene (23-5 Ma) played a crucial role in defining Africa's current biotic composition. During the Early Miocene climates were moist throughout equatorial Africa with rain forests extending from coast to coast. Yet, the floristic diversification which had occurred during the Oligocene became reduced during the early Miocene, possibly due to the closure of the Tethys sea, which altered climatic circulation patterns and resulted in more seasonal equatorial climates (Morley 2000). By this point, the pan-African forest had become fragmented into western and eastern blocks under the climatic influences of the Atlantic and Indian Oceans respectively (Morley 2000). Additionally, the general uplift of the African continent (East African Rift) dramatically altered drainage patterns and may have blocked lowland rain forest taxa from east-west dispersal into suitable refuges (Lovett & Wasser 1993). There is considerable debate over the timing of the initiation of the tectonic rifting in the region, but it is clear that major phases of orogeny occurred in the mid Miocene, with the East African Plateau uplifting around 13.5 Ma (Wichura *et al* 2010) and the Ethiopian Plateau uplifting from 20 Ma (Pik *et al* 2008, Corti 2009).

Although Africa had been colliding with Eurasia since ~63 Ma (Sanmartin, 2002), the final thrust, which resulted in the closure of the Tethys sea and the formation of the Mediterranean basin occurred during the Middle Miocene (16-10 Ma). There is ample evidence of an intercontinental faunal interchange at this time, but the same is not true for the floras of the two regions; at least there is no fossil evidence for this in relation to tropical taxa (Morley 2000). This may be the result of strong latitudinal zonation restricting dispersal between higher and lower latitudes.

Palynological data from both the Niger delta and Cameroon suggest a gradual change in the character of African vegetation during the course of the Miocene, with many tropical woody taxa suddenly disappearing from the pollen record during the Mid-Miocene (Legoux 1978). During the same time grass pollen increases, suggesting an expansion of savanna and open woodland in place of rain forest over a wide area. By the Late Miocene (10-5 Ma) periods of savanna expansion were more pronounced and pollen of tropical trees are less well represented in pollen assemblages although few extinctions of rain forest taxa have been noted for this time (Morley & Richards 1993; Morley 2000). Changes to the temperature of ocean currents also had an important impact on the vegetation of western Africa. Most critically, cold water had been introduced into the Benguela current during the Middle-Late Oligocene (~30-24 Ma) and by the Late Miocene (10-5 Ma) the cold up-welling had intensified causing aridification in the Namib region that resulted in the replacement of wooded grassland with desert (Coetzee 1993).

This Late Miocene cold period was followed by a warm, moist episode from 9-3.5 Ma during which lowland rain forest expanded (to approx. 20° lat.) and diversified while savannas contracted. Yet, by 3.3 Ma increased trade wind activity in combination with colder ocean currents is thought to have enhanced continental dryness and desertification.

The drying up of the Mediterranean Sea between 6.4 and 4.6 Ma during the Messinian Salinity Crisis is believed to have increased aridity in Africa resulting in a major expansion of savanna at the expense of forest. Likewise, the closing of the Indonesian seaway 3-5 Ma has also been suggested as a cause for east African aridification due to the subsequent cooling of the Indian Ocean (Cane & Molnar 2001). The onset of this dryness also coincides with glaciation in the northern hemisphere and the subsequent occurrence of Pleistocene glacial-interglacial sequences (Coetzee 1993).

Ice-age aridity had a devastating effect on lowland rain forest in Central Africa. Palynological studies of this period show that rain forest was mostly replaced by dry grassland and savanna. The patchy modern distribution of rain forest in Africa suggests that it was reduced during the Pleistocene to a number of refugia, notably in Cameroon-Gabon, Eastern Democratic Republic of Congo, Upper Guinea, Sierra Leone-Liberia and East Ivory Coast-West Ghana with another site, possibly fragmented near the East African Coast. However, there are differences in opinion as to the exact locations and extent of refugia (Hamilton & Taylor 1992, Coetzee 1993, Morley 2000, Plana, 2004). At around 14,000 B.P. global temperatures rose after the Atlantic monsoon was re-established and African vegetation obtained its present distribution in which tropical rain forest is primarily arranged around the Gulf of Guinea with deserts at the periphery. This pattern is due to the fact that the Atlantic monsoon affects vegetation in the Congo Basin and Central Africa, while the tropical jet stream prevents the northward migration of summer rain into the Sahara region (Coetzee 1993) (Fig. 5.12).

Plana (2004) suggested that “African rain forests (and possibly rain forests worldwide) are an assemblage of relict species from a once widespread Mid-Tertiary (Late Oligocene–

Early Miocene) rain forest and species of recent Plio–Pleistocene origin, born as a result of multiple consecutive glaciations. The Mid–Tertiary element of relict rain forest species is represented by genera or groups of species that are more drought tolerant, ranging from savannah to the more seasonal rain forests of the East African highlands. These groups tend to have fewer West and Central African rain forest representatives, which commonly occupy basal phylogenetic positions. These phylogenies show an increase in drought tolerance among derived taxa, with rare switches back to rain forest."

Faced with aridification caused by cooling climates, wet tropical plant lineages had to either migrate with the climates to which they were adapted, adjust to drier conditions or go extinct. The coast to coast rain forest, which was present during the Eocene may have been fragmented into western (Guineo-Congolia) and eastern (East African) blocks only one time as a result of a cooling and drying climate during the Oligo-Miocene (~33-20 Ma), generating high levels of endemism (Couvreur *et al* 2008, Burgess *et al* 1998, Lovett & Wasser 1993, Morley 2000). Alternatively, rain forest blocks may have expanded and contracted throughout the Oligo-Miocene (~33-2 Ma) following multiple cycles of climate change, resulting in a series of vicariance events (Fig. 5.13) (Couvreur *et al* 2008, Burgess *et al* 1998, Lovett & Wasser 1993, Coetzee 1993). Two recent phylogenetic studies on African taxa exhibit different diversification patterns. Rain forest-adapted genera *Isolona* and *Monodora* (Annonaceae) remained restricted to remnant pockets of wet forest throughout climatic cycles (Couvreur *et al* 2008), whereas *Acridocarpus* (Malpighiaceae) (Davis *et al* 2002) adapted to drier conditions during periods of Oligo-Miocene aridity.

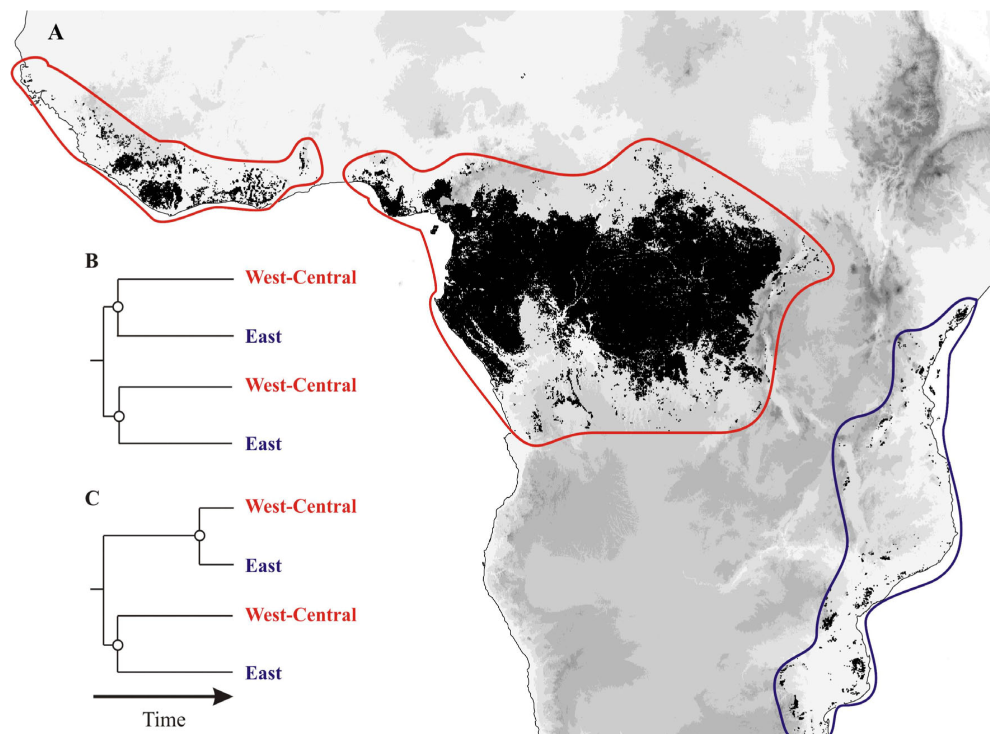


Figure 5.13 Alternative hypotheses of African rain forest origins, excerpted from Couvreur *et al* (2008). (A) Distribution of lowland rain forest in Africa (black) overlaid by altitudinal range (increasing altitude with darker grey). Red lines highlight the Guineo-Congolian region; the blue line highlights the East African region. (B) Phylogenetic tree expected from a single break-up scenario. (C) Phylogenetic tree expected from multiple break-ups at significantly different times scenario. Open circles indicate West-Central/East splits.

Taxa which are restricted to wet forests would be expected to exhibit a pattern of either a single deep phylogenetic split during the Oligocene (30-24 Ma) when aridification initially became prominent across the continent (as shown in Fig. 5.13 B), or multiple splits occurring at different time frames during the Miocene and Pliocene (Fig. 5.13 C) coinciding with climate fluctuation (i.e. the closing of the Tethys sea 16-10 Ma, and the uplift of the East African Rift, 20-13.5 Ma and possibly also 6-4 Ma during the Messinian Salinity Crisis or during Pleistocene glaciations), e.g. Couvreur *et al* 2008 (Fig. 5.14). For those taxa, which have adjusted to drier environments, the timing of diversification into drier biomes would be expected to coincide with these same climatic fluctuations. However, rather than remain restricted to wet forest, adaptation to East African dry scrubland or open woodland forest would have occurred (e.g. Davis *et al* 2002) (Fig. 5.14). Additionally, phylogenetic splits representing Gondwanan vicariance with Madagascar would need to be ~84-66 Ma. Subsequent relationships between the two regions are hypothesized to be due to long distance dispersal.

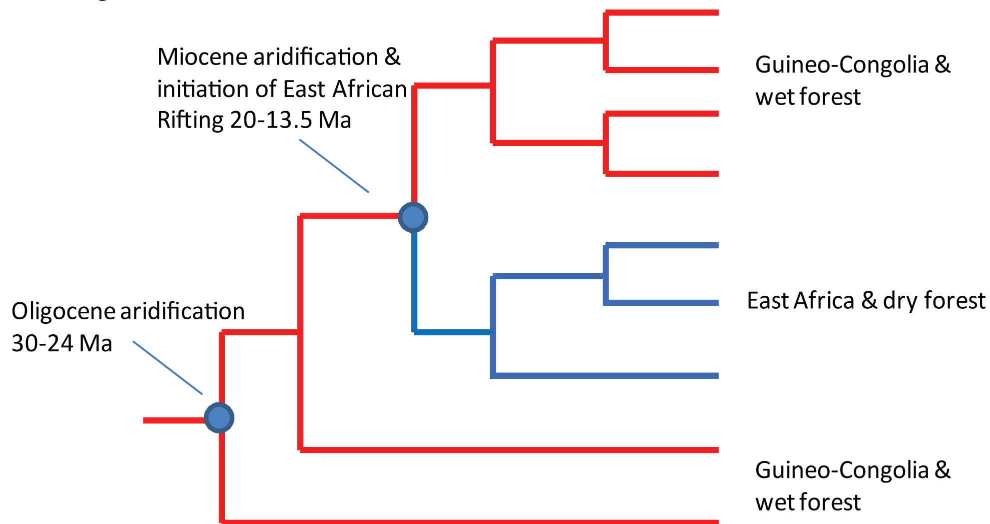


Figure 5.14 Hypothetical area cladogram depicting diversification in African rain forest taxa coinciding with cycles of aridification. Taxa either adapt to a drier biome and radiate therein (i.e. East African dry forest) or remain restricted to pockets of wet forest (i.e. Guineo-Congolia).

5.7.3 Regional history of Asia

5.7.3.1 Southeast Asia

Forest vegetation in Southeast Asia can be categorized into three broad climatic groups: drought-adapted tropical moist forest (monsoon forest), ever-wet tropical rain forest and temperate montane forest. Monsoon forest is distinguished from ever-wet forest in having an average of less than 1,270 mm of rainfall per year and a prolonged period of drought, whereas ever-wet forest receives more than 1,750 mm of rainfall spread evenly throughout the year. Monsoon forest occurs in patches throughout tropical continental Asia and in Java, the Lesser Sunda Islands, Sulawesi, southern New Guinea and northern Australia, whereas ever-wet rainforest stretches across most of Peninsular Malaysia, Sumatra, Borneo, northern Sulawesi and New Guinea. Temperate montane forest occurs in Malesia from 1,000-3,000 meters in New Guinea, northern Borneo, western Sulawesi and western Sumatra (Richardson *et al* 2010) (Fig. 5.15).

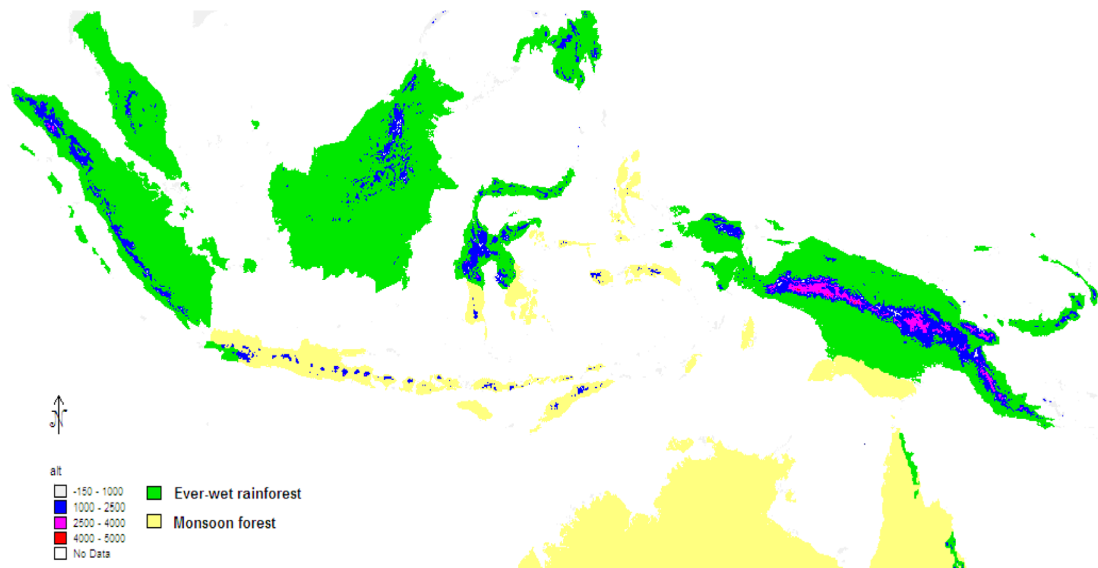


Figure 5.15 Map of forest types in Malesia, excerpted from Richardson *et al* 2010.

The tectonic origins of Southeast Asia are, perhaps, the most complex of any region worldwide. Southeast Asia has, over time, been created by the collision and amalgamation of Gondwanan fragments with Laurasia to form the Sunda shelf, which in turn has been colliding with the Sahul Shelf as Australia–New Guinea migrate northward. These plate movements combined with volcanism and sea-level oscillations have created an ever-changing tapestry of land and sea over the past 60 Ma (Hall 2001, 2009).

During the Early Tertiary, the Sunda shelf formed a wide plateau Sundaland (composed of the Malay Peninsula, Sumatra, Java, Bali, Borneo and now submerged lowland areas between the islands), which was connected to Indochina and East Asia by a broad continuous mountain belt. The surface area of Sundaland was probably greater during the Paleocene than at any time since. When the Indian plate collided with Laurasia during the Middle Eocene (~50-45 Ma) the force of the continuing thrust caused Sundaland to subside, increasing lowland areas. New pollen types occurring in the Javanese Middle Eocene, which are also known from the earlier Palaeocene in India are interpreted as representing a major dispersal event following the collision of the Indian and Laurasian plates (Morley 2000, 2003). Few Southeast Asian taxa dispersed to India during this period and the Indian flora is believed to have been the more successful colonizer of the two since its invasion appears to have resulted in the extinction of many Palaeocene and Early Eocene Southeast Asian taxa as evidenced by the Kayan sandstone formation of Sarawak. Interestingly, Sapotaceae is cited as one of the families with modern affinities which was not wiped out by this Indian invasion (Morley 2000). The role of India as a potential vector for dispersal of Gondwanan taxa to Southeast Asia is discussed in more detail in the following section 5.7.3.2.

By the Middle Eocene, the Sundanese flora reached as far east as Borneo. The southeast arm of Sulawesi was directly connected to Borneo during this time and it has been suggested that the Sundanese flora remained on this fragment of Sulawesi long after the Late Eocene formation of the Makassar Strait (and Wallace's Line) acting as a source of Laurasian taxa for the east Malesian islands (Morley 2000, 2003, Hall 2009, Collinson & Hooker 2003).

Throughout the Late Eocene, Sundaland's climate became cooler and drier. By the Early Oligocene basins had opened up on the sinking Sunda shelf and new ocean-floor was created in the eastern South China Sea, while the thrust of the Indian subcontinent uplifted the Indo-Burman ranges and the Andaman-Nicobar islands. Meanwhile, the Philippines and other Pacific Plate island arcs continued to form ocean ridges and drift west towards Sundaland (Hall 2002).

By the Late Oligocene, the Philippine and Australian plates had begun to collide and the Makassar Strait continued to deepen, forming a major biogeographic barrier in Wallacea (Hall 2002). At this point most of the islands to the east of southwest Sulawesi were either unformed or submerged. Morley (2000) posits that from the Oligocene onwards, the Southeast Asian flora continued to diversify due to rapid geological and environmental change in combination with the dominance of moist, but episodically dry climates.

At the beginning of the Miocene, a large proportion of Sundaland (from Vietnam to Natuna) became submerged following a sudden sea level rise. During this time, rain forests characterized by megathermal angiosperms (including Sapotaceae) extended as far north as the Pearl River in southern China (Morley 2000). Miocene floras from North India also indicate a clear affinity with the Indo-Malayan region (Tiffney & Manchester 2001). Throughout the Miocene, tectonic compression from the south and east by the Pacific and Australian plates resulted in increased flooding in Sundaland.

By the Middle Miocene, the Australian Plate had collided with the Asian Plate forming the Banda Arc and uplifting New Guinea and many of the East Indonesian islands above sea level. The collision of these two plates restricted the Indonesian through-flow resulting in the cooling of the Indian Ocean and warming of the Pacific as well as increased land temperatures in Australia and decreased temperature in northeast Asia. This also dramatically affected rainfall patterns throughout the region (Robert Hall pers. comm. 2009) and is believed to have contributed to the aridification of East Africa (Cane & Molnar 2001). Throughout the Mio-Pliocene, climates were not uniformly moist. The sudden and sporadic appearance of fossil grass pollen suggests that during periods of low sea level, drier episodes supporting more open vegetation intermittently replaced rain forests (Morley 2000).

During the Pleistocene, the climate cooled dramatically resulting in the downward movement of montane forest zones and glacial cycles characterised by dry savannas alternated with moist interglacial cycles during which rain forest expanded. This repeated expansion and contraction of the region's flora into and out of refugia has been suggested as a driver for the diversification of Southeast Asia's flora (Morley 2007). Yet Morley (2000) cautions that in Southeast Asia the refuge theory also needs to take into account the successive drowning and re-exposure of the continental shelves due to an oscillating climate (Voris 2000). He also adds that the time scale for the creation of refugia in Southeast Asia needs to be considered on a broader scale – not just the Quaternary, but over the whole Neogene. This theory is supported by evidence from pollen assemblages, which show that biodiversity accumulated uniformly throughout the Neogene and speciation is likely the result of this continued expansion, contraction and mixing spurred by climate and sea-level changes.

Figure 5.16 Wallace's, Weber's and Lydekker's Lines which divide faunal distributions in Wallacea. Excerpted from Roberts & Motes 2009.

Alfred Russell Wallace (1860, 1863, 1869, 1876) identified a major faunistic divide between the islands of Borneo/Bali and Sulawesi/Lombok, correlating with the Makassar Strait, a deep-water trench separating the Sahul and Sunda shelves. He was the first to propose that this biogeographical division was the result of geological history. Modifications have been made to Wallace's Line by other authors: Huxley's Line (Huxley 1868), Weber's Line (Pilsneer 1904) and Lydekker's Line (Lydekker 1896) (Fig. 5.16), but all concur that this region, known as Wallacea, is a filter to taxa between the Sahul and Sunda shelves (Simpson 1977). The importance of this biogeographical divide for plants has also been the subject of conjecture (Turner *et al* 2001, van Welzen *et al* 2005) and has only recently begun to be studied with molecular phylogenetic tools. Current studies suggest that Wallace's Line is not a strong barrier for the migration of plants (e.g. dispersals across Wallace's Line have occurred in: *Cyrtandra*, Cronk *et al* 2005; *Garcinia*, Saleh 2006; Aglaieae, Muellner *et al* 2008; *Pseuduvaria*, Su & Saunders 2009; *Bridelia*, Li *et al* 2009; *Diospyros*, Duangjai *et al* 2009; *Begonia*, Thomas 2010; *Etlingera*, A. Poulsen pers. comm. 2010; *Rhododendron* section Vireya, Twyford & Richardson in prep.).

Dated phylogenies indicate that many taxa have migrated from west to east during the Miocene as the Asian and Australian plates converged and land became available in New Guinea, but seemingly fewer taxa have migrated from east to west. This is possibly due to phylogenetic niche conservatism amongst dry adapted Australian taxa and the deficit of similarly dry biomes in the predominantly wet Sunda shelf region (Richardson *et al* 2010). Examples of taxa which exhibit a west to east migration pattern in phylogenetic studies are: *Pseuduvaria*, Annonaceae (Su & Saunders 2009); Aglaieae, Meliaceae (Muellner *et al* 2008); at least four separate lineages of *Begonia*, Begoniaceae (Thomas 2010); *Rhododendron*, Ericaceae (Twyford & Richardson in prep, Brown 2006); *Cyrtandra* Gesneriaceae (Atkins *et al* 2001, Cronk *et al* 2005); *Etlingera*, Zingiberaceae (A. Poulsen pers. comm. 2010); *Garcinia*, Clusiaceae (Saleh 2006), and six lineages of Isonandreae, Sapotaceae (Bakar 2009). Four of these studies also demonstrated evidence of back dispersal to the west.

Although dispersal from the Sahul to the Sunda shelf does not appear to be as common, this pattern is exemplified in Myrtaceae (Sytsma *et al* 2004), Proteaceae (Barker *et al* 2007) and Chrysophylloideae, Sapotaceae (J. Richardson pers. comm. 2010).

In relation to the geological history of Southeast Asia, taxa are predicted to disperse across Wallace's Line during the Miocene, when the Sahul and Sunda shelves came into close proximity. This would have been aided by the emergence of New Guinea from ~15-10 Ma. If Sahul shelf taxa are nested within clades from the Sunda shelf, this would support west to east movement. Likewise, if Sundanese taxa from the west are nested within an eastern Sahul shelf clade, this would support east to west migration (Fig. 5.17). An alternative theory posited by Morley (2003), states that Eocene age fossil pollen assemblages recorded from the southwest arm of Sulawesi show a distinct affinity with Sundanese taxa. This is due to the fact that Southwestern Sulawesi was connected to Borneo during the Mid-Eocene and after splitting, drifted east to become amalgamated with other fragments in Wallacea, and hypothetically could have acted as a raft for Eocene age taxa across Wallace's Line. Phylogenetic splits, concordant with this theory would be expected between Sunda and Sahul shelf taxa at ~45-40 Ma (Fig. 5.17).

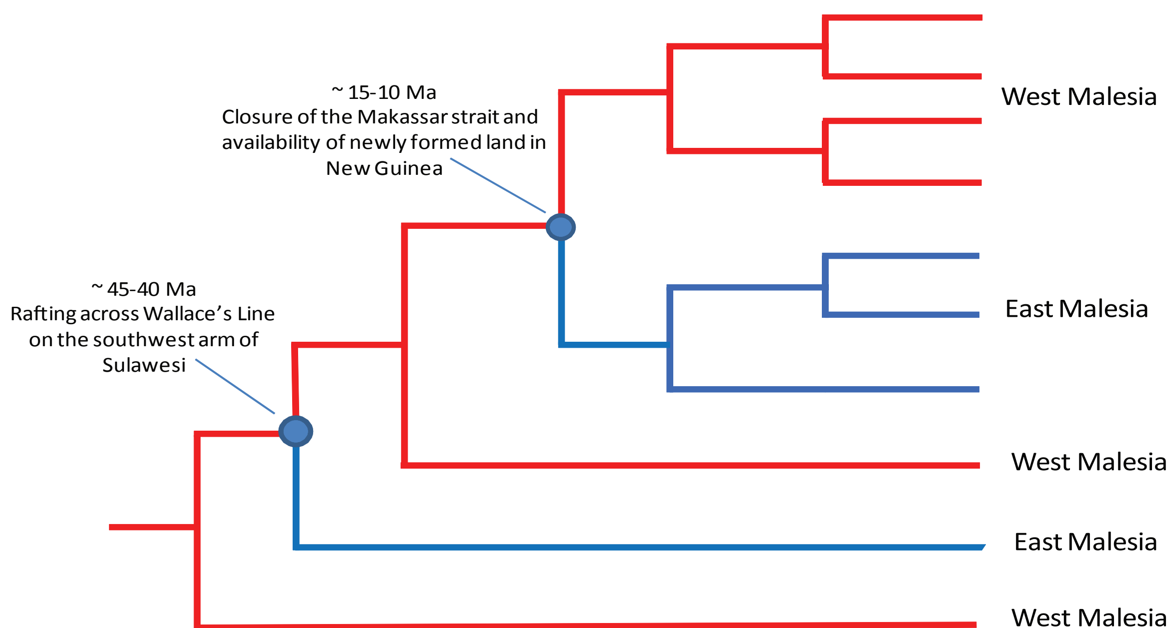


Figure 5.17 Area cladogram depicting hypothetical dispersal across Wallace's Line in accordance with two paleogeological scenarios: rafting on the southwest arm of Sulawesi and after the closure of the Makassar Strait and the emergence of New Guinea.

5.7.3.2 Indian subcontinent

The role of the Indian subcontinent in rafting Gondwanan taxa to Laurasia has been hotly debated (Lakhanpal 1970, Briggs 1987, 1989, Thewissen & McKenna, 1992, Rage 1995, McKenna 1995, Karnath 2006, Ali & Aitchison 2008, Briggs 2003, Datta-Roy & Karnath 2009). It has been hypothesized that the Indian flora had been weakened and impoverished by extensive volcanic activity and prolonged isolation and therefore would not have contributed to the proto-Southeast Asian flora (Briggs 1987). This was primarily attributed to

the vast Deccan trap lava flows which erupted during the Paleocene (68-60 Ma) and covered over one million cubic kilometres, about one-fourth of the Indian subcontinent, making the region uninhabitable. It has also been suggested that the rate at which India travelled north across the equator was so great, that its climate changed drastically over a relatively short period of time, causing mass extinctions of a flora that could not cope with such a rapidly changing environment (Briggs, 1987).

Those theories are beginning to be overturned with fossil evidence and dated phylogenies. Although India would have been located too far south during the initial radiation of the angiosperms to have hosted a significant angiosperm flora, by the mid Cretaceous (~95-84 Ma) India rafted close enough to Africa that some taxa were able to disperse to the subcontinent via Madagascar (Morley 2000, 2003). This is evidenced by Indian fossil taxa from this period, the nearest relatives of which are African (Awasthi & Mehrota 1993). In a recent paper, Prasad *et al* (2009) state that "rich palynofloral assemblages from coal and lignite bearing sedimentary successions of [the] western and northeastern Indian region show [the] existence of [a] well diversified and widely distributed tropical rain forest community during [the] late Palaeocene-early Eocene time interval." Furthermore, the fossil record also provides strong evidence for dispersal of novel angiosperm pollen types from India into Southeast Asia by the Middle Eocene, including numerous palms, Bombacaceae, Sapindaceae, Thymeleaceae, Proteaceae, Linaceae, Olacaceae, Polygalaceae and Ctenolophonaceae (Morley 1998, 2000, Lelono 2000, Collinson & Hooker 2003). Additionally, a chemical signature for Dipterocarpaceae resin has been found in oil deposits (Curiale *et al* 1994, van Aarssen *et al* 1990), which track the family's movement from India to Southeast Asia (Myanmar) (Morley 2003). During the Mid-Eocene (50-39 Ma) when the Indian plate collided with Laurasia, both India and Southeast Asia would have been positioned at similar latitudes and within the same climatic zone, making an intercontinental floristic transfer straightforward (Morley 2003).

There are very few remaining wet-forest refugia on the Indian subcontinent and according to Morley (2003) this may be the reason why it has taken so long for the scientific community to realise the importance of the Indian plate as a vector for the dispersal of tropical taxa from Africa to Southeast Asia. To date only a few molecular studies on plants (Table 5.7) have corroborated evidence of this pathway, most notably Crypteroniaceae (Conti *et al* 2002, 2004).

Divergence times between lineages in a phylogeny representing dispersal/migration between Africa, Madagascar and India should be evident from ~84-65 Ma, marking the time when Madagascar and India split (Ali & Aitchison 2008). Whether there were other vectors for stepping-stone migration between the continental fragments after this date such as the Chagos-Laccadive Plateau and the Mascarene Plateau remains debatable (Schatz 1996). A possible phylogenetic signal for migration/dispersal from India and into Southeast Asia or vice versa should be indicated by splits between lineages in these regions from ~50 Ma when the Indian subcontinent began to make contact with Laurasia. These phylogenetic patterns are illustrated in Fig 5.18.

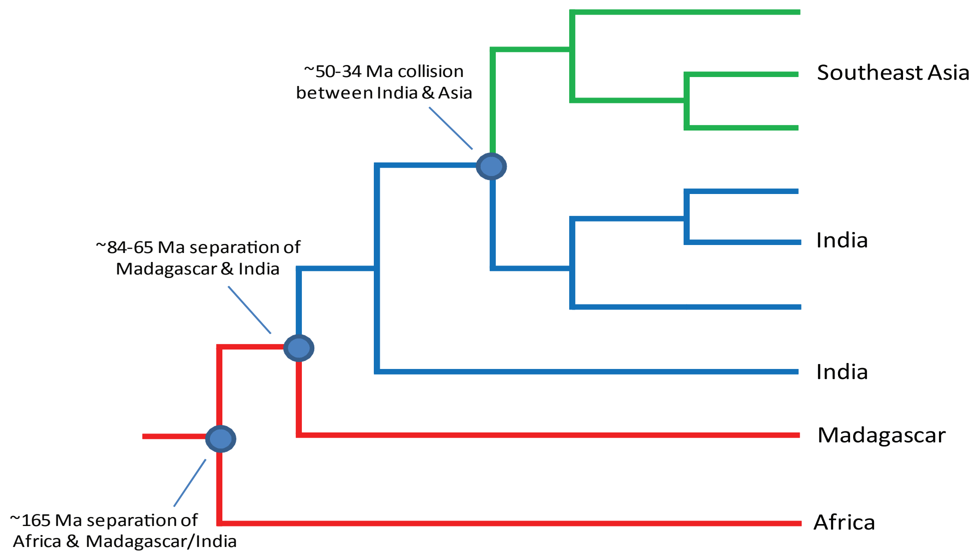


Figure 5.18 Area cladogram depicting a scenario for the migration of taxa from Gondwana to Southeast Asia via Indian rafting . (1) Separation of Madagascar/India from Africa ~165 Ma, (2) separation of India from Madagascar ~95-80 Ma, (3) collision of Indian plate with Asia ~50-34 Ma.

Table 5.7 Examples of (predominantly dated) phylogenetic studies of taxa which conform to a scenario of dispersal into Asia following Indian rafting.

Family/Taxon	Inferred biogeographical history	Reference
Crypteroniaceae *Axinandra, Crypteronia* & *Dactylocladus*	Cretaceous Gondwanan origin, rafting on India and subsequent migration to Southeast Asia. Although according to Morley (2000) pollen first appeared earlier in Australia than India suggesting a southerly Gondwanan route.	Conti *et al* 2002, 2004, Rutschmann *et al* 2004
Dipterocarpaceae (Asian taxa)	Cretaceous Gondwanan origin, rafting on India and mid-Eocene migration to Southeast Asia. (phylogeny not dated, but based on DNA evidence, mycorrhizal relationships, fossil pollen & resin in oil deposits)	Dayanandan *et al* 1999, Ducousso *et al* 2004, Morley 2000
Lauraceae *Beilschmedia* & *Cryptocarya*	Gondwanan migration plus possible rafting on India to Southeast Asia	Chanderbali *et al* 2001
Annonaceae MPM-inaperturate clade	Origin in Africa/Madagascar followed by possible rafting on India to Laurasia	Doyle *et al* 2004, Richardson *et al* 2004
Nepenthaceae *Nepenthes*	Tertiary origin with possible migration through the boreotropics to the Indian Ocean Basin or possible rafting on India to Southeast Asia and subsequent Miocene expansion in Asia eastwards to New Guinea – phylogeny not dated	Meimberg *et al* 2001

Chapter VI - The Sapotaceae fossil record

6.1 Introduction

Prior to the advent of dated molecular phylogenies, the sole method of reconstructing the biogeographic history of a taxon was through an examination of the fossil record. Yet, since it is incomplete, the fossil record can only provide part of the story. Now, reliably identified fossils are also a valuable resource in the calibration of phylogenies, enabling the incorporation of time into molecular studies.

In order to choose appropriate fossils for temporal calibration of the phylogeny in Chapter VII, a survey of Sapotaceae fossil specimens held in museums and published in the literature was conducted. This collated data is also an aid to understanding the historical distribution of the family, the patterns of which, can then be compared to those reconstructed in the phylogeny as an additional validation for or against biogeographic hypotheses. The identification of fossil taxa relevant to this study, specifically those in the tribe Mimusopeae, was checked by Terry Pennington (RBG Kew), who determined the macrofossils and by Madeline Harley, (RBG Kew) who determined the fossil pollen. These identifications, where carried out, are noted in the tables in Appendices 6.1, 6.4, 6.5, 6.6 and 6.7. An overview of the fossil history of the entire family Sapotaceae is presented below in section 6.2, and in the following section, 6.3, fossils placed specifically in the tribe Mimusopeae are highlighted. Three of the surveyed fossils are utilized in dating the phylogeny in Chapter VII.

6.2 Overview of the Sapotaceae fossil record at family level

6.2.1 Cretaceous-Paleocene

A review of the Sapotaceae fossil record points to the early occurrence of the family in both Laurasia and Gondwana. Raven & Axelrod (1974) postulated a West Gondwanan origin for the family before the end of the Cretaceous. If identifications are to be trusted, the earliest recorded Sapotaceae fossil is *Sapotaceoidaepollenites robustus* pollen from the Cretaceous (Senonian)/Paleocene of Borneo (Muller 1968 & 1981). However, Harley (1991) cautions that it is unlikely to be sapotaceous. Likewise, *Sapotaceoidaepollenites occultus* and *S. manifestus* have been recorded from the Senonian of China (Song *et al* 1999 & 2004), but Harley (pers. comm. 2010) believes that these are also not Sapotaceae. Conversely, fossil pollen grains from the Cretaceous/Paleocene of Australia, *Sapotaceoidaepollenites rotundus* (Stoin 2002), are confirmed by Harley (pers. comm. 2010) to be sapotaceous and possibly placed in the Tribes Mimusopeae or Isonandreae. See Figure 6.1.

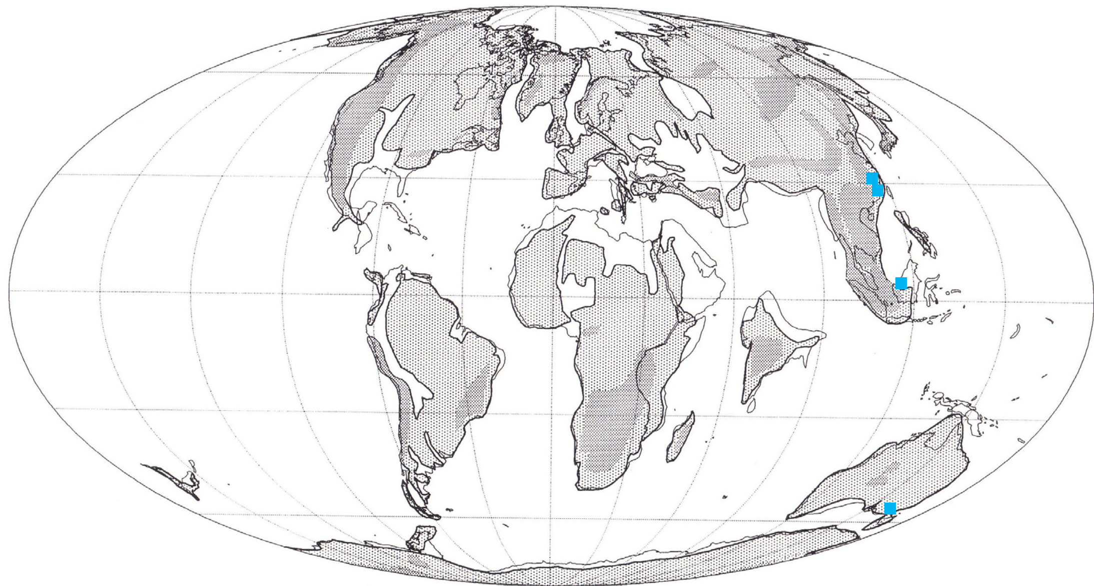


Figure 6.1 Sapotaceae fossils recorded from the Cretaceous-Paleocene. Blue squares denote the locality of cross-checked fossil taxa.

6.2.2 Eocene

By the Eocene, Sapotaceae fossils become more common, particularly in the northern hemisphere. In the United States numerous Sapotaceae pollen taxa are recorded in various localities along the east coast and around the Gulf of Mexico (Taylor 1989; Frederiksen 1980b & 1988; Edson *et al* 2000). *Bumelia* seeds are found in the Clarno Nut Beds of Oregon (Manchester 1994), the seed *Eoachras eocenica* from Mississippi (Berry 1915) and leaves of *Mimusops*, *Bumelia, Sideroxylon, Chrysophyllum* and *Sapotacites* are recorded from Mississippi, Arkansas, Tennessee, Louisiana and Wyoming (Berry 1916; 1924; 1930; Hollick 1899) as well as *Chrysophyllum* and *Lucuma* from Oregon and California (Chaney & Sandborn 1933; Potbury 1935). In the U.K., the Eocene London Clay flora has been made famous by its abundance of fossil fruit forms. Representatives from the Sapotaceae include five taxa (Reid & Chandler 1933) identified by Pennington (Appendix 6.5) as belonging to the genus *Chrysophyllum*. *Tieghemella* (Harley 1991) and other Sapotaceae pollen (Gruas-Cavagnetto 1976) is recorded from southern England as well as possible Sapotaceae wood from East Anglia (Crawley 2001). A summary of Paleocene-Eocene sapotaceous fossil pollen from across France, Germany and Hungary is diagrammed by Kedves (1967). Although these grains occur sporadically, they are broadly dispersed, suggesting a cosmopolitan distribution for the family throughout Europe during this period. Seven different Sapotaceae pollen taxa are recorded from the Eocene of Turkey (Akgün et al 2002). Further east, five pollen taxa are recorded from China (Song *et al* 1999 & 2004), most of which Harley suggests (pers. comm. 2010) are unlikely to be sapotaceous – except for *Sapotaceoidaepollenites sapotoides*, which is probably a member of the Sapotaceae. There are many fewer Eocene fossil records for the family from the equatorial regions and southern hemisphere, but *Malacantha* pollen is recorded from Nigeria (Jan du Chêne *et al* 1978), possible *Chrysophyllum* pollen (Jaramillo & Dilcher 2001, Rull 2000) from Colombia and *Sapotaceoidaepollenites rotundus* pollen (Harris 1972; Stover & Partridge 1973) from Australia. See Figure 6.2.

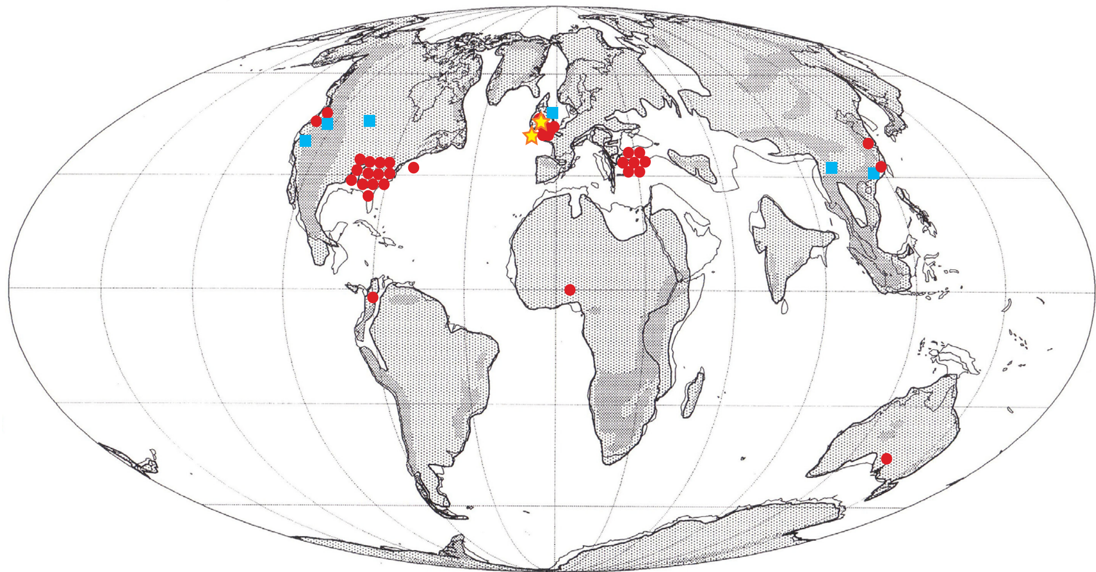


Figure 6.2 Sapotaceae fossils recorded from the Eocene. Blue squares denote the locality of cross-checked fossil taxa. Red circles denote the locality of fossil taxa which have not been cross-checked. Yellow stars denote the locality of cross-checked fossil taxa used in molecular dating analyses in Chapter VII.

6.2.3 Oligocene

The pattern of fossil distribution remains similar throughout the Oligocene. In the eastern United States, *Mimusops/Manilkara* pollen, fruit, wood and possible flowers are recorded from Vermont (Travese 1953, 1955) and along the gulf coast, pollen of *Bumelia*, *Chrysophyllum* and possible *Manilkara* (Frederiksen 1980a) have been documented. In the U.K., Mimusopeae pollen is recorded from the Isle of Wight (Machin 1971). Across Europe, specimens of cf. *Manilkara* wood from Czechoslovakia (Prakash *et al* 1974), pollen from Bulgaria (Ivanov *et al* 2002, 2007) and Turkey (Akgun *et al* 2007) have been collected. In the equatorial regions, *Manilkara/Tieghemella* leaves are recorded from Ethiopia (Jacobs *et al* 2005, Pan pers. comm. 2010) and *Sideroxylon* and *Chrysophyllum* leaves (Hollick 1928) and *Chrysophyllum* pollen (Graham & Jarzen 1969) are recorded from Puerto Rico. Additionally, an increase in Sapotaceae pollen in Egypt (Kedeves 1971) at this time is suggested by Morley (2000) to indicate immigration of taxa from Europe into Africa. See Figure 6.3.

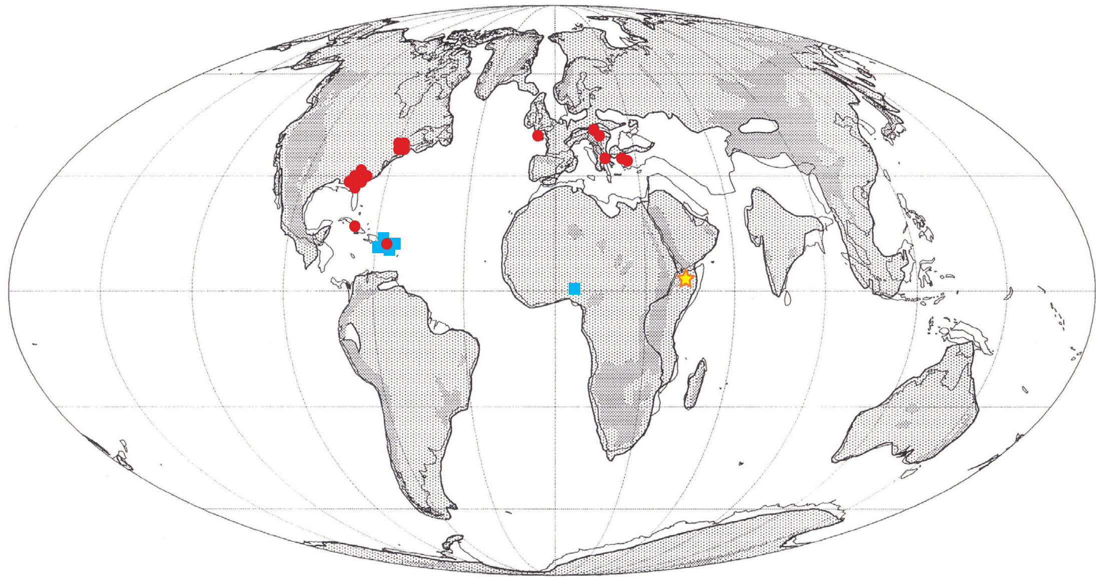


Figure 6.3 Sapotaceae fossils recorded from the Oligocene. Blue squares denote the locality of cross-checked fossil taxa. Red circles denote the locality of fossil taxa which have not been cross-checked. Yellow stars denote the locality of cross-checked fossil taxa used in molecular dating analyses in Chapter VII.

6.2.4 Miocene

Fossils suggest that by the Miocene, Sapotaceae had begun to migrate to the equatorial region due to a cooling climate. In the United States fossil *Bumelia* leaves are recorded from California (Axelrod 1939, 1950) and Nevada (Axelrod 1956). In Europe petrified wood is found in Italy (*Argania* - Biondi 1981), France (*Manilkara* and other unidentified Sapotaceae - Grambast-Fessard 1968), and Germany (*Bumelia* - Selmeier 1991, Gottwald 2004). Pollen is recorded from Austria (Kovar-Eder *et al* 2001), Turkey (Akgün *et al* 2000), China (Song *et al* 1999, 2004) and the Niger Delta (Legoux 1978). Sapotaceae pollen also occurs in Early Miocene pollen assemblages off the coast of Mexico with a marked increase in abundance after the Middle Miocene (Morley 2000). In India there is a sudden prolific occurrence of fossil woods: cf. *Madhuca* and *Manilkara* (Awasthi & Mehrota 1993), cf. *Sideroxylon* (Prakash & Awasthi 1970), cf. *Payena-Palaquium* (Awasthi & Srivastava 1990), *Mimusops/Bassia* (Lakshmanan & Levy 1956; Navale 1973), *Chrysophyllum* (Awasthi 1977) and *Madhuca* (Prakash & Tripathi 1977). Fossil leaves are recorded from Sumatra (Kräusel 1929) and New Zealand (cf. *Pouteria* - Campbell 2002) and pollen from New Guinea (Playford 1982). In Africa petrified woods are recorded from Congo (*Tridesmostemmon* and *Chrysophyllum* - Bande *et al* 1987) and Ethiopia (unidentified Sapotaceae - Prakash *et al* 1982, Wheeler *et al* 2007; *Sideroxylon* Lemoigne *et al* 1974; *Synsepalum/Lecomtedoxa* Lemoigne 1978). In Central and South America, leaf fossils are recorded from: Trinidad (*Mimusops* - Berry 1925 and *Sapota* - Hollick 1928), Haiti (*Mimusops* – Berry 1922), Cuba (*Bumelia* and *Mimusops* - Berry 1939) and Venezuela (*Achras* – Berry 1936). See Figure 6.4.

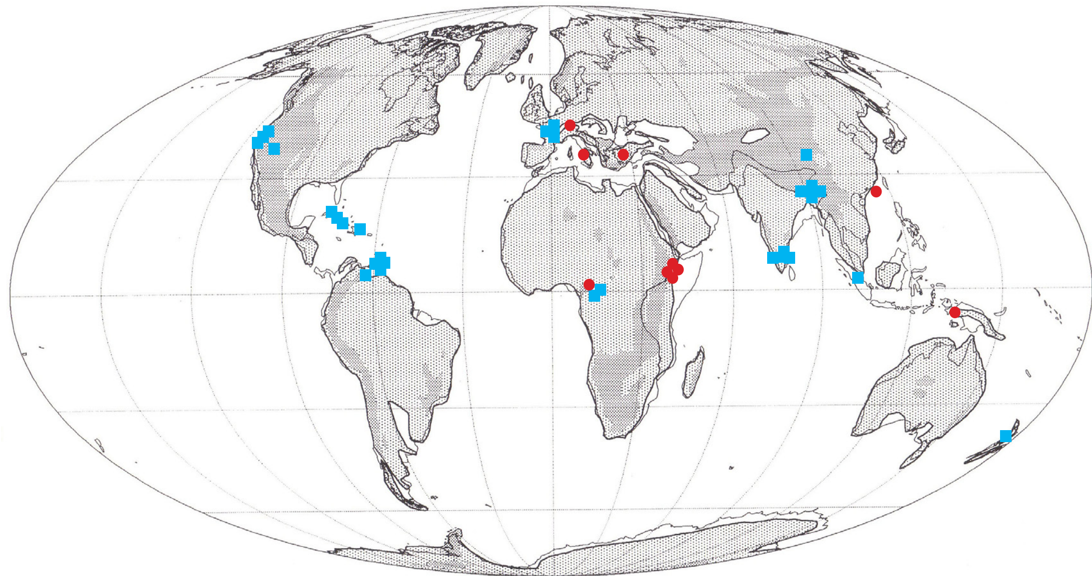


Figure 6.4 Sapotaceae fossils recorded from the Miocene. Blue squares denote the locality of cross-checked fossil taxa. Red circles denote the locality of fossil taxa which have not been cross-checked.

6.2.5 Pleistocene

A few fossils have also been recorded from the Pleistocene and particularly in the case of *Mimusops emarginata* fruit found in Cuba (Berry 1934), the preservation is excellent. *Mimusops preduplicata* leaves from Trinidad (Berry 1924) are also well-preserved and are referable to modern *Manilkara/Mimusops*. However, these fossils are unfortunately, not useful for molecular dating purposes because they are too young. *Pouteria* leaves, fruit, seeds, wood and pollen have been recorded from La Selva Biological Station in Costa Rica (Horn *et al* 2003), *Madhuca* leaves from India (Bande & Srivastava 1990) and *Bumelia* leaves from Maryland in the United States (Hollick 1907) have also been recorded. See Figure 6.5.

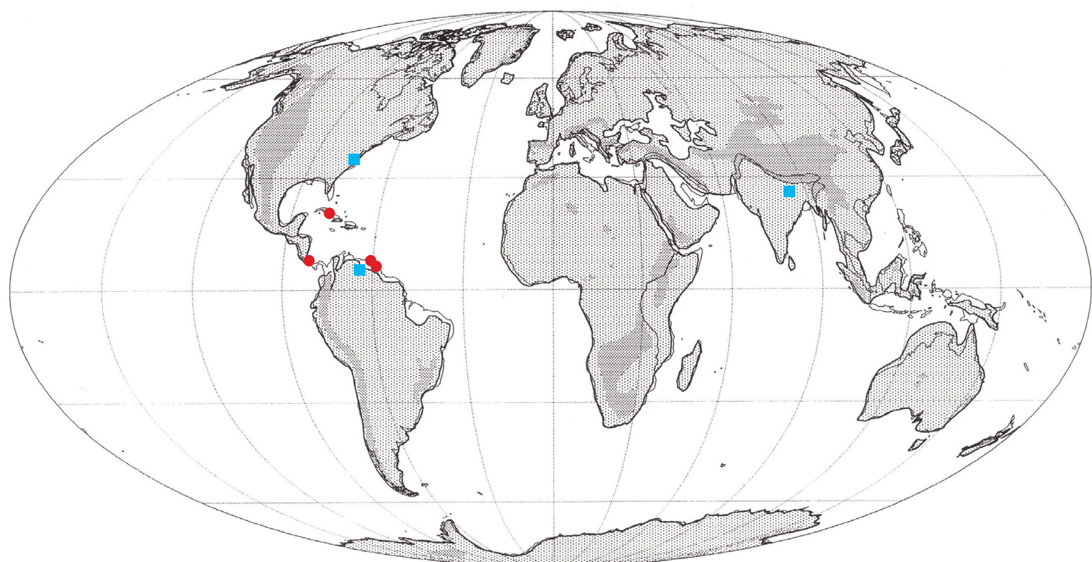


Figure 6.5 Sapotaceae fossils recorded from the Pleistocene. Blue squares denote the locality of cross-checked fossil taxa. Red circles denote the locality of fossil taxa which have not been cross-checked.

