

**Molecular Systematics of the Genus *Phylica* L. with an Emphasis on  
the Island Species**

**James Edward Richardson**

**Doctor of Philosophy  
The University of Edinburgh  
February, 1999**





*Phylica arborea*, Nightingale Island, South Atlantic Ocean.

I declare that this thesis has been composed by myself and the work contained within, unless otherwise stated, is my own.

## ABSTRACT

*Phylica* L. (Rhamnaceae) consists of about 150 species, most of which are found in Cape Province, South Africa. A number of species are found on islands off southern Africa such as St Helena, Tristan da Cunha, New Amsterdam, Mauritius, Réunion and Madagascar. *Phylica* has two close relatives, *Nesiota* Hook. f. (a monotypic genus from St Helena) and *Noltea* Reichb. (a monotypic genus from South Africa). Most of the species on the mainland are ericoid shrubs, whereas some of the island species and the genera *Nesiota* and *Noltea* are broad-leaved trees or shrubs that have retained other putatively primitive characteristics. I assessed tribal relationships in Rhamnaceae and relationships of the family itself using DNA sequences from two regions of the plastid genome, *rbcL* and *trnL-F*. This revealed that the closest relatives of Rhamnaceae are Dirachmaceae and Barbeyaceae. The plastid trees support the monophyly of the family and provide the basis for a new tribal classification. Three major strongly supported clades are identified, but morphological characters could not be found to underpin a formal taxonomic description of these three clades as subfamilies. A morphological phylogenetic analysis of Rhamnaceae using 18 characters provided less resolution than analysis of molecular characters. Sequences of *trnL-F* and internal transcribed spacer nuclear ribosomal DNA (ITS) showed that the genera *Nesiota* and *Noltea* are sister to *Phylica* and palaeoendemic within the context of the tribe Phyliceae and the island species of *Phylica* form an 'island group' embedded within the genus together with the widespread mainland species *P. paniculata*. Within the context of the 'island group' the Mascarene species *P. nitida* is a palaeoendemic sister to the other island species which are recently derived neoendemics. The plesiomorphic, generalist morphology of the island species contrasts with the derived morphological characteristics of the majority of mainland species. Amplified fragment length polymorphisms (AFLPs) reveal higher levels of variation than gene sequences and were therefore used to elucidate relationships between island species and *P. paniculata* from Africa. Parsimony, neighbour joining and PCO analyses performed on the AFLP data set indicate that each of the species forms a distinct group of genotypes, and indicate genetic relationships and possible origins of different island

populations of the same species. The data are consistent with the derivation of *P. arborea* on Gough Island from a single introduction from Tristan da Cunha and on New Amsterdam from a single introduction from Gough Island. AFLPs were used to determine levels of genetic variation in two endangered St Helenan endemic species of Rhamnaceae. No AFLP variation was detected in the four remaining individuals of *Nesiota* indicating that it is effectively clonal. This was contrasted with polymorphism that was detected between populations and among individuals of *P. polifolia*. AFLP data have therefore proved to be useful for developing appropriate conservation strategies for these species.

## ACKNOWLEDGEMENTS

The supervisors: Mark Chase, Quentin Cronk and Mike Fay.

My co-worker on *Phyllica* Frans Weitz, his supervisor Peter Linder and Nicki Allsopp for board and lodging in Cape Town.

Staff in the Molecular Systematics Section, Jodrell Laboratory, Royal Botanic Gardens, Kew: Lola Llédó, Gail Reeves, Vincent Savolainen, Cynthia Morton, Abraham Muasya, Cassio van den Berg, Robyn Cowan, Faridah Qamaraz Zaman, Anette de Bruijn, Jeff Joseph, Martin Powell, Paul Bygrave, Mary Clare Sheahan, Liz Caddick, Trevor Hodgkinson and Doug Goldman. Staff at the Herbarium, Royal Botanic Gardens, Kew, particularly Aaron Davis, Justin Moat, Roger Polhill and Paul Bygrave.

Staff of the Royal Botanic Garden, Edinburgh, particularly Toby Pennington, Pete Hollingsworth, The EGF team, Caroline Guihal, Gill Preston and Rob Cubey.

Staff at The University of Edinburgh.

Collectors of plant material: Frans Weitz, Yves Frenot, Christophe Thébaud, Michelle van der Bank, Ben Erik van Wyk, Yusoof Mungroo, Marc Lebouvier, Peter Goldblatt, Rebecca Rowe and Koos Roux.

The people of Tristan da Cunha particularly Karl and Julia Hagen, The Administrator (Brendan Dalley), The Head Islander (Jimmy Glass) and the rest of the Island Council.

The crews of the SA Agulhas and SAS Drakensberg.

Personal thanks: The Richardson Family, Paul Bygrave, Peter Sims, Justin Moat and Alison Kadleck, The Kings, The Vegas Kids, Helen Carrier, Pam Firth, Liz Caddick, Cymon Cox, Hazel Kennedy and Marjorie Santamaria.

# CONTENTS

## ABSTRACT

## ACKNOWLEDGEMENTS

<b>CHAPTER ONE. Introduction</b>	<b>1</b>
<b>1.1. General Introduction</b>	<b>1</b>
<b>1.2. Comparison of the use of morphological and molecular data in systematics</b>	<b>4</b>
<b>1.3. Introduction to Cladistics and Molecular Systematics</b>	<b>6</b>
1.3.1. Introduction to cladistics	6
1.3.2. Cladistic characters and homology	7
1.3.3. The use of molecular characters in phylogenetic studies	8
1.3.4. Homology of molecular characters	9
1.3.5. Methods for inferring phylogeny	9
1.3.5.1. Distance data	9
1.3.5.2. Discrete data	11
1.3.5.2.1. Maximum parsimony	11
1.3.5.2.2. Maximum likelihood	13
1.3.6. Choice of method used to analyse sequence data	14
<b>1.4. Assessing Variation at Species Boundaries and Between and Within Populations</b>	<b>14</b>
<b>1.5. Molecular markers in population genetics</b>	<b>15</b>
1.5.1. The use of proteins in population genetic studies	16
1.5.2. The use of nucleic acids in population genetic studies	16
1.5.2.1. Homology of DNA segments and alleles	17
1.5.2.2. Restriction fragment length polymorphisms (RFLPs)	18
1.5.2.3. Variable number tandem repeats (VNTRs)	19
1.5.2.4. Randomly amplified polymorphic DNA (RAPD)	20
1.5.2.5. Amplified fragment length polymorphisms (AFLPs)	22

<b>1.6. Conclusions and choice of methods</b>	24
<b>1.7. Phylogenetics of Rhamnaceae and <i>Phylica</i> L.</b>	24
1.7.1. Aims of Rhamnaceae and <i>Phylica</i> Phylogenetic Study	25
<b>1.8. Population Level Studies on <i>Phylica</i> Island Species and their Mainland Relatives</b>	26
1.8.1. Aims of population level study	26
<b>1.9. Conservation Genetics Study</b>	26
1.9.1. Aims of conservation genetic study	27
<b>1.10. Thesis Structure</b>	27
<b>1.11. Bibliography</b>	27
<b>CHAPTER TWO. A Molecular Analysis Of Rhamnaceae Using <i>rbcL</i> And <i>trnL-F</i> Plastid DNA Sequences</b>	39
<b>Abstract</b>	39
<b>2.1. Introduction</b>	39
<b>2.2. Aims</b>	44
<b>2.3. Materials and Methods</b>	44
2.3.1. Material for molecular analysis	44
2.3.2. DNA extraction	44
2.3.3. Gene amplification and purification	45
2.3.4. DNA sequencing	51
2.3.5. Sequence alignment	51
2.3.6. Phylogenetic analysis using parsimony (PAUP)	52
2.3.6.1. Exact methods	52
2.3.6.2. Heuristic methods	52
2.3.6.2.1. Stepwise addition	52
2.3.6.2.2. Branch swapping	53
2.3.6.2.3. Accelerated and delayed transformations	54
2.3.6.2.4. Assessing the reliability of inferred trees	54
2.3.6.2.5. Successive weighting	55



2.3.6.2.6. Combining equally parsimonious trees (consensus techniques)	55
2.3.6.2.7. Confidence measures	56
2.3.6.3. Heuristic search strategy	57
<b>2.4. Results</b>	58
2.4.1. <i>rbcL</i> analysis	58
2.4.2. <i>trnL</i> -F analysis	60
2.4.3. Combined <i>rbcL</i> and <i>trnL</i> -F analysis	61
2.4.4. Molecular evolution	65
<b>2.5. Discussion</b>	69
2.5.1. Molecular evolution	69
2.5.2. Relationships of Rhamnaceae	72
2.5.3. Relationships within Rhamnaceae	72
2.5.4. Biogeography of Rhamnaceae	76
<b>2.6. General Conclusions</b>	78
<b>2.7. Bibliography</b>	79
<b>CHAPTER THREE. Morphological Analysis Of Rhamnaceae</b>	89
<b>Abstract</b>	89
<b>3.1. Introduction</b>	89
<b>3.2. Methods</b>	90
3.2.1. Description of characters	90
3.2.2. Phylogenetic analysis	92
<b>3.3. Results</b>	94
<b>3.4. Discussion</b>	98
<b>3.5. Conclusions</b>	109
<b>3.6. Rhamnaceae Tribal Classification</b>	109
Paliureae	109
Colletieae	110
Phylliceae	111
Gouanieae	111

Pomaderreae	112
Rhamneae	112
Maesopsidaeae	113
Ventilagineae	113
Ampelozizipheae	114
Doerpfeldieae	114
Bathiorhamneae	115
Genera <i>incertae sedis</i>	115
<b>3.7. Bibliography</b>	<b>118</b>

**CHAPTER FOUR. Phylogenetic Studies Of *Phylica* L. With An Emphasis On Island Species: Evidence From Plastid *trnL-F* And Nuclear Transcribed Spacer (Ribosomal DNA) Sequences** 123

<b>Abstract</b>	123
<b>4.1. Introduction</b>	124
4.1.1. Taxonomic history of <i>Phylica</i> island species	124
4.1.2. Biogeographic context of the genus <i>Phylica</i>	125
4.1.2.1. Southern Africa	126
4.1.2.1.1. Pre-Pliocene forest environments (65-5 million years ago, mya)	126
4.1.2.1.2. Pliocene origin of seasonality and the birth of Fynbos (5-0 mya)	131
4.1.2.2. St Helena	132
4.1.2.3. The Tristan da Cunha Group	133
4.1.2.4. Madagascar	134
4.1.2.5. New Amsterdam	135
4.1.2.6. Mauritius	136
4.1.2.7. Réunion	136
4.1.3. Morphology of Phyliceae	136
4.1.4. Phyliceae biogeography	144
<b>4.2. Aims of Study</b>	<b>147</b>

<b>4.3. Materials and Methods</b>	147
4.3.1. Material for molecular analysis	147
4.3.2. DNA extraction	148
4.3.3. Gene amplification and purification	148
4.3.4. DNA sequencing	149
4.3.5. Sequence alignment	149
4.3.6. Phylogenetic analysis	149
4.3.7. Molecular Clock	150
<b>4.4. Results</b>	153
4.4.1. <i>trnL</i> -F analysis	153
4.4.2. ITS analysis	154
4.4.3. Combined analysis	155
<b>4.5. Discussion</b>	155
4.5.1. Origin and paraphyly of <i>Phylica</i>	155
4.5.2. The 'island group' and the origins of the island species	160
4.5.3. Biogeographic history of <i>Phyliceae</i> and its island species	162
4.5.4. Comparative evolution of the island and mainland species of <i>Phylica</i>	167
<b>4.6. Conclusions</b>	168
<b>4.7. Bibliography</b>	169
<b>CHAPTER FIVE. Relationships Of Island Populations Of <i>Phylica</i> L. Based On Amplified Fragment Length Polymorphisms</b>	176
<b>Abstract</b>	176
<b>5.1. Introduction</b>	176
<b>5.2. Aims of Study</b>	178
<b>5.3. Island Vegetation and Demographic Status of Species Involved in the Study</b>	178
5.3.1. Island vegetation	178
5.3.1.1. Tristan da Cunha Group	178
5.3.1.2. New Amsterdam	179
5.3.1.3. Mauritius	181

5.3.1.4. Réunion	181
5.3.1.5. St Helena	181
5.3.2. Taxonomic history and demographic status of species involved in the study	182
5.3.2.1. <i>Phylica arborea</i>	182
5.3.2.2. <i>Phylica polifolia</i>	190
5.3.2.3. <i>Phylica nitida</i>	191
5.3.2.4. <i>Phylica emirnensis</i>	191
5.3.2.5. <i>Phylica bathiei</i>	192
5.3.2.6. <i>Phylica paniculata</i>	193
5.3.3. Morphological differences between members of the island group	194
<b>5.4. Methods</b>	194
5.4.1. Sampling strategies	194
5.4.2. Material for analysis	195
5.4.3. DNA extraction	195
5.4.4. Amplified fragment length polymorphisms (AFLPs)	195
5.4.5. Running AFLPs on gels and band scoring	196
5.4.6. Data analysis: methods for analysing restriction fragment data and expectations for performance	201
5.4.6.1. Unweighted pair group method with arithmetic means (UPGMA)	201
5.4.6.2. Parsimony	201
5.4.6.3. Neighbour joining (NJ)	201
5.4.6.4. Principal co-ordinates analysis (PCO)	202
5.4.6.5. Software	202
<b>5.5. Results</b>	203
5.5.1. UPGMA	203
5.5.2. Parsimony	203
5.5.3. Neighbour joining	204
5.5.4. Principal co-ordinates analysis	208
<b>5.6. Discussion</b>	213

5.6.1. Number of species in the 'island group' and the monophyly of these taxa	213
5.6.2. Genetic variation within and between populations of island species of <i>Phylica</i> and the possible origins of island species and populations	213
5.6.3. Dispersal of <i>Phylica</i> island species	218
<b>5.7. Conclusions</b>	221
<b>5.8. Bibliography</b>	221

**CHAPTER SIX. Conservation Genetics Of Threatened St Helenan Species Of Rhamnaceae** 226

<b>Abstract</b>	226
<b>6.1. Introduction</b>	226
6.1.1. <i>Nesiota elliptica</i>	227
6.1.2. <i>Phylica polifolia</i>	228
6.1.3. Conserving rare plants - genetic variability and species viability	230
6.1.4. Examples of the use of AFLPs in conservation genetics	231
<b>6.2. Aims of Study</b>	232
<b>6.3. Methods</b>	232
<b>6.4. Results</b>	232
<b>6.5. Discussion</b>	237
<b>6.6. Conclusions</b>	239
<b>6.7. Bibliography</b>	240

**CHAPTER SEVEN. Conclusions On The Use Of Molecular Data In Solving Systematic Problems At Different Hierarchical Levels In Rhamnaceae** 245

<b>7.1. Rhamnaceae Study</b>	245
<b>7.2. Phylliceae Study</b>	246
<b>7.3. AFLP Study on the Island Species Of <i>Phylica</i></b>	248

<b>7.4. Conservation genetics</b>	250
<b>7.5. General conclusions</b>	251
<b>7.6. Bibliography</b>	252

## List of Tables

### Chapter Two

2.1. Taxonomic history of relationships of Rhamnaceae and related families	41
2.2. Taxonomic history of suprageneric classifications in Rhamnaceae	46
2.3. Taxon accession data	48
2.4. Performance of each codon position in the <i>rbcL</i> analysis	65
2.5. Performance of <i>trnL</i> -F indel characters	68
2.6. Tree scores for transitions and transversion on an SW tree from the combined <i>rbcL/trnL</i> -F analysis	68
2.7. Comparison of number of steps for the separate analyses versus the combined trees	69
2.8. Summary of the revised classification of Rhamnaceae	76

### Chapter Three

3.1. Characters used in a morphological phylogenetic analysis of Rhamnaceae	93
3.2. Matrix of character states for a morphological analysis of Rhamnaceae	95
3.3. CI and RI values for each of the individual morphological characters on the trees from the morphological analysis and the combined morphological/molecular analysis	98

### Chapter Four

4.1. Taxonomic history of <i>Phyllica</i>	128
4.2. Sequences of AB101 and AB102 primers (G. Sheridan)	149
4.3. Taxon accession data	151

## Chapter Five

5.1. Samples used in a study of AFLPs in island species of <i>Phyllica</i>	197
5.2. Eigenvalues for PCO analyses of AFLP data sets	209

## List of Figures

### Chapter Two

Figure 2.1. *Ziziphus jujuba* flowers showing stamens and petals alternating with sepals (Figure taken from Suessenguth, 1953) 40

Figure 2.2. Example of one optimal SW tree from the *rbcL* analysis, with its Fitch lengths (above branches; ACCTRAN optimisation) and bootstrap values (below). Branches, which are not present in the strict consensus tree are indicated by an arrow. Heuristic search under the Fitch criterion produced more than 6000 equally parsimonious trees with a length of 1174 steps. The consistency index (CI) for these trees was 0.52 and the retention index (RI) was 0.66. There were only seven SW trees with a length of 423378 steps, CI=0.84, and RI=0.86 (Fitch length, 1174 steps) 62

Figure 2.3. Example of one optimal SW tree from the *trnL-F* analysis, with its Fitch lengths (above branches; ACCTRAN optimisation) and bootstrap values (below). Branches not present in the strict consensus tree are indicated by an arrow. Heuristic search under the Fitch criterion produced more than 6000 equally parsimonious trees with a length of 1339 steps, CI=0.67, and RI=0.75. SW produced more than 5000 trees and a length of 652105 steps, CI=0.87, and RI=0.91 (Fitch length, 1339 steps) 63

Figure 2.4. The single optimal SW tree from the combined *rbcL/ trnL-F* analysis, with its Fitch lengths (above branches; ACCTRAN optimisation) and bootstrap values (below). Branches, which are not present in the strict consensus tree are

indicated by an arrow. Heuristic search under the Fitch criterion produced 324 Fitch trees with a length of 2559 steps, CI=0.59 and RI=0.70. SW produced one tree with two trichotomies and a tree length of 1068277 steps, CI=0.85, RI=0.88 (Fitch length, 2559 steps) 64

Figure 2.5. Number of changes per character based on the single SW tree from the combined *rbcL/trnL*-F analysis 66

Figure 2.6. Number of steps each of the variable sites produced on the single SW tree from the combined *rbcL/trnL*-F analysis 67

### Chapter Three

Figure 3.1. One of the 5000 trees from a morphological analysis of Rhamnaceae, using 18 characters. Branch lengths are above branches and bootstrap values are below. Branches that collapse in the strict consensus tree are indicated by an arrow. The length of the trees is 50 steps, CI=0.44 and RI=0.83. The tribal placement of each genus according to Suessenguth (1953) is indicated 99

Figure 3.2. One of the 942 trees from a molecular analysis of Rhamnaceae. Branch lengths are above branches and bootstrap values are below. Branches that collapse in the strict consensus tree are indicated by an arrow. The length of the trees is 1660 steps, CI=0.65 and RI=0.76. The tribal placement of each genus according to Richardson *et al.* (submitted) is indicated 100

Figure 3.3. Strict consensus trees: left = morphological analysis of Rhamnaceae and right = combined *rbcL/trnL*-F molecular analysis (Chapter Two) 101

Figure 3.4. One of the 216 trees from a combined morphological and molecular analysis of Rhamnaceae using 18 morphological characters and *rbcL* and *trnL*-F characters. Branch lengths are above branches and bootstrap values are below. Branches that collapse in the strict consensus tree are indicated by an arrow. The length of the trees is 1726, CI=0.64 and RI=0.76. The tribal placement of each genus according to Richardson *et al.* (submitted) is indicated 102



Figure 3.5. Morphological character states mapped onto a combined morphological and molecular tree. Thick bars represent character state changes 103

## Chapter Four

Figure 4.1. Distribution Of *Phylica* 127

Figure 4.2. *Phylica pubescens*: A. Inflorescence; B. Flower; C. Transverse section of flower; D. Fruit; E. Capsule; Cross section of capsule; F. Seed with elaiosome. *Phylica virgata*: G. Inflorescence; H. Flower. *P. oleaefolia*: J. Fruit; K. Inflorescence (from Suessenguth, 1953) 138

Figure 4.3. Floral morphology of a selection of *Phylica* species (from Pillans, 1942) 139

Figure 4.4. *Noltea africana* (from Sim, 1907) 141

Figure 4.5. *Nesiota elliptica* (from Hooker, 1870) 142

Figure 4.6. One of 190 optimal SW trees from the *trnL-F* analysis, with Fitch lengths (above branches; ACCTRAN optimisation) and bootstrap values (below). Branches not present in the Fitch strict consensus tree are indicated by a solid arrow, and those not present in the SW strict consensus tree are indicated by an open arrow. Heuristic search under the Fitch criterion produced 7000 trees with length 220, CI=0.87 and RI=0.86. SW produced 190 trees with length 172.73, CI=0.97 and RI=0.98 (Fitch length, 220) 157

Figure 4.7. One of 3 optimal SW trees from the ITS analysis, with Fitch lengths (above branches; ACCTRAN optimisation) and bootstrap values (below). Branches not present in the Fitch strict consensus tree are indicated by a solid arrow, and those not present in the SW strict consensus tree are indicated by an open arrow. Heuristic search under the Fitch criterion produced 18 trees with length 732, CI=0.66 and

RI=0.76. SW produced three trees with length 359.82, CI=0.87 and RI=0.92 (Fitch length, 732) 158

Figure 4.8. One of 3 optimal SW trees from the combined ITS/*trnL*-F analysis, with Fitch lengths (above branches; ACCTRAN optimisation) and bootstrap values (below). Branches not present in the Fitch strict consensus tree are indicated by a solid arrow, and those not present in the SW strict consensus tree are indicated by an open arrow. Heuristic search under the Fitch criterion produced six trees with length 916, CI=0.72 and RI=0.76. SW produced three trees with length 513.26, CI=0.92 and RI=0.93 (Fitch length, 916) 159

Figure 4.9. Hypothetical biogeographical development of the tribe Phyliceae based on nucleotide sequence data 166

## Chapter Five

Figure 5.1. The Tristan da Cunha Group (distances between islands not drawn to scale). The black spots indicate sites where *Phyllica arborea* was collected (more detail in Figure 5.2). Map taken from Groves (1981) 186

Figure 5.2. Tristan da Cunha. Black spots indicate sites where *Phyllica arborea* was collected. Map taken from Groves (1981). Settlement Plain samples were collected from heights of between 100 and 250m between the Settlement and Burntwood 187

Figure 5.3. New Amsterdam. Black spots indicate sites from where *P. arborea* was collected. Map taken from Tréhen *et al.* (1990) 190

Figure 5.4. Mauritius. *Phyllica nitida* was collected from the Pétrin Nature Reserve. Map taken from White (1983) 192

Figure 5.5. Distribution of collected samples of *Phyllica paniculata* 193

Figure 5.6. UPGMA analysis of the *Phylica* 'island group' with bootstrap percentages shown below branches. Populations: 1. Réunion; 2. Mauritius; 3. High Hill, St Helena; 4. Lot, St Helena; 5. Settlement Plain, Tristan da Cunha; 6. Big Gulch, Tristan da Cunha; 7. Gough; 8. New Amsterdam; 9. The Ponds, Tristan da Cunha; 10. Nightingale; 11. "Hybrid" genotypes 205

Figure 5.7. Parsimony analysis of the *Phylica* 'island group' showing one of the Fitch trees. The length of a five step branch on the tree is indicated. Note the uneven rates of change between lineages. Bootstrap percentages are indicated below branches. Populations: 1. Réunion; 2. Mauritius; 3. High Hill, St Helena; 4. Lot, St Helena; 5. Settlement Plain, Tristan da Cunha; 6. Big Gulch, Tristan da Cunha; 7. Gough; 8. New Amsterdam; 9. The Ponds, Tristan da Cunha; 10. Nightingale; 11. "Hybrid genotypes" 206

Figure 5.8. Neighbour joining analysis of the *Phylica* 'island group' with bootstrap percentages shown below branches. Note the uneven rates of change between lineages. Populations: 1. Réunion; 2. Mauritius; 3. High Hill, St Helena; 4. Lot, St Helena; 5. Settlement Plain, Tristan da Cunha; 6. Big Gulch, Tristan da Cunha; 7. Gough; 8. New Amsterdam; 9. The Ponds, Tristan da Cunha; 10. Nightingale; 11. "Hybrid" genotypes 207

Figure 5.9. Principal co-ordinates analysis of the *Phylica* 'island group'. The percentage of variance is 19.8% in the first axis and 13.2% in the second 210

Figure 5.10. Principal co-ordinates analysis of *Phylica arborea*. The percentage of variance is 19.9% in the first axis and 10.6% in the second 211

Figure 5.11. Principal co-ordinates analysis of *Phylica nitida*. The percentage of variance is 45.9% in the first axis and 30.8% in the second 212

Figure 5.12. Timing of dispersal of island populations of *Phylica* based on sequence and AFLP data 220

## Chapter Six

- Figure 6.1. St Helena. Black spots indicate sites of remaining populations of *Phyllica polifolia* and the site of the last tree of *Nesiota elliptica*. After Cronk (1984) 230
- Figure 6.2. AFLP profiles of 50-100 bp fragments from *Nesiota elliptica* 233
- Figure 6.3. AFLP profiles of 150-180 bp fragments from *Nesiota elliptica* including profiles from the original surviving tree (*Nesiota* 500). The extra bands in the original tree are suspected to have been amplified from fungal contaminants 234
- Figure 6.4. AFLP profiles of 50-100 bp fragments from two populations of *Phyllica polifolia* (first two rows are the Lot population, the second two rows are the High Hill population). Arrows indicate polymorphisms 235
- Figure 6.5. Principal co-ordinates analysis (using Jaccard's similarity co-efficient) of *P. polifolia*. Percentage variance of axis 1 = 43.8 and axis 2 = 24.8 236
- Figure 6.6. Tree for *P. polifolia* taken from the overall neighbour joining analysis on the island group presented in Chapter 5. Bootstrap percentages are shown below branches 237

### List of Plates

- Plate 1. *Phyllica montana*, Cape Province 143
- Plate 2. *P. pubescens*, Cape Province 143
- Plate 3. *P. cryptandroides*, Cape Province 143
- Plate 4. *P. plumigera*, Cape Province 143

Plate 5. *P. ericoides*, Cape Province 143

Plate 6. *P. buxifolia*, Cape Province 143

## Appendices

Appendix 1. Matrix of *trnL-F* sequences with insertions and deletion characters indicated by asterices i

Appendix 2. Binary matrix of AFLP characters (0 = band absent, 1 = band present) xxvi

# **CHAPTER ONE. INTRODUCTION**

## CHAPTER ONE. Introduction

### 1.1. General Introduction

This study was initiated as an investigation into the relationships between continental and island species of the genus *Phyllica* L. (Rhamnaceae). This genus occurs predominantly in the fynbos of Cape Province, South Africa but it is also found on oceanic islands around southern Africa. The project also incorporates a study that determines the phylogenetic context of the genus within the family and a study of population and conservation genetic aspects of some of the island species. Knowledge of the affinities among oceanic island species and their continental relatives can give information about differing evolutionary processes on islands and continents. The production of a robust estimate of phylogeny incorporating such groups is a vital part of these studies. In cases for which the sister group of an isolated taxon is known vicariance biogeography can be studied, but relationships with possible sister groups are often poorly understood due to rapid morphological change caused by adaptive radiation, which results in sister taxa not resembling each other. The production of robust molecular analyses has greatly aided the study of island taxa and their mainland relatives (e.g. Baldwin, 1992; Fay *et al.*, 1997).

An important question in the study of endemic oceanic island taxa is whether these taxa are relicts or products of more recent dispersal events (i.e. palaeoendemics or neoendemics). Lincoln *et al.* (1982) described a neoendemic species as a species having a limited geographical range attributable only to its recent origin and a palaeoendemic as one with a limited geographical range but of considerable evolutionary age (i.e., a relict). Myers and Giller (1988) stated that neoendemic species are those that have resulted from *in situ* speciation and palaeoendemics are species with a formerly wider distribution which have been reduced by *ex situ* extinction. According to Stace (1989) a neoendemic taxon is one that is evolutionarily young that has been unable to spread to other areas and a palaeoendemic taxon is one that is now restricted but once exhibited a far wider distribution. Stace (1989) also identified holoendemics which are not of recent origin but have retained a narrow distribution, (i.e. there has been no range contraction) and

active epibiotics which are palaeoendemic taxa that have recently diverged to produce new species after a long period of range contraction. Cronk (1997) stated that the concept of relict endemism is independent of adaptive radiation on islands. It is concerned with the source and coloniser lineages rather than post-colonisation speciation events. Whether a species may always be described as either neo- or palaeoendemic is not clear, and many taxa fall between these two extremes. In many instances endemic island taxa are rare and under threat of extinction. An understanding of their biological status as palaeo- or neoendemics is considered important if decisions about conservation strategies need to be made (Vane-Wright *et al.*, 1991).

Studies of the adaptive radiation of closely related insular species, which are characterised by high levels of phenotypic diversity, are useful in learning about rates and mechanisms of evolution. Adaptive radiation on islands may be the result of a release from competition and the utilisation of new niches and ecological opportunities. The Hawaiian silversword alliance is an example of a monophyletic group of neoendemic species, which have arisen through adaptive radiation following a single or a few founder events onto an isolated group of oceanic islands. Molecular data have been used in the study of relationships between these island plants and their mainland sister groups. Baldwin *et al.* used plastid RFLPs (1990) and sequences of the internal transcribed spacer (ITS) of nuclear ribosomal DNA (1992) to elucidate relationships between the silversword alliance (*Argyroxiphium*, *Dubautia* and *Wilkesia*) and the Californian tarweeds (*Adenothamnus*, *Madia*, *Raillardiella* and *Raillardiopsis*). The monophyletic, species-rich silversword alliance was found to have a Californian tarweed sister group.

In a study of *Pelargonium* Bakker *et al.* (1998) suggested that *P. grossularioides* from the Tristan da Cunha group is the result of a relatively recent long distance dispersal event. They used ITS rDNA sequences to show that this species was recently derived from within a clade containing the South African species of the genus, i.e. it is a neoendemic taxon.

It may also be argued that some plants may survive unchanged on islands for long periods because of a lack of competition, low rates of immigration of new species and climatic buffering and hence may in some cases be considered palaeoendemic



(e.g. the Canarian genus *Dendrosonchus* of Compositae and *Lactoris* of Aristolochiaceae). Fossil evidence (Cronk, 1990) indicates that the composition of the St Helenan flora has remained in a similar state for the last nine million years. Mainland sister taxa may be subject to extreme events such as glacial cycles or more long-term climatic changes, which could result in either their extinction or adaptive radiation. Cronk (1992) suggested a relictual series of palaeoendemism, the components of which were distinguished by the relative contribution of *in situ* evolution and *ex situ* extinction to the resulting endemism. *Petrobium* (Compositae), *Commidendrum* (Compositae), *Lachanodes* (Compositae) and *Trochetiopsis* (Sterculiaceae) are considered to be examples of palaeoendemic genera on St Helena. There are numerous other examples of palaeoendemic taxa on oceanic islands including *Dendrosicyos* (Cucurbitaceae) and *Socotranthus* (Apocynaceae) from Socotra in the Indian Ocean. Fay *et al.* (1997) used sequence data for the plastid gene *rbcL* (which codes for the large subunit of the photosynthetic enzyme ribulose biphosphate carboxylase oxygenase) to establish the closest relatives of the endangered endemic *Medusagyne oppositifolia* from the island of Mahé in the Seychelles. Morphological and anatomical analyses of this species had failed to firmly establish a phylogenetic link to any other group. The *rbcL* sequence data showed that *Medusagyne* is a member of a monophyletic group also containing Ochnaceae and Quiinaceae. These three, collectively have a pantropical distribution and include numerous taxa that are localised relict endemics (Fay *et al.*, 1997). On the basis of sequence data, morphology and anatomy, *Medusagyne* is considered to be a relict endemic island taxon.

Wider taxonomic application of nucleotide sequence data has been used to determine suprageneric relationships in a number of taxonomically problematic groups (Dipsacales, Donoghue, 1992; Geraniaceae, Price and Palmer, 1993; Cornaceae, Xiang *et al.* 1993; Saxifragaceae *sensu stricto*, Soltis *et al.*, 1993; Droseraceae, Williams *et al.* 1994, Zygophyllaceae, Sheahan and Chase, 1996; Themidaceae, Fay and Chase, 1996; Lecythidaceae, Morton *et al.*, 1997; Plumbaginaceae, Lledó *et al.*, 1998). The suprageneric or tribal classification of Rhamnaceae (Suessenguth, 1953) had previously been based largely on fruit characters which resulted in the circumscription of two large heterogeneous tribes

and three smaller relatively homogeneous tribes. Rhamnaceae were shown to be part of a weakly supported group which also contains Rosaceae, Urticales, and Fagales based on an analysis of sequences of *rbcL* for 499 species of angiosperms (Chase, Soltis, Olmstead *et al.* 1993). Another study using *rbcL* (Soltis *et al.* 1995) indicated a close relationship between Elaeagnaceae and Rhamnaceae. Further analyses using nuclear 18S rDNA (which codes for the small subunit of nuclear ribosomal RNA), plastid *atpB* (which codes for the beta subunit of ATP synthase) and *rbcL* sequence data (Savolainen *et al.*, submitted; Soltis *et al.*, 1998) supported the link between Rhamnaceae and Elaeagnaceae. Sequence data from *rbcL* and plastid *trnL-F* (which consists of an intron and an intergenic spacer between transfer RNA genes) has also placed the families Barbeyaceae and Dirachmaceae in association with Rhamnaceae (Thulin *et al.*, 1998).

The use of nucleotide sequence data is often limited within species by the low levels of sequence divergence between closely related individuals and taxa. This has resulted in the use of other techniques for detecting polymorphism to allow resolution of relationships between close relatives. Relationships between closely related organisms that are still interbreeding are complicated by gene flow. Attempts at solving the problems in determining relationships between close relatives and in linking population genetics with phylogenetics have so far proved unsatisfactory because the methods and the markers used in either discipline cannot be readily applied to the other. I investigate here the potential of a fingerprinting method (amplified fragment length polymorphism; AFLP; Vos *et al.*, 1995) for determining relationships between some closely related island species of *Phyllica*. The use of this method in determining levels of genetic variability and the application of such information to assessing conservation priorities in some endangered island species is also investigated.

## **1.2. Comparison of the Use of Molecular and Morphological Data in Systematics**

Morphological and molecular characters each have advantages and disadvantages when being used in the reconstruction of phylogeny or assessing variation between populations. Because there are differences, the use of both types of data will

maximise the amount of information and therefore produce more robust overall estimates of relationships.

Potentially all nucleotides in the DNA of an organism are useful characters. At present only a small amount of this potential has been sampled. A molecular approach is useful in cases in which morphological variation is limited. Different parts of the genome evolve at different rates and therefore can be used to answer questions about evolution in different levels of the taxonomic hierarchy ranging from recent changes within populations to the origin of life on earth. Phylogenetic characters have to be heritable, and molecular characters fit this criterion. Many morphological characters are quantitative and are difficult to code in phylogenetic analyses whereas the majority of molecular characters are qualitative or discrete and easier to code. Once the infrastructure is in place, large amounts of molecular data can be gathered in a relatively short time. Most studies have looked at sequences of single loci in the genome leading to the production of gene trees which may not be representative of the organism as a whole. Ideally studies should incorporate more than a single region in the genome and also different genomes within individuals, i.e. plastid and nuclear regions.

Morphology is readily studied using herbarium specimens. The DNA in these specimens often does not persist as well as the morphological features. However, methods for extracting and sequencing DNA from dried specimens have improved greatly in recent years. Morphological methods are also cheaper than molecular methods. In some cases morphological analyses are hindered by a lack of characters suitable for phylogenetic studies. Also, environmental effects are often non-heritable (unless garden or reciprocal transplant experiments are performed this is impossible to assess). Morphological evolution may obscure phylogenetic relationships that can be determined by looking at molecular data. Some morphological characters may evolve at a faster rate than molecular characters in response to stronger selection pressures resulting in parallel evolution of similar character states in different phyletic lines. In other words morphological characters are often not selectively neutral in the way that molecular characters are often reputed to be. Also a small genetic change can result in large phenotypic differences in characters such as flower colour and shape.

### **1.3. Introduction to Cladistics and Molecular Systematics**

Systematics is the term given to the process of detecting, describing and explaining diversity in the biological world. Linnaeus formulated a hierarchical system of classification (1758) prior to the development of theories of evolution. This hierarchical system has subsequently been found to be useful within the contexts of evolutionary theory. Attempts at objective methods for estimating phylogeny based on shared attributes were formulated by Zimmermann (1930; 1931; 1934; 1943), Hennig (1950; 1966) and Wagner (1961). Accurate estimates are necessary to provide the basis for studies that would answer a range of questions about biological change.

#### 1.3.1. Introduction to cladistics

Classifications have been produced for many purposes. Special purpose or artificial classifications utilise one or a few characters. An example of this is Linnaeus' sexual system (1753) based on number of floral parts. This resulted in species from different families being placed in the same group. A general purpose classification is one that utilises many characters and groups together plants having many attributes in common. As more information becomes available, the chances of a natural classification being produced increases. Turrill (1940) introduced the idea of an early 'alpha-' taxonomy which may be successively modified in the light of new information to produce improved 'omega-' taxonomies. Stace (1989) stated that:

“‘omega-taxonomy’ is almost by definition unattainable, but it is the distant goal at which taxonomists should aim” (p. 20).

Special purpose classifications may still be produced which focus study on the development of a particular character, but those utilising a large number of characters are of more general use.

As well as the increasing availability of character information, there has also been a continuous development of ways in which this information is treated. The biggest development was the introduction of the use of computers in the 1960s. Numerical

taxonomists attempted to produce natural classifications using objective methods. Phenetic classifications were produced on the basis of overall similarity between living plants with equal weight being applied to all the characters (and character states) used. The phylogeny of the group could be inferred from the resulting classification, but estimates of phylogeny did not play a necessary role in its production.

Attempts at the modelling of phylogenetic patterns also became increasingly possible. Phylogeny had previously been inferred intuitively. The aim with computers was to produce analyses using objective procedures. Cladistic methodology was first introduced by Hennig (1950) in the book *Grundzüge einer Theorie der Phylogenetischen Systematik* later translated into English in 1966 under the title *Phylogenetic Systematics*. Mayr (1969) coined the term cladism or cladistics. One of the principal aims of cladistics is to determine monophyletic groups on the basis of shared, derived character states. Monophyletic groups are those which arise by the diversification of a single ancestor. Polyphyletic groups are those arising from more than one ancestral group and paraphyletic groups possess a single ancestor in common but do not include all the descendants of that ancestor.

### 1.3.2. Cladistic characters and homology

In cladistic analyses the polarity of change in a group of organisms may be determined, i.e. different character states are assigned primitive or derived status. Shared derived character states are termed synapomorphies, shared ancestral character states are symplesiomorphies and a unique derived character state is an autapomorphy. Prior to determining whether character states are primitive or derived it is vital to determine homology. Homology is similarity due to common descent and is usually considered to be synonymous with synapomorphy. Analagous structures are similar in appearance or function but have different origins, e.g. phyllodes are analagous to leaf-blades but are derived from petioles. Independent lineages may evolve characters or character states that are similar but not homologous. Homoplasy is character conflict within an analysis resulting from misidentified homologies. Homoplasy arises through character state reversal,

character state convergence or parallelism. Parallel or convergent evolution of character states that are analagous may result in the identification of a group which is seemingly monophyletic but in reality is polyphyletic.

Characters can be discrete (qualitative) or continuous (quantitative). For example DNA nucleotide sites are discrete characters whereas DNA:DNA reassociation kinetic studies yield continuous characters. Continuous characters need to be coded into discrete character states for cladistic analysis which is problematic because it is not always clear where to draw boundaries between character states.

Analysis of a data matrix can either result in the production of an unrooted tree (network) or a rooted tree (cladogram). The rooting of an unrooted tree imparts polarity on at least one character transformation. Rooting is usually achieved by outgroup comparison which involves the choice of the sister group or another closely related taxon. The inclusion of an outgroup in an analysis roots a cladogram and determines monophyletic groups and apomorphic and plesiomorphic character states. The assignment of an outgroup is an assumption made outside of the analysis itself.

### 1.3.3. The use of molecular characters in phylogenetic studies

The development of the polymerase chain reaction (PCR; Kleppe *et al.*, 1971; Mullis and Faloona, 1987) has resulted in large amounts of data being made available for DNA sequencing and DNA fingerprinting techniques such as microsatellites (Weber and May, 1989) and RAPDs (Williams *et al.*, 1990). DNA sequences provide us with precisely comparable characters that can be used to examine mechanisms of evolution of molecules by using knowledge of evolutionary history of species. The evolution of molecules can conversely be used to infer the evolutionary history of taxa. The greater availability of molecular data has resulted in improvements in the analysis of such data with the result that the development of phylogenetic analysis as a whole has expanded greatly.

#### 1.3.4. Homology of molecular characters

It is necessary to define two types of homology when referring to sequence regions. Two sequences are said to be orthologous if they can be traced back to a speciation event. If the common ancestry of the sequences can be traced back to a gene duplication event they are said to be paralogous. Only orthologous sequences should be used to infer phylogeny of species. Paralogous sequences within the same genome will evolve along parallel lines, but they are only sources of data for comparative studies if both copies can be identified and analysed separately.

A further level of homology must also be recognised as a potential problem. Once orthology of two sequences has been confirmed it is necessary to confirm the positional homology of individual nucleotides. This is usually not a problem when comparing protein-coding sequences, but insertion and deletion events in non-coding regions can result in uncertainty over the homology of individual nucleotide sites.

Phylogenetic analysis of orthologous sequences results in the production of gene trees (Doyle, 1992). A major concern is whether these gene trees reflect the true overall phylogeny of the organisms under study (Pamilo and Nei, 1988). Retention of ancestral polymorphisms, hybridisation or horizontal gene transfer can result in differences between a gene tree and the organismal phylogeny. Surveying a large number of loci dispersed throughout the genome is more likely to detect evidence of reticulation. If two types of data produce results that are incongruent then it is necessary to explain why.

#### 1.3.5. Methods for inferring phylogeny

The methods for building phylogenetic trees can be divided according to the type of data used, i.e. distance or discrete data.

##### 1.3.5.1. Distance data

Distance methods calculate the genetic distances between pairs of taxa by measuring the amount of evolutionary change between them. A tree is produced from

a matrix of pairwise distances between the taxa. Sequence data could give distances based on the fraction of sites that differ between the two sequences. Examples of distance methods include:

1. UPGMA or average linkage method (Sokal and Sneath, 1963). This method assumes a molecular clock, i.e. a constant rate of evolution in different lineages. Because this method does not take into account rate heterogeneity it can produce an incorrect topology if some lineages have evolved faster than others.
2. Distance Wagner method (Farris, 1972). Farris argued that because of the likelihood of rate heterogeneity among phyletic lines it is not advisable to use phenetic similarity clustering techniques to estimate evolutionary trees. Farris's method was originally applied to immunological distance data and takes into account rate heterogeneity over different phyletic lines.
3. Li's Method (Li, 1981). This method is similar to UPGMA but it also corrects for unequal rates of evolution.
4. Modified Farris method (Tateno *et al.*, 1982). Farris's method ignored stochastic effects and it therefore led to overestimates of branch lengths. Tateno *et al.* argued that their method reduces the effect of random errors.
5. Neighbour joining (Saitou and Nei, 1987). This method operates on the same principle as the minimum evolution method but a comparison of different topologies is built into the algorithm. The principle is to find pairs of OTUs (neighbours) that minimise the total branch length at each stage of clustering of OTUs starting with a star-like tree.

The running time of distance methods increases more slowly with added taxa than discrete methods (Felsenstein, 1984). However there is a loss of information in transforming sequences to distances, and it is unclear what the distances mean biologically.



### 1.3.5.2. Discrete data

Other methods use discrete characters to infer evolutionary change (e.g. character state changes such as nucleotide substitutions) directly on trees. The ancestral states of taxa can be inferred, and the amount of evolutionary change that has taken place can be determined. These methods operate directly on the characters rather than on pairwise distances between taxa. A loss of information can occur when converting characters into distances. There are two main ways of using discrete characters, maximum likelihood and maximum parsimony.

#### 1.3.5.2.1. Maximum parsimony

Maximum parsimony selects as optimal the tree or trees that require the fewest changes. The most parsimonious tree minimises the number of ad hoc hypotheses required to explain the occurrence of homoplasy. Parsimony maximises the amount of evolutionary similarity that can be explained as homologous similarity, i.e. due to common ancestry. Edwards and Cavalli-Sforza (1964) introduced the concept of a "method of minimum evolution", Camin and Sokal (1965) introduced the term parsimony into systematics and the principles of parsimony were first applied to the evolution of molecular sequences by Eck and Dayhoff (1966). Parsimony makes few assumptions about the evolutionary process, it has been extensively studied mathematically and it does not require powerful hardware. Problems with parsimony include the fact that it does not use all the available information (i.e. it ignores what under the assumptions of parsimony are considered to be uninformative sites) and it is supposedly inconsistent, i.e. when rates of change are unequal it doesn't always converge on the right answer as more data are added (Felsenstein, 1978). However, Graybeal (1998) has demonstrated that the accuracy of reconstruction of a four taxon tree using parsimony improved dramatically with the addition of more taxa and also improved with the addition of more characters. Parsimony is also supposedly only reliable when rates of change are slow. However, Hillis (1998) simulated an increase in the expected amount of change along all branches of a particular tree and

demonstrated that various methods for inferring phylogeny, including parsimony, performed better when rates of change were higher.

There are three steps to finding the most parsimonious tree: 1. Determining the optimality criterion used to infer the tree that specifies the restrictions imposed on character-state changes; 2. Specifying the algorithm that is used to search for optimal trees under the conditions imposed by the optimality criterion and 3. The measures used to evaluate the result. Optimality criteria are discussed here and the latter two steps are discussed in the methods for Chapter Two.

Choice of parsimony optimality criterion can depend on the kind of data being analysed. The following optimality criteria have been described:

1. Wagner Parsimony (Wagner, 1961). For a binary character a change from state 0 to state 1 is given equal weight to a change from state 1 to state 0. This means that an unrooted tree can be rooted at any point without changing its length.

2. Fitch Parsimony (Fitch, 1971). Characters with three or more states are unordered, i.e. they can be transformed directly into any other state. This criterion was formulated for DNA sequences which have four character states.

Wagner and Fitch parsimony criteria are appropriate whenever the probabilities of any character state change are unknown or where they are symmetrical i.e. a change from 0 to 1 has the same probability as a change from 1 to 0. Only the Fitch criterion is appropriate for DNA sequences.

3. Dollo Parsimony (Dollo, 1893). This is appropriate when the probability of a reverse change (1 to 0) is zero. In other words character polarity is specified. Every derived character state is uniquely defined (parallel gains of the derived condition are not allowed). All homoplasy must be accounted for by reversal and not parallelism. DeBry and Slade (1985) considered Dollo parsimony was appropriate for analysing restriction fragment data because the probability of gaining a new site is a lot less likely than that of losing an existing site, but Dollo parsimony is too extreme because it permits no parallelism.

4. Camin-Sokal Parsimony (Camin and Sokal, 1965). Character evolution is irreversible (equivalent to ordered but not reversible). Under this criterion all homoplasy must be accounted for by parallel or convergent change. Characters

optimised under Dollo or Camin-Sokal parsimony criteria are examples of directed characters.

#### 1.3.5.2.2. Maximum likelihood

This method chooses the tree that maximises the likelihood, or the probability that the observed data would have occurred. In DNA sequence data nucleotides at each nucleotide site are considered separately, and the log likelihood for having these nucleotides are computed for a given topology by using a particular probability model. This log likelihood is added for all nucleotide sites, and the sum of the log likelihood is maximised to estimate the branch length of the tree. This procedure is repeated for all possible topologies, and the topology that shows the highest likelihood is chosen as the optimal one. Edwards and Cavalli-Sforza (1964) were the first to attempt applying maximum likelihood to estimating phylogenies using gene frequency data. Felsenstein (1981) gave methods for computing the likelihood of a tree with an arbitrary number of species, and of finding branch lengths that maximise the likelihood. Problems with maximum likelihood arise due to computational intensity (matrices containing more than 40 taxa cannot be analysed) and because there is empirical evidence refuting all molecular models (Savolainen *et al.*, submitted; Siddall and Kluge, 1997). Maximum likelihood methods are based on explicit models of evolutionary change. They make more complete use of all available information i.e. all sites are informative and they are supposedly more consistent and efficient than parsimony (Felsenstein, 1988). However, Siddal (1998) demonstrates cases in which maximum likelihood is inconsistent and inaccurate. Also, maximum likelihood requires an explicit model of evolutionary change and the methods are therefore supposedly more 'assumption laden' than parsimony. There is also a lack of empirical evidence to support proposed models of evolution. Also, these methods are relatively slow because currently available hardware is not powerful enough to deal with large data sets.

### 1.3.6. Choice of method used to analyse sequence data

I chose to use discrete methods to analyse my data sets because the use of distance methods involves a loss of information when converting character state matrices into distance matrices. Maximum parsimony was chosen to analyse my data because it makes only a few assumptions about the evolutionary process, has been extensively studied mathematically and does not require powerful hardware. Maximum likelihood methods require an explicit model of evolutionary change and are thus more assumption laden than parsimony, and there is also a lack of hardware that is powerful enough to be able to deal with the large data sets. The Fitch criterion was used in this study because it makes only one assumption about the probability of change, i.e. that there are no lineage specific rate biases.

### **1.4. Assessing Variation at Species Boundaries and Among and Within Populations**

So far I have concentrated on the use of nucleotide sequence data in the reconstruction of phylogeny. If individuals or groups of individuals are interbreeding, these methods are hindered by insufficient sequence divergence and are complicated by recombination. As a result of these difficulties population genetics (which includes the study of groups or individuals still interbreeding) and phylogenetics (which includes the study of reproductively isolated taxa forming unique lineages) have largely persisted as separate disciplines even though speciation processes falls between the two. Attempts to bridge the gap have been made either from a systematic standpoint or a population biology standpoint. Problems have arisen when students of a particular discipline have attempted to bridge the gap because the two use different terminology, which is not surprising since the patterns and processes are different.

Phylogenetic relationships can only be determined when two taxa are isolated, i.e. they are not interbreeding. Bifurcating trees cannot be produced because mating between terminals complicates the patterns produced. There are no algorithms currently available to elucidate reticulate branching patterns. If there is mating taking

place, there is no phylogeny. The use of parsimony in these individuals will produce many trees with mutually incongruent topologies (i.e. polytomies in the strict consensus tree). Distance methods would also detect little or no population structure if panmixis occurs. Different methods of analysis need to be considered for the study of patterns and processes of molecular changes within taxa that are still interbreeding as opposed to those which are independent. This project encompasses molecular studies from suprageneric to population level and has provided the opportunity to assess some of the different molecular techniques available at each of these levels. There are a variety of types of molecular data presently being used but it is not clear whether any of these can be used in both population and phylogenetic studies.

### **1.5. Molecular Markers in Population Genetics**

The following sections review the use of molecular markers in population genetics. The ideal molecular marker should be highly polymorphic, co-dominantly inherited, frequently and evenly distributed throughout the genome, easily and quickly assayed, highly reproducible and easily exchanged between laboratories.

Co-dominantly inherited markers allow the distinction of homo- and heterozygotic states in diploid organisms, which may then be used to interpret population genetic structure via models such as the Hardy-Weinberg equilibrium. If this equilibrium is not in effect, this indicates that phenotypic variation has a non-genetic basis, individuals may not be randomly mating, some selection is present, or there may be migration into the study population from neighbouring sites. Dominantly inherited markers do not allow distinction between homo- and heterozygotic states and therefore cannot be used to evaluate population genetic processes in as much detail as co-dominantly inherited markers. Structure (i.e. the distribution of genotypes) can be detected using dominant markers, and many studies ask questions that only require knowledge of how populations are inter-related. In these cases the use of dominant markers or unknown mixtures of dominant and co-dominant markers are acceptable. The principle aim of the study of infra-specific variation in island species undertaken here is to determine whether there is any detectable structure and if variation in genotypes is discovered how this is

distributed. Markers should be frequently and evenly distributed throughout the genome in order to get an adequate estimate of genotypic variation.

#### 1.5.1. The use of proteins in population genetic studies

Protein studies involve the utilisation of varying electrophoretic mobilities with different primary structure or peptide sequence. Isozymes are functionally similar forms of enzymes, including all polymers of subunits produced by different gene loci or by different alleles at the same locus. Allozymes are a subset of isozymes that are variants of polypeptides representing different alleles at the same locus. The use of isozymes in plant systematics is reviewed in Crawford (1989).

Allozymes are good markers because they are co-dominantly inherited, easy, quick and cheap to assay and highly reproducible. They are reliable and have a well documented history of high performance. However, they are limited to a small part of the genome even if a large number of systems are investigated, so they will consistently underestimate genotypic variation in a population. Also they are not necessarily selectively neutral. Bands with identical electrophoretic mobility cannot be assumed to represent identical alleles if species are distantly related. Changes in nucleotide sequence may have no effect on isozyme phenotype i.e. the amino acid does not change. For example, the F and S alleles of the *Adh-1* gene in maize showed many differences in sequence rather than a single base-pair substitution as had been previously postulated (Sachs *et al.*, 1986). This means that allozymes always underestimate the degree of genotypic variation present within a population.

#### 1.5.2. The use of DNA in population genetic studies

The use of DNA in population genetic studies has a number of advantages over proteins. The genotype rather than the phenotype is assayed, which means that changes in nucleotide sequence are detected which may not have any effect on the phenotype i.e. amino acid sequence, so assessing only phenotypes leads to substantial underestimates of genotype variation and population structure. One or more sequences appropriate to the problem can be selected on the basis of

evolutionary rate or mode of inheritance. The methods are usually general to any type of DNA. DNA can be prepared from small amounts of tissue and is relatively stable, even in non-cryogenically stored tissues. DNA markers covering large parts of the genome can be found whereas allozyme markers focus on individual loci in the nuclear genome.

Nucleotide sequence data can potentially be used to investigate patterns of variation within plant populations as well as between species. Nucleotide sequence data can be studied directly or indirectly. Indirect methods such as analysis of restriction fragment patterns can provide estimates of DNA variation over entire genomes. Direct methods such as sequencing focus on a particular gene or non-coding region which are often not polymorphic enough to resolve relationships among close relatives. Fragment analyses tend to be cheaper and faster than sequence analyses, allowing large numbers of individuals and loci to be screened.

#### 1.5.2.1. Homology of DNA segments and alleles

One of the major problems with the comparison of DNA fragments is determination of homology. Two fragments that have identical mobility are generally assumed to be homologous stretches of DNA. However, fragments of identical length may be from a totally different part of the genome and have entirely different sequences. Homology of fragments from different organisms can be verified either by using the fragment from one organism as a hybridisation probe against the other fragment, by cleaving gel isolated products with restriction enzymes and observing band profiles or by sequencing the fragment. The characters (fragments) need to show enough variation to allow population or phylogenetic analysis but not so much that the level of ambiguity in the homology of fragments is unacceptable. Generally, if the individuals being screened are closely related, estimates of homology are not problematic, but at some unknown level of divergence homology becomes more difficult to assess. Collection of other data types should reveal bands that are incorrectly assessed so that major problems occur only when too few data are being collected.

### 1.5.2.2. Restriction fragment length polymorphisms (RFLPs)

There are two approaches to RFLPs. Either, digestion of total DNA with a restriction enzyme followed by gel electrophoresis, Southern blotting of the gel and hybridisation on the blot using labelled probes or PCR amplification of specific DNA sequences followed by restriction digestion and gel electrophoresis. Restriction of total DNA often produces so many fragments that individual homologous bands have to be identified with a probe. Cloned segments of conservative parts of ribosomal genes hybridise to homologous regions from many species and have been used to demonstrate restriction site variation in nuclear ribosomal DNA within and among populations (Schaal and Learn, 1988). Restriction site variation has also been demonstrated in plastid genomes (e.g. Riesberg *et al.*, 1988). Probes can also be from the genome that is to be analysed ('homologous probes') or from related species ('heterologous probes'). Nuclear RFLP markers can be treated as co-dominant if the study of restriction fragments involves the use of known probes that hybridise to these fragments, thus allowing all alleles to be determined.

RFLP polymorphism should be due to substitution in a restriction site resulting in the gain or loss of a restriction site, but it is often instead due to insertions or deletions, which is one of the reasons why parallel site gains and losses are more frequent than predicted in many studies (Chase and Palmer, 1989). Advantages of RFLPs include the fact that they are often co-dominant markers they are highly reproducible and they are often evenly distributed throughout the genome. However, study of RFLPs of total DNA requires a good supply of probes, and if heterologous probes are unavailable, cDNA or genomic DNA probes must be developed. Also, blotting and hybridisation techniques are time consuming and difficult to automate, and large quantities of good quality DNA are required. Data from RFLP analyses are also difficult to exchange accurately between laboratories.

The use of RFLPs at the level of populations and individuals has been reviewed by Bachman (1994) and Qamaraz-Zaman (1998). Riesberg *et al.* (1988) undertook a molecular re-examination of introgression between *Helianthus annuus* and *H. bolanderi* (Compositae) and distinguished between wild and serpentine races of the



wild sunflower using RFLPs of nuclear DNA. Plastid DNA was also found to be useful in the same group to distinguish between these two races even though cpDNA generally evolves at a slower rate than nuclear DNA. Jansen and Palmer (1988) used plastid RFLPs to demonstrate the parphyly of the tribe Mutiseae (Asteraceae).

#### 1.5.2.3. Variable number tandem repeats (minisatellites and microsatellites)

There are two classes of variable number tandem repeat (VNTR). Minisatellites are short tandem repeated sequences of more than eight basepairs in which the number of repeats between flanking restriction sites is highly variable. Microsatellites are shorter two to eight basepair repeats, which are variable in number. This variation in the number of repeats causes variation in the length of restriction fragments containing the repeats. Minisatellite loci are usually examined in multi-locus profiles via hybridisation methods. Microsatellite loci are usually examined one at a time via PCR. Specific primers for unique locus specific sequences flanking a VNTR are designed and used to detect length alleles of individual VNTR loci. Plastid VNTRs are distinct from nuclear VNTRs, and their use is also very different (see examples below).

VNTRs are extremely variable and have many alleles at each locus, and so they can therefore be used to detect close relatives. They are also co-dominant and automatable if PCR based. However, they can require a relatively large amount of DNA. They often require a labelled probe and produce anonymous bands (if they are not PCR based). One probe can be used to detect VNTRs at many highly variable loci in the genome. This can produce a many-banded DNA fingerprint, but the homology of bands cannot be definitely proved by hybridisation with a common probe because the repeat sequences are ubiquitous on account of their short length. The examination of single microsatellite loci via PCR requires the design of primers that are specific to the organisms in the study which can be time consuming and expensive.

Rogstad *et al.* (1988a) used a human minisatellite probe to reveal RFLPs among individuals of *Populus deltoides* and *P. tremulodes* and (1988b) M13 phage probes to detect DNA minisatellite-like sequences in gymnosperms and angiosperms. Weising

*et al.* (1989) demonstrated the presence of polymorphic simple GATA/GACA repeats in plant genomes. Weising *et al.* (1991) then developed plant DNA fingerprinting with radioactive and digoxigenated probes complementary to simple repetitive DNA sequences.

Polymerase chain reaction of specific microsatellite loci has been used to map polymorphisms in the human and rodent genomes (Weissenbach *et al.*, 1992; Serikawa *et al.*, 1992). Microsatellites seem to have a relatively low abundance in plant genomes, however methods for efficient isolation of microsatellites are now available (Edwards *et al.*, 1996).

Strieff *et al.* (1998) assessed within-population genetic structure in *Quercus robur* and *Q. petraea* using isozymes and microsatellites and used these data to cautiously conclude that greater seed dispersal in *Q. robur* has led to a weaker spatial genetic structure in this species compared with *Q. petraea*. Vendramin and Ziegenhagen (1997) have identified polymorphic plastid microsatellites in *Abies* for use in paternity studies. Plastid microsatellites have also revealed population genetic diversity in red pine, *Pinus resinosa* (Echt *et al.*, 1998) a species which has not shown any allozyme diversity and very little RAPD diversity. When using plastid microsatellite data in phylogenetic studies it is important to be aware of the possibility of size homoplasy as demonstrated by Doyle *et al.* (1998) in wild perennial relatives of soybean (*Glycine* subgenus *Glycine*) in which fragments of the same size were found to be non-homologous.

#### 1.5.2.4. Randomly amplified polymorphic DNA (RAPD)

This method developed by Williams *et al.* (1990, 1993) involves amplification of DNA between two primer sites by PCR using single arbitrary short primers. This procedure relies on the chance that the complementary primer sites occur somewhere in the genome as inverted repeats enclosing a relatively short stretch of DNA. This may produce a series of DNA fragments that can be separated by gel electrophoresis on an agarose gel and visualised by staining with ethidium bromide. The levels of polymorphism produced by the method may be adjusted by using different primers. RAPD polymorphisms should be due to substitutions in primer sites causing loss of

bands, length variation between primer sites or sequence rearrangements, i.e. inversions or translocations, but it is often instead due to insertions or deletions (Chase and Palmer, 1989).

Advantages of RAPDs include the fact that they produce low to moderate levels of polymorphism, are likely to be evenly and frequently distributed throughout the genome, are probably selectively neutral, have no requirement for DNA probes or sequence information for the design of specific primers, are technically simple, require small amounts of genomic DNA and are automatable. However, the amplification products are anonymous pieces of DNA, which could potentially have been amplified from any organic source. This problem applies to any technique that employs PCR (e.g. microsatellites, AFLPs). They also suffer from amplification irregularities because varying PCR conditions can produce different banding patterns. Also, homology of co-migrating bands is uncertain. There could also be length alleles at homologous sites, i.e. bands that migrate at different speeds, which are in fact homologous. There is also the possibility of the presence of paralogous loci i.e. multiple homologous RAPD sites in a genome (various members of a gene family). Unless RAPDs are run on an automated sequencer with size standards in each lane they are hard to exchange between laboratories.

Crawford *et al.* (1991) studied *Lactoris fernandeziana* (Aristolochiaceae) on the Juan Fernandez Islands using enzyme electrophoresis. They studied 83 plants in 12 populations of this polygamo-dioecious shrub of the island Masatierra in the Juan Fernandez Archipelago using 22 allozyme loci and found no variation. Brauner *et al.* (1992) looked at ribosomal DNA and RAPD variation in *L. fernandeziana*. Twenty seven plants from 15 populations were examined for RFLPs in the 18S-25S rDNA and for RAPDs. Three length variants and four restriction site variants were found in the 18S-25S rDNA. Of 106 RAPD bands per plant produced with 16 primers, 26 were polymorphic. RAPDs were therefore considered to be more effective in finding residual variation than isozymes or RFLPs of rDNA.

Van Heusden and Bachmann (1992a,b,c) looked at three annual species in Asteraceae: *Microseris elegans* and *M. bigelovii* from North America and *M. pygmaea* from Chile and attempted a cladistic analysis, which they thought feasible because of the inbreeding, almost clonal, nature of the populations. The *M. elegans*

populations containing closely related biotypes were found to be interspersed with genetically very different plants. The Chilean populations of *M. pygmaea* were suggested as being the result of long distance dispersal from North America with spread from the point of establishment into two genetically isolated series of populations, one coastal and one inland. *Microseris bigelovii* is distributed along the Pacific Coast from southern California to mid-Oregon with disjunct populations near Victoria, British Columbia, which were suggested to be the result of a single colonisation event. RAPD markers were randomised amongst the closer populations to produce a polytomy. Therefore gene flow was thought to be rare enough to allow local populations to evolve characteristic biotypes through inbreeding and selection but still sufficient to randomise allele distributions throughout the range.

#### 1.5.2.5. Amplified fragment length polymorphisms (AFLPs)

Amplified fragment length polymorphisms (AFLPs; Vos *et al.*, 1995) are a multilocus DNA fingerprinting technique. The use of AFLPs is based on the selective PCR amplification of restriction fragments from a digest of total genomic DNA. This process involves 3 main steps: 1. restriction of DNA and ligation of oligonucleotide adaptors; 2. selective amplification of some of the restriction fragments; 3. gel analysis of the amplified fragments. Selective PCR is achieved using primers with a target site consisting of the adaptor and restriction site. Selective PCR is carried out using primers that extend from the restriction fragment sequence and thereby only amplifying fragments that match this extension. Two rounds of PCR are carried out each decreasing the number of fragments amplified. The second round utilises dye-labelled primers that may be visualised on polyacrylamide gels using an automated format. AFLPs produce 10-100 times more markers per primer than some other fingerprinting techniques such as RAPDs. AFLPs therefore screen loci faster than isozymes, RAPDs and RFLPs. There is no chance of primer mismatches using this technique, and therefore unlike RAPDs, AFLP fingerprinting is reproducible between labs (Jones *et al.*, 1997). AFLPs do not require the design of specific primers. Once all equipment is in place (i.e. automated sequencer) a large amount of data can be generated in a small amount of time. There

is greater accuracy in sizing of bands due to size standards being run in each lane. AFLPs have been shown to be distributed throughout the rice genome and not confined to any chromosome or chromosomal region (Zhu *et al.*, 1998), and there is no reason to suspect that they would not have similar distributions throughout other plant genomes. Disadvantages of AFLPs include the fact that they are dominant markers, are technically more demanding and require slightly more DNA than RAPDs. However, the large number of bands gives a good measure of variation across the genome, which may be all that is required if population structure is the question of interest.

