

JCU ePrints

This file is part of the following reference:

Harrington, Mark (2008) *Phylogeny and evolutionary history of Sapindaceae and Dodonaea*. PhD thesis, James Cook University.

Access to this file is available from:

<http://eprints.jcu.edu.au/7893>



Phylogeny and evolutionary history of Sapindaceae and *Dodonaea*

Thesis submitted by

Mark Gerard Harrington BSc (Hons) JCU

in May 2008

for the degree of Doctor of Philosophy

in the School of Marine & Tropical Biology

James Cook University

STATEMENT OF ACCESS

I, the undersigned, author of this work, understand that James Cook University will make this thesis available for use within the University Library and, via the Australian Digital Theses network, for use elsewhere.

I understand that, as an unpublished work, a thesis has significant protection under the Copyright Act and;

I do not wish to place any further restriction on access to this work.

Signature

Date

STATEMENT OF SOURCES

DECLARATION

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Signature

Date

ELECTRONIC COPY

I, the undersigned, the author of this work, declare that the electronic copy of this thesis provided to the James Cook University Library is an accurate copy of the print thesis submitted, within the limits of the technology available.

Signature

Date

STATEMENT ON THE CONTRIBUTION OF OTHERS

This project was funded with grants from the Australian Biological Resources Study, the Cooperative Research Centre for Tropical Rainforest Ecology and Management (Rainforest CRC) and James Cook University (JCU). I was supported by a Australian Postgraduate Award stipend and a completion scholarship from JCU.

Supervision was provided by Prof. Paul Gadek and Dr. Will Edwards (JCU) with editorial assistance from Christopher Quinn, Honorary Research Associate, Botanic Gardens Trust, Sydney.

Acknowledgements

I would like to thank my supervisors Prof. Paul Gadek and Dr. Will Edwards, and Dr. Judy West from the Australian National Herbarium, Christopher Quinn, and Ed Biffin from Adelaide University for further discussions and assistance.

I acknowledge Lynne Jones and Mathew Pye for advice on molecular techniques.

I extend my thanks to Peter Raven for exhorting his many friends to help get plant samples for this project. I thank the following people and institutions from around the world for supplying plant material – Chris Birkinshaw, George Yatskievych and Gregory Gust Missouri Botanical Garden and DNA Bank, Herbarium, Dr. Victor Barradas University National avtonoma of Mexico, Dr. Brinsley Burbidge St. George Village Botanical Garden, St. Croix Virgin Islands, Suhua shi, Sun Yat-sen University, Guangzhou, Cassius Kayombo National Herbarium of Tanzania, Pedro Acevedo United States National Herbarium, Smithsonian Institution, National Tropical Botanical Garden, Kauai, Hawaii, F. Billiet Living Plant Collections, National Botanic Garden of Belgium, Botanic Garden, University of Copenhagen, Dr. Jenny Xiang, North Carolina State University, Tony Roberts. Mary Gandini, Stuart Worboys, James Cook University, Judy West, Kirsten Cowley, Lyn Craven and Ed Biffin Centre for Plant Biodiversity Resaearch, Australian National Herbarium (CANB), Australian National Botanic Gardens living collection, Darren Crayn, Royal Botanical Gardens Sydney, Dr. Shyam Phartyal Forest Research Institute India, Cassius Kayombo National Herbarium of Tanzania, Amanda Shade Botanic Gardens and Parks Authority Perth Western Australia, Chien-I Huang, Royal Botanic Gardens Edinburgh, F. Billiet Living Plant Collections, National Botanic Garden of Belgium, Ching I Peng, Research Centre for Biodiversity. Academia Sinica, Taipei, Taiwan and Christine Stucley; and DNA samples – Shelly James Pacific Centre for Molecular Biodiversity, Bishop Museum, Hawaii Kew Botanical Gardens – DNA Bank, Dr. Clifford Morden Hawaii Plant DNA Library.

Finally, I can now get back to painting the house.

ABSTRACT

The Sapindaceae (soapberry family) is a large family of trees, shrubs and lianas comprising 133 genera (49 monotypic) and ca.1450 species, most with a tropical to subtropical distribution. The circumscription of Sapindaceae has varied, particularly with regard to inclusion of genera from the closely related, predominantly temperate families Aceraceae and Hippocastanaceae. Sequence data from two plastid genes, analysed separately and together using parsimony and Bayesian analysis support a broadly defined Sapindaceae incorporating Aceraceae, Hippocastanaceae and *Xanthoceras*. A division into four subfamilies is proposed: Sapindoideae, Hippocastanoideae, Dodonaeoideae, and Xanthoceroideae. Tribal groupings are critically evaluated in light of the analyses.