6.2.6 Summary of Sapotaceae fossil records

Not all of these records have been verified, but a number of them, which have been taxonomically scrutinized, suggest that the trend is real and that Sapotaceae may have evolved during the late Cretaceous/Early Paleocene with reliable fossils first appearing during the Late Paleocene/Early Eocene. By the Eocene, Sapotaceae fossils were relatively abundant in the northern hemisphere (particularly in the eastern United States and Europe as well as some in China) with only a few records from equatorial regions (Colombia and Nigeria). While this may be due to collection bias, it appears that Sapotaceae flourished in the northern hemisphere and was already a widespread, pantropical family during the time of the Eocene thermal maximum (~55 Mya). Additionally, pollen similar to all major modern groups was present in the fossil record by this point (Harley in Pennington 1991). The Oligocene trend is similar, with numerous fossils in North America and Europe as well as a couple in the Caribbean (Puerto Rico) and North Africa (Ethiopia). Thereafter, in the Miocene, Sapotaceae appears to have spread southwards. In addition to its northern hemisphere enclave, fossils begin to appear in India, Indonesia, New Guinea, New Zealand and more extensively in Ethiopia and Congo. Although, due to collection bias, it is not possible to concretely determine the history of the Sapotaceae based on the fossil record, the verified fossil data we do possess suggests that the family migrated from the northern to the southern hemisphere as the climate cooled during the Oligocene and new land connections became available (i.e. India's contact with Laurasia, Africa's contact with Europe and the close proximity of South America with Central/North America).

6.3 Fossils from the tribe Mimusopeae

6.3.1 Cretaceous-Paleocene and Eocene

Some of the earliest records of Sapotaceae fossils are of *Sapotaceoidaepollenites rotundus* from the Cretaceous/Paleocene of Australia (Stoian 2002), which Harley (pers. comm. 2010) placed possibly in the Mimusopeae or Isonandreae (Fig. 6.6). During the Eocene, Mimusopeae fossils become abundant in the southeastern United States; *Tetracolporpollenites brevis* (Taylor 1989) and possible *Manilkara* pollen (Frederiksen 1980a) are recorded, as well as numerous putative *Mimusops* leaf fossils (Berry 1915, 1916, 1924, 1930). Additionally *Tetracolporpollenites sp.*, which Harley (1991) compares with modern *Tieghemella heckelii* pollen is reported from the Isle of Wight, England. See Figure 6.7.

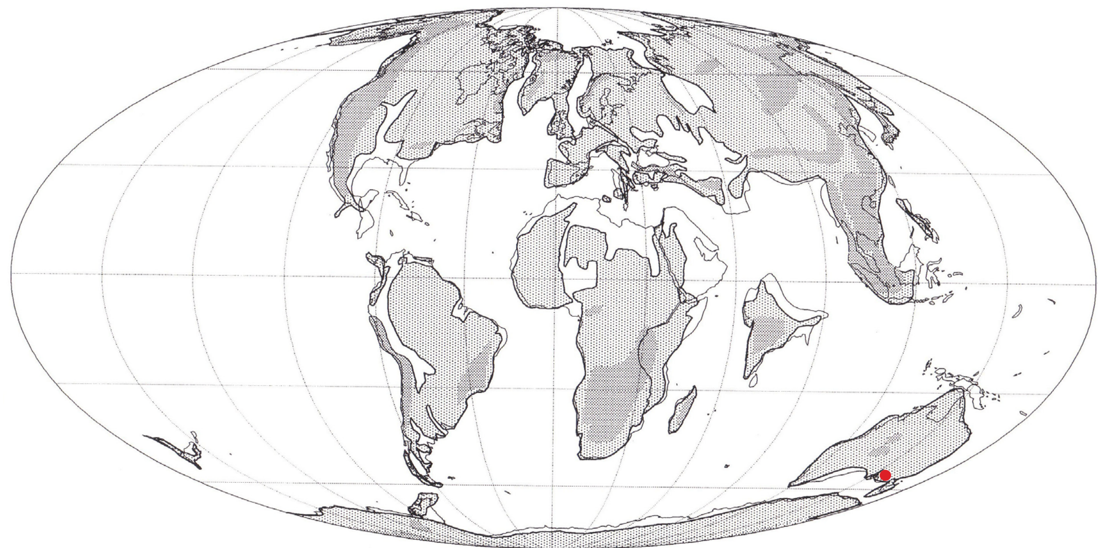


Figure 6.6 Mimusopeae fossils recorded from the Cretaceous-Paleocene. Red circles denote the locality of fossil taxa which have not been cross-checked.

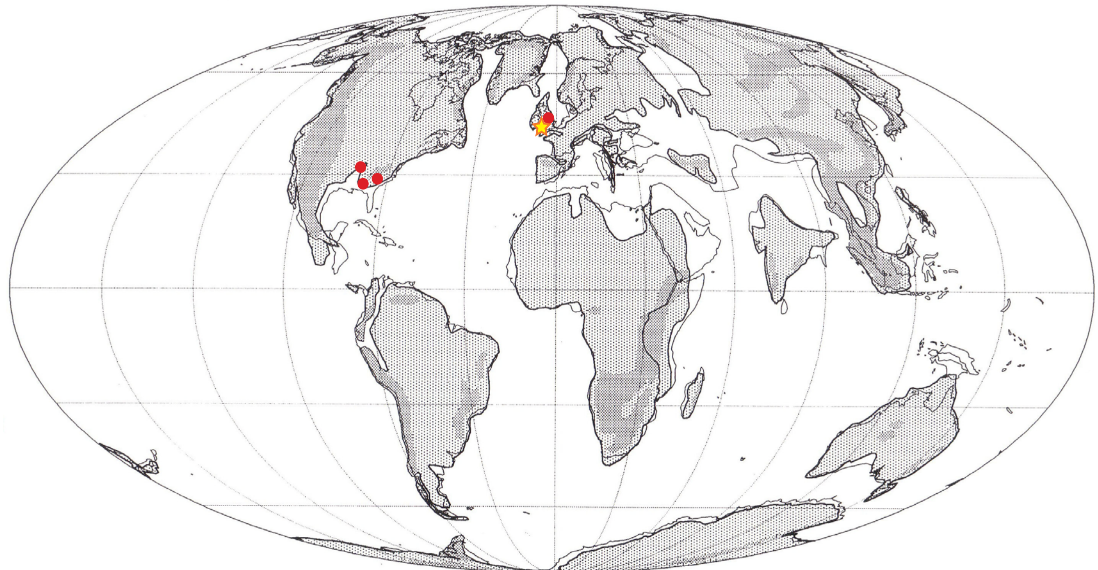


Figure 6.7 Mimusopeae fossils recorded from the Eocene. Red circles denote the locality of fossil taxa which have not been cross-checked. Yellow stars denote the locality of cross-checked fossil taxa used in molecular dating analyses in Chapter VII.

6.3.2 Oligocene

Mimusopeae pollen is still recorded from the Isle of Wight (Machin 1971) during the Oligocene and in the Eastern United States *Manilkara brevipollinia, Manilkara lesquereuxiana, Manilkara longipollinia, Mimusops mirabilis* appear in the Brandon Lignite of Vermont (Traverse 1953 & 1955). Oligocene petrified wood is recorded in Eastern Europe (*Manilkaroxylon bohemicum*, Czechoslovakia, Prakash, Brezinova & Awasthi 1974) and *Sapoteae* sp. leaves resembling *Manilkara or Tieghemella* were discovered in Ethiopia by Jacobs *et al* (2005). See Figure 6.8.

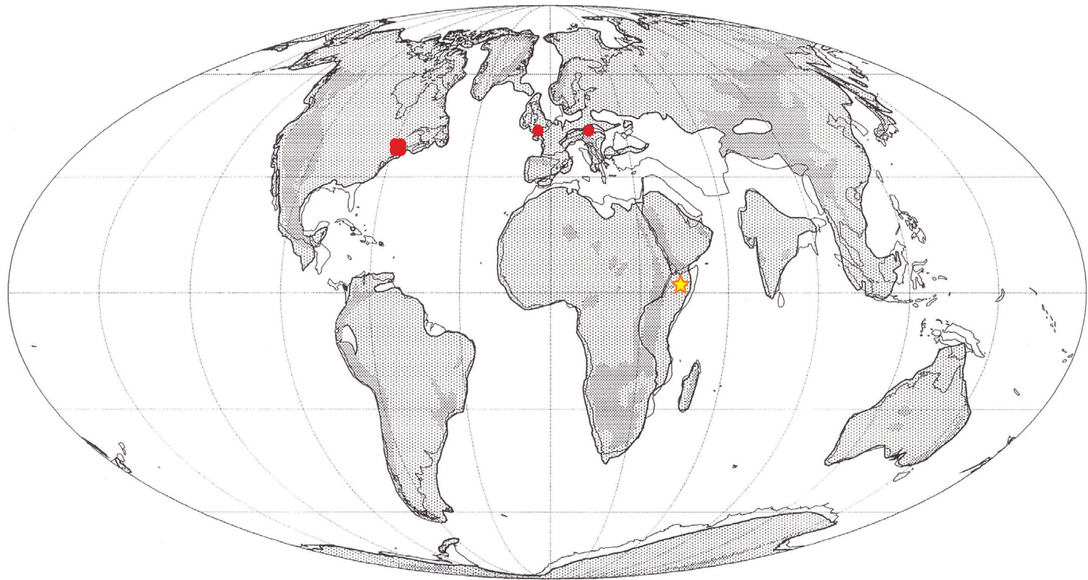


Figure 6.8 Mimusopeae fossils recorded from the Oligocene. Red circles denote the locality of fossil taxa which have not been cross-checked. Yellow stars denote the locality of cross-checked fossil taxa used in molecular dating analyses in Chapter VII.

6.3.3 Miocene and Pleistocene

While the occasional Mimusopeae fossil is still recorded from the northern hemisphere during the Miocene (*Manilkaroxylon crystallophora* wood, France, Grambast-Fessard 1968) there appears to be a geographical shift to the south with the tribe becoming more common at lower latitudes in the modern tropics (Fig. 6.9). Taxa and localities include:

- *Psilastephanocolporites perforates* pollen: compares with *Vitellaria* Cameroon (Salard-Cheboldaeff 1978, 1979, 1981).
- *Sapotaceoidaepollenites kirchheimeri* pollen: possible Mimusopeae or Isonandreae according to Harley (pers. comm. 2010), South China Sea (Song *et al* 1999 & 2004)
- *Manilkara cacharense* wood: compares with cf. *Manilkara hexandra* & *M. littoralis* Northeast India (Awasthi & Mehrota 1993)
- *Mimusops* (or *Bassia*) wood: India (Lakshmanan & Levy 1956; Navale 1973)
- *Sapotoxylon multiporosum* wood: compares with *Mimusops*, *Manilkara*, *Payena*, *Englerophytum* or *Synsepalum*, Ethiopia (Prakash *et al* 1982)
- Numerous putative *Mimusops* leaf fossils recorded from the Caribbean (Haiti, Trinidad & Cuba) as well as Venezuela (Berry 1922, 1925, 1937, 1939; Hollick 1928)

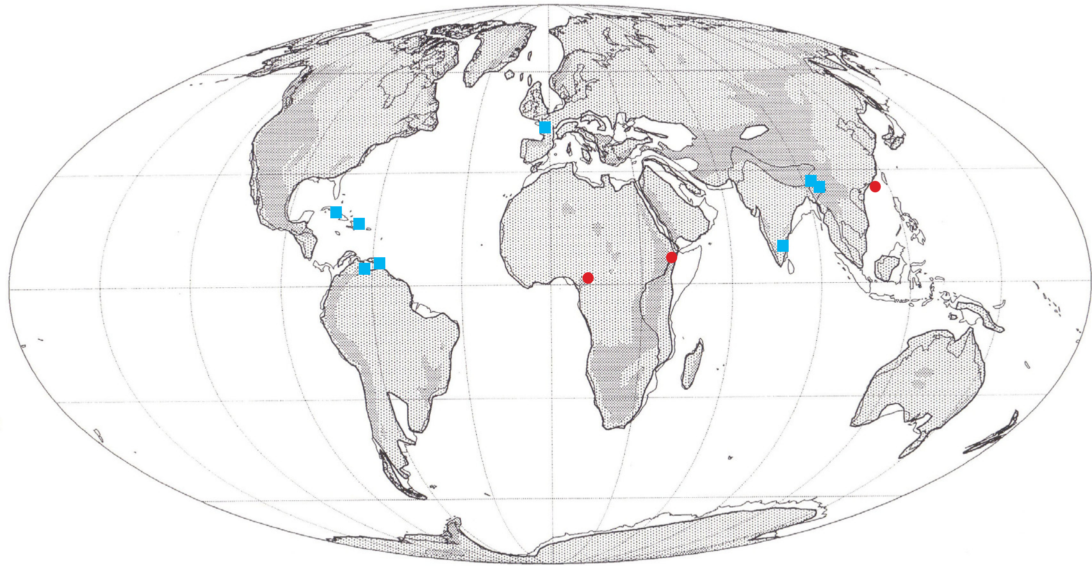


Figure 6.9 Mimusopeae fossils recorded from the Miocene. Blue squares denote the locality of cross-checked fossil taxa. Red circles denote the locality of fossil taxa which have not been cross-checked.

Well-preserved Pleistocene *Manilkara* fossils have been recorded from the Caribbean including *Manilkara jaimiqui* fruit from Cuba (Berry 1934) and *Mimusops preduplicata* leaves from Trinidad (Berry 1925) and Quarternary wood of *Manilkaroxylon diluviale* is recorded from Ecuador (Hofmann 1948). See Figure 6.10.

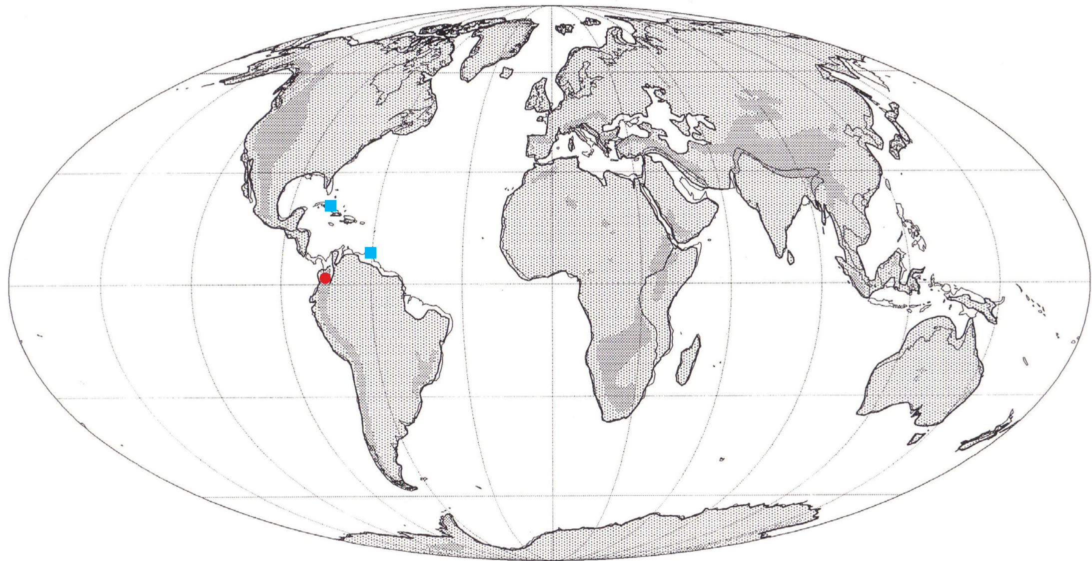


Figure 6.10 Mimusopeae fossils recorded from the Pleistocene. Blue squares denote the locality of cross-checked fossil taxa. Red circles denote the locality of fossil taxa which have not been cross-checked.

6.3.4 Summary of Mimusopeae fossil records

Although some identifications are questionable (particularly those of the Berry leaf specimens), Mimusopeae fossils broadly tell the same story as the rest of the Sapotaceae fossil record, occurring in the Eastern United States and Western Europe during the Paleocene-Eocene Thermal Maximum and beginning to spread to lower latitudes during the Oligocene with fossils predominantly in the modern tropics (Africa/India) by the Miocene.

Because there is significant overlap in Sapotaceae pollen types, Harley errs on the side of caution when identifying fossil pollen (pers. comm. 2010). Therefore, often one cannot be certain that a pollen specimen is exclusively Mimusopeae, because some Isonandreae pollen is similar (Harley 1991). Likewise, many of the putative *Mimusops* leaf fossils collected by Berry and Hollick (and deposited in the Smithsonian Institute and the Peabody Museum) are not detailed enough for reliable identification and so their use in molecular dating should be approached with great caution.

In the appendices below are tables of all the fossil records (pollen, wood, leaf, fruit and seed) for the Sapotaceae encountered during this study. This is a preliminary list and many entries have not been cross-checked for identification, but it is hoped that it will provide the basis for further investigation into suitable Sapotaceae fossils for use in phylogeny calibration and provide a better understanding of what is currently known about the fossil history of the family.

Records have been organised into two basic categories. Appendices 6.1, 6.2, 6.3, and 6.4 are based on references from the literature only, whereas Appendices 6.5, 6.6, 6.7 and 6.8 are composed of fossil collections held in museums.

6.3.5 Appendices

Appendix 6.1 Fossil pollen records from the literature

Taxon	Determination	Age	Locality	Reference
Sapotaceoidaepollenites occultus	Not Sapotaceae - Harley	Senonian	Shubei Basin, Taizhou Formation, China	Song *et al* 1999 & 2004
Sapotaceoidaepollenites manifestus	Not definitively Sapotaceae - Harley	Senonian	Shubei Basin, Taizhou Formation, China	Song *et al* 1999 & 2004
Sapotaceoidaepollenites robustus	Possibly not Sapotaceae – Harley	Senonian & Paleocene	Kayan Sandstone formation, Borneo, Indonesia	Muller 1968
Sapotaceoidaepollenites rotundus	Possibly Mimusopeae - Harley	Late Cretaceous-Paleocene	Australia	Stoian 2002
Sapotaceoidaepollenites granulatus	Possibly Sapotaceae - Harley	Paleocene-Eocene	Buxin Formation, Sanshui Basin, Guangdong Province, China	Song *et al* 1999 & 2004
Sapotaceoidaepollenites megasporus	Unidentifiable - Harley	Paleocene-Eocene	Buxin Formation, Sanshui Basin, Guangdong Province, China	Song *et al* 1999 & 2004
Sapotaceoidaepollenites		Paleocene -Early Eocene	Exxon Lydonia Canyon Block 133 No. 1 Well, Georges Bank Basin, N. Atlantic	Edson *et al* 2000
Tetracolporopollenites brevis	Compared to modern *Palaquim ellipticum*	Paleocene -Early Eocene	Rajpardi lignite, Cambay Basin, Gujarat, India	Prasad *et al* 2009
Tricolporopollenites latizonatus	*Pouteria* (= *Planchonella*) *novo-zeylandica* According to Harley 1991 is comparable to her Pollen Type I or II.	Middle to Upper Eocene	New Zealand	McIntyre 1965
Sapotaceoidaepollenites rotundus	Compared to *Tricolporopollenites latizonatus* in McIntyre (1965). According to Harley 1991 is comparable to her Pollen Type I or II	Middle Eocene	Wilkatana Formation, Australia	Harris 1972, Stover & Partridge 1973
Tetracolporpollenites	Compares with *Tieghemella heckelii* according to Harley	Middle Eocene	Younger Leaf Beds, Isle of Wight, England	Harley *et al* 1991
Sapotaceoidaepollenites neyvelii	Not definitively Sapotaceae -Harley (pers. comm. 2010)	Eocene	Lower Ganchaigou Formation, Qaidam Basin, Qinghai Province, China	Song *et al* 1999 & 2004
Sapotaceoidaepollenites sapotoides	Probably Sapotaceae – Harley (pers. comm. 2010)	Eocene	Sanduo Formation, North Jiangsu Basin, Jiangsu Province, China	Song *et al* 1999 & 2004

Taxon	Determination	Age	Locality	Reference
Tetracolporpollenites brevis	Affinities with Mimusopeae & Madhucoideae (According to Harley 1991 this pollen is reminiscent of *Tsebona macrantha* but the fossil grains are much smaller.)	Eocene	Claiborne Formation, Tennessee, U.S.A.	Taylor 1989
Psilastephanocolporites malacanthoides	*Malacantha alnifolia*	Eocene	Ogwashi-Asaba Formation Nigeria	Jan du Chêne *et al* 1978
Tetracolporpollenites transversalis	cf. *Chrysophyllum*	Eocene	Colombia	Jaramillo & Dilcher 2001, Rull 2000
Tetracolporpollenites maclosus	*Chrysophyllum argenteum*	Eocene	Colombia	Jaramillo & Dilcher 2001, Rull 2000
Chrysophyllum brevisulcatum		Eocene	Jackson Group, Lake Somerville, Texas	Raymond *et al* 1997
Tetracolporpollenites brevis		Eocene	Jackson Group, Lake Somerville, Texas	Raymond *et al* 1997
cf. *Chrysophyllum* (or *Pouteria* or *Micropholis*)	According to Harley (1991) this pollen is "decidedly *Pouteria*-like" and similar to her Pollen Type VIIA	Eocene	Panama	Graham 1985
Tetracolporopollenites oblongus	cf. *Nesoluma*	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites occultus	According to Harley (1991) probably *Mimusops* or *Palaquium*	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites halimbaense	According to Harley (1991) corresponds to her cf. Pollen Type VIA & IIA	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites kirchheimeri	According to Harley (1991) not *Madhuca* as indicated, corresponds to her Pollen Type VA	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites megadolium	cf. *Mimusops* According to Harley (1991) this could be Sapotaceous, corresponding to her Pollen Type VI, but not *Mimusops* as suggested by Gruas-Cavagnetto	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites obscurus	cf. *Isonandra* According to Harley (1991) this may not be *Isonandra* but some similarities to her Pollen Type VI	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites biconus	Sapotaceae	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites boureaui	Sapotaceae or Santalaceae	Eocene	England	Gruas-Cavagnetto 1976

Taxon	Determination	Age	Locality	Reference
Tetracolporopollenites elliptus	Sapotaceae	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites hungaricus	Sapotaceae	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites folliformis	Sapotaceae	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites manifestus	Sapotaceae	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites microellipsus	cf. *Palaquium*	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites microrhombus	Sapotaceae	Eocene	England	Gruas-Cavagnetto 1976
Psilastephanocolporites mimusopsoides	Sapotaceae (*Mimusops*)	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites fsp.	cf. *Chrysophyllum*	Eocene	England	Gruas-Cavagnetto 1976
Tetracolporopollenites fsp.	cf. *Palaquium*	Eocene	England	Gruas-Cavagnetto 1976
Iugopollis sp.	*Pouteria*-like pollen, (possible similarity with Sonneratiaceae?)	Lower Eocene	Cambay, India	Venkatachala & Rawat 1972, Rawat, Mukherjee & Venkatachala 1977
Sapotaceoidaepollenites oblongus	not comparable with pollen of any recent Sapotaceae genera seen by Harley (1991)	Lower Eocene	India	Venkatachala & Rawat 1972
Sapotaceoidaepollenites sp.	not comparable with pollen of any recent Sapotaceae genera seen by Harley 1991	Lower Eocene	India	Venkatachala & Rawat 1972
Sapotaceoidaepollenites oblongatus	not comparable with pollen of any recent Sapotaceae genera seen by Harley 1991	Lower Eocene	India	Venkatachala & Rawat 1972
Iugopollis tetraporites	According to Harley (1991) the pollen is *Pouteria*-like and corresponds to her cf. Pollen Type VIIA	Lower Eocene	Kauvery, Krishna-Godvari and Cambay, India	Venkatachala & Rawat 1972, Venkatachala 1974 Rawat, Mukherjee & Venkatachala 1977
lugopollis sp.	*Pouteria*-like pollen, (similarity with Sonneratiaceae?)	Lower Eocene	Cambay, India	Venkatachala & Rawat, Rawat, Mukherjee & Venkatachala 1977
"sapotaceous" pollen types	According to Harley 1991, compares with her cf. Pollen Types IA-C or IIA and Type VIIA	Lower Eocene	Claiborne Flora	Fairchild & Elsik, 1969
Tetracolporopollenites obscurus		Middle to Late Eocene	Yoncali formation, Turkey	Akgün *et al* 2002
Tetracolporopollenites abditus		Middle to Late Eocene	Yoncali formation, Turkey	Akgün *et al* 2002
Tetracolporopollenites microellipsus		Middle to Late Eocene	Yoncali formation, Turkey	Akgün *et al* 2002

Taxon	Determination	Age	Locality	Reference
Tetracolporopollenites sapotoides		Middle to Late Eocene	Yoncali formation, Turkey	Akgün *et al* 2002
Tetracolporopollenites manifestus		Middle to Late Eocene	Yoncali formation, Turkey	Akgün *et al* 2002
Tetracolporopollenites cf. *oblongus*		Middle to Late Eocene	Yoncali formation Turkey	Akgün *et al* 2002
Tetracolporopollenites sp.		Middle to Late Eocene	Yoncali formation, Turkey	Akgün *et al* 2002
Tetracolporopollenites brevis		Eocene –Early Oligocene	Eastern Gulf Coast, U.S.A.	Frederiksen 1988
Tetracolporopollenites prolatus		Eocene –Early Oligocene	Eastern Gulf Coast, U.S.A.	Frederiksen 1988
Sapotaceoidaepollenites lesquereuxianus = *Tetracolporopollenites lesquereuxianus*		Eocene –Early Oligocene	Eastern Gulf Coast, U.S.A.	Frederiksen 1980b &1988
Sapotaceoidaepollenites megadolium = *Tetracolporopollenites megadolium*		Eocene –Early Oligocene	Eastern Gulf Coast, U.S.A.	Frederiksen 1980b & 1988
?*Bumelia*, *Chrysophyllum* & ?*Manilkara*	*Chrysophyllum brevisulcatum* According to Harley 1991 - either a *Chrysophyllum* or a *Pouteria*, Pollen Type VIIA	Mid-Late Eocene-Early Oligocene	Southeastern United States	Frederiksen 1980a
Tetracolporopollenites microellipsus		Paleocene-Middle Eocene	Germany	Thomson & Pflug 1953
Tetracolporopollenites biconus		Paleocene-Lower Eocene	Germany	Thomson & Pflug 1953
Tetracolporopollenites microrhombus		Paleocene-Lower Eocene	Germany	Thomson & Pflug 1953
Tetracolporopollenites folliformis		Paleocene-Lower Eocene	Germany	Thomson & Pflug 1953
Tetracolporopollenites kirchheimeri		Paleocene, Eocene, Oligocene, Miocene	Germany	Thomson & Pflug 1953
Tetracolporopollenites oblongus		Middle Eocene	Germany	Thomson & Pflug 1953
Tetracolporopollenites sapotoides	comparable to Harley's (1991) Pollen Types I- III	Middle Eocene - Oligocene	Germany	Thomson & Pflug 1953
Tetracolporopollenites obscurus	Noticeable differences in the 20 examples according to Harley 1991, some categorized in her types I & II *s.l.*	Eocene & Miocene	Germany	Thomson & Pflug 1953
Tetracolporopollenites manifestus	*Sideroxylon* and/or *Vitellaria, Vitellariopsis* according to Harley 1991	Paleocene, Eocene, Oligocene	Germany	Thomson & Pflug 1953
Tetracolporopollenites abditus		Middle Eocene	Germany	Thomson & Pflug 1953

Taxon	Determination	Age	Locality	Reference
Tetracolporopollenites occultus		Middle Eocene	Germany	Thomson & Pflug 1953
Tetracolporopollenites sapotoides	cf. *Planchonella*	Mid-Eocene to Mid-Miocene	Southern Germany	Seitner 1987
Tetracolporopollenites sapotoides		Oligo-Miocene	Turkey	Akgün *et al* 2007
Tetracolporopollenites abditus		Oligo-Miocene	Turkey	Akgün *et al* 2007
Tetracolporopollenites manifestus/ellipsoides		Oligo-Miocene	Turkey	Akgün *et al* 2007
Tetracolporopollenites microellipsus		Oligo-Miocene	Turkey	Akgün *et al* 2007
Tetracolporopollenites microrhombus		Oligo-Miocene	Turkey	Akgün *et al* 2007
Tetracolporopollenites obscurus		Oligo-Miocene	Turkey	Akgün *et al* 2007
Tetracolporopollenites ericus		Oligo-Miocene	Turkey	Akgün *et al* 2007
Tetracolporopollenites sp.		Oligo-Miocene	Turkey	Akgün *et al* 2007
Mimusops-Manilkara	Similar to Mimusopeae, Isonandreae, Sideroxyleae &/or *Capurodendron* - Harley	Oligocene	Isle of Wight, England	Machin 1971
cf. *Chrysophyllum*	According to Harley 1991this could be *Chrysophyllum* or small *Pouteria* and is similar to her Pollen Type VIIA or VIIB	Oligocene	Puerto Rico	Graham & Jarzen 1969
Manilkara brevipollinia	According to Harley 1991, this matches her Pollen Type IA	Oligocene	Brandon Lignite, Vermont	Traverse 1953, 1955
Manilkara lesquereuxiana	According to Harley 1991, this matches her Pollen Type IA	Oligocene	Brandon Lignite, Vermont	Traverse 1953, 1955
Manilkara longipollinia	According to Harley 1991, this matches her Pollen Type IA	Oligocene	Brandon Lignite, Vermont	Traverse 1953, 1955
Mimusops mirabilis	According to Harley 1991, this pollen is difficult to reconcile since, although it has a number of Sapotaceous features, the verrucate exine has not been seen previously	Oligocene	Brandon Lignite, Vermont	Traverse 1953, 1955
Un-named Sapotaceae		Oligo-Miocene (Sapotaceae pollen more abundant in Oligocene)	Euxinian Basin (Eastern & Central Paratethys) Bulgaria	Ivanov *et al* 2002, 2007

Taxon	Determination	Age	Locality	Reference
Tetracolporopollenites longipollinius		Paleogene	South Carolina	Frederiksen 1980b
Tetracolporopollenites cf. *manifestus*	Sapotaceae	Miocene	Southern Germany	Seitner 1987
Tetracolporopollenites cf. *folliformis*	Sapotaceae, but Meliaceae also possible	Miocene	Southern Germany	Seitner 1987
Sapotaceae gen indet.		Early Miocene	Austria	Kovar-Eder *et al* 2001
Tetracolporopollenites microellipsus		Upper Miocene	Incesu formation, Central Anatolia	Akgün *et al* 2000
Sapotaceoidaepollenites obscurus	Possibly Sapotaceae - Harley	Miocene	Upper Youshashan Formation, Qaidam Basin, Qinghai Province, China	Song *et al* 1999, 2004
Sapotaceoidaepollenites kirchheimeri	Mimusopeae or Isonandreae - Harley	Miocene	Hanjiang Formation, North Continental Shelf of South China Sea	Song *et al* 1999, 2004
Belskipollis elegans	cf. *Chrysophyllum*	Miocene	Niger Delta	Legoux 1978
Sapotaceoidaepollenitesneyvelii (*S. arcotense* in the plate legends)		Miocene	Madras, India	Ramanujam 1996
Psilatricolporites maculosus	*Chrysophyllum argenteum*	Lower & Middle Miocene	Venezuela	Lorente 1986
Psilatricolporites pachydermatus	*Omphalocarpum* (According to Harley 1991 the general appearance is more reminiscent of Burseraceae pollen)	Lower & Middle Miocene	Venezuela	Lorente 1986
Psilastephanocolporites perforatus	compares with *Butyrospermum* = *Vitellaria* & *Manilkara* – according to Salard-Cheboldaeff	Lower Miocene	Cameroon	Salard-Cheboldaeff 1978, 1979, 1981
Chrysophyllum		Lower Miocene	Valley of Klodnica, Poland	Macko 1957
Sapotaceoidaepollenites africana	According to Harley (1991) corresponds to her Pollen Type VIF	Miocene - Pliocene	Burundi	Sah 1967
Sapotaceoidaepollenites obscurus	Uncertain affinity but possibly Sapotaceous according to Harley 1991	Miocene - Pliocene	Burundi	Sah 1967
Sapotaceoidaepollenites parvus	Uncertain affinity but possibly Sapotaceous according to Harley 1991	Miocene - Pliocene	Burundi	Sah 1967
Sapotaceoidaepollenites communis	According to Harley (1991) corresponds to her Pollen Type VIA	Miocene - Pliocene	Burundi	Sah 1967
Tetracolporopollenites sp.	similar to *Sapotaceoidaepollenites rotundus* according to Harris 1972	Miocene or Pliocene	Huon Peninsula, Papua New Guinea	Playford 1982
Tricolporopollenites glaber	*Sarcosperma* or Burseraceae – according to Harley (1991)	Miocene	Bakony Mountains, Hungary	Deak 1960
Tricolporopollenites globus	*Sarcosperma* or Burseraceae – according to Harley (1991)	Miocene	Bakony Mountains, Hungary	Deak 1960
Chrysophyllum, Madhuca & *Palaquium*		Pleistocene	Kerala, India	Farooqui *et al* 2009

Appendix 6.2 Fossil wood records from the literature

Taxon	Determination	Age	Locality	Reference
Apocynoxylon sapotaceoides	possible Sapotaceae	Upper Paleocene/ lowermost Eocene	East Anglian Crags, Suffolk, England	Crawley 2001
Faureoxylon princeps	family unknown, possibly Sapotaceae, compared with *Mimusops*	post Eocene?	Massif de Termit, Sahara, Niger	Koeniguer & Faure 1967
Paleosideroxylon flammula	*Sideroxylon*	Upper Miocene	Castellane, France	Grambast-Fessard 1968
Manilkaroxylon crystallophora	*Manilkara* (or more likely *Madhuca* according to Awasthi & Mehrota1993)	Upper Miocene	Castellane, France	Grambast-Fessard 1968
Bumelioxylon holleisii	*Bumelia/Sideroxylon*	Upper Miocene	Attenfeld, Germany	Selmeier 1991, Gottwald 2004, Böhme *et al* 2007
Chrysophylloxylon reticulatum	*Chrysophyllum*	Middle Miocene	Hungary	Müller-Stoll & Mädel-Angeliewa 1984
Arganioxylon sardum	similar to *Argania sideroxylon*	Miocene	Sardinia, Italy	Biondi 1981
cf. Sapotaceae	*Mimusops* or *Bassia*	Miocene	Cauvery Basin, Pondicherry, India	Lakshamanan & Levy 1956, Navale 1971
Sapotoxylon sp. 1 & 2		Miocene	Ethiopia	Wheeler *et al* 2007
Sapotoxylon aethiopicum	*Sideroxylon*	Miocene	Ethiopia	Lemoigne *et al* 1974
Sapotoxylon lecomtedoxoides	*Afroseralisia (=Synsepalum)* & *Lecomtedoxa*	Lower? Miocene	Ethiopia	Lemoigne 1978
Chrysophyllum zairense	*Chrysophyllum roxburghii*	Lower Miocene?	Manzandi Point VII, Lower Zaire	Bande *et al* 1987
Tridesmostemon tertiarum	*Tridesmostemmon claessensi*	Lower Miocene?	Karugamania beds, Lake Albert, Zaire	Bande *et al* 1987
Madhuca palaeolongifolia	cf. *Madhuca longifolia*	Mio-Pliocene	Deomali, Northeast India	Awasthi & Mehrota 1993
Manilkara cacharense	cf. *Manilkara hexandra* & *M. littoralis*	Mio-Pliocene	Hailakandi, Northeast India	Awasthi & Mehrota 1993
Siderinium deomaliense	cf. *Sideroxylon grandifolium*	Mio-Pliocene	Deomali, Northeast India	Prakash & Awasthi 1970
Sapotoxylon prepayena	cf. *Payena-Palaquium*	Mio-Pliocene	Kerala Coast	Awasthi & Srivastava 1990
Sapotoxylon multiporosum	*Mimusops, Manilkara, Payena, Englerophytum oblanceolatum,* or *Synsepalum*	Mio-Pliocene	Blue Nile Valley, Ethiopia	Prakash & Awasthi 1982
Sapotoxylon pacltovae		Oligocene?	South Bohemia	Prakash, Brezinova & Awasthi 1974
Manilkaroxylon bohemicum	*Manilkara*	Oligocene?	South Bohemia, Czechoslovakia	Prakash, Brezinova & Awasthi 1974, Petrescu 1978
Paleosideroxylon densiporosum	*Sideroxylon*	Paleogene	Transylvania	Petrescu 1978
Sapotoxylon atkinsoniae		Lower Tertiary	UK	Crawley 1989
Chrysophyllum pondicherrisense		Neogene	Cauvery Basin, Pondicherry, India	Awasthi 1977
Madhucoxylon cacharense	*Madhuca butyracea* Roxb.	Neogene	Hailakandi, Northeast India	Prakash & Tripathi 1977
Sapotoxylon taeniatum		Quaternary?	Bavaria, Germany	Felix 1882
Sapotoxylon gumbelii		Quaternary?	Bavaria, Germany	Felix 1882
Manilkaroxylon diluviale	*Manilkara*	Quaternary	Ecuador	Hofmann 1948

Appendix 6.3 Fossil leaf records from the literature

Taxon	Determination	Age	Locality	Reference
Chrysophyllum tertiarum	unverified	Upper Palaeocene	Nangwalbibra, India	Mehrota 2000
Sapoteae sp.	*Manilkara or Tieghemella*	Oligocene	Chigla, Ethiopia	Jacobs *et al* 2005, A. Pan pers. comm. 2010
Illipophyllum thomsoni	Resembles *Illipe* (= *Madhuca*)	Upper Oligocene to Miocene	Germany	Kräusel & Weyland 1959
Siderophyllum glandulosum	Resembles *Achras* (= *Manilkara*) *Sideroxylon* & *Chrysophyllum*	Upper Oligocene to Miocene	Germany	Kräusel & Weyland 1959
?*Pouteria* ex gr. *costata*		Early Miocene	New Zealand	Campbell 2002
Sapotacites ovatus		Upper Miocene	Sumatra	Kräusel 1929
Pouteria	also fossil fruits/seeds, pollen & wood at this site	Pleistocene	La Selva Biological Station, Costa Rica	Horn *et al*, 2003
Madhuca indica		Late Tertiary or Quaternary	Bihar, India	Bande & Srivastava 1990
Sapotacites crassipes		Tertiary	Sumatra	Heer 1876
Sapoteites ackneri		Tertiary	Transylvania	Andrae 1853
Pouterlabatia lanceolata		Tertiary	Rio Turbio Argentina	Hunicken 1955
Pouterlabatia clarcki		Tertiary	Rio Turbio, Argentina	Hunicken 1955
Chrysophyllum sp.		Tertiary	Rio Turbio, Argentina	Hunicken 1955
Bumelia? *rhomboidea*	cf. *Bumelia oreadum*	Cenomanian	Kansas	Lesquereux 1891

Appendix 6.4 Fossil fruit and seed records from the literature

Taxon	Determination by T.D. Pennington	Age	Locality	Reference
Bumelia? *globosa*	Check against *Sargentodoxa*	Eocene	Clarno Nut Beds, Oregon	Manchester 1994
Bumelia? *subangularis*	Check against *Sargentodoxa*	Eocene	Clarno Nut Beds, Oregon	Manchester 1994

Appendix 6.5 British Museum London Clay specimens

Taxon	BM accession number and determination by TD Pennington (TDP)	Age	Locality	Form	Reference
Sapoticarpum latum	V.23056 - TDP ID No mature seed so difficult to ID V.23057 - same as V.23058 TDP ID Not *Eberhardtia* because it has a narrow scar.	Eocene	London Clay Flora, Sheppey, England	fruit	Reid & Chandler 1933
Sapoticarpum latum	V.23058 - TDP ID Fruit 5-lobed and 5-seeded. Could be *Chrysophyllum* section Aneuchrysophyllum but need to see seed. Could also be *Eberhardtia*.	Eocene	London Clay Flora, Sheppey, England	fruit	Reid & Chandler 1933
Sapoticarpum rotundatum	V.23054 Holotype - TDP ID *Chrysophyllum* section Ragala. Seed scar round base of seed is consistent with species in this section (*C. bangweolense* or *C. sanguinolentum*)	Eocene	London Clay Flora, Sheppey, England	fruit	Reid & Chandler 1933
Sapoticarpum rotundatum	V.23055 - TDP ID *Pouteria* possible sect. Franchetella (small fruited), could also be *Synsepalum*. Based on seed being only down adaxial side.	Eocene	London Clay Flora, Sheppey, England	fruit	Reid & Chandler 1933
Sapoticarpum dubium	V.23059 - TDP ID *Chrysophyllum* section Aneuchrysophyllum? Lobed fruit. Embyo with flat cotyledons but cannot see seed scar. V.23060 - TDP ID *Chrysophyllum* sect. Aneuchrysophyllum	Eocene	London Clay Flora, Sheppey, England	fruit	Reid & Chandler 1933
Sapotispermum sheppeyense	V.23061, 23062, 23063, 23064, 23065, 23066, 23067 TDP ID *Chrysophyllum* or *Pouteria*. Angular seed	Eocene	London Clay Flora, Sheppey, England	seed	Reid & Chandler 1933
Sapotispermum sp. 2	V.40875 Indeterminate	Eocene	London Clay Flora, Sheppey, England	seed	Reid & Chandler 1933
Bumelia	V.18572 TDP ID not Sapotaceae				
Sapotaceae	V.18575 not Sapotaceae				

Appendix 6.6 Smithsonian Institute (United States National Museum) fossil specimens (*reference not recovered for specimens with asterisk)

Taxon and accession number	Determination by T.D. Pennington	Age	Locality	Form	Reference
Eoachras eocenica USNM 35470	Scar area (if that's what it is) is far too large and protruding for *Manilkara zapota* - TD Pennington	Eocene	Claiborne group, Lexington Mississippi	seed	Berry 1915
Mimusops claibornensis USNM 38333	Not enough detail for identification - TD Pennington	Eocene	Claiborne group, Cherry Valley, Arkansas	leaf	Berry 1924
Mimusops eolignitica USNM 35892	Not *Manilkara* or *Mimusops* (venation is eucamptodromous) - TD Pennington	Eocene	Ackerman Formation, Mississippi & Lagrange Formation, Tennessee	leaf	Berry 1916c
Mimusops mississippiensis USNM 35983 & 36305	This could be a possible *Manilkara*/*Mimusops* - TD Pennington	Lower Eocene	Grenada formation, Mississippi	leaf	Berry 1916c
Mimusops praenuntia USNM 39928	Not enough detail for identification - TD Pennington	Lower Eocene	Holly Springs sands, Bradley Pit, Tennessee	leaf	Berry 1930
Mimusops praenuntia USNM 39929	Possibly a *Sideroxylon* (*Bumelia*) - TD Pennington	Lower Eocene	Holly Springs sands, Bradley Pit, Tennessee	leaf	Berry 1930
Mimusops praenuntia USNM 39930	Secondary veins too convergent for *Mimusops/Manilkara* - TD Pennington	Lower Eocene	Holly Springs sands, Bradley Pit, Tennessee	leaf	Berry 1930
Mimusops sieberifolia USNM 35980 & 35981	35980: Not enough detail, 35981: Probably not *Manilkara/Mimusops* - TD Pennington	Lower Eocene	Lagrange Formation, Tennessee	leaf	Berry 1916c
Achras calcicolafolia USNM 39307	Not enough detail for identification - TD Pennington	Miocene	La Victoria, Zulia, Venezuela	leaf	Berry 1936
Mimusops anomala USNM 320555	The size and leaf outline is OK for *Manilkara/Mimusops* but no venation is visible - TD Pennington	Miocene	Trinidad	leaf	Berry 1925c
Mimusops praeparvifolia USNM 36613	No venation visible - TD Pennington	Miocene	Haiti	leaf	Berry 1922b
Mimusops leei USNM 315212	Not enough detail for identification - TD Pennington	Miocene	Fyzabad, Trinidad	leaf	Berry 1937
Mimusops leonii USNM 315027	Leaves have brochidodromous venation so could be *Manilkara/Mimusops*	Miocene	Yumari, Cuba	leaf	Berry 1939
Mimusops miocenica USNM 315028	Not enough detail for identification - TD Pennington	Miocene	Yumari, Cuba	leaf	Berry 1939
Mimusops miocenica USNM 320175 to 320178	Not enough detail for identification - TD Pennington	Miocene	Trinidad	leaf & fruit	Berry 1925c

Taxon and accession number	Determination by T.D. Pennington	Age	Locality	Form	Reference
Sapota agnitionalis USNM 315217	Venation OK for *Manilkara/Mimusops* - TD Pennington	Miocene	Fyzabad, Trinidad	leaf	Hollick 1928
Mimusops miocenica USNM 315133	Not enough detail for identification - TD Pennington	Pliocene	Anzoategui, Venezuela	leaf	Berry 1939
Mimusops preduplicata USNM 37024 to 37030	Size, outline and venation are good for *Mimusops/Manilkara* - TD Pennington	Pleistocene	Trinidad	leaf	Berry 1925b
Mimusops preduplicata USNM 320571 to 320578	Size, outline and venation are good for *Mimusops/Manilkara* - TD Pennington	Pleistocene	Trinidad	leaf	Berry 1925b
Mimusops emarginata (= *Manilkara jaimiqui*) USNM 316365	The size, multilocular fruit and narrow seeds are good for this species - TD Pennington	Pleistocene	Santa Clara, Cuba	fruit	Berry 1934
Bumelia americana P 35967, 39926	Similar to *Bumelia wilcoxiana* (fossil species)	Lower Eocene	Lagrange Formation, Tennessee	leaf	Berry 1916c
Bumelia apalachicolensis P 38286		Miocene	Florida	leaf	Berry 1916d
Bumelia australis P 40478		Tertiary	Rio Pichileufu, Argentina	leaf	Berry 1938
Bumelia coloradensis P 36853		Eocene	Green River, Colorado	leaf	Cockerell 1908, 1925
Bumelia cuneatafolia P 36615		Miocene	Haiti	leaf	Berry 1922b
Bumelia florissanti P 1797, 1798, 39358		Miocene	Florissant, Colorado	leaf	Lesquereux 1883, Cockerell 1908
Bumelia grenadensis P 3597	resembles *Bumelia oreadum* (European fossil specimen)	Lower Eocene	Grenada Formation, Mississippi	leaf	Berry 1916c
Bumelia hurleyensis P 35969		Lower Eocene	Ackerman Formation, Mississippi	leaf	Berry 1916c
Bumelia lojana PAL 313928		Tertiary	Ecuador	leaf	Berry 1945
Bumelia marahiana P 38183		Late Tertiary	Brazil	leaf	Hollick & Berry 1924
Bumelia oklahomensis P 35289		Late Tertiary	Oklahoma	leaf	Berry 1918
Bumelia preangustifolia PAL 321500		Miocene-Pliocene	Citronelle Formation, Alabama	leaf	Berry 1916b
Bumelia prewilcoxiana P 37168, 37173		Maastrichtian	Ripley Formation	leaf	Berry 1925a

Taxon and accession number	Determination by T.D. Pennington	Age	Locality	Form	Reference
Bumelia pseudohorrida P 35971	cf. *Bumelia horrida*	Lower Eocene	Lagrange Formation, Tennessee	leaf	Berry 1916c
Bumelia pseudolycioides P 39927		Eocene	Wilcox Formation, Tennessee & Kentucky	leaf	Berry 1930
Bumelia pseudotenax P 35951	cf. *Bumelia tenax* & *B. lanuginosa*	Lower Eocene	Akerman Formation, Mississippi & Calaveras Creek, Texas	leaf	Berry 1916c
Bumelia reclinatafolia P 35459		Tertiary	Dominican Republic	leaf	Berry 1921
Bumelia retusafolia PAL 315013	Not Sapotaceae, looks more like Piperaceae - TD Pennington	Miocene	Yumuri, Cuba	leaf	Berry 1939
Bumelia ripleyensis P 37169		Maastrichtian	Ripley Formation	leaf	Berry 1925a
Bumelia trinitense PAL 320204		Tertiary	Trindad	leaf	Berry 1925c
Bumelia vicksburgensis P 40032		Oligocene	Texas?	leaf	Berry 1916a
Bumelia wilcoxiana P 35970, P 35968, P 35973	cf. *Bumelia retusa*	Lower Eocene	Marshall County, Mississippi & Lagrange Formation, Tennessee	leaf	Berry 1916c
Sideroxylon ellipticus P 35974	cf. *Sideroxylon surinamense*	Lower Eocene	Lagrange Formation, Tennessee	leaf	Berry 1916c
Sideroxylon mastichodendroides PAL 320190-4		Tertiary	Trindad	leaf	Berry 1925c
Sideroxylon pliocenicum PAL 320665		Pliocene	Bolivia	leaf	Berry 1922a
Sideroxylon premastichodendron P 35975	cf. *Sideroxylon mastichodendron*	Lower Eocene	Lagrange Formation, Tennessee	leaf	Berry1916c
Chrysophyllum cahobasensis P 36614		Tertiary	Haiti	leaf	Berry 1922b
Chrysophyllum cainitoformis PAL 315145, PAL 315146	cf. *Chrysophyllum cainito*	Late Tertiary	Anzoategui, Venezuela	leaf	Berry 1939
Chrysophyllum crassum PAL 320674		Pliocene	Bolivia	leaf	Berry 1922a
Chrysophyllum ficifolia PAL 321570	*Chrysophyllum*	Lower Eocene	Granada formation, Mississippi & Lagrange Formation, Tennessee	leaf	Berry 1916c
Chrysophyllum parvum P 37162					Berry 1925*

Taxon and accession number	Determination by T.D. Pennington	Age	Locality	Form	Reference
Chrysophyllum rolloti PAL 316855		Tertiary	Colombia	leaf	Berry 1929
Sapotacites alaskensis P 37665		Upper Cretaceous	Alaska	leaf	Hollick 1930
Sapotacites ettingshauseni P 34943		Upper Cretaceous	U.S. Gulf region	leaf	Berry 1919
Sapotacites formosus P 34950		Upper Cretaceous	U.S. Gulf region	leaf	Berry 1919
Sapotacites haydenii P 703				leaf	Heer 1858*
Sapotacites haydenii P 974, PAL 311502		Cretaceous	Dakota Formation, Nebraska	leaf	Lesquereux 1883
Sapotacites miraflorianus P 38389		Eocene		leaf	Berry 1924
Sapotacites shirleyensis P 34936, P 34940		Upper Cretaceous	U.S. Gulf region	leaf	Berry 1919
Sapotacites spathulatus P 38287	*Mimusops/Bumelia*	Miocene	Alum Bluff, Florida	leaf	Berry 1916d
Sapotacites sp indet. P 50067		Cenomanian	Dakota Formation, Kansas	leaf	Lesquereux 1891

Appendix 6.7 Yale Peabody Museum fossil specimens. (*reference not recovered for specimens with asterisk)

Taxon	Determination	Age	Locality	Form	Reference
Sapotacites americanus TYPE - YPM 27102 & 152609		Paleocene, Eocene	Louisiana, Red River County, Coushatta	leaf	Harris *et al* 1899
Mimusops claibornensis YPM PU 155819, 155820, 155821	image not avaible	Eocene	Wyoming, Crook County, Squaw Buttes, Tatman Formation	leaf	Berry 1924
Sideroxylon aequale TYPE - YPM 27213		Middle Oligocene	West Indies, Puerto Rico. Collazo River, San Sebastian Formation	leaf	Hollick 1928
Sideroxylon aequale? YPM 27143	used by Smedmark & Anderberg (2007) but specimen not seen	Middle Oligocene	West Indies, Puerto Rico, Collazo River, San Sebastian Formation	leaf	Hollick 1928
Chrysophyllum comparabile TYPE - YPM 27819		Middle Oligocene	West Indies. Puerto Rico, Collazo River, San Sebastian Formation	leaf	Hollick 1928
Chrysophyllum pseudargenteum oblongum TYPE - YPM 27818, 27813, 27824, 27493, 27502		Middle Oligocene	West Indies, Puerto Rico, Collazo River, San Sebastian Formation	leaf	Hollick 1928

Taxon	Determination	Age	Locality	Form	Reference
Bumelia reclinatafolia YPM 23913		Tertiary	Cuba, Matanzas	leaf	Hollick 1924
Mimusops jumuriensis TYPE - YPM 23862-3	Not enough detail in image for determination –TD Pennington	Tertiary	Cuba, Matanzas	leaf	Hollick 1924
Mimusops leonii TYPE - YPM 23815	Not enough detail in image for determination –TD Pennington	Tertiary	Matanzas, Cuba	leaf	Hollick 1924
Bumelia pseudo-lanuginosa TYPE - YPM 23972 & 27793		Pleistocene	Maryland, Calvert County, Island Creek, Sunderland Formation	leaf	Hollick 1907
Sapota agnitionalis TYPE - YPM 27192	Not enough detail in image for determination –TD Pennington	Middle Oligocene	San Sebastian Flora, Puerto Rico	leaf	Hollick 1928
Phyllites mimusopsoideus YPM 152976				leaf	Lesquereux 1883
Sapotacites millicanensis YPM PU 157143 & YPM PU 157144		Eocene	Wyoming, Yellowstone National Park	leaf	Berry*
Sapotacites retusus YPM 152169 & YPM 24853			New Jersey, Amboy Clays	leaf	Heer*

Appendix 6.8 University of California Berkeley Museum of Palaeontology fossil specimens

Taxon	Age	Locality	Form	Reference
Lucuma standleyi 208, 296, 297	Upper Eocene	Oregon, Fisher Formation, Goshen Flora	ovulate cone scale and leaves	Chaney & Sanborn 1933
Chrysophyllum conforme 942	Eocene	La Porte, California	leaf	Potbury 1935
Bumelia florissanti 1514, 1515	Miocene	Tehachapi, California, Kinnie Formation	leaf	Axelrod 1939
Bumelia beaverana 3304	Miocene	Anaverde Formation, California	leaf	Axelrod 1950
Bumelia beaverana 4165	Miocene	Aldrich station, Nevada	leaf	Axelrod 1956

Chapter VII – Molecular dating and biogeographical analysis

7.1 Introduction

Following the biogeographical scenarios presented in Chapter V and the fossil history presented in Chapter VI, this chapter utilizes dated phylogenies to investigate how *Manilkara* has evolved in response to geo-climatic changes on global and regional scales. Different fossil calibration points and methods of relaxing the molecular clock are tested on the dataset, as are the applicability of various DNA substitution models and clock models to find the best fit to the data.