Kardolus *et al.* (1998) investigated the potential application of AFLPs in biosystematics to *Solanum* (Solanaceae) taxonomy in a study of *Solanum* section *Petota*. Quantitative morphological characters and geographical distribution had been used to group taxa. Phylogenetic analysis of this group was difficult because there are few easily scorable qualitative characters and hybridisation and polyploidisation have also made species boundaries unclear. An increase in the number of AFLP fragments with ploidy level was discovered. Inbreeding genotypes had lower levels of polymorphism than outbreeders. Different primer combinations produced more or less the same topology, and the different methods of analysis also produced similar topologies. The high level of variation detected in one of the outbreeding species introduced some conflict in the interspecific analysis. The heterozygosity of *S. microdontum* lead to clustering of its individual genotypes between OTU's of species of a different section. They concluded that they needed to sample more than the one genotype from what is a variable population to get a more conclusive result. They also stated that biosystematic analyses based on molecular markers such as AFLPs are more informative and reliable than those based on morphological traits because of the abundance of discrete binary characters obtained and the exclusion of environmental factors having a substantial influence on quantitative characters.

Roupe van der Voort *et al.* (1997) looked at the use of allele specificity of co-migrating AFLP markers to align genetic maps from different potato genotypes. They sequenced co-migrating fragments, and 19 out of 20 were found to be identical indicating that most co-migrating bands in this study were homologous. Van Eck *et*

*al.* (1995) showed that AFLP markers map genome-wide, hitting several loci on all 12 linkage groups of potato with every primer combination tested.

## 1.6. Conclusions and Choice of Methods

There are advantages and disadvantages in all of these techniques, and given an adequate amount of time a combination of approaches would be the most desirable option. However, the time limit on this particular study was a factor in the choice of technique. AFLPs were chosen because they are highly polymorphic, there is greater control over the degree of polymorphism and they are found throughout the genome. Also they have been shown to be reproducible between labs (Jones *et al.*, 1997), and they do not require the design of specific primers. The main aim of the population genetic aspect of this project was to determine the spatial distribution of genotypes, which could not be resolved using DNA sequences due to lack of polymorphism. This did not require the use of co-dominant markers.

## 1.7. Phylogenetics of Rhamnaceae and *Phylica* L

*Phylica* L. (Rhamnaceae) is an interesting genus as a case study in assessing relationships between oceanic island and continental taxa. According to the last revision by Pillans (1942), *Phylica* consists of about 150 species, most of which are found in Cape Province, South Africa. A number of species are found on islands around southern Africa such as St Helena (*P. polifolia*), Tristan da Cunha and New Amsterdam (*P. arborea*), Mauritius and Réunion (*P. nitida*) and Madagascar (*P. emirnensis* and *P. bathiei*). *Phylica* has two closely related genera *Nesiota* Hook. f. (a monotypic genus from St Helena) and *Noltea* Reichb. (a monotypic genus from South Africa). Most of the species on the mainland are ericoid shrubs, whereas some of the island species and the genera *Nesiota* and *Noltea* are broad-leaved trees or shrubs that have retained other putatively primitive characteristics.

A study of other genera in Rhamnaceae was undertaken to ascertain the evolutionary context of the genus *Phylica* within the family and to determine the sister group to *Phylica* so that this group could be used as an outgroup for the study

of the genus. Rhamnaceae are a cosmopolitan family of about 50 genera and 900 species. They are a good example of a group that requires extra data because of the problems associated with a classification system based on a small number of morphological characters. Suessenguth (1953) divided the family into five tribes largely on the basis of fruit characters, and three of these appeared to be natural groups based on several characters. On the basis of morphological characters however, the two largest tribes appeared to be fairly heterogeneous. Additional data in the form of DNA sequences were desirable to confirm the monophyly of these tribes and the monophyly of the family.

#### 1.7.1. Aims of Rhamnaceae and *Phylica* phylogenetic study

1. To determine whether Rhamnaceae are monophyletic.
2. To determine relationships among genera in Rhamnaceae. Are Suessenguth's tribes monophyletic?
3. To determine the sister group of *Phylica*.
4. To investigate the biogeography of *Phylica*.
5. To determine whether the genus originated in Africa or on the islands and if on the islands, on which island did it arise.
6. To determine whether the island species of *Phylica* are palaeo- or neo-endemic taxa.
7. To determine the nearest mainland relatives of the island species.
8. To determine whether the island taxa are monophyletic.
9. To determine how many species there are on the islands and their biogeographic history.

A morphological phylogenetic analysis of Rhamnaceae was also undertaken. The aim of this study was to demonstrate the relative usefulness of morphological and molecular characters in phylogenetic reconstruction of Rhamnaceae.

## **1.8. Population Level Studies on Island Species of *Phylica* and their Mainland Relatives**

As well as determining the closest mainland relatives of the island species of *Phylica* I have studied how molecular variation is partitioned between and within populations of some of the island species. The origin of island species and populations is of interest. Given the isolated position of the islands, it is possible that some of the island populations were derived from single introductions. Knowledge of the genetic variation within these species is also of interest with regard to the conservation status of those that are endangered. I have used a DNA fingerprinting technique (AFLPs) to attempt to deduce relationships among the island species, among populations of these species on different islands and within populations to answer some of these questions. The effectiveness of AFLPs in answering these questions will be assessed.

### 1.8.1. Aims of population level study

1. To resolve relationships between island species.
2. To determine how genetic variation is structured within and between populations of island species of *Phylica*.
3. To determine the origins of island populations.

## **1.9. Conservation Genetics Study**

Vane-Wright *et al.* (1991) suggested that taxa should be evaluated on the basis of phylogenetic position. Greater conservation efforts should be put towards those taxa which appear to be more isolated members of less species-rich clades. A way of defining biodiversity for prioritising conservation based on the number of species and amount of diversity among species was described by Williams *et al.* (1991). The production of an estimate of phylogeny including endangered island species of



*Phyllica* along with the closely related genus *Nesiota* and information concerning the conservation genetic status of individual species derived from AFLP fingerprint studies will help to set conservation priorities.

#### 1.9.1. Aims of conservation genetics study

1. To determine the conservation genetic status of island species, particularly those which are rare or endangered.
2. To determine the usefulness of AFLP data in conservation genetics studies.

#### **1.10. Thesis Structure**

In Chapter Two I present a molecular analysis of Rhamnaceae using *rbcL* and *trnL-F* plastid DNA sequences. In Chapter Three the results of a morphological analysis of Rhamnaceae allows the comparison of the use of morphological and molecular characters in phylogenetic reconstruction of the group. Chapter Three also includes a revision of the tribal classification of the family. Chapter Four is composed of a molecular analysis of *Phyllica* with an emphasis on island species based on *trnL-F* plastid DNA sequences and sequences of the internal transcribed spacer (ITS) of nuclear ribosomal DNA. A study on the population genetics of some island species of *Phyllica* based on amplified fragment length polymorphisms is presented in Chapter Five and Chapter Six is a study of the conservation genetics of St Helenan species of Rhamnaceae. In Chapter Seven I conclude with a summary of the results from each chapter and discussions on the use of molecular techniques at various hierarchical levels within Rhamnaceae.

#### **1.11. Bibliography**

Bachmann, K. 1994. Molecular markers in plant ecology. *New Phytologist* 126: 403-418.

Bakker, F.T., D. Hellbrügge, A. Culham, & M. Gibby. 1998. Phylogenetic relationships within *Pelargonium* sect. *Peristera* (Geraniaceae) inferred from nrDNA and cpDNA sequence comparisons. *Plant Systematics and Evolution* 211: 273-286.

Baldwin, B.G., Kyhos, D.W. & Dvorak, J. 1990. Chloroplast DNA evolution and adaptive radiation in the Hawaiian Silversword alliance (Madiinae, Asteraceae), *Annals of the Missouri Botanic Garden* 77: 96-109.

Baldwin, B.G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. *Molecular Phylogenetics and Evolution* 1: 3-16.

Brauner, S., D.J. Crawford & T.F. Stuessy. 1992. Ribosomal DNA and RAPD variation in the rare plant family Lactoridaceae. *American Journal of Botany* 79: 1436-1439.

Camin, J.H. & R.R. Sokal. 1965. A method for deducing branching sequences in phylogeny. *Evolution* 19: 311-326.

Chase, M.W. & J.D. Palmer. 1989. Chloroplast DNA systematics of lilioid monocots - resources, feasibility and an example from the Orchidaceae. *American Journal of Botany* 76: 1720-1730.

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

Crawford, D.J. 1989. Enzyme electrophoresis and plant systematics. *In* Soltis, D.E. and P.S. Soltis [eds.], *Isozymes in Plant Biology*. Dioscorides Press, Portland, Oregon.

Crawford, D.J., T.F. Stuessy, D. Wiens & D.W. Haines. 1991. The Lactoridaceae on the Juan Fernandez Islands: enzyme electrophoresis, and new observations on number and sizes of populations. *American Journal of Botany* 78: 453 [abstract].

Cronk, Q.C.B. 1990. The history of the endemic flora of St Helena: late Miocene 'Trochetiopsis - like' pollen from St Helena and the origin of *Trochetiopsis*. *New Phytologist* 114: 159-165.

Cronk, Q.C.B. 1992. Relict floras of Atlantic islands: patterns assessed. *Botanical Journal of the Linnean Society* 46: 91-103.

Cronk, Q.C.B. 1997. Islands: stability, diversity, conservation. *Biodiversity and Conservation* 6: 477-493.

DeBry, R.W. & N.A. Slade. 1985. Cladistic analysis of restriction endonuclease cleavage maps within a maximum likelihood framework. *Systematic Zoology* 34: 21-34.

Dollo, L. 1893. Les lois de l'évolution. *Bulletin de la Société Belge de Géologie, de Paléontologie et d'Hydrologie* 7: 164-66.

Donoghue, M.J., R.G. Olmstead, J.F. Smith & J.D. Palmer. 1992. Phylogenetic relationships of Dipsacales based on *rbcL* sequences. *Annals of the Missouri Botanical Garden* 79: 333-345.

Doyle, J.J. 1992. Gene trees and species trees – molecular systematics as one-character taxonomy. *Systematic Botany* 17: 144-163.

Doyle, J.J., M. Morgante, S.V. Tingey & W. Powell. 1998. Size homoplasy in chloroplast microsatellites of wild perennial relatives of soybean (*Glycine* subgenus *Glycine*). *Molecular Biology and Evolution* 15: 215-218.

Echt, C.S., L.L. DeVerno, M. Anzide & G.G. Vendramin. 1998. Chloroplast microsatellites reveal population genetic diversity in red pine, *Pinus resinosa* Ait. *Molecular Ecology* 7: 307-316.

Eck, R.V. & M.O. Dayhoff. 1966. Atlas of Protein Sequences and Structure 1966. National Biomedical Research Foundation, Silver Spring, Maryland.

Edwards, A.W.F. & L.L. Cavalli-Sforza. 1964. The reconstruction of evolution. *Annals of Human Genetics* 27:105.

Edwards, A.W.F. & L.L. Cavalli-Sforza. 1964. Reconstruction of evolutionary trees. In V.H. Heywood, J. McNeill [eds.], *Phenetic and Phylogenetic Classification*. Systematic Association. Publ. no. 6, London.

Edwards, K.J., J.H.A. Barker, A. Daly, C. Jones & A. Karp. 1996. Microsatellite libraries enriched for several microsatellite sequences in plants. *Biotechniques* 20: 758.

Farris, J.S. 1972. Estimating phylogenetic trees from distance matrices. *American Naturalist* 106: 645-648.

Fay, M.F. & M.W. Chase. 1996. Resurrection of Themidaceae for the *Brodiea* alliance, and recircumscription of Alliaceae, Amaryllidaceae and Agapanthoideae. *Taxon* 45: 441-451.

Fay, M.F., S.M. Swensen & M.W. Chase. 1997. Taxonomic affinities of *Medusagyna oppositifolia* (Medusagynaceae). *Kew Bulletin* 52: 111-120.

Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology* 27: 401-410.

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

Felsenstein, J. 1984. Distance methods for inferring phylogenies: A justification. *Evolution* 38: 16-24.

Felsenstein, J. 1988. Phylogenies from molecular sequences: inference and reliability. *Annual Review of Genetics* 22: 521-565.

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

Graybeal, A. 1998. Is it better to add taxa or characters to a difficult phylogenetic problem? *Systematic Biology* 47: 9-17.

Hennig, W. 1950. *Grundzüge einer Theorie der Phylogenetischen Systematik*. Deutsche Zentralverlag, Berlin.

Hennig, W. 1966. *Phylogenetic Systematics*. University of Illinois Press, Urbana, Illinois.

Hillis, D.M. 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Systematic Zoology* 47: 3-8.

Jansen, R.K. & J.D. Palmer. 1988. Phylogenetic implications of chloroplast DNA restriction site variation in Mutiseae (Asteraceae). *American Journal of Botany* 75: 751-764.

Jones, C.J., K.J. Edwards, S. Castaglione, M.O. Winfield, F. Sala, C. vandeWiel, G. Bredemeijer, B. Vosman, M. Matthes, A. Daly, R. Brettschneider, P. Bettini, M. Buiatti, E. Maestri, A. Malcevski, N. Marmiroli, R. Aert, G. Volckaert, J. Rueda, R. Linacero, A. Vazquez & A. Karp. 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding* 3: 381-390.

Kardolus, J.P., H.J. Van Eck & R.G. Van Den Berg. The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy (Solanaceae). *Plant Systematics and Evolution* 210: 87-103.

Kleppe, K., E. Ohtsuka, R. Kleppe, I. Molineux & H.G. Khorana. 1971. Studies on polynucleotides XCVI. Repair replication of short synthetic DNA's as catalyzed by DNA polymerases. *Journal of Molecular Biology* 56: 341-361.

Li, W.H. 1981. Simple methods for constructing phylogenetic trees from distance matrices. *Proceedings of the National Academy of Sciences USA* 78: 1085-1089.

Lincoln, R.J., G.A. Boxshall & P.F. Clark. 1982. A dictionary of ecology, evolution and systematics. Cambridge University Press, Cambridge.

Linnaeus, C. 1753. *Species Plantarum*. Stockholm.

Linnaeus, C. 1758. *Systema Naturae*. Ed. 10. Stockholm.

Lledó, M.D., Crespo, M.B., Cameron, K.M., Fay, M.F. & Chase, M.W. 1998. Systematics of Plumbaginaceae based upon cladistic analysis of *rbcL* sequence data. *Systematic Botany* 23: 21-29.

Mayr, E. 1969. *Principles of Systematic Zoology*. McGraw Hill, New York.

- Morton, C.M., S.A. Mori, G.T. Prance, K.G. Carol & M.W. Chase. 1997. Phylogenetic relationships of Lecythidaceae: a cladistic analysis using *rbcL* sequence and morphological data. *American Journal of Botany* 84: 530-540.
- Mullis, K.B. & F.A. Faloon. 1987. Specific synthesis of DNA in vitro via a polymerase catalyzed chain reaction. *Methods in Enzymology* 155: 335-350.
- Myers A.A. & P.S. Giller [eds.]. 1988. *Analytical Biogeography, an integrated approach to the study of animal and plant distribution*. Chapman & Hall, London.
- Pamilo, P. & M. Nei. 1988. Relationships between gene trees and species trees. *Molecular Biology and Evolution* 5: 568-583.
- Pillans, N.S. 1942. The genus *Phyllica* Linn., *Journal of South African Botany* 8: 1-164.
- Price, R.A. & J.D. Palmer. 1993. Phylogenetic relationships of the Geraniaceae and Geraniales from *rbcL* sequence comparisons. *Annals of the Missouri Botanical Garden* 80: 661-671.
- Qamaraz-Zaman, F., M.F. Fay, J.S. Parker & M.W. Chase. 1998. Molecular techniques employed in the assessment of genetic diversity. *Lindleyana* 13: 259-283.
- Riesberg, L., D.E. Soltis & J.D. Palmer. 1988. A molecular re-examination of introgression between *Helianthus annuus* and *H. bolanderi* (Compositae). *Evolution* 42: 227-238.
- Rogstad, S.H., J.C. Patton & B. Schaal. 1988a. A human minisatellite probe reveals RFLPs among individuals of two angiosperms. *Nucleic Acids Research* 16: 11378.

Rogstad, S.H., Patton, J.C. & Schaal, B. 1988b. M13 probe detects DNA minisatellite-like sequences in gymnosperms and angiosperms. *Proceedings of the National Academy of Sciences USA* 85: 9176-9178.

Roupe van der Voort, J.N.A.M., P.M. Van Zandvoort, H.J. Van Eck, R.T. Folkertsma, R.C.B. Hutten, J. Draaiistra, F.J. Gommers, E. Jacobsen, J. Helder & J. Bakker. 1997. Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotype. *Molecular and General Genetics* 255: 438-447.

Sachs, M.M., E.S. Dennis, W.L. Gerlach & W.J. Peacock. 1986. Two alleles of maize alcohol dehydrogenase-1 have 3' structural and poly(A) addition polymorphisms. *Genetics* 113: 449-467.

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

Savolainen, V., C.M. Morton, S.B. Hoot & M.W. Chase. (submitted). An examination of phylogenetic patterns of plastid *atpB* sequences among eudicots. *American Journal of Botany*.

Schaal, B.A. & G.H. Learn, Jr. 1988. Ribosomal DNA variation within and among plant populations. *Annals of the Missouri Botanical Garden* 75: 1207-1216.

Serikawa, T., T. Kuramoto, P. Hilbert, M. Mori, J. Yamada, C.J. Dubay, K. Lindpainter, D. Ganten, J.-L. Guénet, G.M. Lathrop & J.S. Beckmann. 1992. Rate gene mapping using PCR-analyzed microsatellites. *Genetics* 131: 701-721.

Sheahan, M.C. & M.W. Chase. 1996. A phylogenetic analysis of Zygophyllaceae R.Br. based on morphological, anatomical and *rbcL* DNA sequence data. *Botanical Journal of the Linnean Society* 122: 279-300.



- Siddall, M.E. & A.G. Kluge. 1997. Probabilism and phylogenetic inference. *Cladistics* 13: 313-336.
- Siddall, M.E. 1998. Prior agreement: Arbitration or arbitrary? Success of parsimony in the four-taxon case: Long-branch repulsion by likelihood in the Farris Zone *Cladistics*. 14: 209-220 .
- Sokal, R.R. & P.H.A. Sneath. 1963. *Numerical Taxonomy*. Freeman & Co., San Francisco.
- Soltis, D.E., D.R. Morgan, A. Grable, P.S. Soltis, & R. Kuzoff. 1993. Molecular systematics of Saxifragaceae *sensu stricto*. *American Journal of Botany* 80: 1056-1081.
- Soltis, D.E., P.S. Soltis, D.R. Morgan, S.M. Swensen, B.C. Mullins, J.M. Dowd & P. Martin. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proceedings of the National Academy of Sciences, USA* 92: 2647-2651.
- Soltis, D.E., P.S. Soltis, M.E. Mort, M.W. Chase, V. Savolainen, S.B. Hoot, C.M. Morton. 1998. Inferring complex phylogenies using parsimony: An empirical approach using three large DNA data sets for angiosperms. *Systematic Biology* 47: 32-42.
- Stace, C.A. 1989. *Plant Taxonomy and Biosystematics*. Edward Arnold, London.
- Strieff, R., T. Labbe, R. Bacilieri, H. Steinkellner, J. Glössl & A. Kremer. 1998. Within population genetic structure in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. *Molecular Ecology* 7: 317-328.
- Suessenguth, K. 1953. Rhamnaceae, Vitaceae, Leeaceae. *In* Engler, A. & K. Prantl [eds.], *Die Natürlichen Pflanzenfamilien* 2, Aufl., 20d. Duncker & Humblot, Berlin.

Tateno, Y, M. Nei & F. Tajima. 1982. Accuracy of estimated phylogenetic trees from molecular data. 1. Distantly related trees. *Journal of Molecular Evolution* 18: 387-404.

Thulin, M., B. Bremer, J.E. Richardson, J. Niklasson, M.F. Fay & M.W. Chase. 1998. Family relationships of the enigmatic rosid genera *Barbeya* and *Dirachma* from the Horn of Africa region. *Plant Systematics and Evolution* 213: 103-119.

Turrill, W.B. 1940. Experimental and synthetic plant taxonomy. In J.S. Huxley [ed.], *The New Systematics*. Clarendon Press, Oxford.

Vane-Wright, R.I., C.J. Humphries & P.H. Williams. 1991. What to protect? – Systematics and the agony of choice. *Biological Conservation* 55: 235-254.

Van Eck, H.J., J. Rouppe van der Voort, J. Draaistra, P. van Zandewoort, E. van Enkevort, B. Segers, J. Peleman, E. Jacobsen, J. Helder & J. Bakker. 1995. The inheritance and chromosomal location of AFLP markers in a non-inbred potato offspring. *Molecular Breeding* 1: 397-410.

Van Heusden, A.W. & K. Bachmann. 1992. Genetic differentiation of *Microseris pygmaea* (Asteraceae, Lactuceae) studied with DNA amplification from arbitrary primers (RAPDs). *Acta Botanica Neerlandica* 41: 385-395.

Van Heusden, A.W. & K. Bachmann. 1992. Genotype relationships in *Microseris elegans* (Asteraceae, Lactuceae) revealed by DNA amplification from arbitrary primers (RAPDs). *Plant Systematics and Evolution* 179: 221-233.

Van Heusden, A.W. & K. Bachmann. 1992. Nuclear DNA polymorphisms among strains of *Microseris bigelovii* (Asteraceae, Lactuceae) amplified from arbitrary primers. *Botanica Acta* 105: 331-336.

Vendramin, G.G. & B. Ziegenhagen. 1997. Characterisation and inheritance of polymorphic plastid microsatellites in *Abies*. *Genome* 40: 857-864.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, N. Hornes, A. Freitjers, J. Pot, J. Peleman, M. Kuiper & M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407-4414.

Wagner, W.H. 1961. Problems in the classification of ferns. *Recent Advances in Botany* 1: 841-844.

Weber, J.L. & P.E. May. 1989. Abundant class of human DNA polymorphism which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* 44: 388-396.

Weising, K., B. Beyermann, J. Ramser & G. Kahl. 1991. Plant DNA fingerprinting with radioactive and digoxigenated probes complementary to simple repetitive DNA sequences. *Electrophoresis* 12: 159-168.

Weising, K., F. Weigand, A.J. Driesel, G. Kahl, H. Zischler & J. Eppelen. 1989. Polymorphic simple GATA/GACA repeats in plant genomes. *Nucleic Acids Research* 17: 10128.

Weissenbach, J., G. Gyapay, C. Dib, A. Vignal, J. Morissette, P. Millasseau G. Vaysseix, M. Lathrop. 1992. A second generation linkage map of the human genome. *Nature* 359: 794-801.

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski & S.V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18: 6531-6535.

Williams, J.G.K., M.K. Hanafey, J.A. Rafalski & S.V. Tingey. 1993. Genetic analysis using randomly amplified polymorphic DNA markers. *Methods in Enzymology* 218: 704-740.

Williams, S.E., V.A. Albert & M.W. Chase. 1994. Relationships of Droseraceae - a cladistic analysis of rbcL sequence and morphological data. *American Journal of Botany* 81: 1027-1037.

Zhu, J., M.D. Gale, S. Quarrie, M.T. Jackson & G.J. Bryan. 1998. AFLP markers for the study of rice biodiversity. *Theoretical and Applied Genetics* 96: 602-611.

Zimmermann, W. 1930. *Die Phylogenie der Pflanzen*. Fischer, Jena, Germany.

Zimmermann, W. 1931. Arbeitsweise der botanischen Phylogenetik und andere Gruppierungswissenschaften. *In* E. Abderhalden [ed.], *Handbuch der biologischen Arbeitsmethoden*, 941-1053. Urban & Schwarzenberg, Berlin.

Zimmermann, W. 1934. Research on phylogeny of species and of single characters. *American Naturalist* 68: 381-384.

Zimmermann, W. 1943. Die Methoden der Phylogenetik. *In* G. Heberer [ed.], *Die Evolution der Organismen*, 20-56. Fischer, Jena, Germany.

**CHAPTER TWO. A MOLECULAR  
ANALYSIS OF RHAMNACEAE USING *rbcL*  
and *trnL-F* PLASTID DNA SEQUENCES**

## CHAPTER TWO. A Molecular Analysis Of Rhamnaceae Using *rbcL* And *trnL-F* Plastid DNA Sequences

### Abstract

Previous tribal classifications of Rhamnaceae have been based on fruit characters, resulting in the delimitation of large and otherwise heterogeneous groups. The last treatment of the tribal classification of the family by Suessenguth recognised five tribes. This classification was evaluated with DNA sequences from two regions of the plastid genome, *rbcL* and *trnL-F*, from 42 genera of Rhamnaceae and representatives of the related families Elaeagnaceae, Barbeyaceae, Dirachmaceae, Urticaceae, Ulmaceae, Moraceae and Rosaceae. The closest relatives of Rhamnaceae are Dirachmaceae and Barbeyaceae. The plastid trees support the monophyly of the family and provide the basis for a new tribal classification. Three major strongly supported clades are identified, but morphological characters could not be found to underpin a formal taxonomic description of these three clades as subfamilies. Therefore only those groups which are also defined by morphological characters are recognised. The biogeography of Rhamnaceae is discussed with reference to the molecular trees. The *trnL-F* trees have higher consistency and retention indices than the *rbcL* trees. The molecular evolution and use of *rbcL* and *trnL-F* in phylogenetic analysis is compared.

### 2.1. Introduction

Rhamnaceae are a cosmopolitan family of trees, shrubs, climbers and one herb consisting of about 50 genera and about 900 species. Rhamnaceae are characterised by simple leaves, small flowers with four or five sepals, which are valvate in bud, four or five stamens, which alternate with the sepals and oppose the petals (see Figure 2.1), anthers, which are frequently enfolded by the hooded petal apices, ovaries, which are usually 2- or 3-locular (sometimes 4- or 5-locular), an intrastaminal, nectariferous disc and a tendency towards xeromorphism. The sepals

often have a fleshy layer on the inner side, which usually forms a keel and ends as a tubercle. This layer is histologically similar to the intrastaminal, nectariferous disc (Cronquist, 1981). The alternation of petals and stamens with sepals is a relatively rare feature in angiosperms, and this has resulted in the family being associated with other families such as Vitaceae and Cornaceae, which also exhibit this character. The xeromorphic adaptations, which some members of the family exhibit, include reduced or absent leaves, crowding of leaves, shortening of branch axes, presence of thorns or spines and a low, shrubby habit. There are few plants of economic value in Rhamnaceae, the most notable being the jujube (*Ziziphus jujuba*), a fruit tree, and the ornamental shrubs *Ceanothus* and *Colletia*.

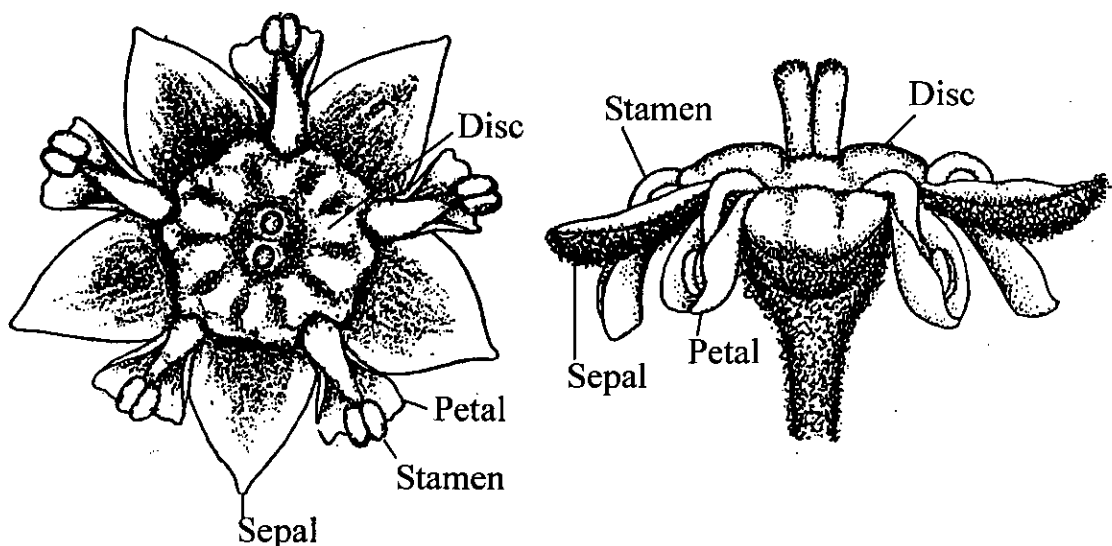


Figure 2.1. *Ziziphus jujuba* flowers showing stamens and petals alternating with sepals (Figure taken from Suessenguth, 1953).

Two patterns have generally been followed in the placement of Rhamnaceae in relation to other families: either they have been placed with groups such as Vitaceae on the basis of the shared feature of petals and stamens alternating with sepals (Takhtajan, 1980; Cronquist, 1988) or with Elaeagnaceae on the basis of shared

vegetative characteristics (Thorne, 1992; Takhtajan, 1997). These systems are summarised in Table 2.1.

Table 2.1. Taxonomic history of relationships of Rhamnaceae and related families

AUTHOR	ORDER	FAMILIES
Hutchinson (1959)	Rhamnales	Rhamnaceae, Heteropyxidaceae, Elaeagnaceae, Vitaceae
	Urticales	Barbeyaceae
	Tiliales	Dirachmaceae
Takhtajan (1980)	Rhamnales	Rhamnaceae, Vitaceae, Leeaceae
	Elaeagnales	Elaeagnaceae
	Barbeyales	Barbeyaceae close to Hammamelidales
	Geraniales	Dirachmoideae, a subfamily of Geraniaceae
Cronquist (1988)	Rhamnales	Rhamnaceae, Vitaceae, Leeaceae
	Proteales	Elaeagnaceae, Proteaceae
	Urticales	Urticaceae, Ulmaceae, Cannabaceae, Moraceae, Cecropiaceae, Barbeyaceae
	Geraniales	Dirachmaceae
Thorne (1992)	Rhamnales	Rhamnaceae, Elaeagnaceae
	Geraniales	Dirachmaceae - as Dirachmoideae, a subfamily of Geraniaceae
	incertae sedis	Barbeyaceae
Takhtajan (1997)	Barbeyales as superorder	Barbeyaceae
	Barbeyanae	
	Malvales	Dirachmaceae
	Rhamnales in superorder	Rhamnaceae
	Rhamnanae	
	Elaeagnales in superorder	Elaeagnaceae
Rhamnanae		

The taxonomic history of suprageneric relationships of genera now placed within Rhamnaceae is presented in Table 2.2. Adanson (1763) was the first to delimit what was to become part of Rhamnaceae under the name Jjubiers. Many of the genera that he included in this group, however, have since been placed in Rosaceae, Aquifoliaceae or Celastraceae. Jussieu (1789) divided Adanson's Jjubiers into six



groups. Brown (1814) merged Jussieu's first two groups to form Celastraceae and a second pair to form Rhamnaceae, which he characterised by features, which still describe the present familial circumscription. The Jujubiers were separated by Brongniart (1827) into the families Celastraceae, Ilicineae (=Aquifoliaceae) and Rhamnaceae, which at this stage included 18 genera. Subsequent treatments included those by Endlicher (1840), Hooker (1862), Baillon (1875), Weberbauer (1895) and Suessenguth (1953).

The most recent suprageneric or tribal classification of Rhamnaceae (Suessenguth 1953) was based largely on fruit characters and generally followed Hooker (1862). Suessenguth listed 58 genera in five tribes. Four genera have been described since Suessenguth's monograph. The first of these, *Oreohertzogia* (Vent 1962), was split from *Rhamnus* but is generally considered to be congeneric with *Rhamnus*. *Bathiorhamnus* Capuron from Madagascar did not appear to have a close affinity with any other group in the family (Capuron 1966). *Alvimiantha* Grey-Wilson from Brazil has been tentatively ascribed to the tribe Gouanieae Reiss. ex Endl. (Grey-Wilson 1978). *Disaster* Gilli (1980) was ascribed to Rhamnaceae but subsequently transferred to Sterculiaceae (Steenis 1982).

The genus *Tzellementinia* Chiov. has been transferred to Euphorbiaceae and synonymised with *Bridelia* Willd. (Friis and Vollesen, 1980). Some of the genera treated by Suessenguth are now regarded as congeneric with other genera in Rhamnaceae. These include *Cormonema* Reiss. ex Endl. (= *Colubrina* Rich. ex Brongn., Standley, 1925 and Cowan, 1952), *Microrhamnus* A. Gray (= *Condalia*, Johnston, 1962), *Hybosperma* Urb. (= *Colubrina*, Johnston 1963), *Sarcomphalus* P. Browne (= *Ziziphus* Mill., Johnston 1964), *Phyllogeiton* (Weberb.) Herzog (= *Berchemia* Neck. ex DC), *Chaydaia* Pit. (= *Rhamnella* Miq.), *Macrorhamnus* H. Perr. (= *Bathiorhamnus* Capuron, 1966), *Talguenea* Miers (= *Trevoa* Miers ex Hook., Tortosa 1992), *Lamellisepalum* Engl. (= *Sageretia* Brongn.), and *Oreorhamnus* Ridl. (= *Rhamnus* L.). Previous to this molecular analysis Rhamnaceae therefore comprised five tribes and 49 genera.

The suprageneric or tribal classification of Rhamnaceae had been based largely on fruit characters. In Suessenguth's system this resulted in the circumscription of two

large heterogeneous tribes, Rhamneae Hook. f. and Zizipheae Brongn. (=Paliureae Reiss. ex Endl.). An example of this heterogeneity can be found when comparing the genera *Ziziphus* and *Berchemia*, which were placed in the tribe Zizipheae because they both have drupaceous fruits. However, there are a number of other characters, which these two genera do not share with each other, such as ovary position and leaf venation, which might indicate relationships to genera in other tribes. The other tribes recognised by Suessenguth, Colletieae Reiss. ex Endl., Gouanieae Reiss. ex Endl. and Ventilagineae Hook. f., appeared on the basis of morphology to be natural groups.

An analysis of sequences of the plastid gene *rbcL* for 499 species of angiosperms (Chase, Soltis, Olmstead *et al.*, 1993) showed that Rhamnaceae are part of a weakly supported group which also contained Rosaceae, Urticales, and Fagales. Further studies using *rbcL* (Soltis *et al.*, 1995) indicated a close relationship between Elaeagnaceae and Rhamnaceae. Other studies using 18S rDNA, *atpB* and *rbcL* sequence data (Savolainen *et al.*, 1996, Soltis *et al.*, 1997) supported the link between Rhamnaceae and Elaeagnaceae. Sequence data from *rbcL* have placed Barbeyaceae and Dirachmaceae in association with Rhamnaceae (Thulin *et al.*, 1998). The occurrence of nitrogen-fixing symbioses in some Rhamnaceae, Elaeagnaceae, Ulmaceae, and Rosaceae offers further support for a close relationship between these families (Soltis *et al.*, 1995; Swensen *et al.*, 1996).

Taxa from the families listed above were included in this analysis in an attempt to refine further the ideas about relationships among them and between genera within Rhamnaceae. Sequences were obtained from two regions of the plastid genome for 66 taxa in Rhamnaceae and related families. Sequence data from *rbcL* at the intra-familial level have been widely applied such as in Dipsacales (Donoghue, 1992), Geraniaceae (Price and Palmer, 1993), Cornaceae (Xiang *et al.*, 1993), Saxifragaceae *sensu stricto* (Soltis *et al.*, 1993), Rosaceae (Morgan, 1994), Droseraceae (Williams *et al.*, 1994), Zygophyllaceae (Sheahan and Chase, 1996), Themidaceae (Fay and Chase, 1996), and Lecythidaceae (Morton *et al.*, 1997). Another plastid region was sequenced which consists of non-coding regions between transfer RNA genes. The *trnL* (UAA) 5' intron and the intergenic spacer between the *trnL* (UAA) 3' exon and

*trnF* (GAA; Taberlet *et al.*, 1991) were sequenced. This region will subsequently be referred to as *trnL-F*. This region has been used in suprageneric phylogenetic analysis of Iridaceae (Reeves *et al.*, 1997). The results of the analysis of these data were used in part to re-define the suprageneric classification of Rhamnaceae.

## **2.2. Aims**

1. To determine the monophyly of Rhamnaceae.
2. To determine relationships among genera in Rhamnaceae. Are Suessenguth's tribes monophyletic?
3. To determine the sister group of the genus *Phyllica* for subsequent phylogenetic analysis of this genus.

## **2.3. Materials and Methods**

### 2.3.1. Material for molecular analysis

Sources of plant material and vouchers used in this analysis are listed in Table 2.3. Forty-two genera of Rhamnaceae were sampled, including at least one representative of each of Suessenguth's five tribes. All genera of Elaeagnaceae, Barbeyaceae and Dirachmaceae and nine genera from Urticales and Rosaceae were also included. Rosaceae were chosen as the outgroup because earlier analyses (Chase, Soltis, Olmstead *et al.*, 1993; Soltis *et al.*, 1995; Thulin *et al.*, 1998) had shown this family to be more distantly related to Rhamnaceae.

### 2.3.2. DNA extraction

DNA was extracted from *c.* 1.0g fresh, 0.2-0.25g silica gel-dried leaves or 0.1-0.2g of material from herbarium sheets using a 2X CTAB method modified from Doyle and Doyle (1987). DNA was extracted from herbarium specimens for 21 of the 66 taxa. DNA was precipitated using isopropanol instead of ethanol because it is

more reliable for Rhamnaceae. DNA extracted from herbarium material was left to precipitate for at least three weeks at -20°C as this has been shown to give better yields (Fay *et al.*, 1998). The reasons for this are unclear, but it could be due to the presence of altered secondary compounds which form as a result of the degradation associated with drying which make the DNA more difficult to precipitate, or simply because the DNA from herbarium specimens is degraded and therefore takes longer to precipitate. All samples were purified on caesium chloride/ethidium bromide gradients (1.55g ml<sup>-1</sup>).

### 2.3.3. Gene amplification and purification

For most taxa the *rbcL* exon was amplified in two overlapping halves using forward primers beginning at position 1 and 636 and reverse primers beginning at position 724 and a downstream ribosomal control site (Lledó *et al.*, 1998; Table 2.4). DNA from some herbarium specimens had to be amplified in shorter pieces using forward primers beginning at position 636 and 895 and reverse primers beginning at position 1024 and the downstream site. The *trnL-F* region (Taberlet *et al.*, 1991) was amplified using the forward primer *c* and the reverse primer *f*. Again some of the DNA from herbarium specimens had to be amplified in shorter pieces using the primer pairs *c* and *d* and *e* and *f*. The *d* and *e* primers are exact complements so these sequences have a 20 base pair gap where the primer site is located. Amplification products were purified using Magic mini-columns (Promega, Southampton, Hampshire, UK) or QIAquick columns (Qiagen, Crawley, West Sussex, UK), following protocols provided by the manufacturers.

PCR amplification of *rbcL* and *trnL-F* involved 28 cycles of denaturation at 94°C for one minute; annealing of primer at 50°C for 30 seconds and nucleic acid extension at 72°C for one minute.

Table 2.2. Taxonomic history of suprageneric classifications in Rhamnaceae. Taxa which have been sampled are indicated with an \* in the Suessenguth system of this table.

AUTHOR	TRIBE/GROUP	GENERA
Jussieu (1789)	1	<i>Celastrus, Euonymus, Polycardia, Staphylea</i>
	2	<i>Cassine, Goupia, Ilex, Myginda, Prinos, Rubentia, Schrebera</i>
	3	<i>Mayepea, Rhamnus, Paliurus, Samara, Ziziphus</i>
	4	<i>Ceanothus, Colletia, Hovenia, Phylica</i>
	5	<i>Brunia, Bumalda</i>
	6	<i>Aucuba, Carpodetus, Gouania, Plectronia, Votomita</i>
Brongniart (1827)	n/a	<i>Berchemia, Ceanothus, Colletia, Colubrina, Condalia, Crumenaria, Cryptandra, Gouania, Hovenia, Paliurus, Phylica, Pomaderris, Retanilla, Rhamnus, Sageretia, Scutia, Ventilago, Ziziphus</i>
Endlicher (1840)	Colletieae	<i>Adolphia, Colletia, Discaria, Retanilla</i>
	Franguleae	<i>Alphitonia, Berchemia, Ceanothus, Colubrina, Condalia, Hovenia, Karwinskia, Noltea, Rhamnus, Sageretia, Scutia, Ziziphus</i>
	Gouanieae	<i>Crumenaria, Gouania, Helinus, Reissekia</i>
	Paliureae	<i>Paliurus, Ventilago</i>
	Phyliceae	<i>Cryptandra, Phylica, Spyridium</i>
	Pomadereae	<i>Pomaderris, Trymalium</i>
Hooker (1862)	Colletieae	<i>Adolphia, Colletia, Discaria, Retanilla, Trevoa</i>
	Gouanieae	<i>Crumenaria, Gouania, Helinus, Reissekia</i>
	Rhamneae	<i>Alphitonia, Ceanothus, Colubrina, Cryptandra, Hovenia, Lasiodiscus, Nesiota, Noltea, Phylica, Pomaderris, Rhamnidium, Rhamnus, Sageretia, Scutia, Spyridium, Trymalium</i>
	Ventilagineae	<i>Smythea, Ventilago</i>
	Zizipheae	<i>Berchemia, Condalia, Microrhamnus, Karwinskia, Paliurus, Ziziphus</i>

Baillon (1877)	Colletieae	<i>Adolphia, Colletia, Discaria, Retanilla, Trevoa</i>
	Gouanieae	<i>Crumenaria, Cryptandra, Gouania, Helinus, Lasiodiscus, Nesiota, Phyllica, Pomaderris, Reissekia, Spyridium, Trymalium</i>
	Rhamneae	<i>Alphitonia, Berchemia, Ceanothus, Colubrina, Condalia, Emmenosperma, Hovenia, Karwinskia, Noltea, Paliurus, Rhamnidium, Rhamnus, Sageretia, Scutia, Smythea, Ventilago, Ziziphus</i>
Weberbauer (1895)	Colletieae	<i>Adolphia, Colletia, Discaria, Retanilla, Trevoa</i>
	Gouanieae	<i>Crumenaria, Gouania, Helinus, Pleuranthodes, Reissekia</i>
	Maesopsidae	<i>Maesopsis</i>
	Rhamneae	<i>Alphitonia, Ceanothus, Colubrina, Cryptandra, Emmenosperma, Hovenia, Lasiodiscus, Nesiota, Noltea, Phyllica, Pomaderris, Rhamnus, Sageretia, Schistocarpaea, Spyridium, Trymalium</i>
	Ventilagineae	<i>Smythea, Ventilago</i>
	Zizipheae	<i>Berchemia, Condalia, Microrhamnus, Karwinskia, Paliurus, Reynosia, Rhamnella, Rhamnidium, Ziziphus</i>
Suessenguth (1953)	Colletieae	<i>Adolphia*, Colletia*, Discaria*, Kentrothamnus, Retanilla, Talguenea, Trevoa*</i>
	Gouanieae	<i>Crumenaria*, Gouania*, Helinus*, Pleuranthodes*, Reissekia*</i>
	Rhamneae	<i>Ampelozizphus*, Alphitonia*, Ceanothus*, Colubrina*, Cormonema, Cryptandra*, Emmenosperma*, Hovenia*, Hybosperma, Lasiodiscus*, Macrorhamnus, Nesiota*, Noltea*, Oreorhamnus, Phyllica*, Pomaderris*, Rhamnus*, Sageretia*, Schistocarpaea*, Scutia*, Siegfriedia*, Spyridium*, Trymalium*, Tzellemtinia</i>
	Ventilagineae	<i>Smythea, Ventilago*</i>
	Zizipheae	<i>Auerodendron, Berchemia*, Berchemiella, Chaydaia, Condalia*, Condaliopsis, Dallachya, Doerpfeldia*, Lamellisepalum, Microrhamnus, Karwinskia*, Krugiodendron*, Maesopsis*, Paliurus*, Phyllogeiton, Reynosia*, Rhamnella*, Rhamnidium*, Sarcomphalus, Ziziphus*</i>

Table 2.3. Taxon Accession data. (K) = Kew, (WS) = Washington State, (UPS) = Uppsala, (MICH) = Michigan.

Family/ Tribe	Species	Provenance	Voucher	Age of material	GenBank <i>rbcL/trnL-F</i>
Colletieae	<i>Adolphia infesta</i> (H.B.K.) Meisn.	Mexico	<i>McVaugh 7506</i> (K)	1945	
Colletieae	<i>Colletia ulicina</i> Gill. & Hook.	Chile	<i>Chase 608</i> (K)	fresh	
Colletieae	<i>Discaria chacaye</i> (G. Don) R.D. Tortosa	s South America	<i>Chase 914</i> (K)	fresh	U59826
Colletieae	<i>Trevoa trinervis</i> Miers	Chile	<i>Wall &amp; Sparre 2430</i> (K)	1947	
Gouanieae	<i>Crumenaria erecta</i> Reiss.	Brazil	<i>Ratter &amp; Rocha R.5015</i> (K)	1984	
Gouanieae	<i>Gouania mauritiana</i> Lam.	Mauritius	<i>Chase 904</i> (K)	fresh	
Gouanieae	<i>Helinus integrifolius</i> Kuntze	East Africa	<i>Thulin &amp; Warfa 5865</i> (K)	1986	
Gouanieae	<i>Pleuranthodes hillebrandii</i> (Oliver) Weberb.	Hawaii	<i>Hutchinson 2776</i> (K)	1967	
Gouanieae	<i>Reissekia smilacina</i> Endl.	Brazil	<i>Arbo et al. 4921</i> (K)	1991	
Rhamneae	<i>Alphitonia excelsa</i> Reiss.	Australia	<i>Chase 2179</i> (K)	fresh	
Rhamneae	<i>Ampeloziziphus amazonicus</i> Ducke	Brazil	<i>Vilhena &amp; Taylor 1004</i> (K)	1983	
Rhamneae	<i>Bathiorhamnus cryptophorus</i> Capuron	Madagascar	<i>Labar &amp; DuPuy 2044</i> (K)	1990	
Rhamneae	<i>Ceanothus thyrsiflorus</i> Esch. (2)	sw USA	<i>Chase 3177</i> (K)	fresh	
Rhamneae	<i>Ceanothus coeruleus</i> Lag. ( <i>trnL-F</i> ) (1)	sw USA	<i>Chase 2413</i> (K)	fresh	
Rhamneae	<i>Ceanothus sanguineus</i> Nutt. ( <i>rbcL</i> ) (1)	sw USA	<i>Morgan 2155</i> (WS)	fresh	U06795
Rhamneae	<i>Colubrina asiatica</i> Brongn. (1)	W Australia	<i>Chase 905</i> (K)	fresh	
Rhamneae	<i>Colubrina reclinata</i> (L'Hér.) Brongn. (2)	W Australia	<i>Chase 2115</i> (K)	fresh	
Rhamneae	<i>Cryptandra cf. spyridioides</i> F. Muell.	Australia	<i>Chase 2180</i> (K)	fresh	
Rhamneae	<i>Emmenosperma alphitonioides</i> F. Muell.	Australia	<i>Clarkson 8826</i> (K)	1990	
Rhamneae	<i>Hovenia dulcis</i> Thunb.	Japan	<i>Chase 968</i> (K)	fresh	
Rhamneae	<i>Lasiodiscus mildbraedii</i> Engl.	S. Tomé & c Africa	<i>Figueiredo et al. 29</i> (K)	1993	
Rhamneae	<i>Nesiota elliptica</i> (Roxb.) Hook. f.	St Helena	<i>Chase 500</i> (K)	fresh	
Rhamneae	<i>Noltea africana</i> (L.) Reichb.	South Africa	<i>Bayliss BS6824 49</i> (K)	1974	

Rhamneae	<i>Paliurus spina-christi</i> Mill.	Asia	Chase 969 (K)	fresh
Rhamneae	<i>Phylica nitida</i> Lam. (1)	Mauritius	Soorer 64-5 (MICH)	1964
Rhamneae	<i>Phylica polifolia</i> (Vahl) Pillans ( <i>rbcL</i> ) (2)	St Helena	Chase 1751 (K)	fresh
Rhamneae	<i>Phylica polifolia</i> (Vahl) Pillans ( <i>trnL-F</i> ) (2)	St Helena	Chase 2269 (K)	fresh
Rhamneae	<i>Phylica pubescens</i> Ait. (3)	South Africa	Chase 859 (K)	fresh
Rhamneae	<i>Pomaderris rugosa</i> Cheeseman	Australia	Chase 857 (K)	fresh
Rhamneae	<i>Schistocarpaea johnsonii</i> F. Muell.	Australia	Gray 1247 (K)	1979
Rhamneae	<i>Siegfriedia darwinoides</i> C.A. Gardner	Australia	Chase 2181 (K)	fresh
Rhamneae	<i>Spyridium globulosum</i> (Labill.) Benth. (3)	Australia	Chase 2021 (K)	fresh
Rhamneae	<i>Spyridium complicatum</i> F. Muell. (2)	Australia	Chase 2182 (K)	fresh
Rhamneae	<i>Spyridium</i> cf. <i>forrestianum</i> (1)	Australia	Chase 2183 (K)	fresh
Rhamneae	<i>Trymalium ledifolium</i> Fenzl (1)	Australia	Chase 2184 (K)	fresh
Rhamneae	<i>Trymalium floribundum</i> Steudel (2)	Australia	Chase 2185 (K)	fresh
Rhamneae	<i>Rhamnus cathartica</i> L. (2)	North America	Chase 100 (UNC)	fresh
Rhamneae	<i>Rhamnus lycioides</i> L. (1)	Spain	Chase 1884 (K)	fresh
Rhamneae	<i>Sageretia thea</i> (Osbeck) M.C. Johnston	Asia	Collenette 9/93 (K)	
Rhamneae	<i>Scutia buxifolia</i> Reiss.	South America	Kew 1973-12719 (K)	fresh
Zizipheae	<i>Berchemia discolor</i> (Klotch) Hemsley	Asia	Collenette 14/93 (K)	fresh
Zizipheae	<i>Condalia microphylla</i> Cav.	Argentina	Kiesling et al. 5967 (K)	1986
Zizipheae	<i>Doerpfeldia cubensis</i> Urban	Cuba	Howard et al. 246 (K)	1950
Zizipheae	<i>Karwinskia humboldtiana</i> (Roem. & Schult) Zucc.	Mexico, Cuba, Haiti	Brennan 14483 (K)	1977
Zizipheae	<i>Krugiodendron ferreum</i> (Vahl) Urban	West Indies, SW USA, Mexico	Lundell 17449 (K)	1963
Zizipheae	<i>Maesopsis eminii</i> Engl.	Tropical Africa	Chase 1338 (K)	fresh
Zizipheae	<i>Reynosia uncinata</i> Urban	Cuba	Chase 363 (K)	fresh
Zizipheae	<i>Rhamnella franguloides</i> (Maxim.) Weberb.	China, Japan, Korea	Chase 912 (K)	fresh



Zizipheae	<i>Rhamnidium elaeocarpum</i> Reiss.	n South America	<i>Santos et al. 693</i> (K)	1983	
Zizipheae	<i>Ziziphus ornata</i> Miq. (2)	Java	<i>Chase 2117</i> (K)	fresh	
Zizipheae	<i>Ziziphus glabrata</i> Roxb. (1)		<i>Chase 472</i> (K)	fresh	U60313
Ventilagineae	<i>Ventilago viminalis</i> Hook. (1)	Australia	<i>Kenneally 9507</i> (K)	1985	
Ventilagineae	<i>Ventilago leiocarpa</i> Benth. (2)	se Asia	<i>Hu 11890</i> (K)	1972	
Elaeagnaceae	<i>Elaeagnus angustifolia</i> L. ( <i>trnL</i> -F)		referenced in GenBank	fresh	U17038
Elaeagnaceae	<i>Elaeagnus</i> sp. ( <i>trnL</i> -F)		<i>Chase 2414</i> (K)	fresh	
Elaeagnaceae	<i>Hippophae salicifolia</i> D. Don	Nepal	<i>Chase 856</i> (K)	fresh	U59821
Elaeagnaceae	<i>Shepherdia canadensis</i> (Pursh.) Nutt. ( <i>rbcL</i> )	USA	referenced in GenBank	fresh	U17039
Elaeagnaceae	<i>Shepherdia argentea</i> L. ( <i>trnL</i> -F)	USA	<i>Chase 3176</i> (K)	fresh	
Barbeyaceae	<i>Barbeya oleoides</i> Schweinf.	southern Arabia, Socotra	<i>Collenette s.n.</i> (K)	fresh	
Dirachmaceae	<i>Dirachma socotrana</i> Schweinf.	Socotra	<i>Thulin &amp; Gifri 8812</i> (UPS)	fresh	
Moraceae	<i>Dorstenia psilurus</i> Welw.	Tropical Africa	<i>Chase 2416</i> (K)	fresh	
Moraceae	<i>Artocarpus heterophyllus</i> Lam.	SE Asia	<i>Chase 2415</i> (K)	fresh	
Moraceae	<i>Ficus pretoriae</i> Burt-Davy	South Africa	<i>Chase 2412</i> (K)	fresh	
Cannabaceae	<i>Cannabis sativa</i> L. ( <i>trnL</i> -F)	Pantropical	<i>Chase 2992</i> (K)	fresh	
Ulmaceae	<i>Gironniera subaequalis</i> Benth. ( <i>trnL</i> -F)	Indomalaysia	<i>Chase 1384</i> (K)	fresh	
Urticaceae	<i>Boehmeria biloba</i> Hooker	Java	<i>Chase 2532</i> (K)	fresh	
Rosaceae	<i>Dryas drummondii</i> L.	Siberia or American Arctic	<i>Chase 917</i> (K)	fresh	U59818
Rosaceae	<i>Spiraea x vanhouttei</i> ( <i>rbcL</i> )	Garden origin	<i>Morgan 2130</i> (WS)	fresh	L11206
Rosaceae	<i>Spiraea betulifolia</i> L. ( <i>trnL</i> -F)	ne Asia to Japan	<i>Chase 2503</i> (K)	fresh	
Rosaceae	<i>Pyrus serotina</i> Rehder ( <i>trnL</i> -F)	China	<i>Chase 1018</i> (K)	fresh	

#### 2.3.4. DNA sequencing

Standard dideoxy methods using S<sup>35</sup> or modified dideoxy cycle sequencing with dye terminators run on an ABI 373A or 377 automated sequencer (according to the manufacturer's protocols; Applied Biosystems, Inc., Warrington, Cheshire, UK) were used to sequence the amplification products directly. Sequences were edited and assembled using Sequence Navigator and Autoassembler (Applied Biosystems Inc.) or manually. All sequences will be submitted to GenBank (for accession numbers see Table 2.3).

#### 2.3.5. Sequence alignment

The *rbcL* sequences were easily aligned because of the absence of insertions or deletions. An initial alignment for five *trnL-F* sequences was performed using Clustal version 1.61 (Higgins, Bleasby and Fuchs, 1992). Subsequent sequences were aligned manually.

After alignment of the *trnL-F* matrix, a matrix of insertion/deletion characters was prepared (characters were coded as present or absent; see Appendix 1). These characters were given weight equal to that of all other characters in the matrix because there was no basis for giving these characters extra weight over substitutions. A large deletion can mask other smaller deletions and taxa, which have these larger deletions, are coded as unknown for deletions that occur entirely within them. For example there is a deletion between positions 891 and 941 for some taxa, and in other taxa there are smaller deletions between these positions, which are coded as missing. The CI and RI values of each of these characters were calculated and are presented in Table 2.5.

A total of 1408 *rbcL* and 1191 *trnL-F* characters were used. The ends were clipped from the sequences to remove primer sites (i.e. 20bp from beginning of *rbcL*, 24bp from the beginning of *trnL-F* and 28bp from the end). Two regions of 59 and 16 bp of the *trnL-F* matrix were too ambiguous to be confidently aligned and so were excluded from all analyses.



### 2.3.6. Phylogenetic analysis using Parsimony (PAUP)

Data matrices were analysed using the parsimony algorithm of the software package PAUP version 3.1.1 for Macintosh (Swofford, 1993). Searches were conducted on the separate *rbcL* and *trnL-F* data sets (which included the matrix of 16 *trnL-F* indel characters) and on both data sets combined. PAUP provides 2 methods for searching for optimal (most parsimonious) trees:

#### 2.3.6.1. Exact methods

An exact method guarantees to find most parsimonious trees but cannot be used for matrices of over 20 terminals because it evaluates every possible tree. In data sets with more than 20 taxa, heuristic methods are implemented because they reduce the number of trees that need to be assessed, but they cannot guarantee finding the shortest tree(s). Because of the large number of taxa in this study heuristic methods were used.

#### 2.3.6.2. Heuristic methods

When applied to the search for most parsimonious trees there are two stages to heuristic methods:

2.3.6.2.1. Stepwise addition - An initial tree is obtained. Taxa are connected one at a time to a developing tree. The optimal tree is saved after each addition.

There is a choice of three ways in which taxa may be added:

i) *As is* - In the order of the data matrix

ii) *Closest* - The closest three taxa make up starting tree - at each successive step all remaining taxa are considered for connection to each branch of the tree - the combination requiring the smallest increase in tree length is chosen.

iii) *Simple* - The distance between each taxon and a reference taxon is calculated (this distance is termed the advancement index). Taxa are added in order of increasing

advancement. The reference taxon could be a hypothetical ancestor possessing the assumed ancestral state for each character.

iv) *Random* - Taxa for the distance calculation are added in a random order using a pseudorandom number generator.

2.3.6.2.2. Branch swapping - Stepwise addition does not often find the most parsimonious trees because one placement of a taxon may be best given the taxa currently on the tree, but that placement may become sub-optimal upon the addition of subsequent taxa. This results in the production of sub-optimal or less parsimonious trees. Improvements can be made by performing sets of pre-defined rearrangements ('branch-swapping'). PAUP uses three branch-swapping algorithms:

- a. nearest neighbour interchanges (NNI) – this is the fastest method, performing the fewest number of swaps per tree
- b. subtree pruning-regrafting (SPR) – this method is slower, but performs more swaps per tree
- c. tree bisection reconnection (TBR) – this is the slowest method, but it performs the most swaps per tree

If a rearrangement is successful in finding a better tree, a round of rearrangements is initiated on this new tree. However if in the process of arriving at the global optimum, we have to pass through trees that are inferior to the one(s) already obtained, we may again be trapped in a local optima unless we can carry out branch swapping on suboptimal trees, which is not feasible since there are too many of these with most matrices. The path to the optimal tree may also require that we pass through trees which are equal to the current tree. This problem is described as 'plateaus' on the optimality surface. This problem is alleviated by performing a number of analyses (replicates) using random stepwise addition of taxa. Taxa are added randomly to the distance calculation using a randomly selected taxon, and branch swapping is undertaken. When swapping is complete a new starting tree is generated by adding taxa randomly i.e. in a different order from the previous replicate. The more replicates that are performed the greater chance of finding the

most parsimonious trees and thus ignoring local sub-optimal trees. The most thorough of the branch swapping algorithms is TBR, and this is the one chosen for these analyses.

If a particular character or character state is missing (e.g. if there has been an insertion or a deletion of a nucleotide or nucleotide sequence) that state which is most parsimonious given the position of the taxon on the tree is assigned for the missing character.

2.3.6.2.3. Accelerated and delayed transformations - Character state changes may be placed on the tree as close to the root as possible. Homoplasy is therefore explained in terms of more distal reversals to plesiomorphic conditions. This procedure is known as the accelerated transformation option (ACCTRAN; Swofford and Maddison, 1987; Swofford, 1990). Conversely parallelisms may be favoured by postponing changes as far as possible from the root of the tree. Delayed transformation optimisation (DELTRAN; Swofford and Maddison, 1987; Swofford, 1990) maximises the proportion of homoplasy that is explained by parallelism. With DNA, ACCTRAN is the usual optimisation mode.

2.3.6.2.4. Assessing the reliability of inferred trees - The consistency index (CI) and the retention index (RI) are measures of how well a data set fits a particular tree. The consistency index (CI) is  $m/s$  where  $m$  is the minimum amount of changes possible and  $s$  is the actual amount of changes on a particular tree. Actual change,  $s$ , will exceed minimum possible change,  $m$ , to the extent that extra steps, or homoplasy, are required to account for the character on the tree. So for a given data set  $CI = 1$  when there is no homoplasy, and decreases as homoplasy increases. CI is negatively correlated with number of terminal taxa and number of characters, which makes its use in comparing trees with different numbers of taxa or characters less useful. Also, CI is inflated as the number of uninformative characters in the data set increases, but this problem can be avoided by using informative characters only.

The RI avoids the problem of uninformative characters by expressing the amount of synapomorphy in a data set by examining the actual amount of homoplasy as a

fraction of the maximum possible homoplasy (symplesiomorphies and autapomorphies do not contribute to RI as they admit no possibilities of homoplasy). The RI is  $(g-s)/(g-m)$  where  $g$  is a measure of how many changes it would take to explain evolution within the transformation series under the worst possible conditions. The RI is low when state changes mostly occur on internal nodes and high when changes mostly occur on branches leading to terminal taxa. The RI is the most important measure of performance for a matrix of characters.

The rescaled consistency index (RC) is the product of the CI and the RI. This figure averages out the performance of characters against worst case and best case scenarios.

2.3.6.2.5. Successive weighting (SW) - Successive weighting (Farris, 1969) is a way of down-weighting characters that are found to be highly homoplasious in an initial heuristic search. An initial cladogram(s) is produced under the Fitch criterion (i.e. equal weights), and the RC for each character on the initial cladogram is determined. In the case of multiple, equally parsimonious cladograms these are average values. The RC is then used to re-weight the initial character matrix. A new analysis is then performed on this altered matrix, and new unit character indices are calculated for the resulting cladogram(s) and the characters are re-weighted again. This process continues until the lengths of trees on successive iterations are identical. This technique produces a cladogram that is based on the most consistent characters.

2.3.6.2.6. Combining equally parsimonious trees (consensus techniques) - Most heuristic searches produce multiple most parsimonious trees. Consensus techniques are ways of combining equally parsimonious trees. These techniques do not always give the best estimate of phylogenetic relationships among groups. They provide evidence given by all equally parsimonious trees for patterns of ingroup relationships.

1. Strict consensus trees (Sokal and Rohlf, 1981) contain only those monophyletic groups that are common to all trees. Semi-strict consensus trees show monophyletic clades which are not contradicted by any of the equally parsimonious trees.
2. Adams consensus trees (Adams, 1972) are designed to give the highest resolution possible between two trees. Taxa that are placed in different positions in some trees are moved to the most resolved node common to all trees. This can result in the occurrence of clades that may not exist in any of the original trees, but it gives an idea of which taxa are most greatly affecting resolution.
3. Majority rule (Margush and Morris, 1981) states that if you have one tree that hypothesises that A and B are more closely related than either is to C and two trees that hypothesise that B and C are more closely related to each other than either is to A then the majority consensus tree will have the latter topology. Davis and Nixon (1996) have shown that groups that produce the greatest number of variable trees are supported by the weakest characters, so majority rule effectively produces consensus trees that perform in the opposite way from what is desired.

The production of strict consensus trees is the most stringent method, and this is the one chosen to combine the equally most parsimonious trees in these analyses.

2.3.6.2.7. Confidence measures - Bootstrap and jackknife methods provide support values for nodes in phylogenetic trees. Bootstrapping involves the random resampling of data to simulate a new data set for tree construction. The process is usually repeated 1000 times. The percentage of times that a clade appears is taken as a measure of support for that grouping.

The bootstrap involves random resampling of taxa or characters from the data set and random replacement until a data set the same size as the original is obtained. This resampling is performed a number of times (in this case 1000 replicates). A particular clade will have a 95% bootstrap value if it appears in 95% of trees. Branches that have less than 50% support are collapsed. Kluge and Wolf (1993) have suggested that bootstrap frequencies rely on the false assumption that each character evolves independently, and Carpenter (1994) demonstrated that the addition of

uninformative characters can result in a decrease in the number of significant groups as quantified by bootstrap frequencies.

Jackknifing involves random deletion without replacement of taxa or characters from a matrix. Jackknife values on branches indicate the percentage of replicates that retain that particular branch. Any branches that have less than 50% support are collapsed.