The evolutionary history of Sapindaceae is evaluated using a variety of Bayesian relaxed clock molecular estimates of divergence times, which either incorporate the dates implied by the fossil record, fossil constraints from outside Sapindaceae, and no fossil constraints. Analyses with fossil constraints from outside Sapindaceae imply a Pliocene-Miocene (6-28 Mya) origin for *Acer* and *Aesculus* rather than a Paleocene (ca 64 Mya) origin implied by the earliest attributed fossils of these genera, and an evolutionary rate for *Acer* and *Aesculus* consistent with the majority of other genera of Sapindaceae sampled. Including the fossil dated Paleocene origin of *Acer* and *Aesculus* as hard bound priors and using four constraints from outside Sapindaceae result in a potentially biologically implausible rapid change in the mean evolutionary rate on the stem branch leading to the split between *Acer* and *Aesculus* and their respective sister genera. These conflicting scenarios suggest the need for a substantial re-evaluation of our understanding of the tempo and mode of evolution of these lineages.

Intergeneric relationships within Dodonaeoideae, Hippocastanoideae and Xanthoceroideae were also assessed by firstly, generating secondary structure predictions for ITS and partial ETS sequences, and then using these predictions to assist alignment of the sequences. Secondly, the alignment was analysed using RNA specific models of sequence evolution that account for the variation in nucleotide evolution in the independent loops and covarying stems regions of the ribosomal spacers. The

phylogeny drawn from the analyses was compared with that from analyses using ‘traditional’ 4-state models and the plastid analyses.

To further our understanding of the origins of the Australian flora, and in particular plant adaptation and evolution in increasingly arid environments, the phylogeny and evolutionary history of Hopbushes (*Dodonaea* – Sapindaceae) and Pepperflowers (*Diplopeltis* – Sapindaceae) were evaluated based on nuclear ITS and partial ETS sequences and models of nucleotide evolution that incorporate secondary structure. The tempo and mode of evolution of these sister genera were evaluated using Bayesian relaxed clock molecular estimates of divergence times, Bayesian relative rates test, lineage through time plots and estimations of diversification rates (speciation minus extinction). The dry and temperate adapted genera of Sapindaceae (*Diplopeltis* species and *Dodonaea* including *Distichostemon* species) are relatively recent radiations in the Australian flora, and are most likely to be in response to increased aridity and seasonality from the late Miocene 14 Mya to Recent. There is evidence of long distance dispersal from northern Australia to Madagascar in the early Pliocene.

The cosmopolitan, polymorphic species *Dodonaea viscosa* (hop bush, varnish tree) has been the subject of taxonomic and ecological enquiry for over 150 years. ITS and partial ETS sequences for 50 samples from across its worldwide distribution were used to evaluate the evolutionary and biogeographic history of this species complex. *Dodonaea viscosa* is not an old lineage as has often been speculated based primarily on its vast distribution, but diverged from its most recent common ancestor and subsequently dispersed around the world within the last two million years. Results also indicate that there are at least two evolutionary lineages within *D. viscosa*.

Table of contents

Chapter 1: Introduction	1
Chapter 2: Phylogenetic inference in Sapindaceae sensu lato using plastid matK and rbcL DNA sequences	4
Materials and Methods	8
Sampling.....	8
Combined Sequence Data.....	9
Phylogenetic Analyses	10
Results	12
Sequence Data.....	12
Phylogenetic Analyses	17
Discussion.....	19
Circumscription of Sapindaceae	19
Aceraceae and Hippocastanaceae	21
Dodonaeoideae.....	22
Cossinieae/Dodonaeeae/Harpullieae.....	23
Doratoxyleae.....	25
Diplokeleba floribunda-Averrhoidium.....	25
Sapindoideae.....	26
Clade A (Lepisantheae–Sapindeae.....	27
Clade B (Nephelieae)	28
Clade C (Cupanieae).....	29
Clade D (Paulinieae-Thouinieae).....	30
Conclusion	31
Chapter 3: Tempo of evolution in Acer and Aesculus (Sapindaceae)	33
Methods	35
Sampling and Sequencing	35
Molecular dating and phylogenetic analysis.....	36
Results and Discussion	40
Chapter 4: Comparative study of the evolution of nuclear ribosomal spacers incorporating secondary structure analyses within Dodonaeoideae, Hippocastanoideae and Xanthoceroideae (Sapindaceae)	54
Methods	57