Node ages and ancestral areas of the tribe Mimusopeae, subtribe Manilkarinae and the genus *Manilkara* are reconstructed in order to investigate where the groups originated and which biogeographic processes have shaped their current distribution. Moreover, hypothetical lineage splits outlined in Chapter V for Gondwanan vicariance (section 5.3, Fig. 5.3), boreotropical migration (section 5.4, Fig. 5.5), and long distance dispersal (section 5.6) are assessed. The diversification of *Manilkara* in each continental region is also investigated in relation to geo-climatic factors as outlined in Chapter V for the Neotropics (Fig. 5.11), Africa (Fig. 5.14) and Asia (Figs. 5.17 and 5.18). Monophyly of the African, Madagascan, Neotropical and Asian *Manilkara* lineages is also tested.

A comparison of the reconstructed mean substitution rate for the Sapotaceae dataset and other published rates is made as an additional validation that the Sapotaceae rate is within the normal range for angiosperms. It is predicted to be similar to that of other long-lived, woody eudicots.

7.2 Dating phylogenies

7.2.1 Molecular clock theory

Dating speciation events and studying their mode and tempo is a primary objective of evolutionary biology, as it is a step towards understanding their causes. If the timing of speciation events coincides with geological or climatic phenomena these may then be implicated as being part of the processes, which resulted in those events. The molecular clock approach allows for the incorporation of time into our reconstruction of the evolution of lineages. As originally conceived, it is based on the principle that DNA evolves at a relatively constant rate and that the difference between sequences of two species is a function of the time since their separation, assuming a neutral model of molecular evolution. Because molecular change is believed to accumulate steadily through time, it has the potential to provide a means of temporal calibration in a phylogeny (Bromham & Penny 2003). Therefore, if the age of a single node can be estimated, for example using a fossil calibration point, the age of all others can also be determined by extrapolation. (While the concept of a strict molecular clock was originally widely accepted, it is now recognized not to be biologically accurate and methods using a relaxed clock are preferred as discussed later in section 7.2.3.)

Zuckerkandl & Pauling (1962) postulated the existence of a molecular clock when they noticed that the rate of amino acid changes in protein sequences amongst lineages of animals was linear over evolutionary timescales. Kimura & Ohta (1971) further suggested that this constant rate of amino acid change was due to neutral mutations and predicted that the rate of molecular evolution of a species should be the same as the neutral mutation rate in individuals. However, while some datasets do exhibit a clock-like accumulation of substitutions, it has become apparent that there is no universal clock across all taxa (Near *et al* 2005; Bromham & Penny 2003; Renner 2005) and recent studies have shown extensive rate heterogeneity amongst lineages of vascular plants (Kay *et al* 2006). Nonetheless, when variability of rates is accounted for in phylogenetic reconstruction, the use of a molecular clock is still a valid method for distinguishing between competing biogeographic hypotheses.

A simple method for calculating divergence rate is to estimate the mutation rate by dividing genetic distance (the number of base changes between two DNA sequences, corrected with a nucleotide substitution model) by the age of a fossil which exhibits synapomorphies for a specific clade in a phylogeny (Bromham & Penny 2005, Shields 2004, Renner 2005) or by geological events associated with speciation of lineages *a priori* (Richardson *et al* 2001b; Plana *et al* 2004, Renner 2005). The rate is then used to convert the genetic distances between taxa of interest into estimates of their absolute ages (Renner 2005). It is assumed that this rate applies universally to the entire phylogeny and dates for the remaining nodes are extrapolated.

Molecular clocks are a powerful, yet controversial tool and there have been many arguments for and against their use in the dating of phylogenies (Smith & Peterson 2002, Graur & Martin 2004, Hedges & Kumar 2004, Cranston & Rannala 2005, Renner 2005, Ho & Larson 2006, Pulquerio & Nichols 2007). It is advisable to be cautious when using molecular dating approaches, realising that they are only as reliable as the precision of their estimated genetic distance and the calibration rate (Bromham & Penny 2005, Renner 2005). The entire process of molecular dating has come under fire as being subject to various sources of error, which must be taken into consideration (Graur & Martin 2004, Near *et al* 2008, Shields 2004).

Errors in dating can be produced, for example, by incorrectly specifying genetic distances due to inappropriate nucleotide substitution models. Rate heterogeneity between datasets can also lead to difficulty in calibration. Phylogenetic placement of fossil taxa is particularly challenging and a potential source of error because calibration points based on incorrect assessments of taxonomic placement and age can produce exceedingly incorrect dates, which are then propagated throughout the phylogeny. Fossil choice should, therefore, be as rigorous and unambiguous as possible (Parham & Irmis 2008). Additionally, while errors associated with isotope-based dating are typically small, they should be taken into consideration if the age of the fossil strata is in any doubt (Smith & Peterson 2002, Hedges & Kumar 2004). Despite all of these difficulties, the use of dated phylogenies has increased dramatically over the past decade and they are employed extensively to test amongst competing biogeographic and evolutionary hypotheses. Provided that potential error margins are taken into account, the approach remains a powerful tool in evolutionary studies.

7.2.2 Approaches for temporal calibration of phylogenies

7.2.2.1 Use of fossils as calibration points

A fossil to be used for calibration of a node in a phylogeny should exhibit at least one synapomorphic character of the crown group it will date and can be assigned to the stem or crown node of that clade. This is an important consideration because it is assumed that the synapomorphic characters of a group evolved at some point in time along its stem lineage. It has been argued that a fossil with the synapomorphies for a particular clade should be placed to constrain the node where the stem lineage splits from its sister group, i.e. node 1 in Fig. 7.1. Since it is not known where along the stem lineage a particular character evolved, the most conservative placement is at the stem node, i.e. the earliest point at which the synapomorphy could have evolved (Renner 2005). However, according to Smedmark & Anderberg (2007), for example, this approach leads to consistent underestimation of divergence times, caused by fossils always being younger than the taxa they represent and being assigned to nodes which are too old. Therefore, in some instances depending upon its taxonomic affinity, a fossil is best placed at the crown node, biasing towards older age estimates, i.e. nodes 2 and 3 in Fig. 7.1. Both approaches have been utilized in recent studies; i.e. stem node (Davis *et al* 2004, Richardson *et al* 2004, Pirie *et al* 2006, Couvreur *et al* 2008) and crown node placement (e.g., Zerega *et al* 2005, Muellner *et al* 2006, Smedmark & Anderberg 2007), as well as a combination of stem and crown (Muellner *et al* 2008) and testing of both placements (Forest *et al* 2005).

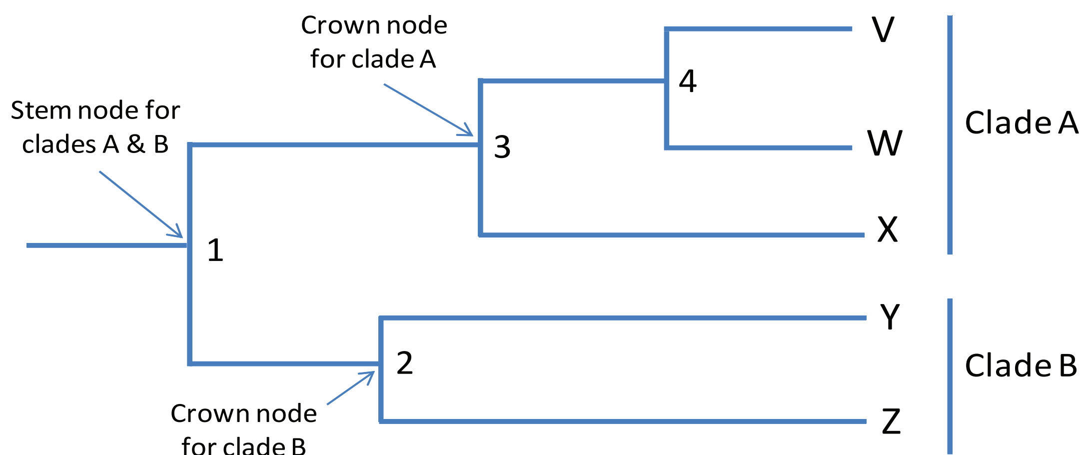


Figure 7.1 In this example, node 3 is the crown node for Clade A (taxa V, W & X) and node 1 is the stem node for Clade A. The stem lineage of this clade is the branch connecting nodes 1 & 3. Node 1 is also the stem node for Clade B (taxa Y & X).

Because the fossil record is incomplete, the first occurrence of a fossil taxon can only establish the time by which that clade must have come into existence, but this often significantly underestimates the true divergence time, as the fossilization process may considerably post-date the evolution of the higher taxon the fossil represents. Fossil calibration can, therefore, only provide an upper bound (i.e. a minimum age estimate) for a clade. The lower bound (when the taxon first originated) is unknown and must be estimated (Marshall 1990, Shields 2004, Smith & Peterson 2002, Hedges & Kumar 2004). Rates based on fossil calibration, therefore, have an inherent imprecision which can be magnified

throughout the analysis, generating further inaccurate date estimates for other splits derived from the original estimate (Shields 2004).

When deciding between several fossils, it is considered best practice to choose the oldest fossil or to investigate the reliability of their effects sequentially. It is also believed that the use of multiple fossils has the potential to reduce calibration errors (Near & Sanderson 2004) and that it is good to have at least one constraint near the root of the tree and one close to the tips (Renner 2005, Smith and Peterson 2002, Soltis *et al* 2002). However, Hedges & Kumar (2004) argue that greater accuracy can be obtained through the use of "tightly constrained fossil calibrations close to the speciation event, rather than many calibrations that are poorly constrained." Linder *et al* (2005) and Milne (2009) found that undersampling of taxa in a dated phylogenetic analysis can also affect the estimated age of a lineage, sometimes dramatically, depending on the dating methods applied. They also determined that the undersampling effect is positively related to distance from the calibrated node, and that calibration points should, therefore, be situated within the study group (Linder *et al* 2005).

Near *et al* (2005) suggest a cross-validation method for discerning between consistent (and, therefore, useful) fossils in an analysis and those which are inconsistent and potentially giving erroneous date estimates. This method is described further and utilized in the analyses presented later in this chapter. Variance in the age estimate for a node using different fossil dates can be used to calculate confidence intervals around age estimates (Near *et al* 2005). However, Smith & Peterson (2002) caution that "it is important to remember that the distribution of error is not normal around the correct date (fossil dates are always underestimates), and so the more wrong dates that are included the further the mean will depart from the true date."

7.2.2.2 Secondary calibration methods in the absence of fossil data

For groups without a reliable fossil history, secondary calibrations are also used. In this instance, the age is taken from another, usually much broader, analysis, which incorporates fossils and at least one member of the focus group. For example, Wikström *et al* (2001) dated a phylogeny of angiosperms using a single well characterized fossil. This gave stem node age estimates for nearly all angiosperm families that have subsequently been used as secondary calibration points to date, for example, Rhamnaceae (Richardson *et al* 2004).

7.2.2.3 Calibration using geological events

When fossil ages are unavailable, another method of dating is to constrain nodes with the age of a geological event, for instance, the age of a volcanic island which supports an endemic species of interest (Richardson *et al* 2001, Plana *et al* 2004). However, this method overlooks various speciation scenarios, i.e. that the species could have diverged from its sister before arrival on the island, that it could have gone extinct on the mainland but only now remains on the island, or that it could have evolved millions of years after the island emerged. The only appropriate use of an island age estimate is in relation to an original colonist, which has diversified and the crown node of this diversification is the earliest point at which it could have arrived on the island.

Some vicariance events could also have been assumed to have been caused by continental break-up and hence to have occurred at the same time (e.g. Becerra 2003). However, when this approach is used in biogeographic analyses it can be circular because it assumes that the geological event caused the divergence. Plants in particular seem to be extremely capable trans-oceanic dispersers (Renner 2004, Cody *et al* 2010), so the use of geological events to date diversification events in this group of organisms is especially risky. This method can also be extended to using the age of a habitat to which a group is strongly adapted, but this can also be imprecise due to the fact that ages of habitats are, for the most part, poorly understood (Renner 2005).

7.2.2.4 Applying rates from other studies

Another commonly employed method is to apply a known substitution rate from a well-calibrated phylogeny for a group which has a similar life history; i.e. a slow growing tree, or a herb with a fast generation time. However, applying a rate from another taxon should be done with caution, as numerous factors including generation time, metabolic rate, efficiency of DNA repair and population size can affect mutation rate (Kay *et al* 2006). Additionally, different genes evolve at different rates depending upon their function and selection pressures; non-coding regions evolve faster than coding regions (Small *et al* 1998). It is, therefore, inadvisable to apply a rate from one gene to that of another. Among-lineage rate heterogeneity, probably caused by differences in life history traits amongst species, has been shown to exist in *Sidalcea* (Andreasen & Baldwin 2001). The prevalence of rate heterogeneity is the basis for applying a relaxed clock model in molecular dating analyses, with rates drawn from an underlying distribution (e.g. lognormal) (Lemey & Posada 2009). These methods are discussed further below.

To summarise, calibration with fossils is the preferred approach because it has the most direct link to the taxa being studied and when done carefully and critically is the best option we currently have for dating lineages in the absence of known absolute ages.

7.2.3 Using a relaxed molecular clock

Contrary to Kimura & Ohta's (1971) original assumption, a strict molecular clock is not biologically realistic. Because rates of sequence divergence have been shown to differ significantly amongst lineages of organisms (Kay *et al* 2006) and sometimes even within a group of closely related species (Andreasen & Baldwin 2001), a strict clock is often too simplistic a model in phylogenetic reconstruction. Despite this, molecular clock theory remains an important concept and methods have been developed which relax the clock, allowing the rate to vary amongst lineages. Two main methods have been proposed. In the first method the variation of rate occurs around an average value, whereas in the second method, the rate is allowed to evolve – i.e. to change over time, based on the assumption that it is linked to other evolutionary characteristics such as metabolic rate or generation time (Ho 2008, Lepage *et al* 2007). These differing methods have been reviewed in Welch & Bromham (2005) and Rutschmann (2006). Relaxed clock models offer a middle ground between strict clock models, which assume a constant rate of evolution across lineages

and time-free models which completely lack a model of evolutionary rates (Wertheim *et al* 2010).

Sanderson (1997 & 2002) was the first to implement a relaxed molecular clock in phylogeny reconstruction using nonparametric rate smoothing (NPRS) and penalized likelihood (PL), which "smooth" the differences between rates across a tree by allowing them to vary on branches, though autocorrelation of rates is assumed. Rate autocorrelation is meant to reflect the fact that closely related lineages share biological characteristics (such as generation time) and are thus expected to have similar evolutionary rates, which are only likely to change over long time frames (Lemey & Posada 2009).

NPRS uses substitution rate combined with an optimality criterion to provide estimates of divergence time. Through this process, it tests how clock-like or non-clock-like the data are and implements the appropriate smoothing parameter across each node in the tree. PL is similar to NPRS, but utilizes a smoothing parameter estimated through a cross-validation procedure. This procedure removes each terminal branch sequentially, estimates the remaining parameters of the model without the branch, predicts the expected number of substitutions along the pruned branch and calculates a cross-validation score based on the difference between the actual and predicted branch lengths (Sanderson 2002, Near & Sanderson 2004, Rutschmann 2006). NPRS and PL are applied in the program r8s (Sanderson 2003).

A different method, using local clocks, was employed by Yoder and Yang (2000). This assigns separate rate parameters to specified branches in a tree. However, rate and time can be confounded (i.e. the inability to separate the contribution of rate and time) when too many priors are specified near the root (Yoder & Yang 2000, Lemey & Posada 2009). This method is implemented in BASEML (in PAML, Yang 1997).

Various Bayesian parametric models have also been proposed by Thorne *et al* (1998), Huelsenbeck *et al* (2000) and Drummond *et al* (2006). Thorne's relaxed clock model assumes that rates are autocorrelated across a tree and designates new rates to descendant lineages from a lognormal distribution, where the mean is equal to the rate of the ancestral lineage. This model is implemented in the software MULTIDIVTIME (Thorne & Kishino 2002).

Huelsenbeck *et al*'s (2000) Bayesian parametric model relaxes the molecular clock by allowing rates to vary across lineages according to a compound Poisson process. Rate changes are modified by multiplying the current rate by a gamma distributed random variable. This differs from other models proposed by Sanderson (1997) and Thorne *et al* (1998) in allowing rates to change anywhere on a tree (rather than only at a node).

Drummond *et al*'s (2006) Bayesian uncorrelated relaxed clock model proposes an alternative to autocorrelation of rates across a tree. Instead branch-specific rates are independently chosen from an underlying rate distribution (i.e. exponential or log normal) depending upon the priors specified. As such, it explicitly models the rate of molecular evolution for each

branch in a phylogeny. This method is applied in the software package BEAST (Bayesian Evolutionary Analysis by Sampling Trees) (Drummond & Rambaut 2007). BEAST is the only program which implements an uncorrelated relaxed clock in a Bayesian framework and simultaneously estimates divergence times, tree topology and evolutionary rates as part of the same calculation. Estimating all of these variables in concert is considered a positive development in dating methodology because, since a rate is drawn independently for each prior, if an underlying distribution of rates exists, a Bayesian search will find it.

7.3 Methodological approach

7.3.1 Introduction to methodological approaches used in this chapter

In the remainder of this chapter a range of methods are tested for reconstructing node ages using various fossil calibration scenarios and three different methods of relaxing the molecular clock in the programs BEAST (relaxed uncorrelated lognormal) and r8s (penalized likelihood and nonparametric rate smoothing). Substitution models and clock-likeness of the data are tested on the dataset using Bayes factors in MrBayes and BEAST. Additionally, ancestral areas are reconstructed in BEAST and hypotheses of area monophyly are tested in MrBayes.

7.3.2 Choice of fossil calibration points

The Sapotaceae fossil record is vast and encompasses micro fossils (pollen) and macro fossils (leaves, wood, fruit and seed). An extensive literature search was conducted in order to understand the range of fossil types and their spread in space and time (presented in Chapter VI). A selection of fossil pollen images from publications (Machin 1971, Song *et al* 1999 & 2004, Stoian 2002) were evaluated with advice from Madeline Harley (RBG Kew) and selected macro fossils held in museums were surveyed with determinations from Terry Pennington (RBG Kew). These latter included the London Clay fossils at the British Museum (Table 6.5) and images of Berry and Hollick collections (originally placed in the Mimusopeae: *Mimusops*, *Achras*, *Eoachras* & *Sapota*) held in collections at the Smithsonian Institution (Table 6.6) and the Peabody Museum (Table 6.7). After investigating a number of fossils, three were chosen as calibration points, as their morphological features were deemed sufficient to allow confident placement at specific nodes in the phylogeny.

Some of the earliest putative fossil pollen grains of Sapotaceae are recorded from the Cretaceous of China (Song *et al* 1999, 2004) and were viewed as potentially useful calibration points for constraining the basal node of the phylogeny, but after careful review of the fossil images, Harley (pers. comm. 2010) questioned their validity as sapotaceous. Cretaceous Sapotaceae pollen reported from Borneo (Muller 1968) was also determined to be of questionable affinity and was not used. Therefore, in order to constrain the maximum age of the crown node of the family (node A in Fig. 7.2), a secondary calibration age of 102 Ma was used, as determined by Bremer *et al* (2004) in a study on the age of the Asterid clade. Bremer's estimate was determined through testing various dating methods in r8s (penalised likelihood, non-parametric rate smoothing and Langley-Fitch) with a maximum

likelihood input tree based on a dataset of 111 taxa in 84 Asterid families and six chloroplast genes (*rbcL*, *ndhF*, *matK, trnL-trnF, trnV-atpE*, *rps16*) calibrated with six well-characterized fossils, which were placed above the node and on the branch leading to the family of the fossil. In the Bremer *et al* (2004) dataset *Manilkara* was the only representative taxon for the Sapotaceae and, hence, this study only provided the age of the stem node of the family.

Sideroxyleae pollen from the early Eocene of England dated at ~49 Ma (Gruas-Cavagnetto, 1976) was used to constrain the minimum age of the Sideroxyleae crown node (node B in Fig. 7.2). This fossil, along with two other fossil taxa was previously used in a study of the Sideroxyleae by Smedmark & Anderberg (2007). The two other fossils: *Sideroxylon aequale* leaves from the Mid-Oligocene of Puerto Rico (Hollick 1928) and *Bumelia retusaefolia* leaves from the Mid-Miocene of Cuba (Berry 1939), were also considered for placement in the phylogeny. However, the *Sideroxylon aequale* specimen, held by the Peabody Museum, was unavailable for scrutiny, and the *Bumelia retusaefolia* specimen, held at the Smithsonian Institute, was determined by Terry Pennington to more closely resemble Piperaceae than Sapotaceae. Neither fossil was, therefore, included as a calibration point in these analyses.

A Mid-Eocene, ~45 Ma, *Tetracolporpollenites* pollen grain from the Isle of Wight was used to constrain the minimum age of the node for the tribe Mimusopeae. This pollen grain was described by Harley (1991) and determined to closely resemble *Tieghemella heckelii* (a monotypic genus in the Mimusopeae). Harley later suggested (pers. comm. 2010) that it would be appropriate to err on the side of caution with the identification and use the fossil to constrain the age of the tribe Mimusopeae rather than the genus itself. This fossil was, therefore, used to constrain the age of the crown node of Mimusopeae (node C in Fig. 7.2).

Three additional fossil pollen grains were cross-examined by Harley for this study as potential candidates for calibrating the Mimusopeae node, but they were not as well characterized or studied as the chosen *Tetracolporpollenites* pollen grain and so were not used. Those fossil pollen taxa, from a range of localities and epochs, were:

1. *Sapotaceoidaepollenites rotundus* (Stoian 2002) Late Cretaceous-Paleocene, Australia (possible Mimusopeae according to Harley)

2. *Sapotaceoidaepollenites kirchheimeri* (Song *et al* 1999 & 2004) Miocene, China (Mimusopeae or Isonandreae according to Harley)

3. *Mimusops-Manilkara* (Machin 1971) Oligocene, Isle of Wight, England (Mimusopeae, Isonandreae, Sideroxyleae and/or *Capurodendron* according to Harley)

The final calibration point is based on a series of Oligocene (~28 Ma) fossil leaves from Ethiopia (Jacobs *et al*, 2005). Aaron Pan (Fort Worth Museum) has described these specimens as *Sapoteae* sp. and suggested possible placement in either *Manilkara* or *Tieghemella* (pers. comm. 2010) based on the occurrence of stoma surrounded by fimbricate periclinal rings, a character present in these genera, but absent from the related genera *Autranella* and *Mimusops*. According to Pan, a number of fossil Sapotaceae pollen types are also known from the same area (Yemane *et al* 1987, Kappelman *et al* 2003), which

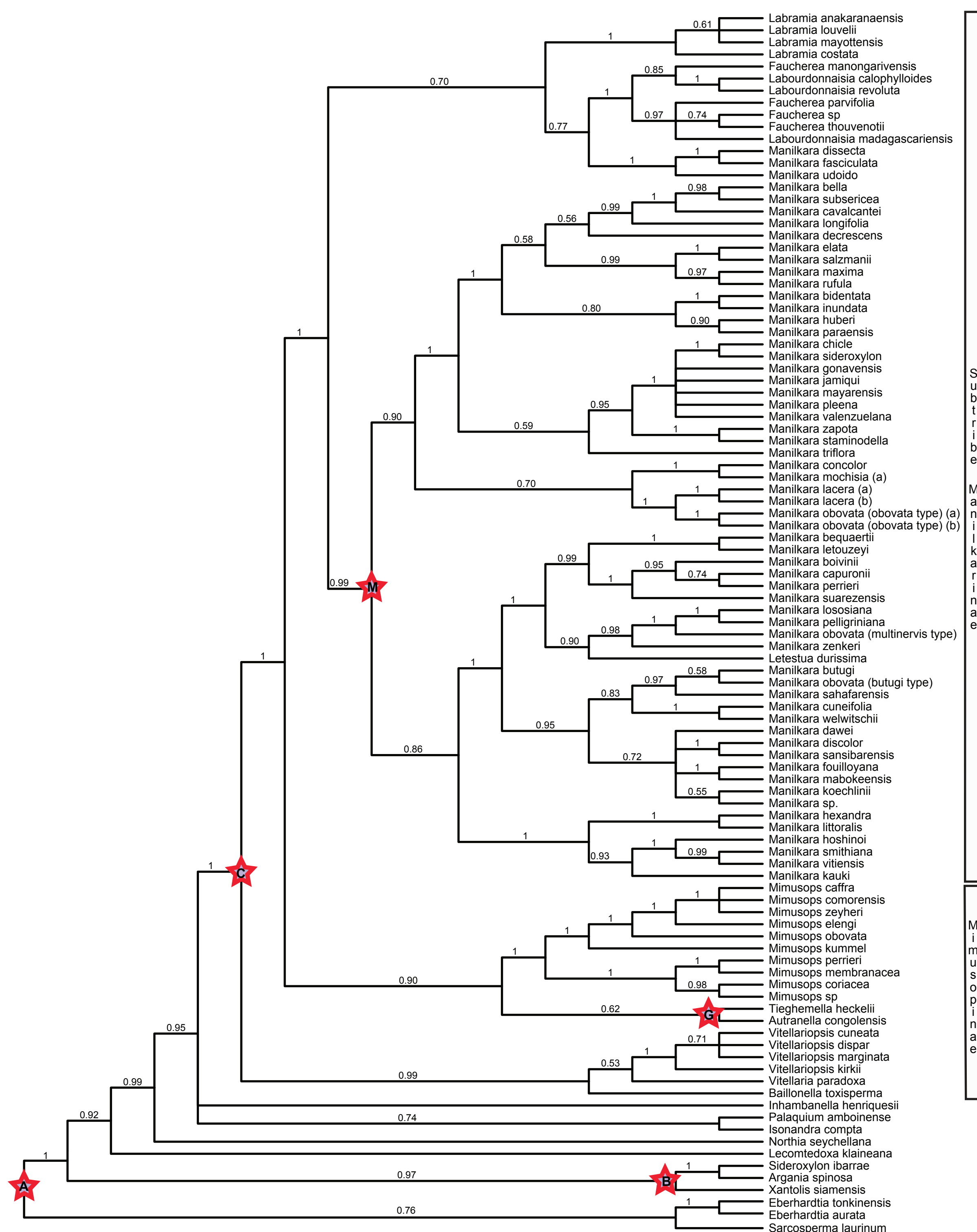


Figure 7.2 Cladogram, adapted from Fig. 4.2, depicting placement of fossil taxa for testing calibration scenarios. Note that analyses in this chapter include many more outgroups than are illustrated in this cladogram, so the depth of node B is not accuratetly portrayed here. See table 7.1 for details of calibration points and taxon sets as defined in BEAST.

were originally dated as 8 Ma but are now considered to be the same age as the macrofossil material, corroborating the presence of the family in Ethiopia during the Oligocene. As these Oligocene fossil leaves had been well-studied and categorized, they were a strong candidate for use in fossil calibration. However, placement in the phylogeny was complicated by the fact that *Manilkara* and *Tieghemella*, although in the same tribe (Mimusopeae), are not sister taxa and placing the fossil at the node of the most recent common ancestor (the entire Tribe Mimusopeae) seemed illogical for such a young date when a 45 Ma fossil pollen grain of cf. *Tieghemella* was a better fit for the same node. It was, therefore, decided to test the utility of this fossil (being the youngest calibration point) alternatively for *Manilkara* and *Tieghemella* in order to determine whether placement on either genus made a significant difference to the age estimates. The fossil was, therefore, placed alternatively on the *Manilkara* crown node (node M in Fig. 7.2) and on the node of the split between *Tieghemella* and *Autranella* (node G in Fig. 7.2).

Lastly, the well-preserved fossil fruit of *Mimusops emarginata* Berry (= *Manilkara jaimiqui*) from the Pleistocene of Trinidad was determined by Terry Pennington to be similar to extant species and of potential value as a calibration point for this study. However, as the fossil is very young, its placement at the tip of the phylogeny would be unlikely to affect the age estimates and so was not used in this analysis.

Because these calibration points are primarily based on fossil ages, they are considered to be minimum age estimates and where possible were, therefore, assigned to the crown group, rather than the stem group to bias in favour of older age estimates. Calibration points and the nodes upon which they were placed are listed in Table 7.1 and illustrated in Fig. 7.2.

7.3.3 Dataset for dating analyses: taxa and outgroup selection

Divergence times were calculated using an expanded ITS dataset with 170 accessions. Chloroplast data were not included because they were not informative enough to discern between alternative hypotheses and because fewer taxa were sampled. The ITS taxon set includes the ingroup tribe Mimusopeae as well as multiple representatives of the tribes Isonandreae and Sideroxyleae to accommodate calibration of fossils related to those groups. Sideroxyleae and some Mimusopeae sequences were donated by Jenny Smedmark and Arne Anderberg at the Natural History Museum in Stockholm or were taken from Smedmark & Anderberg (2007). Sequences designated as being from Geneva were contributed by Yamama Naciri and Laurent Gautier. The tree was rooted using *Sarcosperma*, which has been shown in previous studies to be sister to the rest of the family (Anderberg & Swenson 2003). See Appendix 7.1 for a list of specimens used in the analyses in this chapter.

7.3.4 Bayes factor tests for model selection and clock-likeness in MrBayes & BEAST

Prior to estimating node ages in BEAST, the applicability of a molecular clock to the ITS dataset was tested. An assessment of the clock-likeness of the data was made in MrBayes v.3.1 (Huelsenbeck & Ronquist 2001) by comparing Bayes factors (as previously outlined in Chapter IV) between an unconstrained non-clock, a uniform clock and a birth-death strict

clock model. Substitution models were also tested on the expanded ITS dataset in MrBayes using Bayes factors, with a difference >10 indicating significant support of one model over another (Kass & Rafferty 1995) (see Table 7.2). Partitioning strategies were tested on the dataset in a previous analysis in Chapter IV, and so were not replicated here.

Two independent runs of four MCMCMC chains each (three heated and one cold) were run with a temperature setting of 0.10 for 8,000,000 generations. Trees were sampled every 8,000 generations and a 10% burn-in was removed from the sampled set of trees, leaving a final sample of 800 trees. Convergence of models was determined to have occurred when the standard deviation of split frequencies for two runs reached 0.01 (Ronquist *et al* 2005). This was backed-up by visual confirmation of parameter convergence in Tracer v.1.5 (Rambaut & Drummond 2009).

Subsequent to testing clock-likeness of the data in MrBayes, a further test was carried out in BEAST v.1.5.3 to determine whether a strict clock or a relaxed clock would be a more appropriate model to apply to the dataset (Table 7.5). Analyses were run following the methods outlined in section 7.3.5.1 below and Bayes factors were compared between the two models as outlined above.

7.3.5 Molecular dating analyses

7.3.5.1 BEAST program and settings

Table 7.1 Taxon sets defined in BEAST. Nodes are referenced in the phylogeny in Figure 7.7. Note that *Manilkara dissecta*, *M. fasciculata* and *M. udoido* were not included in the monophyly-constrained *Manilkara* taxon set as these species were shown in a previous analysis in Chapter IV to make *Manilkara* paraphyletic and to be morphologically more similar to *Faucherea/Labourdonnaisia*.

Taxon set name	Calibration point reference	Age	Node	Taxa included	Distribution
Sapotaceae s.s.	Bremer *et al* 2004 Estimate	102 Ma	A	All taxa except *Sarcosperma* (outgroup)	Normal
Sideroxyleae	Gruas-Cavagnetto 1976 fossil pollen	49 Ma	B	*Argania* and all *Sideroxylon/Bumelia* except *S. oxycanthum*	Lognormal
Mimusopeae	Harley 1991 fossil pollen	45 Ma	C	*Autranella Baillonella Faucherea Labourdonnaisia Labramia Letestua Manilkara Mimusops Tieghemella Vitellariopsis Vitellaria*	Lognormal
Manilkara	Jacobs *et al* 2005 fossil leaf	28 Ma	M	*Letestua* and all *Manilkara* (except for *M. dissecta, M. fasciculata, M. udoido*)	Lognormal
Tieghemella	Jacobs *et al* 2005 fossil leaf	28 Ma	G	*Tieghemella* & *Autranella*	Lognormal

The software package BEAST v.1.5.3 (Drummond & Rambaut 2007) was used to analyse divergence times. An XML (eXtensible Mark-up Language) input file was created in BEAUti (Bayesian Evolutionary Analysis Utility software) version v.1.5.3. The ITS dataset was partitioned into three segments: ITS1 (374 bases long), 5.8s (164 bases long) and ITS2 (343 bases long). Substitution models were unlinked across partitions, but clock models and tree topologies were kept on the linked default setting.

Five taxon sets were generated in order to define nodes for placement of fossil calibration points (Table 7.1). They were based on known monophyletic clades from previous analyses in MrBayes (Chapter IV) and were constrained to be monophyletic.

The GTR + I + G model (general time reversible model, plus gamma distributed rate variation, plus a proportion of invariant sites) was applied to each partition (Table 7.3). The mean substitution rate was not fixed and base frequencies were estimated.

In order to relax the assumption of a molecular clock and allow for rate heterogeneity between lineages, an uncorrelated lognormal model was selected. As mentioned above, a strict clock model was also tested (Table 7.5). The tree prior was set to Speciation: Birth-Death Process (because this model most closely resembles the process of speciation) with a randomly generated starting tree.

The most recent common ancestor (tmrca) node age priors were set to define calibration points for taxon sets as outlined in Table 7.1. A normal distribution models non-directional uncertainty (both younger and older ages) and is, therefore, appropriate for modelling age estimates from secondary calibrations, whereas a lognormal distribution is used to model the assumption that a speciation event is likely to have occurred before the actual appearance of the fossil. The first calibration point (Sapotaceae *s.s.*), being based on a secondary calibration from another phylogeny rather than a fossil age, was set using a normal distribution around the mean age with a standard deviation of one. The remaining three fossil-based calibration points were set using a lognormal distribution offset to the minimum age of the fossil with a mean and standard deviation of one. All other priors were left at default settings which were either uniform or gamma.

Posterior distributions for each parameter were estimated using a Metropolis Coupled Monte Carlo Markov Chain (MCMCMC) run for 40,000,000 generations with parameters logged every 5,000 generations, equalling 8,000 generations per run. The BEAUti XML file was executed in BEAST v.1.5.3 (released December 2009). Two separate analyses were run and the output log files were reviewed in Tracer v.1.5 (Rambaut & Drummond 2009) to check for convergence of the two runs and that effective sampling size (ESS) values for all parameters were sufficient (i.e. > 200) (Drummond *et al* 2007).

The tree files from the two runs were combined in LogCombiner v.1.5.3 (Drummond & Rambaut 2007) with a conservative burn-in of 4,000 generations specified (this is half of the 8,000 generations, leaving 4,000 samples per run). When the post burn-in trees from the two runs are combined this leaves a final sample size of 8,000 trees. The combined tree files

were input into TreeAnnotator v.1.5.3 (Drummond & Rambaut 2007). The Maximum Clade Credibility Tree was selected with mean node heights; this option summarises the tree node height statistics from the posterior sample with the maximum sum of posterior probabilities. The output file was visualised in FigTree v.1.3.1.

7.3.5.2 r8s program settings

The program r8s (Sanderson 2003) was used as an additional test of age estimates for nodes against results from BEAST. Two different methods were implemented in r8s: penalized likelihood (PL) and nonparametric rate smoothing (NPRS).

The penalized likelihood method uses a cross-validation procedure to find the optimum rate-smoothing parameter for the transition of substitution rate between ancestor and descendent lineages. Four cross-validation analyses were run in order to test whether they would converge on the same solution; one with the consensus tree and the remaining three with randomly selected, single post-burn-in trees from the MrBayes analysis. Cross-validation analyses resulted in an optimum smoothing parameter of 100 for each of the three individual Bayesian trees, while the cross-validation procedure on the consensus tree suggested a smoothing parameter of 10. These two smoothing parameters were tested on the dataset and resulted in very similar age estimates. For the majority of nodes the age estimates in each validation were nearly identical. Four estimates differed by only one million years (two nodes were 1My younger, and two nodes were 1My older), which is not significant enough to bias the outcome of hypothesis testing. Since support for the two alternative smoothing parameters was equivocal, the rate-smoothing parameter of 10 based on the three single tree analyses was chosen and applied to the penalized likelihood analysis.

Both the penalized likelihood and nonparametric rate smoothing analyses were then run in r8s with 100 trees, comprising the last 50 post-burn in trees from each of two runs of MrBayes, combined into a block. All calibration points (Table 7.1) were fixed and tested in turn to compare with the BEAST analysis. *Sarcosperma laurinum* was used to root the trees, but then was automatically pruned (i.e. excluded) prior to analysis. Additionally, in the PL analyses small non-zero bounds were imposed by the program for zero-length branches in the Bayesian input trees because the program cannot handle branches with no length. The number of iterations was set to 2000 and the number of time guesses (i.e. estimates for the age of a node) was set to three, while all other settings were left on default values.

7.3.5.3 Testing the utility of different fossil calibration points in BEAST and r8s

In both the programs BEAST and r8s, each of the selected fossils were tested in turn singly and in combination to determine whether one fossil placement resulted in significantly different age estimates than the others. This test also included the effect of placement of the *Manilkara*/*Tieghemella* fossil on the crown node of *Manilkara* and on the stem node of *Tieghemella*. (Being monotypic, *Tieghemella* does not have a crown node.)

Fossil calibration combinations comprised:

1. Sapotaceae independently (Sap)
2. Sideroxyleae independently (Sid)
3. Mimusopeae independently (Mim)
4. *Manilkara* independently (Man)
5. *Tieghemella* independently (Tie)
6. Sapotaceae + Sideroxyleae (SapSid)
7. Sapotaceae + Mimusopeae (SapMim)
8. Sapotaceae + *Manilkara* (SapMan)
9. Sapotaceae + *Tieghemella* (SapTie)
10. Sapotaceae + Sideroxyleae + Mimusopeae (SapSidMim)
11. Sapotaceae + Sideroxyleae + *Manilkara* (SapSidMan)
12. Sapotaceae + Sideroxyleae + *Tieghemella* (SapSidTie)
13. Sapotaceae + Mimusopeae + *Manilkara* (SapMimMan)
14. Sapotaceae + Mimusopeae + *Tieghemella* (SapMimTie)
15. Sapotaceae + Sideroxyleae + Mimusopeae + *Manilkara* (SapSidMimMan)
16. Sapotaceae + Sideroxyleae + Mimusopeae + *Tieghemella* (SapSidMimTie)

7.3.5.4 Fossil cross-validation

Choice of calibration strategy was further tested with the fossil cross-validation method outlined in Near & Sanderson (2004) and Near *et al* (2005), which is a procedure used to identify the impact of different individual calibrations on overall age estimation. It is particularly useful in pinpointing fossils with a large error effect which may skew the outcome of analyses. This calculation was based on the BEAST analyses using single calibration points (1-5 in the list above).

Near & Sanderson's procedure entails that in a phylogeny with multiple fossil-dated nodes, the age of a single fossil-dated node is fixed and the difference between molecular and fossil estimates for all other fossil-dated nodes in the phylogeny are calculated. To begin, the difference between fossil and molecular ages were calculated using a single fossil-dated node, χ, which was defined as $D_i = (MA_i - FA_i)$, where FA_i is the fossil age estimate and MA_i is the molecular age estimate for node *i* (Table 7.4).

The mean percentage deviation ($\overline{D}\chi$) between molecular and fossil age estimates is then calculated as follows:

$$\overline{D}_\chi = \frac{\sum_{i \neq \chi} D_i}{n-1}$$

This is followed by a further two steps to identify and remove inconsistent fossils from the analysis. Step one involves determination of the sum of the squared differences:

$$SS_\chi = \sum_{i \neq \chi} D_i^2$$

Each calibration point can then be ranked based on the magnitude of SS. The greater the SS value, the more inconsistent the fossil is with respect to the other fossils in the analysis.

In the second step, the average squared deviation (s), for all fossils in the analysis was calculated as follows:

$$s = \frac{\sum_{\chi=1}^{n} \sum_{i \neq \chi} D_i^2}{n(n-1)}$$

Near *et al* (2004) recommend continuing this process of removing the fossil with the greatest SS value and recalculating with the remaining calibration points until the magnitude of s decreases by only a small fraction as fossils are removed. They further suggest that “removal from the analysis of extreme outliers that provide very inaccurate calibrations with respect to other fossils should cause an appreciable drop in s”

7.3.6. Ancestral area reconstruction in BEAST

Ancestral area states were reconstructed utilizing the same dataset and basic methodology as in the dating analysis above in section 7.3.5.1. Areas were coded for each accession by editing the BEAUti XML file according to the phylogeographic tutorials available on the BEAST website (http://beast.bio.ed.ac.uk/Discrete_Phylogeographic_Analysis).

Regions were coded as follows:
1. Sahul Shelf: Malesia west of Wallace’s Line (from Malaysia to Bali)
2. Sunda Shelf: Malesia east of Wallace’s Line (from Sulawesi east to Fiji)
3. East Asia: Continental Asia east of the Himalayas, including Burma & Indochina
4. South Asia: India & Sri Lanka
5. Middle East: Iran to Turkey and the Arabian Peninsula
6. Seychelles
7. Madagascar: Madagascar, Reunion and the Comoros
8. Africa: Continental Africa, the Canary Islands and the Cape Verde islands
9. North America: U.S.A., Mexico, Central America and the Caribbean
10. South America

Areas are coded according to continent, based predominantly on tectonic plate margins and then on floristic regions (Fig. 7.3). In Southeast Asia, the Sahul and Sunda Shelves mark the boundary between continental Asia and Australia-New Guinea, whereas Malesia is a floristic region, which stretches from the Isthmus of Kra on the Malay Peninsula to Fiji. East Asia is segregated as being east of the Himalayas and south as far as the Malay Peninsula, with a predominantly Indo-Chinese flora. South Asia is delineated by the margin of the Indian subcontinent. The countries of Iran, Turkey and the Arabian Peninsula support a drier Irano-Turanian flora and were, therefore, designated as being part of the Middle-Eastern region. The remaining regions (the Seychelles, Madagascar, Africa and North and South America) are all on separate continental tectonic plates and are floristically unique from one another. See Appendix 7.1 for species-specific area codes.

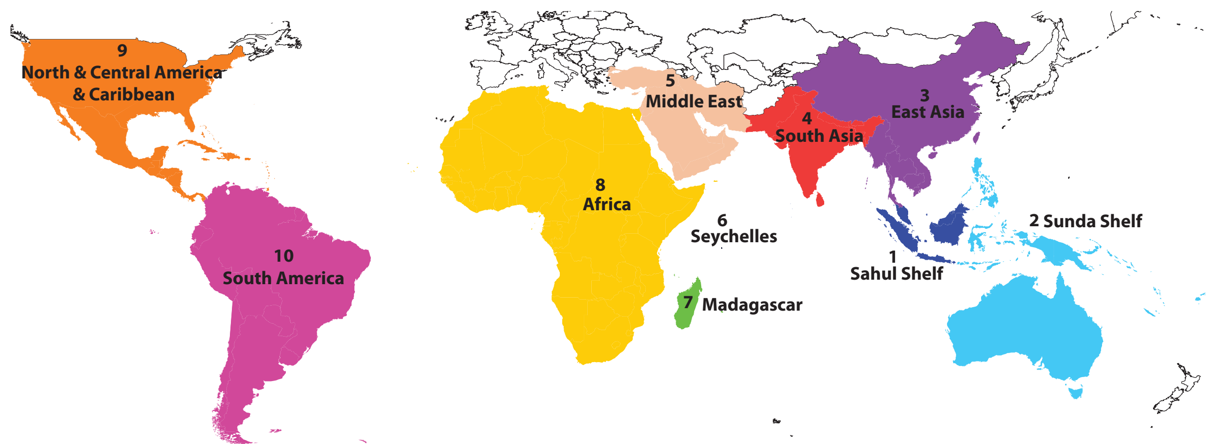


Figure 7.3 Map of regions coded for taxa in the BEAST ancestral area analysis. North America & Caribbean = orange, South America = pink, Africa = yellow, Madagascar = green, Seychelles = slate blue, Middle East = peach, India = red, East Asia = purple, Sunda shelf = royal blue, Sahul shelf = aqua blue.

The SapMim fossil combination (section 7.3.5.3) was chosen to calibrate the phylogeny, although calibration choice should not have a large impact on the outcome, given that the primary aim of this analysis is to reconstruct ancestral states rather than ages. Two separate analyses were run in BEAST and the output log files were reviewed in Tracer v.1.5 (Rambaut & Drummond 2009). Tree files were then combined in LogCombiner v.1.5.3 (Drummond & Rambaut 2007) and a maximum clade credibility tree was generated in TreeAnnotator v.1.5.3. (Drummond & Rambaut 2007) before being visualised in FigTree v.1.3.1.

7.3.7 Hypothesis testing with area constraints in MrBayes

Support for area monophyly was tested in MrBayes by constraining species which occur in a particular region to be monophyletic and making Bayes factor comparisons with the unconstrained model. Using a reduced dataset, only the *Manilkara* species were constrained (not outgroups) and only major continental areas were coded as follows: Africa, Madagascar, Asia and the Neotropics. See Appendix 7.1 for species-specific area codes.

Partitioning strategies were tested on the ITS dataset in a previous analysis in Chapter IV and so were not replicated here. Results indicated that a three-partition model was most appropriate: ITS1, 5.8S and ITS2. Model fitness and clock-likeness were tested with Bayes factor comparison on the dataset (Table 7.9). As a result the GTR+I+G model was applied to ITS1 and ITS2, whereas a GTR+I model was applied to the 5.8S region and a birth-death strict clock was selected. Two independent runs of four MCMCMC chains each (three heated and one cold) were run with a temperature setting of 0.10 for 8,000,000 generations. Trees were sampled every 8,000 generations and a 10% burn-in was removed from the sampled set of trees, leaving a final sample of 800 trees. Convergence of models was determined to have occurred when the standard deviation of split frequencies for two runs reached 0.01. This was backed-up by visual examination of stationarity of traces in Tracer v.1.5 (Rambaut & Drummond 2009).

7.4 Results

7.4.1 Molecular dating results

7.4.1.1 Bayes factor tests for model selection

Bayes factor tests revealed that a GTR+I+G model (Tables 7.2 & 7.3) was the most appropriate for each of the ITS dataset partitions (ITS1, 5.8S and ITS2). An Akaike Information Criterion test in MrModeltest ver. 2.3 (Nylander 2008) favored the same model selection. A birth-death strict clock model was supported over non-clock and uniform clock models in a Bayes factor comparison (Table 7.2). This model and partitioning strategy were applied to analyses in BEAST.

Table 7.2 Bayes factor comparison in MrBayes of different models and clock settings. The two competing best fit models are emboldened. A Bayes factor value of 11.82 shows significant support for the model with three partitions all with a GTR+I+G and a birth-death clock.

Models per partition and clock strategy imposed	**Ln HML**	**Bayes factor: 2ΔlnHML comparison of values shown between the chosen model and each alternative**
ITS1 - GTR+I+G 5.8S - GTR+I+G ITS2 - GTR+I+G No clock	-13284.58	168.34
ITS1 - GTR+I+G 5.8S - GTR+I+G ITS2 - GTR+I+G Uniform clock	-13226.61	52.4
ITS1 - GTR+I+G **5.8S - GTR+I+G** **ITS2 - GTR+I+G** **Birth-death clock**	**-13200.41**	**Chosen model**
ITS1 - GTR+I+G 5.8S - GTR+I ITS2 - GTR+I+G Birth-death clock	-13228.00	55.18
ITS1 - GTR+I+G **5.8S - GTR+G** **ITS2 - GTR+I+G** **Birth-death clock**	-13206.32	11.82
ITS1 - GTR+I+G 5.8S - GTR+I ITS2 - GTR+I+G No clock	-13396.85	392.88
ITS1 - GTR+I+G 5.8S - GTR+G ITS2 - GTR+I+G No clock	-13344.89	288.98

Table 7.3 Summary of sequence data and settings in BEAST

partition	aligned length	parsimony informative sites	variable sites	model	clock setting
ITS 1	374 bases	187	247	GTR + I + G	birth-death clock
5.8S	164 bases	24	48	GTR + I + G	birth-death clock
ITS2	343 bases	181	242	GTR + I + G	birth-death clock

7.4.1.2 Testing the utility of different fossil calibration points in BEAST and r8s

Three nodes at different depths in the phylogeny, which were not set as fixed calibration points, were chosen to illustrate the difference in age estimates between alternative calibration strategies and analyses; they are: the subtribe Manilkarinae (Eocene-Oligocene age) at node H in Fig. 7.7, the clade containing *Faucherea*, *Labourdonnaisia* and three Asian *Manilkara* species (Eocene-Miocene) at node J in Fig. 7.7 and the Neotropical *Manilkara* clade (Oligocene-Miocene) at node N in Fig. 7.7. The range of ages reconstructed for each of these nodes is represented in Figures 7.4, 7.5 and 7.6 respectively. For any one calibration strategy, the mean values given by the different analyses are typically very similar, being approximately 1-4 My different. The exceptions to this are the single basal calibrations for which ages can differ by up to 10 My (e.g. the difference in age estimate between BEAST/ NPRS and PL in Fig. 7.4 for the Sapotaceae calibration point). In summary, the different analyses in BEAST, PL and NPRS return generally similar values, but the choice of calibration points affects the outcome more than the type of analysis. See Appendices 7.2, 7.3 & 7.4 for the full range of ages reconstructed for each profiled node using each of the fossil calibration scenarios and dating methods.

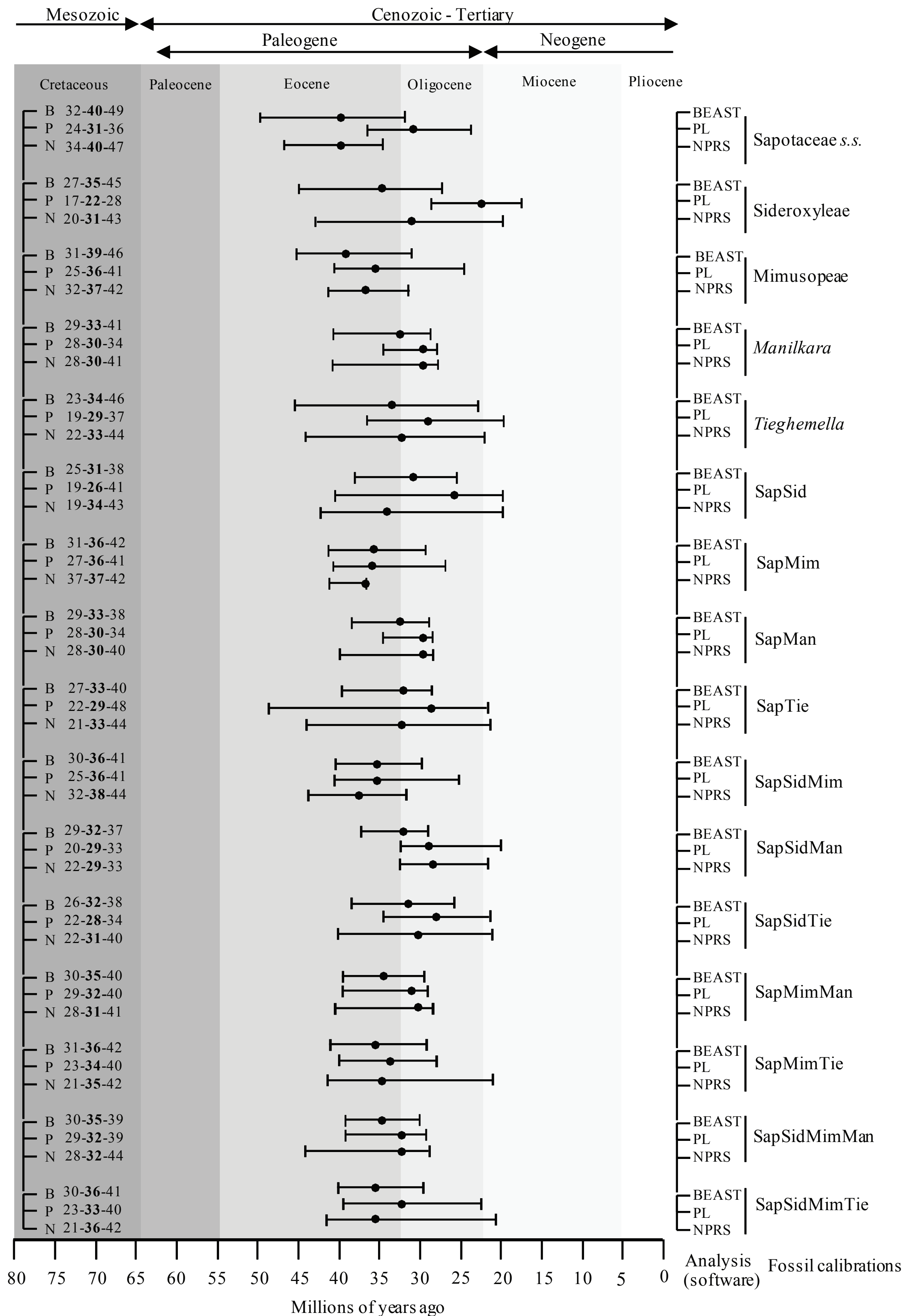


Figure 7.4 Age ranges reconstructed for the subtribe Manilkarinae, node H, using each of the fossil calibration scenarios and each of the molecular dating methods: a relaxed uncorrelated lognormal clock in BEAST, penalized likelihood in r8s and nonparametric rate smoothing in r8s. Circles represent mean node ages and lines represent upper and lower age bounds. Calculated ages are given on the left side of the diagram with the mean age emboldened. Geological epochs are represented by vertical bars in different shades of grey. The overall mean age range for node H is 22-31-40.