#### 2.3.6.3. Heuristic search strategy

1. Heuristic searches were performed under the equal weights criterion (Fitch, 1971) with 1000 random sequence additions and TBR (tree bisection-reconnection) branch-swapping, but saving only 10 trees per replicate. This means that 10 trees of a particular length were saved and each one was swapped on. If a shorter tree was found swapping was conducted on this tree and the others were discarded. Swapping continued until all 10 trees had been swapped on and no shorter trees were found. These trees were saved and a new replicate was initiated. The limit on the number of trees held at each step was implemented to cut down the computer time spent searching on sub-optimal trees.
2. All the shortest trees collected in the 1000 replicates were then used as starting trees for another round of heuristic search. These trees were swapped to completion using TBR until more than 6000 trees were produced, at which point the number of trees was limited and swapping to completion was performed on the 6000 trees collected.
3. Successive approximations weighting (SW; Farris, 1969) was then carried out. Characters were re-weighted according to their re-scaled consistency indices (RC), with a base weight of 1000. A new heuristic search was then carried out with 10 random addition replicates, saving 10 trees per replicate.
4. All trees found in step 3 were used as starting trees and swapped to completion using TBR, saving no more than 5000 trees.
5. Steps 3 and 4 were repeated, and again as needed until two rounds of successive weighting found trees of the same length.



6. At least some of the trees from the last round of SW were saved so that the final re-scaled weight could be readily re-implemented for use in bootstrap analysis.
7. The strict consensus tree was produced.
8. Bootstraps were performed after the final round of successive weighting. If this is done in a new PAUP session, the final weight-set was first re-established by loading the trees saved from step 6, then re-weighting characters by the re-scaled consistency index.
9. One thousand replicates of the bootstrap (Felsenstein, 1985) were carried out with the successive weights applied, using TBR swapping, saving 20 trees per replicate. The following scheme of support was applied: bootstrap values of 50-74% weak support, 75-84% moderate support, and 85-100% strong support.

MacClade (Maddison and Maddison, 1992) was used to calculate the number of steps and CI and RI for different codon positions in the *rbcL* analysis (Table 2.4), and CI and RI values of indel characters from the *trnL-F* matrix (Table 2.5). MacClade was also used to plot the number of unambiguous steps per character optimised on the most parsimonious SW tree from the combined analysis and the number of characters per number of steps on both the *trnL-F* and *rbcL* trees. The CI and RI values were calculated for transitions and transversions using step matrices on the successively-weighted tree of the combined analysis. The transitions were downweighted to zero via a step matrix and the CI and RIs of transversions were thus calculated by PAUP on the combined tree. These could then be used to calculate CI and RIs of transitions (Table 2.6). PAUP was used to calculate the number of steps in different trees for a given data set (Table 2.7).

## 2.4. Results

### 2.4.1. *rbcL* analysis

The *rbcL* data matrix had 1171 variable characters and 674 potentially informative characters out of a total of 1408 characters used, i.e. 48% of characters

were variable in two or more taxa. The heuristic search under the Fitch criterion produced more than 6000 equally parsimonious trees with a length of 1174 steps. The consistency index (CI) for these trees was 0.52 and the retention index (RI) was 0.66. With SW, there were seven trees with a length of 423378 steps, CI was 0.84, and RI was 0.86. The Fitch lengths for these trees was also 1174 steps, i.e. the weighted trees were a subset of the Fitch trees. Figure 2.2 shows one of the SW trees with its Fitch branch lengths (ACCTRAN optimisation) above the branches and SW bootstrap percentages below; branches, which collapse in the strict consensus tree of the weighted analysis, are marked with an arrow.

The trees produced indicate that Rhamnaceae are not a monophyletic group because Elaeagnaceae, Barbeyaceae and Dirachmaceae are all nested within it. The sister group to this clade includes members of the families Moraceae, Ulmaceae and Cannabaceae. However, there is little morphological evidence to indicate that Elaeagnaceae, Dirachmaceae and Barbeyaceae should be included within Rhamnaceae, and support for this grouping from the molecular data is weak. The tribes Rhamneae Hook. F., and Zizipheae Brongn., are paraphyletic, but Colletieae Reiss. ex Endl., and Gouanieae Reiss. ex Endl., are strongly supported monophyletic groups.

Within Rhamnaceae strongly supported major groups are identified: a ziziphoid group which has Elaeagnaceae as a sister group; a rhamnoid group which has *Ampeloziziphus*, *Doerpfeldia*, *Bathiorhamnus* and *Ventilago* as a sister group; and an ampeloziziphoid group which contains the genera *Ampeloziziphus*, *Doerpfeldia* and *Bathiorhamnus*. The inclusion of *Ventilago* in this group is weakly supported.

Other strongly supported groups within these larger groups include:

1. in the ziziphoid group: (i) a group of Australian taxa which had formerly been placed in the tribe Pomaderrieae Reiss. ex Endl.; (ii) *Ceanothus*; (iii) a group with a southern African center of distribution which had formerly been placed in Phylliceae Reiss. ex Endl.; (iv) *Colubrina*; (v) *Ziziphus*, *Paliurus* and *Hovenia*. 2. in the rhamnoid group: (i) a clade composed of *Karwinskia*, *Condalia*, *Krugiodendron*, *Reynosia*, *Rhamnella*, *Rhamnidium*, *Berchemia*, *Sageretia*, *Rhamnus*, *Frangula* and *Scutia*.

### 2.4.2. *trnL*-F analysis

The aligned *trnL*-F data matrix had 1156 variable characters and 566 potentially informative characters out of a total of 1239 characters (i.e. 46%). The heuristic search produced more than 6000 equally parsimonious Fitch trees with 1339 steps, CI =0.67, and RI=0.75. Application of SW produced more than 5000 trees with a length of 652105 steps, CI=0.87, and RI=0.91. The Fitch length of the SW tree was 1339, i.e. the weighted trees were a subset of the Fitch trees. Figure 2.3 shows one of the weighted trees with Fitch branch lengths (ACCTRAN optimisation) and SW bootstrap percentages; branches, which collapse in the strict consensus tree are marked with an arrow. The performance of the indel characters is shown in Table 2.5. The average CI was 0.84 and the average RI was 0.90, indicating that in general the levels of homoplasy for these characters are low.

Rhamnaceae are a strongly supported monophyletic group with a clade containing Dirachmaceae and Barbeyaceae as sister. Elaeagnaceae form a sister group to a clade containing Rhamnaceae, Barbeyaceae, Dirachmaceae and Urticales. Therefore the main differences between trees produced by the separate *rbcL* and *trnL*-F matrices were that the *rbcL* trees placed Elaeagnaceae, Dirachmaceae and Barbeyaceae within Rhamnaceae but with weak bootstrap support, whereas the *trnL*-F trees placed these families outside Rhamnaceae with strong bootstrap support for the monophyly of Rhamnaceae.

Within Rhamnaceae, the strongly supported major groups identified in the *rbcL* analysis here receive further support, i.e. the ziziphoid, rhamnoid and ampeloziziphoid groups. The inclusion of *Ventilago* in the rhamnoid group and not the ampeloziziphoid group is strongly supported. Within these the groups which are strongly supported in the *rbcL* analysis are given further support here. Generally speaking, the generic relationships and the larger clades identified are highly congruent with the *rbcL* results.

### 2.4.3. Combined *rbcL* and *trnL-F* analysis

The combined matrix produced 324 Fitch trees with a length of 2559 steps, a CI=0.59 and RI=0.70. With SW there was only one tree with two trichotomies. The SW tree length was 1068277 steps, CI=0.85, and RI=0.88. Figure 2.4 shows the single SW tree with Fitch branch lengths (ACCTRAN optimisation) and SW bootstrap values; branches which collapse in the strict consensus tree are marked with an arrow. The Fitch length of this tree was 2559 steps (i.e. it was one of the trees found with equal weights).

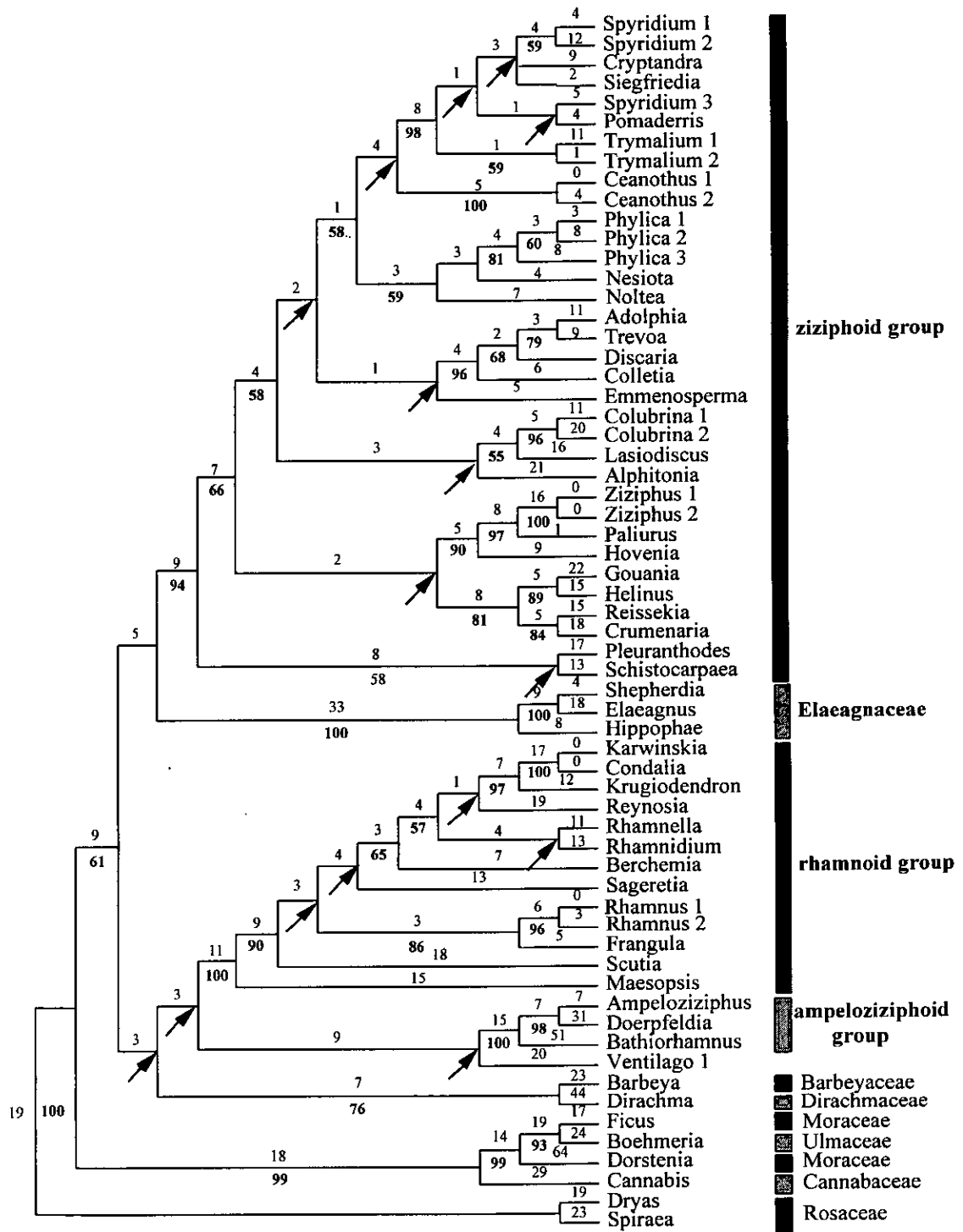


Figure 2.2. Example of one optimal SW tree from the *rbcL* analysis, with its Fitch lengths (above branches; ACCTRAN optimisation) and bootstrap values (below). Branches, which are not present in the strict consensus tree are indicated by an arrow. Heuristic search under the Fitch criterion produced more than 6000 equally parsimonious trees with a length of 1174 steps. The consistency index (CI) for these trees was 0.52 and the retention index (RI) was 0.66. There were only seven SW trees with a length of 423378 steps, CI=0.84, and RI=0.86 (Fitch length, 1174 steps).

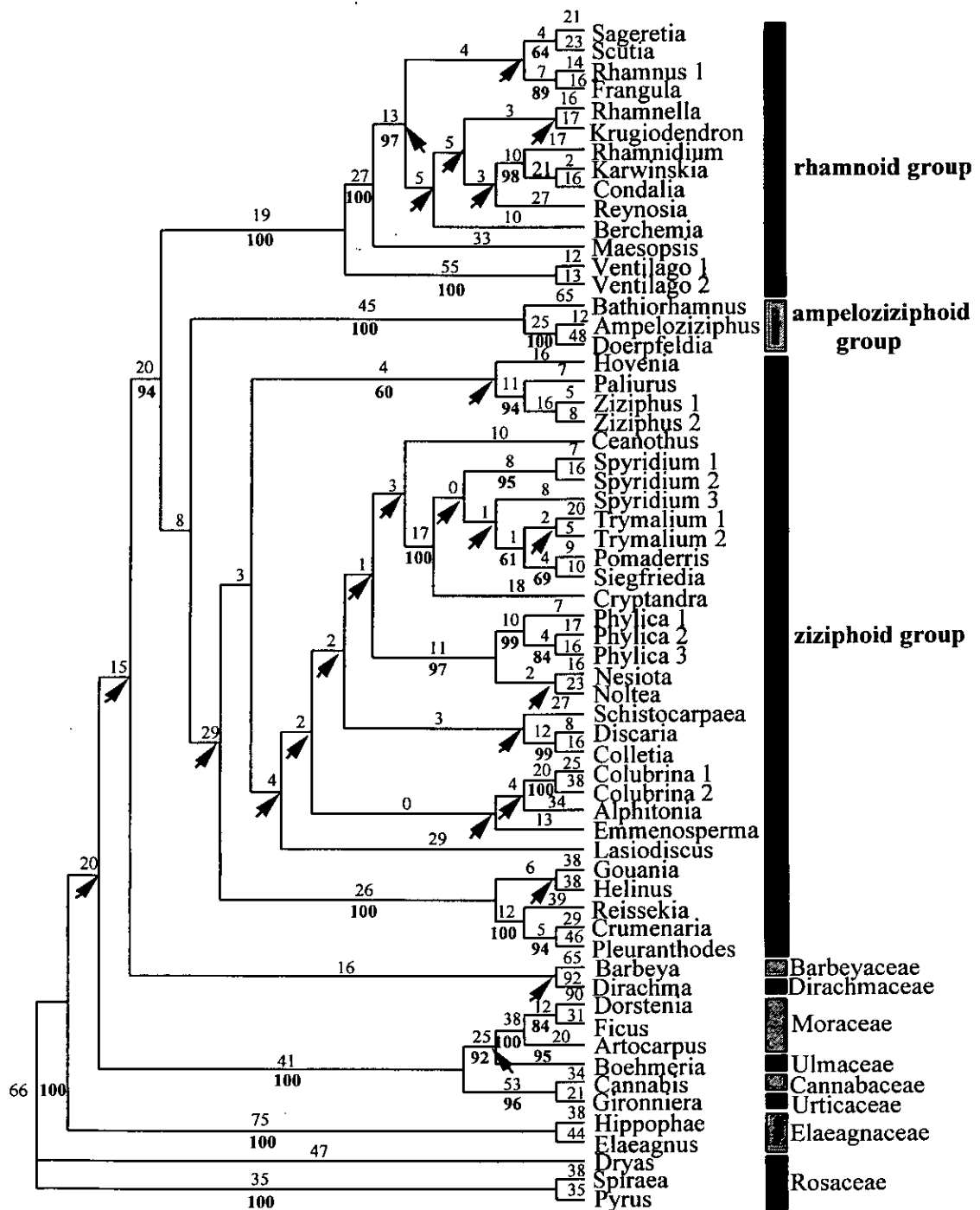


Figure 2.3. Example of one optimal SW tree from the *trnL-F* analysis, with its Fitch lengths (above branches; ACCTRAN optimisation) and bootstrap values (below). Branches not present in the strict consensus tree are indicated by an arrow. Heuristic search under the Fitch criterion produced more than 6000 equally parsimonious trees with a length of 1339 steps, CI=0.67, and RI=0.75. SW produced more than 5000 trees and a length of 652105 steps, CI=0.87, and RI=0.91 (Fitch length, 1339 steps).

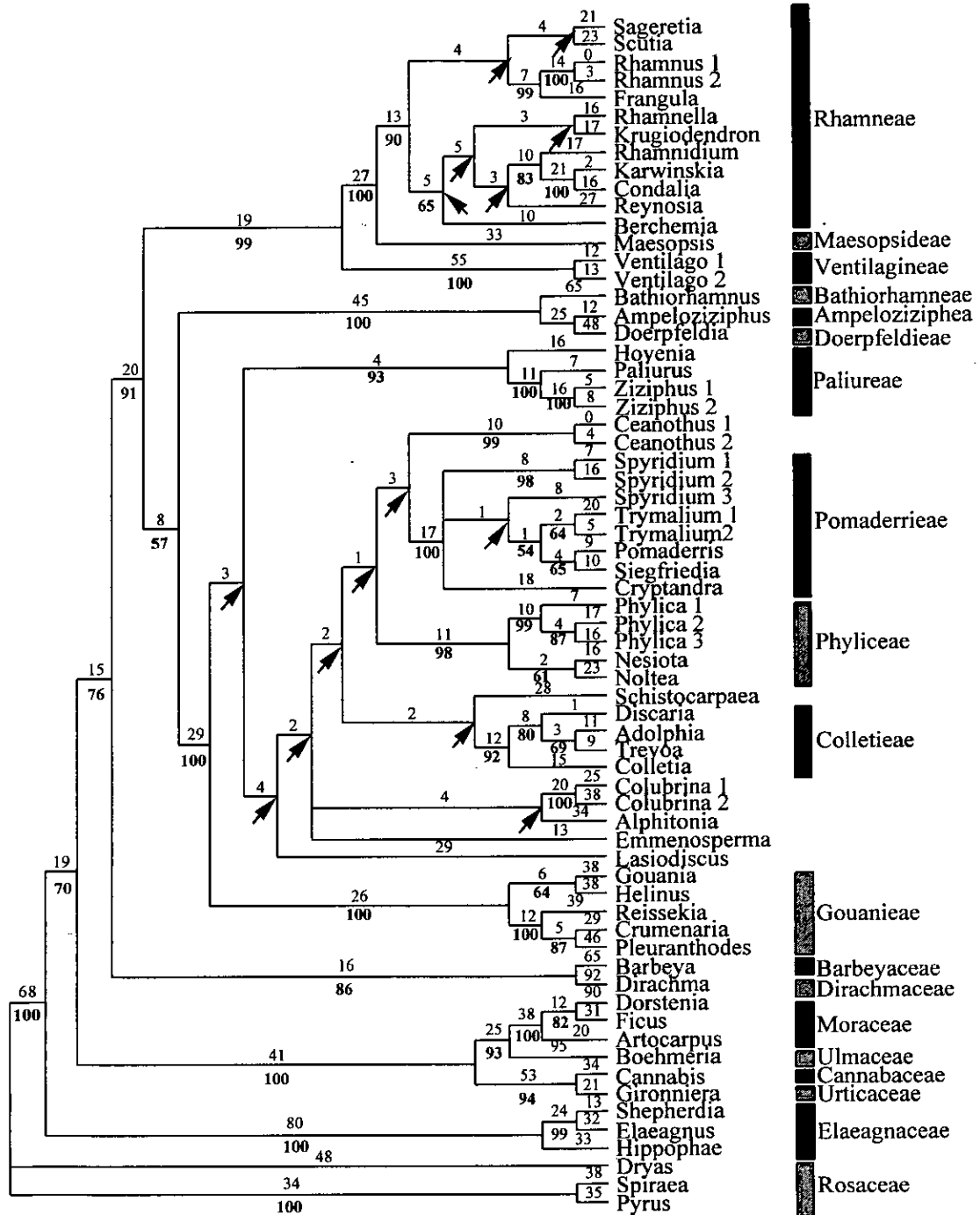


Figure 2.4. The single optimal SW tree from the combined *rbcL/trnL-F* analysis, with its Fitch lengths (above branches; ACCTRAN optimisation) and bootstrap values (below). Branches, which are not present in the strict consensus tree are indicated by an arrow. Heuristic search under the Fitch criterion produced 324 Fitch trees with a length of 2559 steps, CI=0.59 and RI=0.70. SW produced one tree with two trichotomies and a tree length of 1068277 steps, CI=0.85, RI=0.88 (Fitch length, 2559 steps).

The combined trees show a greater similarity to the *trnL-F* tree than to the *rbcL* tree. Rhamnaceae are monophyletic with a clade consisting of Dirachmaceae and Barbeyaceae forming their sister group. Elaeagnaceae fall on a long-branch nearest the outgroup. The ziziphoid, rhamnoid and ampeloziziphoid groups are again strongly supported as are the groups within these clades, which were strongly supported in the separate analyses.

#### 2.4.4. Molecular Evolution

Figure 2.5 shows a plot of the number of changes per character optimised on the single most parsimonious SW tree from the combined analysis. The *trnL-F* plot has a more even distribution of substitutions along its length than *rbcL*. Figure 2.6 shows the number of characters per number of steps on both the *trnL-F* and *rbcL* trees. The *rbcL* graph indicates that some characters change up to 16 times on the combined SW tree whereas the *trnL-F* characters change up to nine times only. This justifies the use of SW which downweights only those characters which change frequently.

Table 2.4 shows that in the *rbcL* analysis the third position of codons has by far the greatest number of steps followed by the first position, followed by the second. The CI value is highest for the second position followed by the first and the third. However, the RI value is highest for the third position, followed by the first, followed by the second. Table 2.5 shows that most of the *trnL-F* indel characters have maximum CI and RI values.

Table 2.4. Performance of each codon position in the *rbcL* analysis.

Codon position	Number of steps	CI	RI	RC
1	277	0.47	0.50	0.24
2	167	0.57	0.44	0.25
3	747	0.51	0.70	0.36



*trnL-F*

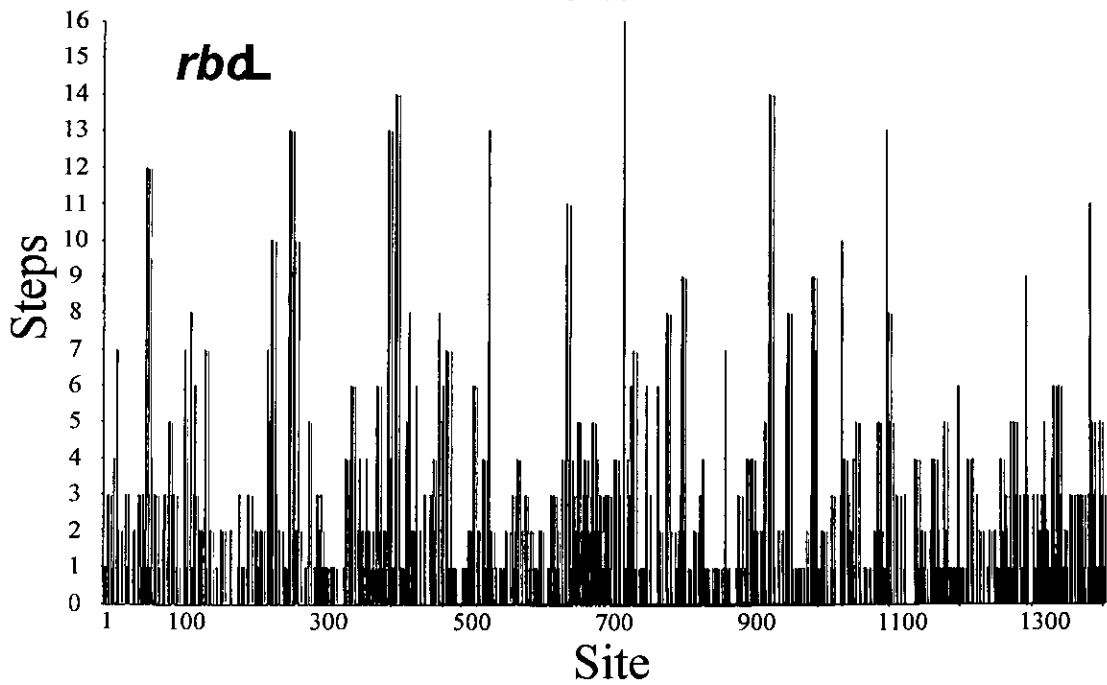
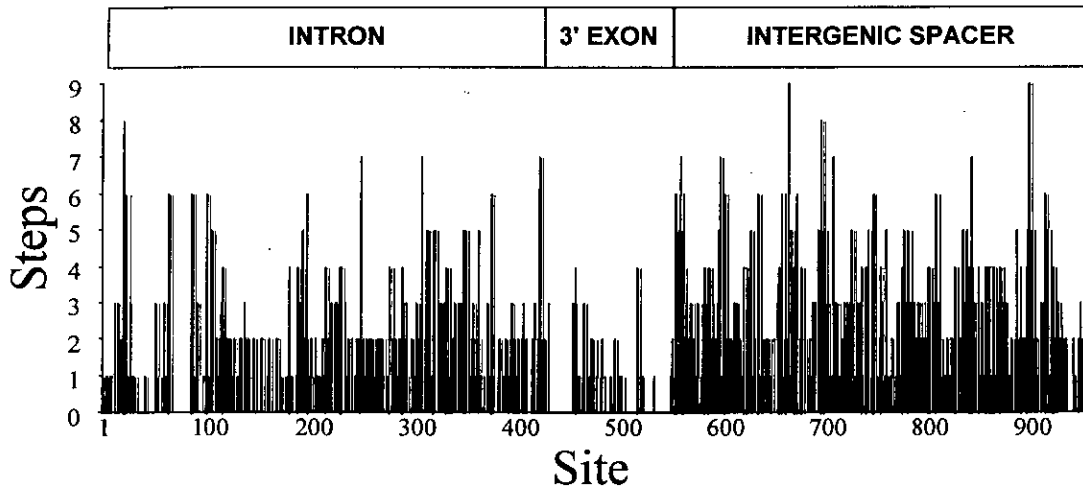


Figure 2.5. Number of changes per character based on the single SW tree from the combined *rbcL/trnL-F* analysis.

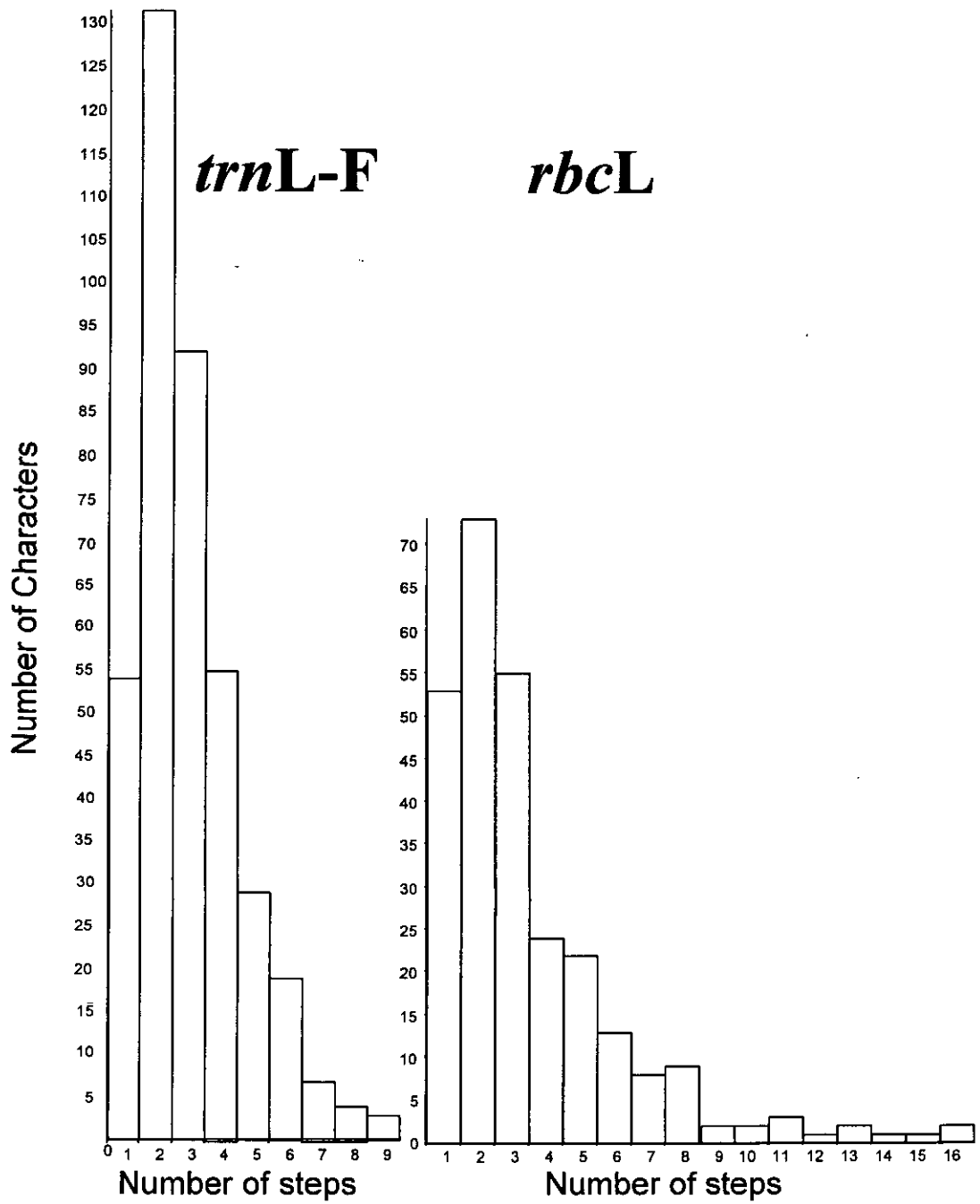


Figure 2.6. Number of steps for each of the variable sites produced on the single SW tree from the combined *rbcL/trnL-F* analysis.

Table 2.5. Performance of *trnL*-F indel characters.

Indel Character	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
CI	1	1	1	1	1	1	1	1	0.25	0.50	1	1	1	1	0.50	0.25
RI	1	1	1	1	1	1	1	1	0.50	0.67	1	1	1	1	0.80	0.40

The transition/transversion ratio for the *rbcL* data matrix calculated on the combined SW tree was 1.17. The transition/transversion ratios for the *trnL*-F data matrix on the combined SW tree were calculated separately for the intron, exon and non-coding regions. The intron ratio was (258/282) 0.91, the exon ratio was (6/1) 6 and the non-transcribed intergenic spacer region ratio was (340/342) 0.99. The exon ratio cannot be considered significant for such a small number of informative sites. For *rbcL* there is a bias for transitions, whereas the more or less one to one ratio in the non-coding regions of *trnL*-F indicate a lack of such bias. Transitions have higher CI and RI values (Table 2.6) than transversions in both *rbcL* and *trnL*-F when optimised on the combined tree.

Table 2.6. Tree scores for transitions and transversions on an SW tree from the combined *rbcL*/*trnL*-F analysis.

	<i>rbcL</i>			<i>trnL</i> -F		
	transitions	transversions	ratio	transitions	transversions	ratio
Number of steps	646	548	1.17	677	664	1.02
CI	0.553	0.465		0.694	0.620	
RI	0.721	0.567		0.786	0.690	

Table 2.7 shows the tree lengths when analysed alone for *rbcL* and *trnL*-F as well as the number of steps for *rbcL* and *trnL*-F data sets optimised on the combined SW tree. Both of the separate analyses underestimate the number of substitutions indicated on the combined tree. The *trnL*-F region had a 1339/1347 difference in number of steps on the *trnL*-F tree compared to the combined tree, which is a 0.6% underestimate of change in the *trnL*-F tree compared to the combined tree. The *rbcL* gene had a 1174/1194 difference in number of steps on the *rbcL* tree compared to the

combined tree, which is a 1.7% underestimate of change in the *rbcL* tree compared to the combined tree. Thus *rbcL* has a greater underestimate of change than does *trnL-F*.

Table 2.7. Comparison of number of steps for the separate analyses versus the combined trees.

tree	<i>rbcL</i> tree length	<i>trnL-F</i> tree length	length on combined tree	difference	%difference
<i>rbcL</i>	1174		1194	+20	1.7
<i>trnL-F</i>		1339	1347	+8	0.6

## 2.5. Discussion

### 2.5.1. Molecular evolution

In *rbcL* trees Rhamnaceae are paraphyletic with Barbeyaceae, Dirachmaceae and Elaeagnaceae nested within, but this is weakly supported. The *trnL-F* analysis indicates that Rhamnaceae are a strongly supported monophyletic group. There are two possible explanations for this result: either the two data sets are really incongruent, or the nesting of Barbeyaceae, Dirachmaceae and Elaeagnaceae in the *rbcL* tree is an artefact, perhaps the result of a long branch attraction. When two or more branches undergo extensive substitution after taxa diverge, the changes in these long branches may display many parallel changes (homoplasy) which provide support for the wrong tree. Hence Elaeagnaceae, Barbeyaceae and Dirachmaceae are attracted to other branches within Rhamnaceae.

High levels of homoplasy are expected in DNA matrices because the possibility for change at each position is limited to only three options. What is important is not the amount of homoplasy, but rather the distribution or structure of homoplasy. Phylogenetic signal is assumed to be present in all sequence matrices, but overlying this there may be other patterns. Functional constraints exist in protein-coding genes such as *rbcL* (Albert *et al.*, 1994), and third positions in codons are expected to be

more variable than first or second positions, as is the case with this *rbcL* data set (Table 2.4). Because of the degenerate nature of the genetic code, the first and second positions in a codon are under higher levels of direct selection, and therefore fewer of them can change than third positions. In non-coding regions such as *trnL-F* there is probably less functional constraint than there is in *rbcL* (constraints in non-coding regions could involve ribosomal RNA processing control sites and other structural aspects). Rates of change for each of the non-coding characters should be more similar, and this is what was found: *trnL-F* has a more even pattern of change than *rbcL* (Fig. 2.5). Also, a plot of number of characters against number of steps shows that *rbcL* has many more hypervariable positions than *trnL-F* (Fig. 2.6). This uneven pattern of variation in *rbcL* makes it harder to detect all changes (i.e. all the homoplasy) in such positions and is therefore more likely to produce misrepresentations of relationships in the form of branch attractions (i.e. underestimates in the actual amount of change). This uneven pattern of change has led to the differential weighting of different codon positions in phylogenetic analyses (e.g. Birstein and DeSalle, 1998). However, as Table 2.4 indicates the performance of third positions in the *rbcL* analysis, in terms of CI and RI values, is more or less equal to if not better than that of first and second positions, so differential weighting of these characters is therefore not justified.

As discussed above, different matrices contain different degrees of functional constraint and combining them should strengthen only the shared signal present, which is likely to be the phylogenetic one. In general, similar weakly supported patterns of separate data sets would be expected to be more strongly supported when combined. Finally, combining data sets detects evidence for additional substitutions that are not detected in one matrix but are detected when combined with another, thus permitting more accurate overall character reconstruction. As a result combined trees might be expected to be longer than any of the individual matrix trees because combined matrices should recover more of the unobserved substitutions in each individual matrix. This is the case when combining the *rbcL* and *trnL-F* data sets in this study (Table 2.7). The greater underestimate in change for *rbcL* compared to *trnL-F* may have resulted in a branch attraction in the *rbcL* trees. This is the most

probable explanation for the nesting of the families Elaeagnaceae, Barbeyaceae and Dirachmaceae within Rhamnaceae in the analysis of the *rbcL* tree.

A further *rbcL* analysis was run in which the monophyly of Rhamnaceae was constrained. This analysis produced a tree with a Fitch length of 1175, i.e. only one step longer than the non-constrained analysis. The most parsimonious *rbcL* tree is only slightly more optimal than the more accurate one, the combined tree, which has much higher levels of internal support. Such underestimates on single matrices highlight the limitations of too little data in which patterns are too weak for accurate reconstruction, not the unreliability of parsimony as an optimality criterion. The following sections of the discussion will focus mainly on the combined tree which should be more accurate for the reasons explained above.

Thirteen of the 16 indel characters from the *trnL-F* data set were non-homoplasious synapomorphies. Therefore in this analysis indel characters appear to be good phylogenetic markers. Of these characters half appear to be unique sequence and the other half are copies or near copies of adjacent regions.

Differential rates of transitions and transversions have been used to justify differential weighting of character state changes in phylogenetic analyses (Zink and Blackwell, 1998; Smith, 1998; Fu, 1998). In this data set, coding regions have a transition bias whereas introns or non-transcribed spacers have no apparent bias. Transitions (purine-purine and pyrimidine-pyrimidine changes) are expected to occur more readily than transversions (purine-pyrimidine) because they are less likely to be detected by correction mechanisms. The transition bias in *rbcL*, but not in non-coding *trnL-F* (Table 2.6), is consistent with the findings of Morton (1995) who demonstrated that substitutions in non-coding regions of the plastid genome were affected by the two, immediately flanking bases. When both the 5' and 3' flanking nucleotides are G or C only 25% of the observed substitutions are transversions whereas if the flanking nucleotides are both A or T 57% of the substitutions are transversions. Because non-coding regions of the plastid genome are A/T rich, the relative proportion of transversions increases, resulting in a more even transition/transversion ratio. The nearly one to one ratio in *trnL-F* indicates that the application of greater weights to transversions in non-coding regions would not be

justified. Also, the better performance in terms of CI and RI values of transitions over transversions in *trnL-F* and *rbcL* (Table 2.6) indicates that differential weighting of these character state changes is not reasonable.

### 2.5.2. Relationships of Rhamnaceae

The *Dirachma/Barbeya* alliance is strongly supported by the bootstrap. This clade is a sister group to Rhamnaceae in the combined tree in a moderately supported clade. Thulin *et al.* (1998) suggested that the families Barbeyaceae and Dirachmaceae should be retained because they differ so significantly in morphology. The results here also indicate that this would be the best circumscription for these families given the large number of morphological and molecular differences between them, Rhamnaceae, Elaeagnaceae, and other families. Greater sampling from within the urticalean families and Rosaceae may result in a better placement of Barbeyaceae and Dirachmaceae, but their combination of traits otherwise restricted to either Rhamnaceae or the urticalean families would appear to indicate either a position as obtained here or as sister to the urticalean families.

### 2.5.3. Relationships within Rhamnaceae

Classification based solely on DNA sequence data should be treated with caution unless backed up by evidence from other sources. It has however, indicated patterns which were not apparent from studies of morphological and anatomical characteristics. The single SW tree from the combined analysis shows that Rhamnaceae are a well supported monophyletic group and also provides support for some of Suessenguth's tribes. However, these molecular data show a division of Rhamnaceae into three clades which are supported by bootstrap values of 99 or 100, but for which there are no obvious morphological apomorphies. Such groups were described as "cryptic clades" (Wojciechowski *et al.*, 1993) in a study that identified a strongly supported clade of aneuploid North American *Astragalus* which was found to be supported by three different lines of genotypic evidence (chromosomal, nuclear

rDNA and plastid DNA). However, there were no morphological characters to support this grouping, and the authors suggested that the group should be given an informal name. I have likewise chosen to adopt informal names for the three cryptic clades identified here.

Group 1: rhamnoid clade - This clade is divided into three strongly supported subgroups. The first of these comprises the tribe Rhamneae Hook. f. and includes genera such as *Rhamnus* and *Berchemia* which have drupaceous fruits, superior ovaries and a nectariferous disc either partly or totally adnate to the calyx tube. The inter-relationships of the genera within this group are not particularly well supported. The second subgroup, Maesopsidae Weberb., consists of the monotypic genus *Maesopsis* which is a sister to Rhamneae and forms the monotypic tribe Maesopsidae. Ventilagineae Hook. f. is the third distinct subgroup with strong support as sister to the *Maesopsis*-Rhamneae alliance. All members of this tribe are climbers with apically winged fruits and semi-inferior ovaries. No sequence data have been gathered for *Smythea*, which is the only other genus previously placed in this tribe. However, this genus is morphologically very similar to *Ventilago* and should be included in the tribe Ventilagineae.

Group 2: ampeloziziphoid clade - This group consists of three highly divergent genera, which have palmately veined leaves and drupaceous fruits: *Ampeloziziphus*, a monotypic genus from Brazil, which is a climber with semi-inferior ovaries and a thick nectariferous disc; *Doerpfeldia*, a monotypic genus from Cuba which is a tree with small leaves and a superior ovary thinly covered by the nectariferous disc; and *Bathiorhamnus*, a genus of two species from Madagascar which are trees with a superior ovary and a thick nectariferous disc. There are, however, no obvious exclusive morphological similarities linking these genera. The high levels of molecular divergence between these genera indicate that they are only distantly related, and it is likely that they are remnants of groups, which were formerly more diverse and widespread. These three should be placed at tribal level because of their highly divergent nature.

Group 3: ziziphoid clade - The third major clade within Rhamnaceae comprises genera which usually have semi-inferior to inferior ovaries and capsular fruits. There



are, however, exceptions to this, e.g. *Ziziphus* and *Paliurus* have drupaceous fruits. In addition some genera of the tribe Colletieae Reiss. ex Endl. have superior ovaries or drupaceous fruits. This ziziphoid group may be further split into seven subgroups. Suessenguth's more derived tribes Colletieae Reiss. ex Endl. and Gouanieae Reiss. ex Endl. are strongly supported monophyletic groups. Gouanieae are climbers with tendrils and longitudinally winged fruits; Colletieae are a group of strongly armed trees or shrubs. An Australian tribe, Pomaderrieae Reiss. ex Endl. are characterised by the presence of stellate hairs. *Ziziphus*, *Paliurus*, and *Hovenia* make up another strongly supported tribe, Paliureae Reiss. ex Endl. *Hovenia* appears to have a close relationship with *Ziziphus* and *Paliurus* in that they all have palmately veined leaves, cymose inflorescences, a base chromosome number of  $x=12$  and a similar pollen exine structure. This relationship is also strongly supported in the combined tree. On the basis of this evidence *Hovenia* is placed in Paliureae. A strongly supported, predominantly South African clade, Phyliceae Reiss. ex Endl., consisting of *Phylica*, *Nesiota*, and *Noltea* also appears distinct and is generally characterised by having an ericoid shrubby habit, inferior ovaries, and leaves with revolute margins and tomentose undersurfaces.

A further distinct clade comprises *Colubrina* which includes trees or shrubs with the nectariferous disc filling the receptacle and surrounding the ovary. The genus *Lasiodiscus* was always thought to be closely related to *Colubrina* (Johnston, 1971), but only the *rbcL* matrix produced trees in which *Colubrina* and *Lasiodiscus* form a clade. Further sampling of the genus *Lasiodiscus* and studies of other sequences might be necessary to lend more molecular support for a *Colubrina/Lasiodiscus* grouping. The two genera are similar morphologically (Figueiredo, 1995) and may eventually be treated as a distinct tribe. However there is insufficient evidence to recognize this group at the present.

The affinities of a number of other genera are unclear. The arborescent genus *Alphitonia* from Malaysia, Australia, and the western Pacific have exocarps that are thick, spongy, and friable at maturity. *Emmenosperma* is similar to *Alphitonia* in that it shares the characteristic of having red arillate seeds persisting on the receptacle after dehiscence. Again further evidence is needed to place these two genera in a

separate tribe. According to the *trnL-F* and combined analyses, *Schistocarpha* appears to be reasonably closely related to the tribe Colletieae. However, there are few morphological characters which support this link.

The North American genus *Ceanothus* is characterised by having receptacles and nectariferous discs persisting on the pedicel and its relationship with the other clades is unresolved. *Ceanothus* and Colletieae engage in root nodular fixation of nitrogen in a symbiotic association with the cyanobacterium *Frankia*. Soltis *et al.* (1995) stated that although all members of a particular clade may have the ability to form such an association, only a few actually do. The positions of these two groups within the ziziphoid clade interspersed with genera that do not form such associations supports this idea. However, the relationships between the nitrogen fixing groups are not well resolved and it is possible that *Ceanothus* and Colletieae are sisters in which case nitrogen fixation may have arisen only once in the family.

A re-classification of tribes in Rhamnaceae is summarised in Table 2.8 and presented in full in Chapter Three. Eleven tribes are now recognised, three of which are new (Ampelozizipheae, Doerpfeldieae and Bathiorhamneae), the constitution of Rhamneae Hook. f. has been emended and the name of one tribe has been corrected (Zizipheae Brongn. to Paliureae Reiss. ex Endl.) as suggested by Schirarend *et al.* (1994) and emended. Ventilagineae Hook.f., Colletieae Reiss. ex Endl. and Gouanieae Reiss. ex Endl. are retained. Pomaderreae Reiss. ex Endl. and Maesopsidae Weberb. have been resurrected, as has Phylliceae Reiss. ex Endl. which has also been emended. The distribution of these tribes is also presented in Table 2.8.

Table 2.8. Summary of revised tribal classification of Rhamnaceae.

Tribe	Genera included	Distribution
Paliureae	<i>Paliurus, Ziziphus, Hovenia</i>	tropics and warm temperate regions
Colletieae	<i>Adolphia, Colletia, Discaria, Kentrothamnus, Retanilla, Trevoa</i>	South America, New Zealand, Australia
Phylliceae	<i>Nesiota, Noltea, Phylica</i>	southern Africa, Atlantic and Indian Ocean islands
Gouanieae	<i>Alvimiantha, Crumenaria, Gouania, Helinus, Pleuranthodes, Reissekia</i>	tropical and warm Americas, Africa, Madagascar, NW India, Indian Ocean Islands
Pomaderreae	<i>Blackallia, Cryptandra, Pomaderris, Siegfriedia, Spyridium, Trymalium</i>	Australia, New Zealand
Rhamnaceae	<i>Auerodendron, Berchemia, Berchemiella, Condalia, Dallachya, Karwinskia, Krugiodendron, Reynosia, Rhamnella, Rhamnidium, Rhamnus, Sageretia, Scutia</i>	tropics to northern temperate regions
Maesopsidaeae	<i>Maesopsis</i>	tropical Africa
Ventilagineae	<i>Smythea, Ventilago</i>	Old World tropics
Ampelozizipheae	<i>Ampeloziziphus</i>	Brazil
Doerpfeldieae	<i>Doerpfeldia</i>	Cuba
Bathiorhamnaceae	<i>Bathiorhamnus</i>	Madagascar
Genera incerta sedis	<i>Ceanothus, Emmenosperma, Schistocarpaea, Alphitonia, Colubrina, Lasiodiscus</i>	

#### 2.5.4. Biogeography of Rhamnaceae

Raven and Axelrod (1974) stated that:

“Rhamnaceae are so well represented both in tropical and temperate regions that it is difficult to trace the history of the family.”

Also the lack of a significant fossil record makes assessments of previous distributions speculative. The distributions of the tribes as circumscribed in Chapter Three and Richardson *et al.* (submitted) are given in Table 2.8.

Two general patterns in the distribution of the three major groups within Rhamnaceae can be observed. The ampeloziziphoid group illustrates a pattern of disjunct distribution also found in other groups between northern South America and Madagascar (e.g. Fay *et al.*, 1998). In this case there are long branch lengths and a lack of morphological similarities, indicating that this group has a long history and probably had a much wider distribution that has subsequently been reduced by extinction, particularly in Africa. The other major groups have similarly wide distributions but were not reduced by extinction to the same extent as the ampeloziziphoid group. Overlaid on this pattern, is another, presumably post-Gondwanan, in which groups are more or less restricted to individual plates. Thus I hypothesize that in spite of the lack of a fossil record Rhamnaceae are an old group well distributed before continental drift separated the components of Gondwanaland.

The ziziphoid group is cosmopolitan with a predominantly southern hemisphere distribution and could be of Gondwanan origin with the exception of *Ceanothus* which has a western North American distribution. This indicates that either this whole southern group had a much greater range throughout Gondwanaland and parts of Laurasia (in what is now North America) and has been subsequently restricted in its distribution or that ancestors of *Ceanothus* arrived at their present location by long distance dispersal. California has many relictual taxa from lineages that are otherwise restricted to the Old World or the southern hemisphere; these include species of *Paeonia* (Paeoniaceae), *Odontostomum* (Tecophilaeaceae) and *Fremontodendron* (Bombacaceae of Malvaceae; Bayer *et al.*, in press). Because *Ceanothus* is sister to other clades within the ziziphoid group I do not consider it to be a recent derivative of one of these clades and thus the most likely explanation for its present distribution is that it is relictual and its clade is reasonably old (c. greater than 65 million years).

Gouanieae have a similar distribution to the ampeloziziphoid group, with some genera of the group also being found in Africa. *Colubrina* is predominantly found in northern South America, although species are also found in Asia, Hawaii, Madagascar and South Africa. *Lasiodiscus* is found in Africa and Madagascar, and this distribution may represent the remnants of previously more widespread groups which are now only found on Madagascar or in rain and coastal forests in the tropical

parts of sub-Saharan Africa and east Africa. *Alphitonia*, Pomaderreae and *Schistocarpha* are Australasian taxa, which represent isolated clades. Colletieae are a mostly South American group, but two species of *Discaria* are found in Australia and New Zealand. This is a southern hemisphere disjunction which is also found in other groups such as *Orthrosanthus* (Iridaceae), *Libertia* (Iridaceae), *Berberidopsis* (Flacourtiaceae) and *Eucryphia* (Eucryphiaceae), and these are probably relicts of formerly more widespread groups which were present through southern South America, east Antarctica, Tasmania, New Zealand, and eastern Australia.

Within Rhamnaceae, relationships are not clearly resolved by *trnL-F* and *rbcL* sequence data. A more in-depth molecular study using a more variable region such as ITS and additional taxon sampling is needed to clarify relationships before any biogeographical conclusions can be drawn. However it does form a strongly supported monophyletic unit which has a wide distribution throughout the tropics into northern temperate regions. Ventilagineae are found in the Old World tropics but with a center of diversity in India. Ventilagineae could have had a Gondwanan origin and subsequently spread into Asia when India collided with Asia. More species in each genus throughout the family need to be analyzed to make a fine-scale biogeographic assessment of the family.

More conclusive proof of the origin of Rhamnaceae and its tribes could come from the discovery of Cretaceous fossils from different continents. However the most recent discoveries reviewed by Muller (1981) are from Oligocene deposits. This means that alternative hypotheses such as more recent dispersal over land bridges cannot be completely discounted.

## 2.6. General Conclusions

According to the combined molecular data set Rhamnaceae are a monophyletic group. Further research is necessary to find more evidence from other fields such as anatomy or chemistry, which could provide added support for the “cryptic clades” which are strongly supported by the molecular data. Although there is strong

molecular support for three major divisions in Rhamnaceae, I have been unable to compile a morphological character set which could adequately describe these groups.

What is clear from these results is that the tribes Rhamneae and Zizipheae as circumscribed by Suessenguth are unnatural and a reclassification of some tribes in Rhamnaceae is necessary. The molecular data indicate that many morphological character states have evolved in parallel (e.g. leaf venation patterns, fruit type, and pollen exine architecture), but it is not a simple matter of morphology versus molecules. Classifications based on one particular morphological character (such as Suessenguth's reliance on fruit characters) often do not compare well with those based on other morphological characters. A classification based on molecular data with the support of some morphological characters seems to be a better solution.

The sister groups of *Phyllica* in the molecular analysis were chosen as outgroups for subsequent studies on the genus. *Phyllica* formed a strongly supported monophyletic group with *Nesiota* and *Noltea*. Members of groups closely related to Phylliceae, such as *Ceanothus*, *Colubrina*, *Lasiodiscus*, *Pomadereae*, and *Alphitonia* were used as outgroups for the analysis of *Phyllica*.

## 2.7. Bibliography

Adams, E.N. 1972. Consensus techniques and the comparison of taxonomic trees. *Systematic Zoology* 34: 46-58.

Adanson, M. 1763. *Familles des Plantes*. 2: 304. Paris.

Albert, V.A., A. Backlund, K. Bremer, M.W. Chase, J.R. Manhart, B.D. Mishler & Nixon, K.C. 1994. Functional constraints and *rbcL* evidence for land plant phylogeny. *Annals of the Missouri Botanical Garden* 81: 534-567.

Baillon, H. *Monographie des Celastracées et des Rhamnacées*. 1875. *Histoire des plantes*, vol. 6. Hachette, Paris.

Bayer, C., M.F. Fay, A.Y. de Bruijn, V. Savolainen, C.M. Morton, K. Kubitzki & M.W. Chase. (in press). Support for an expanded family concept of Malvaceae within a recircumscribed order Malvales: a combined analysis of plastid *atpB* and *rbcL* DNA sequences. *Botanical Journal of the Linnean Society*.

Birstein V.J. & R. DeSalle. 1998. Molecular phylogeny of Acipenserinae. *Molecular Phylogenetics and Evolution* 9: 141-155.

Brongniart, A.T. 1827. Mémoire sur la famille des Rhamnées. *Annales des Sciences Naturelles, Séries 1*, 10: 320-386.

Brown, R. 1814. General remarks, geographical and systematical on the botany of Terra Australis. In Flinders, M. [ed.], *A Voyage to Terra Australis* 2, App. 3: 533-613. London.

Capuron, R. 1966. Notes sur quelques Rhamnacées arbustives ou arborescentes de Madagascar. *Adansonia* 6: 116-141.

Carpenter, J.M. 1995. Successive weighting, reliability and evidence. *Cladistics* 10: 215-220.

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

- Cowan, R.S. 1952. Plant explorations of G. Wilson-Browne, S.J in British Guiana. I. Kanunku Mountains. *Brittonia* 7: 404-405.
- Cronquist, A. 1981. *An Integrated System of Classification of Flowering Plants*. Columbia University Press, New York.
- Cronquist, A. 1988. *The Evolution and Classification of Flowering Plants*, 2nd ed. New York Botanical Garden, New York.
- Donoghue, M.J., R.G. Olmstead, J.F. Smith & J.D. Palmer. 1992. Phylogenetic relationships of Dipsacales based on *rbcL* sequences. *Annals of the Missouri Botanical Garden* 79: 333-345.
- Doyle, J.J. & J.L. Doyle. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11-15.
- Endlicher, S. 1836-40. *Genera Plantarum secundum Ordines Naturales disposita*, 1094-1101.
- Farris, J.S. 1969. A successive approximations approach to character weighting. *Systematic Zoology* 18: 374-385.
- Farris, J.S. 1988. Hennig86 version 1.5; manual; software and MSDOS programme.
- Fay, M.F. & M.W. Chase. 1996. Resurrection of Themidaceae for the *Brodiea* alliance, and recircumscription of Alliaceae, Amaryllidaceae and Agapanthoideae. *Taxon* 45: 441-451.
- Fay, M.F., C. Bayer, W.S. Alverson, A.Y. de Bruijn & M.W. Chase. 1998. Plastid *rbcL* sequence data indicate a close affinity between *Diegodendron* and *Bixa*. *Taxon* 47: 43-50.



Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783-791.

Figueiredo, E. 1995. A revision of *Lasiodiscus* (Rhamnaceae). *Kew Bulletin* 50: 495-526.

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

Friis, I. & Vollesen, K. 1980. The identity of the Ethiopian monotypic genus *Tzellemtinia* Chiov. *Botaniska Notiser* 133: 347-349.

Fu, J. 1998. Toward the phylogeny of the family Lacertidae: implications from mitochondrial DNA 12S and 16S gene sequences (Reptilia: Squamata). *Molecular Phylogenetics and Evolution* 9: 118-130.

Gilli, A. 1980. Beiträge zur Flora von Papua-Neu Guinea II. Dicotyledones. *Annalen der Naturhistorischen Museums, Wien* 8: 454-455.

Grey-Wilson, C. 1978. *Alvimiantha*, a new genus of Rhamnaceae from Bahia, Brazil. *Bradea* 2: 287-290.

Higgins, D.G., A.J. Bleasby & R. Fuchs. 1992. CLUSTAL: a new multiple sequence alignment program. *Computer Applications in Biosciences* 8: 189-191.

Hooker, J.D. 1862. Rhamnaceae. In Bentham, G. & Hooker, J. D. [eds.], *Genera Plantarum* 1: 371-386. Reeve & Co., London.

Hutchinson, J. 1959. *The Families of Flowering Plants* volume 1 Dicotyledons. Clarendon Press, Oxford.

Johnston, M.C. 1962. Revision of *Condalia* including *Microrhamnus* (Rhamnaceae). *Brittonia* 14: 332-368.

Johnston, M.C. 1963. Novelties in *Colubrina* including *Cormonema* and *Hybosperma* (Rhamnaceae). *Wrightia* 3: 91-92.

Johnston, M.C. 1964. The fourteen species of *Ziziphus* including *Sarcomphalus* (Rhamnaceae) indigenous to the West Indies. *American Journal of Botany* 51: 1113-1118.

Johnston, M.C. 1971. Revision of *Colubrina* (Rhamnaceae). *Brittonia* 23: 2-53.

Jussieu, A.L. de. 1789. *Genera Plantarum*. Paris.

Kluge, A.G. & A.J. Wolf. 1993. Cladistics - whats in a word? *Cladistics* 9: 183-199.

Lledó, M.D., Crespo, M.B., Cameron, K.M., Fay, M.F. & Chase, M.W. 1998. Systematics of Plumbaginaceae based upon cladistic analysis of *rbcL* sequence data. *Systematic Botany* 23: 21-29.

Maddison, W.P. & D.R. Maddison. 1992. MacClade version 3.04.

Margush, T. & F.R. McMorris. 1981. Consensus n-trees. *Bulletin of Mathematical Biology* 43: 239-244.

Medan, D. 1988. Gynoecium ontogenesis in the Rhamnaceae: A comparative study. *In* Leins, P., S.C. Tucker & P.K. Endress [eds.], *Aspects of Floral Development*, pp. 133-141. Cramer, Berlin, Stuttgart.

Morgan, D.R., Soltis, D.E. & Robertson, K.R. 1994. Systematic and evolutionary implications of *rbcL* sequence variation in Rosaceae. *American Journal of Botany* 81: 890-903.

Morton, B.R. & M.T. Clegg. 1995. Neighbouring base composition is strongly correlated with base substitution bias in a region of the chloroplast genome. *Journal of Molecular Evolution* 41: 597-603.

Morton, C.M., S.A. Mori, G.T. Prance, K.G. Carol & M.W. Chase. 1997. Phylogenetic relationships of Lecythidaceae: a cladistic analysis using *rbcL* sequence and morphological data. *American Journal of Botany* 84: 530-540.

Muller, J. 1981. Fossil pollen records of extant angiosperms. *Botanical Review* 47: 1-142.

Price, R. & J.D. Palmer. 1993. Phylogenetic relationships of the Geraniaceae and Geraniales from *rbcL* sequence comparisons. *Annals of the Missouri Botanical Garden* 80: 661-671.

Raven, P.H. & D.I. Axelrod. 1974. Angiosperm biogeography and past continental movements. *Annals of the Missouri Botanical Garden* 61: 539-673.

Reeves, G., M.W. Chase, T. de Chies, A.V. Cox, P. Goldblatt, B. Lejeune, P.J. Rudall & M.F. Fay. 1997. Molecular systematics of Iridaceae: a combined analysis of three plastid DNA sequence matrices. *American Journal of Botany* 84: 656 [abstract].

Richardson, J.E., M.F. Fay & M.W. Chase. (submitted). A revision of the tribal classification of Rhamnaceae. *Kew Bulletin*.

Savolainen, V., C.M. Morton, S.B. Hoot & M.W. Chase. 1996. An examination of phylogenetic patterns of plastid *atpB* sequences among eudicots. *American Journal of Botany* 83: 190 [abstract].

Schirarend, C. & M.N. Olabi. 1994. Revision of the genus *Paliurus* Tourn. ex Mill. (Rhamnaceae). *Botanische Jahrbücher* 116: 333-359.

Sheahan, M.C. & M.W. Chase. 1996. A phylogenetic analysis of Zygophyllaceae R.Br. based on morphological, anatomical and *rbcL* DNA sequence data. *Botanical Journal of the Linnean Society* 122: 279-300.

Smith, M.F. 1998. Phylogenetic relationships and geographic structure in pocket gophers in the genus *Thomomys*. *Molecular Phylogenetics and Evolution* 9: 1-14.

Sokal, R.R. & F.J. Rohlf. 1981. *Biometry*, Second Edition. Freeman & Co., San Francisco.

Soltis, D.E., P.S. Soltis, D.L. Nickrent, L.A. Johnson, W.J. Hahn, S.B. Hoot, J.A. Sweere, R.K. Kuzoff, K.A. Kron, M.W. Chase, S.M. Swensen, E.A. Zimmer, S.M. Chaw, L.J. Gillespie, W.J. Kress, K.J. Sytsma. 1997. Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. *Annals of the Missouri Botanical Garden* 84: 1-49.

Soltis, D.E., P.S. Soltis, D.R. Morgan, S.M. Swensen, B.C. Mullins, J.M. Dowd & P. Martin. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proceedings of the National Academy of Sciences, USA* 92: 2647-2651.

Soltis, D.E., D.R. Morgan, A. Grable, P.S. Soltis, & R. Kuzoff. 1993. Molecular systematics of Saxifragaceae *sensu stricto*. *American Journal of Botany* 80: 1056-1081.

Standley, P.C. 1925. New plants from Central America-III. *Journal of the Washington Academy of Science* 15: 285.

Steenis, C.G.G.J. van 1982. Critical notes on New Guinea plants described by A. Gilli. *Blumea* 28: 165-169.

Suessenguth, K. 1953. Rhamnaceae, Vitaceae, Leeaceae. *In* Engler, A. & K. Prantl [eds.], *Die Natürlichen Pflanzenfamilien* 2, Aufl., 20d. Duncker & Humblot, Berlin.

Swensen, S.M. 1996. The evolution of actinorhizal symbioses: evidence for multiple origins of the symbiotic association. *American Journal of Botany* 83: 1503-1512.

Swofford, D.L. 1993. PAUP: phylogenetic analysis using parsimony, version 3.1.1. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.

Swofford, D.L. & W.P. Maddison. 1987. Reconstructing ancestral character states under Wagner parsimony. *Mathematical Bioscience* 87: 199-229.

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

Takhtajan, A. 1980. Outline of the classification of flowering plants (Magnoliophyta). *Botanical Review* 46: 226-359.

Takhtajan, A. 1997. *Diversity and Classification of Flowering Plants*. Columbia University Press, New York.

Thorne, R.F. 1992. Classification and Geography of Flowering Plants. *Botanical Review* 58: 225-348.

Thulin, M., B. Bremer, J.E. Richardson, J. Niklasson, M.F. Fay & M.W. Chase. 1998. Family relationships of the enigmatic rosid genera *Barbeya* and *Dirachma* from the Horn of Africa region. *Plant Systematics and Evolution* 213: 103-119.

Tortosa, R.D. 1992. El complejo *Retanilla-Talguenea-Trevoa* (Rhamnaceae). *Darwiniana* 31: 223-252.

Vent, W. 1962. Monographie der gattung *Oreoherzogia* W. Vent, *gen. nov.* Feddes Repertorium 65: 3-132.

Weberbauer, A. 1895. Rhamnaceae. In A. Engler & K. Prantl [eds.], Die Natürlichen Pflanzenfamilien. Duncker & Humblot, Berlin.

Williams, E.S., V.A. Albert & M.W. Chase. 1994. Relationships of Droseraceae: a cladistic analysis of *rbcL* sequence and morphological data. *American Journal of Botany* 81: 1027-1037.

Wojciechowski, M.F., M.J. Sanderson, B.G. Baldwin & M.J. Donoghue. 1993. Monophyly of aneuploid *Astragalus* (Fabaceae): Evidence from nuclear ribosomal DNA internal transcribed spacer sequences. *American Journal of Botany* 80: 711-711.

Xiang, Q.-Y., D.E. Soltis, D.R. Morgan & P.S. Soltis. 1993. Phylogenetic relationships of *Cornus* L. *sensu lato* and putative relatives inferred from *rbcL* sequence data. *Annals of the Missouri Botanical Garden* 80: 723-734.

Zink, R.M. & R.C. Blackwell. 1998. Molecular systematics and biogeography of aridland gnatcatchers (Genus *Polioptila*) and evidence supporting species status of

the California gnatcatcher (*Polioptila californica*). *Molecular Phylogenetics and Evolution* 9: 26-32.

**CHAPTER THREE. MORPHOLOGICAL  
PHYLOGENETIC ANALYSIS OF  
RHAMNACEAE**



## **CHAPTER THREE. Morphological Phylogenetic Analysis Of Rhamnaceae**

### **Abstract**

A morphological phylogenetic analysis of Rhamnaceae using 18 characters provided less resolution than analysis of molecular characters. Mapping characters onto a tree from a combined analysis provides more accurate information on how particular morphological characters have evolved, e.g. the apparently parallel development of nitrogen fixation. The molecular study from the previous chapter when used in conjunction with certain morphological characters provides the basis for a new tribal classification of the family. The tribes are described on the basis of their molecular groupings and morphology. Eleven tribes are now recognised, three of which are new (Ampelozizipheae, Doerpfeldieae and Bathiorhamneae), the constitution of Rhamneae Hook. f. has been emended and the name of one tribe has been corrected (Zizipheae Brongn. to Paliureae Reiss. ex Endl.) and emended. Ventilagineae Hook.f., Colletieae Reiss. ex Endl. and Gouanieae Reiss. ex Endl. are retained. Pomaderreae Reiss. ex Endl. and Maesopsidaeae Weberb. have been resurrected, as has Phylliceae Reiss. ex Endl. which has also been emended.