Sampling, DNA extraction, amplification and sequencing	57
Secondary structure prediction and sequence alignment.....	58
Phylogenetic analysis - model selection.....	59
Phylogenetic analyses.....	62
Results and discussion	63
Alignment and secondary structure.....	63
Spacer molecular evolution comparison.....	68
Model selection.....	72
Phylogenetic analyses.....	73
Conclusion.....	76
Chapter 5: Understanding the origins of the Australian flora: plant adaptation and evolution in increasingly arid environments - phylogeny and evolutionary history of Hopbushes (<i>Dodonaea</i> – Sapindaceae) and Pepperflowers (<i>Diplopeltis</i> – Sapindaceae) based on nuclear ITS and partial ETS sequences	77
Materials and Methods	81
Taxon sampling and outgroup selection.....	81
DNA extraction, amplification and sequencing.....	81
Alignment and secondary structure prediction	82
Model selection and phylogeny estimation	82
Relative evolutionary rates analysis.....	84
Relative rate test.....	84
Diversification rate.....	85
Biogeographical analyses	86
Results	87
Model selection.....	90
Bayesian MCMC estimation of phylogeny	91
Evolutionary rates analyses	94
Rate of extant lineage accumulation	94
Discussion.....	101
Delimitation of <i>Diplopeltis</i>	101
Phylogenetic reconstruction within <i>Dodonaea</i>	103
Clade I	105
Clade II - <i>Distichostemon</i>	105
Node A	108

Evolutionary history of hopbushes and pepperflowers	109
Transoceanic disjunctions and dispersal biology	113
Rate heterogeneity over time	115
Diversification and geographical phylogenetic structure	118
Conclusion	121
Chapter 6: A species well travelled – the <i>Dodonaea viscosa</i> (Sapindaceae) complex based on phylogenetic analyses of nuclear ribosomal ITS and ETSf sequences	123
Materials and methods	129
Study species biological profile	129
Sampling	131
Model selection	131
Phylogeny estimation	132
Relative evolutionary rates analysis	133
Fossils	134
Results	134
Evolutionary rates analyses	136
Discussion	139
Group I	141
Group II	143
Dispersal	144
One species or many?	146
Conclusion	147
Conclusion	148
BIBLIOGRAPHY	150
APPENDIX 2.1	168
APPENDIX 3.1	172
APPENDIX 4.1	174
APPENDIX 4.2A	175
APPENDIX 4.2B	176
APPENDIX 5.1	177
APPENDIX 5.2	178
APPENDIX 5.3	179
APPENDIX 6.1	181
APPENDIX 6.2	184

List of tables

Chapter 1 Introduction

Chapter 2 Phylogenetic inference in Sapindaceae

Table 2.1 Enumeration schemes for subfamilies and tribes assigned to Sapindaceae.....	7
Table 2.2 List of new <i>matK</i> primers used in this study.....	9
Table 2.3 Species combined as placeholder taxa.....	10

Chapter 3 Tempo of evolution in *Acer* and *Aesculus* (Sapindaceae)

Table 3.1 Prior probability distributions and posterior probability densities from analyses of various combinations of priors	38
Table 3.2 Coefficient of variation and cupdate fieldsovariance statistics for analyses B of the combined dataset and for the single gene analyses	43
Table 3.3 Mean substitution rate per site from the Bayesian lognormal relaxed clock phylogeny estimations of the combined and single gene datasets performed without any time constraints	46

Chapter 4: Comparative study of Dodonaeoideae, Hippocastanoideae and Xanthoceroideae

Table 4.1 Glossary of RNA secondary structure terms	56
Table 4.2 General description of RNA specific models implemented in software package PHASE version 2.0b 2005	60
Table 4.3 Hierarchical selection process to determine the most appropriate model/s to analyse the combined ITS and ETSf datasets	63
Table 4.4 Empirical nucleotide frequencies across the three spacer regions	70
Table 4.5 Base pair frequency, mutability of base pairs and substitution rate parameters inferred using the RNA7C model.....	71
Table 4.6 Comparison of AICc scores for various partitioning schemes.....	73

Chapter 5: Evolutionary history of Hopbushes and Pepperflowers

Table 5.1 Secondary structure description of all indels for the combined data matrix	89
Table 5.2 Comparison of AICc scores for various partitioning schemes	90
Table 5.3 Extrapolation of the substitution rates from the Bayesian uncorrelated log normal estimation of phylogeny and substitution rates for the combined dataset to units per time	99