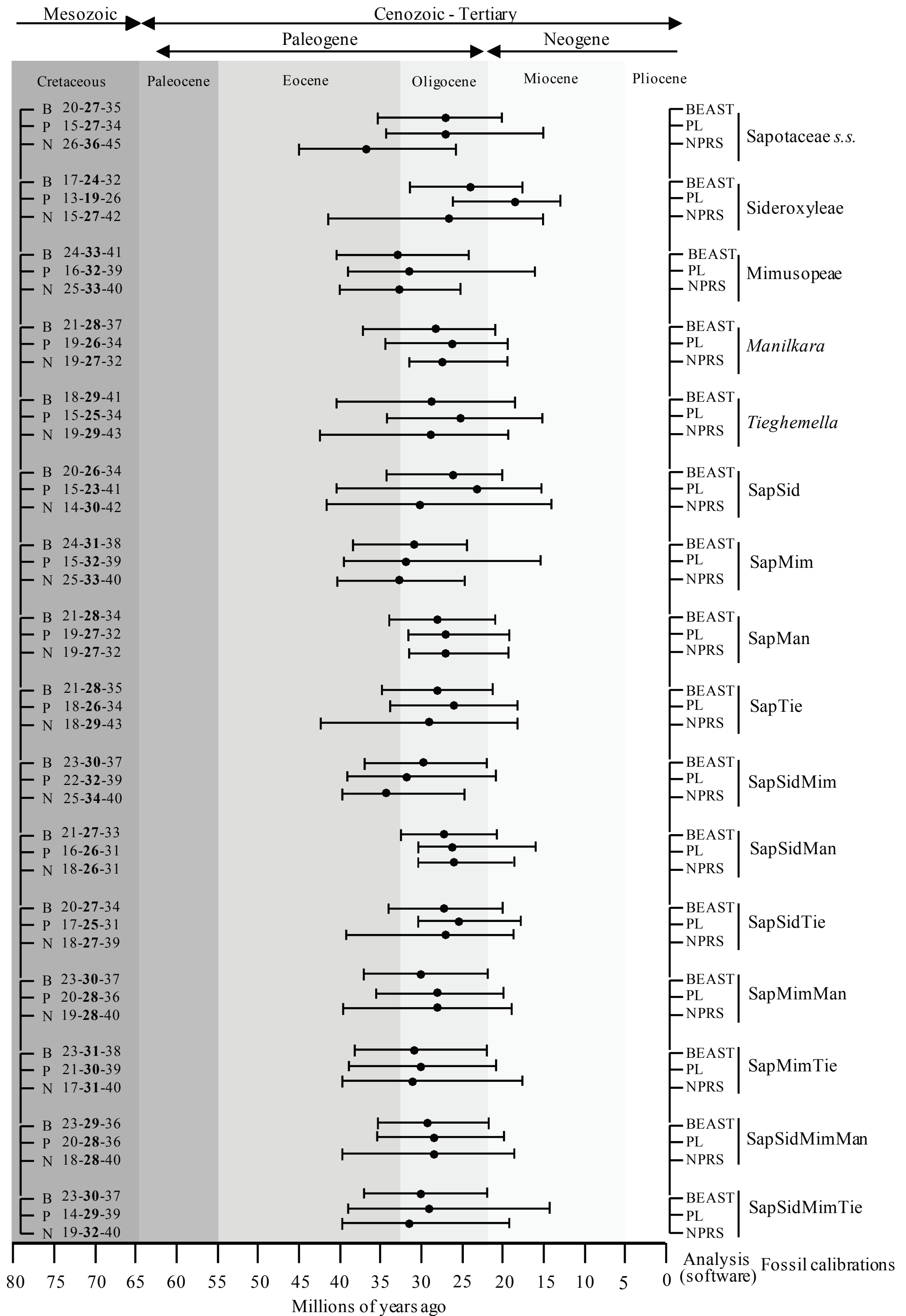


Figure 7.5 Age ranges reconstructed for the *Faucherea*, *Labourdonnaisia* & small *Manilkara* clade, node J, using each of the fossil calibration scenarios and each of the molecular dating methods: a relaxed uncorrelated lognormal clock in BEAST, penalized likelihood in r8s and nonparametric rate smoothing in r8s. Circles represent mean node ages and lines represent upper and lower age bounds. Calculated ages are given on the left side of the diagram with the mean age emboldened. Geological epochs are represented by vertical bars in different shades of grey. The overall mean age range for node J is 19-27-36.

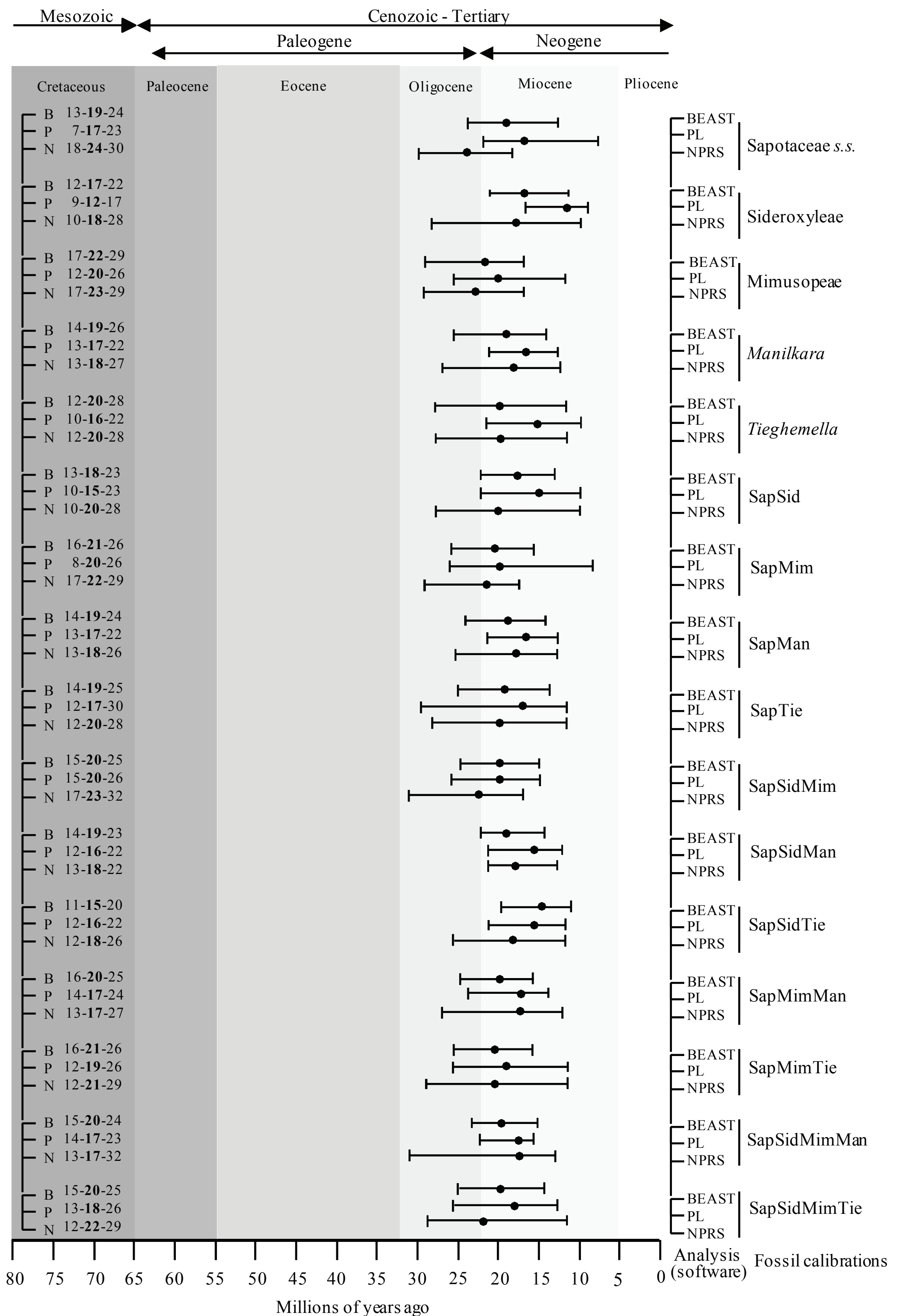


Figure 7.6 Age ranges reconstructed for the Neotropical *Manilkara*, node N, using each of the fossil calibration scenarios and each of the molecular dating methods: a relaxed uncorrelated lognormal clock in BEAST, penalized likelihood in r8s and nonparametric rate smoothing in r8s. Circles represent mean node ages and lines represent upper and lower age bounds. Calculated ages are given on the left side of the diagram with the mean age emboldened. Geological epochs are represented by vertical bars in different shades of grey. The overall mean age range for node N is 12-18-24.

7.4.1.3 Choice of fossil calibration scenario and analytical method

The different calibration strategies were tested with the fossil cross-validation method outlined in Near & Sanderson (2004).

Step 1a. The difference between the molecular and fossil ages calculated using a single fossil-dated node. χ was defined as $D_i = (MA_i - FA_i)$, where MA_i is the fossil age estimate and FA_i is the node age estimate for node *i* (Table 7.4).

Table 7.4 The molecular and the fossil age for each calibration point

Calibration MA_i/FA_i	Sapotaceae *s.s.*	Sideroxyleae	Mimusopeae	*Manilkara*	*Tiegehmella*
Sapotaceae *s.s.*	113/102	59/49	40/45	30/28	29/28
Sideroxyleae	99/102	52/49	35/45	26/28	25/28
Mimusopeae	133/102	69/49	48/45	35/28	35/28
Manilkara	116/102	60/49	41/45	31/28	30/28
Tiegehmella	116/102	60/49	42/45	31/28	31/28

Step 1b. The mean percentage deviation ($\overline{D}\chi$) between molecular and fossil age estimates:

$$\overline{D}_\chi = \frac{\sum_{i \neq \chi} D_i}{n-1}$$

Sapotaceae *s.s.* node $\overline{D}\chi = \frac{10+5+2+1}{5-1} = 4.5$

Sideroxyleae node $\overline{D}\chi = \frac{3+10+2+3}{5-1} = 4.5$

Mimusopeae node $\overline{D}\chi = \frac{31+20+7+7}{5-1} = 16.25$

Manilkara node $\overline{D}\chi = \frac{14+11+4+2}{5-1} = 7.75$

Tieghemella node $\overline{D}\chi = \frac{14+11+3+3}{5-1} = 7.75$

Step 1c. Determination of the sum of the squared differences:

$$SS_\chi = \sum_{i \neq \chi} D_i^2$$

Sapotaceae *s.s.* node SSχ = 100+25+4+1 = 130

Sideroxyleae node SSχ = 9+100+4+9 = 122

Mimusopeae node SSχ = 961+400+49+49 = 1459

Manilkara node SSχ = 196+121+16+4 = 337

Tieghemella node SSχ = 196+121+9+9 = 335

This calculation makes it clear that the Mimusopeae calibration point is potentially skewing the outcome of the analysis, followed by the *Manilkara* and *Tieghemella* fossils.

Step 2. Average squared deviation (s), for all fossils in the analysis:

$$s = \frac{\sum_{\chi=1}^{n} \sum_{i \neq \chi} D_i^2}{n(n-1)}$$

Average squared deviation for all fossils in analysis: $\frac{2383}{20} = 119.15$

Average squared deviation for Mimusopeae: $\frac{924}{20} = 46.2$

Average squared deviation for *Manilkara*: $\frac{587}{20} = 29.35$

Average squared deviation for *Tieghemella*: $\frac{252}{20} = 12.6$

Both the fossil cross-validation calculations and the node age profiles in Figures 7.4, 7.5 and 7.6 make it clear that the Mimusopeae fossil biases towards older age estimates and according to Near & Sanderson's (2004) method should be removed from the analysis. However, as the Mimusopeae pollen grain had been well studied and characterised, it was deemed to be the most reliable fossil in the analysis. Therefore, rather than remove it, age estimates are provided for three different calibration scenarios, taking into consideration this bias. The chronogram in Figure 7.7 represents age estimates from the Sapotaceae *s.s.* + Mimusopeae (SapMim) calibration point only. In the text, this calibration point as well as the Sapotaceae *s.s.* + Sideroxyleae + *Manilkara* (SapSidMan) and Sapotaceae *s.s.* + Sideroxyleae + Mimusopeae + *Manilkara* (SapSidMimMan) calibrations are discussed. Note that for the most part, age estimates do not vary enough to affect the choice of which hypothesis best fits the age estimates (see Figs. 7.4, 7.5 & 7.6 of profiled nodes), except in instances where the node falls on the boundary of an epoch. This is the case in the subtribe Manilkarinae (Fig. 7.4), which may be Eocene or Oligocene in age and, therefore, the data is equivocal in terms of providing evidence for boreotropical migration versus long distance dispersal (see section 7.5.1.3 below).

Choosing to use the *Manilkara* rather than *Tieghemella* rate is arbitrary as both give very similar age estimates for nodes (in BEAST analyses, typically only 1My different if not less). As there are more extant species of *Manilkara* than *Tieghemella*, *Manilkara* was chosen as the representative fossil calibration point out of the two.

Lastly, age estimates from BEAST are represented in the chronogram and were chosen for further discussion. This is because BEAST is the only program out of the three in which estimates of ages are uncorrelated across the tree and divergence times, tree topology and evolutionary rates are estimated as part of the same calculation. NPRS has been shown to overestimate ages and has, therefore, recently fallen out of favour methodologically. According to Rutschmann (2006), "a serious limitation of NPRS is its tendency to over-fit the data, leading to rapid fluctuations in rate in regions of a tree that have short branches."

Additionally, PL often appears to give younger age estimates. This trend was also found by Renner 2005 and by Goodall-Copestake *et al* 2009, who tested different age calibration methods. Zero length branches can also be problematic in the cross validation process in PL, but the program gets past this issue by imposing small non-zero branch lengths on the Bayesian input trees. In both r8s programs, PL & NPRS, the most basal outgroup is pruned, and in this way the analyses are not exactly analogous to those in BEAST.

7.4.1.4 Relaxed uncorrelated lognormal clock versus a strict clock in BEAST

Both a strict clock and a relaxed uncorrelated lognormal clock model were tested for each of the three chosen calibration scenarios (SapMim, SapSidMan & SapSidMimMan) in BEAST. In a Bayes factor comparison, a relaxed clock was shown to significantly better fit the data (Table 7.5). Following Kass & Rafferty (1995), an Ln Bayes factor difference >10 was used to indicate decisive support of one model over another. In this test the difference between clock models is greater than 10, indicating very strong support for the use of a relaxed clock.

Table 7.5 Relaxed versus strict clock Bayes factor test

Calibration strategy	Ln HML	Bayes Factor: 2ΔlnHML
SapMim Strict clock	-13203.88	2ΔlnHML = 129.38 a relaxed clock is significantly better than a strict clock
SapMim Relaxed clock	-13139.19	
SapSidMan Strict clock	-13198.67	2ΔlnHML = 112.5 a relaxed clock is significantly better than a strict clock
SapSidMan Relaxed clock	-13142.42	
SapSidMimMan Strict clock	-13257.88	2ΔlnHML = 223.2 a relaxed clock is significantly better than a strict clock
SapSidMimMan Relaxed clock	-13146.28	

7.4.1.5 Node ages

The BEAST analysis calibrated with the SapMim fossil points resolves the mean crown age of the tribe Mimusopeae as 46 Ma, of Eocene age. The mean age of subtribe Manilkarinae is estimated to be 36 Ma, of late Eocene age and the genus *Manilkara* is shown to have originated during the Oligocene, 33 Ma. Results of the analysis also reveal that cladogenesis and inter-continental dispersal within *Manilkara* occurred throughout the Miocene, although most intensively from the mid-late Miocene. All three calibration scenarios (SapMim, SapSidMan, SapSidMimMan) estimate the profiled node ages within six million years of each other and are in agreement on geological epochs. Mean age estimates for profiled nodes of interest are represented in Table 7.6.

Figure 7.7 Maxium clade credibility chronogram of the ITS dataset derived from an analysis in BEAST using the GTR+I+G model and a relaxed clock for all three partitions, calibrated with the SapMim fossil scenario. Dashed lines indicate branches which lead to nodes with a posterior probability of <0.95. Mean ages are given for profiled nodes and node bars indicate 95% HPD age ranges. Lettered nodes are discussed in the text. Starred nodes indicate the placement of the Sapotaceae and Mimusopeae (SapMim) fossils (Table 7.1). Lineages are colored according to their distribution: Yellow = Africa, Green = Madagascar, Blue = Asia, Pink = South America, Orange = Central America & the Caribbean. Geological epochs are indicated in a scale at the bottom of the chronogram. Tribal and subtribal classification is represented to the right of the chronogram. Outgroups have been reduced to grey bars at the base of the chronogram. Age ranges of all profiled nodes are detailed in Appendix 7.2.

Table 7.6 Mean crown and stem node ages in BEAST for the profiled calibration scenarios. Note that for node R/S under the SapMim calibration point, the maximum clade credibility tree resolved two small African clades rather than one. In this case, the ages of those two clades are 6-27 & 9-29 Ma respectively.

Node in Fig. 7.7	Posterior probability (SapMim)	Clade	SapMim age in Ma	SapSidMan age in Ma	SapSid MimMan age in Ma	Epoch
C	1	Mimusopeae	46-57	40-49	46-56	Paleocene-Eocene
H	1	Manilkarinae	36-42	32-37	35-40	Eocene-Oligocene
M	0.99	*Manilkara s.s.*	33-36	30-32	31-34	Eocene-Oligocene
I	1	*Labramia*	7-36	6-30	6-35	Eocene-Miocene
J	0.92	*Faucherea/Labourdonnaisia/Manilkara*	31-36	27-30	29-34	Eocene-Oligocene
K	1	*Faucherea/Labourdonnaisia*	11-31	10-27	10-29	Oligocene-Miocene
L	1	Small Asian *Manilkara*	16-31	14-27	16-29	Oligocene-Miocene
T	1	Large African *Manilkara*	17-30	15-27	16-28	Oligocene-Miocene
U	1	Asian *Manilkara s.s.*	25-30	23-27	24-28	Oligocene-Miocene
N	1	Neotropical *Manilkara*	21-27	19-26	20-28	Oligocene-Miocene
P	0.87	C. American & Caribbean *Manilkara*	16-19	15-19	16-18	Miocene
O	1	South American *Manilkara*	14-19	12-17	13-18	Miocene
R/S	1/1	Small African *Manilkara*	6-27/9-29	21-26	22-28	Oligocene-Miocene

7.4.1.6 Estimated substitution rates

Substitution rates across the entire phylogeny estimated for the different fossil calibration scenarios in the program BEAST are represented in Table 7.7. They were recorded to gain a better understanding of how much rates differed amongst calibration scenarios. Rates varied from 1.31 x 10-9 substitutions per site per year (SSY) to 1.76 x 10-9 SSY depending upon the calibration strategies used (i.e. SapMim = 1.41 x 10-9, SapSidMan = 1.57 x 10-9, and SapSidMimMan = 1.48 x 10-9).

Table 7.7 Overall substitution rates calculated in BEAST for different fossil calibration scenarios

Fossil calibration scenarios	Overall substitution rate for trees run 1 & 2 in BEAST analyses
Sapotaceae *s.s.*	1.525E-3 & 1.556E-3
Sideroxyleae	1.761E-3 & 1.762E-3
Mimusopeae	1.310E-3 & 1.316E-3
Manilkara	1.524E-3 & 1.511E-3
Tieghemella	1.541E-3 & 1.527E-3
SapSid	1.612E-3 & 1.616E-3
SapMim	1.419E-3 & 1.418E-3
SapMan	1.525E-3 & 1.527E-3
SapTie	1.522E-3 & 1.508E-3
SapSidMim	1.471E-3 & 1.473E-3
SapSidMan	1.576E-3 & 1.575E-3
SapSidTie	1.586E-3 & 1.588E-3
SapMimMan	1.442E-3 & 1.434E-3
SapMimTie	1.420E-3 & 1.421E-3
SapSidMimMan	1.486E-3 & 1.481E-3
SapSidMimTie	1.475E-3 & 1.474E-3

7.4.2 Ancestral area reconstruction results

Previous analyses indicate that resolution along the backbone of the ITS phylogeny is weak and the area of origin is, therefore, difficult to determine. However, all sister taxa to *Manilkara* are African, this may suggest that the most likely explanation is an African origin for the genus with subsequent inter-continental dispersal during the Oligocene. This hypothesis was tested using ancestral area reconstruction in BEAST.

Figure 7.8 Maxium clade credibility chronogram of the ITS dataset derived from a phylogeographic analysis in BEAST using the SapMim fossil calibration scenario with ancestral areas reconstructed. Ten regions were coded as outlined in section 7.3.6 and illustrated in the map below. Lineages are colored according to their distribution. Dashed lines indicate branches which lead to nodes with a posterior probability of <0.95. Pie charts depict percentage likelihood of the ancestral state at the selected node. For a breakdown of area percentage likelihood per profiled node see Table 7.8.

The inclusion of areas as a prior in the BEAST analysis has a subtle effect on topology and age estimates. While the overall findings are very similar to those from the BEAST dating analysis, the estimates in this analysis are slightly different. In light of this, a separate phylogeny is presented here for the ancestral area reconstruction in Figure 7.8.

Tribe Mimusopeae, subtribe Manilkarinae and the genera *Manilkara*, *Labramia* and *Faucherea/Labourdonnaisia* are all found to be ancestrally African. Within *Manilkara*, Madagascan taxa are derived from an African ancestor. The Neotropical clade is also derived from an African ancestor, which dispersed to South America and subsequently to Central America and the Caribbean. Likewise, the Asian clade is derived from an African ancestor which dispersed east of Wallace's Line to the Sahul shelf. The genus *Mimusops* is also shown to be ancestrally African with two dispersal events to the Mascarenes and a single dispersal to the Sunda shelf. Percentage likelihood values for each area are given in Table 7.8 below.

Table 7.8 Percentage likelihood of ancestral areas for profiled nodes. Node letters relate to those given in Figure 7.8. Note - some areas (such as the Middle East and the Seychelles) are included in the analysis to cover outgroups which occur in those areas. Because outgroups are reduced for the sake of space in Figure 7.8, these areas are not represented in this table.

Node/Area	Sahul shelf	Sunda shelf	East Asia	South Asia	Middle East	Seychelles	Madagascar	Africa	North America	South America
Mimusopeae/ Isonandreae Node Z	0%	3%	1%	0%	0%	0%	4%	92%	0%	0%
Mimusopeae Node C	0%	0%	0%	0%	0%	0%	1%	99%	0%	0%
Mimusopeae Node Δ	0%	0%	0%	0%	0%	0%	1%	99%	0%	0%
Mimusops/ Tieghemella Node Γ	0%	0%	0%	0%	0%	0%	1%	99%	0%	0%
Mimusops Node F	0%	0%	0%	0%	0%	0%	9%	91%	0%	0%
Manilkarinae Node H	0%	0%	0%	0%	0%	0%	4%	96%	0%	0%
Labramia/ Faucherea- Labourdonnaisia Manilkara Node Σ	1%	0%	0%	0%	0%	0%	30%	69%	0%	0%
Faucherea- Labourdonnaisia Manilkara Node J	5%	0%	1%	0%	0%	0%	28%	66%	0%	0%
Manilkara s.s. Node M	0%	0%	0%	0%	0%	0%	2%	98%	0%	0%
Main Africa/Asian *Manilkara s.s.* Node Φ	2%	0%	0%	0%	0%	0%	2%	96%	0%	0%
Main African/ Madagascan *Manilkara* Node T	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%
Asian *Manilkara s.s.* Node U	90%	2%	2%	1%	0%	0%	0%	5%	0%	0%
Neotropical/ small African *Manilkara* Node Ψ	0%	0%	0%	0%	0%	0%	2%	98%	0%	0%
Neotropical *Manilkara* Node N	0%	0%	0%	0%	0%	0%	1%	6%	18%	75%
Small African *Manilkara* Node Q	0%	0%	0%	0%	0%	0%	1%	99%	0%	0%

7.4.3 Area constraints hypothesis testing results

The above analyses show that, while there is a strong geographic structure to the phylogeny, not all areas are monophyletic (see Figures 7.7 & 7.8). Although the Neotropical species of *Manilkara* all form a single, well-supported clade, the Asian, African and Madagascan taxa are paraphyletic. Bayes factor assessments of area monophyly allow for hypothesis testing to discern the difference in support for the constrained and unconstrained models. As a reduced dataset was used in this analysis, models and clock-likeness had to be re-tested. In this case, the smaller number of taxa had an effect on the chosen substitution model, which was slightly different from that chosen in the previous analyses.

A Bayes factor value of 9.76 indicated support for the model with three partitions: ITS1: GTR+I+G, 5.8S: GTR+I, ITS2: GTR+I+G and a birth-death clock (Table 7.9). Using this standard model, areas were constrained to be monophyletic and compared against the unconstrained model. In all cases area monophyly was strongly rejected, except for the Neotropical clade, the monophyly of which was narrowly rejected (Table 7.9). Because the Neotropical clade was resolved as monophyletic with strong support (pp 1) in this and previous analyses (Fig. 7.7), the rejection of monophyly in this analysis is a surprising outcome. This result may be due to the unstable placement of *Manilkara triflora*, which can vary from being sister to the entire Neotropical clade (node N), to being sister to the South American subclade (node O) (see Figs. 7.7 & 7.8).

Table 7.9 Bayes factor test for area constraints. The two competing best fit models are emboldened.

Model, clock and area monophyly imposed	**Ln HML**	**Bayes factor: 2ΔlnHML comparison of values shown between the chosen model and each alternative**
ITS1 - GTR+I+G 5.8S - GTR+I+G ITS2 - GTR+I+G Birth-death clock	-5864.1	15.82
ITS1 - GTR+I+G 5.8S - GTR+I+G ITS2 - GTR+I+G Uniform clock	-5875.37	37.82
ITS1 - GTR+I+G 5.8S - GTR+I+G ITS2 - GTR+I+G No clock	-5953.39	193.86
ITS1 - GTR+I+G **5.8S - GTR+G** **ITS2 - GTR+I+G** **Birth-death clock**	-5861.34	9.76
ITS1 - GTR+I+G **5.8S - GTR+I** **ITS2 - GTR+I+G** **Birth-death clock**	**-5856.46**	**Chosen model**
ITS1 - GTR+I+G 5.8S - GTR+I ITS2 - GTR+I+G Birth-death clock Area constraint - Madagascar	-5886.86	Paraphyletic in Figure 7.7 Monophyly rejected Bayes Factor 60.8
ITS1 - GTR+I+G 5.8S - GTR+I ITS2 - GTR+I+G Birth-death clock Area constraint - Africa	-5950.82	Polyphyletic in Figure 7.7 Monophyly rejected Bayes Factor 188.72
ITS1 - GTR+I+G 5.8S - GTR+I ITS2 - GTR+I+G Birth-death clock Area constraint - Neotropics	-5861.95	Monophyletic in Figure 7.7 Posterior probability value: 1 Monophyly rejected Bayes Factor 10.98
ITS1 - GTR+I+G 5.8S - GTR+I ITS2 - GTR+I+G Birth-death clock Area constraint - Asia	-5875.94	Paraphyletic in Figure 7.7 Monophyly rejected Bayes Factor 38.96

7.5 Discussion

7.5.1 Origin and means by which *Manilkara* achieved its pantropical distribution

In the following sections age ranges for taxa are reported as mean ages of the stem and crown nodes respectively.

7.5.1.1 Evidence for origin in Africa

Resolution along the backbone of the clade comprising the subtribe Manilkarinae is poor and the area of origin for *Manilkara* is, therefore, difficult to determine. However, *Manilkara* is nested within a grade of other representatives of the tribe Mimusopeae (*Labramia*, *Faucherea/Labourdonnaisia, Mimusops*, *Tieghemella*, *Autranella*, *Baillonella*, *Vitellaria* and *Vitellariopsis*), which is predominantly composed of African or Madagascan taxa and this suggests that the genus may have had its origin there. In the ancestral area reconstruction *Manilkara* is resolved as having a 98% likelihood of an African origin and a 2% likelihood of being ancestrally Madagascan, whereas the subtribe Manilkarinae is resolved as having a 96% likelihood of being ancestrally African and a 4% likelihood of being Madagascan. Likewise, the tribe Mimusopeae is reconstructed as having a 99% likelihood of being African and 1% Madagascan. As such, there is very strong support for an African ancestry for the genus *Manilkara*, the subtribe Manilkarinae and the tribe Mimusopeae. See Table 7.8.

7.5.1.2 Splits in the phylogeny consistent with Gondwanan vicariance

According to age estimates generated in this analysis intercontinental disjunctions in *Manilkara* are too young (33-5 Ma SapMim) to have been caused by Gondwanan break-up, which would have had to occur before 70 Ma (Fig. 5.3). This result is consistent with intercontinental disjunctions in numerous other tropical groups, whose ages are also too young to have been affected by vicariance of the former super-continent (a small sample includes: Myrtaceae & Vochisiaceae, Sytsma *et al* 2004; Cucurbitaceae, Schaefer *et al* 2008; Aglaieae, Meliaceae, Muellner *et al* 2008; Melastomataceae, Renner *et al* 2001, Renner 2004a,b; *Adansonia*, Bombacaceae, Baum *et al* 1998; *Acridocarpus*, Malpighiaceae, Davis *et al* 2002a,b; Burseraceae, Weeks *et al* 2005; *Poitea* & *Pictetia,* Leguminosae, Lavin *et al* 2001; *Sideroxylon*, Sapotaceae, Smedmark & Anderberg 2007). The only splits in the phylogeny (Fig. 7.7), which are of sufficient age to be Gondwanan are those between the outgroups: *Sarcosperma* (115 Ma SapMim, 113 Ma SapSidMan, 114 Ma SapSidMimMan), *Eberhardtia* (102-115 Ma SapMim, 101-113 Ma SapSidMan, 102-114 Ma SapSidMimMan), the tribe Sideroxyleae (80-102 Ma SapMim, 73-101 Ma SapSidMan, 77-102 Ma SapSidMimMan) and the *Xantolis/Englerophytum* clade (73-80 Ma SapMim, 66-73 Ma SapSid Man, 68-77 Ma SapSidMimMan).

7.5.1.3 Splits in the phylogeny consistent with the boreotropics hypothesis

Age estimates also indicate that *Manilkara s.s.* is too young (33-36 Ma SapMim, 30-32 Ma SapSidMan, 31-34 Ma SapSidMimMan) for its pantropical distribution to be the result of migration through the boreotropics, which would have had to occur between 65-45 Ma, with

intercontinental floristic exchange being most likely during the Paleocene-Eocene Thermal Maximum from 55-50 Ma (Zachos 2002) (Fig. 5.5). Although the North Atlantic land bridge connection between North America and Europe probably existed up until 30 Ma, at least as island stepping stones, the climate was probably no longer suitable for tropical vegetation at that time (Milne & Abbott 2002).

Additionally for further evidence of boreotropical migration through Laurasia one would expect South American lineages to be nested within Central American lineages and in Southeast Asia, lineages east of Wallace's Line to be nested within those west of Wallace's line. *Manilkara* does not exhibit this nested boreotropical migration syndrome, nor is it concordant with the age of the Paleocene-Eocene thermal maximum. The only suitable explanation for *Manilkara*'s disjunct pantropical distribution is, therefore, long-distance dispersal from Africa to Madagascar, Asia and the Neotropics.

The subtribe Manilkarinae is estimated to be 36-42 Ma (SapMim), 32-37 Ma (SapSidMan) and 35-40 Ma (SapSidMimMan), which is just on the cusp of being the appropriate age for boreotropical migration. (Also see Figure 7.4 of Manilkarinae age estimates for a full range of calibration scenarios and analyses.) These ages fall on the boundary between the Eocene and Oligocene. Crown node ages coincide with Oligocene cooling and the closing of the boreotropical route, whereas stem node age estimates are consistent with the hypothesis that subtribe Manilkarinae may have existed in the boreotropics and then migrated towards the equator as the climate in the northern hemisphere cooled at the culmination of the Eocene thermal maximum. This transition from the northern hemisphere to equatorial latitudes is also reflected in the putative Manilkarinae fossil record, where during the Oligocene, there is still a strong representation of fossils in the northern hemisphere, i.e. Isle of Wight, U.K. (Machin 1971), Vermont, U.S.A. (Traverse 1953 & 1955) and Czechoslovakia (Prakash, Brezinova & Awasthi 1974), but fossils also begin to appear in Africa, i.e. *Sapoteae* sp. leaves in Ethiopia (Jacobs *et al* 2005) (Fig. 6.8). Age estimates suggest that the subtribe Manilkarinae began to diversify into genera during the Oligocene. This diversification may have been spurred on by the onset of global aridification and cooling climates. Alternatively, Manilkarinae may have originated in Africa, as suggested by the ancestral area analysis. Because there are no living members of the Mimusopeae/Manilkarinae in the northern hemisphere to include in phylogenetic analysis, it is difficult to discern an unequivocal place of origin for the group.

The tribe Mimusopeae evolved approximately 46-57 Ma (SapMim) 40-49 Ma (SapSidMan) 46-56 Ma (SapSidMimMan) during the Eocene when global climates were warmer and wetter and a megathermal flora occupied the northern hemisphere. These age estimates also coincide with the first occurrence of putative Mimusopeae fossils recorded from North America and Europe (Fig. 6.7), e.g. *Tetracolporpollenites brevis* (Taylor 1989), *Manilkara* pollen (Frederiksen 1980a), and *Mimusops* leaf fossils (Berry 1915, 1916, 1924, 1930) in addition to the *Tetracolporpollenites* sp., pollen grain (Harley 1991), used in this study, which give further weight to the hypothesis that the tribe Mimusopeae was present in the boreotropics and may have originated there.

Previous studies (Smedmark & Anderberg 2007) implicate the break-up of the boreotropics in creating intercontinental disjunctions in the tribe Sideroxyleae and data from the present study are consistent with this hypothesis. Smedmark & Anderberg's (2007) estimate for the age of Sideroxyleae was 68 Ma and in this study ages are reconstructed as being 62-80 (SapMim), 54-73 (SapSidMan) and 56-68 (SapSidMimMan).

7.5.1.4 Splits in the phylogeny consistent with long distance dispersal

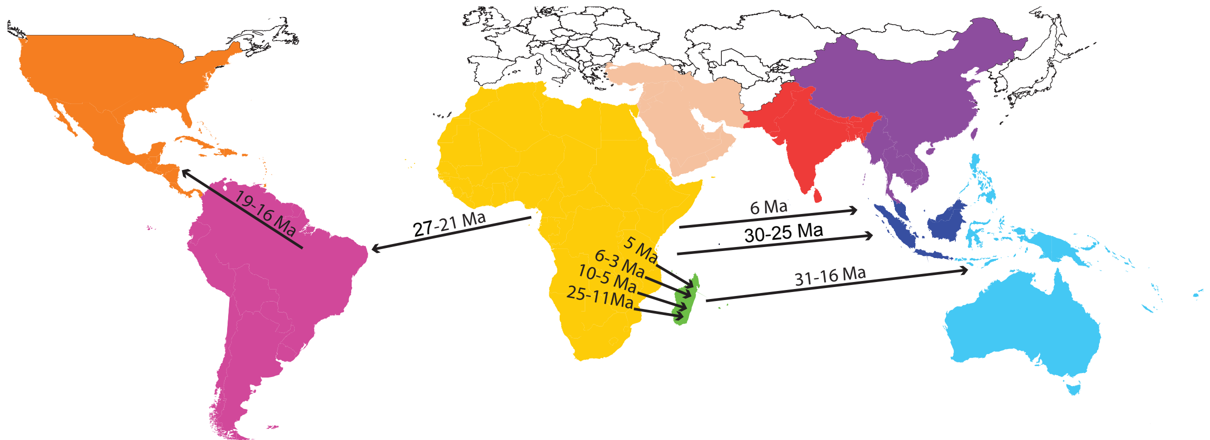


Figure 7.9 Map indicating age and direction of long distance dispersal events in the tribe Mimusopeae, including four in *Manilkara s.s.*, one in the *Manilkara fasciculata* lineage, and three in *Mimusops.* See Table 7.10 for details of individual dispersal events.

Table 7.10 Inferred long distance dispersal events in *Manilkara* and *Mimusops* as reconstructed in the chronograms in Figures 7.7 and 7.8. Also see Figure 7.9 for an illustration of dispersal patterns.

Taxon	Direction of dispersal	Nodes (stem & crown)	Age (stem & crown)	Geological epoch
Manilkara s.s.	Africa to Madagascar	T3	5 Ma	Miocene-Pliocene
	Africa to Madagascar	T4	10-5 Ma	Miocene
	Africa to Asia	Φ-U	30-25 Ma	Oligocene
	Africa to South America	Ψ1-N	27-21 Ma	Oligocene-Miocene
	South to Central America	N1-P	19-16 Ma	Miocene
Manilkara dissecta/udoido/ fasciculata	Africa/Madagascar to Asia	J-L	31-16 Ma	Oligocene-Miocene
Mimusops	Africa to Madagascar	F2-F3	6-3 Ma	Miocene-Pliocene
	Africa to Madagascar	F-F1	25-11 Ma	Oligocene-Miocene
	Africa to Asia	F2	6 Ma	Miocene

7.5.1.4.1 Origin of *Manilkara* in Africa followed by long distance dispersal

As previously discussed, an origin in Africa and subsequent intercontinental dispersal is the most likely scenario to explain *Manilkara*'s current pantropical distribution. The same is true for disjunctions in the *Faucherea/Labourdonnaisia/Manilkara* clade and the genus *Mimusops* (Fig. 7.9 and Table 7.10).

Manilkara s.s. (clade M in Fig. 7.7) is estimated to be of Oligocene age (33-36 Ma SapMim, 30-32 Ma SapSidMan, 31-34 Ma SapSidMimMan), with cladogenesis predominantly occurring during the Mid-Miocene. Following its origin in Africa *Manilkara* subsequently spread via long distance dispersal to Madagascar twice, Asia once and the Neotropics once during the Oligocene-Miocene (Fig. 7.9 and Table 7.10).

Both the *Faucherea/Labourdonnaisia/Manilkara* clade (J in Fig. 7.7) (31-36 Ma SapMim, 27-30 Ma SapSidMan, 29-34 Ma SapSidMimMan) and the genus *Mimusops* (clade F in Fig. 7.7) (25-31 Ma SapMim, 22-33 Ma SapSidMan, 24-41 SapSidMimMan) also exhibit a similar pattern, having originated in Africa during the Late Eocene-Oligocene and dispersed to Madagascar and Asia during the Oligocene-Miocene.

Manilkara species, along with other taxa in the tribe Mimusopeae, have fleshy, sweet fruit ranging in size from 1.5 – 10 cm which are consumed and dispersed by a wide variety of mammals and birds. In the Neotropics the fruit is commonly eaten by primates such as spider monkeys, howler monkeys, capuchins and tamarins (Estrada & Coates-Estrada 1984, Julliot 1996, Chauvet *et al* 2004, Chapman 1989, Oliviera & Ferrari 2000) as well as fruit bats (Uriarte *et al* 2005), tapirs (O'Farrill *et al* 2006) and otters in the Atlantic forests of Brazil (Quandros & Monteiro-Filho 2000). Frugivorous birds such as pigeons and doves have been recorded as eating *Manilkara* in Asia and Africa (Corlett 1998, Snow 1981). Baboons in West Africa (Kunz & Linsenmair 2008), fruit bats in Northern Queensland, Australia (Richards 1990) and wild pigs in Southeast Asia (Corlett 1998) have also been documented as dispersers of *Manilkara*.

With large fruit and seeds, which are too bulky to be wind dispersed, it is more likely that seeds of some *Manilkara* species could travel in the gut-contents of birds or that propagules could have rafted across water barriers in large mats of vegetation. Houle's (1998) study demonstrated that during the Miocene intercontinental rafting could have occurred in less than two weeks on the North and South Equatorial currents between Africa and the Neotropics and on the North and South Equatorial counter current between East Africa and South India/Southeast Asia. Graham (2006) states that biological debris to a maximum size of 300-500μm is occasionally transported by wind under current climatic conditions, but that larger material (700-1200μm) would be transported during periods of maximum warmth, such as during the Middle Miocene (23-12 Ma). Warmer climates also contribute to an increase in hurricanes and tornados, which are capable of transporting much larger objects and are known for dispersing great quantities of plant debris, including seeds and propagules (Nathan *et al* 2008). As unlikely as long distance dispersal of large seeds may sound, such an event would only need to be successful once in many millions of years to support this hypothesis, making the scenario much more plausible. Phylogenetic studies of many other groups have demonstrated their capability for intercontinental dispersal, including other taxa with similarly large, fleshy fruit: Annonaceae (Su & Saunders 2009); *Adansonia*, Bombacaceae (Baum *et al* 1998); *Atelia*, Leguminosae (Ireland *et al* 2010); *Andira*, Leguminosae (Skema 2003); *Commiphora*, Burseraceae (Weeks *et al* 2007); *Macherium*, Leguminosae (Lavin *et al* 2000); *Symphonia*, Clusiaceae (Dick *et al* 2003); Simaroubaceae (Clayton *et al* 2009); Cucurbitaceae (Schaefer *et al* 2008); Chrysophylleae, Sapotaceae (Bartish *et al* 2005), and also those which are smaller and potentially more volant: *Begonia*, Begoniaceae (Copestake *et al* 2009); *Cyrtandra*, Gesneriaceae (Cronk *et al* 2005); *Exacum*, Gentianaceae (Yuan *et al* 2005); *Acridocarpus*, Malpighiaceae (Davis *et al* 2002a, 2004b, 2004); Melastomataceae (Renner *et al* 2001); Rhamnaceae (Richardson *et al* 2004); *Gossypium*, Malvaceae (Cronn *et al* 2004); *Ceiba*, Bombacaceae (Dick *et al* 2007);

Maschalocephalus, Rapateaceae (Givnish 2004); *Pitcairnia*, Bromeliaceae (Givnish 2004) and *Renealmia,* Zingiberaceae (Sarkinen *et al* 2007).

7.5.1.4.2 Dispersal to Madagascar and the Mascarenes

Long-distance dispersal from Africa to Madagascar and surrounding islands has occurred on multiple occasions in the tribe Mimusopeae: twice in *Manilkara s.s.* (clades T3 and T4 in Fig. 7.7), either once or twice for the clade comprising *Labramia* (clade I in Fig 7.7) and *Faucherea/Labourdonnaisia* (clade K in Fig. 7.7) depending on support values, and twice in *Mimusops* (clade F1 and lineage F3 in Fig. 7.7) (also see Fig. 7.9 and Table 7.10).

Manilkara s.s. dispersed from Africa to Madagascar twice during the late Miocene (5-10 & 5-9 Ma SapMim, 4-9 & 5-8 Ma SapSidMan, 5-9 & 5-8 Ma SapSidMimMan), once from wet central Africa to wet eastern Madagascar (clade T3 in Fig. 7.7) and a second time from dry eastern Africa to dry western Madagascar (clade T4 in Fig. 7.7). Age estimates (31-36 Ma SapMim, 27-30 Ma SapSidMan, 29-34 Ma SapSidMimMan) for the *Faucherea/Labourdonnaisia/Manilkara* clade (J in Fig. 7.7) imply that the group originated during the late Eocene-Oligocene through long distance dispersal from Africa to Madagascar and the Mascarenes. In support of this theory, the ancestral area reconstruction analysis gives a 69% chance of the ancestral lineage being African, 30% Madagascan and 1% Asian. Within this clade, the Madagascan *Faucherea/Labourdonnaisia* subclade (clade K in Fig. 7.7) is Oligocene-Miocene in age (11-31 Ma SapMim, 10-27 Ma SapSidMan, 10-29 Ma SapSidMimMan). Although following a similar pattern of dispersal from Africa to Madagascar, diversification in *Labramia* (clade I in Fig. 7.7) may be the result of a separate event during the Oligocene-Miocene (7-36 Ma SapMim, 6-30 Ma SapSidMan, 6-35 Ma SapSidMimMan) due to poor support for the monophyly of clade Σ. Likewise, the genus *Mimusops* also dispersed from Africa to Madagascar and the Mascarenes twice (clades F1 & F3 in Fig. 7.7) during the Late Oligocene-Miocene: 11-25 & 3-6 Ma (SapMim), 10-22 & 3-5 Ma (SapSidMan), 11-24 & 3-6 Ma (SapSidMimMan).

In an investigation of the historical composition of the Madagascan biota, Yoder & Nowak (2006) determined that recent dispersal of taxa from Africa was by far the most important contributing source to the Madagascan flora. Likewise, Wikström *et al* (2010) found that the two Rubiaceae tribes Knoxieae and Vanguerieae have their origins in eastern tropical and southern Africa and dispersed to Madagascar numerous times. Bartish *et al* (2010) also demonstrated that multiple dispersals from Africa to Madagascar have taken place in the Sapotaceae tribe Chrysophylloideae during the Tertiary. *Manilkara*, *Labramia*, *Faucherea/Labourdonnaisia* and *Mimusops* data are highly consistent with the findings from these other studies and all support this hypothesis of recent dispersal from Africa to Madagascar.

7.5.1.4.3 Dispersal to South America – followed by dispersal to Central America and the Caribbean islands

The Neotropical *Manilkara* clade (N in Fig. 7.7) is also derived from an African ancestor, which dispersed to South America during the Oligocene-Miocene (21-27 Ma SapMim, 19-

26 Ma SapSidMan, 20-28 Ma SapSidMimMan). A pattern of Miocene African-Neotropical intercontinental dispersal is evident in numerous other taxa as well, namely *Commiphora* (Weeks *et al* 2007), *Symphonia* (Dick *et al* 2003), *Renealmia* (Sarkinen *et al* 2007), *Cassipourea*, *Rhizophora*, *Raphia* and *Elaeis* (Renner 2004c).

From South America and the Caribbean, further dispersal occurred to Central America twice 6-16 & 2-8 Ma (SapMim), 5-15 & 2-7 Ma (SapSidMan), 5-16 & 2-8 Ma (SapSidMimMan) and from South America throughout the Caribbean islands starting from 12-16 (SapMim), 10-15 Ma (SapSidMan), 11-16 Ma (SapSidMimMan). These age estimates place the New World spread of *Manilkara* prior to the closing of the Isthmus of Panama ~3.5 Ma (Pennington & Dick 2004, Coates & Obando 1996), suggesting further over-water dispersal from South into Central America and the Caribbean, a scenario which has also been demonstrated in numerous other taxa (Cody *et al* 2010).

7.5.1.4.4 Dispersal to Asia

Within the tribe Mimusopeae, long-distance dispersal from Africa to Asia has occurred on three separate occasions: once in *Manilkara s.s.* (clade U in Fig. 7.7), once in the small *Manilkara fasciculata* clade (L in Fig. 7.7), and once in *Mimusops* (lineage F2 in Fig. 7.7) (also see Fig. 7.9 and Table 7.10). All three of these events are too young to have been the result of rafting from Africa to Asia via the Indian subcontinent, which would have had to occur during the Eocene ~50-34 Ma, as outlined in Fig. 5.18.

The Asian lineage of *Manilkara s.s.* (clade U in Fig. 7.7) is estimated to have originated following a long-distance dispersal event from Africa to Asia, west of Wallace's Line during the Oligocene (25-30 Ma SapMim, 23-27 Ma SapSidMan, 24-28 Ma SapSidMimMan). Likewise, the small Asian *Manilkara* subclade (clade L in Fig. 7.7), which is situated in a clade with *Faucherea/Labourdonnaisia* is resolved as being Oligocene-Miocene in age (16-31 Ma SapMim, 14-27 Ma SapSidMan, 16-29 Ma SapSidMimMan). The progenitor of these three Asian taxa is likely to have been African (66%) or Madagascan (28%) and dispersed to Asia, east of Wallace's line. There has also been a single Late Miocene dispersal from Africa to Asia in the genus *Mimusops,* 6-9 Ma (SapMim), 5-8 Ma (SapSidMan), 6-8 Ma (SapSidMimMan). The only Asian species in this genus, *M. elengi*, is extremely widespread from India to New Caledonia and as only a single accession has been included in this analysis, it is not possible to discern where in Asia the species originated, although India is hypothesized based on morphology (leaf shape and indumentum) and proximity to Africa.

A similar scenario has been recorded in *Acridocarpus* (Malpighiaceae) (Davis *et al* 2002), with dispersal between Madagascar and New Caledonia occurring approximately 15-8 Ma. At this time, intercontinental rafting in the southern Indian Ocean would have taken less than two weeks (Houle 1998) and could have been aided by the onset of intense monsoon activity in Asia around 8 Ma (Zachos *et al* 2001, Jacques *et al* 2010). Likewise, Miocene dispersal between Africa/Madagascar and Asia west of Wallace's Line is a common pattern as exemplified by *Exacum* (Yuan *et al* 2005), *Nepenthes* (Meimberg *et al* 2001), *Begonia* (Thomas 2010) and Annonaceae (Richardson *et al* 2004).

In order to explain the large proportion of angiosperm taxa with disjunct distributions across the Indian Ocean, Schatz (1996) put forward a hypothesis of dispersal via "Lemurian Stepping Stones," which suggests that various, now submerged, Indian Ocean archipelagos may have narrowed the over-water distance between Africa/Madagascar and Asia at various times from the Eocene to the Oligocene, particularly ~ 30 Ma when global sea levels dropped. The possible emergence of these stepping stone ridges during this period is corroborated by tectonic studies by Ali & Aitchison (2008). More specifically, portions of the Chagos-Laccadive Plateau and the Mascarene Plateau, which extends 2,000 km between the Seychelles and Mauritius (Fisher *et al* 1967), and the Ninetyeast Ridge which extends 5,000 km along the 90° east meridian from 32° S to 9° N, were above water. The later was vegetated, as evidenced by fossil pollen assemblages (Kemp & Harris 1975, Renner 2010). *Sapotaceoidaepollenites rotundus* pollen (comparable to Mimusopeae or Isonandreae according to Harley 1991) has been recorded from Oligocene sediments on Ninetyeast Ridge (Kemp & Harris 1975), providing evidence that the archipelago was a viable stepping stone to dispersal for Sapotaceae during the period when the tribe Mimusopeae was diversifying into genera.

7.5.1.4.5 Further dispersal events suggested by hard incongruence between the nuclear and chloroplast datasets

In addition to the dispersal events reconstructed in the dated ITS phylogeny described above, further dispersals can be inferred based on instances of hard incongruence between the ITS and the chloroplast analyses, as outlined in Chapter IV sections 4.3.4.2 and 4.4.5. In these analyses, dispersal is inferred between East Africa and South Asia in the chloroplast clade W, which includes the African species *Manilkara concolor* and *M. mochisia* and the Asian species *M. littoralis* and *M. hexandra*. Additional dispersal events occurred between South America, Africa and Madagascar in the chloroplast clade V, which includes the Madagascan species *Manilkara suarezensis*, the Congolese species *M. yangambensis* and the Brazilian species *M. triflora*. Lastly, dispersal between Madagascar and the Mascarenes may have occurred in clade K, which is comprised of *Faucherea* and *Labourdonnaisia* species.

7.5.2 Regional diversification in *Manilkara*

7.5.2.1 Regional diversification patterns in the Neotropics

As discussed in the previous section, the Neotropical *Manilkara* lineage is derived from an African ancestor, which dispersed to South America during the Oligocene-Miocene (21-27 Ma SapMim, 19-26 Ma SapSidMan, 20-28 Ma SapSidMimMan). The Neotropical clade (N, Fig. 7.7) is divided into a Caribbean/Central American subclade (P, Fig. 7.7) (16-19 Ma SapMim, 15-19 Ma SapSidMan, 16-18 Ma SapSidMimMan) and a South American subclade (O, Fig. 7.7), not including *M. triflora* (14-19 Ma SapMim, 12-17 Ma SapSidMan, 13-18 Ma SapSidMimMan).

The Central American/Caribbean clade is further geographically divided with a small subclade (P1, Fig. 7.7) comprising the Central American species *M. staminodella* and *M.*

zapota and another subclade (P2, Fig. 7.7) comprising the Caribbean species *M. mayarensis, M. sideroxylon, M. pleena, M. jamiqui, M. gonavensis* and *M. valenzuelana*. The only exception to this geographical structure is the single Central American species, *M. chicle* (P3 Fig. 7.7), which is nested in the Caribbean clade, suggesting Pliocene dispersal (2 Ma) back to the continent (Table 7.11).