### **3.1. Introduction**

A preliminary morphological phylogenetic analysis of Rhamnaceae was undertaken to determine the usefulness of the available morphological characters in reconstructing phylogeny in this family. Problems with the use of morphological characters in Rhamnaceae were outlined in the previous chapter. One of these problems has been reliance on a small number of morphological characters to delimit tribes, such as the use of fruit characters by Suessenguth (1953). Other characters used by Suessenguth (1953) are also potentially prone to developmental plasticity, e.g. disc and ovary position. There is a lack of morphological characters that can be used for phylogenetic analyses at the supra-generic level. The aim of this chapter is to illustrate the use of morphological characters in phylogenetic analysis in

comparison with the use of molecular characters from the previous chapter and to combine morphological and molecular data in a total evidence approach. Subsequent mapping of morphological characters onto a tree from a combined molecular and morphological analysis will be used to illustrate how they have evolved. For example a close relationship between species with nitrogen-fixing, bacterial symbioses in some Rhamnaceae, Elaeagnaceae, Ulmaceae, and Rosaceae has been demonstrated by Soltis *et al.* (1995) and Swensen *et al.* (1996). The number of times this feature has arisen in Rhamnaceae could be determined by mapping this character onto a combined morphological/molecular tree.

### **3.2. Methods**

I scored eighteen unordered characters for members of each of the genera in Rhamnaceae and *Dirachma* and *Barbeya*. The eighteen characters used in the analysis are presented in Table 3.1 and the character-state matrix in Table 3.2. Most of the characters chosen were those which had previously been used by Suessenguth (1953). The operational taxonomic units for this study were the individual species in Rhamnaceae that were included in the molecular analysis from the previous chapter, plus *Barbeya* and *Dirachma*. Information about character states was derived from studies of literature (e.g. Suessenguth, 1953 and monographs of individual genera listed in the taxonomic section of this chapter) and herbarium specimens. Fruit type, fruit appendages, number of locules per ovary and ovary position were all used by Suessenguth (1953) to delimit tribes in his system.

#### **3.2.1. Description of characters**

1. In some genera the seed remains attached to the torus after dehiscence. This is coded as a two-state character.
2. Disc present/absent is a simple two-state character.

3. In cases in which a disc is present, there are three character states. The disc may be 1. adnate to the calyx tube and the ovary, i.e. filling the calyx tube, 2. adnate to the calyx tube only, i.e. the ovary is free, or 3. adnate to the ovary only.
4. Leaf margins can be revolute or more or less flat in Rhamnaceae.
5. Some groups in Rhamnaceae form symbiotic associations with bacteria, a two-state character.
6. In most genera of Rhamnaceae, the number of locules per ovary is usually either two or three. *Maesopsis* is an exception with one locule per ovary. In certain instances individuals or species which have two locules per ovary may also have four locules per ovary and individuals or species which have three locules per ovary may have four locules per ovary. However, in the majority of cases taxa have either two or three locules per ovary so this character is given three states, number of locules per ovary one, two or four, or usually three.
7. Presence/absence of endosperm.
8. Rhamnaceous fruits are either drupes or capsules.
9. Leaf venation is either pinnate or palmate.
10. Rhamnaceous hairs are either simple or stellate.
11. Longitudinal wings in the tribe Gouanieae are derived from the ovary wall. This character has two states: fruit longitudinal wings present/absent.
12. Apical wings in the tribe Ventilagineae are derived from the ovary wall and the style. This character has two states: fruit apical wings present/absent.
13. The scoring of ovary position is problematic because it is often not clear which state to assign for each taxonomic unit. Within some genera these characters are not discrete due to developmental plasticity. A more detailed study of ovary development, similar to that undertaken by Soltis *et al.* (1992) on *Lithophragma* (Saxifragaceae), may be necessary to properly code these characters. However, as such a study is beyond the scope of this project and because of the limited number of suitable characters available for the Rhamnaceae study I have decided to include these characters in the analysis, with three states: inferior, semi-inferior or superior.
14. The habit character is coded as either trees/shrubs or climbers/herbs. In many rhamnaceous genera different species can be either trees or shrubs (the distinction

between which is arbitrary). The only herb in the family is *Crumenaria* in tribe Gouanieae. The herbaceous habit of this species is a reduction from the climbing form present in all other genera in this tribe (Suessenguth, 1953). I therefore coded habit as a two-state character i.e. trees/shrubs or climbers/herbs.

15. Leaves may be arranged alternately, opposite or in whorls.

16. Tendril presence/absence is a simple two-state character.

17. Sepals may have a keel running along their midrib or not.

18. Stamens and petals may be arranged alternate to the sepals or the arrangement of floral parts may be otherwise.

### 3.2.2. Phylogenetic analysis

*Barbeya* (Barbeyaceae) and *Dirachma* (Dirachmaceae) were used as outgroups in this analysis because they are the sister group to Rhamnaceae in the molecular analysis from the previous chapter. I analysed three data sets: 1. the morphological matrix, 2. the combined *rbcL/trnL-F* molecular data set including only those taxa which were included in the morphological analysis to enable a more accurate comparison with the morphological trees and 3. morphological and molecular data sets combined. For all three matrices data were analysed using the parsimony algorithm of the software package PAUP version 3.1.1 for Macintosh (Swofford, 1993). Tree searches were conducted under the equal weights criterion (Fitch, 1971) with 1000 random taxon additions and TBR (tree bisection-reconnection) swapping, but permitting only five trees to be held at each step. All shortest trees collected in the 1000 replicates were then used as starting trees for another round of heuristic search, and all these trees were swapped on to completion. One thousand replicates of the bootstrap (Felsenstein, 1985) were then carried out applying the same strategy and scheme of support as for the molecular analysis (Chapter Two) except that successive weights were not applied. This was done because bootstrapping with SW applied is potentially unreliable if there is little variability in the data set (as is the case with the morphological data set).

Table 3.1. Characters used in a morphological phylogenetic analysis of Rhamnaceae.

Character	Character state
1. seed attachment	1. attached to torus after dehiscence 2. falling from torus after dehiscence
2. disc presence/absence	1. disc present 2. disc absent
3. disc position	1. adnate to calyx tube and ovary 2. adnate to calyx tube or free 3. adnate to ovary only
4. leaf margin	1. revolute 2. not revolute
5. nitrogen fixation	1. present 2. absent
6. number of locules per ovary	1. usually 3 2. 2 or 4 3. 1
7. endosperm	1. present 2. absent
8. fruit	1. capsule 2. drupe
9. leaf venation	1. palmate 2. pinnate
10. stellate hairs	1. present 2. absent
11. fruit with longitudinal wings	1. absent 2. present
12. fruit with apical wings	1. absent 2. present
13. ovary position	1. superior 2. semi-inferior 3. inferior
14. habit	1. trees or shrubs 2. climbers or herbs
15. leaf position	1. alternate 2. opposite 3. whorled
16. tendrils	1. present 2. absent
17. calyx keel	1. present 2. absent
18. arrangement of floral parts	1. stamens and petals alternating with sepals 2. stamens and petals not alternating with sepals

### 3.3. Results

The morphological analysis produced 5000 trees with a length of 50 with CI=0.44 and RI=0.83. One of the trees from the heuristic search is shown in Figure 3.1. These trees do not show the three major groups evident in the molecular trees. However, they do identify most of the tribal groups (*sensu* Richardson *et al.*, submitted) within Rhamnaceae although support for these groups is low or less than 50%, and relationships between them are not resolved in the strict consensus tree. Suessenguth's tribes Rhamneae and Zizipheae are not monophyletic but Gouanieae, Colletieae and Ventilagineae are (although the latter is monogeneric here).

The molecular analysis produced 942 trees with a length of 1660 with CI=0.65 and RI=0.76. One of the trees from the heuristic search is shown in Figure 3.2. These results are nearly identical to those in the previous chapter (i.e. slightly different sampling does not affect the trees produced). The strict consensus trees for the morphological analysis and the combined *rbcL* and *trnL-F* molecular analysis are shown in Figure 3.3.

The combined morphological and molecular analysis produced 216 trees with a length of 1726, CI=0.64 and RI=0.76. One of the trees from the heuristic search is shown in Figure 3.4. The topology of the combined morphological/molecular trees is more or less the same as that of the molecular analysis (Chapter Two). Individual morphological characters were mapped onto one of the combined trees to visualise their evolution (Figure 3.5). Table 3.3 shows the CI and RI values for each of the individual morphological characters in the morphological analysis and in the combined morphological and molecular analysis. Bootstrap values in the combined morphological/molecular analysis are slightly higher (with one exception) than in the molecular analysis alone.

Table 3.2. Matrix of character states for a morphological analysis of Rhamnaceae.

Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
<i>Sageretia thea</i> (Osbeck) M.C. Johnston	2	1	2	2	2	2	1	2	2	2	1	1	3	1	2	2	1	1
<i>Rhamnus lycioides</i> L.	2	1	2	2	2	2	1	2	2	2	1	1	1	1	1	2	1	1
<i>Rhamnus cathartica</i> L.	2	1	2	2	2	2	1	2	2	2	1	1	1	1	1	2	1	1
<i>Rhamnus frangula</i> L.	2	1	2	2	2	2	1	2	2	2	1	1	1	1	1	2	1	1
<i>Rhamnella franguloides</i> (Maxim.) Weberb.	2	1	2	2	2	2	1	2	2	2	1	1	3	1	1	2	1	1
<i>Krugiodendron ferreum</i> (Vahl) Urban	2	1	2	2	2	2	1	2	2	2	1	1	?	1	1	2	1	1
<i>Rhamnidium elaeocarpum</i> Reiss.	2	1	2	2	2	2	2	2	2	2	1	1	1	1	2	2	1	1
<i>Karwinskia humboldtiana</i> (Roem. & Schult) Zucc.	2	1	2	2	2	2	1	2	2	2	1	1	1	1	2	2	1	1
<i>Condalia microphylla</i> Cav.	2	1	2	2	2	2	1	2	2	2	1	1	1	1	3	2	1	1
<i>Scutia buxifolia</i> Reiss.	2	1	2	2	2	2	2	2	2	2	1	1	1	1	1	2	1	1
<i>Berchemia discolor</i> (Klotch) Hemsley	2	1	2	2	2	2	1	2	2	2	1	1	1	1	1	2	1	1
<i>Reynosia uncinata</i> Urban	2	1	2	2	2	3	1	2	2	2	1	1	1	1	2	2	1	1
<i>Maesopsis eminii</i> Engl.	2	1	2	2	2	3	1	2	2	2	1	1	1	1	2	2	1	1
<i>Ventilago viminalis</i> Hook.	2	1	1	2	2	2	2	?	2	2	1	2	2	2	1	2	1	1
<i>Ventilago leiocarpa</i> Benth.	2	1	1	2	2	2	2	?	2	2	1	2	2	2	1	2	1	1
<i>Bathiorhamnus cryptophorus</i> Capuron	2	1	1	2	2	1	1	2	1	2	1	1	1	1	1	2	1	1
<i>Ampeloziziphus amazonicus</i> Ducke	2	1	1	2	2	1	2	2	1	2	1	1	3	2	1	2	1	1
<i>Doerpfeldia cubensis</i> Urban	2	1	3	2	2	2	1	2	1	2	1	1	1	1	1	2	1	1
<i>Ceanothus coeruleus</i> Lag.	2	1	1	1	1	1	1	1	2	2	1	1	1	1	2	2	1	1

<i>Ceanothus thyrsiflorus</i> Esch.	2	1	1	1	1	1	1	1	2	2	1	1	1	1	2	2	1	1
<i>Gouania mauritiana</i> Lam.	2	1	1	2	2	1	1	1	1	2	2	1	3	2	1	1	1	1
<i>Reissekia smilacina</i> Endl.	2	1	1	2	2	1	1	1	2	2	2	1	3	2	1	1	1	1
<i>Crumenaria erecta</i> Reiss.	2	1	1	2	2	1	1	1	?	2	2	1	3	2	1	2	1	1
<i>Helinus integrifolius</i> Kuntze	2	1	1	2	2	1	1	1	2	2	2	1	3	2	1	1	1	1
<i>Pleuranthodes hillebrandii</i> (Oliver) Weberb.	2	1	1	2	2	2	1	1	?	2	2	1	3	2	1	1	1	1
<i>Schistocarpaea johnsonii</i> F.v. Muell.	2	1	?	?	2	1	1	1	2	2	1	1	1	1	1	2	1	1
<i>Colubrina asiatica</i> Brongn.	2	1	1	2	2	1	1	1	2	2	1	1	3	1	1	2	1	1
<i>Colubrina reclinata</i> (L'Hér.) Brongn.	2	1	1	2	2	1	1	1	2	2	1	1	3	1	1	2	1	1
<i>Lasiodiscus mildbraedii</i> Engl.	2	1	1	2	2	1	1	1	2	2	1	1	3	1	2	2	1	1
<i>Emmenosperma alphitonioides</i> F.Muell.	1	1	1	2	2	2	1	1	2	2	1	1	?	1	2	2	1	1
<i>Alphitonia excelsa</i> Reiss.	1	1	1	2	2	2	1	1	2	2	1	1	3	1	1	2	1	1
<i>Smythea lanceolata</i> Summerhayes	2	1	1	2	2	2	2	?	2	2	1	2	2	2	1	2	1	1
<i>Paliurus spina-christi</i> Mill.	2	1	1	2	2	2	2	2	1	2	1	1	2	1	1	2	1	1
<i>Ziziphus glabrata</i> Roxb.	2	1	1	2	2	2	2	2	1	2	1	1	2	1	3	2	1	1
<i>Ziziphus ornata</i> Miq.	2	1	1	2	2	2	2	2	1	2	1	1	2	1	3	2	1	1
<i>Hovenia dulcis</i> Thunb.	2	1	1	2	2	1	1	1	1	2	1	1	1	1	1	2	1	1
<i>Phylica pubescens</i> Ait.	2	1	1	1	2	1	1	1	2	2	1	1	3	1	1	2	1	1
<i>Phylica polifolia</i> (Vahl) Pillans	2	1	1	1	2	1	1	1	2	2	1	1	3	1	1	2	1	1
<i>Phylica stipularis</i> L.	2	1	1	1	2	1	1	1	2	2	1	1	3	1	1	2	1	1



<i>Phyllica nitida</i> Lam.	2	1	1	1	2	1	1	1	2	2	1	1	3	1	1	2	1	1
<i>Phyllica tropica</i> Baker	2	1	1	1	2	1	1	1	2	2	1	1	3	1	1	2	1	1
<i>Nesiota elliptica</i> (Roxb.) Hook. f.	2	1	1	1	2	1	1	1	2	2	1	1	3	1	2	2	1	1
<i>Noltea africana</i> (L.) Reichb.	2	1	1	2	2	1	1	1	2	2	1	1	2	1	1	2	1	1
<i>Discaria chacaye</i> (G. Don) R.D. Tortosa	2	1	2	2	1	1	1	1	1	2	1	1	2	1	2	2	1	1
<i>Colletia ulicina</i> Gill. & Hook.	2	1	2	2	1	1	1	1	2	2	1	1	2	1	2	2	1	1
<i>Adolphia infesta</i> (H.B.K.) Meisn.	2	1	2	2	1	1	1	1	?	2	1	1	1	1	2	2	1	1
<i>Trevoa trinervis</i> Miers	2	2	?	2	1	1	1	2	1	2	1	1	2	1	2	2	1	1
<i>Spyridium</i> cf. <i>forrestianum</i>	2	1	1	1	2	1	1	1	2	1	1	1	3	1	1	2	1	1
<i>Spyridium globulosum</i> (Labill.) Benth.	2	1	1	1	2	1	1	1	2	1	1	1	3	1	1	2	1	1
<i>Spyridium complicatum</i> F.Muell.	2	1	1	1	2	1	1	1	2	1	1	1	3	1	1	2	1	1
<i>Cryptandra</i> cf. <i>spyridioides</i> F. Muell.	2	1	1	1	2	1	1	1	2	1	1	1	3	1	1	2	1	1
<i>Trymalium ledifolium</i> Fenzl	2	1	1	1	2	1	1	1	2	1	1	1	3	1	1	2	1	1
<i>Trymalium floribundum</i> Steudel	2	1	1	1	2	1	1	1	2	1	1	1	3	1	1	2	1	1
<i>Pomaderris rugosa</i> Cheeseman	2	1	1	1	2	1	1	1	2	1	1	1	2	1	1	2	1	1
<i>Siegfriedia darwinoides</i> C.A. Gardner	2	1	1	1	2	1	1	1	2	1	1	1	3	1	2	2	1	1
<i>Barbeya oleoides</i> Schweinf.	2	2	?	2	2	3	2	?	2	1	1	1	?	1	2	2	2	2
<i>Dirachma socotrana</i> Schweinf.	2	?	?	2	2	?	1	?	2	?	1	1	?	1	1	2	2	2

Table 3.3: CI and RI values for each of the individual morphological characters on the trees from morphological (M) and the combined morphological and molecular analyses (C).

Character	CI (C)	CI (M)	RI (C)	RI (M)
1. seed attachment	1	1	1	1
2. disc presence/absence	0.50	0.50	0.93	0.98
3. disc position	0.60	0.60	0.88	0.88
4. leaf margin	0.50	0.50	0.93	0.98
5. nitrogen fixation	0.50	1	0.80	1
6. number of locules per ovary	0.25	0.40	0.70	0.86
7. endosperm	0.17	0.38	0.37	0.75
8. fruit	0.33	0.50	0.88	0.90
9. leaf venation	0.25	0.20	0.66	0.50
10. stellate hairs	1	1	1	1
11. fruit with longitudinal wings	1	1	1	1
12. fruit with apical wings	1	1	1	1
13. ovary position	0.17	0.30	0.60	0.80
14. habit	0.33	0.50	0.70	0.86
15. leaf position	0.17	0.18	0.40	0.47
16. tendrils	0.50	1	0.67	1
17. calyx keel	1	1	1	1
18. arrangement of floral parts	1	1	1	1
Average values	0.55	0.62	0.80	0.88

### 3.4. Discussion

The morphological analysis does not show the three major groups evident in the molecular trees indicating that these morphological characters, some of which were previously used in sub-familial classification systems (Suessenguth, 1953) cannot identify deep clades within Rhamnaceae. Convergent morphological evolution of these characters obscures these relationships, which are determined using molecular data. Suessenguth's reliance on fruit characters to delimit tribes was understandable given the lack of other characters for use at this hierarchical level.

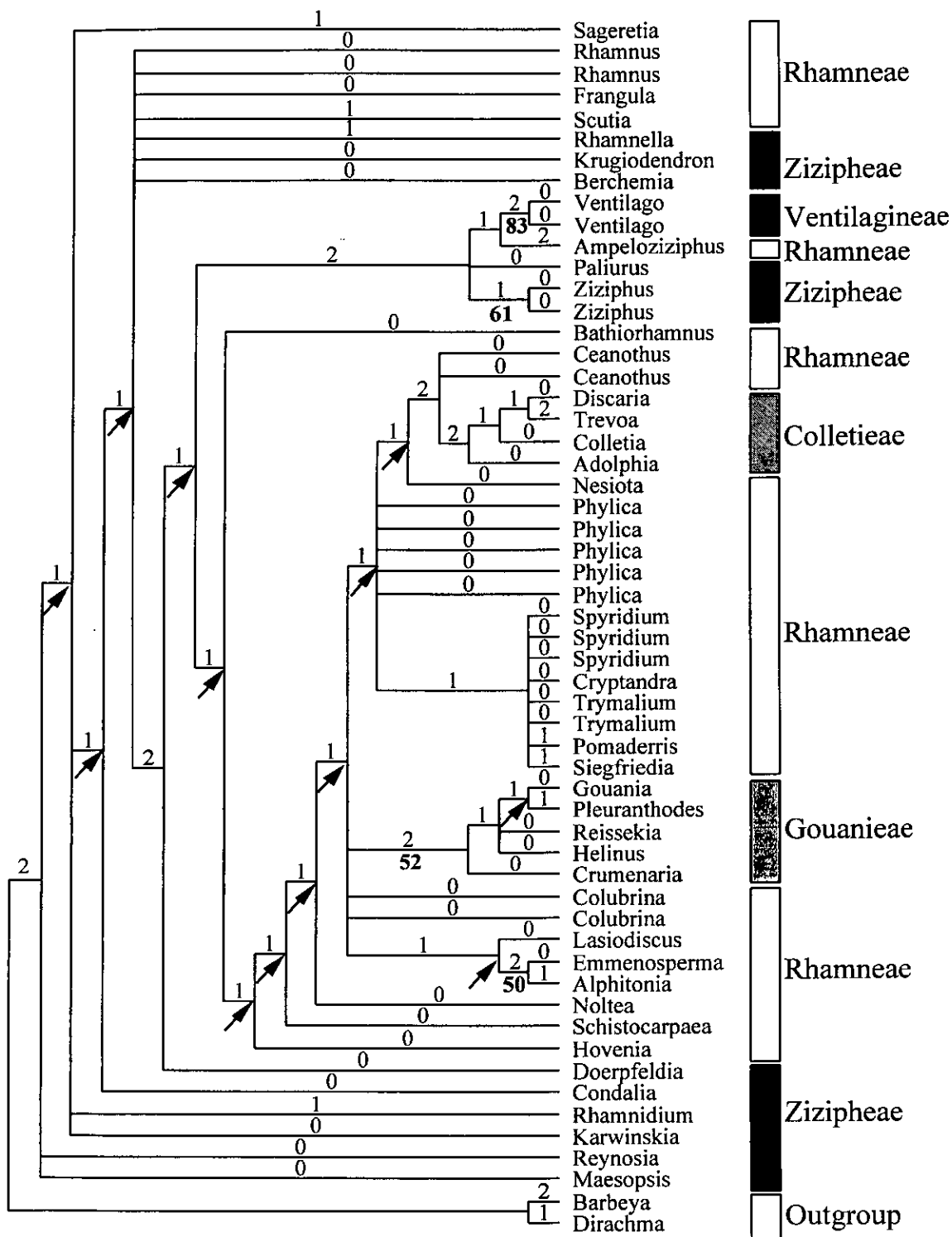


Figure 3.1. One of the 5000 trees from a morphological analysis of Rhamnaceae, using 18 characters. Branch lengths are above branches and bootstrap values are below. Branches that collapse in the strict consensus tree are indicated by an arrow. The length of the trees is 50 steps, CI=0.44 and RI=0.83. The tribal placement of each genus according to Suessenguth (1953) is indicated.

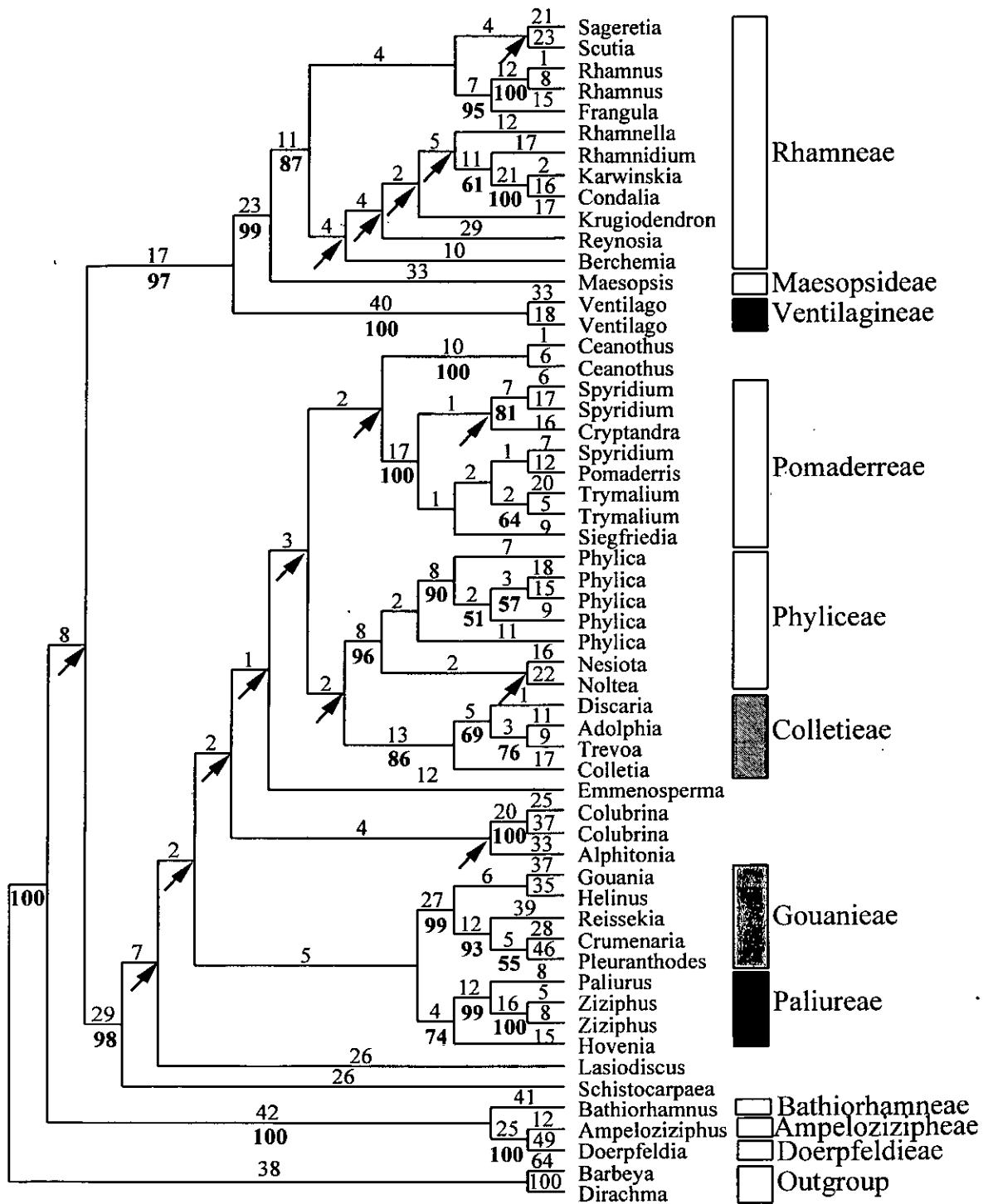


Figure 3.2. One of the 942 trees from a molecular analysis of Rhamnaceae. Branch lengths are above branches and bootstrap values are below. Branches that collapse in the strict consensus tree are indicated by an arrow. The length of the trees is 1660 steps, CI=0.65 and RI=0.76. The tribal placement of each genus according to Richardson *et al.* (submitted) is indicated.

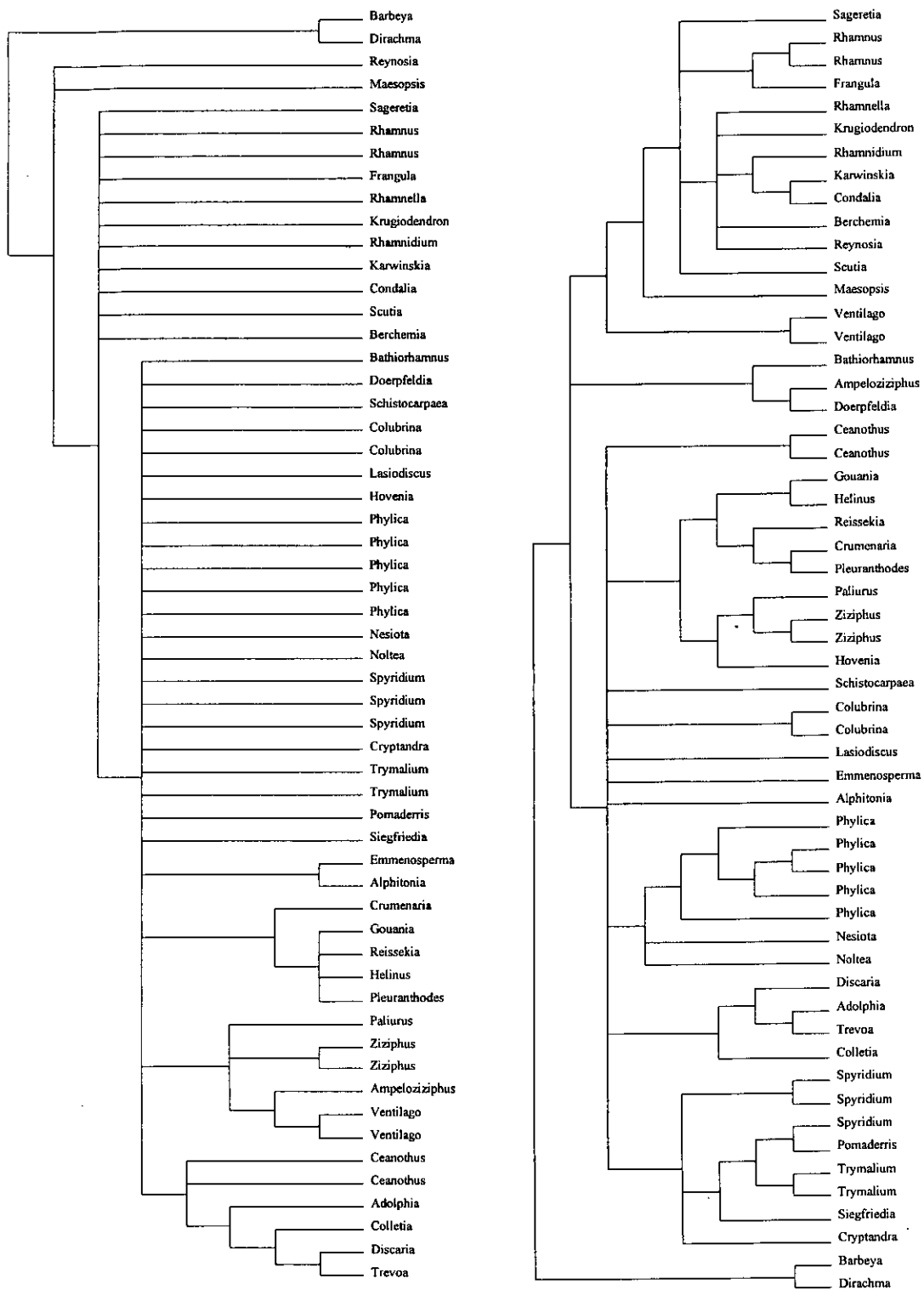


Figure 3.3. Strict consensus trees: left = morphological analysis of Rhamnaceae and right = combined *rbcL/trnL-F* molecular analysis (Chapter Two).

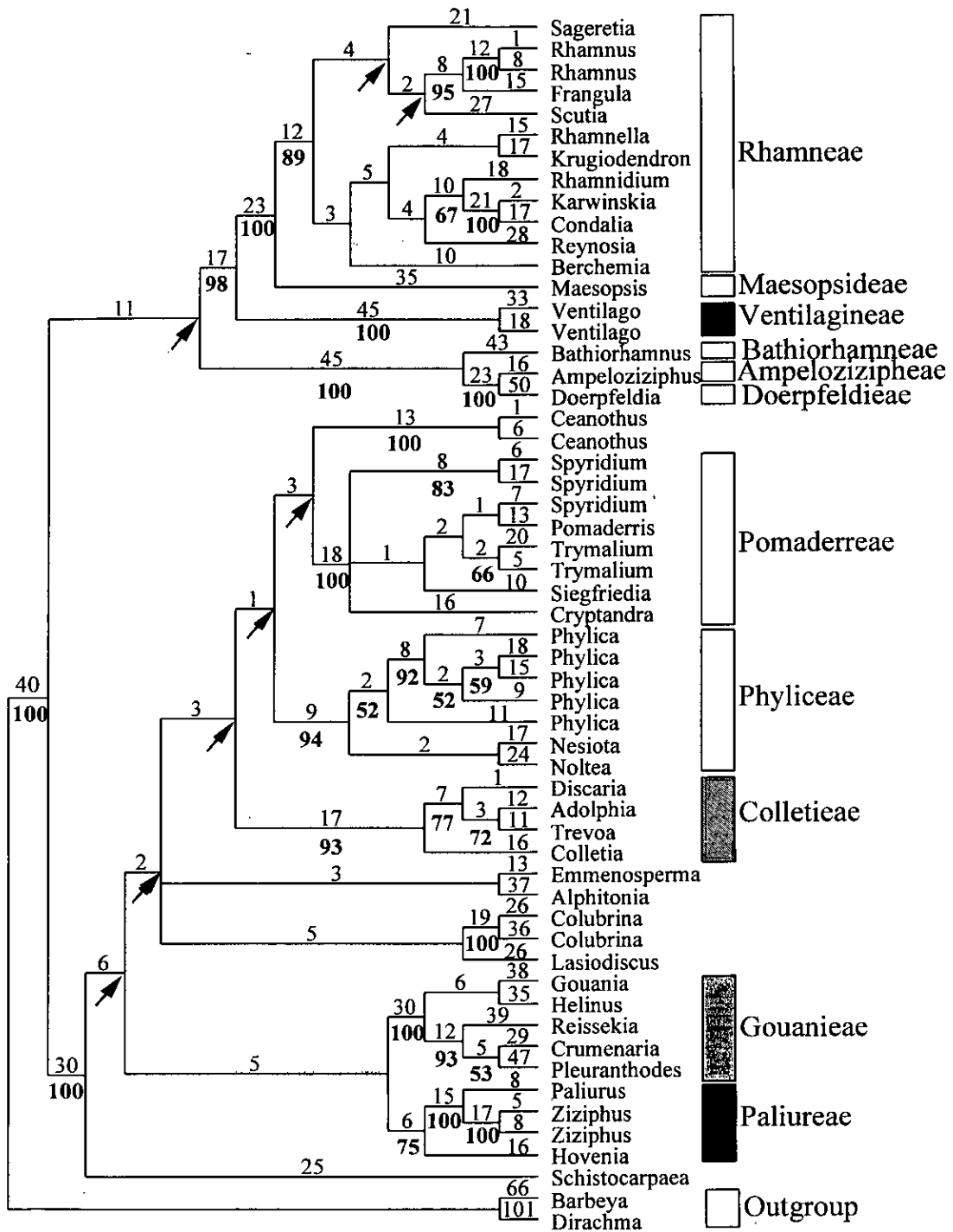


Figure 3.4. One of the 216 trees from a combined morphological and molecular analysis of Rhamnaceae using 18 morphological characters and *rbcL* and *trnL-F* characters. Branch lengths are above branches and bootstrap values are below. Branches that collapse in the strict consensus tree are indicated by an arrow. The length of the trees is 1726, CI=0.64 and RI=0.76. The tribal placement of each genus according to Richardson *et al.* (submitted) is indicated.

Figure 3.5.1. Nitrogen fixation.

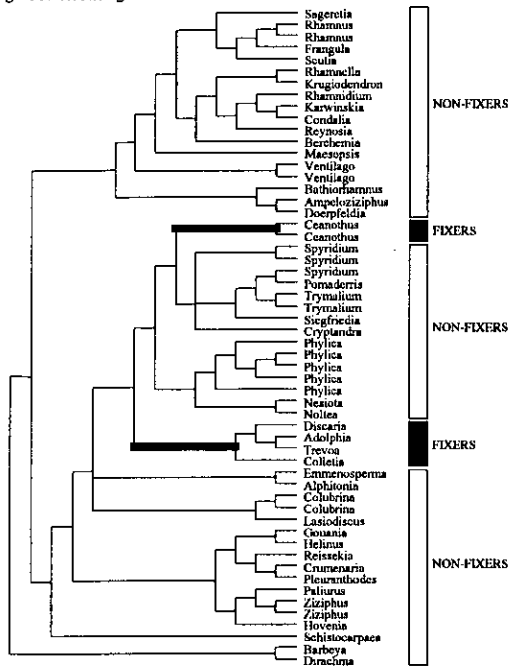


Figure 3.5.3. Ovary position.

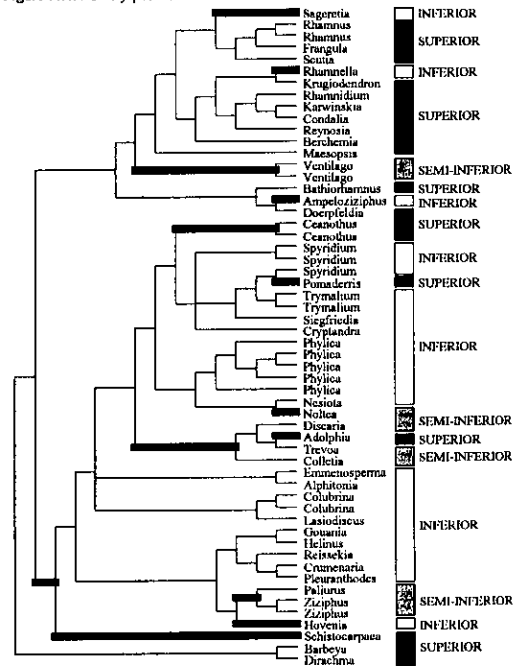


Figure 3.5.2. Fruit appendages.

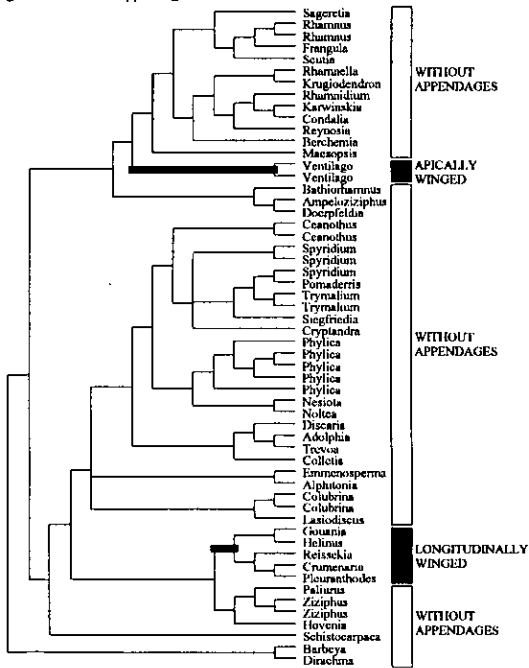


Figure 3.5.4. Fruit.

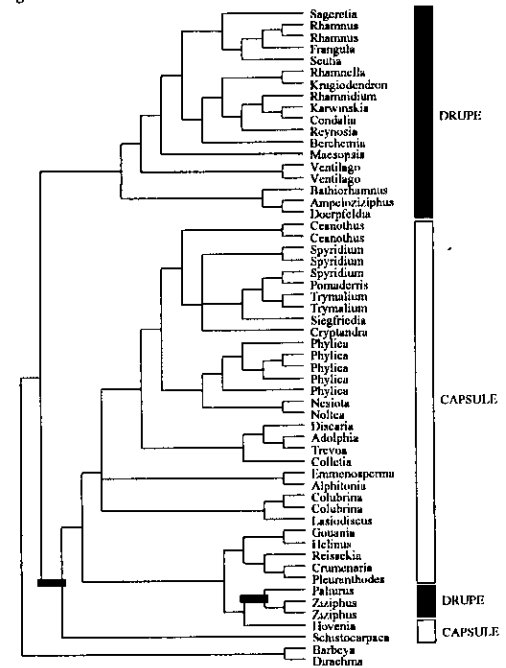


Figure 3.5. Morphological character states mapped onto a combined morphological and molecular tree. Thick bars represent character state changes.

Figure 3.5.5. Habit.

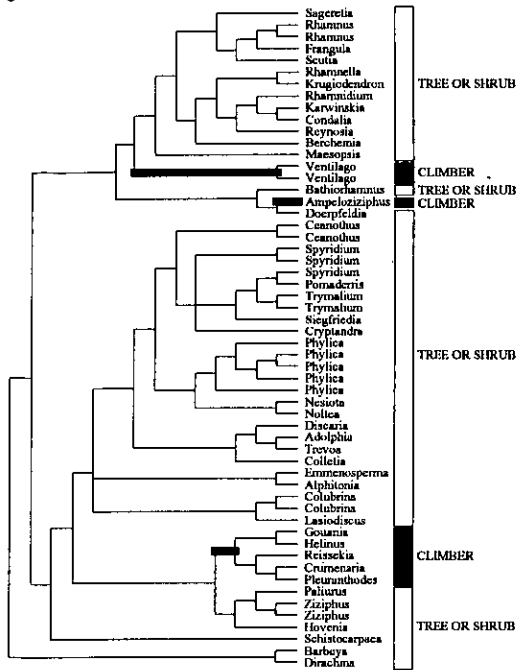


Figure 3.5.6. Plant pubescence.

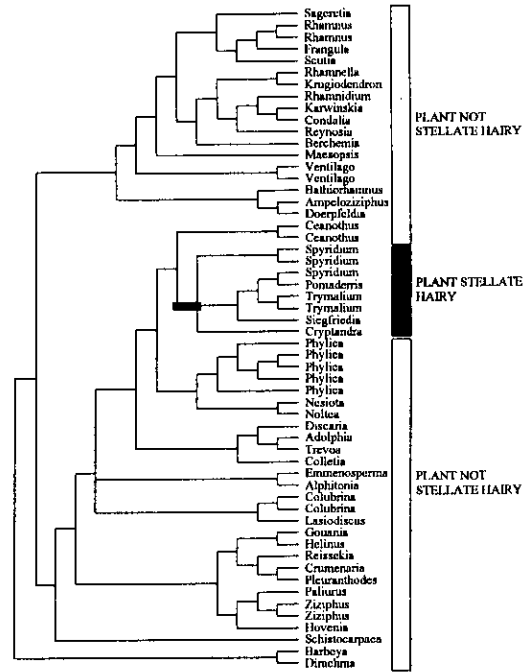


Figure 3.5. Morphological character states mapped onto a combined morphological and molecular tree. Thick bars represent character state changes.

The molecular data indicate the need for reassessment of certain morphological characteristics. More in-depth morphological studies may indicate differences in structure confirming multiple development of certain features. These results help to illustrate the difficulties involved in estimating phylogeny using only a few morphological characters. The CI for the morphological analysis was 0.44 and that of the combined analysis was 0.64. This indicates that the overall levels of homoplasy in the morphological analysis are higher than in the combined analysis and consequently indicates that molecular data are a superior source of information for estimating phylogeny in this group. The CI is negatively correlated with number of terminal taxa and number of characters and is also inflated as the number of uninformative characters in the data set increases (Siebert, 1993) and a better measure of support for molecular data in comparison to morphological data is the RI. Many characters which are of potential use in the estimation of phylogeny are liable to be homoplasious, but homoplasy is also a source of evidence. If a trait evolves twice but in widely separated taxa, then its RI is high even though its CI is low. The



morphological analysis performs better in terms of RI with a value of 0.83 compared to 0.76 for the combined analysis. The lower RI in the molecular analysis could be due to a greater frequency of state changes on branches leading to terminal taxa compared with the morphological trees.

The level of resolution of the strict consensus tree in the morphological analysis is low in comparison to the molecular analysis from Chapter Two (see Figure 3.2). The greater resolving power of molecular data is due to the larger number of molecular characters.

The length of the morphological tree is 50 steps but the number of steps these characters take on the combined morphological and molecular tree is 66. This shows that the addition of molecular data detects more changes than the morphological data alone indicates. As mentioned above, there are various problems in determining the choice and the coding of morphological characters. A more detailed study using more morphological characters and better coding than here may result in better resolution, bootstrap support, CI and RI values in morphological analyses.

The fact that the topology of the combined morphological/molecular analysis is more or less identical to that of the separate molecular analysis is expected, as more molecular characters (2864, 480 of which were informative) were used than morphological ones (18). Differential weighting of morphological and molecular characters could be tried with greater weight being assigned to morphological characters however, this is a highly subjective procedure. Certain morphological characters may be useful in providing added support for some weakly supported or unsupported groups indicated by the molecular trees. For example morphology indicates a closer relationship between *Alphitonia/Emmenosperma* and *Colubrina/Lasiodiscus* although these relationships still have no bootstrap support. Support for clades which had bootstrap support in the molecular analysis alone was slightly increased (with one exception) in the combined morphological/molecular analysis, indicating that addition of morphological characters results in more robust trees.

Because of the better performance in terms of CI, RI and bootstrap values of the combined morphological/molecular analysis morphological characters were mapped

onto one of these trees. Keeled calyces and petals and stamens alternating with sepals are synapomorphies for Rhamnaceae which have arisen once and therefore have an  $RI=1.0$ . A number of other characters are synapomorphies for suprageneric groupings in Rhamnaceae. Attachment of the seed to the torus after dehiscence is a synapomorphy for a weakly supported group containing the genera *Alphitonia* and *Emmenosperma*. The presence of stellate hairs is a synapomorphy for Pomaderreae, presence of apically winged fruits is a synapomorphy for Ventilagineae, and presence of longitudinally winged fruits is a synapomorphy for Gouanieae. Individual CIs and RIs of morphological characters (Table 3.3) do not compare unfavourably with some molecular characters in the molecular analyses from Chapter Two. The problem with the morphological analysis is not that these characters are worse than molecular characters but that there are not enough of them to adequately resolve relationships in this group.

Previous molecular analyses (Soltis *et al.*, 1995; Swensen *et al.*, 1996; Soltis *et al.*, 1998; Savolainen *et al.*, 1996) have indicated that families containing members with the ability to form nitrogen fixing symbioses can be found within the rosoid I clade as described by Chase *et al.* (1993). This was contrary to previous systems which considered nitrogen-fixing species as taxonomically diverse. The fact that the majority of taxa in the rosoid I clade are not nitrogen fixers means that there are two possible scenarios regarding the development of this feature. There could have been a single common origin of this feature that was subsequently lost by members of this clade. Alternatively the ancestor of the nitrogen-fixing clade may have evolved the genetic components that would allow the evolution of nitrogen fixation, and parallel evolution of nitrogen fixation could have occurred during diversification of this clade. This study has allowed a closer investigation of the origins of nitrogen fixation within Rhamnaceae. Figure 3.5.1 shows the distribution of nitrogen fixation within the tree indicating that the ability to fix nitrogen appears to either have developed twice in parallel within the ziziphoid group or to have been present in the ancestor of this group and subsequently lost. However, relationships between clades within this group are not supported by bootstrap, and *Ceanothus* and Colletieae may actually be closest relatives, in which case this phenomenon may have developed only once in

Rhamnaceae. Also, the ability to fix nitrogen has not been extensively investigated in other groups in the ziziphoid clade, and it may be that some of these groups also have nitrogen-fixing capabilities. The molecular tree could therefore be predictive in that it might direct the search for other taxa that fix nitrogen.

Figure 3.5.2 shows the distribution of fruit appendage types within the tree indicating that apically and longitudinally winged fruits have each arisen once. Figure 3.5.3 shows the distribution of ovary position character states. This illustrates that the evolution of highly adaptive or developmentally plastic characters such as ovary position is often likely to be homoplasious. The development of these characters needs to be well studied before any definite conclusions about homology can be made. The molecular results could lead to more in-depth studies of such characters in Rhamnaceae. The only potential morphological evidence for the “cryptic clades” described in the previous chapter comes from possible studies of gynoeceium ontogenesis. Restriction site variation of plastid DNA and nuclear rDNA has been used to assess phylogenetic relationships among the nine species of the taxonomically complex genus *Lithophragma* (Saxifragaceae; Soltis *et al.* 1992), and these agree in part with those based on morphological data. *Lithophragma* infrageneric classification was partly based on ovary position, and groups defined on the basis of ovary position were not found to be monophyletic according to molecular analyses. Comparison of the DNA-based analyses with evidence from morphology indicated that fusion of the hypanthium to the ovary wall has occurred independently several times in the genus or that hypanthium fusion occurred early in the radiation of the genus and was subsequently lost. The molecular phylogenetic study of *Lithophragma* indicated that the presence of either an inferior or a superior ovary might not always represent a homologous character state. A study of gynoeceium ontogenesis revealed that patterns in the initial development of the ovary were consistent with the molecular tree. Monophyletic groups within the genus could be defined on whether they have a floral apex that is initially more or less flat or whether they have a floral apex that initially has a circular depression. Subsequent ontogenetic development leads to the production of either superior or inferior ovaries regardless of the initial developmental state. These character states are therefore not

homologous. Without ontogenetic investigation, this would seem to represent a case of parallel evolution but could in fact be regarded as a case of parallel development of similar character states. A similar phenomenon could be occurring in Rhamnaceae. Medan (1988) has studied the shape of the floral apex and the degree of intercalary growth at carpellary bases in 17 genera of Rhamnaceae. In some taxa the floral apex is more or less flat at the time of primordia differentiation, (*Condalia*, *Rhamnus*, in the rhamnoid clade of the molecular analysis, Chapter Two). These taxa go on to form superior ovaries. In other taxa the floral apex shows a circular depression at the time of primordia differentiation, (*Colletia*, *Noltea*, *Phyllica* and *Pomaderris*, in the ziziphoid clade of the molecular analysis, Chapter Two). These taxa go on to form inferior or semi-inferior ovaries. Studies of more genera in Rhamnaceae could show that there is a situation similar to that in *Lithophragma* in which the latter stages of development of the ovary may obscure the initial patterns leading to character states, which represent false homologies. For example Colletieae in the ziziphoid clade have inferior, semi-inferior or superior ovaries. It would be interesting to determine whether the taxa with a superior ovary developed from a floral apex with a circular depression. The limited sampling in this study could potentially be expanded and provide morphological character support for the cryptic clades defined by the molecular data. The rhamnoid clade could possibly be defined by having a flat floral apex, and other clades could be defined by having an indented floral apex. A study of floral development in Rhamnaceae is feasible, but it is beyond the scope of this project.

Figure 3.5.4 indicates that drupes are the ancestral fruit form within Rhamnaceae with a single development of capsules and a single reversal back to drupes in *Ziziphus* and *Paliurus*. Figure 3.5.5 shows the distribution of habit types indicating that the climbing habit has developed three times from an arborescent ancestral state. The presence of stellate hairs seems to be a derived character that has developed once in Pomaderreae (Figure 3.5.6).

### 3.5. Conclusions

Analysis of the molecular characters used here results in more highly resolved trees than analysis of the morphological characters used because of the larger number of characters. Individual morphological characters do not perform badly in comparison to individual molecular characters and have higher RIs. There are not enough morphological characters to be successful on their own although the addition of morphological data to the molecular analysis does improve bootstrap values slightly for all clades (with one exception) that are supported in the molecular trees. The use of both molecular and morphological data will lead to a better understanding of the developmental biology of the group.

### 3.6. Rhamnaceae Tribal Classification

The following taxonomic account of a revision of the tribal classification of Rhamnaceae is based on the molecular analysis presented in Chapter Two. Seven of the proposed tribes are strongly supported by bootstrap values of 92 or more in the separate and combined molecular analyses. Tribes that are well supported in the molecular analysis with the additional support of morphological characteristics are defined. Those genera that according to molecular and morphological data have no well supported affinities are left as *incertae cedis*.

Some chromosome numbers were taken from Raven (1975), Darlington and Wylie (1982), Kumar and Subramaniam (1986) and Jarolimova (1994).

1. Tribe *Paliureae* Reiss. ex Endl., Benth. and Hook. f. Gen. Pl., 1095 (1840) [=Zizipheae Brongn.]. Type: *Paliurus* Mill. Some characteristics taken from Chun and Tsiang (1939); Johnston (1963, 1964); Schirarend and Olabi (1994).

Trees or shrubs. Branches spinose or unarmed. Leaves alternate or fasciculate, venation palmate. Stipules persistent or caducous. Inflorescences axillary or terminal cymes, inflorescence-axis sometimes becoming succulent (*Hovenia*). Calyx tube widely spreading, scarcely concave; limbs spreading, more or less triangular, midrib

keeled on the inside. Petals usually present. Filaments cylindrical; anthers introrse, 2-locular. Ovary semi-inferior or superior (*Hovenia*), 2- (3- or 4-) locular. Nectariferous disc adnate to ovary and calyx-tube and filling calyx tube, sometimes hairy (*Hovenia*). Style bi- or trifid. Fruit dry with a wide membranous ring around the top (*Paliurus*), a drupe (*Ziziphus*) or a capsule (*Hovenia*). Seed with or without endosperm, coat membranaceous or papery. Chromosome numbers  $2n=12, 24, 26, 36, 40, 48, 72$ . New and Old World tropics and warm temperate regions, southern Europe to Japan.

Three genera: *Paliurus* Mill., *Ziziphus* Mill. (= *Sarcomphalus* R. Br.) and *Hovenia* Thunb.

2. Tribe *Colletieae* Reiss. ex Endl., Benth. and Hook. f. Gen. Pl., 1099 (1840). Type: *Colletia* Comm. ex Juss. Some characteristics taken from Johnston (1973); Tortosa (1983, 1989, 1992, 1993).

Strongly armed trees or shrubs, branches decussate. Spines frequently green. Roots of most genera bearing nitrogen-fixing nodules. Leaves opposite, small, often caducous, venation palmate or pinnate. Stipules absent or present and persistent or falling early. Inflorescences axillary, with flowers solitary or in cymes. Petals present or absent. Filaments filiform or cylindric, erect or subulate; anthers 1- or 2-locular. Ovary 3-(2-) locular, inferior, semi-inferior or superior. Nectariferous disc annular, 5-lobed, adnate to calyx tube or absent. Style 2- or 3-lobed or trifid. Fruit a capsule or a drupe. Seed coat leathery, endosperm present. Chromosome number  $2n=22$  (*Colletia*, *Discaria*). Predominantly South American but also found in North America, New Zealand and Australia.

Six genera: *Adolphia* Meisn., *Colletia* Comm. ex Juss., *Discaria* Hook., *Kentrothamnus* Suess. and Overkott, *Retanilla* (DC.) Brongn. and *Trevoa* Miers ex Hook. (= *Talguenea* Miers ex Endl.).

3. Tribe *Phyliceae* Reiss. ex Endl. emend. J.E. Richardson, M.W. Chase and M.F. Fay, Gen. Pl., 1100 (1840). Type: *Phylica* L. Some characteristics taken from Pillans (1942).

Unarmed ericoid shrubs or trees. Branches often clustered, parallel and erect. Leaves alternate or opposite, usually densely tomentose beneath, leaf margins usually revolute (sometimes toothed and not revolute, *Noltea*), venation pinnate. Stipules absent in all but one species of *Phylica* or present and caducous (*Nesiota*) or present and persistent (*Noltea*). Inflorescences capitate to spicate, paniculate or flowers solitary, terminal or axillary. Bracts leafy or short and scarious. Flowers 5-merous (sometimes 4-merous in *Nesiota*). Calyx persistent, usually topping fruit or deciduous. Filaments subulate, usually short, often curved; anthers 1- or 2-locular. Ovary usually inferior (sometimes semi-inferior), completely or mostly fused to the receptacle, 3-(4-)locular. Nectariferous disc epigynous or slender and covering the inside of the calyx tube, sometimes hairy (*Nesiota*). Style obscurely 3-lobed or trifid. Fruit a capsule, 3-locular; locules 1-seeded, dehiscent. Seeds arillate (at least in *Phylica*), endosperm present. Chromosome number not known. South Africa, St Helena, Tristan da Cunha, Malawi, Tanzania, Mozambique, Zimbabwe, Madagascar, Mauritius, Réunion and New Amsterdam.

Three genera: *Nesiota* Hook. f., *Noltea* Rchb. and *Phylica* L.

4. Tribe *Gouanieae* Reiss. ex Endl., Benth. and Hook. f. Gen. Pl., 1102 (1840). Type: *Gouania* Jacq. Some characteristics taken from Grey-Wilson (1978).

Unarmed climbers or herbs (*Crumenaria*), tendrils present. Leaves alternate, petiolate, entire, base subcordate, apex mucronate, venation pinnate or palmate. Stipules usually caducous. Inflorescences small cymes. Filaments subulate, apex incurved; anthers introrse, 2-locular, longitudinally dehiscent. Ovary inferior, 3-(2- or 4-)locular with one ovule per locule. Nectariferous disc epigynous, fleshy, stellate or margins 5-angled. Style trifid. Fruit a capsule, 3-locular, loculicidally dehiscent, usually with longitudinal wings which lie above the septum of the locules; locules 1-seeded. Seed coat leathery; endosperm present, fleshy. Chromosome number  $2n=22$

(*Helinus*). Tropical and warm America, Africa, Madagascar, Indian Ocean islands and Asia.

Six genera: *Alvimiantha* Grey-Wilson, *Crumenaria* Mart., *Gouania* Jacq., *Helinus* E.Mey. ex Endl., *Pleuranthodes* Weberb. and *Reissekia* Endl.

5. Tribe *Pomadereae* Reiss. ex Endl., Benth. and Hook. f. Gen. Pl., 1101 (1840). Type: *Pomaderris* Labill.. Some characteristics taken from Gardner (1932, 1941); Keighery (1978).

Shrubs or small trees with stellate hairs. Leaves opposite or alternate, venation pinnate. Stipules caducous or persistent. Inflorescence with flowers solitary in axils, cymose or clustered into glomerules. Filaments inflexed. Ovary usually inferior or semi-inferior (rarely superior, *Blackallia*), 3-(or 4-)locular. Nectariferous disc surrounding base of ovary and adnate to calyx tube. Style 3-lobed or trifid. Fruit a capsule, exocarp thin; locules 1-seeded, dehiscent. Seed with a tiny aril, endosperm present. Chromosome numbers  $2n=24, 36, 48$  (*Pomaderris*). Australia and New Zealand.

Six genera: *Blackallia* C.A. Gardner, *Cryptandra* Sm., *Pomaderris* L., *Siegfriedia* C.A. Gardner, *Spyridium* Fenzl. and *Trymalium* Fenzl.

6. Tribe *Rhamnaceae* Hook.f. emend. J.E. Richardson, M.W. Chase and M.F. Fay, Benth. and Hook. f. Gen. Pl. 1: 373 (1862). Type: *Rhamnus* L. Some characteristics taken from Grisebach (1866); Mueller (1875); Urban (1902-03, 1924); Nakai (1923); Wolf (1938); Johnston (1962, 1974).

Trees, shrubs or climbers, sometimes armed. Leaves opposite, sub-opposite or alternate, entire or serrate, venation pinnate. Stipules sometimes absent, often caducous. Inflorescence solitary, fasciculate, umbellate or racemose to cymose, axillary or terminal. Petals present or absent. Ovary superior (rarely inferior), free, usually 2-(1- or 4-) locular. Nectariferous disc lining base of calyx tube or free. Styles 2, often persistent on apex of fruit. Fruit a drupe, 1-4-celled. Seeds without endosperm or endosperm thin or fleshy. In mature seeds hilum next to radicle.



Chromosome number  $2n=12, 20, 24, 26$ . Found throughout the range of the family except southern South America.

Thirteen genera: *Auerodendron* Urb., *Berchemia* Neck. ex DC (= *Phyllogeiton* (Weberb.) Herzog, *Berchemiella* Nakai, *Condalia* Cav., (= *Condaliopsis* (Weberb.) Suess., *Microrhamnus* A. Gray), *Dallachya* F. Muell., *Karwinskia* Zucc., *Krugiodendron* Urb., *Reynosia* Griseb., *Rhamnella* Miq. (= *Chaydaia* Pit.), *Rhamnidium* Reiss., *Rhamnus* L. (= *Oreoherzogia* Vent, *Oreorhamnus* Ridl.), *Sageretia* Brongn. (= *Lamellisepalum* Engl.), *Scutia* (DC) Brongn.

7. Tribe *Maesopsideae* Weberb., Engler and Prantl, Nat. Pflanzenfam. 128: 399 (1895). Type: *Maesopsis* Engl. Some characteristics taken from Schirarend and Süß (1985).

Unarmed trees. Leaves opposite or alternate, strongly toothed with glands at tips of teeth, venation pinnate. Stipules present, small. Inflorescence an axillary pseudoraceme. Petals present. Ovary superior, free, 1-celled, without a prominent placenta. Nectariferous disc lining the inside of the calyx-tube. Style laterally attached to the fruit, tetrafid. Fruit a drupe, 1-seeded. In mature seeds radicle opposite to hilum, endosperm copious, taking up most of volume of seed. Chromosome number  $2n=18$ . Tropical Africa.

One genus: *Maesopsis* Engl.

8. Tribe *Ventilagineae* Hook. f., Benth. and Hook.f. Gen. Pl. 1: 371 (1862). Type: *Ventilago* Gaertn. Some characteristics taken from Banerjee and Mukerjee (1970).

Climbers or rarely small trees, unarmed, tendrils absent. Branches rigid, glabrous. Leaves alternate, stalked, secondary nerves ascending and converging along the margin, venation pinnate. Stipules caducous. Flowers in umbellate cymes or fascicled, arranged in panicles, lateral or terminal. Calyx spreading. Filaments cylindrical; anthers introrse, 2-locular, longitudinally dehiscent, connective long, apiculate. Ovary semi-inferior to inferior, more or less sunk into nectariferous disc, 2-locular; ovules 1 per locule. Nectariferous disc fleshy, tuberculate. Style with 2

short stigmatic lobes. Fruit samaroid and indehiscent. Seed without endosperm. Chromosome number  $2n=24$ . Old World tropics.

Two genera: *Ventilago* Gaertn., *Smythea* Seem. ex A.Gray.

9. Tribe *Ampelozizipheae* J.E.Richardson, M.W.Chase and M.F.Fay *tribus nov.* Type: *Ampeloziziphus* Ducke. Some characteristics taken from Ducke (1935).

Unarmed climbers, tendrils absent. Leaves distichous, alternate, large, venation palmate, 5-nerved, the two outer veins slender, sometimes almost obsolete. Stipules small, setaceous, caducous. Inflorescences axillary cymes, on previous year's growth, often elongate with upper part leafless, forming interrupted racemes to 30 cm long, often with several cymes forming a large panicle. Calyx tube shortly turbinate; lobes subequal. Ovary semi-inferior, included in and united to calyx tube and nectariferous disc, 3-locular. Ovules solitary. Nectariferous disc thick, filling calyx tube and closely adnate to it and the ovary, flat on surface, annular. Style trifid at apex. Fruit a drupe, 3-locular with one seed per locule, base stipitate, stalk surrounded by persistent lobes of calyx; exocarp thick and fleshy; stone hard but thin walled. Seeds sometimes not well developed, coat thick, leathery, smooth, shiny; endosperm and aril absent. Chromosome number unknown. Northern South America.

One genus: *Ampeloziziphus* Ducke.

10. Tribe *Doerpfeldieae* J.E. Richardson, M.W. Chase and M.F. Fay *tribus nov.* Type: *Doerpfeldia* Urb. Some characteristics taken from Urban (1924).

Trees, unarmed. Leaves alternate, often emarginate, otherwise entire, venation palmate, 3-nerved. Stipules at base of petioles, caducous. Flowers axillary, solitary. Flower bud globose. Petals absent. Ovary superior, pseudo-2-locular. Nectariferous disc thinly covering the ovary and not attached to the calyx-tube. Style bifid. Receptacle short. Stamens deeply inserted around the base of the ovary. Fruit a drupe, more or less unequally 2-locular, smaller locule empty; exocarp thin; calyx-tube remaining attached to lower quarter of fruit; endocarp bony. Seed with endosperm. Chromosome number unknown. Cuba.

One genus: *Doerpfeldia* Urb.

11. Tribe *Bathiorhamneae* J.E. Richardson, M.W. Chase and M.F. Fay *tribus nov.*  
Type: *Bathiorhamnus* (H. Perr.) Cap. Some characteristics taken from Capuron (1966).

Unarmed trees. Leaves alternate, 3-nerved with nerves converging at apex, margins entire to toothed, venation palmate. Stipules small. Inflorescences fasciculate, axillary. Sepals punctate-pellucid. Petals small, clawed, lamina cucullate. Ovary superior, 3-locular. Nectariferous disc thick, broadly attached to ovary. Style trifid. Fruit a drupe, base encircled with an annular scar, (1- or 2-) 3-locular, septicidally dehiscent; locules indehiscent. Seed with endosperm, without aril, coat leathery. Chromosome number unknown. Madagascar.

One genus: *Bathiorhamnus* Capuron (= *Macrorhamnus* H. Perr.).

Genera *incertae sedis*:

The following taxa are treated *incertae sedis* because their placement in the molecular tree is ambiguous and because any morphological affinities they show are not strong enough to support their inclusion in any other group. Further sequencing of taxa around these genera should give a clearer idea of their relationships to other groups.

*Ceanothus* L. Some characteristics taken from Van Rensselaer and McMinn (1942).

Shrubs or small trees, sometimes spinescent. Roots of most species bearing nitrogen-fixing nodules. Leaves alternate or opposite, venation palmate or pinnate, deciduous or evergreen. Stipules caducous or persistent. Flowers in terminal composite panicles or axillary racemes. Petals present. Filaments thread-like; anthers introrse, 2-locular. Ovary 3-(4-) locular, superior, more or less immersed in nectariferous disc which is adnate to ovary and calyx tube, annular, subpentagonal, glandular. Style trifid. Fruit a capsule, 3-locular, base of calyx tube circumsissile around base of capsule, 3-ribbed, separating at maturity into three parts, exocarp leathery to weakly fleshy; locules dehiscent, crustaceous, bivalved, 1-seeded.

Receptacle and disc persistent on the pedicel, remaining intact during endocarp dehiscence. Seeds smooth, convex at one side, sometimes arillate, endosperm present. Chromosome number  $2n=24$ . North America. A genus of 55 species.

*Emmenosperma* F. Muell. Some characteristics taken from Mueller (1862-63).