Table 5.4 Net rate of diversification per million years for <i>Dodonaea</i> and <i>Diplopeltis</i> in the absence of extinction ($\epsilon = 0.0$) and at relatively high extinction rate ($\epsilon = 0.9$) for stem and crown groups	100
Chapter 6: <i>Dodonaea viscosa</i>	
Table 6.1 Seven subspecies of Australian <i>Dodonaea viscosa</i> from West (1984), with general description of habit and distribution	127
Table 6.2 Comparison of AICc scores for three different partitions of the ITS/ETSf concatenated alignment.....	135
Table 6.3 Morphological intergradation between subspecies of <i>Dodonaea viscosa</i> , and the regions where it has been recorded.....	144

List of figures

Chapter 1: Introduction

Chapter 2: Phylogenetic inference in Sapindaceae

- Figure 2.1 Consensus tree obtained from the Bayesian *rbcL* analysis. 13
- Figure 2.2 Consensus tree obtained from the Bayesian *matK* analysis..... 15
- Figure 2.3 Consensus tree obtained from the combined Bayesian analysis 16

Chapter 3: Tempo of evolution in Acer and Aesculus

- Figure 3.1 Edited consensus phylogram from Bayesian analysis of combined plastid dataset for Sapindaceae with mean branch lengths measured in expected substitutions per site 41
- Figure 3.2 Bayesian relative rates test 95% credibility intervals 44
- Figure 3.3 Edited chronogram from the Bayesian relaxed clock analyses using four fossil constraints 48
- Figure 3.4 Edited chronogram showing substitution rates 49
- Figure 3.5 Edited chronogram showing substitution rates 51

Chapter 4: Comparative study of Dodonaeoideae, Hippocastanoideae and Xanthoceroideae

- Figure 4.1 Relationships between the RNA7 classes of models 61
- Figure 4.2 RNA transcript secondary structure for *Majidea fosterii*: a) ITS1 b) ITS2 c) ETS 64
- Figure 4.3 Nucleotide frequencies comparisons 69
- Figure 4.4 Phylogeny for 40 taxa of Xanthoceroideae, Hippocastanoideae and Dodonaeoideae of combined ITS and ETS secondary structure alignment 75

Chapter 5: Evolutionary history of Hopbushes and Pepperflowers

- Figure 5.1 Australian biomes 87
- Figure 5.2 Representative secondary structures for helix II of ITS1 RNA transcript 88
- Figure 5.3 Phylogeny for *Dodonaea* and *Diplopeltis* derived from combined ITS and ETS secondary structure alignment 92
- Figure 5.4 Ratogram from the Bayesian MCMC phylogeny and molecular evolutionary rates estimations of the combined aligned matrix without indels..... 96
- Figure 5.5 Chronogram from the Bayesian MCMC phylogeny and molecular evolutionary rates estimations of the combined aligned matrix without indels 98
- Figure 5.6 Lineage through time plot for *Dodonaea* 99

Figure 5.7 Lineage accumulation over time of extant <i>Dodonaea</i> and <i>Diplopeltis</i> in Australian biomes over one million year time intervals	100
Figure 5.8 Confidence intervals of expected species diversity according to age of crown group	101
Figure 5.9 Nucleotide alignment positions 47-80 from helix II ITS1 for selected taxa ..	103
Figure 5.10 Seed types for <i>Dodonaea</i> discussed in text and selection of outgroups	107
Figure 5.11 Contrasting interpretations of the evolutionary history of crown group <i>Diplopeltis</i> (four species) with a long stem from spit with <i>Dodonaea</i>	111
Figure 5.12 Bayesian relative rates test 95% credibility intervals	117
Figure 5.13 Geographical zonation of <i>Dodonaea</i> depicting areas with similar numbers of species.....	120
Chapter 6: <i>Dodonaea viscosa</i>	
Figure 6.1 Worldwide distribution of <i>Dodonaea viscosa</i> and climatic factors of habitat	124
Figure 6.2 Polygamo-dioecious (♀ and bisexual) <i>Dodonaea viscosa</i> ssp. <i>viscosa</i> on front dune next to mangroves Half Moon Bay Beach, Cairns, Australia.....	130
Figure 6.3 One of the 12,000 trees from the Bayesian MCMC analyses of the combined dataset	137
Figure 6.4 The posterior distribution of relative branching times	138
Figure 6.5 Current distribution of <i>Dodonaea viscosa</i> in Australia	141