The South American clade (O, Fig. 7.7) is also further divided into two subclades, which correspond to regional ecology, with one clade comprised of Atlantic coastal forest species (clade O1: *M. salzmanii, M. elata, M. maxima, M, rufula, M. decrescens, M. bella, M. subsericea,* and *M. longifolia*) and the other of Amazonian species (clade O2: *M. huberi, M. bidentata, M. paraensis* and *M. inundata*). *Manilkara cavalcantei*, an Amazonian species in the Atlantic coastal forest clade, is the only inconsistency to this geographic pattern. The phylogenetic split between these two regions occurred approximately during the Mid-Miocene (14-19 Ma SapMim, 12-17 Ma SapSidMan, 10-13 Ma SapSidMimMan), when the Andes were being elevated and drainage systems in the Amazon basin began to shift eastwards (Table 7.11). This phylogenetic pattern is largely congruent with hypothetical lineage splits as outlined in Figure 5.11.

There is considerable debate about the age of the Andean orogeny. Its geology is complex, with different sections uplifting at different times from the Early Miocene ~25 Ma (Parra *et al* 2009) to the Pliocene. It has also recently come to light that a significant portion of the range has been elevated more recently, with the mountains having reached only approximately half their current elevation by 10 Ma (Graham 2009; Gregory-Wodzicki 2000). The uplift of the Andes has had far-reaching effects on the biogeography of the continent from coast to coast, creating the Amazon basin and significantly altering the drainage patterns in the central shield area of the continent (Hoorn *et al* 2010; Mora *et al* 2010), causing rivers to flow northeast rather than west and also dramatically influencing the climate (Sepulchre *et al* 2010).

Atlantic coastal species in clade O1 (Figure 7.7) and Amazonian species in clade O2 are geographically separated by the dry biomes of the Cerrado and the Caatinga, as well as the higher relief of the Brazilian shield. Simon *et al* (2009) and Fritsch *et al* (2004) found that the origin of dry-adapted Cerrado Leguminosae and Melastomataceae lineages span the Late Miocene to the Pliocene (from 9.8 to 0.4 Ma), broadly coinciding with the expansion of C4 grass-dominated savanna biomes. However, it is likely that a dry environment would have been present just prior to this time to allow for adaptation of these groups to the new biome. Such timing is exhibited by the Microlicieae (Melastomataceae), where the reported age is that of the conservative crown node at 9.8 Ma, but the stem node is estimated to be 17 Ma. This suggests that the Cerrado could have been in existence, at least in part, by the time the South American *Manilkara* subclades O1 & O2 diverged ~14 Ma, and may have been the driving factor behind the geographical split in the South American lineage of *Manilkara*. Likewise, a phylogenetic study of *Coursetia* (Leguminosae) (Lavin 2006) reveals that species which inhabit the dry forest of the Brazilian Caatinga are 5-10 Ma. The appearance of dry biomes in the Mid-Miocene, as suggested by these phylogenetic studies and others

(Pennington *et al* 2004), also roughly corresponds with the Miocene climatic optimum between 15-17 Ma (Zachos *et al* 2001).

7.5.2.2 Regional diversification patterns in Africa and Madagascar

African *Manilkara* species are resolved in two or three clades depending upon support values. All clades are Oligocene-Miocene in age. The main African/Madagascan clade (T in Fig. 7.7) is estimated to be (17-30 Ma SapMim, 15-27 Ma SapSidMan, 16-28 Ma SapSidMimMan). When the two smaller clades (R and S in Fig. 7.7) are resolved separately (as in the SapMim calibration analysis, pp values: 0.17 & 0.98), they are estimated to be 6-27 and 9-29 Ma respectively. When they are resolved as a single clade (Q) with poor support (as in the SapSidMan and SapSidMimMan – both pp 0.76), the age estimates are significantly older at 21-26 Ma (SapSidMan) and 22-28 Ma (SapSidMimMan) but still occurring in the Oligocene-Miocene.

As mentioned in Chapter V, section 5.7.2, Africa has been affected by widespread aridification during the Tertiary. The response to this changing climate could have been migration, adaptation or extinction. The pan-African forest, which stretched from nearly the west to the east coast during the late Eocene (Coetzee 1993) may have been segregated into western (Guineo-Congolian) and eastern (East African) sections only once as a result of Oligocene-early Miocene cooling and drying ~33-20 Ma, with the isolation and fragmentation of the two forest blocks generating high levels of endemism (Couvreur *et al* 2008; Burgess *et al* 1998, Lovett & Wasser 1993, Morley 2000). Alternatively, it is also possible that the expansion and contraction of rain forest occurred in cycles throughout the Oligo-Miocene, ~33-2 Ma, enabling intermittent gene flow and driving allopatric speciation (Fig. 5.14) (Couvreur *et al* 2008, Burgess *et al* 1998, Lovett & Wasser 1993, Coetzee 1993).

A study of rain forest-adapted Annonaceae genera (*Isolona* and *Monodora*) (Couvreur *et al* 2008) found that throughout climatic cycles, taxa remained restricted to remnant pockets of wet forest. They are, therefore, an example of a group that migrated or changed its distribution to track wetter climates. Another study of the rain forest genus *Acridocarpus* (Malpighiaceae) (Davis *et al* 2002) indicated an east African dry forest adapted lineage nested within a wet forest lineage. The dry adapted lineage was dated to periods of Oligo-Miocene aridity, and is, therefore, an example of a wet forest lineage, which has adapted to changing environmental conditions rather than becoming restricted to areas of favourable climate. These two differing scenarios of restriction versus adaptation were investigated in *Manilkara*.

The timing of diversification and evolution of dry-adapted species versus wet-restricted species in the three African *Manilkara* clades suggests a combination of both scenarios (as hypothesized in Fig. 5.14). The main split between the three African clades occurs between 33-27 Ma (node M in Fig. 7.7) during a period of dramatic continent-wide cooling, which fragmented the Eocene coast to coast forest block, potentially isolating the three lineages. A second wave of diversification within the main African/Madagascan clade (node T in Fig. 7.7) coincides with the Mid-Miocene climatic optimum 17-15 Ma, when global temperatures

warmed (Zachos 2001). During the same period the collision of the African and Eurasian plates closed the Tethys Sea, instigating further aridification. The resulting drier and warmer climates caused the spread of savannas and the retraction of rain forest, as evidenced by an increase in grass pollen during this period (Morley 2000, Jacobs 2004). Nonetheless, cladogenesis in the main African/Madagascan group gained pace from the Mid-Miocene onwards. In particular, a third wave of diversification from rain forest into drier shrubland environments in eastern and southern Africa occurred subsequent to the main uplift of the Tanganyikan plateau in the East African Rift System ~10 Ma, which had a significant impact on further regional aridification (Lovett & Wasser 1993, Sepulchre *et al* 2006) (Table 7.11).

The main African/Madagascan *Manilkara* clade (T in Fig. 7.7) is predominantly composed of Guineo-Congolian rain forest species. This is almost exclusively the case in subclade T1 (*M. letouzeyi, M. bequaertii, M. zenkeri, M. obovata multinervis*-type*, M. lososiana, M. pelligriniana* and *Letestua durissima*), aside from the Madagascan taxa (*M. suarezensis, M. bovinii, M. perrieri,* and *M. capuronii*), which are also rain forest species. However, within subclade T2, a transition from wet to dry environments can clearly be seen. Among the Guineo-Congolian species, *Manilkara fouilloyana, M. koechlinii, M. mabokeensis* and *M.* sp. 1 are all rain forest taxa, whereas *M. welwitschii* and *M. cunefolia*, are coastal woodland/thicket species which are also found on savanna margins and occasionally in gallery forest. *Manilkara dawei* is transitional between western and eastern forest blocks, being found in mixed lowland rain forest and rain forest edge in Uganda, Cameroon and northern Tanzania, whereas *M. sansibarensis*, *M. discolor* and *M. obovata* (*butugi*-type) are all eastern and southern African species which inhabit shrubland and semi-open woodland on dry sandy soils (Plana unpublished). The sole Madagascan taxon in this lineage (*M. sahafarensis*) is also a dry, deciduous forest species. These dry eastern-southern taxa all evolved between 10-5 Ma subsequent to the main uplift of the East African Rift System and it is likely that the ancestor of these species was associated with drier, more marginal forest/savanna environments due to climatic oscillations which caused fluctuations in the extent of forest versus savanna cover during the Oligo-Miocene. Likewise, the ancestor of the smaller African clade composed of *M. mochisia* (a) and *M. concolor* also diversified into these two dry-adapted eastern/southern species at the same time (~6 Ma). Rather than become extinct during Miocene aridification, some African *Manilkara* lineages adapted to a drying climate, while others were maintained in their ancestral rain forest habitat. These findings are concordant with both the rain forest refuge scenario found in *Isolona* and *Monodora* (Couvreur *et al* 2008) and the scenario of adaptation to a drier environment found in *Acridocarpus* (Davis *et al* 2002).

Madagascan species of *Manilkara* are not monophyletic. Instead, they are resolved in two separate subclades (T3 and T4 in Fig. 7.7) within the main African *Manilkara* clade (T in Fig. 7.7) and appear to follow a pattern of phylogenetic niche conservatism, where rain forest species in Madagascar (*M. suarezensis, M. bovinii, M. perrieri,* and *M. capuronii*) are derived from an African rain forest ancestor and the Madagascan dry forest species (*M. sahafarensis*) is derived from a dry-adapted African clade.

7.5.2.3 Regional diversification patterns in Southeast Asia

As previously discussed, the clade of Asian *Manilkara s.s.* (U in Fig. 7.7) originated following dispersal from Africa during the Oligocene (25-30 Ma SapMim, 23-27 Ma SapSidMan, 24-28 Ma SapSidMimMan). Within the main Asian clade of the chloroplast phylogeny presented in Chapter IV, the Indian species *Manilkara roxburghiana* is sister to the other species and the two Fijian species are the most derived, consistent with the hypothesis that the founding dispersal event was from Africa to India with subsequent spread eastward into Malesia (see Figs. 4.3 & 4.4). However, ancestral area reconstruction suggests that migration within Asia was from east to west (Sahul Shelf to Sunda Shelf), with a 96% likelihood (Figure 7.8, Table 7.8). However, *Manilkara* sp. 2 is a cultivated accession from Purwodadi Botanic Garden in east Java, which was originally collected in Sulawesi. Therefore, this individual, which is sister to the other species in this otherwise Sunda Shelf lineage, could be coded as being either Sunda or Sahul Shelf due to the composite nature of the island of Sulawesi. An alternative coding may suggest an origin of this clade on the Sunda Shelf. Additionally, there are five Asian taxa which were not included in the analysis, which may have affected the outcome: *M. samoensis* (Samoa), *M. napali* (New Guinea), *M. kanosiensis* (New Guinea), *M. celebica* (Sulawesi) and *M. roxburghiana* (India). What makes a scenario of east to west migration less probable is the fact that there was very little emergent land available for colonisation on the Sahul Shelf and in the Pacific until approximately 10 Ma, (Hall 2009) when the New Guinea central range began to be uplifted (Fig.5.17).

If the east to west migration scenario at ~25 Ma is taken to be true, then the complex fluctuating patterns of land and sea in this highly tectonically active region need to be taken into consideration. While the conservative biogeographical view is that significant land was not available for colonisation east of Wallace's Line until at least 10 Ma, the reality is that the extent of land versus sea coverage is still very imperfectly known. The first significant collision between the Sunda and Sahul shelves occurred during the Oligocene (30-25 Ma), uplifting the Bird's Head peninsula of New Guinea. This orogeny combined with a drop in sea levels caused central and western New Guinea to emerge above water (Pubellier *et al* 2003, van Ufford & Cloos 2005). In his reconstruction of the tectonics of Southeast Asia, Hall (1998, 2001) demonstrated that from 30 Ma to the present, the only land which was consistently above sea level was on the Sunda shelf and that New Guinea remained largely submerged until approx 5 Ma. However, some small island arcs were emergent as the tectonic landscape changed. At 25 Ma a portion of southern proto-Sulawesi as well as New Guinea's Bird's Head peninsula and the basin between southern New Guinea and northern Australia were above sea level. By 20 Ma plate boundaries and sea levels had shifted slightly, exposing more land in Sulawesi and the New Guinea-Australia basin, but submerging the Bird's Head peninsula. During the Miocene climatic optimum at around 15 Ma, sea levels rose again and the New Guinea-Australia basin became submerged, but as the Sunda and Sahul shelves converged further, mountain building was stimulated in northern New Guinea. Likewise, Sulawesi continued to increase its land area. Further orogeny and later Miocene aridification after 10 Ma exposed more land for colonisation (Hall 1998,

2001). Given this hypothesis, there could have possibly been intermittent and shifting land areas for *Manilkara* to colonise on the Sahul shelf prior to the late Miocene uplift of New Guinea.

The two youngest (10 Ma) Asian species (*M. vitiensis* & *M. smithiana*) are both Fijian. From the early Eocene to Late Miocene, Fiji formed part of the continuous Outer Melanesian Arc, which extended from Papua New Guinea to New Zealand (Rodd 2010). The Mid-Miocene Colo orogeny thrust up a large landmass, which is now part of Viti Levu. Consequently, the oldest land available for colonisation in Fiji is between 14-5 Ma (Johnson 1991, Heads 2006) and there is evidence for mangrove-fringed coasts during this time (Rodda 1976, 1994). Hence, the age of these two Fijian taxa coincides with the first occurrence of terrestrial land in the archipelago.

All of the above scenarios demonstrate that Wallace's Line has not been an absolute barrier to the dispersal of *Manilkara*. In fact, as discussed above, *Manilkara*'s current pantropical distribution is shown to be the result of dispersal over much greater distances than the Makassar Strait, which delineates Wallace's Line. This finding is in agreement with data from other angiosperm groups, which have, likewise, crossed Wallace's Line: *Pseuduvaria*, Annonaceae (Su & Saunders 2009); Aglaieae, Meliaceae (Muellner *et al* 2008); at least four separate lineages of *Begonia*, Begoniaceae (Thomas 2010); *Rhododendron*, Ericaceae (A. Twyford pers. comm. 2010, Brown 2006); *Cyrtandra*, Gesneriaceae (Cronk *et al* 2005) and *Etlingera*, Zingiberaceae (A. Poulsen pers. comm. 2010). Likewise, elsewhere in Sapotaceae four lineages of Isonandreae have migrated from west to east across Wallace's Line (Bakar 2009), whereas evidence from the tribe Chrysophylloideae suggests recent movement in the opposite direction, from Sahul to Sunda Shelf (J. Richardson pers. comm. 2010).

Table 7.11 Summary of regional diversification events within *Manilkara* in relation to climate and geology.

Region	Diversification event	Node	Crown node age	Epoch	Hypothesized driver of diversification
Neotropics	Amazon-Atlantic coastal forest split	O	14 Ma	Miocene	aridification of Cerrado & Caatinga, possibly triggered by Andean uplift and climate change
	Caribbean-Central American split	P	16 Ma	Miocene	dispersal event
	Central American lineage	P1	6 Ma	Miocene	dispersal event
	Caribbean lineage	P2	12 Ma	Miocene	dispersal event
	M. chicle Caribbean to Central America	P3	2 Ma	Miocene	dispersal event
Africa	1st split of rainforest lineage	M	33 Ma	Early Oligocene	cooling climates & aridification, separating eastern & western forest blocks
	2nd split of rainforest lineage	T	17-15 Ma	Miocene	cooling climates & aridification, separating eastern & western forest blocks
	3rd split of rainforest lineage and diversification to dry shrubland in southern and eastern Africa	Subclades in clade T2	from 10 Ma	Miocene	aridification and uplift of the Tanganyikan plateau
Asia	split between Indian/continental Asian and Malesian lineages	U	22 Ma	Miocene	dispersal event
	diversification from west to east across Wallace's Line and into the Pacific	Subclades of clade U2	13 & 16 Ma	Miocene	dispersal event

7.5.3 Comparison of mean estimated substitution rates across the entire phylogeny

It has been commonly observed that different lineages evolve at different rates and this shift is often linked to life history traits (Table 7.12). A discrepancy in substitution rate can occur even within a single genus, i.e. *Sidalcea* (Andreasen & Baldwin 2001). Estimation of substitution rates can also be sensitive to the methods used. For instance, Kay *et al* (2006) caution that "fossil data typically provide a minimum age of a particular lineage, and therefore may be more likely to *overestimate* a substitution rate, while biogeographic or climatic events provide a maximum age and may be more likely to *underestimate* a rate." In comparison to an overview of ITS substitution rates published in Kay *et al* (2006) (Table 7.12), this dataset (SapMim = 1.41 x 10^{-9}, SapSidMan = 1.57 x 10^{-9}, and SapSidMimMan = 1.48 x 10^{-9} substitutions per site per year) is most similar to *Aesculus* (1.72 x 10^{-9}) in the family Sapindaceae (Xiang *et al* 1998), a similarly long-lived tree, yet phylogenetically only very distantly related. (Sapotaceae is in the Asterid clade, whereas Sapindaceae is in the Rosid clade in the Angiosperm Phylogeny Group III classification (2009)). It has been suggested that generation time and habit (woody versus herbaceous) are potentially important factors governing the rate of evolution, but that phylogenetic relatedness is inconsequential (Page & Holmes 1998, Andreasen & Baldwin 2001). Kay *et al* (2006) found that rates from the family Asteraceae, alone, span almost the entire range studied. Findings from this study in comparison with that of Kay *et al* (2006) seem to support the argument that life history is a key factor in determining substitution rate.

Table 7.12 ITS substitution rates for other angiosperm genera, adapted from Kay *et al* 2006. Only rates of genera which were estimated using fossil calibrations are included so as to be comparable with the Sapotaceae dataset.

Genus	Family	Habit	Mutation Rate (substitutions per site per year)	Reference
Nothofagus	Nothofagaceae	woody	0.50	Manos 1997
Alnus	Betulaceae	woody	1.10	Savard *et al* 1993
Aesculus	Sapindaceae	woody	1.72	Xiang *et al* 1998
Gaertnera	Rubiaceae	woody	1.99	Malcomber 2002
Adansonia	Bombacaceae	woody	2.48	Baum *et a*l 1998
Lupinus	Fabaceae	herbaceous	3.46	Kass & Wink 1997
Astragalus	Fabaceae	herbaceous	3.50	Wojciechowski *et al* 1999
Gentianella	Gentianaceae	herbaceous	4.52	von Hagen & Kadereit, 2001
Soldanella	Primulaceae	herbaceous	8.34	Zhang *et al* 2001

7.6 Conclusions

In the previous chapter on historical biogeography (Chapter V, section 5.1) three main hypothetical scenarios were put forward to explain the current pantropical distribution of *Manilkara*: West Gondwanan vicariance (~110-70 Ma), boreotropical migration (~65-45 Ma) or recent long distance dispersal (since ~33 Ma). The data clearly support an Oligocene origin in Africa for the genus *Manilkara* with cladogenesis and intercontinental long-distance dispersal occurring during the Oligocene-Miocene. Ages estimated through all calibration scenarios and analytical methods tested reveal that the genus *Manilkara* is too young for its pantropical distribution to be the result of Gondwanan vicariance or boreotropical migration and recent long-distance dispersal is, instead, supported. It is, however, possible that the tribe

Mimusopeae and subtribe Manilkarinae may have originated in the northern hemisphere (or were at least present there) during the Eocene and subsequently migrated to more equatorial regions as climates cooled during the Oligocene. Boreotropical migration has been demonstrated for another pantropical tribe in the Sapotaceae, the Sideroxyleae (Smedmark & Anderberg 2007). Interestingly, age estimates for the genus *Mimusops* reveal the same trend as for *Manilkara*, with an origin in Africa during the late Oligocene and subsequent diversification during the Miocene with at least two long distance dispersals to Madagascar and a single dispersal to Asia.

Geo-climatic events, particularly during the Miocene, have been shown to correspond to periods of diversification in *Manilkara* on each continent. Specifically, the formation of the dry biomes of the Cerrado and Caatinga in South America correspond with the split between Atlantic coastal forest and Amazonian lineages of *Manilkara*. In Africa climate oscillations and the uplift of the Tanganyikan plateau are correlated with cladogenesis and the movement of lineages into drier habitats, whereas in Asia *Manilkara* dispersed from west to east across Wallace’s Line as the Asian and Australian plates converged and land was uplifted in New Guinea.

In Chapter VIII, to further investigate whether *Manilkara* can be used as a model for studying the historical assembly of tropical rain forest worldwide, biogeographic patterns reconstructed in this chapter are compared to those of other tropical rain forest taxa with intercontinental disjunct distributions.

7.7 Appendices

Appendix 7.1 Specimen data for taxa included in the different analyses. A ✓ indicates that sequence data for this accession was included in the analysis. A – indicates that sequence data for this accession was unavailable. Accessions marked AA are from Stockholm, G from Geneva and EDNA from Edinburgh.

Species	DNA accession number	Collector's number	Country of origin	MrBayes model selection	BEAST node ages	BEAST ancestral area	MrBayes area constraints
Argania spinosa	AA	Nordenstam 9325	Morocco	✓	✓	✓ Africa	-
Aulandra longifolia	AA	Christensen 1720	Sarawak, Malaysia	✓	✓	✓ Sunda shelf	-
Autranella congolensis	AA	Bokdam 4401	Congo	✓	✓	✓ Africa	✓
Baillonella toxisperma	EDNA09-01453	Bourobou *s.n.*	Gabon	✓	✓	✓ Africa	-
Burckella macropoda	AA	Chase 1359	Java, Indonesia	✓	✓	✓ Sunda shelf	-
Capurodendron androyense	AA	Humbert 28855	Madagascar	✓	✓	✓ Madagascar	-
Diploknema butyracea	AA	Polunin, Sykes & Williams 3975	Nepal	✓	✓	✓ South Asia	-
Diploknema oligomera	AA	Chase 1360	Java, Indonesia	✓	✓	✓ Sahul shelf	-
Eberhardtia aurata	AA	G. Hao 534 Cultivated	S. China Bot. Gard.	✓	✓	✓ East Asia	-
Eberhardtia tonkinensis	Gen Bank AF456258	Yang,S.-X. unpublished	Yunnan, China	✓	✓	✓ East Asia	✓
Englerophytum natalense	AA	C. K. 3483	Tanzania	✓	✓	✓ Africa	-
Faucherea manongarivensis	EDNA06-05896	L. Gautier *et al* 3910	Madagascar	✓	✓	✓ Madagascar	✓
Faucherea parvifolia	EDNA06-05889	L. Gautier 163	Madagascar	✓	✓	✓ Madagascar	✓
Faucherea thouvenotii	EDNA06-05897	L. Gautier 3938	Madagascar	✓	✓	✓ Madagascar	✓
Faucherea sp.	EDNA07-01933	A. Anderberg 233	Madagascar	✓	✓	✓ Madagascar	✓
Inhambanella henriquesii	AA	de Winter & Vahrmeijer 8536	South Africa	✓	✓	✓ Africa	✓
Isonandra compta	AA	Emanuelsson 3039	Sri Lanka	✓	✓	✓ South Asia	-
Isonandra perakensis	AA	Pennington & Wong 10227	Malaysia	✓	✓	✓ Sunda shelf	-
Isonandra sp	AA	Philcox, Weerasooriya & Weerasekera 10744	Sri Lanka	✓	✓	✓ South Asia	-
Labourdonnaisia callophylloides	EDNA08-02271	R. Capuron 28171SF	Reunion	✓	✓	✓ Madagascar	✓
Labourdonnaisia madagascariensis	EDNA07-02212	R. Capuron 27747SF	Madagascar	✓	✓	✓ Madagascar	✓
Labourdonnaisia revoluta	EDNA07-02271	Lorence 1602	Mauritius	✓	✓	✓ Madagascar	✓
Labramia anakaranaensis	EDNA06-05884	L. Gautier 4037	Madagascar	✓	✓	✓ Madagascar	✓
Labramia costata	EDNA07-02272	G. Schatz & A. Gentry 2094	Madagascar	✓	✓	✓ Madagascar	✓
Labramia louvelii	EDNA07-01927	A. Anderberg 245	Madagascar	✓	✓	✓ Madagascar	✓
Labramia mayottensis	AA	Labat *et al* 3309	Mayotte, Comoros	✓	✓	✓ Madagascar	✓
Letestua durissima	AA	Normand *s.n.*	Congo	✓	✓	✓ Africa	✓ Africa
Madhuca crassipes	AA	Jugah ak. Kudi 23757	Sarawak, Malaysia	✓	✓	✓ Sunda Shelf	-
Madhuca hainanensis	AA	G. Hao 530, Cultivated	South China Bot. Gard.	✓	✓	✓ East Asia	-
Madhuca leucodermis	AA	Takeuchi *et al.* 17858	New Guinea	✓	✓	✓ Sahul shelf	-
Madhuca longifolia	AA	G. Hao 531, Cultivated	South China Bot. Gard.	✓	✓	✓ South Asia	-
Madhuca microphylla	AA	Fagerlind 4790	Sri Lanka	✓	✓	✓ South Asia	-
Madhuca motleyana	AA	Pennington and Kochummen 10259	Malaysia	✓	✓	✓ Sunda shelf	-

Species	DNA accession number	Collector's number	Country of origin	MrBayes model selection	BEAST node ages	BEAST ancestral area	MrBayes area constraints
Madhuca palembanica	AA	Triono, Saman & Victobery 11	Indonesia	✓	✓	✓ Sunda shelf	-
Madhuca utilis	AA	Pennington & Asri 10209	Malaysia	✓	✓	✓ Sunda shelf	-
Manilkara bella	EDNA08-02267	Folli 501	Brazil	✓	✓	✓ South America	✓ Neotropics
Manilkara bequaertii	EDNA07-02081	F. Breteler 15348	Gabon	✓	✓	✓ Africa	✓ Africa
Manilkara bidentata (a)	M17116014	Jerome Chave *s.n.* (Bridge project)	French Guiana	✓	✓	✓ South America	✓ Neotropics
Manilkara bidentata (b)	EDNA06-05887	T. Pennington 1203	Peru	✓	✓	✓ South America	✓ Neotropics
Manilkara boivinii	EDNA06-05905	L. Gautier 3278	Madagascar	✓	✓	✓ Madagascar	✓ Madagascar
Manilkara capuronii	EDNA07-02079	R. Capuron 11.377SF	Madagascar	✓	✓	✓ Madagascar	✓ Madagascar
Manilkara cavalcantei	EDNA07-02205	Vicentini *et al* 527	Brazil	✓	✓	✓ South America	✓ Neotropics
Manilkara chicle	AA	Castillo *et al* 2083	Guatemala	✓	✓	✓ North America	✓ Neotropics
Manilkara concolor	AA	Swenson & Karis 635	South Africa	✓	✓	✓ Africa	✓ Africa
Manilkara cuneifolia	EDNA07-02264	G. McPherson 16792	Gabon	✓	✓	✓ Africa	✓ Africa
Manilkara dawei	EDNA07-01928	D.J. Harris 7707	Central African Republic	✓	✓	✓ Africa	✓ Africa
Manilkara decrescens	EDNA08-02268	J.D. & E.G. Chapman 6689	Brazil	✓	✓	✓ South America	✓ Neotropics
Manilkara discolor	EDNA06-05892	K. Vollesen 2460	Tanzania	✓	✓	✓ Africa	✓ Africa
Manilkara dissecta	EDNA06-05883	M. Gardner TNCA 4012	New Caledonia	✓	✓	✓ Sahul shelf	✓ Asia
Manilkara elata	EDNA08-02265	Jardin *et al* 2277	Brazil	✓	✓	✓ South America	✓ Neotropics
Manilkara fasciculata	EDNA08-02258	K. Armstrong 353	West Papua, Indonesia	✓	✓	✓ Sahul shelf	✓ Asia
Manilkara fouilloyana	EDNA07-02267	G. McPherson 16173	Gabon	✓	✓	✓ Africa	✓ Africa
Manilkara gonavensis	EDNA08-02264	Ekman 8741	Haiti	✓	✓	✓ North America	✓ Neotropics
Manilkara hexandra	EDNA07-02053	P.L. Comanor 868	Sri Lanka	✓	✓	✓ South Asia	✓ Asia
Manilkara hoshinoi (a)	EDNA08-02340	M. Hoshino 2138	Pulau	✓	✓	✓ Sahul shelf	✓ Asia
Manilkara hoshinoi (b)	EDNA07-02054	F.H. Damon 217	Papua New Guinea	✓	✓	✓ Sahul shelf	✓ Asia
Manilkara huberi	EDNA07-01926	O. Poncy 1828	French Guiana	✓	✓	✓ South America	✓ Neotropics
Manilkara inundata	EDNA07-02093	Sothers & Saraiva 22	Brazil	✓	✓	✓ South America	✓ Neotropics
Manilkara jamiqui	EDNA07-02201	Urquiola & Dressler 529	Cuba	✓	✓	✓ North America	✓ Neotropics
Manilkara kauki	EDNA08-02260	K. Armstrong 379	Bali, Indonesia	✓	✓	✓ Sunda shelf	✓ Asia
Manilkara koechlinii	EDNA06-05893	J. Casier 443	Dem. Rep. of Congo	✓	✓	✓ Africa	✓ Africa
Manilkara lacera	EDNA07-01095b	D.J. Harris 8200A	Gabon	✓	✓	✓ Africa	✓ Africa
Manilkara letouzeyi	EDNA08-02338	R. Letouzey 4444	Cameroon	✓	✓	✓ Africa	✓ Africa
Manilkara littoralis	EDNA07-02052	Maung Gale 14654	Myanmar	✓	✓	✓ East Asia	✓ Asia
Manilkara longifolia	EDNA07-02092	Thomas *et al* 8076	Brazil	✓	✓	✓ South America	✓ Neotropics
Manilkara lososiana	EDNA07-02088	D. Kenfack 625	Cameroon	✓	✓	✓ Africa	✓ Africa
Manilkara mabokeensis	EDNA07-01094	D.J. Harris 7164	Central African Republic	✓	✓	✓ Africa	✓ Africa
Manilkara maxima	EDNA07-02091	Sant'Ana *et al* 670	Brazil	✓	✓	✓ South America	✓ Neotropics
Manilkara mayarensis	AA	Ekman 9971	Cuba	✓	✓	✓ North America	✓ Neotropics
Manilkara mochisia (a)	EDNA06-05888	L. Gautier 4171	Zambia	✓	✓	✓ Africa	✓ Africa

Species	DNA accession number	Collector's number	Country of origin	MrBayes model selection	BEAST node ages	BEAST ancestral area	MrBayes area constraints
Manilkara obovata (butugi-type)	EDNA07-02262	Friis & Vollesen 740	Sudan	✓	✓	✓ Africa	✓ Africa
Manilkara obovata (multinervis-type)	EDNA07-02263	Schmidt *et al* 3274	Ghana	✓	✓	✓ Africa	✓ Africa
Manilkara obovata (obovata-type) (b)	EDNA08-02261	GAF Malanda 7	Republic of Congo	✓	✓	✓ Africa	✓ Africa
Manilkara paraensis	EDNA07-02206	Zarucchi *et al* 2526	Brazil	✓	✓	✓ South America	✓ Neotropics
Manilkara pellegriniana	EDNA07-02087	D.J. Harris & M. Fay 1843	Cameroon	✓	✓	✓ Africa	✓ Africa
Manilkara perrieri	EDNA07-02082	R. Capuron 28132-SF	Madagascar	✓	✓	✓ Madagascar	✓ Madagascar
Manilkara pleena	EDNA07-02209	A. Lioger & P. Lioger 33453	Puerto Rico	✓	✓	✓ North America	✓ Neotropics
Manilkara rufula	EDNA07-02208	G. Ignacio & A. Caurenio 37	Brazil	✓	✓	✓ South America	✓ Neotropics
Manilkara sahafarensis	EDNA07-02085	R. Capuron 20.965-SF	Madagascar	✓	✓	✓ Madagascar	✓ Madagascar
Manilkara salzmannii	EDNA07-02207	Jardim *et al* 2277	Brazil	✓	✓	✓ South America	✓ Neotropics
Manilkara sansibarensis	EDNA07-01083c	Abeid 272	Tanzania	✓	✓	✓ Africa	✓ Africa
Manilkara sideroxylon	EDNA07-02203	Ekman 16173	Cuba	✓	✓	✓ North America	✓ Neotropics
Manilkara smithiana	EDNA07-02057	A.C. Smith 1450	Fiji	✓	✓	✓ Sahul shelf	✓ Asia
Manilkara sp. 1	EDNA07-02260	P. Sita 4107	Congo	✓	✓	✓ Africa	-
Manilkara sp. 2	EDNA08-02256	P199701118/SO118, Cultivated	Purwodadi B.G., Indonesia	✓	✓	✓ Sahul shelf	✓ Asia
Manilkara staminodella	EDNA07-02204	Anderberg *et al* 50	Costa Rica	✓	✓	✓ North America	✓ Neotropics
Manilkara suarezensis	EDNA07-02259	Randriamampionona 248	Madagascar	✓	✓	✓ Madagascar	✓ Madagascar
Manilkara subsericea	EDNA07-02202	Hatschbach & Souza 51302	Brazil	✓	✓	✓ South America	✓ Neotropics
Manilkara triflora	EDNA08-02343	Fonseca *et al* 2887	Brazil	✓	✓	✓ South America	✓ Neotropics
Manilkara udoido	EDNA07-02058	S. Slappy LR26622	Palau	✓	✓	✓ Sahul shelf	✓ Asia
Manilkara valenzuelana	EDNA07-02211	A. Lioger & P. Lioger 22980	Dominican Republic	✓	✓	✓ North America	✓ Neotropics
Manilkara vitiensis	EDNA08-02345	Smith 1461	Fiji	✓	✓	✓ Sahul shelf	✓ Asia
Manilkara welwitschii	EDNA06-05891	J.J.F.E. de Wilde & R.W. de Wilde-Bakhuizen 11385	Gabon	✓	✓	✓ Africa	✓ Africa
Manilkara zapota	EDNA06-05886	J. Clayton 12	Trinidad	✓	✓	✓ North America	✓ Neotropics
Manilkara zenkeri	EDNA07-02084	Doumenge 526	Cameroon	✓	✓	✓ Africa	✓ Africa
Mimusops caffra	AA	Swenson & Karis 636	South Africa	✓	✓	✓ Africa	-
Mimusops comorensis	AA	Pignal & Ginguette 1065	Comores Islands	✓	✓	✓ Madagascar	-
Mimusops coriacea	Geneva	Bernadi 11891	Madagascar	✓	✓	✓ Madagascar	-
Mimusops elengi	AA	Chantaranothai 2305	Thailand	✓	✓	✓ East Asia	-
Mimusops kummel	Geneva	Kayambo 4996	Tanzania	✓	✓	✓ Africa	-
Mimusops membranacea	Geneva	Randrianaivo 126	Madagascar	✓	✓	✓ Madagascar	-
Mimusops obovata	AA	Swenson & Karis 633	South Africa	✓	✓	✓ Africa	-
Mimusops perrieri	Geneva	S.F. 18297	Madagascar	✓	✓	✓ Madagascar	-
Mimusops sp. (voalala complex)	Geneva	Randrianaivo 583	Madagascar	✓	✓	✓ Madagascar	-
Mimusops zeyheri	AA	Dahlstrand 6386	South Africa	✓	✓	✓ Africa	✓
Northia seychellana	AA	L. Chong-Seng *s. n.*	Seychelles	✓	✓	✓ Seychelles	-
Palaquium amboinense	AA	Iuijesundara *s.n.*	Sri Lanka	✓	✓	✓ South Asia	-
Palaquium formosanum	AA	Chung & Anderberg 1421	Taiwan	✓	✓	✓ East Asia	-
Palaquium microphyllum	AA	Pennington, Kochummen & Wong 10222	Malaysia	✓	✓	✓ Sunda shelf	-
Palaquium ridleyi	AA	P. Wilkie 858	Malaysia (Borneo)	✓	✓	✓ Sunda shelf	-

Species	DNA accession number	Collector's number	Country of origin	MrBayes model selection	BEAST node ages	BEAST ancestral area	MrBayes area constraints
Palaquium stenophyllum	EDNA07-01936	Ferry Slik 9592	Borneo, Indonesia	✓	✓	✓ Sunda shelf	-
Payena acuminata	AA	Chase 1368	Indonesia	✓	✓	✓ Sunda shelf	-
Payena lucida	AA	Ambri *et al* AA1604	Borneo	✓	✓	✓ Sunda shelf	-
Sarcosperma laurinum	AA	Saunders 2000	Hong Kong	✓	✓	✓ East Asia	✓
Sideroxylon americanum	AA	Gillis 11576	Bahamas	✓	✓	✓ North America	✓
Sideroxylon angustum	AA	Ekman 4034	Cuba	✓	✓	✓ North America	-
Sideroxylon beguei	AA	McPherson *et al* 14831	Madagascar	✓	✓	✓ Madagascar	-
Sideroxylon betsimisarakum	AA	Schonenberger *et al* A-102	Madagascar	✓	✓	✓ Madagascar	-
Sideroxylon borbonicum	AA	Bosser 21325	Reunion	✓	✓	✓ Madagascar	-
Sideroxylon capiri	AA	Garcýa 1848	Mexico	✓	✓	✓ North America	-
Sideroxylon capuronii	AA	Capuron 20151-SF	Madagascar	✓	✓	✓ Madagascar	-
Sideroxylon celastrinum	AA	Correll 50467	Bahamas	✓	✓	✓ North America	-
Sideroxylon confertum	AA	Ekman 17405	Cuba	✓	✓	✓ North America	-
Sideroxylon contrerasii	AA	Lundell 20793	Guatemala	✓	✓	✓ North America	-
Sideroxylon cubense	AA	Beurton & Mory 927	Dominican Republic	✓	✓	✓ North America	-
Sideroxylon floribundum	AA	Lundell 20263	Guatemala	✓	✓	✓ North America	-
Sideroxylon foetidissimum	AA	Lundin 638	Cuba	✓	✓	✓ North America	-
Sideroxylon galeatum	AA	Friedman 3288	Rodrigues	✓	✓	✓ Madgascar	-
Sideroxylon gerrardianum	AA	Capuron 28826-SF	Madagascar	✓	✓	✓ Madagascar	-
Sideroxylon grandiflorum	AA	Friedman *et al* 2653,	Mauritius	✓	✓	✓ Madagascar	-
Sideroxylon horridum	AA	Gutierrez & Nilsson 5	Cuba	✓	✓	✓ North America	-
Sideroxylon ibarrae	AA	Lundell 19752	Guatemala	✓	✓	✓ North America	-
Sideroxylon inerme	AA	Nielsen *s.n.,* Cultivated	Denmark	✓	✓	✓ Africa	-
Sideroxylon lanuginosum	AA	Correll & Ogden 28456	Texas	✓	✓	✓ North America	-
Sideroxylon lanuginosum	Gen Bank F174617	R.B. Jackson *et al* 1999	Texas, U.S.A.	✓	✓	✓ North America	-
Sideroxylon leucophyllum	AA	Carter 5706	Mexico	✓	✓	✓ North America	-
Sideroxylon lycioides	AA	Radford *et al* 11453	South Carolina, USA	✓	✓	✓ North America	-
Sideroxylon majus	AA	Capuron 28185SF	Reunion	✓	✓	✓ Madagascar	-
Sideroxylon marginatum	AA	Leyens CV-96–672	Cape Verde	✓	✓	✓ Africa	-
Sideroxylon marmulano	AA	Swenson & Fernandez 581	Canary Islands	✓	✓	✓ Africa	-
Sideroxylon mascatense	AA	Thulin, Beier & Hussein 9774	Yemen	✓	✓	✓ Middle East	-
Sideroxylon obovatum	AA	Garcýa *et al* 5586	Dominican Republic	✓	✓	✓ North America	-
Sideroxylon obtusifolium	AA	Alvarez *et al* 28772	Mexico	✓	✓	✓ North America	-
Sideroxylon occidentale	AA	Carter & Sharsmith 4268	Mexico	✓	✓	✓ North America	-
Sideroxylon oxyacanthum	AA	Wood Y/75/388	Yemen	✓	✓	✓ Middle East	-
Sideroxylon palmeri	AA	Palmer 1513	Mexico	✓	✓	✓ North America	-
Sideroxylon persimile	AA	Veliz 99.7038	Guatemala	✓	✓	✓ North America	-
Sideroxylon picardae	AA	Ekman 15576	Hispaniola	✓	✓	✓ North America	-
Sideroxylon portoricense	AA	Mathew 1	Jamaica	✓	✓	✓ North America	-
Sideroxylon puberulum	AA	Coode 4121	Mauritius	✓	✓	✓ Madagascar	-
Sideroxylon reclinatum	AA	Traverse 592	USA	✓	✓	✓ North America	-
Sideroxylon repens	AA	Greuter & Rankin 24954	Dominican Republic	✓	✓	✓ North America	-
Sideroxylon rotundifolium	AA	Webster *et al* 8458	Jamaica	✓	✓	✓ North America	-

Species	DNA accession number	Collector's number	Country of origin	MrBayes model selection	BEAST node ages	BEAST ancestral area	MrBayes area constraints
Sideroxylon salicifolium	AA	Gutierrez & Nilsson 14	Cuba	✓	✓	✓ North America	-
Sideroxylon saxorum	AA	Jongkind 3500	Madagascar	✓	✓	✓ Madagascar	-
Sideroxylon sessiliflorum	AA	Lorence & Edgerley 2706	Mauritius	✓	✓	✓ Madagascar	-
Sideroxylon stenospermum	AA	Stevens 22935	Nicaragua	✓	✓	✓ North America	-
Sideroxylon stevensonii	AA	Lundell & Contreras 19057	Guatemala	✓	✓	✓ North America	-
Sideroxylon tambolokoko	AA	Capuron 22388-SF	Madagascar	✓	✓	✓ Madagascar	-
Sideroxylon tenax	AA	Radford & Leonard 11519	South Carolina	✓	✓	✓ North America	-
Sideroxylon tepicense	AA	Gentry 2931	Mexico	✓	✓	✓ North America	-
Sideroxylon wightianum	AA	G. Hao 532, Cultivated	S. China Bot. Gard	✓	✓	✓ East Asia	-
Tieghemella heckelii	AA	Jongkind 3936	Ghana	✓	✓	✓ Africa	-
Vitellaria paradoxa	AA	Neumann 1512	Benin	✓	✓	✓ Africa	-
Vitellariopsis cuneata	AA	Thomas 3662	Tanzania	✓	✓	✓ Africa	-
Vitellariopsis dispar	AA	Pentz 2	South Africa	✓	✓	✓ Africa	-
Vitellariopsis kirkii	AA	Robertson 4085	Kenya	✓	✓	✓ Africa	✓
Vitellariopsis marginata	AA	Chase 1122	South Africa	✓	✓	✓ Africa	-
Xantolis cambodiana	AA	Chantaranothai 2307	Thailand	✓	✓	✓ East Asia	-
Xantolis siamensis	AA	Smitairi 1	Thailand	✓	✓	✓ East Asia	-

Appendix 7.2 Mean node age with 95% HPD age range reconstructed for profiled nodes in BEAST using all fossil calibration scenarios.

Fossil calibration scenario/ Profiled node	**Sapotaceae *s.s.* (node A)**	**Sideroxyleae (node B)**	**Mimusopeae (node C)**	**Manilkarinae (node H)**	***Manilkara* (node M)**	***Mimusops* (node F)**	***Tieghemella* (node G)**	***Faucherea/ Labourdonnaisia/ Manilkara* (node J)**
Sapotaceae *s.s.*	100-113-136	46-59-71	32-40-49	32-40-49	24-30-36	15-22-30	18-29-39	20-27-35
Sideroxyleae	74-99-129	49-52-59	27-35-45	21-29-36	19-26-33	13-19-27	16-25-35	17-24-32
Mimusopeae	96-133-173	52-69-87	45-48-54	31-39-46	28-35-42	18-26-35	22-35-45	24-33-41
Manilkara	81-116-153	44-60-78	33-41-52	29-33-41	28-31-36	15-23-31	19-30-42	21-28-37
Tieghemella	72-116-164	39-60-85	31-42-57	23-34-46	22-31-43	14-23-33	28-31-36	18-29-41
SapSid	100-113-131	49-53-62	30-39-46	25-31-38	23-28-35	14-21-28	18-28-37	20-26-34
SapMim	101-115-140	51-62-74	45-46-49	31-36-42	28-33-38	18-25-32	21-33-42	24-31-38
SapMan	100-114-135	47-59-71	34-41-48	29-33-38	28-30-34	15-22-29	20-29-39	21-28-34
SapTie	100-114-135	48-60-72	34-41-49	27-33-40	25-30-36	16-23-30	28-30-35	21-28-35
SapSidMim	100-115-137	49-56-65	45-46-49	30-36-41	26-32-38	17-24-32	21-32-41	23-30-37
SapSidMan	100-113-133	49-54-61	33-40-46	29-32-37	28-30-33	15-22-28	18-28-37	21-27-33
SapSidTie	100-113-133	49-54-62	32-39-42	26-32-38	23-29-35	15-22-29	28-30-34	20-27-34
SapMimMan	100-115-138	50-62-73	45-46-49	30-35-40	28-32-36	17-24-32	22-32-42	23-30-37
SapMimTie	100-115-138	51-62-74	45-46-49	31-36-42	28-33-38	17-25-32	28-31-37	23-31-38
SapSidMimMan	100-114-136	49-56-65	45-46-49	30-35-39	28-31-35	17-24-32	21-32-41	23-29-36
SapSidMimTie	100-114-136	49-56-65	45-46-48	30-36-41	27-32-37	17-24-32	28-31-37	23-30-37

Fossil calibration scenario/ Profiled node	***Labramia* (node I)**	***Faucherea/ Labourdonnaisia* (node K)**	**small Asian *Manilkara* (node L)**	**Neotropical *Manilkara* (node N)**	**S. American *Manilkara* (node O)**	**C. American & Caribbean *Manilkara* (node P)**	**large African *Manilkara* (node T)**	**large Asian *Manilkara* (node U)**	**small African *Manilkara* (node Q or R & S)**
Sapotaceae *s.s.*	2-6-10	5-10-15	8-15-22	13-19-24	8-13-17	10-15-20	11-16-20	17-23-29	13-21-29
Sideroxyleae	2-5-9	5-9-13	7-13-19	12-17-22	7-11-15	9-13-18	9-14-18	15-20-27	11-19-26
Mimusops	3-7-13	7-12-17	9-17-25	17-22-29	10-15-21	12-18-24	13-19-24	20-27-34	13-19-24
Manilkara	2-6-11	6-10-15	8-15-22	14-19-26	9-13-18	10-15-21	11-16-22	17-24-30	14-22-30
Tieghemella	2-6-11	5-10-16	7-15-24	12-20-28	8-13-20	10-16-23	10-16-24	16-24-34	12-22-33
SapSid	2-6-10	5-9-14	8-14-21	13-18-23	8-12-16	10-14-19	10-15-19	16-22-27	13-20-28
SapMim	3-7-12	6-11-16	9-16-23	16-21-26	10-14-19	12-16-21	12-17-22	20-25-31	n/a 2-6-13 & 4-9-16
SapMan	2-6-11	6-10-15	9-15-22	14-19-24	9-13-17	11-15-20	12-16-20	18-23-28	14-21-28
SapTie	2-6-11	5-10-15	8-15-22	14-19-25	9-13-18	11-15-20	11-16-21	18-23-30	14-22-29
SapSidMim	2-6-11	6-11-16	9-16-23	15-20-25	9-14-18	11-16-21	12-17-22	19-24-30	15-23-30
SapSidMan	2-6-10	6-10-14	8-14-21	14-19-23	9-12-16	11-15-19	11-15-19	18-23-27	14-21-28
SapSidTie	2-6-10	5-10-14	8-14-21	11-15-20	14-18-23	10-14-19	11-15-20	17-22-28	13-21-28
SapMimMan	2-7-11	6-11-16	9-16-23	16-20-25	10-14-18	12-16-21	12-17-22	19-24-30	15-23-30
SapMimTie	3-7-12	6-11-16	9-16-24	16-21-26	10-14-19	11-16-21	12-17-22	20-25-31	15-23-31
SapSidMimMan	2-6-11	6-10-15	9-16-23	15-20-24	9-13-18	12-16-20	12-16-21	19-24-29	15-22-29
SapSidMimTie	2-6-11	6-11-16	9-16-24	15-20-25	9-14-18	11-16-20	12-17-21	19-24-30	n/a 1-6-12 & 4-9-15

Appendix 7.3 Mean ages with confidence intervals reconstructed for profiled nodes in r8s using penalized likelihood and all fossil calibration scenarios. Fixed ages are emboldened.

Fossil calibration scenario/ Profiled node	**Sapotaceae *s.s.* (node A) (in 101 trees)**	**Sideroxyleae (node B) (in 101 trees)**	**Mimusopeae (node C) (in 101 trees)**	**Manilkarinae (node H) (in 94 trees)**	***Manilkara* (node M) (in 100 trees)**	***Mimusops* (node F) (in 101 trees)**	***Tieghemella* (node G) (in 101 trees)**	***Faucherea/Labramia/ Manilkara* (node J) (in 101 trees)**
Sapotaceae *s.s.*	**102-102-102**	53-69-84	29-38-48	24-31-36	14-28-35	10-21-27	12-29-39	15-27-34
Sideroxyleae	58-72-92	**49-49-49**	21-27-45	17-22-28	15-20-25	8-15-21	13-21-43	13-19-26
Mimusopeae	58-101-102	41-71-86	**45-45-45**	25-36-41	23-33-39	13-24-32	20-34-46	16-32-39
Manilkara	80-97-102	53-66-84	31-36-43	28-30-34	**28-28-28**	11-20-28	20-29-36	19-26-34
Tieghemella	68-93-102	45-63-85	26-35-45	19-29-37	18-27-35	10-19-27	**28-28-28**	15-25-34
SapSid	**102-102-102**	**49-49-49**	23-32-46	19-26-41	16-24-37	10-18-26	17-25-34	15-23-41
SapMim	**102-102-102**	56-71-86	**45-45-45**	27-36-41	25-33-39	13-24-32	12-34-46	15-32-39
SapMan	**102-102-102**	54-69-84	31-37-45	28-30-34	**28-28-28**	12-20-27	21-29-37	19-27-32
SapTie	**102-102-102**	55-68-85	26-36-57	22-29-48	21-27-38	10-20-42	**28-28-28**	18-26-34
SapSidMim	**102-102-102**	**49-49-49**	**45-45-45**	25-36-41	24-33-39	13-24-31	24-34-45	22-32-39
SapSidMan	**102-102-102**	**49-49-49**	23-35-44	20-29-33	**28-28-28**	11-19-26	19-27-35	16-26-31
SapSidTie	**102-102-102**	**49-49-49**	26-34-45	22-28-34	21-26-32	10-19-26	**28-28-28**	17-25-31
SapMimMan	**102-102-102**	56-71-86	**45-45-45**	29-32-40	**28-28-28**	12-23-37	23-33-46	20-28-36
SapMimTie	**102-102-102**	56-71-86	**45-45-45**	23-34-40	22-31-38	11-21-31	**28-28-28**	21-30-39
SapSidMimMan	**102-102-102**	**49-49-49**	**45-45-45**	29-32-39	**28-28-28**	12-23-36	24-33-45	20-28-36
SapSidMimTie	**102-102-102**	**49-49-49**	**45-45-45**	23-33-40	22-31-38	11-21-31	**28-28-28**	14-29-39

Fossil calibration scenario/ Profiled node	***Labramia* (node I) (in 101 trees)**	***Fauchera/ Labourdonnaisia* (node K) (in 101 trees)**	**small Asian *Manilkara* (node L) (in 101 trees)**	**Neotropical *Manilkara* (node N) (in 97 trees)**	**S. American *Manilkara* (node O) (in 101 trees)**	**C. American & Caribbean *Manilkara* (node P) (in 101 trees)**	**largeAfrican *Manilkara* (node T) (in 101 trees)**	**large Asian *Manilkara* (node U) (in 101 trees)**	**small African *Manilkara* (node Q) (in 101 trees)**
Sapotaceae *s.s.*	2-5-9	5-9-14	9-14-21	7-17-23	6-12-16	6-14-18	10-15-20	8-22-30	9-22-32
Sideroxyleae	1-4-6	3-7-13	6-10-17	9-12-17	5-8-13	7-10-14	8-10-14	12-16-22	10-15-23
Mimusops	2-6-12	6-11-17	11-16-24	12-20-26	8-14-18	10-17-23	12-17-23	12-26-32	13-25-36
Manilkara	2-5-10	5-9-14	9-14-22	13-17-22	8-11-15	9-14-18	9-14-19	17-22-25	13-21-28
Tieghemella	1-5-16	5-9-14	8-13-20	10-16-22	6-11-16	7-13-18	8-14-20	11-21-29	12-20-29
SapSid	2-4-15	4-8-28	8-12-23	10-15-23	7-10-15	8-12-18	9-12-24	14-19-30	12-18-34
SapMim	2-6-15	6-11-17	11-16-24	8-20-26	8-14-18	8-17-23	12-17-23	15-26-32	17-25-35
SapMan	2-5-8	5-9-14	9-14-20	13-17-22	8-11-15	10-14-18	9-14-18	17-22-25	14-21-27
SapTie	1-5-14	5-9-15	9-13-20	12-17-30	7-11-20	9-14-20	9-14-20	16-21-28	13-21-29
SapSidMim	2-6-10	6-11-17	10-16-24	15-20-26	9-14-18	11-17-22	11-17-22	20-26-31	16-25-35
SapSidMan	2-5-8	5-9-13	9-13-19	12-16-22	7-11-15	10-14-18	9-14-18	15-21-25	14-21-27
SapSidTie	1-5-9	5-8-18	8-13-20	12-16-22	7-11-15	8-13-17	9-13-19	15-20-26	12-20-28
SapMimMan	2-6-20	6-10-15	9-15-23	14-17-24	8-12-16	11-14-19	10-15-20	18-22-29	11-21-35
SapMimTie	2-6-21	6-10-17	9-15-24	12-19-26	9-13-20	6-16-21	11-16-26	18-24-30	13-24-34
SapSidMimMan	2-5-12	6-10-15	9-14-21	14-17-23	8-12-16	10-14-19	10-14-23	14-22-28	13-21-34
SapSidMimTie	2-6-16	6-10-17	9-15-24	13-18-26	9-13-21	10-15-21	11-16-22	16-24-30	13-23-34

Appendix 7.4 Mean ages with confidence intervals reconstructed for profiled nodes in r8s using nonparametric rate smoothing and all fossil calibration scenarios. Fixed ages are emboldened.