Unarmed trees. Leaves sub-opposite, entire, leathery, venation pinnate. Stipules absent. Inflorescences repeatedly trichotomous panicles. Calyx 5-lobed; lobes deciduous. Petals 5. Anthers longitudinally dehiscent. Ovary superior, 2-(3-)locular. Nectariferous disc thin, lining the base of the receptacle. Style filiform, bifid. Stigma bi-(tri-)fid. Fruit a capsule, 2-(3-)locular, septicidally dehiscent; locules dehiscent; exocarp thin and leathery; endocarp osseous-crustaceous, splitting unequally. Seeds persisting on receptacle after dehiscence, erect; aril and endosperm present. Chromosome number unknown. Australia. A genus of three species.

*Schistocarpha* F. Muell. Some characteristics taken from Mueller (1891).

Tree, unarmed. Leaves alternate, venation pinnate. Stipules deciduous. Inflorescences terminal and axillary panicles. Bracts small. Calyx deeply 5-lobed; lobes semi-lanceolate, deciduous. Petals 5. Anthers longitudinally dehiscent. Ovary 3-locular, superior, almost fully emerged. Nectariferous disc slightly undulate at margin. Style trifid. Fruit a capsule, calyx-tube persistent and surrounding base; exocarp crustaceous, irregularly trivalved; endocarp receding, thinly papery; locules splitting to base along inner side, ruptured and twisted on outer side. Seeds without albumen, testa chartaceous. Chromosome number unknown. Australia. A monotypic genus.

*Alphitonia* Reiss. ex Endl. Some characteristics taken from Braid (1925).

Trees, sometimes large, unarmed. Branches rust-red, tomentose. Leaves alternate, petiolate, venation pinnate, entire, indumentum weakly to strongly developed, darkening above when dried. Stipules subulate, villose, deciduous. Inflorescences subterminal, paniculate racemes. Ovary semi-inferior, 2- or 3-locular. Nectariferous disc adnate to ovary and calyx tube and filling calyx tube. Style 2- or 3-lobed. Fruit a

drupe; margin of receptacle reaching bottom third or middle half of fruit; exocarp thick, spongy; endocarp of 2 or 3 hard, coriaceous locules; locules dehiscent down the ventral suture and partially down the dorsal suture; exocarp, endocarp and portions of the receptacle fall away; seeds persisting on the remainder of the receptacle, arillate, endosperm cartilaginous, coat hard or tough. Chromosome number unknown. Malaysia, Australia, West Pacific islands, New Caledonia. A genus of six species.

*Colubrina* Rich ex Brongn. Some characteristics taken from Johnston (1971).

Shrubs or trees, armed or unarmed, rarely scandent. Leaves alternate or opposite, venation pinnate or palmate, often glandular. Stipules lateral and basal or interpetiolar, usually caducous. Inflorescence of cymes or small thyrses, sessile and umbel-like or shortly stalked, few-flowered and corymb-like or a compound partial dichasium. Flowerbuds more or less glabrous to densely hairy. Ovary inferior to superior, 3-(4-)locular. Nectariferous disc large, nearly filling the receptacle and often hiding the ovary, remaining united from the lower fifth to the upper half of the fruit. Styles trifid. Fruit a capsule; mesocarp thin, dry, leathery to brittle and flaky; endocarp crustaceous or cartilaginous; locules dehiscent. Receptacle and disc breaking irregularly as endocarp dehisces into separate locules. Seeds with endosperm, sometimes with a small aril. Chromosome number  $2n=16, 24$ . Tropical and warm areas in the Americas and Africa. A genus of thirty one species.

*Lasiodiscus* Hook. f. Some characteristics taken from Figueiredo (1995).

Trees or shrubs, unarmed. Leaves opposite, pinnate or palmate, often with minute, glandular teeth. Stipules interpetiolar, usually caducous. Inflorescences usually a partial dichasium. Flower buds sub-glabrous to densely hairy. Ovary inferior or half-inferior, 3-locular. Nectariferous disc fleshy, covering the ovary from the insertion of the petals and stamens to the base of the style. Fruit a capsule; locules dehiscent. Seeds with endosperm. Chromosome number unknown. Tropical Africa and Madagascar. A genus of twelve species.

### 3.7. Bibliography

Banerjee, S.P. & P.K. Mukerjee. 1970. Studies in the Rhamnaceae 3, a taxonomic revision of Indian Ventilagineae. *Indian Forester* 96: 203-217.

Braid, K.W. 1925. Revision of the genus *Alphitonia*. *Kew Bulletin* 171-186.

Capuron, R. 1966. Notes sur quelques Rhamnacées arbustives ou arborescentes de Madagascar. *Adansonia* 6: 116-141.

Chun, W.Y. & Y. Tsiang. 1939. A new species of *Hovenia*. *Sunyatsenia* 4: 16-17.

Darlington, C.P. & A.P. Wylie. 1982. Chromosome atlas of flowering plants. George Allen & Unwin, London.

Ducke, A. 1935. Plantes nouvelles ou peu connues de la région Amazonienne (ix serie). *Archivos do Instituto de Biologia Vegetal* 2: 157.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783-791.

Figueiredo, E. 1995. A revision of *Lasiodiscus* (Rhamnaceae). *Kew Bulletin* 50: 495-526.

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

Gardner, C.A. 1932. Contributiones florae Australiae occidentalis no. 8. *Journal of the Royal Society of Western Australia* 19: 79-93.

Gardner, C.A. 1941. Contributiones florae Australiae occidentalis no. 9. Journal of the Royal Society of Western Australia 27: 165-210.

Grey-Wilson, C. 1978. *Alvimiantha*, a new genus of Rhamnaceae from Bahia, Brazil. Bradea 2: 287-290.

Grisebach, A. 1866. Catalogus Plantarum Cubensium. Engelmann, Leipzig.

Jarolimova, V. 1994. Chromosome counts of some Cuban angiosperms. Folia geobotanica and phytotaxonomica 29: 101-106.

Johnston, M.C. 1962. Revision of *Condalia* including *Microrhamnus* (Rhamnaceae). Brittonia 14: 332-368.

Johnston, M.C. 1963. The species of *Ziziphus* indigenous to the United States and Mexico. American Journal of Botany 50: 1020-1027.

Johnston, M.C. 1964. The fourteen species of *Ziziphus* including *Sarcomphalus* (Rhamnaceae) indigenous to the West Indies. American Journal of Botany 51: 1113-1118.

Johnston, M.C. 1971. Revision of *Colubrina* (Rhamnaceae). Brittonia 23: 2-53.

Johnston, M.C. 1973. Revision of *Kentrothamnus* (Rhamnaceae). Journal of the Arnold Arboretum 54: 471-473.

Johnston, M.C. 1974. Revision of *Scutia* (Rhamnaceae). Bulletin of the Torrey Botanical Club 101: 64-71.

Keighery, G.J. 1978. *Siegfriedia*. Australian Plants 11: 176.

- Kumar, V. & B. Subramaniam. 1986. Chromosome atlas of flowering plants of the Indian subcontinent volume 1, dicotyledons. Botanical Survey of India, Calcutta.
- Medan, D. 1988. Gynoecium ontogenesis in the Rhamnaceae: A comparative study. *In* Leins, P., S.C. Tucker & P.K. Endress [eds.], Aspects of Floral Development, pp. 133-141. Cramer, Berlin, Stuttgart.
- Mueller, F. 1862-63. Celastrineae. *Fragmenta Phytogeographie Australiae*. 3: 62-63. Government Printer, Melbourne.
- Mueller, F. 1875. Rhamnaceae. *Fragmenta Phytogeographie Australiae*. 9: 140-141. Government Printer, Melbourne.
- Mueller, F. 1891. Descriptions of new Australian plants with occasional other annotations. *The Victorian Naturalist* 7: 180-183.
- Nakai, T. 1923. *Genera nova Rhamnacearum et Leguminosarum ex Asia orientali*. *Japanese Botanical Magazine* 37: 29-36.
- Pillans, N.S. 1942. The genus *Phyllica* Linn. *Journal of South African Botany* 8: 1-164.
- Raven, P. 1975. The bases of angiosperm phylogeny: cytology. *Annals of the Missouri Botanical Garden* 62: 724-764.
- Rensselaer, M. van & H.E. McMinn. 1942. *Ceanothus*. Santa Barbara Botanic Garden, Santa Barbara, California.
- Richardson, J.E., M.F. Fay & M.W. Chase. (submitted). A revision of the tribal classification of Rhamnaceae. *Kew Bulletin*.



Savolainen, V., C.M. Morton, S.B. Hoot & M.W. Chase. 1996. An examination of phylogenetic patterns of plastid *atpB* sequences among eudicots. *American Journal of Botany* 83: 190 [abstract].

Schirarend, C. & M.N. Olabi. 1994. Revision of the genus *Paliurus* Tourn. ex Mill. (Rhamnaceae). *Botanische Jahrbücher* 116: 333-359.

Schirarend, C. & H. Süß. 1985. Zur Holzanatomie und systematischen Stellung der Gattung *Maesopsis* Engler (Rhamnaceae). *Gleditschia* 13: 41-45.

Siebert, D.J. 1993. Tree statistics; trees and 'confidence'; consensus trees; alternatives to parsimony; character weighting; character conflict and its resolution. *In* Forey, P.L., C.J. Humphries, I.J. Kitching, R.W. Scotland, D.J. Siebert & D.M. Williams [eds.], *Cladistics, A Practical Course in Systematics*. Clarendon Press, Oxford.

Soltis, D.E., P.S. Soltis, J.N. Thompson & O. Pellmyr. 1992. Chloroplast DNA Variation in *Lithophragma* (Saxifragaceae). *Systematic Botany* 17: 607-619.

Soltis, D.E., P.S. Soltis, D.R. Morgan, S.M. Swensen, B.C. Mullins, J.M. Dowd & P. Martin. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proceedings of the National Academy of Sciences, USA* 92: 2647-2651.

Suessenguth, K. 1953. Rhamnaceae, Vitaceae, Leeaceae. *In* Engler, A. & K. Prantl, [eds.], *Die Natürlichen Pflanzenfamilien* 2, Aufl., 20d. Duncker & Humblot, Berlin.

Swensen, S.M. 1996. The evolution of actinorrhizal symbioses: evidence for multiple origins of the symbiotic association. *American Journal of Botany* 83: 1503-1512.

Swofford, D.L. 1993. PAUP: phylogenetic analysis using parsimony, version 3.1.1. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.

Tortosa, R.D. 1983. El género *Discaria* (Rhamnaceae). Bolétin de la Sociedad Argentina de Botanica 22: 301-336.

Tortosa, R.D. 1989. El genero *Colletia* (Rhamnaceae). Parodiana 5: 279-332.

Tortosa, R.D. 1992. El complejo *Retanilla-Talguenea-Trevoa* (Rhamnaceae). Darwiniana 31: 223-252.

Tortosa, R.D. 1993. Revision del genero *Adolphia* (Rhamnaceae-Colletieae). Darwiniana 32: 185-189.

Urban, I. 1902-03. Symbolae Antillanae seu Fundamenta Florae Indiae Occidentalis. 3: 313-315. Borntraeger, Berlin.

Urban, I. 1924. Symbolae Antillanae seu Fundamenta Florae Indiae Occidentalis 10: 221-223. Borntraeger, Berlin.

Wolf, C.B. 1938. The North American species of *Rhamnus*. Rancho Santa Ana Botanic Garden, Claremont, California.

**CHAPTER FOUR. PHYLOGENETIC  
ANALYSIS OF *PHYLICA* L. WITH AN  
EMPHASIS ON ISLAND SPECIES:  
EVIDENCE FROM PLASTID *trnL-F* AND  
NUCLEAR INTERNAL TRANSCRIBED  
SPACER (RIBOSOMAL DNA) SEQUENCES**

## CHAPTER FOUR. Phylogenetic Analysis Of *Phylica* L. With An Emphasis On Island Species: Evidence From Plastid *trnL-F* DNA And Nuclear Internal Transcribed Spacer (Ribosomal DNA) Sequences

### Abstract

The tribe Phyliceae consists of *Noltea* Reichb., a monotypic genus from South Africa, *Nesiota* Hook. f., a monotypic genus from St Helena, and *Phylica* L., a genus of about 150 species from southern Africa (mostly Cape Province), St Helena (*P. polifolia*), the Tristan da Cunha Group and New Amsterdam (*P. arborea*), Mauritius and Réunion (*P. nitida*) and Madagascar (*P. emirnensis* and *P. bathiei*). The relationships of the island species were evaluated using sequences for plastid *trnL-F* DNA (intron/spacer) and the internal transcribed spacers of nuclear ribosomal DNA (ITS). Most of the species on the mainland are ericoid shrubs adapted to specific edaphic conditions, a range of different pollinators resulting in diverse inflorescence and floral structures and the increasingly arid climate of the region which has resulted in adaptations in vegetative features such as the reduction in leaf size. In contrast some of the island species and the genera *Nesiota* and *Noltea* are broad-leaved trees or shrubs that have retained other putatively primitive characteristics such as a paniculate inflorescence and a cyathiform calyx tube. The monotypic genera *Nesiota* and *Noltea* were found to be palaeoendemic species within the context of the tribe. The island species of *Phylica* formed a monophyletic group together with the widespread mainland species *P. paniculata*. Within the context of this 'island group', the Mascarene species *P. nitida* was found to be palaeoendemic and the St Helenan, Tristan da Cunha Group and New Amsterdam species were found to be recently derived neoendemic species. The plesiomorphic, generalist morphology of the island species contrasts with the derived morphological characteristics of the majority of mainland species, but the 'island group' occupies a derived position in the phylogenetic trees, thus indicating either a reversal or retention of these primitive traits.

## 4.1. Introduction

*Phylica* L. (Rhamnaceae) was described by Linnaeus in *Species Plantarum* (1753) and has a varied taxonomic history with some authors recognising numerous segregates (Table 4.1). The latest revision of the genus by Pillans (1942) included 150 species. The genera *Soulangia* Brongn., *Trichocephalus* Brongn., *Petalopogon* Reiss., *Tylanthus* Reiss., *Walpersia* Reiss. and *Calophylica* Presl were all sunk into *Phylica* by Pillans because he found that newly discovered morphologically intermediate species meant that these segregates could not be adequately distinguished from *Phylica*. Although Pillans' monograph does not give any ideas concerning the phylogeny of *Phylica*, he placed putatively closely related species together in the order that he listed them. Some of the species which had been placed in *Soulangia*, including many of the island species, were grouped together. On this basis the likely mainland relatives of island *Phylica* species would be those which were placed in this genus.

*Phylica* is distributed through parts of southern Africa including South Africa, Zimbabwe, Tanzania and Malawi, as well as Madagascar, Mauritius, Réunion, New Amsterdam, the Tristan da Cunha Group (Tristan da Cunha, Nightingale, Inaccessible and Gough Islands) and St Helena. The distribution of *Phylica* is shown in Figure 4.1. The vast majority of species occur in Cape Province and are a component of fynbos vegetation. Richardson *et al.* (submitted) found that both *Nesiota* Hook.f. and *Noltea* Reichb. were closely related to *Phylica*. *Nesiota* and *Noltea* are both monotypic genera from St Helena and Cape Province, South Africa, respectively.

### 4.1.1. Taxonomic history of *Phylica* island species

Table 4.1 indicates some of the problems associated with the taxonomy of island species of *Phylica*. For example, Don (1932) considered *Phylica* on St Helena to represent two species in separate genera, *Trichocephalus ramosissima* Don and *Soulangia thymifolia* Brongn. Pillans later lumped these two species into a single

species, *Phylica polifolia*. Hemsley (1885) stated that *Phylica* plants from the Tristan da Cunha Group, New Amsterdam, Bourbon (Réunion) and Mauritius and perhaps Madagascar were one species, *P. nitida*. The Bourbon specimens examined had rather smaller flowers with shorter calyx-lobes; otherwise there is less difference between them and some from the Tristan da Cunha Group than between specimens from the Tristan da Cunha Group alone. Christopherson *et al.* (1937) stated that *P. arborea* was found on the Tristan da Cunha Group, New Amsterdam and the Mascarenes, i.e. he also thought that the *Phylica* species from these islands were conspecific. Pillans (1942) listed five species of *Phylica* found on islands. These were *P. polifolia* from St Helena, *P. arborea* from the Tristan da Cunha Group, Mauritius and New Amsterdam, *P. mauritiana* from Madagascar, Mauritius and Réunion, *P. emirnensis* from Madagascar and Tanzania and *P. bathiei* from Madagascar. Guého (1977) differentiated *P. nitida* from Mauritius and Réunion from *P. arborea* from the Tristan da Cunha Group and New Amsterdam. The current classification of island species of *Phylica* therefore stands as follows: *P. polifolia* Pillans from St Helena, *P. arborea* Thouars from the Tristan da Cunha Group and New Amsterdam, *P. nitida* Lam. from Mauritius and Réunion, *P. emirnensis* Pillans and *P. bathiei* Pillans from Madagascar and *P. emirnensis* var. *nyasae* Pillans from Tanzania. *Phylica tropica* Baker could also be included in this group as an isolated mainland species in mountainous regions of Malawi and Zimbabwe.

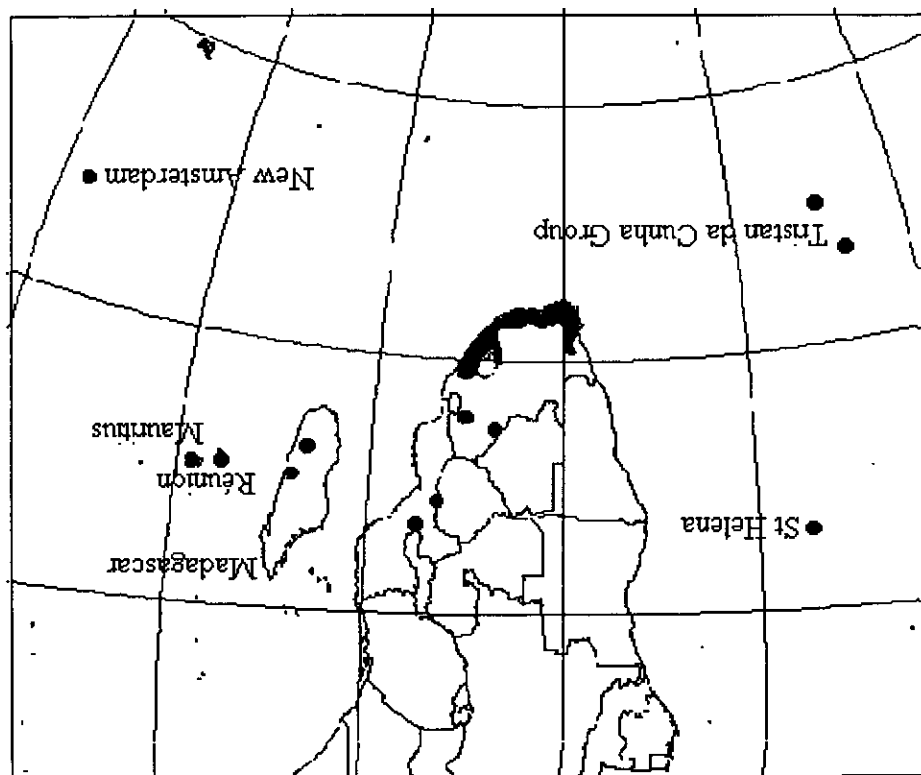
#### 4.1.2. Biogeographic context of *Phylica*

To gain a better understanding of the biological patterns which are apparent today, it is necessary to review the geographic processes partly responsible for them. A chronological history of the geography of southern Africa (particularly the area in which fynbos vegetation is now found, i.e. southwestern Cape Province) and surrounding islands is presented below.

#### 4.1.2.1. Southern Africa

##### 4.1.2.1.1. Pre-Pliocene forest environments (65-5 million years ago; mya)

Pollen remains indicate that tropical rainforest was dominant in the fynbos region 65 mya, including Gondwanan trees of the Podocarpaceae, Proteaceae, Araucariaceae, Casuarinaceae, Cupressaceae, Anacardiaceae, Fabaceae, Euphorbiaceae, Sapindaceae, palms and tree ferns with fynbos elements including members of Proteaceae, Ericaceae, Restionaceae and Rosaceae (Scholtz, 1985). This Gondwanan flora was being enriched by tropical elements entering from the north. Throughout the world the uniformly warm oceans and lack of ice meant that sea levels were much higher than at present and the shoreline in the fynbos region was located near the base of the mountain ranges. About 35 mya a drier phase resulted in the formation of a proto-fynbos with forested areas giving way to a drier type of woodland, which may have included many fynbos elements (Scholtz, 1985). As reconstructed from pollen sequences there was a return to warm wet climates and sub-tropical forests 25 mya with Neogene vegetation, including palms in the Cape region (Coetzee, 1978a,b; Coetzee and Rogers, 1982; Coetzee *et al.*, 1983; Coetzee and Muller 1984; Scott, 1995). The transition from sub-tropical forest to fynbos vegetation has been linked to developments in the southern ocean. Around 16 mya the Antarctic ice sheet began to expand and Antarctica finally separated from South America around 13 mya allowing the development of a cold Circum-Antarctic (Benguela) current which was crucial to the development of the climate of southern Africa (Shackleton and Kennet, 1975; Van Zinderen Bakker, 1975; Coetzee, 1978a,b; Siesser, 1978; Kennet, 1980). This cold ocean current along the west coast aridified southwest Africa (Siesser, 1980). Sea levels also dropped, and sand was blown inland to form the large dunefields that exist today (Coetzee, 1983). Occasional warmer phases allowed a rise in sea levels resulting in deposition of marine sediments which today support alkaline loving endemics. Many plants of the Neogene sub-tropical forest were lost during the increasingly arid Pleiocene and Pleistocene which led to the formation of dry Cape and Karoo vegetation in South

Figure 4.1. Distribution Of *Phyllca*.

Africa (Coetzee, 1978). Like St Helena, the Afro-montane regions and the eastern Cape forests at Kynsna have vegetation which could be relicts of the Miocene wet forest. Putative relict arboresecent composites are found in Afro-Montane regions (e.g. *Dendroseneo*) and Kynsna forests, Madagascar and the Mascarenes (e.g. *Brachylaena*). Plant forms which no longer occur in Africa but still exist on other fragments of Gondwanaland such as Madagascar were still found in fynbos regions. Many of the fynbos plants that exist today still had not appeared. However, by the early Pliocene there was open fynbos vegetation in the region (Scott, 1995).



Table 4.1. Taxonomic history of *Phylica*. Island species are highlighted in bold type.

Author	Tribe	Genus	Species
Brongniart & Dumas (1827)	n/a	<i>Trichocephalus</i>	<i>T. stipularis</i> Brongn. (=Phylica <i>stipularis</i> L.), <i>T. spicatus</i> Brongn. (=P. <i>spicata</i> L.)
		<i>Phylica</i> subgenus <i>Ericoideae</i>	<i>P. parviflora</i> L., <i>P. ericoides</i> L., <i>P. acerosa</i> Willd. (=P. <i>ericoides</i> L.), <i>P. nitida</i> Lam., <i>P. reflexa</i> Lam. (=P. <i>dioica</i> L.)
		<i>Phylica</i> subgenus <i>Strigosae</i>	<i>P. bicolor</i> L. (=P. <i>strigosa</i> Berg., <i>P. pinea</i> Thunb., <i>P. rosmarinifolia</i> Lam. (=P. <i>imberbis</i> Berg.), <i>P. villosa</i> Thunb., <i>P. horizontalis</i> Vent. (=P. <i>plumosa</i> L.), <i>P. plumosa</i> L., <i>P. squarrosa</i> Vent. (=P. <i>plumosa</i> L.), <i>P. capitata</i> Thunb. (=P. <i>pubescens</i> Ait.)
		<i>Soulangia</i>	<i>S. axillaris</i> Brongn. (=P. <i>axillaris</i> Lam.), <i>S. oleaefolia</i> Brongn. (=P. <i>oleaefolia</i> Vent.), <b><i>S. thymifolia</i> Brongn. (=P. <i>polifolia</i> (Vahl) Pillans)</b> , <i>S. paniculata</i> Brongn. (=P. <i>paniculata</i> Willd.), <i>S. buxifolia</i> Brongn. (=P. <i>buxifolia</i> L.), <i>S. cordata</i> Brongn. (=P. <i>buxifolia</i> L.)
Don (1832)	n/a	<i>Trichocephalus</i>	<i>T. stipularis</i> Brongn. (=Phylica <i>stipularis</i> L.), <i>T. spicatus</i> Brongn. (=P. <i>spicata</i> L.), <i>T. elliptica</i> Don (=Nesiota <i>elliptica</i> Hook. F.), <b><i>T. ramosissima</i> Don (=P. <i>polifolia</i> (Vahl) Pillans)</b>
		<i>Phylica</i> sect. <i>Ericoides</i>	<i>P. parviflora</i> L., <i>P. ericoides</i> L., <i>P. glabrata</i> Thunb., <i>P. acerosa</i> Willd. (=P. <i>ericoides</i> L.), <i>P. nitida</i> Lam., <i>P. secunda</i> Thunb. (=P. <i>imberbis</i> Berg.), <i>P. australis</i> Link. (=P. <i>parviflora</i> Berg.), <i>P. pumila</i> Wendl. (=P. <i>plumosa</i> L.), <i>P. excelsa</i> Wendl., <i>P. callosa</i> L. f., <i>P. elongata</i> Willd. (=Staavia <i>globosa</i> Sond.), <i>P. squamosa</i> Willd. (=Raspalia <i>passerinoidea</i> Oliv.), <i>P. rubra</i> Willd., <i>P. microcephala</i> Willd. (=P. <i>ericoides</i> L.)
		<i>Phylica</i> sect. <i>Strigosa</i>	<i>P. bicolor</i> L. (=P. <i>strigosa</i> Berg.), <i>P. pinea</i> Thunb., <i>P. rosmarinifolia</i> Lam. (=P. <i>imberbis</i> Berg.), <i>P. villosa</i> Thunb., <i>P. horizontalis</i> Vent. (=P. <i>plumosa</i> L.), <i>P. plumosa</i> L., <i>P. squarrosa</i> Vent. (=P. <i>plumosa</i> L.), <i>P. capitata</i> Thunb. (=P. <i>pubescens</i> Ait.), <i>P. commelini</i> Spreng. (=P.

*plumosa* L.), *P. lanceolata* Thunb. (= *P. imberbis* Berg.), *P. villosa* Thunb., *P. cylindrica* Wendl., *P. imberbis* L., *P. papillosa* Wendl. (= *P. excelsa* Wendl.), *P. divaricata* Vent., *P. ledifolia* Desf. (= *P. paniculata* Willd.), *P. trichotoma* Thunb. (= *Staavia globosa* Sond.), *P. globosa* Thunb. (= *Staavia globosa* Sond.), *P. atrata* Licht.  
*S. axillaris* Brongn. (= *P. axillaris* Lam.), *S. oleaeifolia* Brongn. (= *P. oleaeifolia* Vent.), *S. thymifolia* Brongn. (= *P. polifolia* (Vahl) Pillans), *S. paniculata* Brongn. (= *P. paniculata* Willd.), *S. buxifolia* Brongn. (= *P. buxifolia* L.), *S. pedicellata* G. Don (= *P. villosa* Thunb.), *S. reclinata* G. Don (= *P. pinea* Thunb.), *S. arborea* G. Don (= *P. arborea* Thouars), *S. orientalis* Sweet (= *P. oleaeifolia* Vent.), *S. dioica* G. Don (= *P. dioica* L.), *W. stipularis* Reiss. (= *P. stipularis* L.), *W. curvifolia* Presl (= *P. curvifolia* (Presl) Pillans, *W. squarrosa* Presl (= *P. meyeri* Sond.), *W. spicata* Reiss. (= *P. spicata* L.), *W. hirtifolia* Presl (= *P. cylindrica* Wendl.), *W. rigida* Presl (= *P. rigida* E & Z) *P. cuspidatus* Reiss. (= *P. cuspidata* E & Z) Numerous Cape species  
***T. emirnensis* Tul. (= *P. emirnensis* (Tul.) Pillans, *T. atratus* Presl (= *P. atrata* Licht ex Roemer & Schultes), *T. gracilis* Presl (*P. gracilis* D. Dietr.), *T. littoralis* Presl (= *P. littoralis* D. Dietr.), *T. comosus* Presl (= *P. comosa* Steud.), *T. brevifolius* Presl (= *P. brevifolia* E & Z), *T. thunbergianus* Presl (= *P. thunbergiana* Sond.), *T. debilis* Presl (= *P. debilis* E & Z), *T. virgatus* Presl (= *P. virgata* D. Dietr.), *T. retrorsus* Presl (= *P. retrorsa* E. Mey. Ex Sond.), *T. diosmoides* Presl (= *P. cephalantha* Sond.), *T. willdenowiana* Presl (= *P. willdenowiana* E & Z), *T. abietinus* Presl (= *P. abietina* E & Z), *T. imberbis* Presl (= *P. imberbis* Berg.), *T. lanceolatus* Presl (= *P. imberbis* Berg.), *T. excelsus* Presl (= *P. excelsa* Wendl.), *T. eriophorus* Presl (= *P. ericoides* L.), *T.***

Soulangia

Walpersia

Petalopogon

Phyllia

Tylanthus

Phylliceae

Endlicher  
(1840)

		<i>atratus</i> Presl (= <i>P. atrata</i> Licht. ex Roem & Schultes), <i>T. parviflorus</i> Presl (= <i>P. parviflora</i> Berg.), <i>T. distichus</i> Presl (= <i>P. disticha</i> E & Z), <i>T. callosus</i> Presl (= <i>P. callosa</i> L.), <i>T. gracilis</i> Presl (= <i>P. gracilis</i> D. Dietr.), <i>T. litoralis</i> Presl (= <i>P. litoralis</i> D. Dietr.), <i>T. comosus</i> Presl (= <i>P. comosa</i> Steud.), <i>T. virgatus</i> Presl (= <i>P. virgata</i> D. Dietr.)
	<i>Soulangia</i>	<i>S. paniculata</i> Brongn. (= <i>P. paniculata</i> Willd.), <i>S. oleaefolia</i> Brongn. (= <i>P. oleaefolia</i> Vent.), <b><i>S. thymifolia</i> Brongn. (= <i>P. polifolia</i> (Vahl) Pillans)</b> , <b><i>S. arborea</i> G. Don (= <i>P. arborea</i> Thouars)</b> , <i>S. buxifolia</i> Brongn. (= <i>P. buxifolia</i> L.), <i>S. axillaris</i> Brongn. (= <i>P. axillaris</i> Lam.), <i>S. reclinata</i> G. Don (= <i>P. pinea</i> Thunb.), <i>S. rubra</i> Lindl. (= <i>P. purpurea</i> Sond.), <i>S. subcanescens</i> Presl (= <i>P. crytandroides</i> Sond.), <i>S. plumosa</i> (= <i>P. ambigua</i> Sond.), <i>S. pinea</i> E & Z (= <i>P. villosa</i> Thunb.), <i>S. ledifolia</i> E & Z (= <i>P. lasiocarpa</i> Sond.), <i>S. willdenowiana</i> A. Dietr. (= <i>P. willdenowiana</i> E & Z), <i>S. dioica</i> Don (= <i>P. dioica</i> L.)
	<i>Spyridium</i>	
	<i>Cryptandra</i>	
Not	<i>Trichocephalus</i>	<b><i>T. ramosissimus</i> Don (= <i>P. polifolia</i> (Vahl) Pillans)</b> , <i>T. elongatus</i> E & Z (= <i>P. propinqua</i> Sond.), <i>T. laevis</i> E & Z (= <i>P. laevis</i> Steud.), <i>T. harvey</i> Arnott (= <i>P. harveyi</i> (Arnott) Pillans), <i>T. stipularis</i> Brongn. (= <i>P. stipularis</i> L.), <i>T. spicatus</i> Brongn. (= <i>P. spicata</i> L.), <i>T. trachyphyllus</i> E & Z (= <i>P. trachyphylla</i> D. Dietr.)
Phylliceae		
	<i>Calophyllica</i>	<i>C. gnidioides</i> Presl (= <i>P. gnidioides</i> E & Z)
	<i>Nesiota</i>	<i>N. elliptica</i>
	<i>Noltea</i>	<i>N. africana</i>

#### 4.1.2.1.2. Pliocene origin of seasonality and the birth of Fynbos (5-0 mya)

Six mya the coastal lowlands were covered with open shrubland dominated by grasses, restios, geophytes, and composites (Scott, 1995). Sub-tropical forest vegetation was found near coasts, on sand dunes and along riverbanks as remnants of the previous vegetation of the south-western Cape. The inland plains were grassy woodlands which included many fynbos elements such as proteas, ericas and other ericoid shrubs. Herbivore fossil taxa related to animals of the present day African savannah dating to this period give evidence in support of this type of vegetation (Vrba, 1985). Burnt bones also indicate that fires began to play an important part in the ecology of the landscape. Around five mya fynbos forms increased, the forest declined further, and the first evidence of widespread fire was noted. Four mya saw the inception of a Mediterranean climate with dry summers: rain-bearing westerly winds in winter and dry southeasterly winds in summer. There was also an increased incidence of fire caused by lightning strikes. Three mya fynbos was the predominant vegetation throughout much of western and southern Cape Province with pollen data indicating that *Protea* savanna occurred after the change from sub-tropical forest to more open vegetation around three mya (Scott and Bonnefille, 1986; Scott, 1995). Van Zinderen, Bakker and Muller (1987) studied two offshore boreholes estimated at 250 000 and 550 000 years old which contained high proportions of fynbos elements such as Asteraceae, Ericaceae, Proteaceae and Restionaceae.

One and a half mya saw the start of glacial cycles with a periodicity of 100 000 glacial years and warm interglacials of only 10 000 years. During glacial times, conditions were dry, sea levels dropped, coastal plains were wider, there was less orographic rainfall, frosts were heavy in lowlands and snow was widespread in mountains. Differences in climate between west and east were exaggerated during glacials, and this may explain the greater species diversity in the western region of fynbos compared to the east. The latest glacial was between 75000 and 12000 bp, with grassy vegetation on lowlands with many grazing mammals. The present interglacial period is characterised by shrubby vegetation on lowlands with many browsing mammals.

To summarise, the climate of the Cape has changed from a tropical to a warm temperate forest climate and eventually to a summer dry mediterranean climate. These changes eliminated many taxa, leaving only a few families of xeromorphic plants which now dominate the region. These remaining taxa extended their distributions into areas vacated by the forests. Recent aridification and associated increase in fire has resulted in proliferation of fynbos species. Fire fragments populations promoting evolution of new species. The rapidly changing climate of the region has augmented this process resulting in the adaptation of new features. Significantly there were no catastrophic changes that would have wiped out entire ecosystems. The Cape is subject to two seasonal contrasts with summer droughts and strong dry winds which means fires are easily started when lightning strikes and low winter rainfall and low temperatures which will delay evaporation allowing winter growth. Speciation has also been augmented by the mountainous landscape where virtually every mountain peak has a distinct climate (Linder, 1985). Climatic shifts have allowed certain populations to escape from their particular habitats whereas others remained as they were. The complex geomorphological history of the region has also resulted in a mosaic of different soil types (Partridge, 1997). Many isolated endemic species are closely associated with a particular soil type (Cowling and Richardson, 1995).

The following sections are reviews of the geography and biology of each of the islands on which species of *Phyllica* are found including a summary of the affinities of each of the islands floras.

#### 4.1.2.2. St Helena

St Helena is an island in the southern Atlantic Ocean (15° 56' S, 5° 42' W) with an area of 122 square km. The age of the island has been estimated at 14.3 million years with the main volcanic activity ceasing at about 7.5 mya (Baker *et al.*, 1967). The island has a stable sub-tropical climate which is influenced by the south-east trade wind belt and the Benguela Current. Cronk's work on the St Helenan flora (1987) led to his formulation of a relictual series of island endemics. If endemics are

tabulated in order of increasing taxonomic isolation, the distributions of the hypothetical sister groups form a series. The less isolated endemics generally have closely related species in Africa. The more isolated endemics have related species scattered in the southern hemisphere, often in regions of high endemism such as Andean southern South America and Australasia. These distributions are considered to be relictual.

According to Cronk (1987) the southeast trade wind and southeast Benguela current brought more recently dispersed plants (neoendemics) from southern Africa. They may also have brought palaeoendemics from southern Africa, but subsequent extinction in southern Africa means that the nearest extant relatives are in the New World. Cronk (1987) suggested two main recruitment areas from the east: (i) southern Africa: e.g. *P. polifolia*; (ii) Mascarenes: e.g. *Acalypha rubra* and *Trochetiopsis* spp. These plants were either transported by currents (transported south by Agulhas current and north from Cape Agulhas by the Benguela current) or were once more widespread and have become extinct on the mainland. The St Helenan relict composites have affinities with South America whereas the more recent colonists are southern African (Cronk, 1987). On St Helena neoendemics are generally plants of the arid coastal zone whereas palaeoendemics are generally upland wet-thicket plants (Cronk, 1987).

#### 4.1.2.3. Tristan da Cunha Group

The Tristan da Cunha Group consists of four islands of volcanic origin situated to the west of the mid-Atlantic ridge 2800km from Africa and 3200km from the nearest point in South America. All of these islands differ in size, age and erosional stage. Tristan da Cunha is situated 37° 15' S, 12° 30' W and is the youngest island with the lowest lava flows being about one million years old. The most recent volcanic eruption was in 1961. The oldest Nightingale, is situated 37° 28' S, 12° 32' W and dated at around 18 (+/- 4) million years (this date was taken from Middle Island which is a sea stack near Nightingale). Inaccessible is situated 37° 19' S, 12° 44' W and is 6 (+/-1) million years old. Gough is found 40° 20' S, 10° 00' W and is part of

a separate volcanic mass for which the oldest rocks are dated at about 6 (+/-2) million years old. All of these islands may be regarded as still being volcanically active.

The Tristan da Cunha Group has a cool temperate maritime climate and is under the influence of maritime tropical and maritime polar air masses from the western south Atlantic. The prevailing winds are westerly and consequently rainfall is greater on the western side of islands than on the east. The eastern side is also the warmest part of the islands. The climate on Gough is slightly wetter and cooler than on the other three islands in the Tristan da Cunha Group.

Groves (1981) stated that most of the native and endemic vascular plants of the Tristan da Cunha Group have a South American or south circumpolar distribution or are supposedly closely allied to species that have such a range. Although the islands are geographically closer to southern Africa, the affinity of flowering plants on the archipelago is generally closer to South America. However two thirds of the fern flora have taxa with a greater affiliation with Africa. Cronk's ideas on the St Helenan flora may also apply to the flora of the Tristan da Cunha Group, i.e. palaeoendemics are of a South American or south circumpolar distribution and more recent colonists are South African. According to an ITS sequence analysis, *Pelargonium grossularioides* on the Tristan da Cunha Group is derived from within the South African *Pelargonium* species (Bakker, 1998) indicating that it is a recent introduction. No phylogenetic analyses have been conducted on other taxa from the Tristan da Cunha Group. Apart from *P. arborea* there are no taxa on the islands which appear from their morphology to be palaeoendemic.

#### 4.1.2.4. Madagascar

Madagascar is a continental island in the Indian Ocean. All but the very south of the island is found within the tropics. The initial formation of the Mozambique channel was 250 to 220 mya, and this may have given some isolation from Africa. Between 200 and 155 mya the island split away from Africa (most authors quote 165 mya for the split, but there was still contact through Antarctica at that time). Madagascar, India, Australia and Antarctica split from Gondwana 138 mya.

Madagascar and India split from Antarctica 130 mya and, Madagascar and India split 90-88 mya (Boast and Nairn, 1982; Brenon, 1972; Smith, 1994; Storey *et al.*, 1995).

Leroy (1978) stated that Madagascar has:

“a flora that has differentiated principally through the original Gondwanan stock and has on course of time grown rich through evolution of its members and immigration of newcomers through long distance dispersal.”

Schatz (1996) suggested that the Madagascan flora exhibits a high affinity with Indo-Australo-Malesian floras to the east with three patterns of dispersal/vicariance being identified: (i) Cretaceous dispersal to Madagascar with ensuing distributions from India (and/or South Africa) across Antarctica to South America and Australo-east Malesia during the time of the initial radiation of the angiosperms; (ii) Eocene-Oligocene (and continuing to the present) dispersal to Madagascar (and Africa) from Laurasia and western Malesia via India (pre- and post-collision with India) along ‘Lemurian Stepping Stones’ in the western Indian Ocean; and (iii) continuous (and recent) long-distance dispersal to Madagascar as a function of the prevailing easterly winds and Indian Ocean currents.

#### 4.1.2.5. New Amsterdam

New Amsterdam (37° 47' S; 77° 34' E) is a volcanic island situated roughly midway between Australia and South Africa. The age of the island is estimated as being 690 000 years with the most intense period of volcanic activity being from 400 000 to 200 000 years ago. It is 10 x 7km wide with a land area of *c.* 55km<sup>2</sup>. Steep cliffs from 30-700m skirt most of the island. There are only sixteen flowering plants on New Amsterdam, four of which are endemic, and seventeen cryptogams. Two of the flowering plants are endemic to New Amsterdam and the nearby island of St Paul. One flowering plant is American (also found on the Tristan da Cunha Group, Marion and Kerguelen islands), three are from New Zealand, two are generally dispersed throughout the south temperate zone, one is cosmopolitan, *Spartina arundinacea* and *Uncinia brevicaulis* var. *brevicaulis* are found only on New



Amsterdam, St Paul and the Tristan da Cunha Group, and *Phylica arborea* is found only on the Tristan da Cunha Group and New Amsterdam. Five of the flowering plants on New Amsterdam are also found on the Tristan da Cunha Group indicating a strong affinity between the floras of these two islands.

#### 4.1.2.6. Mauritius

Mauritius is volcanic in origin except for the fringing coral reefs and composed of alkaline olivine basalts. The island is about 7.8 million years old and is located 840km from Madagascar and approximately 200km from Réunion. The island has a varied topography with ranges of peaks, plateaux and low lying plains. There are 800-900 species of plants, roughly one third of which are endemic (Strahm, 1984). According to Cadet (1977) 70% of the genera of flowering plants on Mauritius have closest relatives on Madagascar or the African mainland, 8% are endemic and 8% have oriental indo-pacific relatives.

#### 4.1.2.7. Réunion

Réunion is 2 million years old and situated 780km east of Madagascar and 200km south-west of Mauritius. The centre of the island is composed of a volcanic mountain culminating at Piton de Neiges at 3069m. There are *c.* 500 species of indigenous seed plants. The floral affinities of Réunion are probably similar to those of Mauritius since the two islands are so geographically close. As Réunion is younger than Mauritius, it may have gained at least some of its flora from Mauritius.

#### 4.1.3. Morphology of Phylicae

Pillans (1942) pointed out evidence of three lines of evolution within *Phylica*. These included change from a racemose inflorescence through a spicate to a capitate inflorescence (Figure 4.2), the lengthening of the calyx tube and the reduction in size or complete disappearance of the petals (Figure 4.3). Vegetative changes included

development of a low shrubby habit (see Plates 1-5, page 143; photographs by the author) and narrow, revolute leaves. This can be contrasted with the putatively primitive arborescent, broad-leaved form found in some species. *Phylica* and other genera such as *Erica* (Ericaceae) adapted to the changing environment in the Cape. The evolution of the inflorescence and the calyx tube (Figures 4.2 and 4.3) could be adaptations to different pollinators. There have been few studies on pollination biology of particular plant groups in the fynbos, and there have been no studies on *Phylica*. However, it is clear that competition for the attention of animal pollinators has been one of the major driving forces in the evolution of the great diversity of floral morphology in fynbos. Urn-shaped flowers in *Erica* are pollinated by bees, whereas tubular flowered species are pollinated by long proboscid flies, such as horse flies, tangle winged flies and bee flies (Schumann and Kirsten, 1992). Some of the tubular flowers in *Phylica* could also be pollinated by these insects. Some *Phylica* species appear to be 'generalists' being pollinated by a range of different insects. I observed *Phylica pinea*, which is a fynbos species with similar floral morphology to the island species, being visited by bees, beetles and flies. The development of a low habit and the reduction in leaf size are responses to increased aridity (Plates 1-6 on page 143 illustrate the range of habits of *Phylica* species). Most of the *Phylica* species on the mainland are ericoid shrubs, whereas some of the island species and the genera *Nesiota* and *Noltea* are broad-leaved trees and shrubs that have not developed specialised pollinator relationships (Figures 4.4 and 4.5).

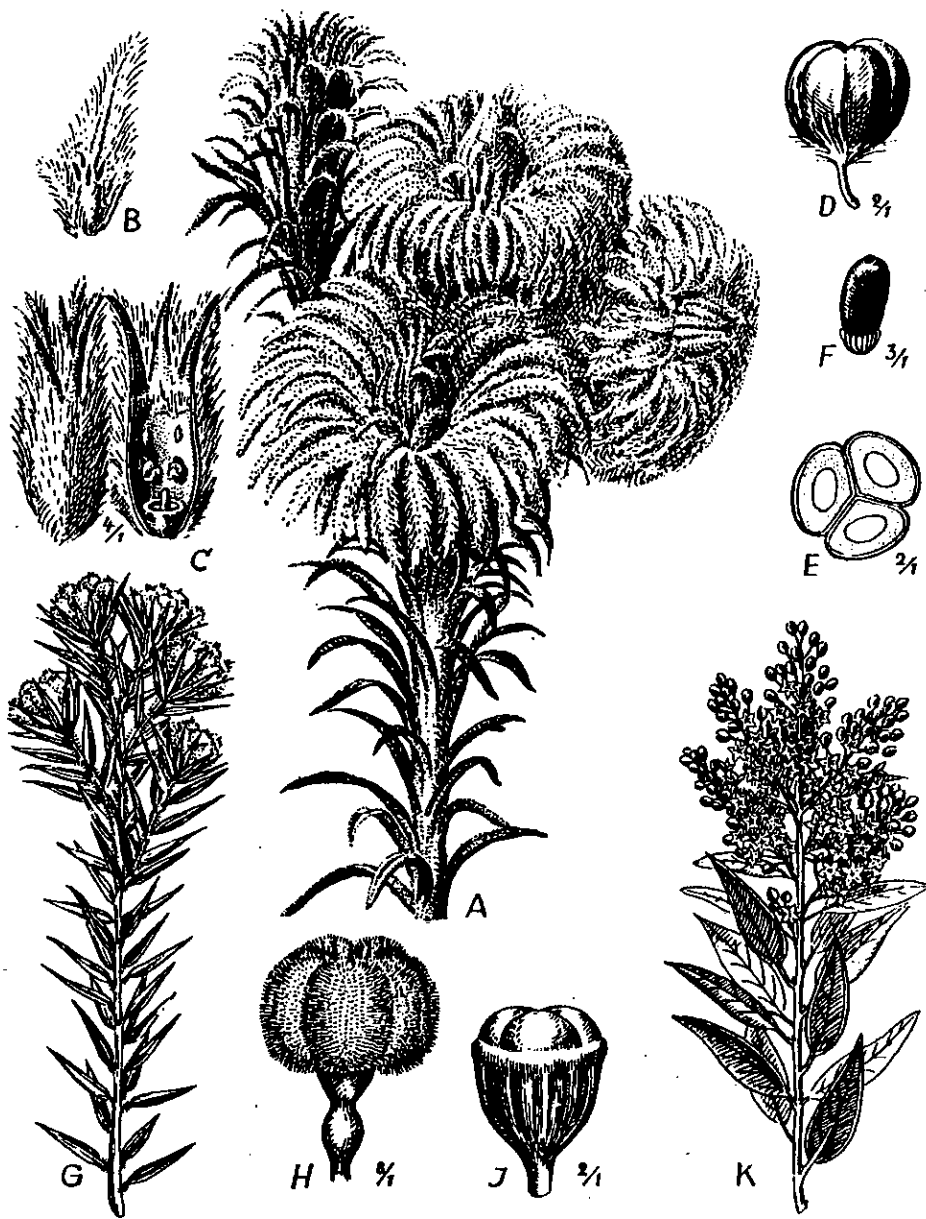


Figure 4.2. *Phylica pubescens*: A. Inflorescence; B. Flower; C. Transverse section of flower; D. Fruit; E. Capsule; Cross section of capsule; F. Seed with elaiosome. *Phylica virgata*: G. Inflorescence; H. Flower. *P. oleaefolia*: J. Fruit; K. Inflorescence (from Suessenguth, 1953).

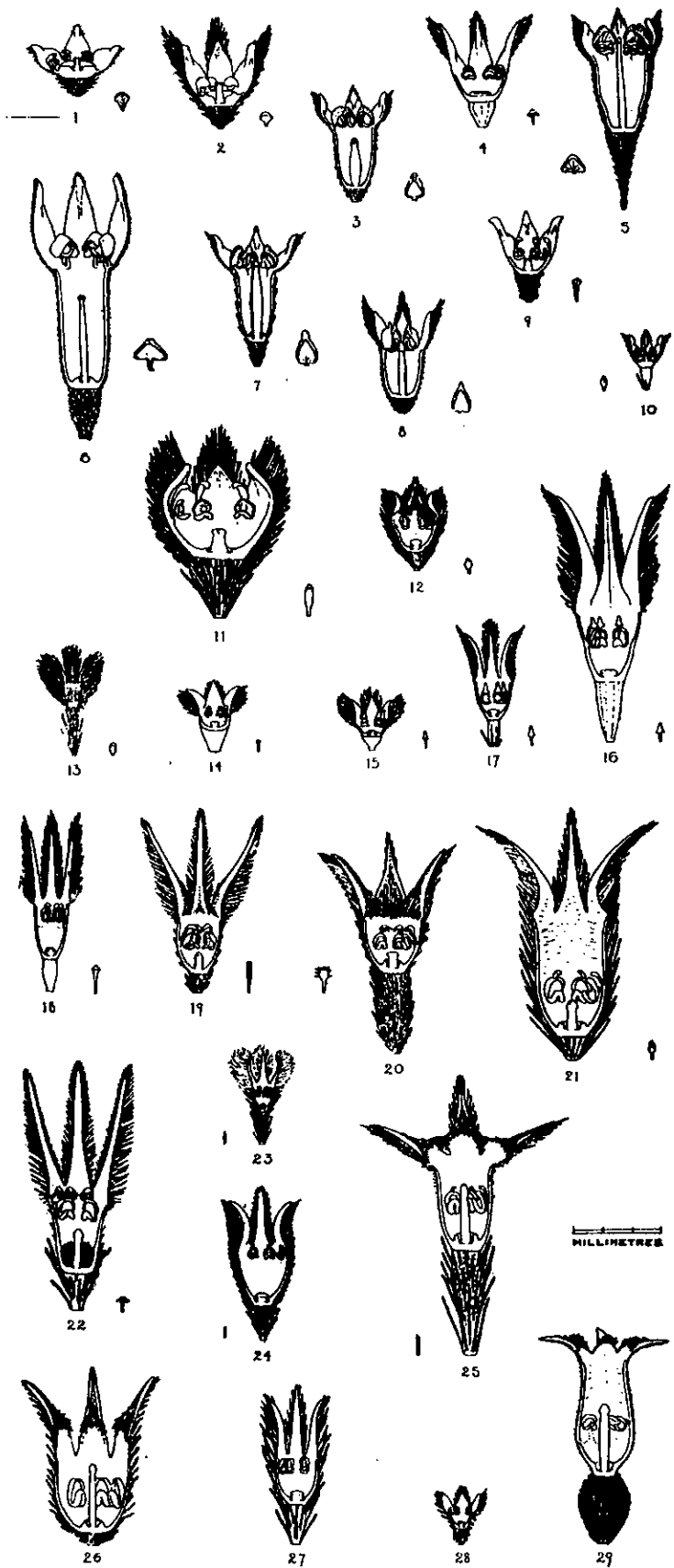
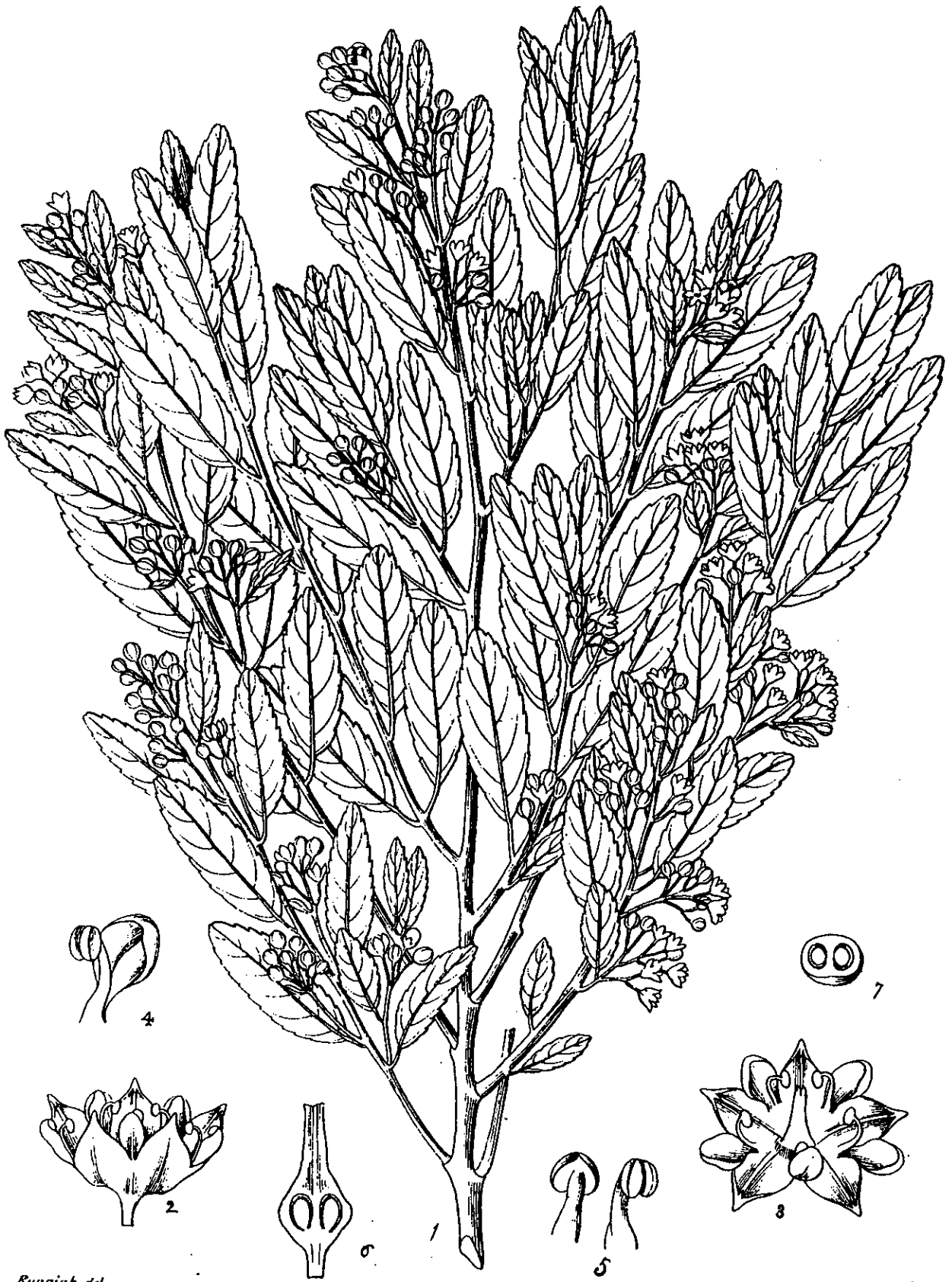


Figure 4.3. Floral morphology of a selection of *Phylica* species (from Pillans, 1942).

Species in Figure 4.3 and their distributions.

1. *P. paniculata*. Southern Africa.
2. *P. arborea*. Tristan da Cunha Group, New Amsterdam.
3. *P. imberbis*. Western Cape.
4. *P. callosa*. Western Cape.
5. *P. willdenowiana*. Western Cape.
6. *P. gnidioides*. Western Cape.
7. *P. velutina*. Western Cape.
8. *P. excelsa*. Western Cape.
9. *P. greyii*. Western Cape.
10. *P. minutiflora*. Western Cape.
11. *P. emirnensis*. Madagascar.
12. *P. thunbergiana*. Western Cape.
13. *P. keetii*. Western Cape.
14. *P. ericoides*. Western Cape.
15. *P. disticha*. Western Cape.
16. *P. propinqua*. Western Cape.
17. *P. gracilis*. Western Cape.
18. *P. amoena*. Western Cape.
19. *P. spicata*. Western Cape.
20. *P. bolusii*. Western Cape.
21. *P. pubescens*. Western Cape.
22. *P. stipularis*. Western Cape.
23. *P. debilis*. Western Cape.
24. *P. odorata*. Western Cape.
25. *P. affinis*. Western Cape.
26. *P. rigida*. Western Cape.
27. *P. constricta*. Western Cape.
28. *P. comptonii*. Western Cape.
29. *P. retorta*. Western Cape.



Rungiah, del.

Dunphy, lith.

Figure 4.4. *Noltea africana* (from Sim, 1907).

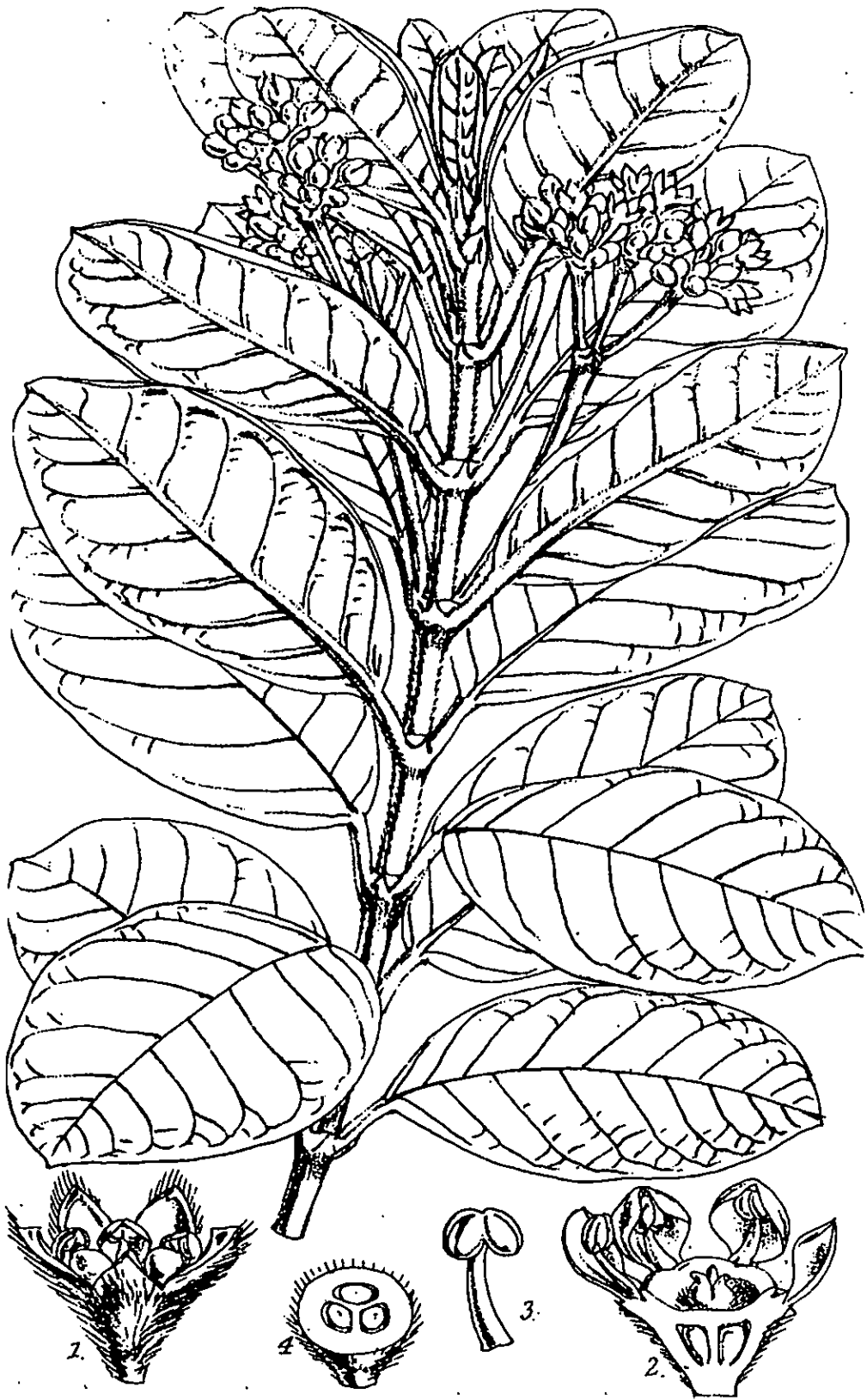


Figure 4.5. *Nesiota elliptica* (from Hooker, 1870).



Plate 1. *Phylica montana*, Cape Province.



Plate 2. *P. pubescens*, Cape Province.



Plate 3. *P. cryptandroides*, Cape Province.



Plate 4. *P. plumigera*, Cape Province.



Plate 5. *P. ericoides*, Cape Province.

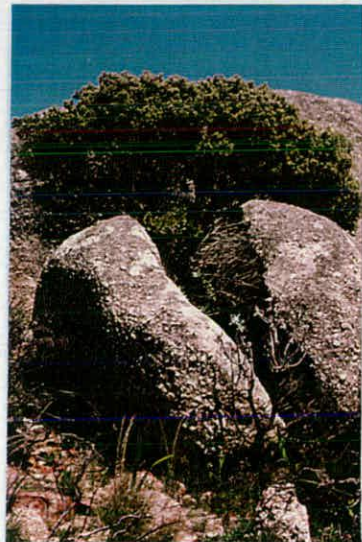


Plate 6. *P. buxifolia*, Cape Province.



#### 4.1.4. Phylliceae biogeography

One of the main objectives of this study was to establish the relationship between the island species of *Phyllica* and those from mainland Africa using nucleotide sequence data. Baldwin *et al.* (1992) studied the Hawaiian silversword alliance as an example of the use of DNA sequencing to illustrate evolution of neoendemic island species in contrast to the 'slower' rates of morphological change exhibited by their nearest relatives on the mainland. This may be contrasted with some of the endemic species of St Helena in which it appears that evolution of their closest relatives on the continent has been progressing more rapidly than on the island. Cronk (1992) suggested a relictual series of palaeoendemics, the components of which were distinguished by the relative contribution of *in situ* evolution and *ex situ* extinction to the resulting endemism. *Petrobium* (Compositae), *Commidendrum* (Compositae), *Lachanodes* (Compositae) and *Trochetiopsis* (Sterculiaceae) are considered to be examples of palaeoendemic genera on St Helena which have retained plesiomorphic morphologies.

The question of whether Phylliceae/*Phyllica* was once more widely distributed in continental Africa and Madagascar or whether it has dispersed to outlying regions more recently is of interest. One could envisage that a Phylliceae/*Phyllica* ancestor had an ancient widespread distribution throughout southern Africa and the characteristics and distribution of the group changed with the changing climate. It is also possible that *Phyllica* originally evolved on islands and dispersed to Africa. The birth of fynbos has been dated at around six mya when elements such as grasses, restios, geophytes, composites, *Protea*, *Erica* and other ericoid shrubs began to dominate (Scott, 1995). All *Phyllica* species and *Nesiota* have some adaptations to drier climates, whereas the related genus *Noltea*, which grows outside of fynbos regions in coastal rainforest, does not. Some attempt could be made to date the emergence and evolutionary development of *Phyllica* and compare it with the appearance of islands. The accurate dating of the emergence of volcanic islands gives a time limit for the dispersal of island species. This study was aimed at determining whether island

species are 'relictual' (i.e. from the previous distribution) or whether they are the product of more recent dispersal events.

The western island species include *P. polifolia* Pillans from St Helena and *P. arborea* Thouars from the Tristan da Cunha Group and New Amsterdam. *Phylica nitida* from Mauritius and Réunion, *P. emirnensis* Pillans and *P. bathiei* Pillans from Madagascar and *P. emirnensis* var. *nyasae* Pillans from Tanzania can be classed as eastern island species. *Phylica tropica* Baker from Malawi could also be included in this group as an isolated mainland species found in the ericaceous belt in mountainous regions of Malawi, Zimbabwe and Mozambique. Axelrod and Raven (1978) have argued that a few Cape genera discontinuous with Madagascar (e.g. *Aristea*, *Philippia*, *Phylica*, *Restio*) had reached Madagascar by long distance dispersal. A continuously favourable habitat between the mountains of the Cape region and those of Madagascar was thought to be unlikely at any time. They supported this idea with the fact that these genera constitute a small proportion of the floras, i.e. there are only two species of *Phylica* in Madagascar. However, molecular data have shown that Madagascar also has relict genera such as *Bathiorhamnus* (Rhamnaceae; Chapter Two) that only have one or two species. Axelrod and Raven (1978) stated that:

“the large number of species of the important genera in the Cape vegetation is a striking feature of the flora as compared with the nearby refugial temperate rainforest to the east where genera have few species.”