Fossil calibration scenario/ Profiled node	**Sapotaceae *s.s.* (node A) (in 101 trees)**	**Sideroxyleae (node B) (in 101 trees)**	**Mimusopeae (node C) (in 101 trees)**	**Manilkarinae (node H) (in 94 trees)**	***Manilkara* s.s. (node M) (in 100 trees)**	***Mimusops* (node F) (in 101 trees)**	***Tieghemella* (node G) (in 101 trees)**	***Faucherea/Labramia/ Manilkara* (node J) (in 101 trees)**
Sapotaceae *s.s.*	**102-102-102**	56-76-87	39-48-59	34-40-47	30-37-44	16-27-37	25-38-49	26-36-45
Sideroxyleae	58-84-101	**49-49-49**	24-36-56	20-31-43	18-29-41	9-20-34	13-29-46	15-27-42
Mimusopeae	101-101-101	59-75-87	**45-45-45**	32-37-42	29-35-40	13-25-34	25-35-46	25-33-40
Manilkara	95-101-101	54-73-86	32-38-52	28-30-41	**28-28-28**	11-20-31	20-29-46	19-27-32
Tieghemella	86-101-101	57-74-87	26-40-58	22-33-44	21-23-42	10-20-32	**28-28-28**	19-29-43
SapSid	**102-102-102**	**49-49-49**	23-40-56	19-34-43	17-31-41	10-22-34	13-32-46	14-30-42
SapMim	**102-102-102**	59-75-87	**45-45-45**	32-37-42	29-35-40	13-25-34	25-35-46	25-33-40
SapMan	**102-102-102**	54-73-86	32-38-51	28-30-40	**28-28-28**	11-20-31	20-29-46	19-27-32
SapTie	**102-102-102**	57-74-87	26-40-58	21-33-44	20-31-42	10-20-32	**28-28-28**	18-29-43
SapSidMim	**102-102-102**	**49-49-49**	**45-45-45**	32-38-44	30-35-40	14-25-40	25-36-45	25-34-40
SapSidMan	**102-102-102**	**49-49-49**	24-36-47	22-29-33	**28-28-28**	11-20-28	20-28-37	18-26-31
SapSidTie	**102-102-102**	**49-49-49**	27-36-54	22-31-40	21-28-38	10-20-29	**28-28-28**	18-27-39
SapMimMan	**102-102-102**	60-75-87	**45-45-45**	28-31-41	**28-28-28**	11-24-34	23-34-46	19-28-40
SapMimTie	**102-102-102**	59-75-87	**45-45-45**	21-35-42	20-33-39	10-21-34	**28-28-28**	17-31-40
SapSidMimMan	**102-102-102**	**49-49-49**	**45-45-45**	28-32-44	**28-28-28**	11-24-40	24-34-45	19-28-40
SapSidMimTie	**102-102-102**	**49-49-49**	**45-45-45**	21-36-42	20-33-40	10-21-34	**28-28-28**	19-32-40

Fossil calibration scenario/ Profiled node	***Labramia* (node I) (in 101 trees)**	***Faucherea/ Labourdonnaisia* (node K) (in 101 trees)**	**small Asian *Manilkara* (node L) (in 101 trees)**	**Neotropical *Manilkara* (node N) (in 97 trees)**	**S. American *Manilkara* (node O) (in 101 trees)**	**C. American & Caribbean *Manilkara* (node P) (in 101 trees)**	**large African *Manilkara* (node T) (in 101 trees)**	**large Asian *Manilkara* (node U) (in 101 trees)**	**small African *Manilkara* (node Q) (in 101 trees)**
Sapotaceae *s.s.*	3-7-11	7-12-19	12-19-28	18-24-30	11-17-25	14-20-28	14-21-27	22-30-37	18-29-39
Sideroxyleae	2-5-11	4-9-17	7-14-25	10-18-28	7-13-24	8-16-26	9-16-25	12-23-35	11-22-35
Mimusops	2-6-10	7-12-19	11-17-26	17-23-29	10-16-22	13-19-25	13-19-25	21-28-33	18-27-37
Manilkara	2-5-8	5-9-14	9-14-21	13-18-27	9-13-19	10-15-22	10-15-21	17-22-32	14-21-37
Tieghemella	1-6-10	5-10-17	9-15-26	12-20-28	7-14-23	10-17-25	10-17-25	16-24-35	13-24-34
SapSid	2-6-11	4-10-17	6-16-25	10-20-28	7-15-24	9-17-26	9-18-25	12-25-35	11-24-35
SapMim	2-6-10	7-12-19	11-17-26	17-22-29	10-16-22	13-19-25	13-19-25	21-28-33	18-27-37
SapMan	2-5-8	5-9-14	9-14-21	13-18-26	9-13-19	10-15-22	10-15-21	17-22-31	14-21-37
SapTie	1-6-10	5-10-17	9-15-26	12-20-28	7-14-23	10-17-25	19-17-25	16-24-35	12-24-34
SapSidMim	2-7-18	7-12-19	11-17-26	17-23-32	11-16-24	14-19-27	13-20-25	21-28-37	18-27-41
SapSidMan	2-5-8	5-9-14	9-14-20	13-18-22	9-13-17	10-15-19	11-15-19	17-22-25	14-21-27
SapSidTie	1-5-9	5-9-17	9-14-23	12-18-26	7-13-21	10-15-22	10-16-23	15-23-32	13-22-32
SapMimMan	2-5-8	5-9-15	9-14-23	13-17-27	8-12-19	10-15-22	10-15-21	16-22-32	13-21-37
SapMimTie	2-6-16	5-11-19	8-16-25	12-21-29	9-15-22	10-18-25	10-18-25	13-26-32	12-25-36
SapSidMimMan	2-5-12	5-9-19	9-14-26	13-17-32	8-12-23	10-15-26	10-15-25	16-22-37	13-21-41
SapSidMimTie	2-6-21	5-11-19	8-16-25	12-22-29	9-15-22	10-18-25	10-19-25	15-26-36	12-26-37

Chapter VIII – Conclusions

8.1 Overview of findings

Systematic and biogeographic research on *Manilkara* is significant because, to date, no other pantropical genus with such an even spread of taxa across the Neotropics, Africa and Asia has been studied with molecular dating techniques and a nearly complete species level sample. This broad distribution and robust sampling has enabled the reconstruction of both finer scale biome level ecological patterns and grosser level intercontinental disjunctions. As such, *Manilkara* has proven to be an ideal model taxon with which to test hypotheses on the historical assembly of tropical forests worldwide.

8.1.1 Systematics

Hypotheses about *Manilkara*'s biogeography and systematic relationships were set out at the beginning of this thesis. Taxonomic questions were primarily concerned with the monophyly of *Manilkara* and related genera as outlined by Pennington (1991). In chapter IV a phylogenetic investigation of Pennington's classification of the tribe Mimusopeae and its constituent subtribes was made. The present research is in agreement with earlier studies (Swenson & Anderberg 2005, Smedmark *et al* 2006), that Mimusopeae is not a natural group due to the inclusion of the subtribe Glueminae and the genus *Northia*. Exclusion of these taxa renders Mimusopeae monophyletic and this change to the classification is backed up by morphological data. The subtribe Mimusopineae is also non-monophyletic and was instead resolved as a grade basal to the subtribe Manilkarinae, which is monophyletic with the exclusion of *Northia*. These findings are also supported by morphology and changes will need to be made to the classification to reflect these phylogenetic relationships.

There is not enough resolution in the basal branching structure of the subtribe Manilkarinae to discern relationships between the constituent clades. According to the phylogenetic trees presented, *Manilkara,* as currently defined, is not monophyletic and will need to be recircumscribed to exclude the clade comprising *M. fasciculata, M. dissecta* and *M. udoido*. In addition, the genus *Letestua* may need to be included within *Manilkara*, though this, as yet, is tentative. Likewise, *Faucherea* and *Labourdonnaisia* were both resolved as paraphyletic. All eighteen species (eleven and seven respectively) in the two genera should be sampled in order to determine their inter-relationships, in conjunction with a morphological study. However, it is likely that there has been hybridization between the two genera, as exemplified by the putative chloroplast capture event indicated by the present study, and, therefore, a tidy classification in relation to morphology may be difficult to reconcile.

In Africa, some *Manilkara* species, i.e. those in the *M. obovata* complex, remain difficult to delimit based on morphology. Despite these more or less minor changes listed above, the subtribe Manilkarinae and core *Manilkara* are good, sound groups upon which to base biogeographical inferences.

Considerable progress has been made in the past seven years in reconstructing evolutionary relationships between genera and tribes in the Sapotaceae, and the present research is a contribution towards this goal. The challenge now remains to reconcile morphology with molecular data.

Suggestions for future systematic research:

- As in any study, there were some taxa which were consistently difficult to amplify and sequence. Therefore, continuing to generate a complete sample of all species in *Manilkara* would be ideal. Of particular interest would be the inclusion in the ITS dataset of the Asian species *M. celebica*, *M. kanosiensis*, *M roxburghiana*, *M. napali* and *M. samoensis*, which were difficult to amplify due to the use of old, alcohol-collected herbarium material. The inclusion of the African species *M. yangambensis* in the ITS dataset would also help to clarify possible dispersal events between Africa and the Neotropics in the chloroplast capture scenario. Additionally, another accession of *Letestua durissima* should be sampled to check whether its placement within *Manilkara* is secure.

- Sequencing of all *Faucherea* and *Labourdonnaisia* species is necessary to determine phylogenetic relationships and aid in generic delimitation. This should be done in conjunction with a morphological study and taxonomic revision.

- The sequencing of additional chloroplast regions would contribute to a more robust dataset and help resolve basal relationships between clades in the subtribe Manilkarinae. It would be ideal to use the same regions (*ndhF, trnC–petN*, *petN–psbM*, *psbM–trnD*, *trnH*–psbA and *psB-psbH*), which have already been used in previous Sapotaceae studies to enable the combination of datasets. Likewise, the use of ETS and other low copy nuclear genes is also likely to help resolve basal relationships between clades.

- Mapping important morphological characters (number of calyx and corolla lobes, presence or absence of petal appendages and presence or absence of staminodes) onto the phylogeny, would help to visualise whether these characters are homoplasious or not and whether they remain valuable for delimiting genera and tribes.

8.1.2 Biogeography

At the beginning of this thesis, questions were also raised about *Manilkara*'s biogeography. Where and when did *Manilkara* originate and what factors have contributed to the pantropical distribution we see today? Has vicariance or dispersal played a more prominent role in creating this intercontinental disjunction? What does the timing of diversification on different continents tell us about the historical assembly of tropical forests in each region?

These questions were addressed in Chapter VII, where historical biogeography and regional

ecological patterns were explored. It was determined that *Manilkara* is likely to have originated in Africa during the Oligocene. From Africa *Manilkara* has dispersed once to Asia and once to the Neotropics. It also appears that there has been a dispersal from Madagascar to Asia in the *M. fasciculata*/*dissecta*/*udoido* lineage. All of these intercontinental dispersals occurred during the Miocene. This study contributes to the growing body of evidence in support of relatively recent intercontinental dispersal as being a major factor in the compilation and development of modern tropical floras (e.g. Renner 2004c).

Furthermore, within each continent, patterns of diversification were shown to coincide with geological and climatological activity. In South America, diversification appears to be coordinated with aridification and the rearrangement of drainage patterns in the Amazon basin as a result of Andean orogeny. The Atlantic coastal forest clade and the Amazonian clade of *Manilkara* split from one another approximately 14 Ma, during the mid-Miocene thermal maximum, at around the same time as diversification of clades restricted to the dry biomes of the cerrado and Caatinga was initiated. In Africa diversification also coincides with cycles of aridification throughout the Oligocene and Miocene as well as with the uplift of the east African plateaux. Phylogenetic niche conservatism was evident in *Manilkara*, with only two subclades switching biome from wet to dry forest. In one clade, an African rain forest ancestor dispersed to Madagascan rain forest, whereas in the other clade an African dry forest ancestor dispersed to Madagascan dry forest, but no dispersals involved a switch in biome and this is also an indicator of a degree of phylogenetic niche conservatism. In Southeast Asia the classical biogeographic divide of Wallace's Line has been shown not to have affected the dispersal of *Manilkara*. However, what is crucial is the appearance of land in New Guinea ~10 Ma as the Asian and Australian plates converged, which coincides with the dispersal and establishment of new taxa east of Wallace's Line.

Suggestions for future biogeographic research:

- A study of diversification rates of clades on each continent would help to discern whether speciation has occurred more rapidly in the Neotropics, Africa or Asia and whether an increase in speciation rate was coordinated with geo-climatic triggers.

- An emerging field in ancestral area reconstruction is the use of geographical and temporal constraints, i.e. the higher likelihood of dispersal/vicariance between neighboring areas than distant areas and during periods of time when such transitions are most probable, such as the emergence and submergence of islands in Southeast Asia or availability of intercontinental land bridges in the northern hemisphere during the Eocene. This approach is outlined in Ree *et al* 2005, Ree & Smith 2008, Ree & Sanmartin 2009 and is implemented in Lagrange (likelihood analysis of geographic range evolution). Although this method was not used in this study, it would be an interesting way to further test biogeographic hypotheses.

8.2 Relevance of this study in relation to other biome level studies

How does this research relate to the bigger picture and what can *Manilkara* tell us about the historical assembly of tropical forests worldwide? This research in combination with data from other tropical disjunct taxa can point towards an overall synthesis. Current methods in reconstructing historical biogeography predominantly focus on the geographical history of a single taxon rather than the combined area histories of all the taxa in a biome. This is mainly due to a lack of phylogenetic data, but that situation is changing with more and more studies emerging. No one has yet used a multi taxon approach to reconstruct the historical assembly of the main tropical rain forest blocks worldwide (i.e. Amazon Basin, Congo Basin and Malesia). However, studies on the floristic composition of the northern and southern hemispheres (Xiang *et al* 1998, 2000, Wen 1999, Donoghue & Smith 2004, Sanmartin & Ronquist 2004), the Andes (Sarkinen 2010), Neotropical seasonally dry forest (Pennington *et al* 2004), the South African Cape flora (Warren & Hawkins 2006, Galley & Linder 2006, Verboom *et al* 2009), the Australian sclerophyllous flora (Crisp *et al* 2004) and others have been carried out. Here they are reviewed to lend context to the study of tropical rain forest biome construction.

8.2.1 Northern and Southern hemisphere meta-analyses

In an attempt to investigate putative Gondwanan distributions, Sanmartin & Ronquist (2004) conducted an analysis of southern hemisphere disjunctions. They found that biogeographic patterns in plants have not been significantly influenced by Gondwanan breakup and suggested that this was either due to the fact that the studied groups are too young, or that more recent dispersal and extinction events have obscured the original Gondwanan vicariance event.

To investigate distribution patterns attributed to the fragmentation of the tertiary relict flora, meta-analyses of Eastern North America - East Asian disjuncts were carried out by Xiang *et al* (1998, 2000), Wen (1999) and Donoghue & Smith (2004). All of these studies suggest that the majority of disjunctions in the warm-temperate genera examined, occurred during the Miocene and are the result of the disintegration of a once continuous mixed mesophytic forest community due to a cooling climate. It should be noted, however, that these studies include very few disjunct taxa with both temperate and tropical lineages and, so, are biased towards representing only part of the story of the break-up of the Tertiary northern hemisphere flora. A meta-analysis of subtropical-tropical disjunct taxa between warm to tropical regions of Asia and the Americas is likely to reveal an earlier chapter in this story with the breakup of the boreotropical flora during the late Eocene-Oligocene.

8.2.2 Neotropical meta-analyses

In the Neotropics, seasonally dry tropical forest has been investigated by Pennington *et al* (2004), who found that this biome is at least mid Miocene to Pliocene age. An analysis of low to high altitude Andean taxa by Sarkinen (2010) suggests that diversification has occurred at a range of ages in different sub-biomes from the Mid Miocene on the low flanks

of the Andes to the late Pliocene in the high altitude grasslands. A study of Cerrado endemic lineages by Simon *et al* (2009), points to this biome also being of Late Miocene to Pliocene age (9.8-0.4 Ma). Lastly, Cody *et al* (2010) have shown that long distance dispersal has been an important mechanism in the assembly of the neotropical flora, as numerous taxa dispersed between Central and South America before the closing of the isthmus of Panama ~ 3Ma.

8.2.3 African meta-analyses

In Africa, the Cape flora has been widely investigated. Richardson *et al* (2001c) suggested that this unique biome may have arisen rapidly and recently, based on diversification in endemic Cape species of *Phylica* (Rhamnaceae), which began ~7 Ma and was coincident with aridification caused by the development of the cold Benguela current. A subsequent study by Warren and Hawkins (2006) showed that estimates for diversification in 14 endemic Cape lineages varied widely from 7-101 Ma and that it is likely that many lineages pre-date the establishment of the Cape Flora as a distinctive entity. Another study by Galley & Linder (2006), which sought to identify the origin and age of endemic Cape clades, backed up these claims. They found that Cape lineages were the result of intercontinental dispersal at various times throughout the Cenozoic from 15-80 Ma and suggested that the process of recruiting lineages to the Cape flora is on-going. Nearly half of their Cape clades showed a trans-Indian Ocean disjunction, while relatively few showed a sister relationship with African or Neotropical lineages and many showed a relationship with the Eurasian temperate flora. Numerous dispersal patterns recovered suggest that the Cape flora has a cosmopolitan heritage and has been assembled over a long period of time.

In a targeted study of the Fynbos and Karoo biomes, Verboom *et al* (2009) found that all succulent karoo-endemic lineages were less than 17.5 Ma old, the majority being younger than 10 Ma. Based on this, they suggest that recent radiation may have been triggered by climatic deterioration since the late Miocene. Fynbos-endemic lineages had a broader age distribution, with some lineages originating in the Oligocene, but the majority being more recent. They go on to argue that the fynbos and succulent karoo floras have rather different diversification histories. Galley *et al* (2007) and Bellstedt *et al* (2008) demonstrate that the migration of temperate taxa from the Cape into tropical East Africa has occurred in the last 17 Ma, consistent with the Mio-Pliocene formation of the East African Rift. Lastly, studies on the African pan-temperate element of the flora (Gehrke & Linder 2009) have shown that it is almost entirely northern hemisphere derived, suggesting that the Holartic is the most important source of angiosperm lineages in the African high mountain flora in both tropical and southern Africa.

8.2.4 Southeast Asian and Australasian meta-analyses

No meta-analyses have yet been performed on Southeast Asian biomes, but a review of migration patterns of angiosperm taxa across Wallace's Line (Richardson *et al* 2010) suggests that successful dispersals from west to east are concordant with the appearance of land in New Guinea during the Miocene ~10 Ma and dispersals from east to west are

probably less common due to phylogenetic niche conservatism of dry adapted lineages in the Australian flora. New Guinea is primarily covered in wet forest and, therefore, assuming the lesser probability of biome switching, is more likely to have been invaded from the west as there is a greater area of wet forest on the Sunda shelf than to the east in Australia, the wet forest of which has been contracting.

In Australia, Crisp *et al* (2004) demonstrate that climatic changes during the late Oligocene and Miocene are reflected in the phylogenies of biome-endemic taxa. A decline in the diversity of rain forest lineages is coordinated with the development of a cooler, drier more seasonal climate from 25-10 Ma, while a concomitant increase in speciation rate was found in sclerophyllous lineages.

In New Zealand, studies suggest that a large proportion of the modern flora was recently recruited from Australia by dispersal across the Tasman Sea, as opposed to being of Gondwanan stock (McGlone *et al* 2005, Winkworth *et al* 2002).

8.3 Theories about the historical construction of tropical rain forest

What can we begin to discern about biome construction worldwide from these initial studies? Have these biomes been shown to be ancient with a gradual accumulation of lineages or recent and rapid, or a combination of rates and modes?

It would appear that the studied biomes span a range of ages, although the majority are relatively recent (Miocene-Pliocene). The Cape flora studies are a prime example of this, where some characteristic taxa have been in place since the Cretaceous (Warren & Hawkins 2006, Galley & Linder 2006), whereas others have radiated recently due to climatic changes initiating specialised sub-biomes such as the Fynbos and Karoo (Verboom *et al* 2009).
The common message here is that, like the taxa of which they are comprised, biomes are constantly evolving, with climatic and geological activity often being the catalysts, which instigate a change in the rate of their evolution.

The obvious missing element in these biome studies is tropical rain forest. Thus far, no other comprehensive, dated studies on pantropical genera have been published and no meta-analyses have been carried out on rain forest restricted taxa on a global scale. Although angiosperms began to evolve in the late Cretaceous, the fossil record in South America (Burnham & Johnson 2004, Wing *et al* 2009) and Africa (Jacobs 2004) suggests that tropical rain forest did not emerge as a biome until ~60 Ma (Pennington *et al* 2006). Many extant angiosperm lineages originated in the tropical zone due to the nearly pole to pole extent of warm and mesic climates during the Paleocene-Eocene Thermal Maximum. Wiens and Donoghue (2004) suggest that if much of the world was tropical for a long period (until 30-40 Ma), then more extant clades should have originated in the tropics than in temperate regions. Since the onset of global climate deterioration at the end of the Eocene, tropical rain forest taxa have had to either move with the changing climate towards the equator or evolve to deal with cooler environments. According to Donoghue (2008), half of angiosperm families have no temperate representatives, implying that it is not easy to evolve tolerance to

freezing temperatures and seasonal environments. However, Haffer (1969) hypothesized that such climate fluctuations could act as an engine for species diversity when forest contracted to isolated refugial pockets during cool and arid periods facilitating allopatric speciation. Alternatively, Stebbins (1974) put forward the idea that tropical rain forest can be considered a museum, rather than a cradle of speciation, due to the steady accumulation of lineages over time in a climatically unperturbed environment with a low extinction rate. These opposing theories, that rain forest diversity is recent and speciation is triggered by climatic cycles or that rain forests are ancient and more or less unchanging, can be tested with dated phylogenies and an examination of the fossil record.

How old are pantropical rain forest lineages such as *Manilkara*, and can their ages be used as a proxy for biome age?

8.4 The bigger picture – comparison with other tropical disjunct taxa

8.4.1 Methods

To gauge whether the age and biogeographic pattern exhibited by *Manilkara* is typical of other rain forest lineages and whether predictions can be made about biome construction based on the results of this study, a brief survey of dated phylogenies (Appendix 8.1) of tropical rain forest genera with intercontinental disjunctions was conducted. These studies vary widely in their level of taxon sampling and dating methods employed. All genera in this survey occur in tropical forest and have ranges that exhibit at least one intercontinental disjunction. (There were numerous studies of other genera, which have disjunctions between Africa and Madagascar or throughout Southeast Asia to the Pacific, but these are not strictly intercontinental disjunctions and were not considered here.) Thirty four genera (twenty-eight woody and six herbaceous) were surveyed in twenty families from a range of clades (i.e. basal dicots, monocots and eudicots). In each case, time of diversification is defined by the crown node age, although the initial appearance of a taxon could be older, given that sampling was usually poor.

8.4.2 Results

Given the different sampling and dating methods as well as the differing amounts of data explicitly presented concerning the timings of disjunctions in each of the studies, it is hard to make an even comparison without significant further study and analysis. However, a preliminary estimate reveals that out of the thirty-four genera, six had crown node ages in the Cretaceous, eleven in the Eocene, eight in the Oligocene and nine in the Miocene (however, six of these taxa had an estimated age spread that overlapped epoch boundaries). There were no evident vicariance/dispersal events during the Cretaceous or Paleocene, whereas nine occurred during the Eocene, five during the Oligocene, twenty-nine during the Miocene and four during the Pliocene. Furthermore, the disjunct distribution in none of the genera was old enough to have been the result of Gondwanan break-up, whereas ten were possibly the result of boreotropical migration, and a minimum of forty long distance dispersal events were inferred. These comprised at least fifteen trans-Atlantic dispersals, at least twenty trans-Indian Ocean dispersals and at least five trans-Pacific dispersals.

8.4.3 Discussion - implications for tropical rain forest biome age

Manilkara, with two trans-Indian and one trans-Atlantic dispersal event, and *Mimusops* with one trans-Indian Ocean dispersal during the Miocene both fit this overall trend. These preliminary results overwhelmingly suggest that Miocene long distance dispersal was an important agent in the development of modern tropical floras. It would appear that the Indian and Atlantic oceans have relatively similar rates of successful dispersal events with fewer successful dispersals occurring across the Pacific (i.e. between Southeast Asia and the Neotropics), but this bias may disappear with the addition of more taxa to the analysis.

Given that long distance dispersal is not limited in time (i.e. presumably dispersal has occurred as long as there have been organisms to disperse) why does the Miocene stand out as a period in which dispersal events are particularly successful? This may to be due to the fact that extensive climatic oscillations and tectonic activity during the Miocene would have weakened existing populations, opening new habitat niches and enabling successful colonization by immigrants. Additionally, it is difficult to differentiate dispersal events, which have occurred since the Oligocene from vicariance events, and this may make the frequency of Miocene dispersals appear more pronounced.

It is clear from dated phylogenies that geological and climatic events are important environmental triggers for speciation, because radiations and innovations in numerous lineages, particularly in *Manilkara*, have been shown to coincide with such events. Miocene climate oscillations and geological upheaval in each of the tropical regions appears to have left a clear signal in many phylogenies, i.e. the uplift of the Andes and consequent re-arrangement of the Amazon drainage basin as well as the closure of the Isthmus of Panama in the Neotropics, the uplift of the East African Rift in Africa, as well as sea level oscillations on the Sunda shelf combined with the closure of the Indonesian through-flow and creation of land in New Guinea.

As shown by this brief survey, another important contributing factor to current intercontinental tropical disjunctions is the possibility that tropical rain forest taxa may have migrated through the northern hemisphere during the Paleocene-Eocene thermal maximum. However, this scenario is difficult to confirm when there are no extant species in the northern hemisphere to sample in a phylogeny; although copious fossil evidence supports many claims to this migration route. As in the Miocene, cooling climates at the end of the Eocene thermal maximum forced the equatorial migration or extinction of taxa.

The curious outlier in this survey is the role of Gondwanan vicariance in the creation of modern intercontinental disjunctions. This and numerous other recent phylogenetic studies have shown that many modern genera are Oligocene or Miocene in age. Genera are, therefore, often too young to reflect evidence of Gondwanan fragmentation because such patterns are overlain with more recent dispersals and radiations. It is unlikely that we will find many modern genera which are Cretaceous or Paleocene in age. For example, while the basal dicot family Monimiaceae is ~102 Ma, all genera are of Miocene age (Renner 2005).

If they haven't gone extinct, most taxa, which underwent Gondwanan vicariance, will have evolved into families and orders. This is because, as evolution progresses, today's species become tomorrow's genera and families and in doing so, some of the more ancient patterns become obscured. Therefore, Gondwanan vicariance is often only detectable at deeper nodes in a phylogeny, such as at tribal and family level. An example of this is in the classic Gondwanan genus *Nothofagus*, where a dated phylogenetic analysis of extant taxa reveals recent long distance dispersal and radiation rather than Gondwanan vicariance (Knapp *et al* 2005, Cook & Crisp 2005b). However, when fossil data are considered, a former Gondwanan distribution is evident (Swenson *et al* 2001, Cook & Crisp 2005b). To explain this incongruence, Sanmartin & Ronquist (2004) suggest that the distribution of *Nothofagus* was initially shaped by Gondwanan vicariance, but subsequent extinctions have resulted in extant species showing a pattern in conflict with geology. Recent dispersal is also overlain on an older Gondwanan vicariance pattern in the Atherospermataceae (Renner *et al* 2000).

Lastly, it is important to remember that the diversification of intercontinental disjunct genera is only part of the picture because it doesn't tell us about intracontinental radiations, particularly in those genera, which are regional or biome endemics. One example of such a genus is *Inga* (Leguminosae), which has undergone rapid speciation in the Amazon and the lower flanks of the Andes, generating ~300 species in the past two million years (Richardson *et al* 2001a). Another is the Southeast Asian genus *Cyrtandra* (Gesneriaceae) with ~600 species in Southeast Asia and the Pacific (Cronk *et al* 2005), which is likely to have diversified at a similar rate.

From these preliminary findings we can infer that the pantropical rain forest biome is composed of taxa of a range of ages, some of which have persisted in-situ since the Paleocene throughout climatic oscillations and others that are recent immigrants, or have recently radiated to fill open niches following geo-climatic events. Thus the tropical rain forest biome is likely to be older than biomes such as the Cerrado, seasonally dry Neotropical forest, the high Andes, the Karoo, the Afro-montane flora and the Australian sclerophyllous flora, but many of the taxa of which the modern rain forest biome is comprised are likely to be young due to a high rate of immigration and recent rapid diversification in response to environmental pressures. Patterns exhibited in numerous phylogenetic studies of tropical taxa suggest that relatively recent (Miocene) intercontinental dispersal and subsequent radiation has played a very significant role in the assembly of modern tropical rain forests. Hence, they are not just the static museums, which Stebbins (1974) hypothesized they might be. Instead age ranges of taxa reveal a mixture of old (long persisting) and young (recently diversified) elements reflecting a combination of both Stebbins (1974) and Haffer's (1969) models.

Finally, *Manilkara* has proven to be a valuable model with which to investigate patterns of diversification in each of the world's major tropical regions. In doing so, this research has provided a better understanding of the historical assembly of the tropical rain forest biome. However, given that a preliminary biome-level comparative analysis suggests that the pantropical rain forest is composed of taxa of various ages with different histories,

Manilkara should be viewed as just one part of the wider story. Nonetheless, other dated phylogenies are increasingly revealing that a significant component of modern floras results from the relatively recent immigration of new taxa via long distance dispersal. In light of this, the pattern reflected in *Manilkara*'s distribution would appear to be a key part of the story. Still, many more taxa need to be studied in concert to gain a better understanding of the historical assembly of the tropical rain forest biome worldwide.

8.4.4 Suggestions for further research

To further investigate biome assembly many more fully sampled, accurately dated phylogenies of tropical rain forest taxa are needed. Additionally, all taxa would need to be re-analyzed using a standardized dating method. There was not enough data in the majority of the surveyed studies to discern whether speciation was coordinated with geo-climatic events on each continent, but future studies should take this into account. Potential questions for investigation are:

- Do pantropical or intercontinental disjunct genera share a similar age structure in their disjunctions? If so, does this coincide with a known vicariance scenario or dispersal event (i.e. Gondwanan vicariance or boreotropical migration) or not?

- Are particular dispersal pathways more common than others (i.e. trans-Atlantic, versus trans-Indian, versus trans-Pacific)? If so, can a vector be identified (i.e. a specific ocean current)?

- Does a shift in distribution correspond to a shift in biome? If so, how often does this occur?

- Do radiations occur within biomes or between biomes (i.e. wet Guineo-Congolian rain forest to dry East African scrub forest or wet Amazon to dry Cerrado)?

- Is diversification in a clade correlated with a geo-climatic trigger (i.e. aridification or the uplift of a mountain range)?

- Is the difference in species richness in the Neotropics, Africa and Asia due to different pressures? For example, does the relative importance of long persistence of clades without extinction versus recent and rapid radiation vary between these regions?

8.5 Appendix

Appendix 8.1 Surveyed dated phylogenetic studies of genera with intercontinental disjunctions.

Family	Genus	Number of species sampled	Distribution	Crown node	Intercontinental disjunction	Reference
Boraginaceae/ Ehretiaceae	*Ehretia*	11 out of ~50 spp.	pantropical	102-93 Ma Cretaceous	Place of origin equivocal, dispersal from the New to Old world 37 Ma	Gottschling *et al* 2004
Boraginaceae/ Cordiaceae	*Cordia*	12 out of 250-300 spp.	pantropical	95-92 Ma Cretaceous	Origin in the Neotropics, a dispersal from the Neotropics to Asia 44-52 Ma and a second dispersal from the Neotropics to Africa 15-16 Ma and from Africa to Asia during the Miocene	Gottschling *et al* 2004
Piperaceae	*Peperomia*	16 out of ~1600 spp.	pantropical	88 Ma Cretaceous	Place of origin equivocal, numerous migration and dispersal events between the Neotropics, Africa, Asia and the Pacific from the Eocene to the Miocene	Smith *et al* 2008
Boraginaceae/ Heliotropiaceae	*Euploca*	3 out of ~185 spp.	pantropical	86-46 Ma Cretaceous-Eocene	Place of origin equivocal - split between New & Old World lineages, hypothesized dispersal from the Neotropics to Africa and subsequently to Australia	Gottschling *et al* 2004
Piperaceae	*Piper*	49 out of ~1050 spp.	pantropical	71 Ma Cretaceous	Origin in the Neotropics (west Gondwana), one dispersal to Africa and the Pacific at 37 Ma and another dispersal from the Neotropics to Africa ~10 Ma, a further dispersal from Asia to Africa ~10 Ma. Two dispersals from Asia to Australia/New Caledonia around 30 and 9 Ma	Smith *et al* 2008
Sapotaceae	*Sideroxylon*	45 out of ~75 spp.	pantropical	68 Ma Cretaceous	Place of origin equivocal, but Boreotropics hypothesized– basal split between Africa/Middle East, Neotropics and Asia. Migration/dispersal from Asia to Africa and the Indian Ocean islands 52 Ma.	Smedmark & Anderberg 2007
Hernandiaceae	*Hernandia*	12 out of 23 spp.	pantropical	50 Ma Eocene	Origin in Australia, dispersal from Australia into the Pacific at 28 Ma, and from Australia/Pacific to South America 16 Ma, further dispersal from South America to Sao Tome/Africa 3 Ma	Michalak *et al* 2010

Family	Genus	Number of species sampled	Distribution	Crown node	Intercontinental disjunction	Reference
Boraginaceae/ Ehretiaceae	*Bourreria*	4 out of ~50 spp.	Neotropics - Africa	45-42 Ma Eocene	Place of origin equivocal - New & Old World split at 42-45 Ma	Gottschling *et al* 2004
Moraceae	*Ficus*	9 out of ~750-850 spp.	pantropical	43 Ma Eocene	Place of origin equivocal - split between New and Old world taxa	Zerega *et al* 2005
Lauraceae	*Cinnamomum*	4 out of ~250 spp.	Neotropics - Asia	~42 Ma Eocene	Place of origin equivocal - split between Asia and South America	Chanderbali *et al* 2001
Annonaceae	*Anaxagorea*	3 out of 21 spp.	Neotropics – Asia	38-33 Mya Eocene	Origin in South America with Oligocene dispersal to Southeast Asia ~25 Ma	Richardson *et al* 2004
Annonaceae	*Xylopia*	4 out of 160 spp.	pantropical	37-23 Mya Eocene - Oligocene	Origin in Asia with dispersal to Africa and South America during the early Miocene ~16 Ma	Richardson *et al* 2004
Cucurbitaceae	*Momordica*	58 out of ~59 spp.	Old World	35 Ma Eocene	Origin in Africa with dispersal to Asia 12 Ma	Schaefer & Renner 2010
Malpighiaceae	*Acridocarpus*	14 out of 29 spp.	Old World	~35 Ma Eocene	Place of origin equivocal - split between Africa and Madagascar, dispersal from Madagascar to New Caledonia ~8 Ma	Davis *et al* 2002
Gentianaceae	*Exacum*	30 out of 64 spp.	Old World	35-8 Ma Eocene - Miocene	Origin in Madagascar, late Miocene dispersals from Madagascar to Asia and Australia	Yuan *et al* 2005
Orchidaceae	*Vanilla*	50 out of 106 spp.	pantropical	34 Ma Eocene	Origin in South America with dispersal to Africa 25 Ma	Bouetard *et al* 2010
Melastomataceae	*Memecylon*	2 out of 150 spp.	Old World	Eocene	Place of origin equivocal - split between Africa-Asia during the Eocene	Renner 2004b
Burseraceae	*Commiphora*	37 out of 150 spp.	pantropical	32-23 Ma Oligocene	Hypothesized origin in Africa, dispersal from Africa to South America 24 Ma and Africa to India 5 Ma	Weeks *et al* 2007
Hernandiaceae	*Gyrocarpus*	4 out of 4 spp.	pantropical	31 Ma Oligocene	Origin in the Neotropics, dispersal from Neotropics to Africa 19 Ma and from Africa to Asia and the Neotropics 2 Ma	Michalak *et al* 2010
Boraginaceae/ Heliotropiaceae	*Tournefortia*	3 out of ~53 spp.	pantropical	30-16 Ma Oligocene - Miocene	Place of origin equivocal - split between New & Old World lineages16-30 Ma	Gottschling *et al* 2004
Hernandiaceae	*Illigera*	10 out of 18 spp.	Old World	27 Ma Oligocene	Place of origin equivocal, split between African and Asian lineages at 27 Ma, with a further dispersal from Asia to Africa 5 Ma	Michalak *et al* 2010

Family	Genus	Number of species sampled	Distribution	Crown node	Intercontinental disjunction	Reference
Begoniaceae	*Begonia*	23 out of ~1400 spp.	pantropical	~26 Ma Oligocene	Origin in Africa, dispersal from Africa to the Neotropics 23 Ma, and Socotra & Asia 15 Ma. Back dispersal from Asia to the Neotropics ~15 Ma and Neotropics to Africa 7 Ma	Thomas 2010
Annonaceae	*Annona*	7 out of 137 spp.	Neotropics-Africa	25-21 Mya Oligocene – Miocene	Origin in South America with dispersal to Africa 21-25 Ma	Richardson *et al* 2004
Melastomataceae	*Medinilla*	2 out of ~375 spp.	Old World	24-15 Ma Oligocene – Miocene	Place of origin equivocal - split between Madagascar and Asia	Renner 2004b
Anisophylleaceae	*Anisophyllea*	12 out of 27 spp.	pantropical	23 Ma Oligocene	Place of origin equivocal - split between Neotropical and Old World, dispersal between Africa and Asian 22 Ma	Zhang *et al* 2007
Simaroubaceae	*Soulamea*	3 out of 14 spp.	Old World	Early Miocene	Origin in Asia-Pacific, dispersal between Asia-Pacific and Madagascar during the Mid Miocene	Clayton *et al* 2009
Zingiberaceae	*Renealmia*	24 out of ~75 spp.	Neotropics - Africa	15-2 Ma Miocene	Origin in Africa and subsequent late Miocene dispersal from Africa to South America	Sarkinen *et al* 2007
Bombacaceae	*Adansonia*	8 out of 8 spp.	Old World	15-2 Ma Miocene	Place of origin equivocal, late Miocene dispersal between Africa/Madagascar and Australia	Baum *et al* 1998
Annonaceae	*Uvaria*	3 out of ~100 spp.	Old World	14-12 Mya Miocene	Origin in Asia with subsequent dispersal to Africa ~10 Ma	Richardson *et al* 2004
Simaroubaceae	*Quassia*	2 out of ~40 spp.	pantropical	Mid Miocene	Place of origin equivocal - split between Africa and South America Mid Miocene	Clayton *et al* 2009
Simaroubaceae	*Picrasma*	5 out of ~8 spp.	Neotropics-Asia	Mid Miocene	Origin in Asia, dispersal from Australia to South America during the Late Miocene	Clayton *et al* 2009
Moraceae	*Dorstenia*	2 out of ~105 spp.	pantropical	~10 Ma Miocene	Place of origin unstudied/unknown, but occurred in Miocene suggesting pantropical distribution is due to intercontinental dispersal	Zerega *et al* 2005
Phyllanthaceae	*Bridelia*	22 out of 37 spp.	Old World	10 Ma Miocene	Origin in Asia, dispersal from tropical Asia to Africa once or twice between 10-1.85 Ma. New Guinea to Australia at least twice ~ 2 Ma	Li *et al* 2009
Rubiaceae	*Gaertnera*	30 out of 70 spp.	Old World	10-5 Ma Miocene	Origin in Africa, dispersal from Africa to Southeast Asia ~ 5 Ma	Malcomber 2002

ACOSTA, M. C. & PREMOLI, A. C. (2010). Evidence of chloroplast capture in South American *Nothofagus* (subgenus Nothofagus, Nothofagaceae). *Molecular Phylogenetics and Evolution* 54(1): 235-242.

ADANSON, M. (1763). *Familles des Plantes vol. 2*. Paris: Chez Vincent, Imprimeur-Libraire de Mgr le Comte de Provence.

ADHIKARI, B. (2010). Systematics and phylogeographic studies of *Berberis* L. (Berberidaceae) in the Nepal Himalaya. Ph.D. thesis, Edinburgh: University of Edinburgh.

AKGUN, F., AKAY, E. & ERDOGAN, B. (2002). Tertiary terrestrial to shallow marine deposition in central Anatolia: a palynological approach. *Turkish Journal of Earth Sciences* 11: 127-160.

AKGUN, F., KAYSERI, M. S. & AUIRAZ, M. S. (2007). Palaeoclimatic evolution and vegetational changes during the Late Oligocene-Miocene period in Western and Central Anatolia (Turkey). *Palaeogeography Palaeoclimatology Palaeoecology* 253: 56-90.

ALI, J. R. & AITCHISON, J. C. (2008). Gondwana to Asia: plate tectonics, paleogeography and the biological connectivity of the Indian sub-continent from the Middle Jurassic through latest Eocene (166-35) Ma. *Earth Science Reviews* 88: 145-166.

ALI, J. R. & HUBER, M. (2010). Mammalian biodiversity on Madagascar controlled by ocean currents. *Nature* 463(7281): 653-680.

ALVAREZ, I. & WENDEL, J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29(3): 417-434.

ANDERBERG, A. A. & SWENSON, U. (2003). Evolutionary lineages in Sapotaceae (Ericales): A cladistic analysis based on ndhF sequence data. *International Journal of Plant Sciences* 164(5): 763-773.

ANDRAE, C. J. (1853). *Beiträge zur Kenntniss der fossilen Flora Siebenbürgens und des Banates*. Wien: Aus der K.K. Hof- und Staatsdruckerei.

ANDREASEN, K. & BALDWIN, B. G. (2001). Unequal evolutionary rates between annual and perennial lineages of checker mallows (*Sidalcea*, Malvaceae): Evidence from 18S-26S rDNA internal and external transcribed spacers. *Molecular Biology and Evolution* 18(6): 936-944.

ANGIOSPERM PHYLOGENY GROUP (2003). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* 141: 399-436.

ANGIOSPERM PHYLOGENY GROUP (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* 161(2): 105-121.

ANTONELLI, A., NYLANDER, J. A. A., PERSSON, C. & SANMARTIN, I. (2009). Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proceedings of the National Academy of Sciences of the United States of America* 106(24): 9749-9754.

ARCHIBALD, J. K., MORT, M. E. & CRAWFORD, D. J. (2003). Bayesian inference of phylogeny: a non-technical primer. *Taxon* 52: 187-191.

Atkins, H., Preston, J. & Cronk, Q. C. B. (2001). A molecular test of Huxley's line: *Cyrtandra* (Gesneriaceae) in Borneo and the Philippines. *Biological Journal of the Linnean Society* 72(1): 143-159.

Aubréville, A. (1964). Les Sapotacées: taxonomie et phytogéographie. *Adansonia* Mémoires.1: 1-157.

Awasthi, N. (1977). On two new fossil woods resembling *Chrysophyllum* and *Holoptelea* from the Cuddalore series near Pondicherry. *Paleobotanist* 24: 21-25.

Awasthi, N. & Mehrota, R. C. (1993). Further contribution to the Neogene flora of northeast India and significance of the occurrence of African element. *Geophytology* 23(1): 81-92.

Awasthi, N. & Srivastava, G. P. (1990). Some new carbonised woods from Neogene of Kerala coast and their bearing on paleoclimate. *Paleobotanist* 38: 285-292.

Axelrod, D. I. (1939). *A Miocene flora from the western border of the Mohave desert*. Washington, D.C.: Carnegie Institution of Washington. 516: 1-128.

Axelrod, D. I. (1950). *The Anaverde flora of southern California.* Washington, D.C.: Carnegie Institute of Washington 590: 119-158.

Axelrod, D. I. (1956). *Mio-Pliocene floras from west-central Nevada*. University of California Publications in Geological Science, vol. 33.

Azevedo, V. C. R., Vinson, C. C. & Ciampi, A. Y. (2005). Twelve microsatellite loci in *Manilkara huberi* (Ducke) Standl (Sapotaceae), an Amazonian timber species. *Molecular Ecology Notes* 5(1): 13-15.

Azuma, H., Garcia-Franco, J. G., Rico-Gray, V. & Thien, L. B. (2001). Molecular phylogeny of the Magnoliaceae: the biogeography of tropical and temperate disjunctions. *American Journal of Botany* 88(12): 2275–2285.

Baehni, C. (1938). Mémoires sur les Sapotacées. I. Système de classification. *Candollea* 7: 394–508.

Baehni, C. (1965). Mémoires sur les Sapotacées 3. Inventaire des genres. *Boissiera* 11: 1-262.

Baillon, M. H. (1891). *Histoire des Plantes: Sapotacées*. Paris: Libraire de L. Hachette.

Bakar, A. M. (2009). A phylogenetic and biogeographic study of Isonandreae (Sapotaceae) with special reference to the genus *Palaquium* using ITS sequence data. M.Sc. Thesis, Edinburgh: University of Edinburgh.

Baldwin, B. G. & Markos, S. (1998). Phylogenetic utility of the external transcribed spacer (ETS) of 18S-26S rDNA: Congruence of ETS and ITS trees of *Calycadenia* (Compositae). *Molecular Phylogenetics and Evolution* 10(3): 449-463.

Baldwin, B. G., Sanderson, M. J., Porter, J. M., Wojciechowski, M. F., Campbell, C. S. & Donoghue, M. J. (1995). The ITS region of nuclear ribosomal DNA - a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden* 82(2): 247-277.

Baldwin, B. G. & Wagner, W. L. (2010). Hawaiian angiosperm radiations of North American origin. *Annals of Botany* 105(6): 849-879.

Bande, M. B., Dechamps, R., Lakhanpal, R. N. & Prakash, U. (1987). Some new fossil woods from the Cenozoic of Zaire. *Musée Royale Afrique Centrale, Tervuren (Belgium), Départment* de Géologie et Minéralogie*, Rapport annuel 1985-1986*: 113-140.

Bande, M. B. & Srivastava, G. P. (1990). Late Cenozoic plant-impressions from Mahuadanr Valley, Palamu District, Bihar. *Paleobotanist* 37(3): 331-366.

Banfer, G., Moog, U., Fiala, B., Mohamed, M., Weising, K. & Blattner, F. R. (2006). A chloroplast genealogy of myrmecophytic *Macaranga* species (Euphorbiaceae) in Southeast Asia reveals hybridization, vicariance and long-distance dispersals. *Molecular Ecology* 15(14): 4409-4424.

Barber, H. N., Dadswell, H. E. & Ingle, H. D. (1959). Transport of driftwood from South America to Tasmania and Macquarie island. *Nature* 184(4681): 203-204.

Barker, N. P., Weston, P. H., Rutschmann, F. & Sauquet, H. (2007). Molecular dating of the "Gondwanan" plant family Proteaceae is only partially congruent with the timing of the break-up of Gondwana. *Journal of Biogeography* 34(2012-2027).

Bartish, I. V., Antonelli, A., Richardson, J. E. & Swenson, U. (2010). Vicariance or long-distance dispersal: historical biogeography of the pantropical subfamily Chrysophylloideae (Sapotaceae). *Journal of Biogeography*. In press.

Bartish, I. V., Swenson, U., Munzinger, J. & Anderberg, A. A. (2005). Phylogenetic relationships among New Caledonian Sapotaceae (Ericales): Molecular evidence for generic polyphyly and repeated dispersal. *American Journal of Botany* 92(4): 667-673.

Baum, D. A., Small, R. L. & Wendel, J. F. (1998). Biogeography and floral evolution of baobobs (*Adansonia*, Malvaceae) as inferred from multiple data sets. *Systematic Biology* 47(2): 181-207.

Becerra, J. X. (2003). Synchronous coadaptation in an ancient case of herbivory. *Proceedings of the National Academy of Sciences of the United States of America* 100(22): 12804-12807.

Becerra, J. X. (2005). Timing the origin and expansion of the Mexican tropical dry forest. *Proceedings of the National Academy of Sciences of the United States of America* 102(31): 10919-10923.

Bell, C. D. & Donoghue, M. J. (2005). Phylogeny and biogeography of Valerianaceae (Dipsacales) with special reference to the South American valerians. *Organisms Diversity & Evolution* 5(2): 147-159.

Bellstedt, D. U., Van Zyl, L., Marais, E. M., Bytebier, B. L. G., De Villiers, C. A., Dreyer, L. L. & Makwarela, A. M. (2007). Evidence of repeated migrations through the African and corridor as revealed by a molecular phylogeny of southern African members of the genus *Zygophyllum*. *South African Journal of Botany* 73(2): 278-279.

Bentham, G. & Hooker, J. D. (1876). *Genera Plantarum vol. 2*. London: Lovell Reeve & Co.

BERRY, E. W. (1915). An Eocene ancestor of the zapodilla. *American Journal of Science* 39(230): 208-213.

BERRY, E. W. (1916a). The Catahoula sandstone. *United States Geological Survey Professional Papers* 98M: 227-251.

BERRY, E. W. (1916b). The flora of the Citronelle formation. *United States Geological Survey Professional Papers* 98L: 193-208.

BERRY, E. W. (1916c). The Lower Eocene floras of southeastern North America. *United States Geological Survey Professional Papers* 91: 1-481.

BERRY, E. W. (1916d). The physical conditions and age indicated by the flora of the Alum Bluff Formation. *United States Geological Survey Professional Papers* 98-E: 1-59.

BERRY, E. W. (1918). Fossil plants from the late Tertiary of Oklahoma. *Proceedings of the United States National Museum* 54: 627–636.

BERRY, E. W. (1919). The Upper Cretaceous floras of the eastern Gulf region. *United States Geological Survey Professional Papers* 112: 117.

BERRY, E. W. (1921). Tertiary fossil plants from the Dominican Republic. *Proceedings of the United States National Museum* 59 117-127.

BERRY, E. W. (1922a). Pliocene fossil plants from eastern Bolivia. *Johns Hopkins University Studies in Geology* 4: 145-204.

BERRY, E. W. (1922b). Tertiary fossil plants from the Republic of Haiti. *Proceedings of the United States National Museum* 62: 1-10.

BERRY, E. W. (1924). The Middle and Upper Eocene floras of southeastern North America. *United States Geological Survey Professional Papers* 92: 1-206.

BERRY, E. W. (1925a). The flora of the Ripley formation. *United States Geological Survey Professional Papers* 36: 1-94.

BERRY, E. W. (1925b). A Pleistocene flora from the island of Trinidad. *Proceedings of the United States National Museum* 66: 1-9.

BERRY, E. W. (1925c). The Tertiary flora of the island of Trinidad. *Johns Hopkins University Studies in Geology* 6: 71-162.

BERRY, E. W. (1929). Tertiary fossil plants from Colombia, South America. *Proceedings of the United States National Museum* 75(24): 1-12.

BERRY, E. W. (1930). Revision of the Lower Eocene Wilcox flora of the southeastern States with descriptions of new species, chiefly from Tennessee and Kentucky. *United States Geological Survey Professional Papers* 156: 1-196.

BERRY, E. W. (1934). Pleistocene plants from Cuba. *Torrey Botanical Club Bulletin* 61: 237-240.

BERRY, E. W. (1936). Tertiary plants from Venezuela. *Proceedings of the United States National Museum* 83: 335-360.

Berry, E. W. (1937). A flora from the Forest clay of Trinidad. *Johns Hopkins University Studies in Geology* 12: 51-68 .

Berry, E. W. (1938). Tertiary flora from the Rio Pichileufu, Argentina. *Geological Society of America, Special Papers* 12: 1-149.

Berry, E. W. (1939). A Miocene flora from the gorge of the Yumari river, Matanzas, Cuba. *Johns Hopkins University Studies in Geology* 13: 95-135.

Berry, E. W. (1945). Fossil floras from southern Ecuador. *Johns Hopkins University Studies in Geology* 14: 93-150.

Biondi, E. (1981). *Arganioxylon sardum* n. gen., n. sp. and *Sclerocaryoxylon chiarugii* n. gen., n. sp. fossil woods from the Miocene in Sardinia (Italy). *Review of Palaeobotany and Palynology* 34(3-4): 301-320.

Bohme, M., Bruch, A. A. & Selmeier, A. (2007). The reconstruction of Early and Middle Miocene climate and vegetation in Southern Germany as determined from the fossil wood flora. *Palaeogeography, Palaeoclimatology, Palaeoecology* 253: 91-114.

Bouetard, A., Lefeuvre, P., Gigant, R., Bory, S., Pignal, M., Besse, P. & Grisoni, M. (2010). Evidence of transoceanic dispersion of the genus *Vanilla* based on plastid DNA phylogenetic analysis. *Molecular Phylogenetics and Evolution* 55(2): 621-630.

Bremer, K., Friis, E. M. & Bremer, B. (2004). Molecular phylogenetic dating of asterid flowering plants shows early Cretaceous diversification. *Systematic Biology* 53(3): 496-505.

Bridson, D. & Forman, L. (2000). *The herbarium handbook*. London: Royal Botanic Gardens Kew.

Briggs, J. C. (1987). *Biogeography and Plate Tectonics*. Amsterdam: Elsevier.

Briggs, J. C. (1989). The historic biogeography of India: islolation or contact? *Systematic Zoology* 38(4): 322-332.

Briggs, J. C. (2003). The biogeographic and tectonic history of India. *Journal of Biogeography* 30(3): 381-388.

Bromham, L. & Penny, D. (2003). The modern molecular clock. *Nature Reviews Genetics* 4: 216-224.

Brown, G. K., Nelson, G. & Ladiges, P. Y. (2006). Historical biogeography of *Rhododendron* section *Vireya* and the Malesian archipelago. *Journal of Biogeography* 33: 1929-1944.

Buckler, E. S., Ippolito, A. & Holtsford, T. P. (1997). The evolution of ribosomal DNA: Divergent paralogues and phylogenetic implications. *Genetics* 145(3): 821-832.

Buffon, G. L. L., Compte de (1776). Servant de suite a l'histoire des animaux Quadrupedes (Supplement III) In: *Histoire naturelle generale et particuliere*. Paris: Imprimerie Royal.

Buffon, G. L. L., Compte de (1761). *Histoire naturelle generale*. Paris: Imprimerie Royal.

Burgess, N. D., Clarke, G. P. & Rodgers, W. A. (1998). Coastal forests of eastern Africa: status, endemism patterns and their potential causes. *Biological Journal of the Linnean Society* 64(3): 337-367.

Burnham, R. J. & Johnson, K. R. (2004). South American palaeobotany and the origins of neotropical rainforests. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 359(1450): 1595-1610.

Cabral, C. M. d. S. (2008). Multidisciplinary taxonomic revision in the genus *Vitex* L. (Lamiaceae, Viticoideae) in Africa. Ph.D. thesis, Coimbra: University of Coimbra.

Campbell, J. D. (2002). Angiosperm fruit and leaf fossils from Miocene silcrete, Landslip Hill, northern Southland, New Zealand. *Journal of the Royal Society of New Zealand* 32(1): 149-154.

Cane, M. A. & Molnar, P. (2001). Closing of the Indonesian seaway as a precursor to east African aridircation around 3-4 million years ago. *Nature* 411(6834): 157-162.

Chanderbali, A. S., van der Werff, H. & Renner, S. S. (2001). Phylogeny and historical biogeography of Lauraceae: evidence from the chloroplast and nuclear genomes. *Annals of the Missouri Botanical Garden* 88(1): 104-134.

Chandler, M. E. J. (1954). Some Upper Cretaceous and Eocene fruits from Egypt. *Bulletin of the British Museum [Natural History] Geology* 2(4): 147-187.

Chaney, R. W. & Sanborn, E. I. (1933). *The Goshen flora of west-central Oregon*. Carnegie Institute of Washington 439: 1-103.

Chapman, C. A. (1989). Primate seed dispersal - the fate of dispersed seeds. *Biotropica* 21(2): 148-154.

Chatrou, L. W., Escribano, M. P., Viruel, M. A., Maas, J. W., Richardson, J. E. & Hormaza, J. I. (2009). Flanking regions of monomorphic microsatellite loci provide a new source of data for plant species-level phylogenetics. *Molecular Phylogenetics and Evolution* 53(3): 726-733.

Chauvet, S., Feer, F. & Forget, P. M. (2004). Seed fate of two Sapotaceae species in a Guianan rain forest in the context of escape and satiation hypotheses. *Journal of Tropical Ecology* 20: 1-9.

Clayton, J. W. (2003). Pantropical Genera: Systematics and Biogeography. A pilot study for the pantropical tree *Manilkara*. M.Sc. Thesis, Edinburgh: University of Edinburgh.

Clayton, J. W., Soltis, P. S. & Soltis, D. E. (2009). Recent long-distance dispersal overshadows ancient biogeographical patterns in a pantropical angiosperm family (Simaroubaceae, Sapindales). *Systematic Biology* 58: 1-16.

Coates, A. G. & Obando, J. A. (1996). The geologic evolution of the Central American isthmus. In: J. B. C. Jackson, A. F. Budd, & A. G. Coates (eds.) *Evolution and environment in tropical America*. pp. 21-56. Chicago: University of Chicago Press.

Cockerell, T. D. A. (1908). The fossil flora of Florissant. *Bulletin of the American Museum of Natural History* 24(4): 71-110.

Cockerell, T. D. A. (1925). Plant and insect fossils from the Green River Eocene of Colorado, No. 2556. *Proceedings of the United States National Museum* 66: 1-13.

Cody, S., Richardson, J. E., Rull, V., Ellis, C. & Pennington, R. T. (2010). The great American biotic interchange revisited. *Ecography* 33: 1-7.

Coetzee, J. A. (1993). African Flora since the terminal Jurassic. In: P. Goldblatt (ed.) *Biological Relationships between Africa and South America*. pp. 37-61. Yale University Press.

Colinvaux, P. A., De Oliveira, P. E. & Bush, M. B. (2000). Amazonian and neotropical plant communities on glacial time-scales: The failure of the aridity and refuge hypotheses. *Quaternary Science Reviews* 19(1-5): 141-169.

Collinson, M. E. & Hooker, J. J. (2003). Paleogene vegetation of Eurasia: framework for mammalian faunas. In: J.F.W. Reumer & W. Wessels (eds.) Distribution and migration of Tertiary mammals in Eurasia, A volume in honour of Hans De Bruijn. *Deinsea* 10: 41-83.

Comes, H.P. & Abbott, R.J. (1999). Reticulate evolution in the Mediterranean species complex of *Senecio* sect. *Senecio*: uniting phylogenetic and population-level approaches. In: P. M. Hollingsworth, R.M. Bateman & R.J. Gornall (eds.) *Molecular systematics and plant evolution*. pp. 171-198. London: Taylor & Francis.

Conti, E., Eriksson, T., Shonenberger, J., Sytsma, K. J. & Baum, D. A. (2002). Early Tertiary Out-of-India dispersal of Crypteroniaceae: evidence from phylogeny and molecular dating. *Evolution* 56(10): 1931-1942.

Conti, E., Rutschmann, F., Eriksson, T., Sytsma, K. J. & Baum, D. A. (2004). Calibration of molecular clocks and the biogeographic history of Crypteroniaceae: a reply to Moyle. *Evolution* 58(8): 1874-1876.

Cook, L. G. & Crisp, M. D. (2005a). Directional asymmetry of long-distance dispersal and colonization could mislead reconstructions of biogeography. *Journal of Biogeography* 32(5): 741-754.

Cook, L. G. & Crisp, M. D. (2005b). Not so ancient: the extant crown group of *Nothofagus* represents a post-Gondwanan radiation. *Proceedings of the Royal Society: Biological Sciences* 272(1580): 2535-2544.

Corlett, R. T. (1998). Frugivory and seed dispersal by vertebrates in the Oriental (Indomalayan) Region. *Biological Reviews* 73(4): 413-448.

Corti, G. (2009). Continental rift evolution: From rift initiation to incipient break-up in the Main Ethiopian Rift, East Africa. *Earth-Science Reviews* 96(1-2): 1-53.

Costin, A. B. (1965). Long distance seed dispersal to Macquarie island. *Nature* 206(4981): 317.

Couvreur, T. L. P., Chatrou, L. W., Sosef, M. S. & Richardson, J. E. (2008). Molecular phylogenetics reveal multiple tertiary vicariance origins of the African rain forest trees. *BMC Biology* 6(54).

Cowie, R. H. & Holland, B. S. (2006). Dispersal is fundamental to biogeography and the evolution of biodiversity on oceanic islands. *Journal of Biogeography* 33: 193-198.

Crane, P. R., Friis, E. M. & Pedersen, K. R. (1995). The origin and early diversification of angiosperms. *Nature* 374: 27-33.

Cranston, K. & Rannala, B. (2005). Closing the gap between rocks and clocks. *Heredity* 94(5): 461-462.

Crawley, M. (1989). Dicotyledonous wood from the lower Tertiary of Britain. *Palaeontology* 32: 597-622.

Crawley, M. (2001). Angiosperm Woods from British Lower Cretaceous and Palaeogene Deposits. *Palaeontological Association, Special papers in Palaeontology* 66: 1-100.

Crisp, M., Cook, L. & Steane, D. (2004). Radiation of the Australian flora: what can comparisons of molecular phylogenies across multiple taxa tell us about the evolution of diversity in present-day communities? *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 359(1450): 1551-1571.

Croizat, L. (1952). *Manual of Phytogeography*. The Hague: Dr. W. Junk.

Croizat, L. (1958). *Panbiogeography* (2 volumes). Caracas: published by the author.

Cronk, Q. C. B., Kiehn, M., Wagner, W. L. & Smith, J. F. (2005). Evolution of *Cyrtandra* (Gesneriaceae) in the Pacific Ocean: the origin of a supertramp clade. *American Journal of Botany* 92(6): 1017-1024.

Cronn, R. & Wendel, J. F. (2004). Cryptic trysts, genomic mergers and plant speciation. *New Phytologist* 161(1): 133-142.

Cuenca, A., Asmussen-Lange, C. B. & Borchsenius, F. (2008). A dated phylogeny of the palm tribe Chamaedoreeae supports Eocene dispersal between Africa, North and South America. *Molecular Phylogenetics and Evolution* 46(2): 760-775.

Curiale, J. A., Kyi, P., Collins, I. D., Din, A., Nyein, K., Nyunt, M. & Stuart, C. J. (1994). The central Myanmar (Burma) oil family - composition and implications for source. *Organic Geochemistry* 22(2): 237-255.

Daly, D.C. & Mitchell, J.D. (2000). Lowland vegetation of tropical South America - an overview. In: lentz, D. (ed.) *Imperfect Balance: Landscape Transformations in the pre-Columbian Americas.* pp. 391-454. Columbia University Press, New York.

Darwin, C. (1855). Effect of Salt-Water on the Germination of Seeds. *Gardener's Chronicle and Agricultural Gazette* 47: 773.

Darwin, C. (1857). On the Action of Sea-Water on the Germination of Seeds. *Journal of the Proceedings of the Linnean Society* 1: 130-140.

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.

Datta-Roy, A. & Karanth, K. P. (2009). The Out-of-India hypothesis: What do molecules suggest? *Journal of Biosciences* 34(5): 687-697.

Davis, C. C., Bell, C. D., Fritsch, P. W. & Matthews, S. (2002a). Phylogeny of *Acridocarpus-Brachylophon* (Malpighiaceae): implications for Tertiary tropical floras and Afroasian biogeography. *Evolution* 56(12): 2395-2405.

Davis, C. C., Bell, C. D., Matthews, S. & Donoghue, M. J. (2002b). Laurasian migration explains Gondwanan disjunctions: evidence from Malpighiaceae. *Proceedings of the National Academy of Sciences of the United States of America* 99(10): 6833-6837.

Davis, C. C., Fritsch, P. W., Bell, C. D. & Matthews, S. (2004). High-latitude Tertiary migrations of an exculsively tropical clade: evidence from Malpighiaceae. *International Journal of Plant Sciences* 165(Supplement 4): S107-S121.

Davis, C. C., Fritsch, P. W., Li, J. H. & Donoghue, M. J. (2002c). Phylogeny and biogeography of *Cercis* (Fabaceae): Evidence from nuclear ribosomal ITS and chloroplast ndhF sequence data. *Systematic Botany* 27(2): 289-302.

Dayanandan, S., Ashton, P. S., Williams, S. M. & Primack, R. B. (1999). Phylogeny of the tropical tree family Dipterocarpaceae based on nucleotide sequences of the chloroplast rbcL gene. *American Journal of Botany* 86(8): 1182-1190.

de Candolle, A. P. (1820). *Essai elementaire de geographie botanique*. De l'imprimerie de F.G. Levrault.

de Candolle, A. P. (1844). *Prodromus Systematis Naturalis Regni Vegetabilis*. Paris: Sumptibus Sociorum Treuttel et Würtz.

de Candolle, A. P. (1855). *Geographie Botanique Raisonee* (2 volumes). Paris: Masson.

de Queiroz, A. (2005). The resurrection of oceanic dispersal in historical biogeography. *TRENDS in Ecology and Evolution* 20(2): 68-73.

Degnan, J. H. & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution* 24(6): 332-340.

Demesure, B., Sodzi, N. & Petit, R. J. (1995). A set of universal primers for amplification of polymorphic noncoding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* 4(1): 129-131.

Dengduangboripant, J., Cronk, Q. B. C., Kokubugata, G., & Moeller, M. (2007). Variation and inheritance of nuclear ribosomal DNA clusters in *Streptocarpus* (Gesneriaceae) and their biological and phylogenetic implications. *International Journal of Plant Science*. 168(4): 455-467.

Dettmann, M. E. (1989). Antarctica: Cretaceous cradle of austral temperate rainforests? In: J. A. Crame (ed.) *Origins and Evolution of the Antarctic Biota*. pp. 89-105. Geological Society Special Publication.

Dick, C. W., Abdul-Salim, K. & Bermingham, E. (2003). Molecular sytematic analysis reveals cryptic tertiary diversification of a widespread tropical rain forest tree. *The American Naturalist* 162(6): 691-703.

Dick, C. W., Bermingham, E., Lemes, M. R. & Gribel, R. (2007). Extreme long-distance dispersal of the lowland tropical rainforest tree *Ceiba pentandra* L. (Malvaceae) in Africa and the Neotropics. *Molecular Ecology* 16: 3039-3049.

Dick, C. W. & Wright, S. J. (2005). Tropical mountain cradles of dry forest diversity. *Proceedings of the National Academy of Sciences of the United States of America* 102(31): 10757-10758.

Donoghue, M. J. (2008). A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences of the United States of America* 15(1): 11549-11555.

Donoghue, M. J., Bell, C. D. & Li, J. H. (2001). Phylogenetic patterns in Northern Hemisphere plant geography. *International Journal of Plant Sciences* 162: S41-S52.

Donoghue, M. J. & Moore, B. R. (2003). Toward an integrative historical biogeography. *Integrative and Comparative Biology* 43(2): 261-270.

Donoghue, M. J. & Smith, S. A. (2004). Patterns in the assembly of temperate forests around the Northern Hemisphere. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 359(1450): 1633-1644.

Doyle, J., Biens, P., Doerenkamp, A. & Jardine, S. (1977). Angiosperm pollen from the pre-Albian Cretaceous of equatorial Africa. *Bulletin des Centres de Recherches Exploration-Production Elf-Aquitaine* 1: 451–473.

Doyle, J. & Doyle, J. (1990). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-11.

Doyle, J. A. (2000). Paleobotany, relationships and geographic history of Winteraceae. *Annals of the Missouri Botanical Garden* 87(3): 303-316.

Doyle, J. A., Sauquet, H., Scharaschkin, T. & Le Thomas, A. (2004). Phylogeny, molecular and fossil dating, and biogeographic history of Annonaceae and Myristicaceae (Magnoliales). *International Journal of Plant Sciences* 165(4): S55-S67.

Doyle, J. J. (1992). Gene trees and species trees: Molecular systematics as one-character taxonomy. *Systematic Botany* 17(1): 144-163.

Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4(5): 699-710.

Drummond, A. J., Ho, S. Y. W., Rawlence, N. & Rambaut, A. (2007). *Beast Manual: A Rough Guide to BEAST* 1.4. Available at: http://beast.bio.ed.ac.uk.

Drummond, A. J. & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7: 214.

Duangjai, S., Samuel, R., Munzinger, J., Forest, F., Wallnofer, B., Barfuss, M. H. J., Fischer, G. & Chase, M. W. (2009). A multi-locus plastid phylogenetic analysis of the pantropical genus *Diospyros* (Ebenaceae), with an emphasis on the radiation and biogeographic origins of the New Caledonian endemic species. *Molecular Phylogenetics and Evolution* 52(3): 602-620.

Dubard, M. (1912). Les Sapotacées du groupe des Sideroxylinées. *Annales de l'Institut Botanico-Geologique Colonial de Marseille* 2(10): 1-90.

Dubard, M. (1915). Les Sapotacées du groupe des Sideroxylinées-Mimusopées. *Annales de l'Institut Botanico-Geologique Colonial de Marseille* 3(3): 1-62.

Ducousso, M., Bena, G., Bourgeois, C., Buyck, B., Eyssartier, G., Vincelette, M., Rabevohitra, R., Randrihasipara, L., Dreyfus, B. & Prin, Y. (2004). The last common ancestor of Sarcolaenaceae and Asian dipterocarp trees was ectomycorrhizal before the India-Madagascar separaction, about 88 million years ago. *Molecular Ecology* 13: 231-236.

Edson, G. M., Olsen, D. L. & Petty, A. J. (2000). OCS Report MMS 2000-033. EXXON Lydonia Canyon Block 133 No.1 Well; Geological and Operation Summary, pp. 1-52.

Edwards, S. V. (2009). Is a new and general theory of molecular systematics emerging? *Evolution* 63(1): 1-19.

Elea, S. A. & Nazar, R. N. (1997). Role of the 5.8S rRNA in ribosome translocation. *Nucleic Acids Research* 25(9): 1788-1797.

Engler, A. (1890). *Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie*. Stuttgart: Schweizerbart.

Engler, A. & Prantl, K. (1891). *Die Natürlichen Pflanzenfamilien*. Leipzig: Verlag von Wilhelm Engelmann.

Erkens, R. H. J., Maas, J. W. & Couvreur, T. L. P. (2009). From Africa via Europe to South America: migrational route of a species-rich genus of Neotropical lowland rain forest trees (*Guatteria*, Annonaceae). *Journal of Biogeography* 36(12): 2338-2352.

Estrada, A. & Coates-Estrada, R. (1984). Fruit eating and seed dispersal by howling monkeys (*Alouatta palliata*) in the tropical rainforest of Los Tuxtlas, Mexico. *American Journal of Primatology* 6(2): 77-91.

Fairchild, W. W. & Elsik, W. C. (1969). Characteristic palynomorphs of the Lower Tertiary in the Gulf Coast. *Paleontographica Abteilung B* 128: 81-89.

Farooqui, A., Ray, J. G., Farooqui, S. A., Tiwari, R. K. & Khan, Z. A. (2009). Tropical rainforest vegetation, climate and sea level during the Pleistocene in Kerala, India. *Quarternary International* 213(1-2): 2-11.

Feliner, G. N. & Rossello, J. A. (2007). Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular Phylogenetics and Evolution* 44(2): 911-919.

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

Felsenstein, J. (1985). Confidence-limits on phylogenies - an approach using the bootstrap. *Evolution* 39(4): 783-791.

Fisher, R. L., Johnson, G. L. & Heezen, B. C. (1967). Mascarene Plateau, Western Indian Ocean. *Geological Society of America Bulletin* 78(10): 1247-1266.

Fitch, W. M. (1971). Towards defining the course of evolution: minimum change for a specific tree topology. *Systematic Zoology* 20(4): 406-416.

Forest, F, Savolainen, V., Chase, M.W., Lupia, R., Bruneau, A. & Crane P.R. (2005). Teasing apart molecular-versus fossil-based error estimates when dating phylogenetic trees: A case study in the birch family (Betulaceae). Systematic Botany 30(1): 118-133.

Frederiksen, N. O. (1980a). Mid-Tertiary climate of Southeastern United-States - The sporomorph evidence. *Journal of Paleontology* 54(4): 728-739.

Frederiksen, N. O. (1980b). Paleogene sporomorphs from South-Carolina USA and quantitative correlations with the Gulf Coast. *Palynology* 4: 125-180.

Frederiksen, N. O. (1988). Sporomorph biostratigraphy, floral changes, and paleoclimatology, Eocene and earliest Oligocene of the Eastern Gulf Coast. *United States Geological Survey Professional Papers* 1448: 53-54.

Frederiksen, N. O. (1994). Middle and Late Paleocene angiosperm pollen from Pakistan. *Palynology* 18: 91-137.

Fritsch, P. W., Almeda, F., Renner, S. S., Martins, A. B. & Cruz, B. C. (2004). Phylogeny and circumscription of the near-endemic Brazilian tribe Microlicieae (Melastomataceae). *American Journal of Botany* 91(7): 1105-1114.

Fritsch, P. W., Morton, C. M., Chen, T. & Meldrum, C. (2001). Phylogeny and biogeography of the Styracaceae. *International Journal of Plant Sciences* 162(Supplement 6): S95-S116.

Fukui, A. (2003). Relationship between seed retention time in bird's gut and fruit characteristics. *Ornithological Science* 2: 41-48.

Galley, C., Bytebier, B., Bellstedt, D. U. & Linder, H. P. (2007). The cape element in the Afrotemperate flora: from Cape to Cairo? *Proceedings of the Royal Society B-Biological Sciences* 274(1609): 535-543.

Galley, C. & Linder, P. H. (2006). Geographical affinities of the Cape flora, South Africa. *Journal of Biogeography* 33: 236-250.

Garzione, C. N., Hoke, G. D., Libarkin, J. C., Withers, S., MacFadden, B., Eiler, J., Ghosh, P. & Mulch, A. (2008). Rise of the Andes. *Science* 320(5881): 1304-1307.

Gehrke, B. & Linder, H. P. (2009). The scramble for Africa: pan-temperate elements on the African high mountains. *Proceedings of the Royal Society B-Biological Sciences* 276(1667): 2657-2665.

Gilly, C. L. (1943). Studies in the Sapotaceae II. The Sapodilla-Nispero Complex. *Tropical Woods* 73: 1-22.

Givnish, T. J., Millam, K. C., Evans, T. M., Hall, J. C., Pires, C. J., Berry, P. E. & Sytsma, K. J. (2004). Ancient vicariance or recent long-distance dispersal? Inferences about phylogeny and South American-African disjunctions in Rapateaceae and Bromeliaceae based on *ndh*F sequnce data. *International Journal of Plant Sciences* 65(Supplement 4): S35-S54.

Givnish, T. J. & Renner, S. S. (2004). Tropical intercontinental disjunctions: Gondwana breakup, immigration from the boreotropics, and transoceanic dispersal. *International Journal of Plant Sciences* 165(Supplement 4): S1-S6.

Goldblatt, P. (1993). Biological relationships between Africa and South America: An overview. In: *Biological relationships between Africa and South America - 37th Annual Systematics Symposium, October 4-6, 1990*, pp. 1-14. St. Louis, Missouri, USA: Yale University Press.

Goodall-Copestake, W. P., Harris, D. J. & Hollingsworth, P. M. (2009). The origin of a mega-diverse genus: dating *Begonia* (Begoniaceae) using alternative datasets, calibrations and relaxed clock methods. *Botanical Journal of the Linnean Society* 159: 363-380.

Gottschling, M., Nadja, D., Hilger, H. H. & Weigend, M. (2004). Testing hypotheses on disjunctions present in the primarily woody Boraginales: Ehretiaceae, Cordiaceae and Heliotropiaceae inferred from ITS1 sequence data. *International Journal of Plant Sciences* 165(Supplement 4): S123-S135.

Gottwald, H. P. J. (2004). Neue taxonomische Untersuchungen an 205 tertiären Hölzern und 2 verkieselten Rindenresten aus der Südlichen Frankenalb und deren Randgebieten-mit Aussagen über Herkunft und Flora, Klima und Alter. *Documenta Naturae* 153: 1–93.

Govaerts, R., Frodin, D. G. & Pennington, T. D. (2001). *World checklist and bibliography of Sapotaceae* London: Royal Botanic Gardens, Kew.

Graham, A. (1985). Studies in Neotropical paleobotany. IV. The Eocene communities of Panama. *Annals of the Missouri Botanical Garden* 72(3): 504-534.

Graham, A. (2006). Modern processes and historical factors in the origin of the African element in Latin America. *Annals of the Missouri Botanical Garden* 93(2): 335-339.

Graham, A. (2009). The Andes: a geological overview from a biological perspective. *Annals of the Missouri Botanical Garden* 96(3): 371-385.

Graham, A. & Jarzen, D. M. (1969). Studies in Neotropical paleobotany. I. Oligocene communities in Puerto Rico. *Annals of the Missouri Botanical Garden* 56(3): 308-357.

Graham, A., Stewart, R. H. & Stewart, J. L. (1985). Studies in Neotopical paleobotany. III. The Tertiary communities of Panama - Geology of the pollen-bearing deposits. *Annals of the Missouri Botanical Garden* 72(3): 485-503.

Grambast-Fessard, N. (1968). Contribution a l'étude des flores Tertiares des régions provençales et alpines: IV. – Deux structures ligneuses nouvelles de Sapotacées. *Naturalia monspeliensia, Series Botanique* 19: 57-74.

Graur, D. & Martin, W. (2004). Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *TRENDS in Genetics* 20(2): 80-86.

Green, P. T. (1999). Greta's Garbo: stranded seeds and fruits from Greta Beach, Christmas Island, Indian Ocean. *Journal of Biogeography* 26(5): 937-946.

Gregory-Wodzicki, K. M. (2000). Uplift history of the Central and Northern Andes: A review. *Geological Society of America Bulletin* 112(7): 1091-1105.

Gruas-Cavagnetto, C. (1976). Etude palynologique du Paleogene du sud de l'Angleterre. *Cahiers de Micropaleontologie* 1: 1–49.

Gunn, C. R. & Dennis, J. V. (1976). *World guide to tropical drift seeds and fruits*.

Haffer, J. (1969). Speciation in Amazonian forest birds. *Science* 165(3889): 131-137.

Hall, R. (1998). The plate tectonics of Cenozoic SE Asia and the distribution of land and sea. In: R. Hall & J.D. Holloway (eds.) *Biogeography and geological evolution of SE Asia*. pp. 99-131. Leiden, The Netherlands: Backhuys Publishers.

Hall, R. (2001). Cenozoic reconstructions of SE Asia and the SW Pacific: Changing patterns of land and sea. In: I. Metcalfe, J.M.B. Smith, M. Moorwood & I.D. Davidson (eds.) *Faunal and Floral Migrations and Evolution in SE Asia-Australasia*. pp. 35-56. Lisse: A.A. Balkema (Swets & Zeitlinger Publishers).

Hall, R. (2002). Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer-based reconstructions, model and animations. *Journal of Asian Earth Sciences* 20(4): 353-431.

Hall, R. (2009). Southeast Asia's changing palaeogeography. *Blumea* 54(1-3): 148-161.

Hall, T. (2005). BioEdit version 7.0.5. Carlsbad, California: Ibis Therapeutics. Available at: http://www.mbio.ncsu.edu/bioedit/bioedit.html

Hamilton, A. C. & Taylor, D. (1992). History of climate and forests in tropical Africa during the last 8 million years. In: N. Meyer (ed.) *Tropical Forests and Climate*. pp. 65-78.: Kluwer Academic Publishers.

Hamilton, M. B. (1999). Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology* 8(3): 521-523.

Harbaugh, D. T. & Baldwin, B. G. (2007). Phylogeny and biogeography of the sandalwoods (*Santalum*, Santalaceace): Repeated dispersals throughout the Pacific. *American Journal of Botany* 94: 1028-1040.

Harbaugh, D. T., Wagner, W. L., Allan, G. J. & Zimmer, E. A. (2009). The Hawaiian Archipelago is a stepping stone for dispersal in the Pacific: an example from the plant genus *Melicope* (Rutaceae). *Journal of Biogeography* 36(2): 230-241.

Harley, M. M. (1991). The pollen morphology of the Sapotaceae. *Kew Bulletin* 46(3): 379.

Harley, M. M., Kurmann, M. H. & Ferguson, I. K. (1991). Systematic implications of comparative morphology in selected fossil and extant pollen from the Palmae and the Sapotaceae. In: S. Blackmore & S. H. Barnes (eds.) *Pollen and Spores: Patterns of Diversification*. pp. 225-238. Oxford: Clarendon Press

Harrington, G. J., Kemp, S., J. & Koch, P. L. (2004). Paleocene-Eocene paratropical floral change in North America: responses to climate change and plant immigration. *Journal of the Geological Society, London* 161: 173-184.

Harris, W. K. (1972). New form species of pollen from southern Australian early Tertiary sediments. *Transactions of the Royal Society of South Australia* 96(1): 53-65.

Hartog, M. M. (1878). On the floral structures and affinities of Sapotaceae. *Journal of Botany* 16: 63-72.

Heads, M. (2006). Seed plants of Fiji: an ecological analysis. *Biological Journal of the Linnean Society* 89(3): 407-431.

Hedges, S. B. & Kumar, S. (2004). Precision of molecular time estimates. *TRENDS in Genetics* 20(5): 242-247.

Heer, O. (1876). Beiträge zur fossilen flora von Sumatra. *Allgemeine Schweizerische Gesellschaft fur die gesammten Naturwissenschaften*: 1-22.

Hemsley, J. H. (1966). Notes on African Sapotaceae. *Kew Bulletin* 20: 461.

Hennig, W. (1966). *Phylogenetic Systematics. 3rd edition*. Translated by D.D. Davis and R. Zanderl. Urbana: University of Illinois Press.

Hershkovitz, M. A. & Lewis, L. A. (1996). Deep-level diagnostic value of the rDNA-ITS region. *Molecular Biology and Evolution* 13(9): 1276-1295.

Hickey, L. J. & Doyle, J. A. (1977). Early Cretaceous fossil evidence for angiosperm evolution. *Botanical Review* 43(1): 3-104.

Hnatiuk, S. H. & Rudall, P. (1985). Driftwood genera found on Aldabra Atoll. *Kew Bulletin* 40(3): 471-477.

Ho, S. (2008). The molecular clock and estimating species divergence. *Nature Education* 1(1).

Hofmann, E. (1948). *Manilkaroxylon diluviale* n. sp. ein fossiles Sapotaceen Holz aus dem Quartar von Sta. Paula in Ekuador. *Palaeobiologica* 8: 280-282.

Holder, M. & Lewis, P. O. (2003). Phylogeny estimation: traditional and bayesian approaches. *Nature Reviews Genetics* 4: 275-284.

Hollick, A. (1907). Species novae Pleistocaeniae Marylandicae ab Arthur Hollick descriptae. *Repertorium novarum specierum regni vegetabilis* 4(7-8): 100-104.

Hollick, A. (1924). A review of the fossil flora of the West Indies with descriptions of new species. *New York Botanical Garden Bulletin* 12: 259.

Hollick, A. (1928). Paleobotany of Porto Rico. *New York Academy of Sciences, Scientific Survey of Porto Rico and the Virgin Islands* 7: 177–393.

Hollick, A. (1930). The Upper Cretaceous floras of Alaska. *United States Geological Survey Professional Papers* 156: 123.

Hollick, A. & Berry, E. W. (1924). *A late Tertiary flora from Bahia, Brazil*. The Johns Hopkins Press.

Hollick, C. A. (1899). A report on a collection of fossil plants from northwestern Louisiana. In: G. D. Harris & A. C. Veatch (eds.) *A preliminary report on the geology of Louisiana*. pp. 276-287. Baton Rouge.

Hooghiemstra, H., van der Hammen, T. & Cleef, A. (2002). Evolution of forests in the Northern Andes and Amazonian lowlands during the Tertiary and Quaternary. In: M. R. Guariguata & G. H. Kattan (eds.) *Ecology of Neotropical rainforests*. pp. 43–58. Carthage: Ediciones Libro Universitario Regional.

Hooker, J. D. (1853). *The botany of the antarctic voyage of H.M.S. Discovery ships "Erebus" and "Terror" in the years 1839-1843*. London: Lovell Reeve.

Hooker, J. D. (1867). Lecture on Insular Floras. Delivered before the British Association for the Advancement of Science, August 27, 1866, Nottingham.

HOORN, C., WESSELINGH, F. P., HOVIKOSKI, J. & GUERRERO, J. (2010). The developement of the Amazonian mega-wetland (Miocene; Brazil, Colombia, Peru, Bolivia). In: C. HOORN & F. P. WESSELINGH (eds.) *Amazonia: landscape and species evolution: A look into the past*. pp. 123-142. Chichester: Wiley-Blackwell.

HORN, S. P., SANFORD JR., R. L., DILCHER, D., LOTT, T. A., RENNE, P. R., WIEMANN, M. C., COZADD, D. & VARGAS, O. (2003). Pleistocene plant fossils in and near La Selva Biological Station, Costa Rica. *Biotropica* 35(3): 434-441.

HOULE, A. (1998). Floating islands: a mode of long distance dispersal for small and medium-sized terrestrial vertebrates. *Diversity and Distributions* 4: 201-216.

HOWARTH, D. G., GUSTAFSSON, M. H. G., BAUM, D. A. & MOTLEY, T. J. (2003). Phylogenetics of the genus *Scaevola* (Goodeniaceae): Implication for dispersal patterns across the Pacific Basin and colonization of the Hawaiian Islands. *American Journal of Botany* 90(6): 915-923.

HUELSENBECK, J. P. & RONQUIST, F. (2001). MrBayes. Bayesian Inference of phylogeny. *Bioinformatics* 17: 754-755.

HUELSENBECK, J. P., RONQUIST, F., NIELSEN, R. & BOLLBACK, J. P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294(5550): 2310-2314.

HUGHES, C. & EASTWOOD, R. (2006). Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences of the United States of America* 103(27): 10334-10339.

HUNICKEN, M. (1955). *Depositos Neocretacicos y Terciarios del extremo SSW de Santa Cruz (Cuenca Carbonifera de Rio Turbio)*. Buenos Aires: Ministerio de Educacion de la Nacion Direccion General de Cultura.

HUXLEY, T. H. (1868). On the classification and distribution of the Alectoromorphae and Heteromorphae. *Proceedings of the Zoological Society of London* (6): 294-319.

ICKERT-BOND, S. M., RYDIN, C. & RENNER, S. S. (2009). A fossil-calibrated relaxed clock for *Ephedra* indicates an Oligocene age for the divergence of Asian and New World clades and Miocene dispersal into South America. *Journal of Systematics and Evolution* 47(5): 444-456.

ICKERT-BOND, S. M. & WEN, J. (2006). Phylogeny and biogeography of Altingiaceae: Evidence from combined analysis of five non-coding chloroplast regions. *Molecular Phylogenetics and Evolution* 39(2): 512-528.

IRELAND, H. E., KITE, G. C., VEITCH, N. C., CHASE, M. W., SCHRIRE, B., LAVIN, M., LINARES, J. & PENNINGTON, T. R. (2010). Biogeographical, ecological and morphological structure in a phylogenetic analysis of *Atelia* (Swartzieae, Fabaceae) derived from combined molecular, morphological and chemical data. *Botanical Journal of the Linnean Society* 162: 39-53.

ITURRALDE-VINENT, M. A. & MACPHEE, R. D. E. (1999). Paleogeography of the Caribbean region: Implications for cenozoic biogeography. *Bulletin of the American Museum of Natural History* (238): 1-95.

IVANOV, D. A., ASHRAF, A. R. & MOSBRUGGER, V. (2007). Late Oligocene and Miocene climate and vegetation in the eastern Paratethys area (northeast Bulgaria), based on pollen data. *Palaeogeography, Palaeoclimatology, Palaeoecology* 255: 342-360.

IVANOV, D. A., ASHRAF, A. R., MOSBRUGGER, V. & PALAMAREV, E. (2002). Palynological evidence for Miocene climate change in the Forecarpathian Basin (Central Paratethys, NW Bulgaria). *Palaeogeography, Palaeoclimatology, Palaeoecology* 178: 19-37.

JACOBS, B. F. (2004). Palaeobotanical studies from tropical Africa: relevance to the evolution of forest, woodland and savannah biomes. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 359: 1573-1583.

JACOBS, B. F., TABOR, N., FESEHA, M., PAN, A., KAPPELMAN, J., RASMUSSEN, T., SANDERS, W., WIEMANN, M. C., CRABAUGH, J. & GARCIA MASSINI, J. L. (2005). Oligocene terrestrial strata of northwestern Ethiopia: a preliminary report on paleoenvironments and paleontology. *Palaeontologica Electronica* 8(1): 1-19.

JACQUES, F. M. B., GUO, S. X., SU, T., XING, Y. W., HUANG, Y. J., LIU, Y. S., FERGUSON, D. K. & ZHOU, Z.-K. (2010). Quantitative reconstruction of the Late Miocene monsoon climates of southwest China: A case study of the Lincang flora from Yunnan Province. *Palaeogeography, Palaeoclimatology, Palaeoecology*. In press.

JAN DU CHENE, R. E., ONYIKE, M. S. & SOWUNMI, M. A. (1978). Some new Eocene pollen of the Ogwashi-Asaba formation Southeastern Nigheria. *Revista Espanola de Micropaleontologia* 10(2): 285-322.

JARAMILLO, C. A. (2002). Response of tropical vegetation to Paleogene warming. *Paleobiology* 28(2): 222-243.

JARAMILLO, C. A. & DILCHER, D. L. (2001). Middle Paleogene palynology of Central Colombia, South America: A study of pollen and spores from tropical latitudes. *Palaeontographica Abteilung B Palaeophytologie* 258(4-6): 87-213.

JOHNSON, H. (1991). Petroleum geology of Fiji. *Marine Geology* 98: 2-4.

JUDD, W. S., CAMPBELL, C. S., KELLOGG, E. A., STEVENS, P. F. & DONOGHUE, M. J. (2002). *Plant Systematics: A Phylogenetic Approach, Second Edition*. Sunderland, Massachusetts: Sinauer Associates, Inc.

JULLIOT, C. (1996). Seed dispersal by red howling monkeys (*Alouatta seniculus*) in the tropical rain forest of French Guiana. *International Journal of Primatology* 17(2): 239-258.

JUSSIEU, A. L. (1789). *Genera Plantarum*. Paris: Herissant et Theophilum Barrois.

KAPPELMAN, J., RASMUSSEN, D. T., SANDERS, W. J., FESEHA, M., BOWN, T., COPELAND, P., CRABAUGH, J., FLEAGLE, J., GLANTZ, M., GORDON, A., JACOBS, B., MAGA, M., MULDOON, K., PAN, A., PYNE, L., RICHMOND, B., RYAN, T., SEIFFERT, E. R., SEN, S., TODD, L., WIEMANN, M. C. & WINKLER, A. (2003). Oligocene mammals from Ethiopia and faunal exchange between Afro-Arabia and Eurasia. *Nature* 426(6966): 549-552.

KARANTH, K. P. (2006). Out-of-India Gondwanan origin of some tropical Asian biota. *Current Science* 90(6): 789-792.

KASS, E. & WINK, M. (1997). Molecular phylogeny and phylogeography of *Lupinus* (Leguminosae) inferred from nucleotide sequences of the rbcL gene and ITS 1+2 regions of rDNA. *Plant Systematics and Evolution* 208(3-4): 139-167.

KASS, R. E. & RAFTERY, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association* 90(430): 773-795.

Kay, K. M., Whittall, J. B. & Hodges, S. A. (2006). A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. *BMC Evolutionary Biology* 6: 36.

Kedeves, M. (1971). Presence de types sporomorphs importants dans les sediments prequarternaires Egyptiens *Acta Botanica Academiae Scientarum Hungaricae* 17: 371-378.

Keeley, S. C., Forsman, Z. H. & Chan, R. (2007). A phylogeny of the “evil tribe” (Vernonieae: Compositeae) reveals Old/New World long distance dispersal: support from separate and combined congruent datasets (*trnL*-F, *ndh*F, ITS). *Molecular Phylogenetics and Evolution* 44: 89-103.

Kemp, E. M. & Harris, W. K. (1975). The vegetaton of Tertiary islands on the Ninetyeast Ridge. *Nature* 258: 303-307.

Kimura, M. & Ohta, T. (1971). On the rate of molecular evolution. *Journal of Molecular Evolution* 1(1): 1-17.

Kluge, A. G. (1989). A concern for evidence and a phylogenetic hypothesis of relationships among Epicrates (Boidae, Serpentes). *Systematic Zoology* 38(1): 7-25.

Kluge, A. G. (2004). On total evidence: for the record. *Cladistics-the International Journal of the Willi Hennig Society* 20(2): 205-207.

Knapp, M., Stockler, K., Havell, D., Delsuc, F., Sebastiani, F. & Lockhart, P. J. (2005). Relaxed molecular clock provides evidence for long-distance dispersal of *Nothofagus* (Southern Beech). *PLoS Biology* 3(1): 0038-0043.

Knobloch, E. & Konzalova, M. (1998). Comparison of the Eocene plant assemblages of Bohemia (Czech Republic) and Saxony (Germany). *Review of Palaeobotany and Palynology* 101(1-4): 29-41.

Koeniguer, J.-C. & Faure, H. (1967). Étude paléoxylologique du Sahara méridional: sur la presence de *Faureoxylon princeps* n.g., n.sp. et de *Opilioxylon nigerinum* n.g. n.sp. *Comptes rendus du 92e Congres National Societes Savantes Strasbourg* 3: 143-152.

Kovar-Eder, J., Haas, M., Hofmann, C. C. & Meller, B. (2001). An early Miocene plant assemblage severely influenced by a volcanic eruption, Styria, Austria. *Palaeontology* 44(4): 575-600.

Krausel, R. (1929). Fossile Pflanzen aus dem Tertiar von Sud-Sumatra. Ein weiterer Beitrag zur Kenntnis der fossilen Flora Niederlandisch-Indiens. *Verhandelingen van het Geologisch-Mijnbouwkundig Genootschap voor Nederland en Kolonien. Geologische Serie*: 1-44 + 47 p.

Krausel, R. & Weyland, H. (1959). Kritische untersuchungen zur kutikularanalyse Tertiarer blatter IV: dicotyledonen aus der braunkohle. *Paleontographica Abteilung B* 105(5-6): 101-124.

Kulju, K. K. M., Sierra, S. E. C., Draisma, S. G. A., Samuel, R. & van Welzen, P. C. (2007). Molecular phylogeny of *Macaranga*, *Mallotus*, and related genera (Euphorbiaceae *s.s.*): insights from plastid and nuclear DNA sequence data. *American Journal of Botany* 94(10): 1726-1743.

KUNZ, B. K. & LINSENMAIR, K. E. (2008). The role of the olive baboon (*Papio anubis*, Cercopithecidae) as seed disperser in a savanna-forest mosaic of West Africa. *Journal of Tropical Ecology* 24: 235-246.

LADIGES, P. Y., UDOVICIC, F. & NELSON, G. (2003). Australian biogeographical connections and the phylogeny of large genera in the plant family Myrtaceae. *Journal of Biogeography* 30: 989-998.

LAKHANPAL, R. N. (1970). Tertiary floras of India and their bearing on the historical geology of the region. *Taxon* 19(5): 675-694.

LAKSHAMANAN, S. M. & LEVY, J. F. (1956). Geology and botany of lignite from South Arcot, Madras. *Fuel* 35(4): 446-450.

LAM, H. J. (1939). On the system of the Sapotaceae, with some remarks on taxonomical methods. *Recueil de Travaux Botaniques Néerlandais* 36: 509-525.

LAM, H. J. (1941). Note on the Sapotaceae-Mimusopoideae in general and on the far-eastern *Manilkara* allies in particular. *Blumea* 4(2): 323-358.

LAM, H. J. & VAN ROYEN, P. (1953). Propositions pour le Congrès de Paris: (Proposal no. 19) Proposal for the conservation of the genus *Manilkara* of the Sapotaceae. *Taxon* 2: 112.

LARGET, B. & SIMON, D. L. (1999). Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. *Molecular Biology and Evolution* 16(6): 750-759.

LAVIN, M. (2006). Floristic and geographical stability of discontinuous seasonally dry tropical forests explains patterns of plant phylogeny and endemism. In: R.T. PENNINGTON, G.P. LEWIS & J.A. RATTER (eds.) *Neotropical Savannas and Seasonally Dry Forests: Plant Diversity, Biogeography, and Conservation*. pp. 433-447. Boca Raton: CRC Press.

LAVIN, M. & LUCKOW, M. (1993). Origins and relationships of tropical North America in the context of the boreotropics hypothesis. *American Journal of Botany* 80(1): 1-14.

LAVIN, M., SCHRIRE, B., LEWIS, G., PENNINGTON, T. R., DELGADO-SALINAS, A., THULIN, M., HUGHES, C., MATOS, A. B. & WOJCIECHOWSKI, M. F. (2004). Metacommunity process rather than continental tectonic history better explains geographically structured phylogenies in legumes. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 359: 1509-1522.

LAVIN, M., THULIN, M., LABAT, J. N. & PENNINGTON, R. T. (2000). Africa, the odd man out: Molecular biogeography of dalbergioid legumes (Fabaceae) suggests otherwise. *Systematic Botany* 25(3): 449-467.

LAVIN, M., WOJCIECHOWSKI, M. F., RICHMAN, A., ROTELLA, J., SANDERSON, M. J. & MATOS, A. B. (2001). Identifying tertiary radiations of Fabaceae in the Greater Antilles: Alternatives to cladistic vicariance analysis. *International Journal of Plant Sciences* 162: S53-S76.

LEBRUN, J. P. (2001). *Introduction á la flore d'Afrique*. Paris: Cirad, Ibis Press.

LEBRUN, J. P. & STORK, A. L. (2003). *Tropical African flowering plants: ecology and distribution, volume 1 Annonaceae-Balanitaceae*. Conservatoire et Jardin Botaniques de la Ville de Geneve.

Legoux, O. (1978). Quelques especes de pollen caracteristiques du Neogene du Nigeria. *Bulletin Centre Recherche Exploration - Production Elf Aquitaine* 2: 265-317.

Lelono, E. (2000). Palynological study of the Eocene Nanggulan Formation Central Java, Indonesia. Ph.D. Thesis, London: University of London.

Lemey, P. & Posada, D. (2009). Molecular clock analysis. In: P. Lemey, M. Salemi & A.-M. Vandamme (eds.) *The Phylogenetic Handbook*. pp. 362-377. Cambridge: Cambridge University Press.

Lemoigne, Y. (1978). Flores Tertiares de la haute vallee dr l'Omo (Ethiopie). *Palaeontographica Abteilung B Palaeophytologie* 165: 89-173.

Lemoigne, Y., Beauchamp, J. & Samuel, E. (1974). Étude paléobotanique des dépots volcaniques d'age Tertiare des bordures est et ouest du système des rifts Éthiopiens. *Géobios* 7(3): 267-288 pl. 241-245.

Lepage, T., Bryant, D., Philippe, H. & Lartillot, N. (2007). A general comparison of relaxed molecular clock models. *Molecular Biology and Evolution* 24(12): 2669-2680.

Lesquereux, L. (1883). Contributions to the fossil flora of the western territories part III, the Cretaceous and Tertiary floras, vol. 8. In: J. A. Allen, E. D. Cope, E. Coues, F. V. Hayden, J. Leidy, L. Lesquereux, F. B. Meek, A. S. Packard, S. H. Scudder & C. Thomas (eds.) *Report of the United States Geological Survey of the territories FV Hayden, United States geologist in charge*. Washington: United States Geological Survey of the Territories, Government Printing Office.

Lesquereux, L. C. (1891). *The flora of the Dakota group, a posthumous work by Leo Lesquereux, Edited by F.H. Knowlton*. Washington: United States Geological Survey, Government Printing Office.

Lewis, P. O. (1998). Maximum likelihood as an alternative to parsimony for inferring phylogeny using nucleotide sequence data. In: D. E. Solits, P. S. Soltis & J. J. Doyle (eds.) *Molecular Systematics of Plants II*. pp. 132–163. Boston: Kluwer.

Li, Y. Q., Dressler, S., Zhang, D. X. & Renner, S. S. (2009). More Miocene dispersal between Africa and Asia-the case of *Bridelia* (Phyllanthaceae). *Systematic Botany* 34(3): 521-529.

Linder, H. P. (2003). The radiation of the Cape flora, southern Africa. *Biological Review* 78: 597-638.

Linder, H. P. (2008). Plant species radiations: where, when, why? *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 363(1506): 3097-3105.

Linder, H. P. & Hardy, C. R. (2004). Evolution of the species-rich Cape flora. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 359(1450): 1623-1632.

Linder, H. P., Hardy, C. R. & Rutschmann, F. (2005). Taxon sampling effects in molecular clock dating: An example from the African Restionaceae. *Molecular Phylogenetics and Evolution* 35(3): 569-582.

Linder, P. H., Eldenas, P. & Briggs, B. G. (2003). Contrasting patterns of radiation in African and Australian Restionaceae. *Evolution* 57(12): 2688-2702.

Linnaeus, C. (1753). *Species Plantarum*. Holmiae: Impensis Laurentii Salvii.

Lomolino, M. V. & Heaney, L. R. (eds.) (2004). *Frontiers of biogeography: New directions in the geography of nature*. Sunderland, Massachusetts: Sinauer Associates, Inc.

Lomolino, M. V., Riddle, B. R. & Brown, J. H. (2006). *Biogeography, Third Edition*. Sunderland, Massachusetts: Sinauer Associates, Inc.

Lorente, M. A. (1986). Palynology and palynofacies of the Upper Tertiary in Venezuela. In: *Dissertation Botanicae 99*. Berlin, Stuttgart: J. Cramer.

Lovett, J. & Wasser, S. (1993). *Biogeography and Ecology of the Rainforests of Eastern Africa.* Cambridge: Cambridge University Press.

Lydekker, R. (1896). *A Geographical History of Mammals*. Cambridge: Cambridge University Press.

Lyell, C. (1830). *Principles of Geology, Being an attempt to explain the former changes of the earth's surface, by reference to causes now in operation*. London: John Murray.

Mabberley, D. J. (2008). *Mabberley's plant-book: a portable dictionary of plants, their classifications, and uses*. Cambridge University Press.

Machin, J. (1971). Plant microfossils from Tertiary deposits of the Isle of Wight. *New Phytologist* 70: 851-872.

Macko, S. (1957). Lower Miocene pollen flora from the valley of Klodnica near Gliwice (Upper Silesia). *Travaux des Societe Sci Lettres Wroclaw, Ser B* 88: 1-314.

Maddison, W. P. (1997). Gene trees in species trees. *Systematic Biology* 46(3): 523-536.

Maddison, W. P. & Maddison, D. R. (2008). MacClade version 4.08. Sinauer Associates, Sunderland. Available at: http://macclade.org/macclade.html

Mai, D. H. (1976). Fossile Fruchte und Samen aus dem Mitteleozan des Geiseltales. *Abhandlungen des Zentralen Geologischen Instituts*: 93–150.

Mai, D. H. (1995). *Tertiare Vegetationsgeschichte Europas*. Jena, Germany: Gustav Fischer Verlag.

Mai, D. H. & Walther, H. (1978). *Die Floren der Haselbacher Serie im Weißelster-Becken (Bezirk Leipzig, DDR)*: Abhandlungen des Staatlichen Museums fur Mineralogie und Geologie zu Dresden 28: 1-200.

Makulbekov, N. M. (1987). The evolution and changes in the Paleogene plant assemblages of Kazakhstan, middle Asia and Mongolia. *Paleontological Journal* 21(3): 69-75.

Malcomber, S. T. (2002). Phylogeny of *Gaertnera* Lam. (Rubiaceae) based on multiple DNA markers: Evidence of a rapid radiation in a widespread, morphologically diverse genus. *Evolution* 56(1): 42-57.

Manchester, S. R. (1981). Fossil plants of the Eocene Clarno Nut Beds. *Oregon Geology* 43(6) 74-86.

Manchester, S. R. (1994). Fruits and seeds of the middle eocene nut beds flora Clarno Formation, Oregon. *Palaeontographica Americana* 58(0): 7-200.

Manchester, S. R. (1999). Biogeographical relationships of North American Tertiary floras. *Annals of the Missouri Botanical Garden* 86(2): 472-522.

Manchester, S. R. & Tiffney, B. H. (2001). Integration of paleobotanical and neobotanical data in the assessment of phytogeographic history of Holarctic angiosperm clades. *International Journal of Plant Sciences* 162(Supplement 6): S19-S27.

Manilal, K. S. (2003). *Van Rheede's Hortus Malabaricus: English Edition*. Thirurananthapuram: University of Kerala.

Manos, P. S. (1997). Systematics of *Nothofagus* (Nothofagaceae) based on rDNA spacer sequences (ITS): Taxonomic congruence with morphology and plastid sequences. *American Journal of Botany* 84(8): 1137-1155.

Marshall, C. R. (1990). The fossil record and estimating divergence times between lineages: maximum divergence times and the importance of reliable divergence times. *Journal of Molecular Evolution* 30: 400-408.

Mason, R. (1961). Dispersal of tropical seeds by ocean currents. *Nature* 191(478): 408-409.

Mast, A. R., Willis, C. L., Jones, E. H., Downs, K. M. & Weston, P. H. (2008). A smaller *Macadamia* from a more vagile tribe: Inference of phylogenetic relationships, divergence times, and diaspore evolution in *Macadamia* and relatives (tribe Macadamieae; Proteaceae). *American Journal of Botany* 95(7): 843-870.

Mathews, J.P. (2009). *Chicle: The chewing gum of the Americas, from the ancient Maya to William Wrigley*. Tucson: University of Arizona Press.