Examples of refugia in and around southern Africa include oceanic islands, mountains, temperate rainforest and riverbanks. The species found in these areas may be the products of more recent dispersal events, or they could have been in refugia for some time but only recently provided stock for dispersal to islands or other favourable areas. Some species of *Phylica*, such as *P. paniculata*, are found only along permanent watercourses or on wet mountains in southern Africa. These species may not be able to withstand dry conditions. Peripheral endemics in the genus might be markers for the recurrent expansions and contractions that are part of the history of every centre of endemism. Members of Phyliceae that presently occupy these

hypothetically refugial distributions could be relictual. *Phylica paniculata* Willd. exhibits putatively primitive characteristics (arborescent habit, paniculate inflorescence, cyathiform calyx, broader than ericoid leaves) and has a wide distribution throughout South Africa and into southern Zimbabwe. This species appears morphologically to be closely related to the island species, *P. arborea*, *P. polifolia* and *P. nitida*. The distribution of *P. paniculata* could be the product of recent dispersal or an older distribution, i.e. these montane regions could be refugia. Other outlying species with a southeast African distribution include *P. natalensis*, *P. thodei*, *P. gnidioides*, *P. simii*, *P. tysoni*, and *P. litoralis*. The Madagascan and Mascarene species of *Phylica* could have been derived from eastern populations of *P. paniculata* or from these other southeast African species.

Selection of outgroups for the study of *Phylica* was based upon the molecular phylogeny of Rhamnaceae (Chapter Two) in which *Phylica* falls in a monophyletic group with two monotypic genera, *Noltea* from Cape Province and *Nesiota* from St Helena. This group was found within the ziziphoid group, which also included representatives of *Ceanothus*, *Colubrina*, *Lasiodiscus*, Pomaderreae and *Alphitonia*. The choice of regions to be sequenced was determined by sequencing two closely related species of *Phylica* and members of outgroup taxa from the ziziphoid group. Sequences of *trnL-F* plastid DNA and the internal transcribed spacer (ITS) nuclear rDNA were found to be divergent enough to resolve relationships between most species of *Phylica*. Sequences of these regions were therefore produced for each of the island species of *Phylica*, representatives of the main groups found on mainland Africa (particularly those with a morphology similar to that of the island species, e.g. *P. paniculata*, *P. buxifolia*, *P. oleaefolia*), *Nesiota*, *Noltea* and the outgroups. The use of the internal transcribed spacers (ITS) of nuclear ribosomal DNA in phylogenetic analyses has been demonstrated in the past by Baldwin (1992, 1993) in his study of the Hawaiian silversword alliance and Californian tarweeds and *Calycadenia* (Compositae). The use of the ITS region in estimating phylogeny in angiosperms has been reviewed by Baldwin *et al.* (1995). The plastid *trnL-F* region has also been used in phylogenetic analyses, e.g. *Gentiana* (Gielly and Taberlet,

1996), Haemodoraceae (Hopper *et al.*, in press), Iridaceae (Reeves *et al.*, 1997), Plumbaginaceae (Llédó *et al.*, 1998) and Rhamnaceae (Richardson *et al.*, submitted).

## 4.2. Aims Of Study

1. To investigate the biogeography of *Phylica*.
2. To determine whether *Phylica* originated in Africa or on the islands.
3. To determine whether the island species of *Phylica* are palaeo- or neo-endemic taxa.
4. To determine the nearest mainland relatives of the island species.
5. To determine whether the island taxa are monophyletic.
6. To determine how many species there are on the islands.
7. To determine the sequence of colonisation events of the islands.

## 4.3. Materials and Methods

### 4.3.1. Material for molecular analysis

Sources of plant material and vouchers or accessions used in this analysis are listed in Table 4.2. Silica gel dried material of *P. arborea* and most of the South African species included in this study were collected during a field trip undertaken in September and October 1996. Total DNA was extracted from fresh or silica gel dried leaves and herbarium specimens. No fresh material of *P. emirnensis*, *P. thodei*, *P. tropica* or *P. natalensis* could be obtained, and some sequencing work was not possible on the DNA obtained from herbarium material because it was too degraded. DNA could not be obtained for *P. bathiei* or *P. emirnensis* var. *nyasae*. The South African species of *Phylica* chosen were used as they represented different infra-generic morphological groupings as suggested by Weitz (pers. comm.). Weitz and Richardson *et al.* (unpubl.) have sampled an additional 30 species of *Phylica* for both ITS and *trnL-F* and the set used here are representative of the phylogenetic

distribution in the genus. No additional species are more closely related to the island group than those used in my study.

#### 4.3.2. DNA extraction

DNA was extracted from *c.* 1.0g fresh, 0.2-0.25g silica gel-dried leaves or 0.1-0.2g of material from herbarium sheets using a 2X CTAB method modified from Doyle and Doyle (1987). DNA was precipitated using isopropanol instead of ethanol because it was found to be more reliable. DNA extracted from herbarium material was found to precipitate better if left for at least three weeks at -20°C (Fay *et al.*, 1998). The reasons for this are unclear, but it could be due to stronger interactions between secondary compounds and DNA in dried herbarium material or because the DNA from herbarium specimens is degraded and therefore takes longer to precipitate. All samples were purified on caesium chloride/ethidium bromide gradients (1.55g/ml).

#### 4.3.3. Gene amplification and purification

I amplified the *trnL-F* region (Taberlet *et al.*, 1991) using the forward primer *c* and the reverse primer *f*. Amplification of *trnL-F* involved 28 cycles, each consisting of: denaturation at 94°C for one minute; annealing of primer at 50°C for 30 seconds and nucleic acid extension at 72°C for one minute. The ITS region was amplified using AB101R and AB102F primers (developed by G. Sheridan, University of Bath; Table 4.3). Amplification of ITS involved 30 cycles of denaturation at 97°C for one minute; annealing of primer at 50°C for one minute and nucleic acid extension at 72°C for three minutes. The production of PCR templates for some samples, particularly those from herbarium specimens, required double amplifications. For ITS, this involved 20 amplification cycles using AB101R and AB102F primers followed by 24 cycles using AB101 and ITS 4 primers. For *trnL-F*, this involved 20 amplification cycles using *a* and *f* primers followed by 24 cycles using *c* and *f* primers. Excessive amplification cycles or the use of the same primer pairs in the

first and second amplification resulted in primer dimers that made interpretation of electropherograms difficult. Amplification products were purified using Magic mini columns (Promega) following protocols provided by the manufacturer.

Table 4.2. Sequences of AB101 and AB102 primers (G. Sheridan, University of Bath).

Primer	Primer sequence
AB101F	ACGAATTCATGGTCCGGTGAAGTGTTTCG
AB102R	TAGAATTCCCCGGTTCGCTCGCCGTTAC

#### 4.3.4. DNA sequencing

Modified dideoxy cycle sequencing with dye terminators run on an ABI 373A or 377 automated sequencer (according to the manufacturer's protocols; Applied Biosystems, Inc.) was used to sequence the amplification products directly. I edited and assembled the sequences using the Sequence Navigator and Autoassembler software programs of Applied Biosystems, Inc.

#### 4.3.5. Sequence alignment

For both ITS and *trnL-F*, I performed an initial alignment for the first five sequences produced using Clustal (Higgins, Bleasby and Fuchs, 1992). Subsequent sequences were aligned by eye.

#### 4.3.6. Phylogenetic analysis

I analysed the data using the parsimony algorithm of the software package PAUP\* version 4.0d64 for Macintosh (Swofford, 1998). Searches were conducted on the separate ITS and *trnL-F* data sets (which included a matrix of 17 *trnL-F* indel characters) and on both data sets combined since I found them to be congruent. The

heuristic search strategy was the same as that used in the previous chapter. For the combined analysis, all taxa either *trnL-F* or ITS missing were removed; if these were retained, the number of trees found was high, resolution was low, and bootstrap values were low (results not shown). Successive weighting was not used in the bootstrap because this procedure is prone to overestimate support with low levels of divergence (such as within *Phyllica*).

#### 4.3.7. Molecular clock

The timing of dispersal to islands may be roughly estimated by the number of nucleotide substitutions per million years based on a reasonably well established geological event. If *P. nitida* on Réunion dispersed there at the earliest possible time, i.e. two mya the ITS tree indicates that *P. nitida* evolved four autapomorphies in the two million years since it arrived on the island or two autapomorphies per million years giving two ITS nucleotide substitutions every one million years. This was used to estimate the divergence times of other lineages. This rate can be compared with that calculated for *Dendroseris* (Asteraceae) (Sang *et al.*, 1995) an endemic genus from the volcanic islands of the Juan Fernandez archipelago. This archipelago consists of two islands, Masatierra which arose four mya and Masafuera which arose 1-2 mya. If *D. regia* which is endemic to Masafuera dispersed there from Masatierra at the earliest possible time, i.e. 1-2 mya the ITS phylogeny indicates that it evolved 6 autapomorphies since it diverged from its closest relative or 3-6 per million years which is a higher rate than that found for *P. nitida* on Réunion. This demonstrates the error if clocks calibrated in distantly related taxa are used. When comparing closely related taxa, as is the case in this instance, it is less likely that there will be large differences in rates of change among them.

Table 4.3. Taxon Accession data.

SPECIES	PROVENANCE	VOUCHER	Date material collected
<i>Alphitonia excelsa</i> Reiss.	Australia	Chase 2179 (K)	silica gel
<i>Ceanothus coeruleus</i> Lag.	SW USA	Chase 2413 (K)	silica gel
<i>Colubrina asiatica</i> Brongn. (1)	W. Australia	Chase 905 (K)	silica gel
<i>Colubrina reclinata</i> (L'Hér.) Brongn. (2)	W. Australia	Chase 2115 (K)	silica gel
<i>Nesiota elliptica</i> (Roxb.) Hook. f.	St Helena	Chase 500 (K)	silica gel
<i>Noltea africana</i> (L.) Reichb. (ITS)	South Africa (Cape Province)	JER48	silica gel
<i>Noltea africana</i> (L.) Reichb. (trnL-F)	South Africa (Cape Province)	Bayliss BS6824 49 (K)	1974
<i>Phylica arborea</i> Thouars	Tristan da Cunha	JER51	silica gel
<i>Phylica arborea</i> Thouars	New Amsterdam	JER166	silica gel
<i>Phylica buxifolia</i> L.	South Africa (Cape Province)	JER1	silica gel
<i>Phylica cryptandroides</i> Sond.	South Africa (Cape Province)	JER28	silica gel
<i>Phylica emirnensis</i> (Tulasne) Pillans	Madagascar	Goldblatt & Schatz 8972	1989
<i>Phylica natalensis</i> Pillans	South Africa (Natal)	Nicholson <i>s.n.</i>	1969
<i>Phylica oleaefolia</i> Vent.	South Africa (Cape Province)	JER25	silica gel
<i>Phylica stipularis</i> L.	South Africa (Cape Province)	JER4	silica gel
<i>Phylica stipularis</i> L. (2)	South Africa (Cape Province)	FMW1080	silica gel
<i>Phylica plumigera</i> Pillans	South Africa (Cape Province)	JER26	silica gel
<i>Phylica ericoides</i> L.	South Africa (Cape Province)	JER13	silica gel
<i>Phylica paniculata</i> Willd.	South Africa (Cape Province)	JER162	silica gel
<i>Phylica paniculata</i> Willd.	South Africa (Cape Province)	FMW950	silica gel
<i>Phylica paniculata</i> Willd.	South Africa (Transvaal)	MvdB1	silica gel
<i>Phylica paniculata</i> Willd.	South Africa (Cape Province)	CFR136	1975
<i>Phylica spicata</i> L. f.	South Africa (Cape Province)	JER46	silica gel
<i>Phylica nitida</i> Lam.	Réunion	Thébaud <i>s.n.</i>	silica gel



<i>Phylica nitida</i> Lam.	Mauritius	Soorer 64-5	1964
<i>Phylica polifolia</i> (Vahl) Pillans (ITS)	St Helena	Chase 1751 (K)	silica gel
<i>Phylica polifolia</i> (Vahl) Pillans ( <i>trnL-F</i> )	St Helena	Chase 2269 (K)	silica gel
<i>Phylica pubescens</i> Ait.	South Africa (Cape Province)	Chase 859 (K)	silica gel
<i>Phylica thodei</i> Phill.	South Africa (Natal)	Hilliard & Burt 15379	1982
<i>Phylica tropica</i> Baker	Malawi	Brass 16739 (NYBG)	1946
<i>Pomaderris rugosa</i> Cheeseman	Australia	Chase 857 (K)	silica gel
<i>Siegfriedia darwinioides</i> C.A. Gardner	Australia	Chase 2181 (K)	silica gel
<i>Spyridium globulosum</i> (Labill.) Benth.	Australia	Chase 2021 (K)	silica gel
<i>Trymalium ledifolium</i> Fenzl	Australia	Chase 2184 (K)	silica gel

## 4.4. Results

### 4.4.1. *trnL*-F analysis

The data matrix had 107 variable characters and 69 potentially informative characters out of a total of 968 characters, i.e. 7% of characters were variable in two or more taxa. The initial 1000 replicate search produced 6910 trees of length 220. These trees were then swapped on until 7000 trees of length 220 were collected. These trees had CI=0.87 and RI=0.89. Successive weighting (SW) produced 170 trees of length 172.73 and with CI=0.97 and RI=0.98. The Fitch length of this tree was 220 (i.e. they were a subset of the Fitch trees). Figure 4.6 shows one of the trees with its Fitch branch lengths (ACCTRAN optimisation) above the branches and Fitch bootstrap percentages below; branches collapsing in the strict consensus tree of the Fitch analysis are marked with a solid arrow and those not present in the strict consensus of the SW trees are marked by an open arrow.

In the *trnL*-F analysis Phylliceae are a strongly supported monophyletic group, but the genus *Phyllica* is paraphyletic with *Nesiota elliptica* sister to *P. stipularis*. If *P. stipularis* is excluded from *Phyllica*, the other species in the genus form a strongly supported monophyletic group. In the tree shown *P. paniculata*, *P. arborea*, *P. polifolia*, *P. tropica*, *P. natalensis* and *P. emirnensis* form a group derived from the mainland, although this relationship breaks down in strict consensus trees and there is less than 50% bootstrap support. The two individuals of *P. nitida* from Mauritius and Réunion form a strongly supported sister group to the rest of this group, within which *P. emirnensis*, *P. natalensis* and *P. tropica* form a weakly supported monophyletic group. *Phyllica thodei* from eastern South Africa is also a member of the 'island group' in some of the shortest trees. The degree of resolution between the other species in this group, *P. paniculata*, *P. polifolia* and *P. arborea* is poor, due to low levels of divergence, and relationships between these species should be considered unresolved. In some trees, *P. paniculata* is paraphyletic.

#### 4.4.2. ITS analysis

The data matrix had 353 variable characters and 210 potentially informative characters out of a total of 821 characters, i.e. 26% of characters were variable in two or more taxa. The trees from the initial 1000 replicate search were swapped on to completion to produce 18 trees of length 732, CI=0.66 and RI=0.76. SW produced three trees with length 359.82 and with CI=0.87 and RI=0.92. The Fitch length of this tree was 732 (i.e. they were a subset of the Fitch trees). Figure 4.7 shows one of the SW trees with its Fitch branch lengths (ACCTRAN optimisation) above the branches and Fitch bootstrap percentages below; branches collapsing in the strict consensus tree of the Fitch analysis are marked with a solid arrow and those not present in the strict consensus of the SW trees are marked by an open arrow.

In the ITS analysis *Phylliceae* are a strongly supported monophyletic group, but again the genus *Phyllica* is paraphyletic with *Nesiota elliptica* sister to *P. stipularis*. The remainder of the *Phyllica* species form a strongly supported monophyletic group. The ITS data set did not include *P. emirnensis*, *P. natalensis* or *P. tropica* because the ITS region could not be sequenced from DNA of herbarium specimens of these species. However, some more island individuals including *P. arborea* from Nightingale and New Amsterdam and another *P. polifolia* from St Helena were added. Apart from the differences in the taxa included, the ITS topology was nearly identical to the *trnL-F* topology. *Phyllica nitida*, *P. paniculata*, *P. arborea* and *P. polifolia* form a strongly supported monophyletic group derived from within the mainland species with *P. thodei* from Natal as sister. The two individuals of *P. nitida* from Mauritius and Réunion form a strongly supported distinct sister group to the rest of this group. The degree of resolution between *P. paniculata*, *P. polifolia* and *P. arborea* is poor, due to low levels of divergence, and assumptions about relationships between these species should be treated with some caution. However, the results indicate that *P. paniculata* could be paraphyletic to both *P. arborea* and *P. polifolia*.

#### 4.4.3. Combined analysis

Given that taxa in each of the separate analyses are nearly perfectly congruent, this justifies the direct combination of the two data sets. The trees from the initial 1000 replicate search were swapped to completion producing 6 trees with length 916, CI=0.72 and RI=0.76. SW produced 3 trees with length 513.27, CI=0.92 and RI=0.93. The Fitch length of this tree was 916 (i.e. they were a subset of the Fitch trees). Figure 4.8 shows one of the SW trees with its Fitch branch lengths (ACCTRAN optimisation) above the branches and Fitch bootstrap percentages below; branches collapsing in the strict consensus tree of the Fitch analysis are marked with a solid arrow, and those not present in the strict consensus of the SW trees are marked by an open arrow.

*Phylica* is again paraphyletic with *Nesiota elliptica* nested as sister to *P. stipularis*. The remainder of the *Phylica* species form a strongly supported monophyletic group. *Phylica thodei* is sister to the 'island group' which form a strongly supported monophyletic group together with the most widespread South African species, *P. paniculata*. This group is derived from within the other Cape species of *Phylica*. The Mascarene species *P. nitida* forms the sister group to the rest of the 'island group'. The rest of this island group will be referred to as the 'paniculata group'. *Phylica paniculata* is again paraphyletic, but the degree of sequence divergence is not great enough to adequately address differences between these species.

### **4.5. Discussion**

#### 4.5.1. Origin and paraphyly of *Phylica*

These analyses show that *Phylica* clearly originated on the African mainland and not on any of the islands because the island taxa form a well supported group derived from deeply within the mainland species. For *Phylica* to be monophyletic, either *Nesiota* should be placed in *Phylica*, or *P. stipularis* should be placed in a separate genus. *Trichocephalus* could be resurrected for *P. stipularis* which had formerly been

placed in *Trichocephalus* along with a few other species (Brongniart and Dumas, 1827). The latter option is considered to be the most reasonable because both *Nesiota elliptica* and *P. stipularis* have a number of morphological differences justifying their treatment as separate genera. *Nesiota elliptica* has broad leaves in comparison to other genera in the tribe, and the leaves are opposite with stipules (Figure 4.5). *Phylica stipularis* also has stipules, but its leaves are narrow and alternate, and it has a unique floral feature in that there is pubescence on the ovary and disc. All other *Phylica* species are exstipulate and have narrow, alternately arranged leaves. That *Nesiota* is a palaeoendemic is supported by its long branch in the molecular trees, putatively plesiomorphic morphological characteristics (Figure 4.5) and distribution on St Helena. *Noltea africana* grows along riversides or streams or is found in southern temperate rainforest. It has attributes which may be regarded as primitive within both the tribe Phyliceae and Rhamnaceae, i.e. arborescent habit, broad leaves, paniculate inflorescence and cyathiform calyx-tube (Figure 4.4), which are all plesiomorphic characteristics. Its position in the trees and the degree of molecular and morphological divergence from its closest relative also indicate that it is a relict taxon. Other than the molecular phylogeny and the presence of stipules, there is little else to indicate that *P. stipularis* is a taxonomic relict. It shares distributions and habitats with other species of *Phylica*, but because its position as sister to *Nesiota* is well supported, their traits must have been derived independently. Only the plesiomorphic presence of stipules marks its isolated phylogenetic position. *Phylica stipularis* grows in the western Cape, and thus it has undergone the same selection as most of the other fynbos species. Only species growing in wetter sites with more neutral soils are able to retain other plesiomorphic traits. Parallel specialisation occurs in several lineages within *Phylica* proper (see below), and this can result in both phylogenetically derived species retaining plesiomorphic traits and parallel modifications occurring in their close relatives that experienced the changing climate of the fynbos over the last six mya.

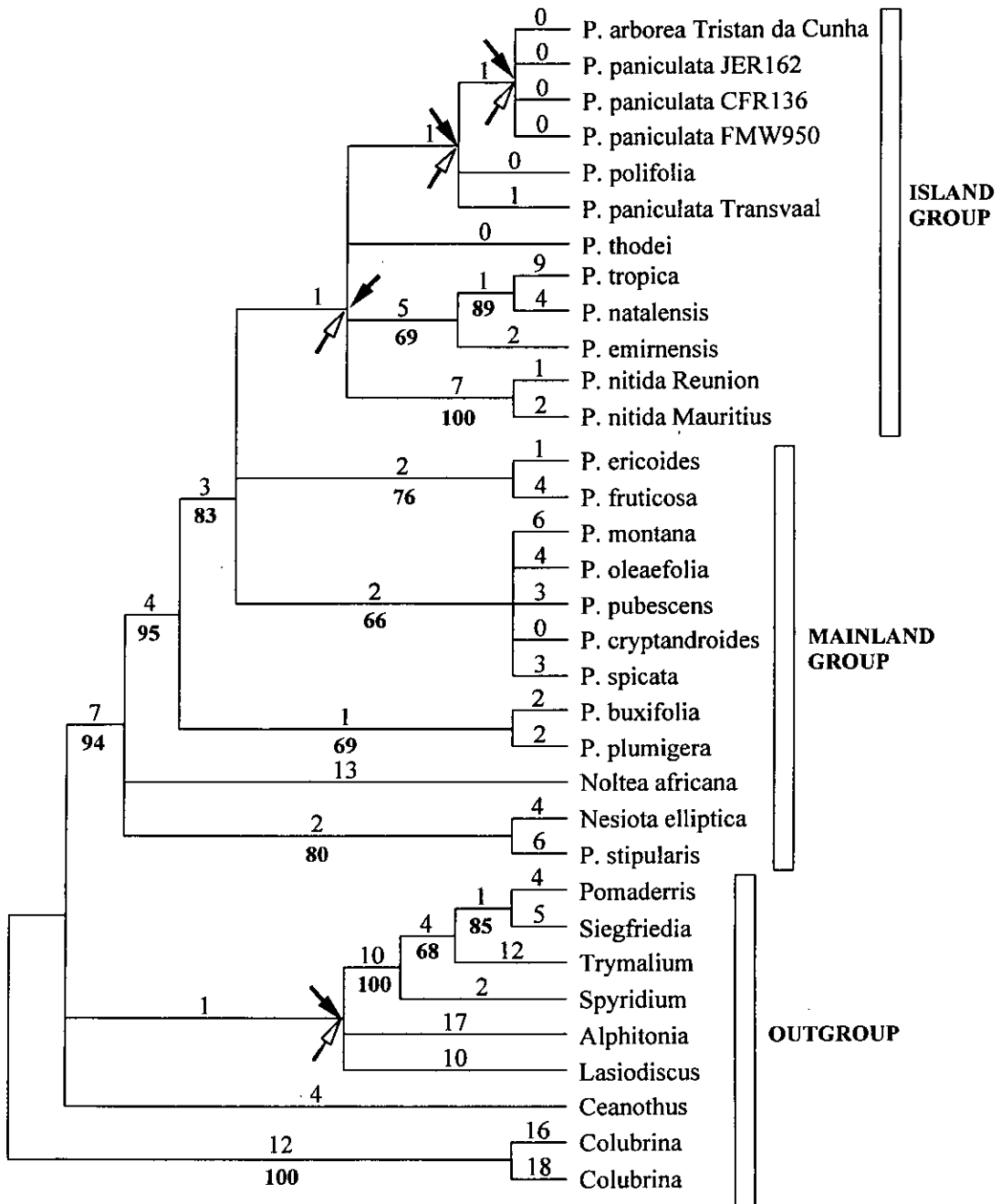


Figure 4.6. One of 190 optimal SW trees from the *trnL-F* analysis, with Fitch lengths (above branches; ACCTRAN optimisation) and bootstrap values (below). Branches not present in the Fitch strict consensus tree are indicated by a solid arrow, and those not present in the SW strict consensus tree are indicated by an open arrow. Heuristic search under the Fitch criterion produced 7000 trees with length 220, CI=0.87 and RI=0.89. SW produced 170 trees with length 172.73, CI=0.97 and RI=0.98 (Fitch length, 220).

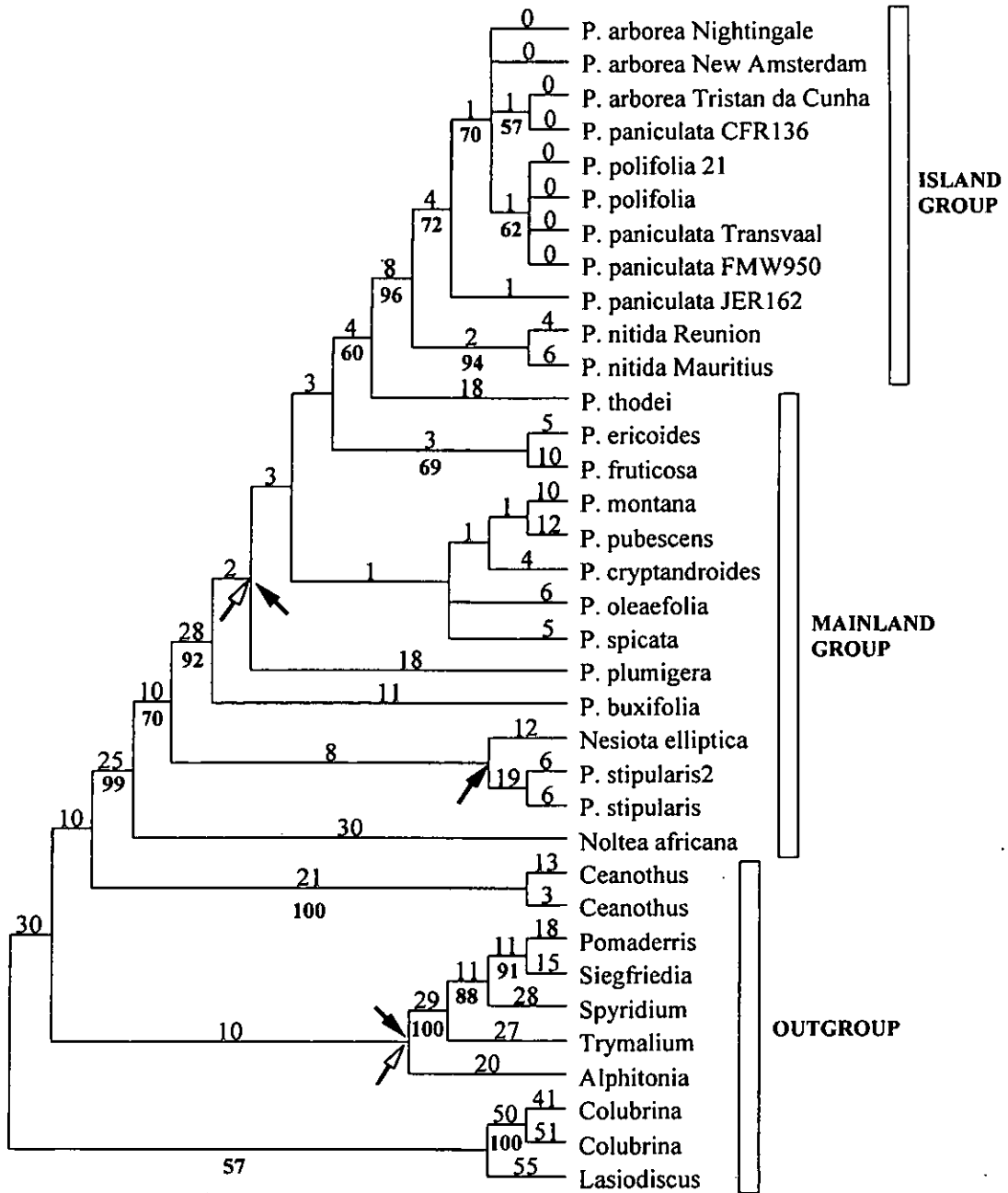


Figure 4.7. One of 3 optimal SW trees from the ITS analysis, with Fitch lengths (above branches; ACCTRAN optimisation) and bootstrap values (below). Branches not present in the Fitch strict consensus tree are indicated by a solid arrow, and those not present in the SW strict consensus tree are indicated by an open arrow. Heuristic search under the Fitch criterion produced 18 trees with length 732, CI=0.66 and RI=0.76. SW produced three trees with length 359.82, CI=0.87 and RI=0.92 (Fitch length, 732).

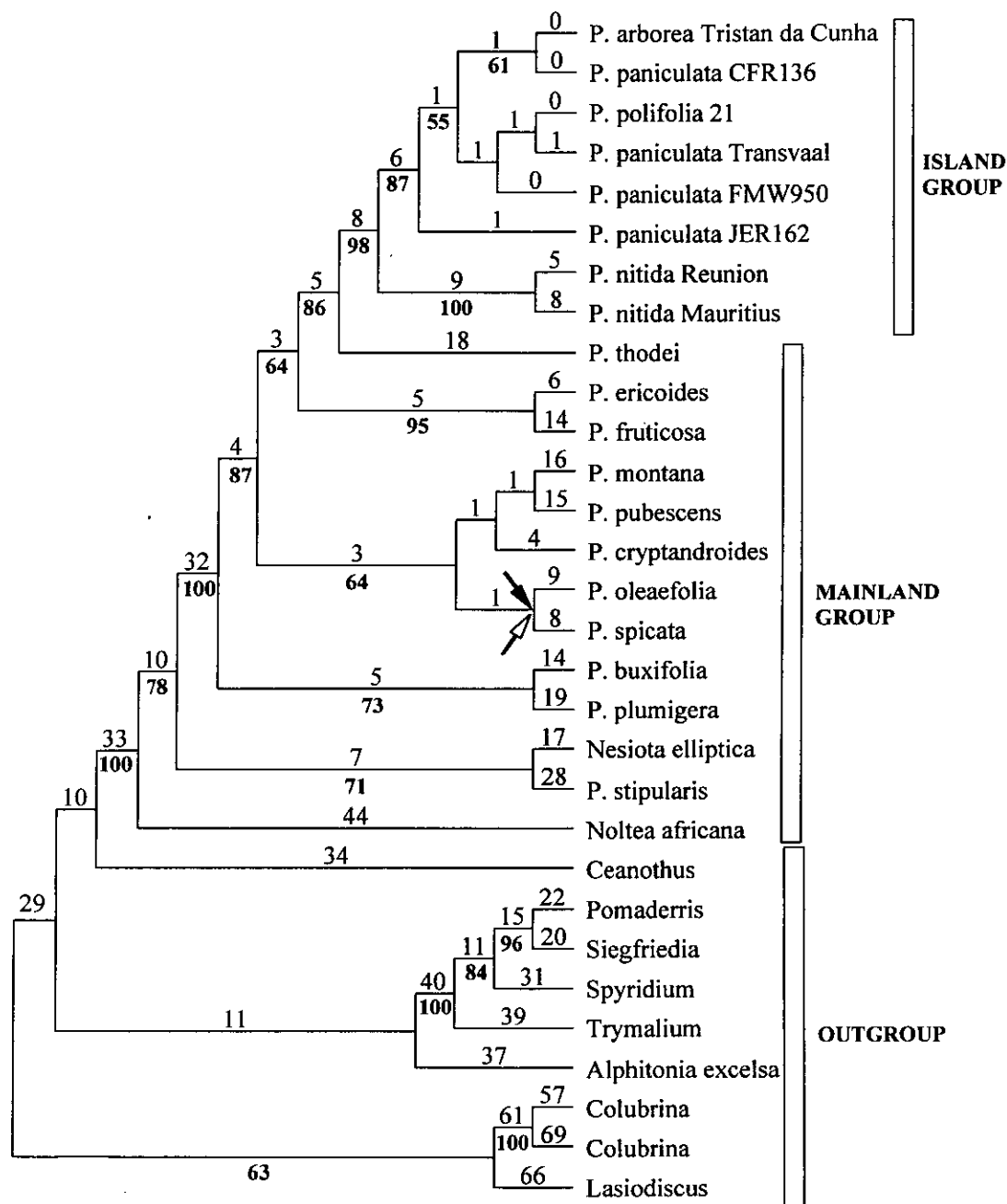


Figure 4.8. One of 3 optimal SW trees from the combined ITS/*trnL*-F analysis, with Fitch lengths (above branches; ACCTRAN optimisation) and bootstrap values (below). Branches not present in the Fitch strict consensus tree are indicated by a solid arrow, and those not present in the SW strict consensus tree are indicated by an open arrow. Heuristic search under the Fitch criterion produced six trees with length 916, CI=0.72 and RI=0.76. SW produced three trees with length 513.27, CI=0.92 and RI=0.93 (Fitch length, 916).



#### 4.5.2. The 'Island Group' and the origins of the island species

All of the island species of *Phylica* form a clade (the 'island group'), which also includes the widespread mainland species *P. paniculata* along with *P. tropica* from Malawi and *P. natalensis* from Natal. In the combined analysis, this group has strong bootstrap support. The 'island group' can itself be split into two clades consisting of (i) Mascarene *P. nitida* and (ii) the 'paniculata group' including *P. paniculata*, the western island species *P. arborea* and *P. polifolia*, and probably *P. emirnensis* from Madagascar, *P. tropica* and *P. natalensis*. A clade containing *P. tropica*, *P. emirnensis* and *P. natalensis* forms a well supported group within the 'paniculata group' in the *trnL-F* analysis, but the production of ITS sequences is necessary to verify their inclusion within this group because in some *trnL-F* trees *P. thodei* is also included in the 'island group', whereas in the ITS and combined analyses it is excluded. All other *Phylica* groupings contain almost exclusively Cape species. The degree of variation found between the species was not high enough to resolve all species within the 'island group' and does not resolve relationships between all the Cape Province species (Weitz and Richardson, unpubl.). The sequence data however provide enough information to distinguish between infra-generic groups of species.

On the basis of the molecular trees, *P. arborea*, *P. polifolia*, *P. nitida* and *P. emirnensis* appear to be relatively recently derived or neoendemic within the context of the genus with the nearest mainland relative being *P. paniculata*. The Mascarene species *P. nitida* may be regarded as palaeoendemic within the 'island group'. *Phylica paniculata* grows almost exclusively alongside streams or in montane regions. These habitats along with moister oceanic islands are possible local refugia for relict taxa. However, it is more likely that this species and the island species occupy their present distributions because of more recent dispersal, perhaps escaping from some sort of refugial site. The lack of sequence divergence between the western island species and *P. paniculata* and their position within the phylogenetic tree indicates that the former have been relatively recently derived. In other words, island species such as *P. polifolia* and *P. arborea* are not palaeoendemic as was suspected from their seemingly plesiomorphic morphological characteristics, but rather they result from a recent long-distance dispersal of a derivative lineage that has retained

plesiomorphic floral and vegetative traits. The low level of *trnL-F* sequence divergence of *P. emirnensis* (Madagascar) and its phylogenetic position relative to *P. paniculata*, *P. arborea* and *P. polifolia* indicates that its development was relatively recent and that it too is a product of a recent dispersal event. This can be contrasted with other Madagascan taxa, such as *Bathiorhamnus*, (Chapter Two) that are relict taxa with high levels of sequence and morphological divergence from their closest relatives.

In assessing the relative positions and level of sequence divergence of *P. nitida* and the 'paniculata group' in the phylogeny, it could be hypothesized that *P. nitida* (or its ancestor) split off from the same ancestral stock as the 'paniculata group' by dispersing to the Mascarenes and diverging. Meanwhile the progenitor of the 'paniculata group' stayed in refugia and retained primitive characteristics and may also have dispersed to other montane regions. The eastern African and Madagascan species seem to have been derived from within the ancestral stock of the 'paniculata group' a little later than the Mascarene species on the basis of their molecular divergence and phylogenetic position. Dispersal to the Tristan da Cunha Group or St Helena from within the 'paniculata group' was even more recent.

All analyses indicate that *P. paniculata* is potentially paraphyletic, but there is no clear evidence that the island species were not derived from an ancestor in common with *P. paniculata*. The paraphyly of *P. paniculata* could be an artefact of low levels of divergence and lineage sorting of ancestral polymorphisms after divergence. Alleles and plastid cytotypes may diversify within a population prior to dispersal, and organismal histories and gene histories can be partly independent. Species trees, which estimate the history of diversification of a group of organisms, should be distinguished from gene trees, which represent the history of the molecular diversification within that organismal tree. Determination of the monophyly of the island taxa and the number of species on the islands could not be properly established with the number of samples that were studied, and the degree of sequence divergence exhibited by ITS and *trnL-F* is too limited to make robust conclusions. Resolution of relationships among these species requires additional sampling as well as data from other more polymorphic sources of information. However, *P. arborea* on the Tristan da Cunha Group and New Amsterdam may be distinguished from *P.*

*paniculata* by its thyriform, rounded or oblong inflorescence which is densely tomentose compared to the more variable inflorescence of *P. paniculata* which has flowers in short spikes assembled in panicles, or in pedunculate or sessile, capituliform spikes arising in the axils of upper leaves, assembled in panicles or solitary at the ends of branchlets. This indicates that *P. arborea* may be a monophyletic derivative of the more variable *P. paniculata*.

#### 4.5.3. Biogeographic history of Phyliceae and its island species

The closest relatives of Phyliceae are the Pomaderriaceae (Australia), Colletiaceae (Australia and southern South America) and *Ceanothus* (western North America) and given the distribution and phylogenetic position (Chapter Two) of these groups I suggest that they each represent refugia for a larger group that was once much more widespread (i.e. these distributions are not the result of long distance dispersals).

A *Noltea*-like ancestor could have been more widespread throughout the Cape region about 25 mya in the area now covered by fynbos when the vegetation consisted of warm, wet sub-tropical forests. Plant groups which made up this flora either no longer occur in Africa or exist in refugia such as in the coastal forest vegetation of the southern Cape where *Noltea* is presently found. *Noltea* has evolved 30 autapomorphies since it diverged from its nearest relative, and this indicates a divergence of 15 mya, i.e. some 2 mya before extensive aridification of southern Africa began. *Noltea* is a tree with none of the adaptations to the dry climate of *Phylica*, such as ericoid habit and revolute leaves, and it has a paniculate inflorescence with a cyathiform calyx-tube (Figure 4.4), which are primitive features within the tribe and within Rhamnaceae. In the molecular trees, it is the sister to the rest of the tribe Phyliceae. It grows predominantly along riverbanks, i.e. in a mesic environment that could be considered refugial.

*Phylica* was most likely once more widely distributed throughout continental Africa, and subsequently its distribution was restricted, but many of the plants which characterise fynbos vegetation did not appear in the fossil record until the Pliocene, indicating a later dispersal/development of the species to give *Phylica* its present widespread distribution. Members of the 'island group' are found almost exclusively

on oceanic islands or on 'islands' of the Afromontane archipelago as described by White (1983). This group could also include *P. thodei* from the Drakensberg. The levels of molecular divergence and phylogenetic positions of these widespread taxa indicate recent dispersal to these 'islands'.

Nightingale Island in the Tristan da Cunha Group was formed 18 +/- 4 mya (Wace and Holdgate, 1958) and this is therefore the earliest possible time for dispersal of *Phyllica* to the Tristan da Cunha Group. However, the degree of sequence divergence between *P. arborea* and its closest mainland relative indicates a more recent dispersal (c. 0.5 mya assuming the molecular clock). Sixteen mya the growth of the Antarctic ice sheet increased as Antarctica was finally separated from South America allowing the development of the cold Circum-Antarctic (Benguela) current (Siesser, 1980). The development of the Benguela current resulted in a cold ocean along the west coast of Africa which speeded up the aridification of SW Africa. The increasingly arid conditions may have begun to force *Noltea* or its ancestor into the refugia of the temperate rainforest in which it is now found. St Helena was formed 14.3 mya and this is the earliest possible time for the dispersal of *N. elliptica* or *P. polifolia* onto the island. The degree of molecular and morphological divergence between *N. elliptica* and other species of *Phyllica* strongly supports the hypothesis that *N. elliptica* arrived on the island long before *P. polifolia* which has low levels of molecular divergence from its closest mainland relative. The fact that *Nesiota* had developed some of the features which characterise *Phyllica* and other plants that grow in arid environments indicates that it may have dispersed after the development of this type of climate in southwestern Africa, i.e. from c. 13 mya. *Nesiota elliptica* has 12 ITS autapomorphies, and assuming the molecular clock this indicates that it diverged from its mainland ancestor 6 mya, i.e. around the time when some morphological adaptations to an arid climate might have occurred.

Mauritius was formed 7.8 mya, and this represents the earliest possible time for dispersal of *P. nitida*. Dispersal to the Mascarenes is most likely to have occurred with an initial dispersal to Mauritius followed by dispersal to Réunion, although it is possible that there may have been an earlier dispersal to the older island of Rodrigues followed by extinction there. Since diverging from its mainland progenitor, *P. nitida* developed two synapomorphies and *P. nitida* on Mauritius developed a further six

autapomorphies giving a total of eight, indicating that assuming the molecular clock dispersal to the Mascarenes occurred *c.* four mya.

The birth of fynbos vegetation began properly six mya (Scott, 1995), and this is the time when fynbos plants such as *Phylica* would have increased the development of adaptations to the dry climate. At this time the coastal lowlands were covered with open shrubland dominated by grasses, restios, geophytes, and composites (Coetzee *et al.*, 1983). The inland plains consisted of grassy woodlands which included many fynbos elements including proteas, ericas and other ericoid shrubs such as *Phylica* (Coetzee *et al.*, 1983). Forest vegetation was becoming restricted near coasts on sand dunes and along riverbanks. Five mya fynbos forms increased, the forests declined, and the first evidence of widespread fire was documented. Four mya saw the inception of a Mediterranean climate with dry summers, rain bearing westerly winds in winter and dry south easterly winds in summer. Fires caused by lightning strikes became increasingly important in the ecology of the region.

The fact that *P. emirnensis*, *P. tropica* and *P. natalensis* are possibly derived from within the 'island group' indicates that the present distributions of *P. emirnensis* and *P. tropica* could only be the result of recent long-distance dispersal events rather than ancient vicariance. Axelrod and Raven's (1978) suggestion that genera shared by the Cape and Madagascar reached Madagascar by long distance dispersal is therefore supported in this case. A continuously favourable habitat between mountains of the Cape region and those of Madagascar at any time was thought to be unlikely, and presumably this finding can be applied to links between the Cape mountains and montane regions of Malawi, Zimbabwe and Mozambique where *P. tropica* is found. Although *P. emirnensis*, *P. tropica* and *P. natalensis* are morphologically similar to the other island species, they have developed more adaptations to an arid climate with decreased size and increased inrolling of leaves and possession of inflorescences which are slightly more advanced. These species were probably derived from a *P. paniculata*-like progenitor, but they have since adapted to the drier habitats to which they dispersed.

The low level of sequence divergence and the phylogenetic position of *P. polifolia* indicates that this species is a recent introduction onto St Helena. Cronk (1987) stated that St Helenan neoendemics from southern Africa arrived on the

southeast trade wind and/or the southeast Benguela current. *Phylica arborea* is also undoubtedly a recent introduction to the Tristan da Cunha Group. A survey of other species on this group of islands is necessary to ascertain whether the species with South American affinities are of a more ancient origin. *Phylica arborea* is also a recent introduction to New Amsterdam because its sequences are nearly identical with those of *P. arborea* on the Tristan da Cunha Group. The sequence data do not permit determination of whether this species arrived on the Tristan da Cunha Group first and then dispersed to New Amsterdam or vice versa. New Amsterdam Island was formed 0.69 mya, and Tristan da Cunha was formed 1 mya so the dispersal of *P. arborea* from the Tristan da Cunha Group to New Amsterdam or vice versa is certainly a recent dispersal as indicated by the lack of sequence divergence between individuals on these islands. There are only one or two substitutions between *P. paniculata*, *P. arborea* and *P. polifolia*, which assuming the molecular clock indicates that each of these taxa diverged from a common ancestor between 1-0.5 mya.

Dispersal times for each of the island species are summarised in Figure 4.9. In the case of *P. arborea*, the question of direction of movement between islands has not been resolved, but movement from the older Tristan da Cunha Group to the much younger New Amsterdam could be hypothesised as more reasonable.

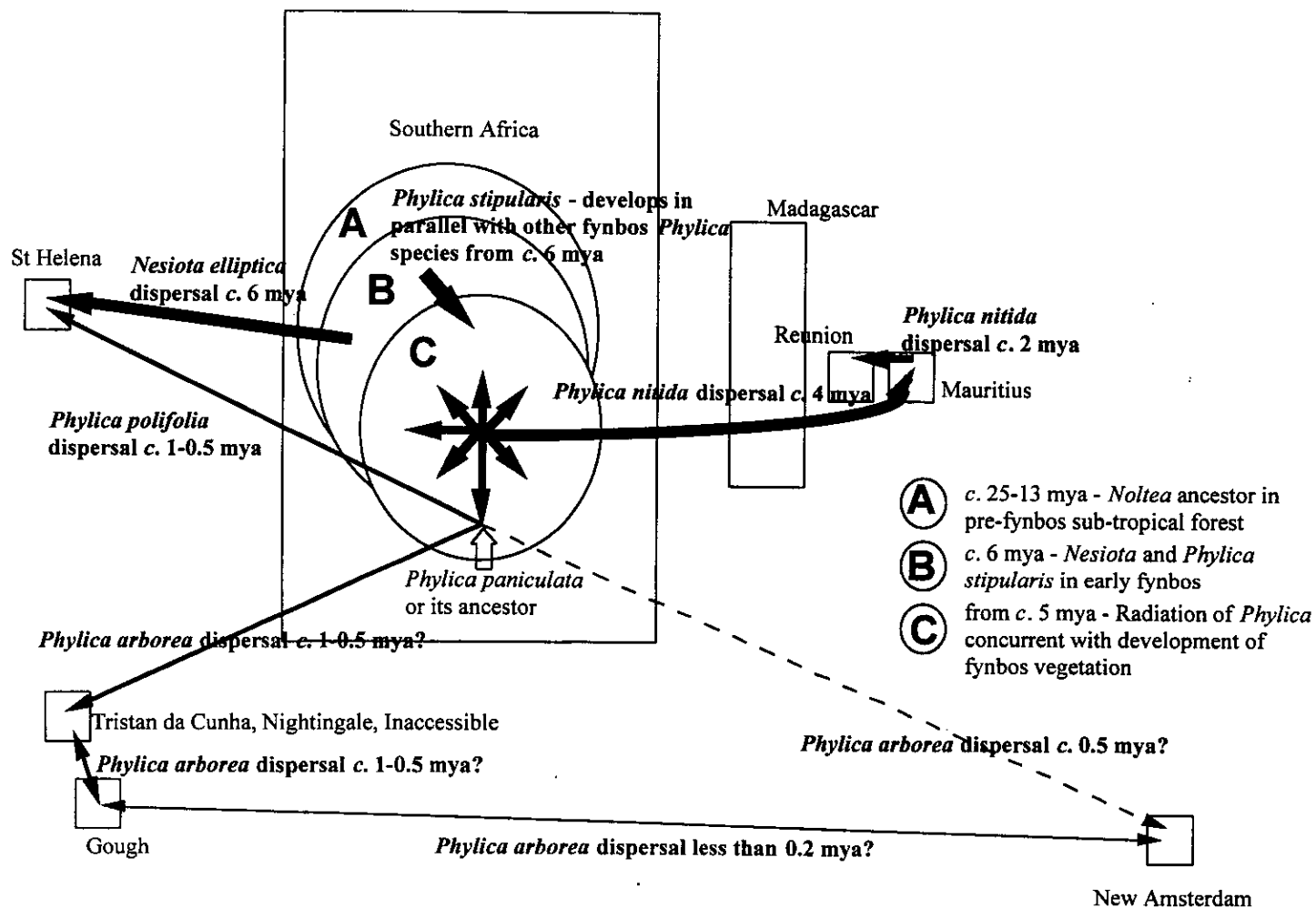


Figure 4.9. Hypothetical biogeographical development of the tribe Phyliceae based on nucleotide sequence data.

#### 4.5.4. Comparative evolution of the island and mainland species of *Phylica*

The 'island group' has retained primitive morphological characteristics whereas most other groups in the western Cape have developed advanced characteristics, such as ericoid habit, increased inrolling of leaves, capituliform inflorescence and elongation of the calyx-tube. Many of the neoendemic species have highly restricted distributions and are associated with particular soil types. The majority of the mainland species have adapted to specific and localised pollinators and are therefore more tied to local conditions and unlikely to be reproductively successful if they disperse. The widespread distribution of the 'island group' is probably due to the fact that morphologically (and ecologically) they are generalists. The floral morphology of these species is of the basic rhamnaceous type with a cyathiform calyx tube. This means that pollination by many wider-ranging species, many of them generalists, is possible, and consequently wherever these species disperse they are more likely to be reproductively successful. The flatter leaves of the island group may also mean that they are more likely to survive wetter conditions than those species which are more highly adapted to the extreme, dry conditions of the Cape. The fact that the members of the island group grow on volcanic soils also shows a tolerance of a wider range of substrates than many mainland species. Pillans grouped several mainland species such as *P. buxifolia* and *P. oleaefolia* that also retained primitive morphological characteristics (as has *P. paniculata*) with the island species, but these species are unrelated to the island group according to the molecular data. These species represent additional evidence for the parallel retention of plesiomorphic morphology.

A more in-depth study of molecular variation in *Phylica paniculata* could answer questions such as whether there were a series of founder events to all other points of its distribution and oceanic islands, but the evidence of which populations were involved in such dispersals could have been erased by a long period of continued interbreeding within *P. paniculata* after these dispersal events. The distribution of *P. paniculata* makes it part of a southern centre of distribution that has been described as one of the five montane centres in Africa. It is likely a peripheral species, which indicates the geographic range of *Phylica*/Phyliceae under different (wetter) climatic conditions, whereas the bulk of the species in the ancestral distribution of the genus



have become more specialised in response to climate change and the evolution of pollinators that are restricted to particular habitats within different regions in the Cape. In such a scenario, multiple lineages more or less simultaneously would have become similarly specialised, so that a more complete sampling of *Phylica* would exhibit species clusters within which the composite species would represent parallel series from different geographic zones (such sampling was not possible in this study). Generalists are good dispersers and not necessarily representative of 'old lineages' within the context of their close relatives. It is possible that they represent reversals, but multiple reversals to plesiomorphic floral structure, habit and habitats seems less likely than multiple specialisations within other lineages in response to the drastic changes in climate and geography that have occurred in the Cape. The progenitor of the 'island group' could have been restricted to more mesic environments at an early stage in the development of *Phylica*. It therefore would not have to have developed adaptations to the dry climate in the way that other lineages have and could have retained a more plesiomorphic vegetative form (this is also the case in other lineages where species have retained plesiomorphic features, e.g. *P. buxifolia*, *P. oleaefolia*). At the time of divergence, all other lineages may have had similar vegetative forms but instead of moving into a more mesic environment they remained in increasingly dry areas in which they were forced to adapt.

#### 4.6. Conclusions

The combined nuclear ITS and plastid *trnL-F* analyses indicate that *Nesiota* and *Noltea* are palaeoendemic genera within the context of *Phyliceae*. *Phylica* originated on the African mainland rather than from any of the islands on which species are presently found although the morphological data was potentially compatible with the hypothesis that *Phylica* originated on the islands because the island species have plesiomorphic morphological features. Within the context of the 'island group' *P. nitida* from the Mascarenes is a palaeoendemic species. The low levels of molecular divergence of the other island species in comparison to their nearest mainland relative (*P. paniculata*) indicates that these species are recently derived, neoendemic species within the context of the genus. The molecular data indicate patterns that

have been masked by rapid morphological divergence and radiation on the mainland. Questions regarding the number of species on the islands, whether these species are monophyletic and the sequence of colonisation of island species or populations could not be answered using ITS or *trnL-F* sequence data due to low levels of sequence divergence, and those must be addressed by using more polymorphic markers, such as DNA fingerprinting.

#### 4.7. Bibliography

Axelrod, D.I. & P.H. Raven. 1978. Late Cretaceous and Tertiary vegetation history of Africa. In M.J.A. Werger [eds.], *Biology and Ecology of Southern Africa*, 77-130. Junk, The Hague.

Baker, I., N.H. Gayle & J. Simons. 1967. The geochronology of the St Helena volcanoes. *Nature* 215: 1451-1456.

Bakker, F.T., D. Hellbrügge, A. Culham, & M. Gibby. 1998. Phylogenetic relationships within *Pelargonium* sect. *Peristera* (Geraniaceae) inferred from nrDNA and cpDNA sequence comparisons. *Plant Systematics and Evolution* 211: 273-286.

Baldwin, B.G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. *Molecular Phylogenetics and Evolution* 1: 3-16.

Baldwin, B.G. 1993. Molecular phylogenetics of *Calycadenia* (Compositae) based on ITS sequences of nuclear ribosomal DNA: chromosomal and morphological evolution reexamined. *American Journal of Botany* 80: 222-238.

Baldwin, B.G., M.S. Sanderson, J.M. Porter, M.F. Wojciechowski, C.S. Campbell & M.S. Donoghue. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden* 82: 247-277.

Boast, J. & A.E.M. Nairn. 1982. An Outline of the Geology of Madagascar. *In* A.E.M. Nairn & F.G. Stehli [eds.], *The Ocean Basins and Margins*, volume 6, The Indian Ocean, 649-696. Plenum Press, New York & London.

Brenon, P. 1972. The Geology of Madagascar. *In* R. Battistini & G. Richard-Vindard [eds.], *Biogeography and Ecology in Madagascar*, 27-86. Junk, The Hague.

Brongniart, A.T. & Dumas. 1827. Mémoire sur la famille des Rhamnées, *Annales des Sciences Naturelles*, Séries 1, 10: 320-386.

Cadet, L.J.T. 1977. La végétation de l'île de la Réunion: Etude phytoécologique et phytosociologique. PhD Thesis, University of Aix-Marseille III, Marseille.

Christopherson, E. 1937. Plants of Tristan da Cunha, Scientific results of the Norwegian Antarctic expeditions 1927-28. I Kommissjon hos Jacob Dybwad, Oslo.

Coetzee, J.A. 1978. Climate and biological changes in south-western Africa during the Late Cainozoic. *Palaeoecology of Africa* 10: 13-29.

Coetzee, J.A. 1978. Late Cainozoic palaeoenvironments of southern Africa. *In* E.M. Van Zinderen Bakker [ed.], *Antarctic Glacial History and World Palaeoenvironments*, 115-127. Balkema, Rotterdam.

Coetzee, J.A. & J. Rogers. 1982. Palynological and lithological evidence for the Miocene palaeoenvironment in the Saldanha region (South Africa). *Palaeogeography, Palaeoclimatology, Palaeoecology* 39: 71-85.

Coetzee, J.A., A. Scholtz & H.J. Deacon. 1983. Palynological studies and the vegetation history of the fynbos. *In* H.J. Deacon, Q.B. Hendey & J.J.N. Lambrechts [eds.], *Fynbos Palaeoecology. A Preliminary Synthesis*, 156-173. South African National Scientific Programmes Report 75, Pretoria, CSIR.

Coetzee, J.A. & J. Muller. 1984. The phytogeographic significance of some extinct Gondwanan pollen types from the Tertiary of the south western Cape (South Africa). *Annals of the Missouri Botanical Garden* 70: 1088-1099.

Cowling, R. & D. Richardson. 1995. *Fynbos, South Africa's Unique Floral Kingdom*. Fernwood Press, Vlaeberg.

Cronk, Q.C.B. 1987. The history of endemic flora of St Helena: a relictual series. *New Phytologist* 109: 509-520.

Cronk, Q.C.B. 1992. Relict floras of Atlantic islands: patterns assessed. *Botanical Journal of the Linnean Society* 46: 91-103.

Don, G. 1832. A general system of gardening and botany 2: 38-43. Gilbert & Rivington, London.

Doyle, J.J. & J.L. Doyle. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11-15.

Endlicher, S. 1836-1840. *Genera Plantarum secundum Ordines Naturales disposita*, 1094-1101. Universitatis Bibliopolam, Vindobonae.

Fay, M.F., C. Bayer, W.S. Alverson, A.Y. de Bruijn & M.W. Chase. 1998. Plastid *rbcL* sequence data indicate a close affinity between *Diegodendron* and *Bixa*. *Taxon* 47: 43-50.

Gielly, L. & P. Taberlet. 1996. A phylogeny of the European gentians inferred from chloroplast *trnL* (UAA) intron sequences. *Botanical Journal of the Linnean Society* 120: 57-75.

- Groves, E.W. 1981. Vascular plant collections from the Tristan da Cunha group of islands. *Bulletin of the British Museum (Natural History)* 8: 333-420.
- Guého, J. 1976. Sur l'identité du *Phyllica* (Rhamnaceae) des isles Mascareignes. *Adansonia* ser. 2, 15: 509-513.
- Hemsley, A.B. 1885. Report on the Scientific results of the H.M.S. Challenger 1(2): botany, insular floras, Atlantic and Indian Oceans. HM Government, London.
- Higgins, D.G., A.J. Bleasby & R. Fuchs. 1992. CLUSTAL: a new multiple sequence alignment program. *Computer Applications in Biosciences* 8: 189-191.
- Hooker, J.D. 1870. *Icones Plantarum* volume 11. Longman, Rees, Orme, Brown, Green & Longmans, London.
- Hopper, S.D., M.F. Fay, M. Rossetto, & M.W. Chase. (in press). A molecular phylogenetic analysis of the bloodroot and kangaroo paw family, Haemodoraceae: taxonomic, biogeographic and conservation implications. *Botanical Journal of the Linnean Society*.
- Kennet, J.P. 1980. Palaeoceanographic and biogeographic evolution of the southern ocean during the Cenozoic, and Cenozoic microfossil datums. *Palaeogeography, Palaeoclimatology, Palaeoecology* 31: 123-152.
- Leroy, J.-F. 1978. Composition, origin, and affinities of the Madagascan vascular flora. *Annals of the Missouri Botanical Garden* 65: 535-589.
- Linder, H.P. 1985. Gene flow, speciation, and species diversity patterns in a species rich area: the Cape Flora. *In Vrba, E.S. [ed.], Species and Speciation. Transvaal Museum Monograph* 4: 53-57.
- Linnaeus, C. 1753. *Species Plantarum*, Ed. 10. Stockholm.

Lledó, M. D., M.B. Crespo, A.V. Cox, M.F. Fay & M.W. Chase. 1998. Polyphyly of *Limoniastrum* (Plumbaginaceae): evidence from sequenced plastid *rbcL*, *trnL* intron and *trnL-F* intergene spacer. *Systematic Botany* 23: 21-29.

Myers A.A. & P.S. Giller [eds.]. 1988. *Analytical Biogeography, an integrated approach to the study of animal and plant distribution*. Chapman & Hall, London.

Partridge, T.C. 1997. Evolution of landscapes. *In* R.M. Cowling, D.M. Richardson & S.M. Pierce [eds.], *Vegetation of Southern Africa*. Cambridge University Press, Cambridge.

Pillans, N.S. 1942. The genus *Phylica* Linn. *Journal of South African Botany* 8: 1-164.

Reichenbach, H.G.L. 1828. *Conspectus Regni Vegetabilis*, 145. Hofmeister, Leipzig.

Reeves, G., M.W. Chase, T. de Chies, A.V. Cox, P. Goldblatt, B. Lejeune, P.J. Rudall & M.F. Fay. 1997. Molecular systematics of Iridaceae: a combined analysis of three plastid DNA sequence matrices. *American Journal of Botany* 84: 656 [abstract].

Richardson, J.E., M.F. Fay, Q.C.B. Cronk, D. Bowman & M.W. Chase. (submitted). A molecular analysis of Rhamnaceae using plastid *rbcL* and *trnL-F* sequences. *American Journal of Botany*.

Sang, T., D.J. Crawford, S.-C. Kim & T.F. Stuessy. 1995. Radiation of the endemic genus *Dendroseris* (Asteraceae) on the Juan Fernandez Islands: evidence from sequences of the ITS region of nuclear ribosomal DNA. *American Journal of Botany* 81: 1494-1501.

- Schatz, G.E. 1996. Malagasy/Indo-Australo-Malesian phytogeographic connections. *In* W.R. Lourenço [ed.], *Biogeography of Madagascar*. Editions ORSTOM, Paris.
- Scholtz, A. 1985. Palynology of the Upper Cretaceous lacustrine sediments of the Arnot pipe, Banke, Namaqualand. *Annals of the South African Museum* 95: 1-109.
- Scott, L. & R. Bonnefille. 1986. A search for pollen from the hominid deposits of Kromdraai, Sterkfontein and Swartkrans: some problems and preliminary results. *South African Journal of Science* 82: 380-382.
- Scott, L. 1995. Pollen evidence for vegetation and climate change in Southern Africa during the Neogene and Quaternary. *In* E.S. Vrba, G.H. Denton, T.C. Partridge & L.H. Burckle [eds.], *Palaeoclimate and Evolution with Emphasis on Human Origins*, 56-76. Yale University Press, New Haven.
- Shackleton, N.J. & J.P. Kennet. 1975. Paleotemperate history of the Cenozoic and the initiation of Antarctic glaciation: oxygen and carbon isotope analysis in DSDP sites 277, 279 and 281. *Initial Reports of the DSDP* 29: 743-755.
- Siesser, W.G. 1978. Aridification of the Namib Desert: evidence from ocean cores. *In* E.M. Van Zinderen Bakker [ed.], *Antarctic Glacial History and World Palaeoenvironments*, 115-127. Balkema, Rotterdam.
- Siesser, W.G. 1980. Late Miocene origin of the Benguela upwelling system off northern Namibia. *Science* 208: 283-285.
- Sim, T.R. 1907. *The forests and forest flora of the colony of the Cape of Good Hope*. Aberdeen.
- Smith A.G. 1994. *Atlas of Mesozoic and Cenozoic Coastlines*. Cambridge University Press, Cambridge.

Storey, M., J.J. Mahoney, A.D. Saunders, R.A. Duncan, S.P. Kelley, M.F. Coffin. 1995. Timing of Hot Spot Related Volcanism and the Breakup of Madagascar and India. *Science* 267: 852-855.

Suessenguth, K. 1953. Rhamnaceae, Vitaceae, Leeaceae. *In* Engler, A. & K. Prantl [eds.], *Die Natürlichen Pflanzenfamilien* 2, Aufl., 20d. Duncker & Humblot, Berlin.

Swofford, D.L. 1998. PAUP: phylogenetic analysis using parsimony (and other methods), version 4.0. Sinaeur Associates, Sunderland, Massachusetts.

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

Van Zinderen Bakker, E.M. 1975. The origin and palaeoenvironment of the Namib Desert biome. *Journal of Biogeography* 2: 65-73.

Van Zinderen Bakker, E.M. & J. Muller. 1987. Pollen studies in the Namib Desert. *Pollen et Spores* 29: 185-205.

Vrba, E.S. 1985. Early hominids in southern Africa: updated observations on chronological and ecological background. *In* P.V. Tobias [ed.], *Hominid Evolution*, 195-200. Liss, New York.

White, F. 1983. *The Vegetation of Africa: a descriptive memoir to accompany the Unesco/AETFAT/UNSO vegetation map of Africa.*



**CHAPTER FIVE. RELATIONSHIPS OF  
ISLAND POPULATIONS OF *PHYLICA* L.  
BASED ON AMPLIFIED FRAGMENT  
LENGTH POLYMORPHISMS**

## CHAPTER FIVE. Relationships Of Island Populations Of *Phylica* L. Based On Amplified Fragment Length Polymorphisms

### Abstract

According to phylogenetic analysis of DNA sequences from the plastid *trnL-F* intron and inter-genic spacer and the nuclear internal transcribed spacer of the large ribosomal genes, the island species of the genus *Phylica* form a monophyletic group together with the South African species *P. paniculata* and *P. natalensis* and the eastern African species *P. tropica*. DNA fingerprints (AFLPs) revealed higher levels of polymorphism than the gene sequences which differed by only one or two substitutions. AFLPs were therefore used to elucidate relationships between the island species and *P. paniculata* from the mainland. Parsimony, neighbour joining, UPGMA and PCO analyses performed on the data set indicated that each of the island group species studied is distinct. AFLPs were useful in elucidating the genetic relationships and possible infra-specific origins of different island populations. *Phylica nitida* on Réunion is likely to have been derived from *P. nitida* on Mauritius. Although the sampling on New Amsterdam is not extensive, the data are also consistent with the hypothesis that *P. arborea* on New Amsterdam was derived from a single introduction of *P. arborea* from Gough Island. Similarly the Gough Island population appears to have been derived from one introduction, but it is so distinct from those on Tristan da Cunha, that there may have been two separate dispersals to Gough and Tristan/Nightingale from different lines of the mainland progenitor. There is also evidence of a reintroduction from Gough to Tristan da Cunha.

### 5.1. Introduction

The previous chapter established that island species of *Phylica* together with three species from mainland Africa form a monophyletic group. The degree of ITS and *trnL-F* sequence variation found between or within species was not high enough to allow a more thorough investigation of the relationships between these species so a study of more polymorphic markers was necessary. A fingerprinting technique, amplified fragment length polymorphism (AFLPs), was chosen to assess variability from individuals up to closely related species. This technique surveys more markers than other available techniques. For example, AFLPs give 10-100 times more markers than RAPDs, which therefore permits a finer scale assessment of levels of variation and distribution of genotypes. This study was undertaken with the primary

aim of resolving the relationships between the 'island group' species and assessing their degree of genotypic distinctness as well as investigating their possible origins. If we assume that genetic polymorphisms are mostly due to neutral sequence variation, characters that define a particular genotype will be maintained by lack of recombination rather than selection. The geographical distribution of genotypes should then reflect the history of colonisation of the geographic range. The distribution of populations of the same species or closely related species on island archipelagos presents a good opportunity for studying genotypic distinctness and levels of gene flow, and the *Phylica* 'island group' is thus a good model for studying genetic patterns involved in speciation or species differentiation. The great geographic distance between some of the island populations of *Phylica* could mean that the opportunity for gene flow between these islands is restricted and studies of the levels of genotypic differentiation between these populations could demonstrate that these populations are in the early stages of speciation through geographic isolation. The low level of DNA sequence divergence demonstrated in Chapter Four and the age of the islands involved limit the timing of some of these dispersal events to within the last million years. Levels of genotypic differentiation since dispersal can therefore be assessed with these time limits in mind. It might be expected that within species population genotypic structure on the same island could indicate panmixis and AFLPs should demonstrate this.

Successive introductions of island taxa onto progressively younger islands have been indicated in phylogenetic analyses using mitochondrial DNA sequences (Juan *et al.*, 1995; 1996a; 1996b; 1998). However, low levels of polymorphism in the plastid genome of plants limit their use in studies among closely related species. The use of AFLPs, which are predominantly nuclear markers, to determine possible successive introductions of *Phylica* onto progressively younger islands and the origin of island populations was therefore investigated. Section 5.3 gives background accounts of the vegetation on each of the islands, with particular reference to species of *Phylica* and is followed by a section detailing the demographic status of *Phylica* on each of the islands.

An example of the utilisation of DNA fingerprint data to study genotypic differentiation within and among species throughout a wide geographical range comes from Van Heusden and Bachmann (1992a,b,c) who used RAPD data, which has similar properties to AFLPs, to study inter- and intraspecific variation in three closely related annual species in Asteraceae: *Microseris elegans* and *M. bigelovii* from North America and *M. pygmaea* from Chile. The *M. elegans* populations containing closely related biotypes were found to be interspersed with genetically

very different plants. The Chilean populations of *M. pygmaea* were suggested to be the result of long distance dispersal from North America with subsequent spread from the point of establishment into two genetically isolated series of populations, one coastal and one inland. *Microseris bigelovii* is distributed along the Pacific Coast from southern California to mid-Oregon with disjunct populations near Victoria, British Columbia, which were suggested to be the result of a single colonisation event, and RAPD markers were randomised amongst the closer populations to produce a polytomy. Therefore gene flow was thought to be rare enough to allow local populations to evolve characteristic biotypes through inbreeding and selection but still sufficient to randomise allele distributions throughout the range of these closer populations.

## 5.2. Aims of Study

1. To determine how many species there are on the islands and whether these taxa are distinct.
2. To determine the spatial distribution of genotypes of island species of *Phyllica*.
3. To evaluate the origins of island populations.

## 5.3. Island Vegetation and Demographic Status of Species Involved in the Study

### 5.3.1. Island vegetation

#### 5.3.1.1. Tristan da Cunha Group

Wace and Holdgate (1958) carried out a vegetation survey of Tristan da Cunha and divided the island into four topographic zones: the Lowland Plain; the Cliffs; the Base; and the Peak. *Phyllica arborea* bush is found on the cliffs in scree and rock communities together with *Blechnum penna-marina* sward, *Rumohra adiantiforme* heath and *Blechnum palmiforme* scrub. The *P. arborea* bush on the cliffs is rather open, although patches with closed canopy do occur. The trees straggle along the ground, rooting into the shallow peat and rarely exceed a height of two metres. On the cliffs above Sandy Point c. 95% of the ground is covered by *P. arborea* bush, above Big Point *P. arborea* occurs only sporadically and above the Settlement *P. arborea* is even less frequent. On the base there are four types of vegetation: *P. arborea* bush on the lower parts of the base, *Blechnum palmiforme* scrub (450-700m), *Empetrum rubrum* heath (above 750m) and peat mires (in several places

where drainage is impeded). At the base above Big Gulch, a *P. arborea* canopy 3-5m above the ground is developed from trunks that lie along the surface for as much as 10m, sending up branches at intervals towards their downhill ends. The branches bear a heavy epiphytic flora. In places where *P. arborea* does not form a continuous canopy a mixed pteridophyte association is found. Above 450m *P. arborea* is mostly confined to sheltered gullies, and exposed trees show stunting and wind-cutting effects. The distribution of *P. arborea* is affected by the interaction of altitude and exposure, and it is not found in coastal plain communities or on the Peak.

Roux *et al.* (1992) studied the vegetation of Inaccessible and Nightingale and defined four vegetation types: tussock grassland, fern bush, wet heath and bogs. *Phyllica arborea* is found in fern bush that covers most of the plateau on Inaccessible and is restricted to regions around the ponds on Nightingale. These communities are composed of *Blechnum palmiforme* heath and *P. arborea* bush (found on the more sheltered eastern part of Inaccessible at 150-250m). Moving from *B. palmiforme* heath there is a gradation from procumbent *P. arborea* to 5m high canopies in sheltered areas. The Serengeti in the centre of Inaccessible consists of open *P. arborea* woodland that also occurs on tussock grassland on the coastal slopes. The trees occur singly, in small groups or occasionally in large groups with closed canopies. Trees off the plateau have a few lichens and an understorey of *Spartina arundinacea*. On Nightingale closed canopy *P. arborea* is found only around ponds with scattered growth on tussock grassland particularly on drainage lines.

Wace and Holdgate (1958) stated that on Tristan da Cunha:

“isolation and growth of the [human] population throughout the 19th century led to a depletion of the natural resources” and that “the island tree [*P. arborea*] was cut from the more accessible northern slopes.”

Wace (1961) reported that on Gough Island *P. arborea* formed dense thickets over broken ground and more sheltered parts of the glens below 300m. Also scattered trees were found on exposed ridges and open slopes in the same zone and among the tussock grass of western cliffs, but no trees were seen above 450m. On Gough *P. arborea* produces a pure, irregular canopy wherever it dominates any community.

#### 5.3.1.2. New Amsterdam

Valentyn (1726) described a continuous belt of forest along the east coast and Hooker (1875) stated that Labillardiere reported New Amsterdam to be covered with trees whereas the neighbouring island of St Paul had not even a shrub in 1799. The

composition of the forest was not reported. The isolated position of the island meant that the only visitors were sailors who stated that a variety of plants grew there, some of which were trees with trunks several inches in diameter. Hooker was informed of a collection by Captain Goodenough of H.M.S. Pearl and Lieutenant Hoskin stated the following in his Admiralty report:

“On the N.E. side, near the coast, on lower ranges small trees struggled for existence, looking stunted in their growth.”

Hooker (1875) stated that the specimens sent were identical to *P. arborea* from Tristan da Cunha, and he suggested that *P. arborea* may have originated from seeds from South Africa but he was unable to offer an explanation for how they had been transported.

The forest had been broken up by successive burnings until in 1874 only nine small patches of trees survived (Velain, 1893). Trehen *et al.* (1990) reported that the composition of the original fauna and flora is virtually unknown. Most present ecological systems on the island have been induced by fire and introduced flora and fauna (especially cattle). Six ecological systems were described. *Phylica* is found in the lowland area from the shoreline to an altitude of 270m. At the moment the remnants are located in the area known as ‘Le Bois’ which has been protected from cattle since 1977. There have been large changes in the soil and vegetation of the lowland over the last two centuries. Micol (1995) reported that from the end of the 18th century several accidental and deliberate peat fires, usually lasting several months, were caused by sealers. The last fire was in 1974, and it covered the whole island except the western cliffs over the course of a year, causing severe damage to the *Phylica* forest. Von Pelzeln (1861) reported that five years after a fire in 1853 thick vegetation had returned indicating rapid regeneration. No regeneration occurred in the same area after an 1899 fire (de la Rue, 1932) probably because of cattle browsing. Micol (1995) compared pictures from 1696 with one from 1875 and noted the decrease in *Phylica* forestland stated that there was a reported decrease from 27% of the island area in 1726 to 5% in 1878. The main threats to *P. arborea* were from feral cattle and alien plant species, and a restoration programme was initiated in 1987, which involved the division of the island in two by a fence. This separated the cattle from the trees and allowed a programme of reintroduction of *P. arborea* to be initiated. The depletion of the vegetation of New Amsterdam can be contrasted with man’s comparatively minimal effect on the flora of the Tristan da Cunha Group.

#### 5.3.1.3. Mauritius

Originally most of Mauritius was covered with dense tropical evergreen forest, with heath and dwarf forest at higher altitudes and palm savannas in dry eastern regions (Procter and Salm, 1975; Vaughan and Wiehe, 1937). The indigenous vegetation of the island has been almost totally cleared for cultivation or has been outcompeted by exotic species. The *Philippia/Phylica* heath formation on Mauritius is restricted to a small area of a few square kilometres at Pétrin and a tiny patch of the north flank of Mont La Selle (Midlands). *Phylica nitida* is found in upland heath (the local name for *P. nitida* is 'la bruyère' which means the heather) or dwarf heath forest at altitudes higher than 650m. These areas are almost devoid of true soil (Parish and Feillafe, 1965), and the soil that is present is nutritionally poor. Mungroo (pers. comm.) states that although receiving high rainfall (4400mm at Pétrin), the heath formation is exposed to constant drying winds, and as a result most of the species possess xeromorphic leaves which are needle-like, sclerophyllous or variously hairy. Dwarf thickets of the ericoid shrub *Philippia brachyphylla* (Ericaceae) together with *P. nitida* and *Helichrysum yuccaefolium* (Asteraceae) form a semi-open stratum 1-3m high. A number of other woody species potentially capable of developing into trees occur here as stunted individuals.

#### 5.3.1.4. Réunion

Because this island is more mountainous, there is less of a threat from human over-exploitation than on Mauritius. Coastal vegetation is badly degraded, and much low altitude forest has disappeared from the western part of the island. Moist, low altitude mixed evergreen forest (up to 1000m) exists as fragments, but the mid-altitude forest and high-altitude ericoid vegetation is better preserved. On Réunion *P. nitida* grows on nearly every mountain and is common at higher altitudes in ericoid vegetation (Thébaud, pers. comm.). It is reported at higher woodland levels, reaching optimal growth at 1500-2000m. At Piton des Neiges plants up to 30cm tall occur sporadically on rocky cliffs at 2500-3000m. Below 1500m it occupies eroded rocky crests exposed to wind.

#### 5.3.1.5. St Helena

An account of the vegetation of St Helena is presented in Chapter Six.

### 5.3.2. Taxonomic history and demographic status of species involved in the study

In his description, Hemsley (1873-76) stated that the *Phylica* plants from the Tristan da Cunha Group, New Amsterdam, Bourbon (Réunion) and Mauritius and perhaps Madagascar were members of the same species, *P. nitida*. He wrote:

“This shrub or small tree...varies considerably in foliage and general appearance at different stages of growth, especially in the Tristan da Cunha group itself. Bourbon (Réunion) specimens which we have examined have rather smaller flowers, with shorter calyx-lobes; otherwise there is little difference between them and some from Tristan da Cunha than between the specimens from Tristan da Cunha alone.”

Christopherson *et al.* (1937) stated that the *Phylica* plants found on the Tristan da Cunha Group, New Amsterdam and the Mascarenes were members of the same species, *P. arborea*. Pillans (1942) in his monograph of *Phylica* stated that *P. arborea* was found on the Tristan da Cunha Group, Mauritius and New Amsterdam. He also described a further species *P. mauritiana* from Mauritius. Guého (1977) differentiated *P. nitida*, which he described as the only species on Mauritius and Réunion, from *P. arborea* on the Tristan da Cunha Group and New Amsterdam. DNA sequence analysis (Chapter Four) has indicated that the Mascarene species is distinct. The taxonomic history and morphological differences between each of the species involved in the study is presented along with details of samples collected and used in this study.

#### 5.3.2.1. *Phylica arborea*

*Phylica arborea* Thouars (*Soulangia arborea* Don; *P. superba* Hort. ex A. Dietr.) was described in the Flora of Tristan d'Acugna (Thouars, 1811). On Tristan da Cunha this plant is known as the ‘island tree’. This species occurs on the Tristan da Cunha Group of islands in the South Atlantic and New Amsterdam Island in the southern Indian Ocean. The following is a summary of information available on *P. arborea* prior to this study.

#### *Phylica arborea* on the Tristan da Cunha Group

Moseley (1875) was on Tristan da Cunha for a very short time so he only visited the shoreline of the settlement and the gully immediately above the settlement (Hottentot Gulch). He reported that the cliffs were scantily covered with grasses,



sedges, mosses and ferns, with darker patches of *P. arborea* and *Empetrum nigrum* var. *rubrum* becoming more and more marked towards the summit. In the gully above the settlement *P. arborea* grew from 150m upwards. Other trees in this locality had been cut down for firewood, but there was still plenty of wood on the island, and the trunks of the trees on the upper plateau reached a diameter of 40cm (according to the inhabitants). On Inaccessible Island the cliffs were densely covered with *Spartina arundinacea* with *P. arborea* growing on the summits of slight elevations. The trees grew thickly together and their branches met overhead. The ground beneath them was covered with ferns, mosses and sedges with *Acaena sanguisorbae* and *Chenopodium tomentosum* (the tea plant). Trunks of the trees were covered in lichens. *Phylica arborea* grew on the base and could grow under the shelter of cliffs to a height of 6m or somewhat more. The trunks were never straight, but usually procumbent and again ascending, with the largest seen being 30cm in diameter (on the upper plateau diameters of 45cm had been reported). Trees in exposed areas were beaten down by gales. The wood of the tree was reported to be brittle, and when exposed rapidly decays, but it was serviceable when dried carefully with the bark present. On Nightingale Island *S. arundinacea* covered the whole island except the summits of ridges and a few patches on the lower tract, which were occupied by *P. arborea*. Many of these trees in one spot were prostrate because of the wind and some were dead. *Phylica arborea* and *S. arundinacea* dominated the conspicuous part of the vegetation of all the islands. *Phylica arborea* occurred in patches or coppices in the midst of large areas of the grass, the ground beneath being covered with a thick growth of mosses, sedges and ferns, *Nertera depressa*, *Acaena sanguisorbae* and *Chenopodium tomentosum*. On all islands the trees were in the same stage of development, bearing fully formed, but green, fruit.

Moseley (1875) gave the following description for *P. arborea*:

“The foliage of the tree is of a dark glossy green, with the undersides of the narrow, almost needle-like leaves white and downy. Hence the tree, which in habit is very like a yew, presents as a whole a mixture of glaucous grey and dark olive green shades; it bears berries of about the size of sweet peas, which are eaten by the finch which lives in the islands.”

He added that:

“the constant heavy gales do not permit the tree to grow erect; the trunk is usually procumbent at its origin for several feet, and then rises again often at a right angle. It is always more or less twisted or gnarled. In sheltered places, as under the cliffs on the north-east of Inaccessible Island, the tree is as high as 25 feet, but it is not nearly so high on the summit of the island, though the trunks are said to reach a length of 30 feet or more.”

Carmichael (1818) reported that the northern extremity of the settlement plain of Tristan da Cunha was largely cleared of its wood. Firing of grass and trees had destroyed the vegetation, but the remains still lay on the ground. The rest of the island was still in a state of nature, covered with an impenetrable copse. In an ascent to the peak Carmichael commented that during the climb they did not rely on any support from *Phylica* bushes because most of them were rotten.

Breytenbach *et al.* (1986) studied the different patterns of regeneration of *P. arborea* on each of the islands in the Tristan da Cunha Group. *Phylica arborea* is a myrmecochorous species, and Breytenbach suggested that the absence of ants on Gough results in seeds not being buried and consequently regeneration is low. Because seeds need to be buried regeneration only occurs on land-slips which are rare. Breytenbach also suggested that predation by introduced mice (*Mus musculus*) could be preventing regeneration of unburied seeds. Ryan *et al.* (1989) compared *Phylica* regeneration on Gough with that on Inaccessible where there are no mice. On Inaccessible regeneration is regular whereas on Gough it is episodic. The author suggested that this is possibly due to the presence of mice on Gough, and he suggested the need to study the factors preventing regeneration there. Milton *et al.* (1993) suggested that the absence of ants and the presence of mice were not as important in tree regeneration as Breytenbach *et al.* (1986) had postulated because of the activities of ground- and burrow-nesting seabirds. There is no evidence that mice destroy ripe seeds. *Phylica arborea* seedlings survive longer on mineral soil than on organic soil, and they colonise bare ground ahead of rhizomatous ferns. Saplings were found in all Tristan da Cunha populations but were more frequent on disturbed sites. On Gough saplings are absent from established populations. Milton *et al.* (1993) hypothesised that the periodic recruitment of *P. arborea* follows disturbance-induced mortality of parent plants, and the patchy distribution and homogeneous age structure of old *P. arborea* populations on Gough Island indicates that recruitment was dependent on disturbances long before the introduction of rodents to the island. On the relatively drier Tristan da Cunha, Inaccessible and Nightingale, regeneration is continuous with seeds germinating beneath dying trees. On wetter Gough the fern *Histiopteris incisa* grows under and around *Phylica* perhaps preventing the establishment of *Phylica* seedlings.

## Account of Samples of *Phyllica arborea* Collected on Tristan da Cunha and Nightingale

The following passage is an assessment of the state of *P. arborea* on Tristan da Cunha and Nightingale Island after a field trip I made in October 1996. Collections were made from four areas on Tristan da Cunha and from a population on Nightingale Island. Details of the samples collected are also given and the distribution of samples collected is shown in Figures 5.1 and 5.2.

### Settlement Plain (JER55-76 and JER133, 134)

Growth of *P. arborea* below about 100m is prevented by cattle grazing. No flowering was observed on cliffs above the Settlement Plain up to an altitude of c. 250m, which was the maximum to which I was permitted to climb. Some ripe fruits were present, possibly from a late winter flowering. A fasciated form was found growing above Donkey Piece at 350m in a population of about 15 individuals including seedlings. Tree heights ranged from 1.5-3.5m, and the fruit size was about 6mm. Samples were collected at c. 400m intervals along the Settlement Plain at altitudes of between 100 and 250m.

### Burntwood (JER82, JER87)

Trees at the edge of the cliffs were up to 2.5m high and covered in lichen. These plants were growing in a hollow so they were protected from wind. Other individuals which were more exposed grew to a maximum height of about 1m, the leaves were yellowish, only growing at tips of branches and they were thinner and more revolute than the leaves on trees on the side of cliffs or in more sheltered areas. Some ripe fruits were found in a similar stage of development to those found on the cliffs above the Settlement Plain. The tree heights ranged from 0.5-2.5m. The fruit size was about 6mm.

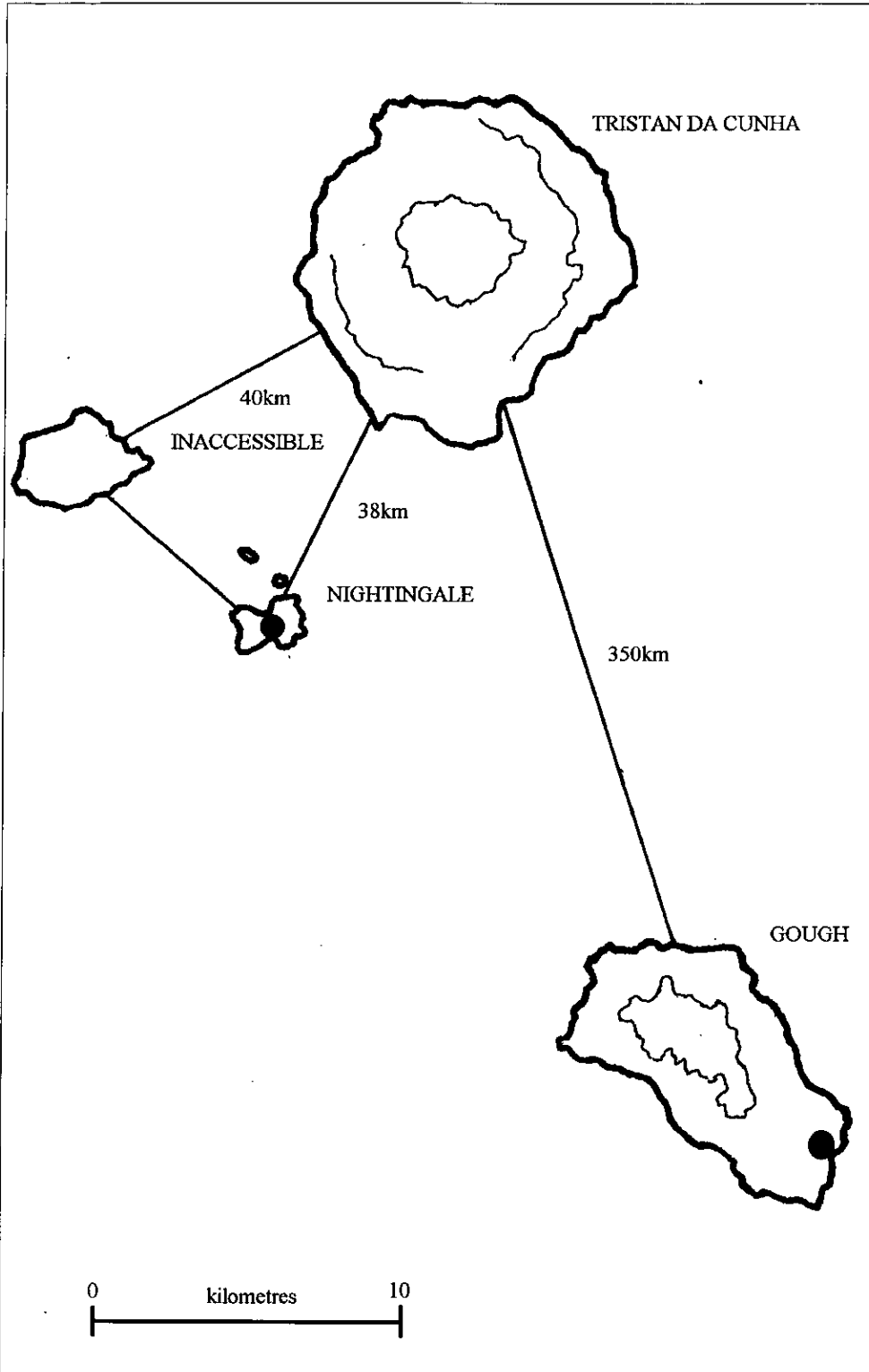


Figure 5.1. The Tristan da Cunha Group (distances between islands not drawn to scale). The black spots indicate sites where *Phyllica arborea* was collected (more detail in Figure 5.2). Map taken from Groves (1981).

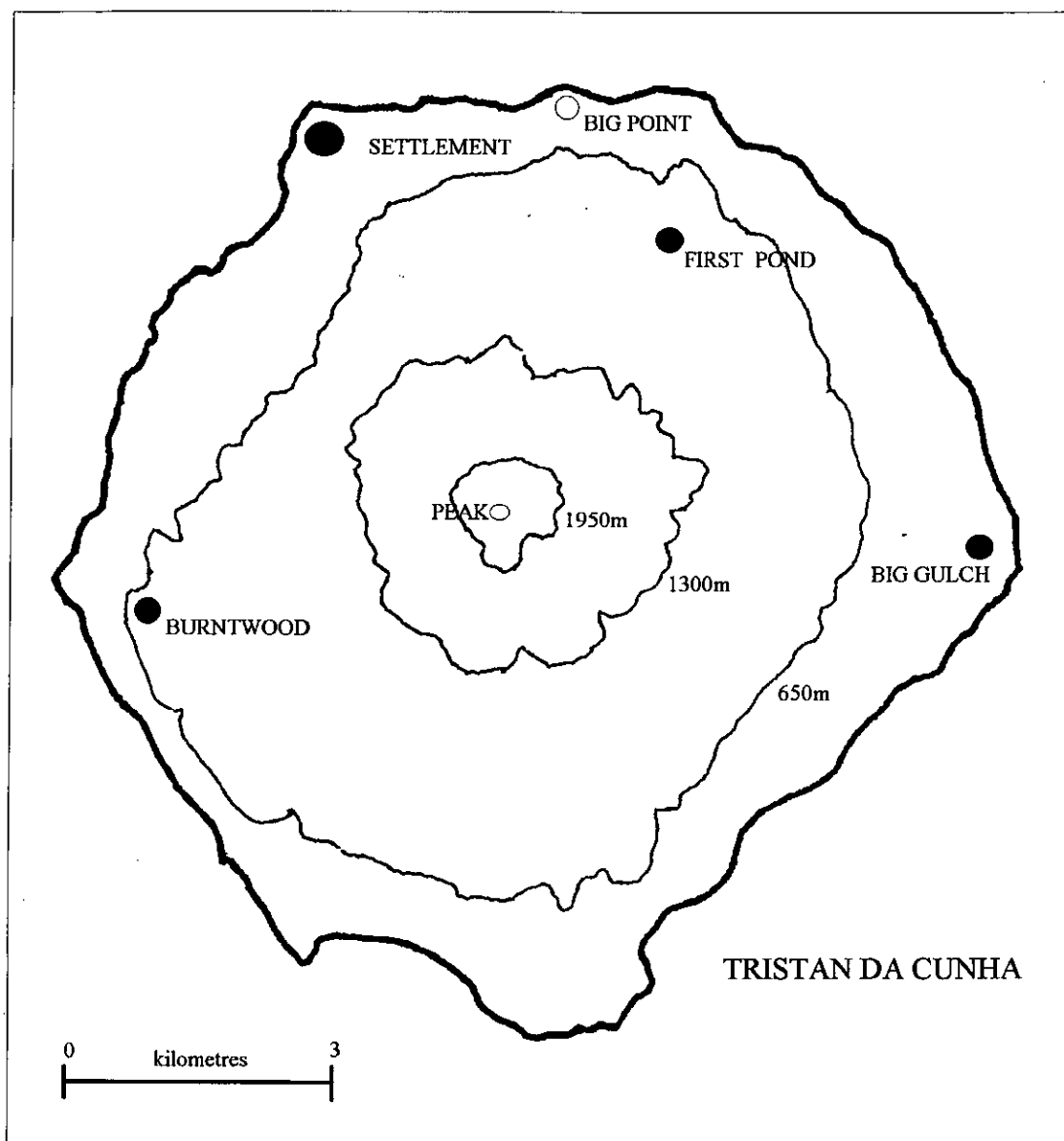


Figure 5.2. Tristan da Cunha. Black spots indicate sites where *Phyllica arborea* was collected. Map taken from Groves (1981). Settlement Plain samples were collected from heights of between 100 and 250m between the Settlement and Burntwood.

#### Big Gulch (JER93-102)

Samples were taken from a population of about 600 individuals on the west-facing side of the gulch just above a penguin rookery. This gulch is probably the warmest part of the island and one of only two places where trees were flowering at the time of my visit. According to my guide, most trees in the population were about 12 years old. Few seedlings were found. All trees older than about five years had grey and red

lichens growing on the bark. The largest trees grew to heights of 4m. The fruit size was 6mm. Samples were collected from trees which had no visible fungal growth.

#### First Pond (JER137-151)

Samples were taken from a population of about 400 trees just above a mossy bog on the edge of the pond in a dense *Blechnum palmiforme* undergrowth. The older trees (4-5 years) had immature, reddish fruits, about 5mm across. Tree heights ranged from 1.5 to 2m. Fruit sizes ranged from 2-5mm, and the colours ranged from red to green, due to differences in maturation stage. The leaves were about 3 to 5mm across and yellowish except for those at the apices, which were reddish.

#### Nightingale Island (JER108-122)

Samples were taken from one population of about 500 individuals. Trees are still being cut by islanders, but there are plenty of regenerating seedlings found in open areas. Tree heights were up to 7m, the green fruits were about 9mm across and leaves were about 7mm across compared to an average of about 4mm on Tristan da Cunha. The greater general size of trees on Nightingale could be due to the more fertile land and lower exposure than on Tristan da Cunha. The fertility on Nightingale may be greater due to the larger number of birds nesting on the island and the consequent increase in bird droppings.

Instructions were left with one of the islanders to collect material from the southern part of Tristan da Cunha and from Inaccessible, which I was unable to reach on my visit. However, this material was not received in time to be included in the study.

Generally speaking *P. arborea* on Tristan da Cunha appeared to be in a reasonably healthy state. Apart from those areas which were in constant human use, *P. arborea* appeared to be growing in a state similar (numbering tens of thousands) to when settlers first colonised the islands. They grow in an altitudinal zone from sea level to about 500m around Tristan da Cunha itself and in isolated populations on Nightingale and Inaccessible. Large-scale use of *Phyllica* wood had been discontinued by the islanders as a result of the greater use of natural gas. Small-scale collection of wood is permitted for use by some of the older islanders, and wood is also still collected by islanders on their annual trips to Nightingale Island, but again

this is in small quantities and does not appear to be having any adverse affect on the population. Seedlings were common throughout nearly all areas visited. The healthy state, in terms of numbers, of Tristan da Cunha Group populations may be contrasted with the relatively unhealthy state of populations on New Amsterdam.

*Phyllica arborea* on Gough Island is in a similar condition to that on Tristan da Cunha in terms of numbers (Roux, pers. comm.). The Gough Island samples (KR1 to 9) were collected by J.P. Roux from a number of populations in the southern part of the island.

#### *Phyllica arborea* on New Amsterdam

Samples of *P. arborea* were collected from four sites on New Amsterdam: Grand Bois, Martin du Viviès, Antonelli Crater, and Grand Tunnel although only samples from Grand Bois and Martin du Viviès were used in the final analysis. Samples were collected by Yves Frenot. These four sites represent fragments of the original distribution.

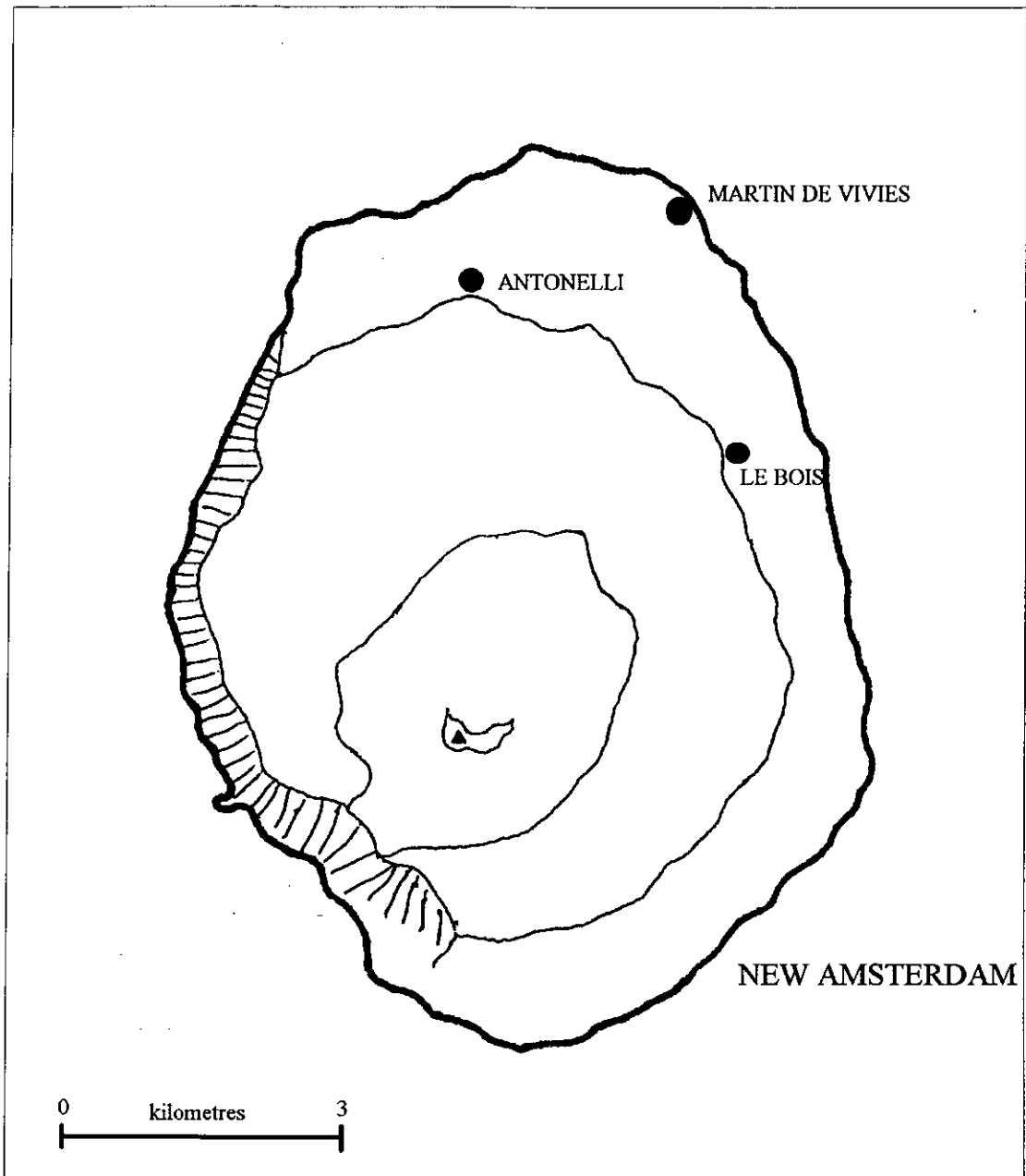


Figure 5.3. New Amsterdam. Black spots indicate sites from where *P. arborea* was collected. Map taken from Tréhen *et al.* (1990).

#### 5.3.2.2. *Phylica polifolia*

*Phylica polifolia* (Vahl) Pillans (*Rhamnus polifolia* Vahl; *P. thymifolia* Vent.; *P. rosmarinifolia* Thunb.; *P. ramosissima* DC; *Soulangia thymifolia* Brongn.; *Trichocephalus ramosissimus* (DC) Don) was first described as *Rhamnus polifolia*



Vahl in *Symbolae Botanicae* 3: 41 (1794). Details of samples collected and the demographic status of this species may be found in Chapter Six.

#### 5.3.2.3. *Phylica nitida*

*Phylica nitida* Lam. (*Blaeria leucocephala* Bory; *P. leucocephala* (Bory) Cordem.; *P. mauritiana* Boj. ex Baker; *P. mauritiana* var. *linearifolia* Pillans) was described in *Tableau Encyclopédie Méthodique Botanique* 2: 77 (1797). Guého (1977) lumped the species of *Phylica* which occur on Mauritius and Réunion into a single species, *P. nitida*. DNA sequence studies (Chapter Four) placed the individuals of *Phylica* from Mauritius and Réunion in a clade in the ‘island group’. The level of divergence between this group and the ‘paniculata group’ is reasonably high, indicating that they are more ancient derivatives of the ancestor of the ‘island group’ than are other members of this group. The level of DNA sequence divergence between the Mauritian and Réunion plants is also much higher than that between the members of the ‘paniculata group’.

The six Mauritian *P. nitida* individuals (collected by Yusoof Mungroo; YM1-6) used in this analysis came from the single remaining population within the Pétrin Conservation Management Area, with each individual being about 2-3 metres apart. The Pétrin Conservation Management Area covers 6.2 hectares fenced in February 1995 and weeded of Chinese guava (*Psidium cattleianum*), privet (*Ligustrum robustum*), ravenale (*Ravenala madagascariensis*), *Eucalyptus* sp., pine (*Pinus* sp.) and wild raspberry (*Rubus alcaefolius*). Material from Réunion was taken from five individuals (collected by Christophe Thébaud; CT1-5) of a population located near the Plateau des Basaltes on the active volcano (Piton de la Fournaise) about 5-10 metres apart.

#### 5.3.2.4. *Phylica emirnensis*

*Phylica emirnensis* (Tulasne) Pillans was first described as *Tylanthus emirnensis* Tulasne in *Annales Sciences Naturelles*, series 4, 8: 128 (1857) and was subsequently placed in *Phylica* by Pillans (1942). This species is from mountains in the province of Emirna, Madagascar.

No fresh or silica-gel dried material of this species was available, and so this species was excluded from this study. A *trnL-F* sequence was produced from a herbarium specimen, which indicated a relationship to *P. tropica* (Malawi) and *P. natalensis* (eastern South Africa) in a clade which is part of the ‘paniculata group’

(see previous chapter). Additional fresh or silica gel dried material for further molecular studies on this species is currently sought.

#### 5.3.2.5. *Phylica bathiei*

*Phylica bathiei* Pillans is from Madagascar but without precise locality. No fresh, silica gel dried or herbarium material of this species was available.

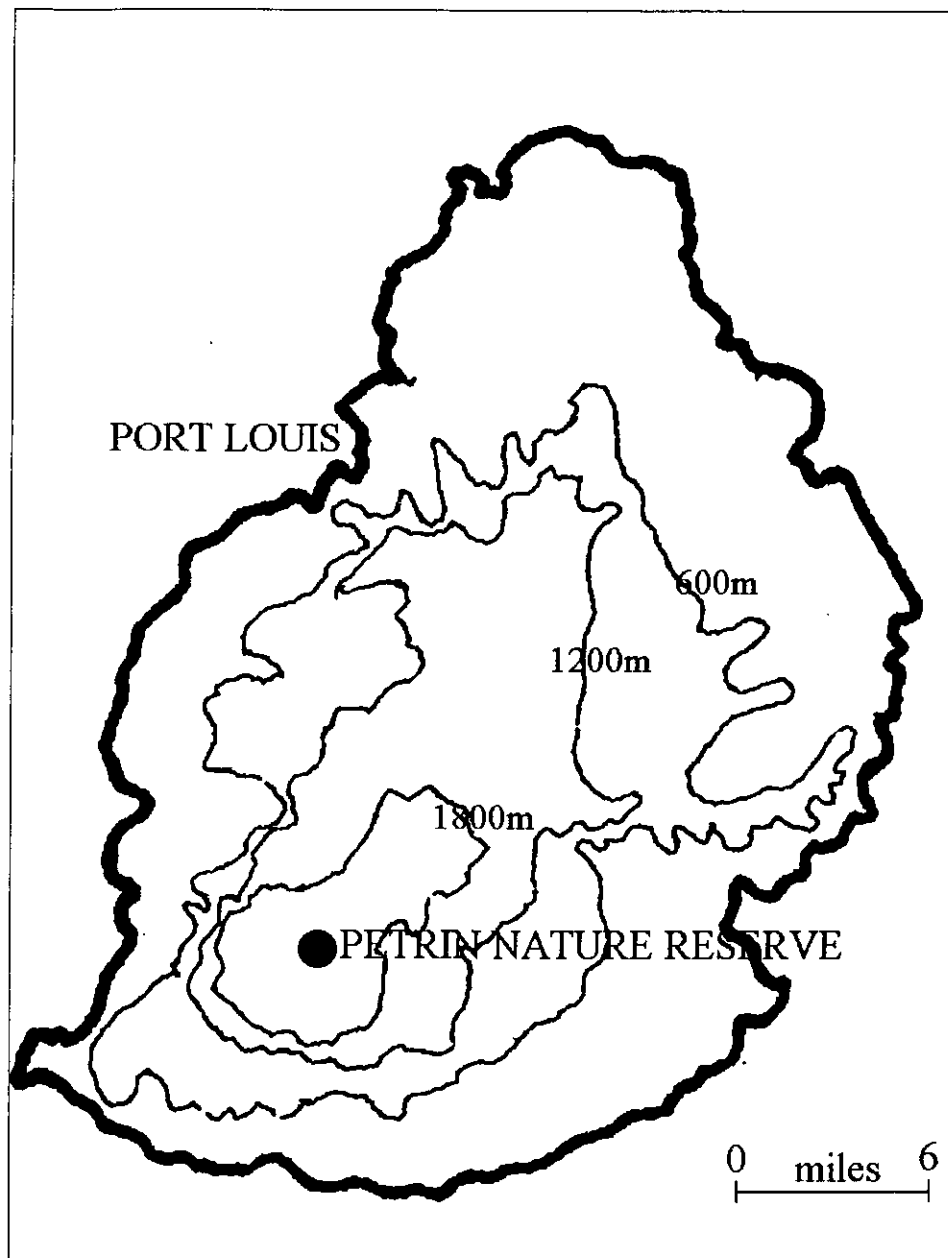


Figure 5.4. Mauritius. *Phylica nitida* was collected from the Pétrin Nature Reserve. Map taken from White (1983).

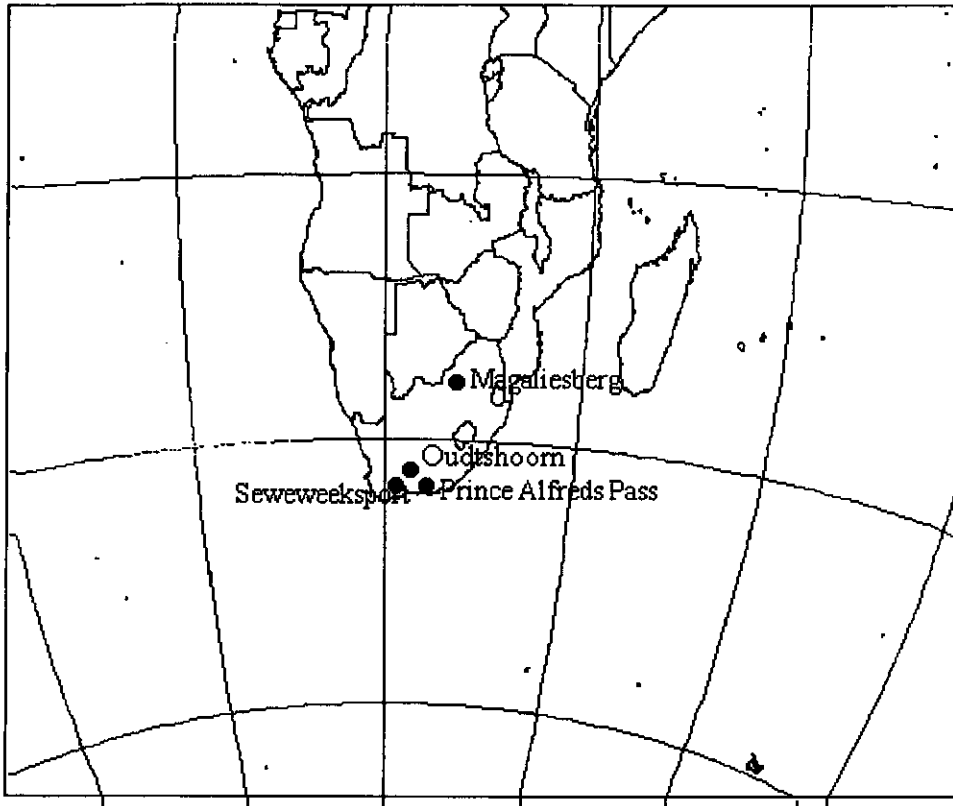


Figure 5.5. Distribution of collected samples of *Phylica paniculata*.

#### 5.3.2.6. *P. paniculata*

*Phylica paniculata* Willd. (*P. oblongifolia* Du Mont de Cours; *P. thymifolia* Vent.; *P. myrtifolia* Poir.; *P. ledifolia* Desf.; *P. angustifolia* Hort. ex Steud.; *Soulangia paniculata* (Willd.) Brongn.; *S. arborescens* Ecklon and Zeyher; *S. rosmarinifolia* Harv.; *S. myrtifolia* A. Dietr.; *S. rubra* A. Dietr.; *S. epacridifolia* A. Dietr.; *P. sessiliflora* Hort. ex Steud.; *P. arborescens* Steud.; *S. marifolia* Bernh. ex Krauss; *S. parviflora* Presl.) was first described in *Species Plantarum* 1: 1112 (1798). *Phylica paniculata* has the widest distribution of any continental species in the genus, from the Worcester to Maclear Divisions (Cape Province) to near Durban (Natal), Barberton, Rustenberg and Lydenberg Divisions (Transvaal) and the Chimanimani Mountains of Zimbabwe. It is found either in montane areas or along river banks. The individuals used in this study were taken from Seweweekspoort (JER162), Prince Alfred's Pass (CFR136) and Oudtshoorn (FMW950) in Cape Province and from Magaliesberg (Transvaal; MvdB1-2). As mentioned in the previous chapter, it has putatively primitive morphological characteristics. This

species has been demonstrated to be related to *P. arborea* and *P. polifolia* (Chapter Four).

### 5.3.3. Morphological differences between members of the island group

The morphological differences between members of the 'island group' are summarised by the following key:

- 1a. Flowers in short spikes assembled in panicles, or in peduncled or sessile clusters in the axils of the upper leaves, or crowded in a thyriform inflorescence 2
- b. Flowers in capitula subtended by several leaves 4
- 2a. Leaves at first with short tomentum upon the upper surface; sepals 1-1.5 mm long, with dorsal hair at least half as long *P. arborea*
- b. Leaves at first pilose upon the upper surface; sepals 0.75-1 mm long, with dorsal hair much less than half as long 3
- 3a. Petals with claw one-third as long as lamina *P. polifolia*
- b. Petals with claw as long as lamina *P. paniculata*
- 4a. Petals with the lamina rotundate, cucullate, deeply concave 5
- b. Petal lamina lanceolate, ovate-lanceolate, concave on the inner side and slightly incurved at the apex or towards the middle but never cucullate 6
- 5a. Flowers pedicellate *P. natalensis*
- b. Flowers stipitate *P. nitida*
- 6a. Flowers about 5 mm long; petals inserted on the upper half of the tube *P. tropica*.
- b. Flowers about 3.5 mm long; petals inserted at the mouth of the tube *P. emirnensis*

## **5.4. Methods**

### 5.4.1. Sampling Strategies

Conditions for sampling on Tristan da Cunha were not ideal. It is preferable to survey a site first and then to sample. However, the time I spent on the island was not sufficient for me to do this due to problems with access. I attempted a nested sampling: 1. Between islands; 2. Between populations on Tristan da Cunha; 3. Within populations. Samples were collected randomly within these subsets. Sampling of other species and the New Amsterdam individuals was undertaken by others employing similar strategies.

#### 5.4.2. Material for analysis

Sources of plant material and vouchers used in this analysis are listed in Table 5.1.

#### 5.4.3. DNA extraction

DNA was extracted in two ways:

1. DNA was extracted from *c.* 1g fresh or 0.2-0.25g silica gel-dried leaves using a 2X CTAB method modified from Doyle and Doyle (1987). DNA was precipitated using isopropanol instead of ethanol because it had been found to be more reliable for these taxa in previous studies (Chapter Two). Some samples were purified on caesium chloride/ethidium bromide gradients (1.55g/ml).
2. Extractions were also performed using a further modified 2X CTAB method in which DNA was purified using QIAquick columns (QIAGEN, Crawley, West Sussex, UK) following protocols provided by the manufacturers.

#### 5.4.4. Amplified Fragment Length Polymorphisms (AFLPs)

Protocols supplied by the Perkin-Elmer Corporation (Applied Biosystems Inc., Warrington, Cheshire, UK) were used to produce amplified fragment length polymorphisms (AFLPs; Vos *et al.*, 1995). DNA was restricted with the endonucleases *EcoRI* and *MseI*, and fragments were ligated to double stranded adaptors. Two rounds of PCR amplification were then performed: pre-selective amplification used primers with a one base (bp) pair extension, and selective amplification used dye labelled primers with a three bp extension. This process reduces fragments to a number that may be visualised. Two different selective primer pairs were used (ACA/CAA and AAC/CAT anchors). These were chosen after an initial study of a range of primer pairs on two closely related individuals and another more distantly related individual (as indicated by the sequence data). The chosen primer pairs gave sufficient variation to allow distinction between closely related individuals and at the same time gave some shared bands between more distantly related individuals.

#### 5.4.5. Running AFLPs on gels and band scoring

The AFLPs were separated and visualised using an ABI 377 automated sequencer (according to the manufacturer's protocols; Applied Biosystems, Inc., Warrington, Cheshire, UK). Fragments were sized by running dye-labelled size standards in each lane. The AFLP profiles were edited using Genescan version 2.0.2 and Genotyper version 1.1 (Applied Biosystems Inc., Warrington, Cheshire, UK). Genescan automatically scores bands ranging from 50-500 bp in length. Bands that were below a cut of 50 arbitrary fluorescence units were not scored. Bands were edited manually because some bands were just below the threshold permitted by the software in some individuals and just above the threshold in others. Non-homologous bands that fell within the same size class were also edited manually. Bands were scored as present/absent and a binary matrix was produced.

Table 5.1. Samples used in a study of AFLPs in island species of the genus *Phyllica* L. JER (J.E. Richardson), KR (J.P.Roux), YF (Yves Frenot), YM (Yusoof Mungroo), CT (Christophe Thébaud), RR (Rebecca Rowe). \* indicates samples not used in the final analyses because the DNA was of insufficient quality.

Sample	Locality	Collector number
<i>P. polifolia</i> HHA	St Helena (High Hill)	RRA
<i>P. polifolia</i> HHB*	St Helena (High Hill)	RRB
<i>P. polifolia</i> LOT1	St Helena (Lot)	RR1
<i>P. polifolia</i> LOT2	St Helena (Lot)	RR2
<i>P. polifolia</i> LOT10	St Helena (Lot)	RR10
<i>P. polifolia</i> LOT11	St Helena (Lot)	RR11
<i>P. polifolia</i> 17	St Helena (High Hill)	RR17
<i>P. polifolia</i> 18	St Helena (High Hill)	RR18
<i>P. polifolia</i> 19*	St Helena (High Hill)	RR19
<i>P. polifolia</i> 20	St Helena (High Hill)	RR20
<i>P. polifolia</i> 21	St Helena (High Hill)	RR21
<i>P. polifolia</i> 22	St Helena (High Hill)	RR22
<i>P. polifolia</i> 23	St Helena (High Hill)	RR23
<i>P. polifolia</i> 25*	St Helena	RR25
<i>P. polifolia</i> 26A*	St Helena	RR26A
<i>P. polifolia</i> 26B*	St Helena	RR26B
<i>P. polifolia</i> 27*	St Helena	RR27
<i>P. polifolia</i> 28*	St Helena	RR28
<i>P. polifolia</i> 29*	St Helena	RR29
<i>P. polifolia</i> 30*	St Helena	RR30
<i>P. polifolia</i> 31	St Helena (High Hill)	RR31
<i>P. polifolia</i> 32*	St Helena	RR32
<i>P. paniculata</i> JER162	South Africa (Seweweekspoort, Cape Province)	JER162
<i>P. paniculata</i> 136(6)	South Africa (Prince Alfred's Pass, Cape Province)	CFR136
<i>P. paniculata</i> 136(7)*	South Africa (Prince Alfred's Pass, Cape Province)	CFR136
<i>P. paniculata</i> Mag1	South Africa (Magaliesburg, Transvaal)	MvdB1
<i>P. paniculata</i> Mag2	South Africa (Magaliesburg, Transvaal)	MvdB2

<i>P. paniculata</i> Mag3*	South Africa (Magaliesburg, Transvaal)	MvdB3
<i>P. paniculata</i> Mag4*	South Africa (Magaliesburg, Transvaal)	MvdB4
<i>P. paniculata</i> FMW950	South Africa (Oudtshoorn, Cape Province)	FMW950
<i>P. arborea</i> 55	Tristan (Spring Gulch)	JER55
<i>P. arborea</i> 56	Tristan (Gulch to north of Spring Gulch)	JER56
<i>P. arborea</i> 57	Tristan (Spring Ridge)	JER57
<i>P. arborea</i> 58	Tristan (Gulch just south of Wash Gulch)	JER58
<i>P. arborea</i> 59	Tristan (Goatridge)	JER59
<i>P. arborea</i> 60	Tristan (Goatridge)	JER60
<i>P. arborea</i> 61	Tristan (Little Sandy Gulch)	JER61
<i>P. arborea</i> 62	Tristan (Big Sandy Gulch)	JER62
<i>P. arborea</i> 63*	Tristan (Between Big Sandy and Wash Gulches)	JER63
<i>P. arborea</i> 65	Tristan (Between Wash and Spring Gulch)	JER65
<i>P. arborea</i> 66*	Tristan (Between Big Sandy and Wash Gulches)	JER66
<i>P. arborea</i> 67	Tristan (330 metres north of 66)	JER67
<i>P. arborea</i> 71	Tristan (Wash Gulch)	JER71
<i>P. arborea</i> 72	Tristan (Between Big Sandy and Wash Gulches)	JER72
<i>P. arborea</i> 74	Tristan (Between Big Sandy and Wash Gulches)	JER74
<i>P. arborea</i> 75	Tristan (north side of Big Sandy Gulch)	JER75
<i>P. arborea</i> 76	Tristan (south side of Big Sandy Gulch)	JER76
<i>P. arborea</i> 82	Tristan (cliff edge, Burntwood)	JER82
<i>P. arborea</i> 87*	Tristan (base at Burntwood)	JER87
<i>P. arborea</i> 91	Tristan (First Gulch)	JER91
<i>P. arborea</i> 92	Tristan (Donkey Piece)	JER92
<i>P. arborea</i> 93*	Tristan (Big Gulch)	JER93
<i>P. arborea</i> 94	Tristan (Big Gulch)	JER94
<i>P. arborea</i> 95*	Tristan (Big Gulch)	JER95
<i>P. arborea</i> 96	Tristan (Big Gulch)	JER96
<i>P. arborea</i> 97	Tristan (Big Gulch)	JER97
<i>P. arborea</i> 98*	Tristan (Big Gulch)	JER98
<i>P. arborea</i> 99*	Tristan (Big Gulch)	JER99
<i>P. arborea</i> 100	Tristan (Big Gulch)	JER100
<i>P. arborea</i> 101*	Tristan (Big Gulch)	JER101
<i>P. arborea</i> 102*	Tristan (Big Gulch)	JER102



<i>P. arborea</i> 108*	Nightingale (Resting Place)	JER108
<i>P. arborea</i> 109*	Nightingale (Resting Place)	JER109
<i>P. arborea</i> 110*	Nightingale (Resting Place)	JER110
<i>P. arborea</i> 111	Nightingale (Resting Place)	JER111
<i>P. arborea</i> 112*	Nightingale (Resting Place)	JER112
<i>P. arborea</i> 113	Nightingale (Resting Place)	JER113
<i>P. arborea</i> 114	Nightingale (Resting Place)	JER114
<i>P. arborea</i> 115	Nightingale (Resting Place)	JER115
<i>P. arborea</i> 116	Nightingale (Resting Place)	JER116
<i>P. arborea</i> 117	Nightingale (Resting Place)	JER117
<i>P. arborea</i> 118	Nightingale (Resting Place)	JER118
<i>P. arborea</i> 119	Nightingale (Resting Place)	JER119
<i>P. arborea</i> 120	Nightingale (Resting Place)	JER120
<i>P. arborea</i> 121	Nightingale (Resting Place)	JER121
<i>P. arborea</i> 122*	Nightingale (Resting Place)	JER122
<i>P. arborea</i> 133*	Tristan (cliffs above volcano)	JER133
<i>P. arborea</i> 134*	Tristan (base of volcano)	JER134
<i>P. arborea</i> 137*	Tristan (First Pond)	JER137
<i>P. arborea</i> 138*	Tristan (First Pond)	JER138
<i>P. arborea</i> 139	Tristan (First Pond)	JER139
<i>P. arborea</i> 140	Tristan (First Pond)	JER140
<i>P. arborea</i> 141*	Tristan (First Pond)	JER141
<i>P. arborea</i> 142	Tristan (First Pond)	JER142
<i>P. arborea</i> 143	Tristan (First Pond)	JER143
<i>P. arborea</i> 144*	Tristan (First Pond)	JER144
<i>P. arborea</i> 145*	Tristan (First Pond)	JER145
<i>P. arborea</i> 146	Tristan (First Pond)	JER146
<i>P. arborea</i> 147*	Tristan (First Pond)	JER147
<i>P. arborea</i> 148	Tristan (First Pond)	JER148
<i>P. arborea</i> 149*	Tristan (First Pond)	JER149
<i>P. arborea</i> 150*	Tristan (Second Pond)	JER150
<i>P. arborea</i> 151	Tristan (hill above Third Pond)	JER151
<i>P. arborea</i> KR1	Gough Island (between Meteorological Station and Seal Beach)	JR1
<i>P. arborea</i> KR3	Gough Island (first trees below Tafel-Koppie)	JR3

<i>P. arborea</i> KR4*	Gough Island (east of helipad above base)	JR4
<i>P. arborea</i> KR5	Gough Island (east of helipad above base)	JR5
<i>P. arborea</i> KR6	Gough Island (between Geese? and Tafel Koppie)	JR6
<i>P. arborea</i> KR7	Gough Island (Ruin Ridge)	JR7
<i>P. arborea</i> KR8*	Gough Island (Meteorological Station)	JR8
<i>P. arborea</i> KR9	Gough Island (Meteorological Station)	JR9
<i>P. arborea</i> AI1*	New Amsterdam (Grand Bois)	YF1
<i>P. arborea</i> AI2	New Amsterdam (Grand Bois)	YF2
<i>P. arborea</i> AI3*	New Amsterdam (Grand Bois)	YF3
<i>P. arborea</i> AI4	New Amsterdam (Grand Bois)	YF4
<i>P. arborea</i> AI5	New Amsterdam (Martin du Viviers)	YF5
<i>P. arborea</i> AI6*	New Amsterdam (Antonelli Crater)	YF6
<i>P. arborea</i> AI7*	New Amsterdam (Antonelli Crater)	YF7
<i>P. arborea</i> AI8*	New Amsterdam (Grand Tunnel)	YF8
<i>P. arborea</i> AI9*	New Amsterdam (Martin du Viviers)	YF9
<i>P. arborea</i> AI10*	New Amsterdam (Martin du Viviers)	YF10
<i>P. nitida</i> CT1	Réunion (Piton de la Fournaise)	CT1
<i>P. nitida</i> CT2	Réunion (Piton de la Fournaise)	CT2
<i>P. nitida</i> CT3	Réunion (Piton de la Fournaise)	CT3
<i>P. nitida</i> CT4	Réunion (Piton de la Fournaise)	CT4
<i>P. nitida</i> CT5	Réunion (Piton de la Fournaise)	CT5
<i>P. nitida</i> MM1	Mauritius (Pétrin Nature Reserve)	YM1
<i>P. nitida</i> MM2	Mauritius (Pétrin Nature Reserve)	YM2
<i>P. nitida</i> MM3	Mauritius (Pétrin Nature Reserve)	YM3
<i>P. nitida</i> MM4	Mauritius (Pétrin Nature Reserve)	YM4
<i>P. nitida</i> MM5	Mauritius (Pétrin Nature Reserve)	YM5
<i>P. nitida</i> MM6	Mauritius (Pétrin Nature Reserve)	YM6
<i>N. elliptica</i> 1	St Helena	RRNes1
<i>N. elliptica</i> 2	St Helena	RRNes2
<i>N. elliptica</i> 3	St Helena	RRNes3
<i>N. elliptica</i> 4	St Helena	RRNes4

#### 5.4.6. Data analysis: methods for analysing restriction fragment data and expectations for performance

Several methods were used to analyse the AFLP data generated in this study. The use of these methods are discussed below.

##### 5.4.6.1. Unweighted pair group method with arithmetic means (UPGMA)

UPGMA (Sokal and Sneath, 1963) involves the production of a similarity matrix in which the most similar units are clustered together sequentially. The distance between two clusters is the average of the distances between members of one cluster and members of the other. The total amount of divergence is divided equally between the two groups, i.e. the lengths of the corresponding branches of the phenogram leading to members of each group is half the total divergence between them. This method can be expected to produce spurious results when rates of change among individuals are heterogeneous.

##### 5.4.6.2. Parsimony

Cladistic and phenetic methods produce divergent branching patterns, and the former will not determine the correct relationships for a group that contains taxa of hybrid origin. The interpretation of relationships between interbreeding individuals using phylogenetic methods is inappropriate but generally would be expected to produce unresolved relationships. Interbreeding results in segregation of alleles, and phylogenetic methods can only be appropriately applied to non-reticulating taxa or clonally inherited molecules such as mtDNA or cpDNA. Segregation would be expected to reveal large amounts of conflict and little consensual support for topologies produced using this method. However, parsimony should work for isolated populations between which no genetic exchange is taking place. This method was used to assess differences in the results compared with the other methods used.

##### 5.4.6.3. Neighbour Joining

Unlike UPGMA, this method permits rate heterogeneity. The principle is to find pairs of OTUs (neighbours) that minimise the total branch length at each stage of clustering of OTUs, starting with a star-like tree.

#### 5.4.6.4. Principal Co-ordinates Analysis (PCO)

Clustering methods such as UPGMA and distance methods such as NJ may be criticised in that they assume that clusters are present within a given data set. This assumption is avoided by using ordination or multi-dimensional scaling methods. Ordination is a way of describing how the experimental units in a study relate to each other if many measurements are made on each of them. Units are represented by points in geometrical space with one dimension for each variable measured.

Principal co-ordinates analysis (PCO; Gower, 1966) represents the distances between units by a map. A similarity matrix is produced which calculates the distances between all possible pairs of units. The process of turning a data matrix into a distance or similarity matrix can be reversed: a matrix of similarities between units can be used to map the units as points in a geometric space with a reduced number of dimensions. The map can reveal hidden patterns in the similarity matrix and show whether any units can be grouped. This method is an example of metric scaling.

#### 5.4.6.5. Software

Data was analysed using two software packages. Parsimony, UPGMA and NJ algorithms of the software package PAUP version 4.0d64 for Macintosh (Swofford, 1998) were used. The heuristic search strategy of the parsimony analysis was the same as that which was used in the previous chapters but without successive weighting. MacClade (Maddison and Maddison, 1992) was used to calculate the number of character states unique to particular individuals or groups of individuals in the trees from the parsimony analysis. *Phylca nitida* was chosen as the outgroup for these studies because in the sequence analysis outlined in the previous chapter it was the sister group to the rest of the taxa included in the AFLP analysis.

The binary matrix was converted into a similarity matrix between pairs of individuals using SIMIL in the R package (Legendre and Vaudor, 1991). This was done using Jaccard's coefficient (Jaccard, 1908) in which shared absence is not treated as similarity. This matrix was then used in a PCO analysis also using the R package. Some of the individuals for which only one primer pair was run and which were included in the tree building methods were excluded from the PCO analyses because the R package does not cope with large amounts of missing data. The PCO analyses were also performed on each of the individual species by splitting up the initial binary matrix into a single one for each species.

## 5.5. Results

The AFLP data matrix (Appendix 2) had 347 potentially informative characters out of a total of 745 characters used, i.e. 47% of characters were variable in two or more accessions.

### 5.5.1. UPGMA

In the UPGMA analysis (Fig. 5.6) the ‘paniculata group’ is moderately supported. There is weak support for a group of genotypes containing *P. arborea*, with *P. polifolia* and most of the *P. paniculata* samples within it. *Phyllica polifolia* forms a strongly supported set of genotypes. *Phyllica nitida* from Mauritius and Réunion each form strongly supported groups (apart from one Mauritian individual, YM3) as does *P. paniculata* (apart from one Cape individual, JER162) and the *P. polifolia* Lot population. One individual within the High Hill population appears to be quite distinct (RR31) from the rest of this population. Within *P. arborea*, apart from the Settlement Plain samples which were collected over a relatively wide geographic range, the different populations generally form distinct sets of genotypes although these are not supported. Each of these sets of genotypes is nested in different positions between Settlement Plain individuals. The island population from Nightingale forms a weakly supported group of genotypes. The Gough and New Amsterdam individuals together have no support, but the New Amsterdam individuals themselves form a moderately supported set of genotypes within the Gough individuals. The populations from the Ponds and Big Gulch do not form clearly distinct groups of genotypes. One group of Tristan da Cunha genotypes (JER91, 92, 94 and 97) from the Settlement Plain and Big Gulch forms a weakly supported set of genotypes. This group (along with a further individual, JER62) will be referred to as “hybrid genotypes” because in the PCO analysis they appear to be intermediate between the Tristan da Cunha/Nightingale and Gough/New Amsterdam genotypes and they could have resulted from a reintroduction from Gough to Tristan da Cunha followed by interbreeding resulting in an intermediate genotype.

### 5.5.2. Parsimony

The search produced 120 trees of length 1735 with CI=0.36 and RI=0.62. Figure 5.7 shows one of these trees with Wagner (equal weights, unordered states) branch lengths (ACCTRAN optimisation) indicated by the lengths of the branches, Wagner

bootstrap percentages below and branches that collapse in the strict consensus tree of the Wagner analysis are marked with an arrow.

The 'paniculata group' forms a distinct group of genotypes in all trees, but there was less than 50% bootstrap support. Within the 'island group' *P. nitida* and *P. polifolia* are strongly supported sets of genotypes, *P. arborea* is a weakly supported set, but the *P. paniculata* individuals are not supported. However, each species forms distinct groups of genotypes in the strict consensus.

The Lot population of *P. polifolia* which is phenotypically distinct from the High Hill population forms a group of genotypes with less than 50% bootstrap support. The Réunion and Mauritian individuals of *P. nitida* form two distinct strongly supported sets of genotypes. The *P. paniculata* individuals sampled form a distinct genotypic group with the two individuals from the Magaliesberg forming a strongly supported group.