Mayle, F. E. (2004). Assessment of the Neotropical dry forest refugia hypothesis in the light of palaeoecological data and vegetation model simulations. *Journal of Quaternary Science* 19(7): 713-720.

McGlone, M. S. (2005). Goodbye Gondwana. *Journal of Biogeography* 32(5): 739-740.

McIntyre, D. J. (1965). Some new pollen species from New Zealand Tertiary deposits. *New Zealand Journal of Botany* 3: 204-214.

McKenna, M. C. (1983). Holarctic landmass rearrangement, cosmic events and Cenozoic terrestrial organisms. *Annals of the Missouri Botanical Garden* 70: 459-489.

McLoughlin, S. (2001). The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Australian Journal of Botany* 49(3): 271-300.

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

Meimberg, H., Wistuba, A., Dittrich, P. & Heubl, G. (2001). Molecular phylogeny of Nepenthaceae based on cladistic analysis of plastid trnK intron sequence data. *Plant Biology* 3: 164-175.

Michalak, I., Zhang, L. B. & Renner, S. S. (2010). Trans-Atlantic, trans-Pacific and trans-Indian Ocean dispersal in the small Gondwanan Laurales family Hernandiaceae. *Journal of Biogeography* 37(7): 1214-1226.

Milne, R. I. (2006). Northern hemisphere plant disjunctions: a window on tertiary land bridges and climate change? *Annals of Botany* 98(3): 465-472.

Milne, R. I. (2009). Effects of taxon sampling on molecular dating for within-genus divergence events, when deep fossils are used for calibration. *Journal of Systematics and Evolution* 47(5): 383-401.

Milne, R. I. & Abbott, R. J. (2002). The origin and evolution of tertiary relict floras. *Advances in Botanical Research* 38: 281-314.

Milne, R. I., Davies, C., Prickett, R., Inns, L. H. & Chamberlain, D. F. (2010). Phylogeny of *Rhododendron* subgenus Hymenanthes based on chloroplast DNA markers: between-lineage hybridisation during adaptive radiation? *Plant Systematics and Evolution* 285(3-4): 233-244.

Möller, M. & Cronk, Q. C. B. (1997). Origin and relationships of *Saintpaulia* (Gesneriaceae) based on ribosomal DNA internal transcribed spacer (ITS) sequences. *American Journal of Botany* 84(7): 956-965.

Moore, B. R. & Donoghue, M. J. (2007). Correlates of diversification in the plant clade Dipsacales: Geographic movement and evolutionary innovations. *American Naturalist* 170: S28-S55.

Mora, A., Baby, P., Roddaz, M., Parra, M., Brusset, S., Hermoza, W. & Espurt, N. (2010). Tectonic history of the Andes and sub-Andean zones: implications for the development of the Amazonian drainage basin. In: C. Hoorn & F. P. Wesselingh (eds.) *Amazonia: landscape and species evolution. A look into the past*. pp. 38-60. Chichester: Wiley-Blackwell.

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

Morley, R. J. (2000). *Origin and evolution of tropical rain forests*. Chichester: John Wiley & Sons, Ltd.

Morley, R. J. (2001). Interplate dispersal path for megathermal angiosperms. *Perspectives in Plant Ecology, Evolution and Systematics* 6: 5-20.

Morley, R. J. (2007). Cretaceous and Tertiary climate change and the past distribution of megathermal rainforests. *Tropical Rainforest Responses to Climatic Change*: 1-31.

Morley, R. J. & Dick, C. W. (2003). Missing fossils, molecular clocks and the origin of the Melastomataceae. *American Journal of Botany* 90(11): 1638-1644.

Morley, R. J. & Richards, K. (1993). Graminae cuticle - a key indicator of late Cenozoic climatic change in the Niger delta. *Review of Palaeobotany and Palynology* 77(1-2): 119-127.

Muellner, A. N., Pannell, C. M., Coleman, A. & Chase, M. W. (2008). The origin and evolution of Indomalesian, Australasian and Pacific island biotas: insights from Aglaieae (Meliaceae, Sapindales) *Journal of Biogeography* 35: 1769-1789.

Muellner, A. N., Pennington, T. D., Koecke, A. V. & Renner, S. S. (2010). Biogeography of *Cedrela* (Meliaceae, Sapindales) in Central and South America. *American Journal of Botany* 97(3): 511-518.

Muller-Stoll, W. R. & Madel-Angeliewa, E. (1984). Fossile Hölzer mit schmalen apotrachealen Parenchymbändern, III. Die Sapotaceae-Gattung *Chrysophylloxylon* gen. nov. *Feddes Repertorium* 95(3): 169-181.

Muller, J. (1968). Palynology of the Pedawan and plateau sandstone formations (Cretaceous-Eocene) in Sarawak, Malaysia. *Micropaleontology* 14(1): 1-37.

Muñoz, J., Felicísimo, A. M., Cabezas, F., Burgaz, A. R. & Martinez, I. (2004). Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science* 304(5674): 1144-1147.

Nathan, R. (2006). Long-distance dispersal of plants. *Science* 313(786): 786-788.

Nathan, R., Schurr, F. M., Spiegel, O., Steinitz, O., Trakhtenbrot, A. & Tsoar, A. (2008). Mechanisms of long-distance seed dispersal. *TRENDS in Ecology & Evolution* 23(11): 638-647.

Navale, G. K. B. (1971). Some contribution to the paleobotany of Neyveli lignite, South India. *Paleobotanist* 20: 179-189.

Near, T. J., Meylan, P. A. & Schaffer, H. B. (2005). Assessing concordance of fossil calibration points in molecular studies: an example using turtles. *The American Naturalist* 165(2): 138-146.

Near, T. J., Meylan, P. A. & Schaffer, H. B. (2008). Caveats on the use of fossil calibration points for molecular dating: a reply to Parham and Irmis. *The American Naturalist* 171(1): 137-140.

Near, T. J. & Sanderson, M. J. (2004). Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 359(1450): 1477-1483.

Nelson, G. (1978). From Candolle to Croizat comments on the history of biogeography. *Journal of the History of Biology* 11(2): 269-305.

Nelson, G. & Rosen, D. E. (1981). *Vicarance biogeography: a critique*. New York Columbia University Press.

Nelson, G. J. (1969). The problem of historical biogeography. *Systematic Zoology* 18: 234-246.

Nelson, G. J. (1974). Historical biogeography: An alternative formalization. *Systematic Zoology* 23: 555-558.

Nie, Z. L., Sun, H., Meng, Y. & Wen, J. (2009). Phylogenetic analysis of *Toxicodendron* (Anacardiaceae) and its biogeographic implications on the evolution of north temperate and tropical intercontinental disjunctions. *Journal of Systematics and Evolution* 47(5): 416-430.

Nylander, J. A. A. (2004). MrModeltest version 2.3. Evolutionary Biology Centre, Uppsala University: Available at: http://darwin.uvigo.es/software/modeltest.html

O'Farrill, G., Calmé, S. & Gonzalez, A. (2006). *Manilkara zapota*: A new record of a species dispersed by tapirs. *Tapir Conservation, The Newsletter of the IUCN/SSC Tapir Specialist Group* 15/1(19): 32-35.

Okuyama, Y., Fujii, N., Wakabayashi, M., Kawakita, A., Ito, M., Watanabe, M., Murakami, N. & Kato, M. (2005). Nonuniform concerted evolution and chloroplast capture: Heterogeneity of observed introgression patterns in three molecular data partition phylogenies of Asian *Mitella* (Saxifragaceae). *Molecular Biology and Evolution* 22(2): 285-296.

Oliveira, A. C. M. & Ferrari, S. F. (2000). Seed dispersal by black-handed tamarins, *Saguinus midas niger* (Callitrichinae, Primates): implications for the regeneration of degraded forest habitats in eastern Amazonia. *Journal of Tropical Ecology* 16: 709-716.

Page, R. D. M. & Holmes, E. C. (1998). *Molecular evolution: A phylogenetic approach*. Blackwell Publishing.

Palamarev, E. (1973). Die Eozane Flora des Burgas-Beckens. *Izvestiya na Geologicheskiya Institut*, *Bulgarska Akademiya na Naukite, Sofiya.* 24: 75-124.

Parham, J. F. & Irmis, R. B. (2008). Caveats on the use of fossil calibrations for molecular dating: a comment on Near et al. *The American Naturalist* 171(1): 132-136.

Parra, M., Mora, A., Jaramillo, C., Strecker, M. R., Sobel, E. R., Quiroz, L., Rueda, M. & Torres, V. (2009). Orogenic wedge advance in the northern Andes: Evidence from the Oligocene-Miocene sedimentary record of the Medina Basin, Eastern Cordillera, Colombia. *Geological Society of America Bulletin* 121(5-6): 780-800.

Peng, D. & Wang, X. Q. (2008). Reticulate evolution in *Thuja* inferred from multiple gene sequences: Implications for the study of biogeographical disjunction between eastern Asia and North America. *Molecular Phylogenetics and Evolution* 47(3): 1190-1202.

Pennington, R. T. & Dick, C. W. (2004). The role of immigrants in the assembly of the South American rainforest tree flora. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 359(1450): 1611-1622.

Pennington, R. T., Lavin, M., Prado, D. E., Pendry, C. A., Pell, S. K. & Butterworth, C. A. (2004). Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 359(1443): 515-537.

Pennington, R. T., Lavin, M., Sarkinen, T., Lewis, G. P., Klitgaard, B. B. & Hughes, C. E. (2010). Contrasting plant diversification histories within the Andean biodiversity hotspot. *Proceedings of the National Academy of Sciences of the United States of America* 107(31): 13783-13787.

Pennington, R. T., Lewis, G. P. & Ratter, J. A. (2006). An overview of the plant diversity, biogeography and conservation of neotropical savannas and seasonally dry forests. In: R.T.

Pennington, G.P. Lewis & J.A. Ratter (eds.) *Neotropical Savannas and Seasonally Dry Forests: Plant Diversity, Biogeography, and Conservation*. pp. 1-29. Boca Raton: CRC Press.

Pennington, R. T., Prado, D. E. & Pendry, C. A. (2000). Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography* 27(2): 261-273.

Pennington, R. T., Richardson, J. E. & Lavin, M. (2006). Insights into the historical construction of species-rich biomes from dated plant phylogenies, neutral ecological theory and phylogenetic community structure. *New Phytologist* 172(4): 605-616.

Pennington, T. D. (1990). *Flora Neotropica Monograph 52. Sapotaceae*. Bronx, New York: New York Botanical Garden.

Pennington, T. D. (1991). *The genera of the Sapotaceae*. Royal Botanic Gardens Kew & New York Botanical Garden.

Percy, D. M., Garver, A. M., Wagner, W. L., James, H. F., Cunningham, C. W., Miller, S. E. & Fleischer, R. C. (2008). Progressive island colonization and ancient origin of Hawaiian *Metrosideros* (Myrtaceae). *Proceedings of the Royal Society B-Biological Sciences* 275(1642): 1479-1490.

Perkins, G. H. (1905). Tertiary lignite of Brandon, Vermont, and its fossils. *Bulletin of the Geological Society of America* 16: 499-516.

Petit, R.J., & Excoffier, L. (2009). Gene flow and species delimitation. *Trends in Ecology and Evolution*. 24(7): 386-393.

Petrescu, J. (1978). Studiul Lemnelor Fosile din Oligocenul din Nord-vestul Transsilvaniei. *Institutul de Geologie si Geofizica, Memorii* 27: 113-183.

Pik, R., Marty, B., Carignan, J., Yirgu, G. & Ayalew, T. (2008). Timing of East African Rift development in southern Ethiopia: Implication for mantle plume activity and evolution of topography. *Geology* 36(2): 167-170.

Pilsneer, P. (1904). La 'Ligne de Weber,' limite zoologique de l'Asie et de L'Australie. *Bulletin de la Classe des Sciences, Académie Royale de Belgique*. 1001-1022.

Pirie, M.D., Chatrou, L.W., Mols, J.B., Erkens, R.H.J. & Oosterhof, J. (2006). 'Andean-centred' genera in the short-branch clade of Annonaceae: testing biogeographical hypotheses using phylogeny reconstruction and molecular dating. Journal of Biogeography 33(1): 31-46.

Plana, V. (2004). Mechanisms and tempo of evolution in the African Guineo-Congolian rainforest. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 359(1450): 1585-1594.

Plana, V. (unpublished). Taxonomic revision of the genus *Manilkara* (Sapotaceae) in Africa.

Plana, V., Gascoigne, A., Forrest, L. L., Harris, D. J. & Pennington, T. R. (2004). Pleistocene and pre-Pleistocene *Begonia* speciation in Africa. *Molecular Phylogenetics and Evolution* 31: 449–461.

Platnick, N. I. & Nelson, G. (1978). A method of analysis for historical biogeography. *Systematic Zoology* 27(1): 1-16.

Playford, G. (1982). Neogene palynomorphs from the Huon peninsula, Papua New Guinea. *Palynology* 6: 29-54.

Poczai, P. & Hyvonen, J. (2010). Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects. *Molecular Biology Reports* 37(4): 1897-1912.

Potbury, S. S. (1935). *The LaPorte flora of Plumas County, California*. Carnegie Institute of Washington Publication Number 465: 29-81.

Prakash, U. & Awasthi, N. (1970). Fossil woods from the Tertiary of Eastern India. *Paleobotanist* 18: 32-44.

Prakash, U. & Awasthi, N. (1982). Fossil dicotyledonous woods from the Tertiary of Blue Nile Valley, Ethiopia. *Paleobotanist* 30(1): 43-59.

Prakash, U., Brezinova, D. & Awasthi, N. (1974). Fossil woods from the Tertiary of South Bohemia. *Paleontographica Abteilung B* 147: 107-133.

Prakash, U. & Tripathi, P. P. (1977). Fossil woods of *Ougenia* and *Madhuca* from the Tertiary of Assam *Paleobotanist* 24: 140-145.

Prasad, V., Farooqui, A., Tripathi, S. K. M., Garg, R. & Thakur, B. (2009). Evidence of Late Palaeocene-Early Eocene equatorial rain forest refugia in southern Western Ghats, India. *Journal of Biosciences* 34(5): 777-797.

Price, J. P. & Clague, D. A. (2002). How old is the Hawaiian biota? Geology and phylogeny suggest recent divergence. *Proceedings of the Royal Society: Biological Sciences* 269: 2429-2435.

Proctor, V. W. (1968). Long-distance dispersal of seeds by retention in digestive tracts of birds. *Science* 160(3825): 321-322.

Pubellier, M., Ali, J. & Monnier, C. (2003). Cenozoic plate interaction of the Australia and Philippine Sea Plates: “hit-and-run” tectonics. *Tectonophysics* 363(3-4): 181-199.

Quadros, J. & Monteiro-Filho, E. L. A. (2000). Fruit occurrence in the diet of the Neotropical otter, *Lontra longicaudis*, in southern Brazilian Atlantic forest and its implication for seed dispersion. *Mastozoología Neotropical* 7(1): 33-36.

Radlkofer, L. (1888). In: Durand (ed.) *Index Generum Phanerogamarum*. Brussels: Sumptibus auctoris.

Rage, J. C. & Jaeger, J. J. (1995). The sinking Indian raft - A response to Thewissen and McKenna. *Systematic Biology* 44(2): 260-264.

Ramanujam, c. G. K. (1996). Floristic diversity in the Neogene and modern forests in Pondicherry-Neyveli area of Tamil Nadu-its impact on climatic and phytogeographic aspects. *Rheedea* 6(1): 29-41.

Rambaut, A. & Drummond, A. J. (2009). Tracer v.1.5. Available at: http://tree.bio.ed.ac.uk/software/tracer/.

Rauscher, J. T. (2002). Molecular phylogenetics of the *Espeletia* complex (Asteraceae): Evidence from nrDNA its sequences on the closest relatives of an Andean adaptive radiation. *American Journal of Botany* 89(7): 1074-1084.

Raven, P. H. & Axelrod, D. I. (1974). Angiosperm biogeography and past continental movements. *Annals of the Missouri Botanical Garden* 61(3): 539-673.

Rawat, M. S., Mukherjee, J. & Venkatachala, B. S. (1977). Palynology of the Kadi formation, Cambray basin, India. *Proceedings of the 4th Colloquium on Indian Micropaleontology and Stratigraphy*: 179-192.

Raymond, A., Phillips, M. K., Gennett, J. A. & Comet, P. A. (1997). Palynology and paleoecology of lignites from the Manning Formation (Jackson Group) outcrop in the Lake Somerville spillway of east-central Texas. *International Journal of Coal Geology* 34(3-4): 195-223.

Ree, R. H., Moore, B. R., Webb, C. O. & Donoghue, M. J. (2005). A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59(11): 2299-2311.

Ree, R. H. & Sanmartin, I. (2009). Prospects and challenges for parametric models in historical biogeographical inference. *Journal of Biogeography* 36(7): 1211-1220.

Ree, R. H. & Smith, S. A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* 57(1): 4-14.

Reid, E. M. & Chandler, M. E. J. (1933). *London Clay Flora*. London: British Museum (Natural History).

Renner, S. S. (2004a). Bayesian analysis of combined chloroplast loci, using multiple calibrations, supports the recent arrival of Melastomataceae in Africa and Madagascar. *American Journal of Botany* 91(9): 1427-1435.

Renner, S. S. (2004b). Multiple Miocene Melastomataceae dispersal between Madagascar, Africa and India. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 359: 1485-1494.

Renner, S. S. (2004c). Plant dispersal across the tropical Atlantic by wind and sea currents. *International Journal of Plant Sciences* 165(Supplement 4): S23-S33.

Renner, S. S. (2005). Relaxed molecular clocks for dating historical plant dispersal events. *Trends in Plant Science* 10(11): 550-558.

Renner, S. S. (2010). Biogeographic insights from a short-lived Palaeocene island in the Ninetyeast Ridge. *Journal of Biogeography* 37(7): 1177-1178.

Renner, S. S., Clausing, G. & Meyer, K. (2001). Historical biogeography of Melastomataceae: the roles of Tertiary migration and long-distance dispersal. *American Journal of Botany* 88(7): 1290-1300.

Renner, S. S., Foreman, D. B. & Murray, D. (2000). Timing transantarctic disjunctions in the Atherospermataceae (Laurales): evidence from coding and noncoding chloroplast sequences. *Systematic Biology* 49(3): 579-591.

Renner, S. S., Schaefer, H. & Kocyan, A. (2007). Phylogenetics of *Cucumis* (Cucurbitaceae): Cucumber (*C. sativus*) belongs in an Asian/Australian clade far from melon (*C. melo*). *BMC Evolutionary Biology* 7: 58

Renner, S. S., Strijk, J. S., Strasberg, D. & Thebaud, C. (2010). Biogeography of the Monimiaceae (Laurales): a role for East Gondwana and long-distance dispersal, but not West Gondwana. *Journal of Biogeography* 37(7): 1227-1238.

RENNER, S. S., ZHANG, L.-B. & MURATA, J. (2004). A chloroplast phylogeny of *Arisaema* (Araceae) illustrates Tertiary floristic links between Asia, North America and East Africa. *American Journal of Botany* 91(6): 881-888.

RICHARDS, G. C. (1990). The spectacled flying fox *Pteropus conspicillatus* Chiroptera Pteropodidae in North Queensland and Australia 2. Diet, seed dispersal and feeding ecology. *Australian Mammalogy* 13(1-2): 25-32.

RICHARDSON, J. E., CHATROU, L. W., MOLS, J. B., ERKENS, R. H. J. & PIRIE, M. D. (2004). Historical biogeography of two cosomopolitan families of flowering plants: Annonaceae and Rhamnaceae. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 359: 1495-1508.

RICHARDSON, J. E., COSTION, C. M. & MUELLNER, A. N. (2010). The Malesian Floristic Interchange: Continental Collision, Dispersal and Phylogenetic Niche Conservatism. In: *Proceedings of the SAGE 2009 - Southeast Asia Gateway Evolution Conference*. Royal Holloway, University of London. (In press).

RICHARDSON, J. E., PENNINGTON, R. T., PENNINGTON, T. D. & HOLLINGSWORTH, P. M. (2001a). Rapid diversification of a species-rich genus of neotropical rain forest trees. *Science* 293(5538): 2242-2245.

RICHARDSON, J. E., WEITZ, F. M., FAY, M. F., CRONK, Q. C. B., LINDER, H. P., REEVES, G. & CHASE, M. W. (2001b). Phylogenetic analysis of *Phylica* L. (Rhamnaceae) with an emphasis on island species: evidence from plastid trnL-F and nuclear internal transcribed spacer (ribosomal) DNA sequences. *Taxon* 50(2): 405-427.

RICHARDSON, J. E., WEITZ, F. M., FAY, M. F., CRONK, Q. C. B., LINDER, H. P., REEVES, G. & CHASE, M. W. (2001c). Rapid and recent origin of species richness in the Cape flora of South Africa. *Nature* 412(6843): 181-183.

RODD, J. A. The petroleum potential of Fiji In: http://www.mrd.gov.fj/gfiji/petroleum/petroleum.html.

RODDA, P. (1976). Geology of northern and central Viti Levu. *Bulletin, Mineral Resources Division Fiji, volume 3.* Suva, Fiji: Ministry of Lands and Mineral Resources.

RODDA, P. (1994). Geology of Fiji. In: A. J. STEVENSON, R. H. HERZER & P. F. BALLANCE (eds.) *Contributions to the marine and on-land geology and resources of the Tonga-Lau-Fiji region* pp. 131-151. SOPAC Technical Bulletin.

RONQUIST, F., HUELSENBECK, J. P. & VAN DER MARK, P. (2005). MrBayes 3.1. Available at: http://mrbayes.csit.fsu.edu/

RONQUIST, F., VAN DER MARK, P. & HUELSENBECK, J. P. (2009). Bayesian phylogenetic analysis using MrBayes. In: P. LEMEY, M. SALEMI & A.-M. VANDAMME (eds.) *The phylogenetic handbook: A practical approach to phylogenetic analysis and hypothesis testing*. pp. 210-266. Cambridge: Cambridge University Press.

RULL, V. (2000). Ecostratigraphic study of Paleocene and early Eocene palynological cyclicity in northern South America. *Palaios* 15(1): 14-24.

RUTSCHMANN, F. (2006). Molecular dating of phylogenetic trees: A brief review of current methods that estimate divergence times. *Diversity and Distributions* 12(1): 35-48.

RUTSCHMANN, F., ERIKSSON, T., SHONENBERGER, J. & CONTI, E. (2004). Did Crypteroniaceae really disperse out of India? Molecular dating evidence from *rbc*L, *ndh*F and *rpl*16 intron sequences. *International Journal of Plant Sciences* 165(Supplement 4): S69-S83.

RYDIN, C., PEDERSEN, K. R. & FRIIS, E. M. (2004). On the evolutionary history of *Ephedra*: Cretaceous fossils and extant molecules. *Proceedings of the National Academy of Sciences of the United States of America* 101(47): 16571-16576.

SAH, S. C. D. (1967). Palynology of an upper Neogene profile from Rusizi valley (Burundi). *Annales Musée Royale Afrique Centrale Belgique Ser IN-8, Sciences Géologiques* 57: 1-175.

SALARD-CHEBOLDAEFF, M. (1978). Sur la palynoflore Maestrichtienne et Tertiaire du bassin sédimentaire littoral du Cameroun. *Pollen et Spores* 20: 215-260.

SALARD-CHEBOLDAEFF, M. (1979). Palynologie Maestrichtienne et Tertiaire du Cameroun. Etude qualitative et repartition verticale des principales especes. *Review of Palaeobotany and Palynology* 28(3-4): 365-388.

SALARD-CHEBOLDAEFF, M. (1981). Palynologie Maestrichtienne et Tertiaire du Cameroun. Resultats botaniques. *Review of Palaeobotany and Palynology* 32: 401-439.

SALEH, M. N. (2006). Taxonomic revision and molecular studies of *Garcinia* section Garcinia (Guttiferae). Ph.D. Thesis, Edinburgh: University of Edinburgh.

SANDERSON, M. J. (1997). A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* 14(12): 1218-1231.

SANDERSON, M. J. (2002). Estimating absolute rates of molecular evolution and divegence times: a penalized likelihood approach. *Molecular Biology and Evolution* 19(1): 101-109.

SANDERSON, M. J. (2003). r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 19(2): 301-302.

SANG, T., CRAWFORD, D. J. & STUESSY, T. F. (1997). Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* 84(8): 1120-1136.

SANMARTIN, I. (2002). A paleogeographic history of the southern hemisphere. Unpublished manuscript.

SANMARTIN, I. & RONQUIST, F. (2004). Southern Hemisphere biogeography inferred by event-based models: Plant versus animal patterns. *Systematic Biology* 53(2): 216-243.

SARKINEN, T. (2010). Historical assembly of seasonally dry tropical forest diversity in the tropical Andes. Ph.D. thesis, Oxford: University of Oxford.

SARKINEN, T. E., NEWMAN, M. F., MAAS, P. J. M., MAAS, H., POULSEN, A. D., HARRIS, D. J., RICHARDSON, J. E., CLARK, A., HOLLINGSWORTH, M. & PENNINGTON, T. R. (2007). Recent oceanic long-distance dispersal and divergence in the amphi-Atlantic rain forest genus *Renealmia* L.f. (Zingiberaceae). *Molecular Phylogenetics and Evolution* 44: 968-980.

SAVARD, L., MICHAUD, M. & BOUSQUET, J. (1993). Genetic diversity and phylogenetic relationships between birches and alders using ITS, 18S rRNA, and rbcL gene sequences. *Molecular Phylogenetics and Evolution* 2(2): 112-118.

Schaefer, H., Heibl, C. & Renner, S. S. (2008). Gourds afloat: a dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proceedings of the Royal Society: Biological Sciences* 276: 843-851.

Schatz, G. E. (1996). Malagasy/Indo-Australo-Malesian phytogeographic connections. *Biogeography and Madagascar*: 73-83.

Scherson, R. A., Vidal, R. & Sanderson, M. J. (2008). Phylogeny, biogeography, and rates of diversification of New World *Astragalus* (Leguminosae) with an emphasis on South American radiations. *American Journal of Botany* 95(8): 1030-1039.

Schrenk, J. (1888). Schweinfurth's method of preserving plants for herbaria. *Bulletin of the Torrey Botanical Club* 15: 292-293.

Scott, R. A. (1954). Fossil fruits and seeds from the Eocene Clarno Formation of Oregon. *Palaeontographica* 96(B): 66-97.

Seitner, L. (1987). Miozane mikrofloren aus sedimenten der sussbrackwassermolasse und der oberen susswassermolasse suddeutschlands. Ph.D thesis, Munchen: Ludwig Maximilians Universitat.

Selmeier, A. (1991). Ein verkieseltes Sapotaceae-Holz, *Bumelioxylon holleisii* n. gen., n. sp., aus jungtertiaren Schichten der Sudlichen Frankenalb (Bayern). *Archaeopteryx* 9: 55-72.

Sepulchre, P., Ramstein, G., Fluteau, F., Schuster, M., Tiercelin, J. J. & Brunet, M. (2006). Tectonic uplift and Eastern Africa aridification. *Science* 313(5792): 1419-1423.

Sepulchre, P., Sloan, L. C. & Fluteau, F. (2010). Modelling the response of Amazonian climate to the uplift of the Andean mountain range. In: C. Hoorn & F. P. Wesselingh (eds.) *Amazonia: landscape and species evolution A look into the past*. pp. 211-222. Chichester: Wiley-Blackwell.

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W. S., Miller, J., Siripun, K. C., Winder, C. T., Schilling, E. E. & Small, R. L. (2005). The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* 92(1): 142-166.

Shaw, J., Lickey, E. B., Schilling, E. E. & 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(3): 275-288.

Shields, R. (2004). Pushing the envelope on molecular dating. *TRENDS in Genetics* 20(5): 221-222.

Shilin, P. V. (2000). Cretaceous flora of Tyul'kili (northeastern Priaral'e) *Paleontological Journal* 34 (2): 211-220.

Simmons, M. P. & Ochoterena, H. (2000). Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* 49(2): 369-381.

Simon, M. F., Grether, R., de Queiroz, L. P., Skema, C., Pennington, R. T. & Hughes, C. E. (2009). Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences of the United States of America* 106(48): 20359-20364.

Simpson, G. G. (1977). Too many lines - Limits of Oriental and Australian zoogeographic regions. *Proceedings of the American Philosophical Society* 121(2): 107-120.

Skema, C. (2003). Phylogeny and biogeography of *Andira*. A case study in Neotropical speciation. M.Sc. Thesis, Edinburgh: University of Edinburgh.

Small, R. L., Cronn, R. C. & Wendel, J. F. (2004). Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* 17(2): 145-170.

Small, R. L., Ryburn, J. A., Cronn, R. C., Seelanan, T. & Wendel, J. F. (1998). The tortoise and the hare: Choosing between noncoding plastome and nuclear ADH sequences for phylogeny reconstruction in a recently diverged plant group. *American Journal of Botany* 85(9): 1301-1315.

Smedmark, J. E. E. & Anderberg, A. A. (2007). Boreotropical migration explains hybridization between geographically distant lineages in the pantropical clade Sideroxyleae (Sapotaceae). *American Journal of Botany* 94: 1491-1505.

Smedmark, J. E. E., Swenson, U. & Anderberg, A. A. (2006). Accounting for variation of substitution rates through time in Bayesian phylogeny reconstruction of Sapotoideae (Sapotaceae). *Molecular Phylogenetics and Evolution* 39(3): 706-721.

Smith, A. B. & Peterson, K. J. (2002). Dating the time of origin of major clades: molecular clocks and the fossil record. *Annual Review of Earth and Planetary Sciences* 30: 65-88.

Smith, J. F., Stevens, A. C., Tepe, E. J. & Davidson, C. (2008). Placing the origin of two species-rich genera in the late Cretaceous with later species divergence in the Tertiary: a phylogenetic, biogeographic and molecular dating analysis of *Piper* and *Peperomia* (Piperaceae). *Plant Systematics and Evolution* 275(1-2): 9-30.

Smith, J. M. B. (1990). Drift disseminules on Fijian beaches. *New Zealand Journal of Botany* 28(1): 13-20.

Smith, S. D. & Baum, D. A. (2006). Phylogenetics of the florally diverse Andean clade Iochrominae (Solanaceae). *American Journal of Botany* 93(8): 1140-1153.

Snow, D. W. (1981). Tropical frugivorous birds and their food plants - a world survey. *Biotropica* 13(1): 1-14.

Soejima, A., Wen, J., Zapata, M. & Dillon, M. O. (2008). Phylogeny and putative hybridization in the subtribe Paranepheliinae (Liabeae, Asteraceae), implications for classification, biogeography, and Andean orogeny. *Journal of Systematics and Evolution* 46(3): 375-390.

Soltis, D. E., Johnson, L. A. & Looney, C. (1996). Discordance between ITS and chloroplast topologies in the *Boykinia* group (Saxifragaceae). *Systematic Botany* 21(2): 169-185.

Soltis, D. E. & Kuzoff, R. K. (1995). Discordance between nuclear and chloroplast phylogenies in the *Heuchera* group (Saxifragaceae). *Evolution* 49(4): 727-742.

Soltis, D. E. & Soltis, P. S. (1998). Choosing an approach and an appropriate gene for phylogenetic analysis. In: *Molecular systematics of plants, II: DNA sequencing*. pp. 1-42. Kluwer Academic Publishers.

Soltis, D. E., Soltis, P. S., Collier, T. G. & Edgerton, M. L. (1991). Chloroplast DNA variation within and among genera of the *Heuchera* group (Saxifragaceae) - Evidence for chloroplast transfer and paraphyly. *American Journal of Botany* 78(8): 1091-1112.

Soltis, P. S., Soltis, D. E., Savolainen, V., Crane, P. R. & Barraclough, T. G. (2002). Rate heterogeneity among lineages of tracheophytes: Integration of molecular and fossil data and evidence for molecular living fossils. *Proceedings of the National Academy of Sciences of the United States of America* 99(7): 4430-4435.

Song, Z. C., Li, M., Wang, W., Zhao, C. & Zhu, Z. (1999). *Fossil spores and pollen of China: The Late Cretaceous and Tertiary spores and pollen*. Science Press.

Song, Z. C., Wang, W. M. & Huang, F. (2004). Fossil pollen records of extant angiosperms in China. *Botanical Review* 70(4): 425-458.

Spennemann, D. H. R. (1997). On the origin of drift materials in the Marshall Islands. In: *Atoll Research Bulletin*. pp. 1-8.

Stebbins, G. L. (1974). *Flowering plants: evolution above the species level*. Cambridge, MA: Harvard University Press.

Stoian, L. M. (2002). *Palynological analyses and dating of drillhole McNamara 1, Gambier–Otway Basins, South Australia*. Primary Industries and Resources South Australia.

Stover, L. E. & Partridge, A. D. (1973). Tertiary and Late Cretaceous spores and pollen from the Gippsland Basin Southeastern Australia. *Proceedings of the Royal Society of Victoria* 85(2): 237-286.

Strong, C. C. & Skolmen, R. G. (1963). Origin of drift-logs on the beaches of Hawaii. *Nature* 197: 890.

Su, Y. C. F. & Saunders, R. M. K. (2009). Evolutionary divergence times in the Annonaceae: evidence of a late Miocene origin of *Pseuduvaria* in Sundaland with subsequent diversification in New Guinea. *BMC Evolutionary Biology* 9(153).

Swenson, U. & Anderberg, A. A. (2005). Phylogeny, character evolution, and classification of Sapotaceae (Ericales). *Cladistics* 21(2): 101-130.

Swenson, U., Bartish, I. V. & Munzinger, J. (2007). Phylogeny, diagnostic characters and generic limitation of Australasian Chrysophylloideae (Sapotaceae, Ericales): evidence from ITS sequence data and morphology. *Cladistics* 23(3): 201-228.

Swenson, U., Hill, R. S. & McLoughlin, S. (2001). Biogeography of *Nothofagus* supports the sequence of Gondwana break-up. *Taxon* 50(4): 1025-1041.

Swenson, U., Lowry, P. P., Munzinger, J., Rydin, C. & Bartish, I. V. (2008). Phylogeny and generic limits in the *Niemeyera* complex of New Caledonian Sapotaceae: evidence of multiple origins of the anisomerous flower. *Molecular Phylogenetics and Evolution* 49(3): 909-929.

Swenson, U., Munzinger, J. & Bartish, I. V. (2007b). Molecular phylogeny of *Planchonella* (Sapotaceae) and eight new species from New Caledonia. *Taxon* 56(2): 329-354.

Swenson, U., Richardson, J. E. & Bartish, I. V. (2008b). Multi-gene phylogeny of the pantropical subfamily Chrysophylloideae (Sapotaceae): evidence of generic polyphyly and extensive morphological homoplasy. *Cladistics* 24(6): 1006-1031.

Swofford, D. L. (1993). PAUP - a computer program for phylogenetic inference using maximum parsimony. *Journal of General Physiology* 102(6): A9-A9.

Swofford, D. L. (2003). PAUP* 4.0b10: Phylogenetic Analysis Using Parsimony (* and other methods) In: Sunderland, Massachusetts: Sinnauer Associates.

Sytsma, K. J., Litt, A., Zjhra, M. L., Pires, C. J., Nepokroeff, M., Conti, E., Walker, J. & Wilson, P. G. (2004). Clades, clocks and continents: historical and biogeographical analysis of Myrtaceae, Vochysiaceae, and relatives in the southern hemisphere. *International Journal of Plant Sciences* 165(Supplement 4): S85-S105.

Taylor, D. J. & Piel, W. H. (2004). An assessment of accuracy, error, and conflict with support values from genome-scale phylogenetic data. *Molecular Biology and Evolution* 21(8): 1534-1537.

Taylor, D. W. (1989). Select palynomorphs from the Midde Eocene Caliborne Formation, Tenn, (USA). *Review of Palaeobotany and Palynology* 58(2-4): 111-128.

Thomas, W. W. (1999). Conservation and monographic research on the flora of tropical America. *Biodiversity and Conservation* 8: 1007–1015.

Thomson, P. W. & Pflug, H. (1953). Pollen and spores from the middle Tertiary of Europe. *Palaeontographica* 94 B.: 1-138.

Thorne, J. L., Kishino, H. & Painter, I. S. (1998). Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* 15(12): 1647-1657.

Tiffney, B. H. (1985a). The Eocene North-Atlantic land-bridge - its importance in Tertiary and modern phytogeography of the northern hemisphere *Journal of the Arnold Arboretum* 66(2): 243-273.

Tiffney, B. H. (1985b). Perspectives on the origin of the floristic similarity between eastern Asia and Eastern North America. *Journal of the Arnold Arboretum* 66(1): 73-94.

Tiffney, B. H. (1994). Re-evaluation of the age of the Brandon Lignite (Vermont, USA) based on plant-megafossils. *Review of Palaeobotany and Palynology* 82(3-4): 299-315.

Tiffney, B. H. & Manchester, S. R. (2001). The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the Northern Hemisphere tertiary. *International Journal of Plant Sciences* 162: S3-S17.

Tosh, J. (2009). Biogeography of Madagascan Rubiaceae: Case studies from the Afro-Madagascan genus *Tricalysia* and the pantropical genus *Ixora*. Ph.D. Thesis, Lueven: K.U. Lueven.

Traverse, A. (1955). *Pollen analysis of the Brandon Lignite of Vermont*. Report of Investigations 5151: United States Department of the Interior, Bureau of Mines.

Traverse, A. (1994). Palynofloral geochronology of the Brandon Lignite of Vermont, USA. *Review of Palaeobotany and Palynology* 82(3-4): 265-297.

Traverse, A. & Barghoorn, E. S. (1953). Micropaleontology of the Brandon lignite, an early tertiary coal in central Vermont. *Journal of Paleontology* 27(2): 289-293.

Triono, T., Brown, A. H. D., West, J. G. & Crisp, M. D. (2007). A phylogeny of *Pouteria* (Sapotaceae) from Malesia and Australasia. *Australian Systematic Botany* 20(2): 107-118.

Tsitrone, A., Kirkpatrick, M. & Levin, D. A. (2003). A model for chloroplast capture. *Evolution* 57(8): 1776-1782.

Turner, H., Hovenkamp, P. & van Welzen, P. C. (2001). Biogeography of Southeast Asia and the West Pacific. *Journal of Biogeography* 28(2): 217-230.

Twyford, A. D. & Richardson, J. E. (In prep). Patterns in the colonization of the Malay Archipelago by *Rhododendron* subgenus Vireya. *Molecular Phylogenetics and Evolution*.

Upchurch, P. (2008). Gondwanan break-up: legacies of a lost world? *Trends in Ecology & Evolution* 23(4): 229-236.

Uriarte, M., Canham, C. D., Thompson, J., Zimmerman, J. K. & Brokaw, N. (2005). Seedling recruitment in a hurricane-driven tropical forest: light limitation, density-dependence and the spatial distribution of parent trees. *Journal of Ecology* 93(2): 291-304.

van Aarssen, B. G. K., Cox, H. C., Hoogendoorn, P. & De Leeuw, J. W. (1990). A cadinene biopolymer in fossil and extant dammar resins as a source for cadinanes and bicadinanes in crude oils from Southeast Asia. *Geochimica et Cosmochimica Acta* 54(11): 3021-3032.

van Rheede tot Drakestein, H. (1678). *Hortus Indicus Malabaricus*. Amsterdam: sumptibus Johannis van Someren, et Joannis van Dyck.

van Royen, P. (1953). Revision of the Sapotaceae of the Malaysian area in a wider sense, volume 1 *Manilkara* Adanson em. Gilly in the Far East. *Blumea* 7: 401-412.

van Ufford, A. Q. & Cloos, M. (2005). Cenozoic tectonics of New Guinea. *American Association of Petroleum Geologists Bulletin* 89(1): 119-140.

van Welzen, P. C., Slik, J. W. F. & Alahuhta, J. (2005). Plant distribution patterns and plate tectonics in Malesia. *Kongelige Danske Videnskabernes Selskab Biologiske Skrifter* 55: 199-217.

Vandamme, A.-M. (2009). Basic concepts of molecular evolution. In: P. Lemey, M. Salemi & A.-M. Vandamme (eds.) *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. pp. 3-29. Cambridge University Press.

Vanderhammen, T. (1991). Paleoecological background - Neotropics. *Climatic Change* 19(1-2): 37-47.

Venkatachala, B. S. (1974). Palynological zonation of the Mesozoic and Tertiary sub-surface sediments in the Cauvery basin. In: R. R. *e. a.* Surange (ed.) *Aspects and Appraisal of Indian Palaeobotany*. Lucknow: Birbal Sahni Institute of Palaeobotany.

Venkatachala, B. S. & Rawat, M. S. (1972). *Palynology of the Tertiary sediments in the Cauvery basin. I. Palaeogene - Eocene palynoflora from the sub-surface*. University of Calcutta.

Verboom, G. A., Archibald, J. K., Bakker, F. T., Bellstedt, D. U., Conrad, F., Dreyer, L. L., Forest, F., Galley, C., Goldblatt, P., Henning, J. F., Mummenhoff, K., Linder, H. P., Muasya, A. M., Oberlander, K. C., Savolainen, V., Snijman, D. A., van der Niet, T. & Nowell, T. L. (2009). Origin and diversification of the Greater Cape flora: Ancient species repository, hot-bed of recent radiation, or both? *Molecular Phylogenetics and Evolution* 51(1): 44-53.

Vimal, K. P. (1952). Spores and pollen from tertiary lignites from Dandot, West Punjab (Pakistan). *Proceedings of the Indian National Science Academy Sect B* 36(4): 135-147.

Visher, S. S. (1925). Tropical cyclones and the dispersal of life from island to island in the Pacific. *American Naturalist* 59: 70-78.

von Hagen, K. B. & Kadereit, J. W. (2001). The phylogeny of *Gentianella* (Gentianaceae) and its colonization of the southern hemisphere as revealed by nuclear and chloroplast DNA sequence variation. *Organisms Diversity & Evolution* 1(1): 61-79.

von Hagen, K. B. & Kadereit, J. W. (2003). The diversification of *Halenia* (Gentianaceae): Ecological opportunity versus key innovation. *Evolution* 57(11): 2507-2518.

von Humboldt, A. (1805). *Essai sur la geographie des plantes accompagne d'un tableau physique des regions equinoxiales, fonde sur les mesures executees, depuis le dixieme degre de latitude boreale j'usqu'au dixieme degre de latitude australe, pendent les annees 1799, 1800, 1801, 1802 et 1803*. Paris: Levrault Schoell.

von Humboldt, A. (1808). *Ansichten der natur mit wissenschaftlichen erlauterungen*. Tubingen, Germany: J. C. Cotta.

Voris, H. K. (2000). Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography* 27(5): 1153-1167.

Wallace, A. R. (1860). On the Zoological Geography of the Malay Archipelago. *Zoological Journal of the Linnaean Society, London* 4: 172-184.

Wallace, A. R. (1863). On the physical geography of the Malay archipelago. *Journal of the Royal Geographical Societyof London* 33: 217-234.

Wallace, A. R. (1869). *The Malay Archipelago*. London: Macmillan.

Wallace, A. R. (1876). *The Geographical Distribution of Animals*. London: Macmillan.

Wang, Y., Fritsch, P. W., Shi, S., Almeda, F., Cruz, B. C. & Kelly, L. M. (2004). Phylogeny and infrageneric classification of *Symplocos* (Symplocaceae) inferred from DNA sequence data. *American Journal of Botany* 91(11): 1901-1914.

Wanntorp, L. & Wanntorp, H.-E. (2003). The biogeography of *Gunnera* L.: vicariance and dispersal. *Journal of Biogeography* 30: 979-987.

Warren, B. H. & Hawkins, J. A. (2006). The distribution of species diversity across a flora's component lineages: dating the Cape's 'relicts'. *Proceedings of the Royal Society B-Biological Sciences* 273(1598): 2149-2158.

Weeks, A., Daley, D. C. & Simpson, B. B. (2005). The phylogenetic history and biogeography of the frankincense and myrrh family (Burseraceae) based on nuclear and chloroplast sequence data. *Molecular Phylogenetics and Evolution* 35: 85-101.

Weeks, A. & Simpson, B. B. (2007). Molecular phylogenetic analysis of *Commiphora* (Burseraceae) yields insight on the evolution of historical biogeography of an "impossible" genus. *Molecular Phylogenetics and Evolution* 42: 62-79.

Wegener, A. (1912). Die Entstehung der Kontinente. *Geologische Rundschau* 3(4): 276-292.

Wegener, A. (1966). *The Origin of Continents and Oceans*. New York: Dover.

Wen, J. (1999). Evolution of Eastern Asian and Eastern North American disjunct distributions in flowering plants. *Annual Review of Ecology, Evolution and Systematics* 30: 421-455.

Wendel, J. F., Schnabel, A. & Seelanan, T. (1995). An unusual ribosomal DNA sequence from *Gossypium gossypioides* reveals ancient, cryptic intergenomic introgression. *Molecular Phylogenetics and Evolution* 4(3): 298-313.

Wertheim, J. O., Sanderson, M. J., Worobey, M. & Bjork, A. (2010). Relaxed molecular clocks, the bias-variance trade-off, and the quality of phylogenetic inference. *Systematic Biology* 59(1): 1-8.

Wheeler, E. A. & Manchester, S. R. (2002). Woods of the Middle Eocene Nut Beds Flora, Clarno Formation, Oregon, USA. *Iawa Journal*: 4-188.

Wheeler, E. A., Wiemann, M. C. & Fleagle, J. G. (2007). Woods from the Miocene Bakate Formation, Ethiopia Anatomical characteristics, estimates of original specific gravity and ecological inferences. *Review of Palaeobotany and Palynology* 146: 193-207.

White, F. (1983). *The vegetation of Africa: a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa*. Paris, UNESCO.

White, T. J., T. Burns, S. Lee, and J. Taylor. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: M. Innis, D. Gelfand, J. Sninsky & T. White (eds.) *PCR protocols: a guide to methods and applications* pp. 315-322. San Diego, California: Academic Press.

Whitmore, T. C. (1998). *An introduction to tropical rainforests, 2nd ed*. Oxford University Press.

Wichura, H., Bousquet, R., Oberhansli, R., Strecker, M. R. & Trauth, M. H. (2010). Evidence for middle Miocene uplift of the East African Plateau. *Geology* 38(6): 543-546.

Wiens, J.J. & Donoghue, M.J. (2004). Historical biogeography, ecology and species richness. Trends in Ecology & Evolution 19(12): 639-644.

Wikström, N., Avino, M., Razafimandimbison, S. G. & Bremer, B. (2010). Historical biogeography of the coffee family (Rubiaceae, Gentianales) in Madagascar: case studies from the tribes Knoxieae, Naucleeae, Paederieae and Vanguerieae. *Journal of Biogeography* 37(6): 1094-1113.

Wikström N., Savolainen, V. & Chase, M. W. (2001). Evolution of the angiosperms: calibrating the family tree. *Proceedings of the Royal Society: Biological Sciences* 268: 2211-2220.

WING, S. L., HARRINGTON, G. J., SMITH, F. A., BLOCH, J. I., BOYER, D. M. & FREEMAN, K. H. (2005). Transient floral change and rapid global warming at the Paleocene-Eocene boundary. *Science* 310: 993-996.

WING, S. L., HERRERA, F., JARAMILLO, C. A., GOMEZ-NAVARRO, C., WILF, P. & LABANDEIRA, C. C. (2009). Late Paleocene fossils from the Cerrejon Formation, Colombia, are the earliest record of Neotropical rainforest. *Proceedings of the National Academy of Sciences of the United States of America* 106(44): 18627-18632.

WINKWORTH, R. C., WAGSTAFF, S. J., GLENNY, D. & LOCKHART, P. J. (2002). Plant dispersal N.E.W.S. from New Zealand. *Trends in Ecology & Evolution* 17(11): 514-520.

WOJCIECHOWSKI, M. F., SANDERSON, M. J. & HU, J. M. (1999). Evidence on the monophyly of *Astragalus* (Fabaceae) and its major subgroups based on nuclear ribosomal DNA ITS and chloroplast DNA trnL intron data. *Systematic Botany* 24(3): 409-437.

WOLFE, A. D. & ELISENS, W. J. (1995). Evidence of chloroplast capture and pollen-mediated gene flow in *Penstemon* sect Peltanthera (Scrophulariaceae). *Systematic Botany* 20(4): 395-412.

WOLFE, J. A. (1975). Some aspects of plant geography of the Northern Hemisphere during the late Cretaceous and Tertiary. *Annals of the Missouri Botanical Garden* 62(2): 264-279.

WRIGHT, S. D., YONG, C. G., DAWSON, J. W., WHITTAKER, D. J. & GARDNER, R. C. (2000). Riding the ice age El Nino? Pacific biogeography and evolution of *Metrosideros* subg. Metrosideros (Myrtaceae) inferred from nuclear ribosomal DNA. *Proceedings of the National Academy of Sciences of the United States of America* 97(8): 4118-4123.

WRIGHT, S. D., YONG, C. G., WICHMAN, S. R., DAWSON, J. W. & GARDNER, R. C. (2001). Stepping stones to Hawaii: a trans-equatorial dispersal pathway for *Metrosideros* (Myrtaceae) inferred from nrDNA (ITS + ETS). *Journal of Biogeography* 28(6): 769-774.

XIANG, Q.-Y., THOMAS, D. T., ZHANG, W., MANCHESTER, S. R. & MURRELL, Z. (2006). Species level phylogeny of the genus *Cornus* (Cornaceae) based on molecular and morphological evidence - implications for taxonomy and Tertiary intercontinental migration. *Taxon* 55(1): 9-30.

XIANG, Q. Y., CRAWFORD, D. J., WOLFE, A. D., TANG, Y. C. & DEPAMPHILIS, C. W. (1998a). Origin and biogeography of *Aesculus* L. (Hippocastanaceae): A molecular phylogenetic perspective. *Evolution* 52(4): 988-997.

XIANG, Q. Y., SOLTIS, D. E. & SOLTIS, P. S. (1998b). The eastern Asian and eastern and western North American floristic disjunction: Congruent phylogenetic patterns in seven diverse genera. *Molecular Phylogenetics and Evolution* 10(2): 178-190.

XIANG, Q. Y., SOLTIS, D. E., SOLTIS, P. S., MANCHESTER, S. R. & CRAWFORD, D. J. (2000). Timing the eastern Asian-Eastern North American floristic disjunction: Molecular clock corroborates paleontological estimates. *Molecular Phylogenetics and Evolution* 15(3): 462-472.

XIE, L., YI, T. S., LI, R., LI, D. Z. & WEN, J. (2010). Evolution and biogeographic diversification of the witch-hazel genus (*Hamamelis* L., Hamamelidaceae) in the Northern Hemisphere. *Molecular Phylogenetics and Evolution* 56(2): 675-689.

YANG, Z. H. & RANNALA, B. (1997). Bayesian phylogenetic inference using DNA sequences: A Markov Chain Monte Carlo method. *Molecular Biology and Evolution* 14(7): 717-724.

Yemane, K., Robert, C. & Bonnefille, R. (1987). Pollen and clay mineral assemblages of a late Miocene lacustrine sequence from the northwestern Ethiopian highlands. *Palaeogeography, Palaeoclimatology, Palaeoecology* 60(1-2): 123-141.

Yoder, A. D. & Nowak, M. D. (2006). Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. *Annual Review of Ecology, Evolution and Systematics* 37: 405-431.

Yoder, A. D. & Yang, Z. H. (2000). Estimation of primate speciation dates using local molecular clocks. *Molecular Biology and Evolution* 17(7): 1081-1090.

You, C. X. & Petty, W. H. (1991). Effects of hurricane Hugo on *Manilkara bidentata*, a primary tree species in the Luquillo Experimental Forest of Puerto Rico. *Biotropica* 23(4): 400-406.

Yuan, Y. M., Wolhauser, S., Moller, M., Klackenberg, J., Callmander, M. W. & Kupfer, P. (2005). Phylogeny and biogeography of *Exacum* (Gentianaceae): a disjunctive distribution in the Indian Ocean basin resulted from long distance dispersal and extensive radiation. *Systematic Biology* 54(1): 21-34.

Zachos, J. C., Pagani, M., Sloan, L., Thomas, E. & Billups, K. (2001). Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292: 686-693.

Zerega, N. J. C., Clement, W. L., Datwyler, S. L. & Weiblen, G. D. (2005). Biogeography and divergence times in the mulberry family (Moraceae). *Molecular Phylogenetics and Evolution* 37: 402-416.

Zhang, J. M., Wang, J. X., Xia, T. & Zhou, S. L. (2009). DNA barcoding: species delimitation in tree peonies. *Science in China Series C-Life Sciences* 52(6): 568-578.

Zhang, L. B., Comes, H. P. & Kadereit, J. W. (2001). Phylogeny and quaternary history of the European montane/alpine endemic *Soldanella* (Primulaceae) based on ITS and AFLP variation. *American Journal of Botany* 88(12): 2331-2345.

Zhang, L. B., Simmons, M. P. & Renner, S. S. (2007). A phylogeny of Anisophylleaceae based on six nuclear and plastid loci: Ancient disjunctions and recent dispersal between South America, Africa, and Asia. *Molecular Phylogenetics and Evolution* 44(3): 1057-1067.

Zuckerkandl, E. & Pauling, L. (1962). Molecular disease, evolution, and genic heterogeneity. *M Kasha, and B Pullman, Editors Horizons in Biochemistry Albert Szent-Gyorgyi Dedicatory Volume*: p. 189-225.