Within *P. arborea*, apart from the Settlement Plain samples, which were collected over a relatively wide geographical range, the different populations generally form weakly supported but distinct sets of genotypes. The population from Nightingale is distinct as are the Gough and New Amsterdam accessions, with those from New Amsterdam being distinct from those on Gough. The Gough/New Amsterdam cluster is generally well separated from the Tristan da Cunha/Nightingale genotypes. However, one set of Tristan da Cunha genotypes (JER 62, 91, 92, 94 and 97 from the Settlement Plain and Big Gulch) cluster with those from Gough and New Amsterdam in a strong association with the latter genotypes rather than with the others from Tristan da Cunha or Nightingale. The population from the Ponds is only slightly distinct, and genotypes from Big Gulch do not form a distinct set of genotypes.

### 5.5.3. Neighbour Joining

Figure 5.8 shows the tree produced by the neighbour joining analysis. The 'paniculata group' and all individual species form strongly supported groups except for *P. paniculata*. Within-species groups of genotypes are identical to those found in the parsimony analysis with similar levels of support.

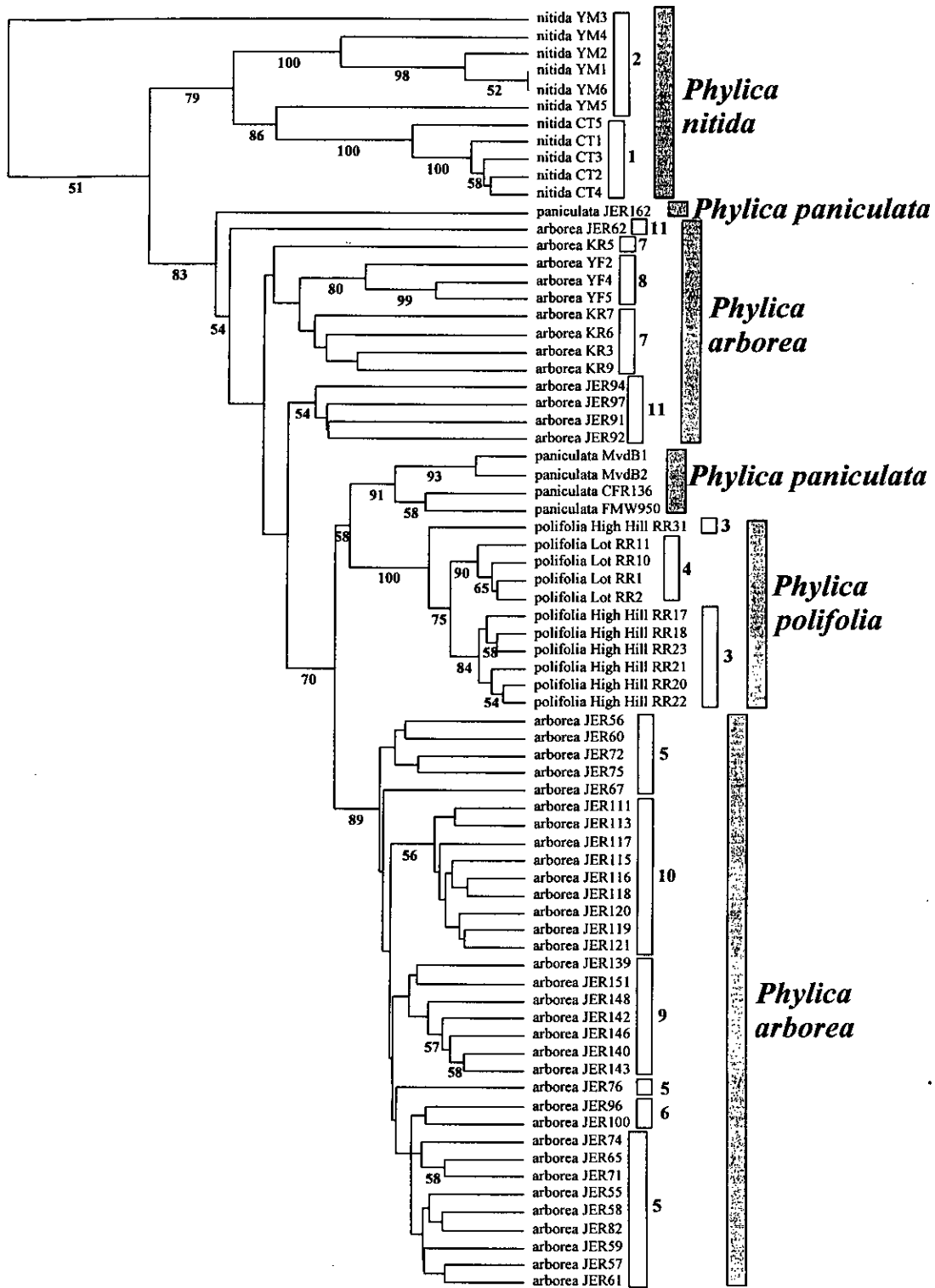


Figure 5.6. UPGMA analysis of the *Phylica* 'island group' with bootstrap percentages shown below branches. Populations: 1. Réunion; 2. Mauritius; 3. High Hill, St Helena; 4. Lot, St Helena; 5. Settlement Plain, Tristan da Cunha; 6. Big Gulch, Tristan da Cunha; 7. Gough; 8. New Amsterdam; 9. The Ponds, Tristan da Cunha; 10. Nightingale; 11. "Hybrid" genotypes.

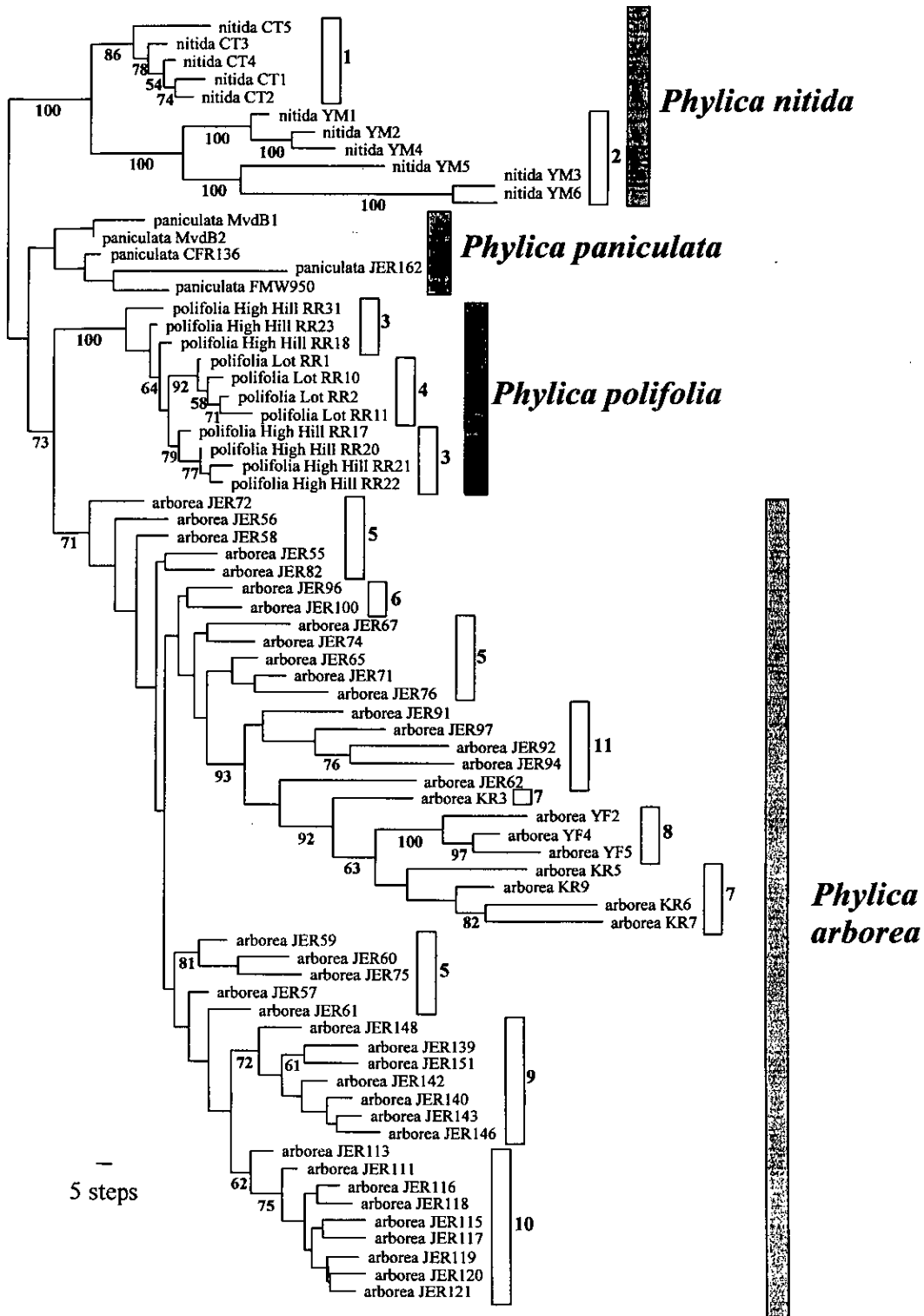


Figure 5.7. Parsimony analysis of the *Phylica* 'island group' showing one of the Fitch trees. The length of a five step branch on the tree is indicated. Note the uneven rates of change between lineages. Bootstrap percentages are indicated below branches. Populations: 1. Réunion; 2. Mauritius; 3. High Hill, St Helena; 4. Lot, St Helena; 5. Settlement Plain, Tristan da Cunha; 6. Big Gulch, Tristan da Cunha; 7. Gough; 8. New Amsterdam; 9. The Ponds, Tristan da Cunha; 10. Nightingale; 11. "Hybrid genotypes".



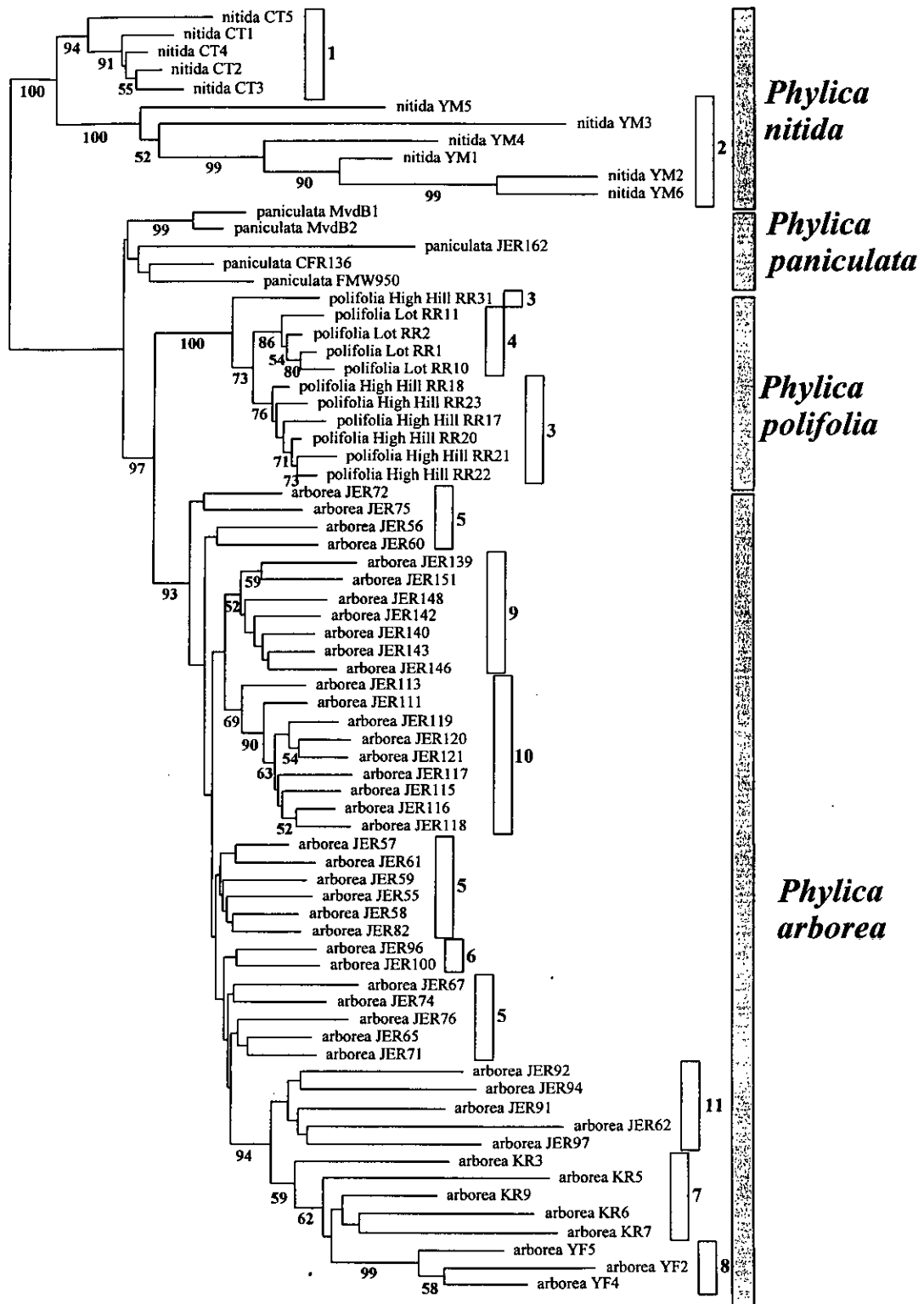


Figure 5.8. Neighbour joining analysis of the *Phylica* 'island group' with bootstrap percentages shown below branches. Note the uneven rates of change between lineages. Populations: 1. Réunion; 2. Mauritius; 3. High Hill, St Helena; 4. Lot, St Helena; 5. Settlement Plain, Tristan da Cunha; 6. Big Gulch, Tristan da Cunha; 7. Gough; 8. New Amsterdam; 9. The Ponds, Tristan da Cunha; 10. Nightingale; 11. "Hybrid"-genotypes.

#### 5.5.4. Principal Co-ordinates Analysis

Figure 5.9 shows the patterns produced by the PCO of the whole AFLP data set. The eigenvalues for all PCO analyses are listed in Table 5.2. Each of the species included in this study are grouped together as distinct sets of genotypes. In the PCO analyses performed for each of the individual species (Figures 5.10, 5.11), the same patterns that were indicated by tree-building methods are revealed. The *P. polifolia* analysis produced two distinct groups of genotypes representing the separate High Hill and Lot populations (Figure 6.2; Chapter Six) with one individual within the High Hill population also appearing to be quite distinct (RR31). The *P. nitida* analysis showed that the Réunion individuals form a tight group of genotypes, which are distinct from the Mauritian individuals which are much more variable in comparison. The *P. arborea* analysis revealed a number of distinct sets of genotypes: one containing the Gough and New Amsterdam individuals, one containing the Nightingale individuals and one containing the rest of the Tristan da Cunha individuals. There is also a group of genotypes from Tristan da Cunha which are intermediate between the Gough and New Amsterdam groups and the other Tristan da Cunha individuals. The eigenvalues and percentage of variance for each analysis (Table 5.2) show a decrease with an increase in taxonomic range. As expected, many of the "distinct" clusters of genotypes shown by the tree-building methods do not have distinct markers and are therefore not separable with PCO analysis. The within-species relationships are in agreement in parsimony, NJ and PCO analyses, but those in the UPGMA tree differ somewhat.

Table 5.2. Eigenvalues for PCO analyses of AFLP data sets.

*Phyllica*

Eigenvalues	% of variance
2.48	19.81
1.64	13.18
1.12	9.01
0.62	5.10

*P. arborea*

Eigenvalues	% of variance
1.02	19.92
0.53	10.56
0.31	6.37
0.26	5.26

*P. nitida*

Eigenvalues	% of variance
0.56	45.90
0.38	30.82
0.16	12.93
0.06	4.74

*P. polifolia*

Eigenvalues	% of variance
0.13	43.83
0.07	24.76
0.03	9.48
0.02	7.30

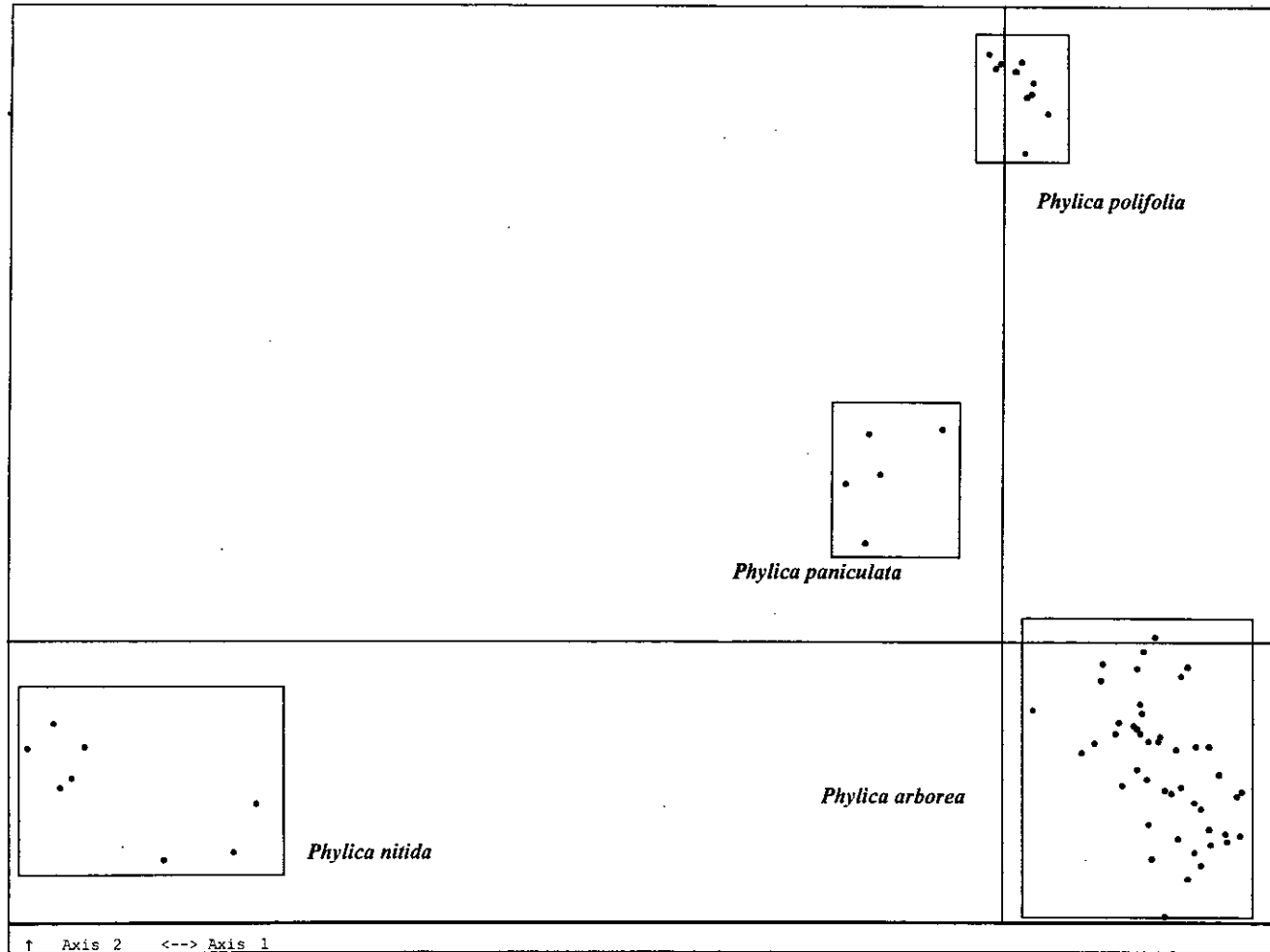


Figure 5.9. Principal co-ordinates analysis of the *Phylica* 'island group'. The percentage of variance is 19.8% in the first axis and 13.2% in the second.

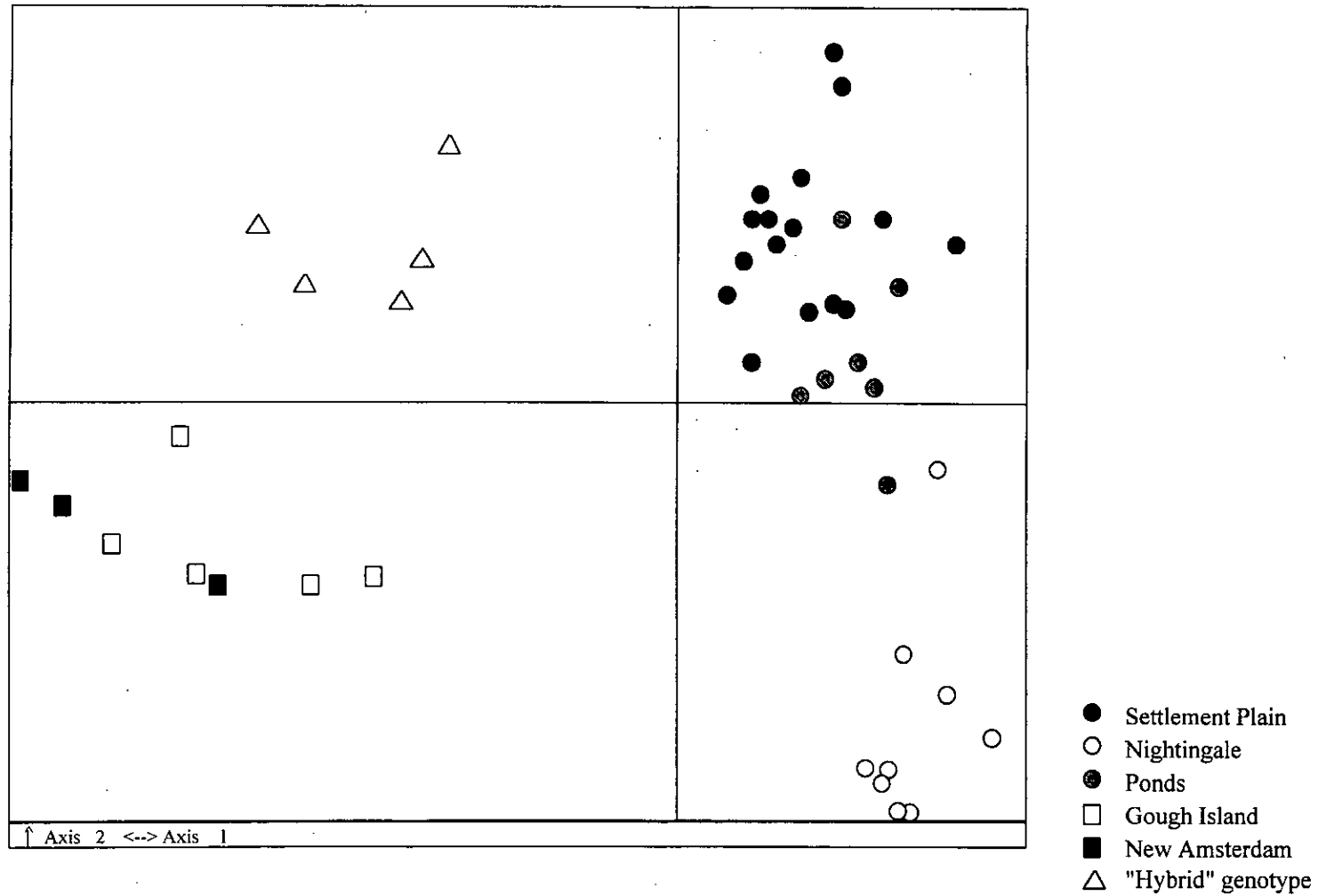


Figure 5.10. Principal co-ordinates analysis of *Phylica arborea*. The percentage of variance is 19.9% in the first axis and 10.6% in the second.

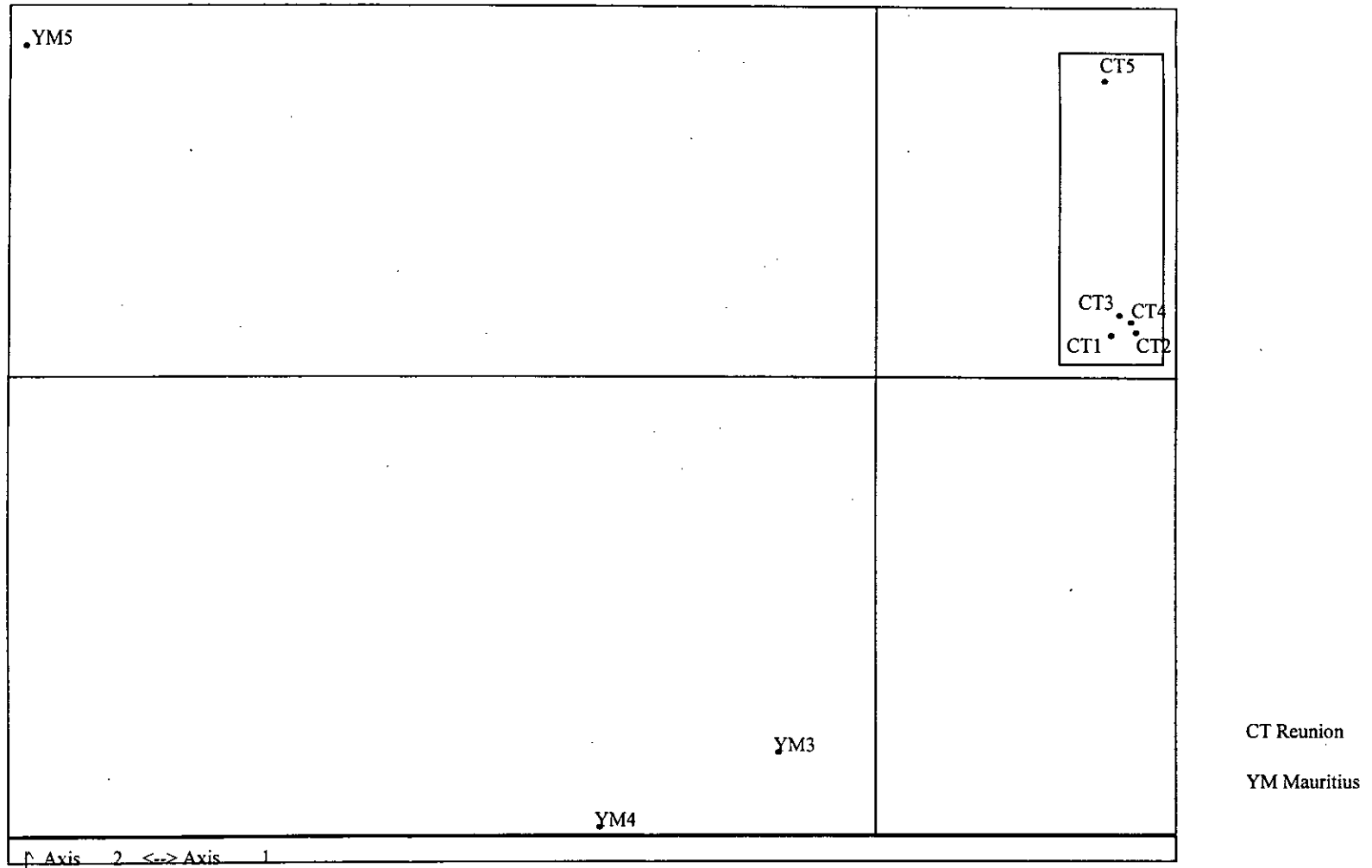


Figure 5.11. Principal co-ordinates analysis of *Phyllica nitida*. The percentage of variance is 45.9% in the first axis and 30.8% in the second.

## 5.6. Discussion

### 5.6.1. Number of species in the 'island group' and the monophyly of these taxa

Because UPGMA assumes a constant rate of evolution in different lineages and does not permit rate heterogeneity, it can produce an incorrect topology if some lineages are evolving faster than others. The main difference between the NJ, parsimony and UPGMA analyses is that with UPGMA *P. polifolia* and *P. paniculata* form strongly supported sets of genotypes that are nested within *P. arborea*. If this UPGMA result were treated as being representative of origin it would indicate that *P. paniculata* and *P. polifolia* were derived from within *P. arborea* and that *P. arborea* is a paraphyletic species. However, the bootstrap support for this situation is weak compared to support for the monophyly of these species in the NJ and parsimony analyses. The PCO analysis of the whole data set also supports the idea that each of the species included in this study has distinctive genotypic markers, which is consistent with each of them being monophyletic.

Given this level of sampling, each of the species should remain taxonomically as they were prior to this study, i.e. *P. polifolia* on St Helena, *P. arborea* on the Tristan da Cunha Group and New Amsterdam, *P. nitida* on Mauritius and Réunion and *P. paniculata* in southern Africa. *Phyllica paniculata* requires better sampling throughout its range because these samples are highly divergent, resulting in incorrect topologies with UPGMA and weak support with the other tree building methods. However, the PCO analysis groups the *P. paniculata* genotypes together as a single distinctive cluster. The eigenvalues and percentage of variance for each analysis (Table 5.2) show the effect of too wide a taxonomic range in the *Phyllica* and *P. arborea* analyses, both of which have low values compared to those in *P. polifolia* and *P. nitida*.

### 5.6.2. Genetic variation within and among populations of island species of *Phyllica* and the possible origins of island species and populations

Given that the monophyly of each of the island species has been established, within-species population genetic architecture, and the evolutionary forces that might have caused this structure can be assessed. These forces might include migration or gene flow, mutation, genetic drift, natural selection, divergence during isolation, assortment and genetic recombination mediated by the mating system. In a panmictic population you might expect low levels of genetic structure within and among

populations as a result of gene flow. Establishment of genetic structure will be the result of differentiation due to geographical isolation, given that island species of *Phyllica* appear to be outbreeding, (Richardson, field observations).

Within *P. nitida* there appears to be a large amount of genetic differentiation between the Mauritian and Réunion populations. This supports the ITS and *trnL-F* sequence data which also indicated that these two populations were quite distinct (Chapter Four). However, increased sampling of other populations from Réunion may indicate a lower degree of genetic differentiation but, with this level of sampling there does not appear to be any evidence for gene flow between the two islands, which is the result of a long period of isolation. These two populations should be considered as subspecies if this level of distinctiveness is maintained with increased sampling.

Within *P. arborea* there appears to be some structure to the samples taken from Nightingale, New Amsterdam, Gough, and at First Pond on Tristan da Cunha. There are morphological characters that might support some of these groups of genotypes. For example, the Nightingale individuals were trees to a height of 7m, with fruits to 9mm across and leaves to about 7mm, i.e. features were generally larger than for any other populations. The First Pond individuals were low growing with smaller fruits and the Settlement individuals were intermediate between First Pond and Nightingale. The structure of populations on Tristan da Cunha, indicated by the tree building methods, could break down with increased sampling around the island and on the other islands. Some of this structure is lost in PCO analyses indicating that there is gene flow between most of these populations. It is also possible that the morphological differences may be the result of different environmental conditions, which is particularly likely on Nightingale where large trees grow in a sheltered area. The soil on Nightingale may also be richer due to the guano produced by the larger bird populations on the island compared to Tristan da Cunha. Controlled growth of different forms of *P. arborea* is necessary to determine whether these morphological differences are genetic or merely the result of phenotypic plasticity. Genotypically the Nightingale population appears reasonably distinct from the Tristan da Cunha population, with low to moderate support in the tree building methods, although one Nightingale individual has a genotype which is similar to some of the Tristan da Cunha genotypes indicating that there is still a certain amount of gene flow, most likely by seed dispersal, between the two islands. The Gough/New Amsterdam genotypes are distinct from the rest of the Tristan da Cunha Group (with moderate support in NJ, strong support in parsimony analyses and highly isolated with PCO) although there are no clear morphological differences between these two groups.



According to the tree building methods the New Amsterdam population is distinct from the Gough population, with strong support, indicating a lack of gene flow between these two islands which you might expect given the large distance between them. The divergent genotypes indicate a long period of isolation between the Gough/New Amsterdam and Tristan da Cunha/Nightingale populations. This can be contrasted with the lack of structure within the continuous population on the Settlement Plain, which might be due to a simple case of nearest-neighbour interbreeding, i.e. the likelihood of breeding between individuals decreases with distance.

There is a group of Tristan da Cunha individuals that seem to have genotypic similarities with Gough. Samples JER62, 91, 92, 95 and 97 often form a weakly supported group of genotypes in the tree building methods and also form a distinct group in the PCO analysis. These individuals, which are "intermediate" between the Gough and Tristan da Cunha groups of genotypes in the PCO analysis, are most likely the result of a re-introduction of Gough genotypes and subsequent "hybridisation" with Tristan da Cunha genotypes.

Within Tristan da Cunha populations there appears to be gene flow across the island from the Settlement plain to Big Gulch as Settlement genotypes are found at Big Gulch and vice versa. There does not appear to be any genetic differentiation within the Settlement Plain population indicating that there is gene flow. The Ponds population is found on the base of the island as opposed to the cliffs where most of the other populations were sampled. The tree building methods give some support to genetic differentiation indicating the absence of gene flow between this population and others on Tristan da Cunha. However, increased sampling between the Ponds and Settlement Plain populations may result in a breakdown in this structure.

Within *P. polifolia* there appears to be some genetic differentiation between the Lot and High Hill populations sampled. These relationships are discussed further in Chapter Six.

The number of *P. paniculata* samples are not extensive enough to make any definite conclusions about population genetic structure within this species. However, there is abundant differentiation between the individuals studied as evidenced by the long branches between them. This indicates a lack of gene flow, which would be expected if isolated populations were sampled over a wide geographic range, as is the case with *P. paniculata*. It is a mountain-dwelling species, and its populations do not form a continuous distribution. Increased sampling might however indicate a lower degree of differentiation.

There are various possibilities concerning the origin of the island species. The sequence data (Chapter Four) did not address species distinctions or relationships due to the low level of variability detected. The sequence results only showed that *P. nitida* diverged some time before the other island species, and the AFLP results are consistent with this. The results of the neighbour joining, parsimony and PCO analyses indicate that *P. paniculata*, *P. polifolia* and *P. arborea* were derived independently from a common ancestor on the mainland. Sampling within *P. paniculata* was perhaps not great enough to draw any conclusions about whether some of the island species were derived from different populations of this species. The derivation of different island species from different populations of *P. paniculata* appears unlikely. The problem of the putative paraphyly of *P. paniculata* or its ancestor cannot be properly addressed here because of its long period of isolation and gene flow among populations subsequent to the dispersal of *P. polifolia* and *P. arborea*, which would make this species appear monophyletic, even though it may not have been. Independent assortment would thus be expected to remove evidence of paraphyly from the nuclear genome. Only uniparentally inherited genomes might be expected to still exhibit evidence of paraphyly, but this would also be difficult to separate from differential inheritance of ancestral polymorphism in *P. paniculata*.

Within species, current results are consistent with successive colonisations from older to younger islands. The Réunion population of *P. nitida* could have been derived from an introduction from a population on the older island of Mauritius. The greater genetic diversity on the Tristan da Cunha Group is consistent with the hypothesis that the original introduction of *P. arborea* (or its ancestor) was to this archipelago, although the relatively sparse sampling on Gough and New Amsterdam precludes saying this with certainty. The introduction to the Tristan da Cunha Group could have been followed by a single introduction to New Amsterdam. All analyses are consistent with a single founder event on New Amsterdam from Gough, and Gough may have only been colonised once (there may have been more events but with this level of sampling there is no evidence for this). This is again consistent with successive colonisation from older to younger islands (assuming that the first colonisation of the Tristan da Cunha Group was on Nightingale which is the oldest island in the archipelago). The estimated time of dispersal is half a million years ago (Chapter Four) so the original founding event could in fact have been on either island (Nightingale is *c.* 18 mya and Tristan da Cunha is *c.* one million years old). There are three unique AFLP bands found in the New Amsterdam genotypes, which lends support to the hypothesis of a single origin for this population. No further gene flow occurred after founder events from Tristan da Cunha to Gough and New Amsterdam

until recently with the possible reintroduction to Tristan da Cunha from Gough which has resulted in "hybrid" genotypes (JER62, 91, 92, 95 and 97) that cluster between the Gough/New Amsterdam and Tristan da Cunha genotypes with PCO (Figure 5.10).

There are problems with the use of extant plants to determine the genetic origin of populations. If we take the putative single introduction to New Amsterdam as an example, it could be hypothesised that the New Amsterdam population was once more significant in terms of numbers and genetic diversity and has recently contracted. It is possible that if the original degree of variation were still present we would have seen a different pattern indicating that the Tristan da Cunha Group populations of *P. arborea* arose from a single or a few founder events from New Amsterdam. This possibility should also be taken into account when making suggestions about other possible founder events within the 'island group'. For example, the two groups of genotypes (Gough/New Amsterdam and Tristan da Cunha/Nightingale) could be due to two separate colonisation events from different source populations. The lack of knowledge about the extent of past variation restricts the ability to make definite conclusions about the origins of populations or species. However, the patterns obtained are consistent with the original population being on Tristan da Cunha and Nightingale as indicated by the greater diversity of genotypes, an early single introduction to Gough, after which isolation of the two groups resulted in the production of distinct genotypes. Following this, a further dispersal from Gough to New Amsterdam occurred, and a recent reintroduction from Gough back to Tristan da Cunha, perhaps with some "hybridisation" between the Gough and Tristan da Cunha genotypes. The hypothesised relationships between island species and populations and the estimated sequence and timings of dispersals are presented in Figure 5.12. This figure provides putative answers to some of the questions left unresolved in Chapter Four. Further evidence for the origin of island species populations could be obtained by looking at other molecular markers such as plastid cytotypes. For example, if the Gough/New Amsterdam cytotype were distinct from that of Tristan da Cunha/Nightingale and this cytotype was found in the "hybrid" populations on Tristan da Cunha, this would be further evidence for the direction of dispersal postulated here. This AFLP study has indicated potential patterns of dispersal and could be used to direct further areas of study using alternative markers.

### 5.6.3. Dispersal of *Phyllica* island species.

There are three possibilities for the mode of dispersal of *Phyllica* to oceanic islands around southern Africa.

#### 1. Human Dispersal

*Phyllica arborea* was noted on the original visits to both Tristan da Cunha and New Amsterdam discounting the possibility that the initial introduction of seeds may have been due to human activity.

#### 2. Ocean current dispersal

Cronk (1987) suggested that *Phyllica* could have been transported to St Helena by currents (south by the Agulhas current and north from Cape Agulhas by the Benguela current). The possibility of ocean current transport may be eliminated by exposing *Phyllica* fruits to seawater for a period longer than would be necessary for a capsule to make the journey from New Amsterdam to the Tristan da Cunha Group or vice versa. The distance between Tristan da Cunha and New Amsterdam is *c.* 7250 km. West wind drift has a movement of 13 km/day. The minimum time taken to travel the distance is therefore  $7250/13 = c. 560$  days. Seawater temperatures around Tristan are 11-13°C in winter and 13-18°C in summer. Germination experiments were set up to see if *Phyllica* fruits could withstand this length of time in seawater at roughly comparable temperatures. This involved the setting up of a control germination and the submerging of *P. arborea* fruits in seawater at *c.* 15°C for a period of 560 days or longer and testing for germination. The results of these germination experiments are not yet available. These fruits however have none of the traits (e.g. indehiscent capsules, good protection by thick ovary or seed coat walls) found in other sea-dispersed taxa (e.g. *Crinum*, *Cocos*), so they appear unlikely to be thus dispersed. Even if capsules could be transported by sea, *Phyllica* species are not plants of the strand, and so the mode of dispersal lacks a way of getting into their preferred sites away from beaches.

#### 3. Bird dispersal

In a report of an expedition to Tristan da Cunha following a volcanic eruption in 1962 Dickson (1965) stated that *P. arborea* berries (actually capsules) are adapted to

bird dispersal and that they are eaten by native land birds or are found in their stomach contents. Hagen (1952) noted that four breeding species or subspecies of sea birds which have not been found breeding in any other part of the world are common to the Tristan da Cunha Group and the New Amsterdam-St Paul group. Birds which frequent both these islands include the yellow-nosed albatross (*Diomedea chlororhynchos*), for which New Amsterdam, St Paul and Prince Edward islands in the southern Indian Ocean together with the Tristan da Cunha archipelago are the principal breeding grounds. Individual birds from the two groups of islands do intermingle, and it is possible that these birds may have been responsible for movement of seed between these islands. This distribution is shared by three flowering plants, *P. arborea*, *Spartina arundinacea* and *Uncinia brevicaulis* var. *rigida*, and floristic links between the two islands are reasonably strong. Christophersen (1937) pointed out that sea birds do not eat fruits and only approach land to breed and that the Tristan da Cunha Group is not on the migration route for any land birds. Furthermore, the time to travel between islands exceeds the time taken for diaspores to be excreted. Taking into account these two observations, it seems unlikely that seed was transported between the Tristan da Cunha Group and New Amsterdam by internal bird dispersal. It is possible however that land birds may have eaten fruits and deposited seed near the nesting sites of sea birds. These seeds could then have been attached to the feet of sea birds and transported externally.

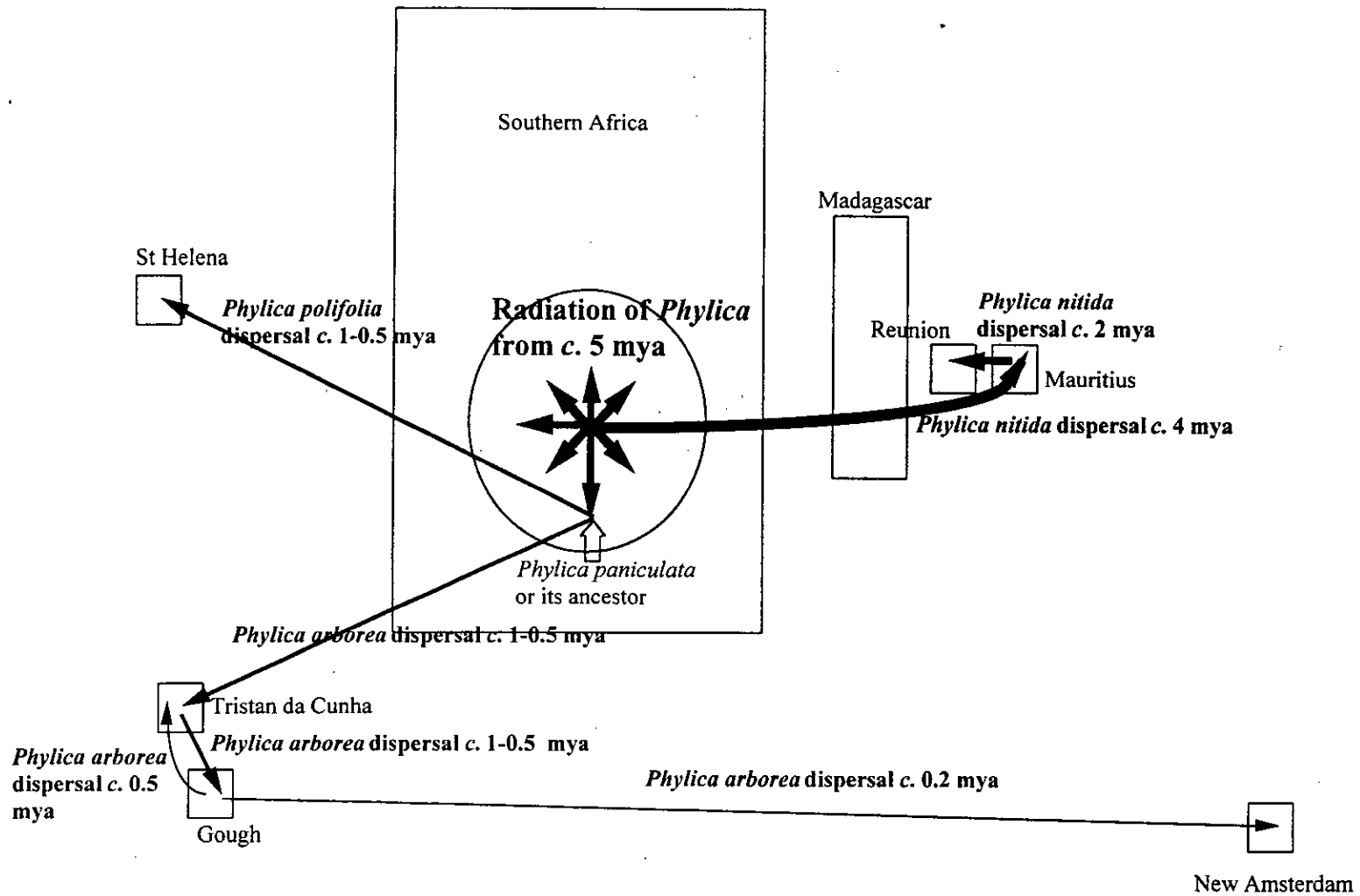


Figure 5.12. Timing of dispersal of island populations of *Phylica* based on sequence and AFLP data.

## 5.7. Conclusions

This study has shown that the island species of *Phylica* form distinct groups, i.e. they are distinct species. The AFLP data also support what was indicated by the sequence analysis in the previous chapter, i.e. *P. nitida* diverged some time before the other island species. Each of the species, *P. arborea*, *P. polifolia* and *P. paniculata*, have been independently derived, probably from a 'paniculata-like' African ancestor.

AFLPs were also useful in elucidating within-species relationships. Gene flow, as would be expected, appears to be more frequent within populations on the same island than among populations on different islands. From the AFLP and sequence data (Chapter Four) it appears that *P. nitida* on Réunion could have been derived from *P. nitida* on Mauritius. The AFLP data also indicate that populations of *P. arborea* on New Amsterdam could have been derived from a single introduction from Gough Island and that the Gough Island population could have been derived from one or more introductions from Tristan da Cunha or from the early dispersal of *P. arborea* to both sites independently. These results are to an extent compatible with the ages of the islands with populations from older islands generally colonising younger ones. There also appears to have been a recent re-colonisation of Tristan da Cunha from Gough and subsequent inter-breeding resulting in genotypes on Tristan da Cunha which are intermediate between those otherwise occupying these two islands.

The results produced by the parsimony analysis are similar to those produced by Neighbour Joining and PCO analyses. The UPGMA analysis produced a different result, but this method is often considered to be unreliable because it does not take into account rate heterogeneity, which is clearly evident in the NJ and parsimony results. The fact that all the methods used produced broadly similar results indicates that there are reasonably clear patterns in this data set. An increased level of sampling of some populations, particularly of *P. paniculata* and Gough/New Amsterdam accessions, and the use of other molecular markers (such as plastid RFLPS or microsatellites) would permit making firmer conclusions.

## 5.8. Bibliography

Applied Biosystems Plant Mapping Protocol (part number 402083, Revision B, May 1996).

- Breytenbach, G.J. 1986. Dispersal: The case of the missing ant and the introduced mouse. *South African Journal of Botany* 52: 463-466.
- Carmichael, P. 1818. Some account of the island of Tristan da Cunha and its natural productions. *Transactions of the Linnean Society of London* 12: 483-513.
- Christopherson, E. 1937. Plants of Tristan da Cunha, Scientific results of the Norwegian Antarctic expeditions 1927-28. I Kommissjon hos Jacob Dybwad, Oslo.
- Cronk, Q.C.B. 1987. The history of endemic flora of St Helena: a relictual series. *New Phytologist* 109: 509-520.
- Daycard, L. 1985. Situation Ecologique de l'Île Amsterdam: Bilan et Avenir. *Memoire de D.E.A., University of Montpellier*.
- Dickson, J.H. 1965. The biology of the Tristan da Cunha Islands. *Philosophical Transactions of the Royal Society of London* 249: 259-271.
- Doyle, J.J. & J.L. Doyle. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11-15.
- Gower, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 53: 325-338.
- Groves, E.W. 1981. Vascular plant collections from the Tristan da Cunha group of islands. *Bulletin of the British Museum (Natural History)* 8: 333-420.
- Guého, J. 1976. Sur l'identité du *Phyllica* (Rhamnaceae) des îles Mascareignes, *Adansonia séries 2*, 15: 509-513.
- Hagen, Y. 1952. Birds of Tristan da Cunha. Results of the Norwegian Scientific Expedition to Tristan da Cunha, 1937-38, No. 20.
- Hemsley, W.B. 1873-76. Report on the Scientific Results of the Exploring Voyage of H.M.S. Challenger, Botany volume 1.



Hooker, J.D. 1875. On the discovery of *Phylica arborea*, Thouars a tree of Tristan D'Acunha, in Amsterdam Island in the S. Indian Ocean; with an enumeration of the phanerogams and vascular cryptogams of that Island and of St. Paul. Botanical Journal of the Linnean Society 14: 474-479.

Jaccard, P. 1908. Nouvelles recherches sur la distribution florale. Bulletin de la Société Vaud. Sciences Naturelles 44: 223-270.

Juan, C., K.M. Ibrahim, P. Oromi & G.M. Hewitt. 1996. Mitochondrial DNA sequence variation and phylogeography of *Pimelia* darkling beetles on the island of Tenerife (Canary Islands). Heredity 77: 589-598

Juan, C., K.M. Ibrahim, P. Oromi & G.M. Hewitt. 1998. The phylogeography of the darkling beetle, *Hegeter politus*, in the eastern Canary Islands. Proceedings of the Royal Society of London B 265: 135-140.

Juan, C., P. Oromi & G.M. Hewitt. 1995. Mitochondrial DNA phylogeny and sequential colonization of Canary Islands by darkling beetles of the genus *Pimelia* (Tenebrionidae). Proceedings of the Royal Society of London 261: 173-180.

Juan, C., P. Oromi & G.M. Hewitt. 1996. Phylogeny of the genus *Hegeter* (Tenebrionidae, Coleoptera) and its colonization of the Canary Islands deduced from cytochrome oxidase I mitochondrial DNA sequences. Heredity 76: 392-403.

Lamarck, J.B.P.A. 1797. Encyclopédie Méthodique Botanique, Tableau Encyclopedique 2: 77.

Legendre, P. & A. Vaudor. 1991. The R Package: Multidimensional analysis, spatial analysis. Department of Biological Sciences, University of Montreal.

Maddison, W.P. & D.R. Maddison. 1992. MacClade version 3.04. Sinaeur Associates Inc., Sunderland, Massachusetts.

Micol, T. & P. Jouventin. 1995. Restoration of Amsterdam Island, South Indian Ocean, following control of feral cattle. Biological Conservation 73: 199-206.

Milton, S.J., P.G. Ryan, C.L. Moloney, J. Cooper, W.R.J. Dean & A.C. Medeiros Jr. 1993. Disturbance and demography of *Phylica arborea* (Rhamnaceae) on the Tristan-Gough group of islands. *Botanical Journal of the Linnean Society* 111: 55-70.

Moseley, H.N. 1875. Plants from Tristan D'Acunha. *Botanical Journal of the Linnean Society* 14: 377-384.

Pillans, N.S. 1942. The genus *Phylica* Linn. *Journal of South African Botany* 8: 1-164.

Procter, J. & R. Salm. 1975. Conservation on Mauritius (1974). IUCN, Morges, Switzerland.

Roux, J.P., P.G. Ryan, S.J. Milton & C.L. Moloney. 1992. Vegetation and checklist of Inaccessible Island, central South Atlantic Ocean, with notes on Nightingale Island. *Bothalia* 22: 93-109.

Rue, A. de la. 1932. La flore et la faune des Iles Saint-Paul et Amsterdam. *Terre et la Vie, Revue d'Histoire Naturelle* 2: 642-662. Paris.

Ryan, P.G., C.L. Maloney & B.P. Watkins. 1989. Concern about the adverse affects of introduced mice on island tree *Phylica arborea* regeneration. *South African Journal of Science* 85: 626-627.

Sokal, R.R. & P.H.A. Sneath. 1963. *Numerical Taxonomy*. Freeman & Co., San Francisco.

Swofford, D.L. 1998. PAUP: phylogenetic analysis using parsimony (and other methods), version 4.0d64. Sinaeur Associates, Sunderland, Massachusetts.

Thouars, A.D.P. 1811. *Flore de Tristan d'Acugna*. Paris.

Trehan, P., Y. Frenot, M. Lebouvier & P. Vernon. 1990. Invertebrate fauna and their role in the degradation of cattle dung at Amsterdam Island. *In: Kerry, K.R. & G. Hemple [eds.], Antarctic Ecosystems, Ecological Change and Conservation*. Springer Verlag, Berlin, Heidelberg.

- Tulasne, L.R. 1857. *Florae Madagascariensis fragmenta*. *Annales des Sciences Naturelles*, series 4, 8: 128.
- Vahl, M. 1794. *Symbolae Botanicae* 3: 41. Nicolaus Möller & Son.
- Valentyn, F. 1726. *Oud en Nieuw Oost Indien Verhandeling der ZeeHorenkens Banda*. Dordrecht, Amsterdam.
- Vaughan, R.E. & P.O. Wiehe. 1937. Studies on the vegetation of Mauritius, 1: A preliminary survey of the plant communities. *Journal of Ecology* 25: 289-343.
- Velain, C. 1893. Les Illes Saint-Paul et Amsterdam. *Annales de Géographie* 2: 329-364.
- Von Pelzeln, A. 1861. *Reise der Osterreichischen Fregatte 'Novarr' um die Erde in der Jahren 1857, 1858, 1859, tome 1*. Vienna.
- Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, N. Hornes, A. Freitjers, J. Pot, J. Peleman, M. Kuiper & M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407-4414.
- Wace, N.M. & M.W. Holdgate. 1958. The vegetation of Tristan da Cunha. *Journal of Ecology* 46: 593-620.
- Wace, N.M. 1961. The vegetation of Gough Island. *Ecological Monographs* 31: 337-367.
- White, F. 1983. *The Vegetation of Africa: a descriptive memoir to accompany the Unesco/AETFAT/UNSO vegetation map of Africa*.
- Willdenow, C.W. 1798. *Species Plantarum* 1: 1112. Berlin.

**CHAPTER SIX. CONSERVATION  
GENETICS OF THREATENED ST  
HELENAN SPECIES OF RHAMNACEAE**

## **CHAPTER SIX. Conservation Genetics Of Threatened St Helenan Species Of Rhamnaceae**

### **Abstract**

Amplified fragment length polymorphisms (AFLPs) were used to determine levels of genetic variability in two endangered endemic species of Rhamnaceae. No AFLP variation was detected in the four remaining individuals of *Nesiota elliptica* indicating that it is effectively clonal. This was contrasted with polymorphism detected between populations and among individuals of *Phyllica polifolia*. AFLP polymorphism was found to be congruent with phenotypic differences between two of the remaining wild populations of *P. polifolia*. It is recommended that seed orchards of these two populations should be kept separately as mixing might disrupt the adaptation of these individuals to their particular habitats. AFLP data have thus proved to be useful for developing appropriate conservation strategies for these species.

### **6.1. Introduction**

Because some of the taxa in this study are extinct in the wild or endangered (*N. elliptica* and *P. polifolia*) I wanted to ascertain the degree of genetic variability within species since this kind of information would be useful in the development of appropriate conservation strategies. Small islands are often characterised by high levels of environmental degradation and species extinction. On Atlantic islands and the Mascarenes these developments date back to European colonial settlement. Severe environmental degradation has taken place on St Helena, and similar problems are also apparent on the Mascarene islands (Mauritius and Réunion). In 1659 the Dutch East India Company settled St Helena and since then environmental degradation has been caused by unmanaged populations of feral livestock, clearing of vegetation to provide crop land and pastures for smallholdings and estates, felling of trees for tanning and timber for small-scale industry, sudden and significant fluctuations of population associated with temporary garrisons, merchant fleets,

introduction of invasive plant species as crops and ornamentals, introduced insect pests, erosion-prone volcanic soils, and modified soil processes resulting from forest clearance and possibly the loss of nesting seabird colonies (Cronk, 1989; Maunder *et al.*, 1995). Surviving populations of endemics are subject to continued threats from inbreeding, stochastic events and invasives/pathogens. All 40 endemic plant species on St Helena are rare or threatened. The St Helenan species included in this study reflect the generally poor state of the endemic flora of the island. The current demographic status of these species is discussed below.

#### 6.1.1. *Nesiota elliptica*

*Nesiota elliptica* (Roxb.) Hook.f. from St Helena is known on the island as the St Helena Olive. It is a small tree, once known from localised populations on the highest parts of the eastern central ridge. This very restricted area represents the only suitable habitat for *N. elliptica* and indicates that the range and population size of this species have probably always been restricted (Cronk, pers. comm.). It became noticeably rare in the nineteenth century, and Melliss (1875) found no more than 12-15 plants in existence in tree fern thicket (*Dicksonia arborescens*) along the central ridge between 700 and 820m on the northern side of Diana's Peak. This species was presumed extinct until 1977 when George Benjamin discovered a single tree near Diana's Peak (Cronk, 1987) on a precipitous cliff. The locality is indicated in Figure 6.1. It was not listed in the IUCN Red Data book (Lucas and Synge, 1978) because it had only just been rediscovered. The plant was healthy in 1980 (Cronk, pers. comm.) with no evidence of fungal infection. The last remaining wild tree died in 1994 and it is therefore given the status EW, i.e. extinct in the wild (Oldfield *et al.*, 1998).

Its status was also evaluated by Jackson (1991; 1994). At the outset of this project there were a total of four individuals *ex situ*: three at Pouncey's and one at the Agriculture and Forestry Department at Scotland, St Helena. A strong self-incompatibility mechanism means that few viable seeds have been set despite many hand pollinations and propagation is extremely difficult. Only one cutting has ever been successfully rooted and attempts at micropropagation have proved unsuccessful due to systemic fungal contamination with 14 species of fungi being isolated from

the wild plant (Fay, 1989). These fungal infections may have resulted from recent introductions (Cronk, pers. comm.). The single successful cutting grew to 2m high at Scotland. It was suffering from a fungal infection and died in 1997. A study of the genetic diversity of this species was considered desirable to assess its conservation genetic status. The three remaining plants and the now dead last wild tree and cutting were included in this analysis (Table 5.1, Chapter Five; RRNes1 to RRNes4 and MWC500). Sample RRNes1 was the cutting derived from the last wild tree, RRNes2 to 4 are seedlings derived from the same tree and sample MWC500 was derived from the original wild tree. Cuttings would be expected to be identical to the wild tree, but seedlings should have some variation due to segregation at heterozygotic loci.

### 6.1.2. *Phylica polifolia*

*Phylica polifolia* (Vahl) Pillans is endemic to St Helena where the common name is wild rosemary. Melliss (1875) described it as occurring at Fairyland, Plantation, Rosemary Hall, Oaklands, Oakbank and Lot, with only 100 plants remaining. Kerr (1970) described it as being extremely rare and in danger of extinction. He described only one old tree several metres tall with a good thick trunk at Blue Hill and one planted in a hedge at Scotland (St Helena). Oldfield *et al.* (1998) have given it a CR C2a status which is defined by IUCN (1994) as critically endangered with total numbers being small and declining, and with either fragmented or localised populations, with a total population estimated to number less than 250 mature individuals and a continuing decline in numbers of mature individuals, observed, projected, or inferred, and also with a severely fragmented population structure (i.e. no subpopulation estimated to contain more than 50 mature individuals). Walter and Gillet (1998) list *P. polifolia* as endangered. The last tree form died more than 20 years ago at Blue Hill. Plants now only occur in dry locations on cliffs. Although there may be up to 100 plants, their distribution is fragmented, and they are vulnerable to competition from introduced plants.

In the wild there are about 50 recorded plants remaining (High Hill, three clumps; Lot, c. six plants; Man's Head 12 plants; cliffs between Distant Cottage and Asse's Ears, one plant). Plants held *ex situ* include two plants at High Peak, plants at St

Paul's school and material at RBG, Kew. The threats and problems to this species include possible genetic depauperacy and loss of major habitat sites. The species was previously known as a large shrub with stems to three metres, but plants today tend to form sprawling bushes (Cronk, pers. comm.). This could be the result of a severe genetic bottleneck, with the remaining individuals all representing cliff ecotypes or alternatively the species may naturally have this habit when young. Originally *P. polifolia* grew in an association with dry or moist gumwood forests at altitudes of 500-650m (Cronk, 1989).

Material for this study was collected by Rebecca Rowe (Table 5.1, Chapter Five; RRA and B and RR1-32) and included samples from High Hill and Lot (Figure 6.1). The High Hill plants were collected on a south-east facing cliff face from a population of 27 plants growing in three main clumps on the cliff face. All plants were in highly branched, interwoven canopies and prostrate growth forms down the cliff face. Samples from Lot were collected from a population of about 6 plants on a south facing cliff face. The plants were large and shrubby with a spread of 1-3m. There are phenotypic differences between the High Hill and Lot populations (Rowe, pers. comm.) with the Lot individuals having a more upright growth form than the prostrate High Hill individuals. I wanted to determine whether these differences were reflected in the genetic data. Reintroduction of individuals into areas to which they are not adapted could lead to an unnecessary loss of material, and therefore seed orchards from the two populations may be best kept separately.





Figure 6.1. St Helena. Black spots indicate sites of remaining populations of *Phylica polifolia* and the site of the last tree of *Nesiota elliptica*. After Cronk (1984).

### 6.1.3. Conserving rare plants – genetic variability and species viability

Species which have experienced a reduction in numbers may be at risk due to demographic, genetic and environmental factors (Schaeffer, 1981). Genetic variation is necessary to maintain adaptive potential and populations lacking genetic variability are therefore more likely to become extinct (Beardmore, 1983; Lande and Barrowclough, 1987; Simberloff, 1988; Salwasser, 1990; Bawa and Ashton, 1991). Genetic variation may be lost from small populations by inbreeding and genetic drift (random changes in gene frequencies that occur due to sampling error, including the loss of alleles; Beardmore, 1983; Simberloff, 1988) and deleterious alleles may become fixed (Wright, 1931). However, there are examples of healthy populations that have low levels of genetic variability as measured by isozyme electrophoretic

studies. For example *Ipomoea purpurea*, introduced to the eastern United States, and *Xanthium strumarium*, are weedy species which show a large amount of phenotypic variation but no detectable electrophoretic variation (Clegg and Brown, 1983). However, species that have been drastically reduced in population numbers recently will be more vulnerable to inbreeding depression and loss of genetic diversity than those species which have larger numbers or have historically maintained small populations (Soulé, 1983; Lande and Barrowclough, 1987). Determination of the structure of genetic variability is important in conservation, and evolutionary history, breeding system, ecology and demography all shape this structure and it should be interpreted with these factors in mind (Holsinger and Gottlieb, 1991; Brown and Schoen, 1992). This kind of information is rarely available, leading to unsuccessful attempts to reinstate species that have become rare for unknown biological reasons (Falk and Olwell, 1992). This study is aimed at adding knowledge of levels of genetic variability to existing knowledge of evolutionary history (see Chapter Four). Information on demography, breeding system and ecology is now needed to determine a more successful approach to conservation of rare St Helenan species of Rhamnaceae.

#### 6.1.4. Examples of the use of AFLPs in conservation genetics

Amplified fragment length polymorphisms (AFLPs) have been used to obtain information on levels of genetic diversity in a number of rare or endangered plants, e.g. *Astragalus cremnophylax* var. *cremnophylax* (Leguminosae; Travis *et al.*, 1996), *Populus nigra* subsp. *betulifolia* (Salicaceae; Winfield *et al.*, 1998), *Isoetes* (Isoetaceae; Hoot *et al.*, 1998), *Orchis simia* (Orchidaceae; Qamaraz-Zaman *et al.*, 1998) and *Populus euphratica* (Salicaceae; Fay *et al.*, in press). This technique is efficient at revealing diversity at and below the species level. For example in a study of *Lactuca* (Compositae) Hill *et al.* (1996) distinguished between previously established taxonomic units at both species and cultivar levels.

## 6.2. Aims of Study

1. To determine the level of genetic diversity within island species particularly those that are rare or endangered (*Nesiota elliptica* and *Phyllica polifolia*).
2. To use AFLP data to help determine conservation management strategies for endangered species of Rhamnaceae on St Helena.

## 6.3. Methods

The individuals used in this study are indicated in Chapter Five, Table 5.1. The protocols for the production and analysis of AFLP data sets are also detailed in Chapter Five. AFLP characters from the *P. polifolia* individuals were subjected to PCO and neighbour joining analyses.

## 6.4. Results

Samples of AFLP profiles for fragments sized between 50 and 100 base pairs from *N. elliptica* and *P. polifolia* are shown in Figures 6.2 and 6.4 respectively. The three seedlings and the cutting derived plant of *N. elliptica* had indistinguishable AFLP profiles throughout the 50-500 bp range of fragment sizes with a total of 80 bands being scored. Figure 6.3 shows AFLP profiles of 100-180 bp fragments from these four plants of *N. elliptica* and the last, now dead, wild tree. The lack of variability in *N. elliptica* can be compared with polymorphism detected within and between the two populations of *P. polifolia* in which a total of 112 bands were scored throughout the 50-500 bp range. The results of a PCO analysis on the *P. polifolia* data set are shown in Figure 6.5. The High Hill population is considerably more diverse than that at Lot and with the exception of one sample (RR31) they are well differentiated. A tree taken from the overall *Phyllica* neighbour joining analysis (Chapter Five) is shown in Figure 6.6.

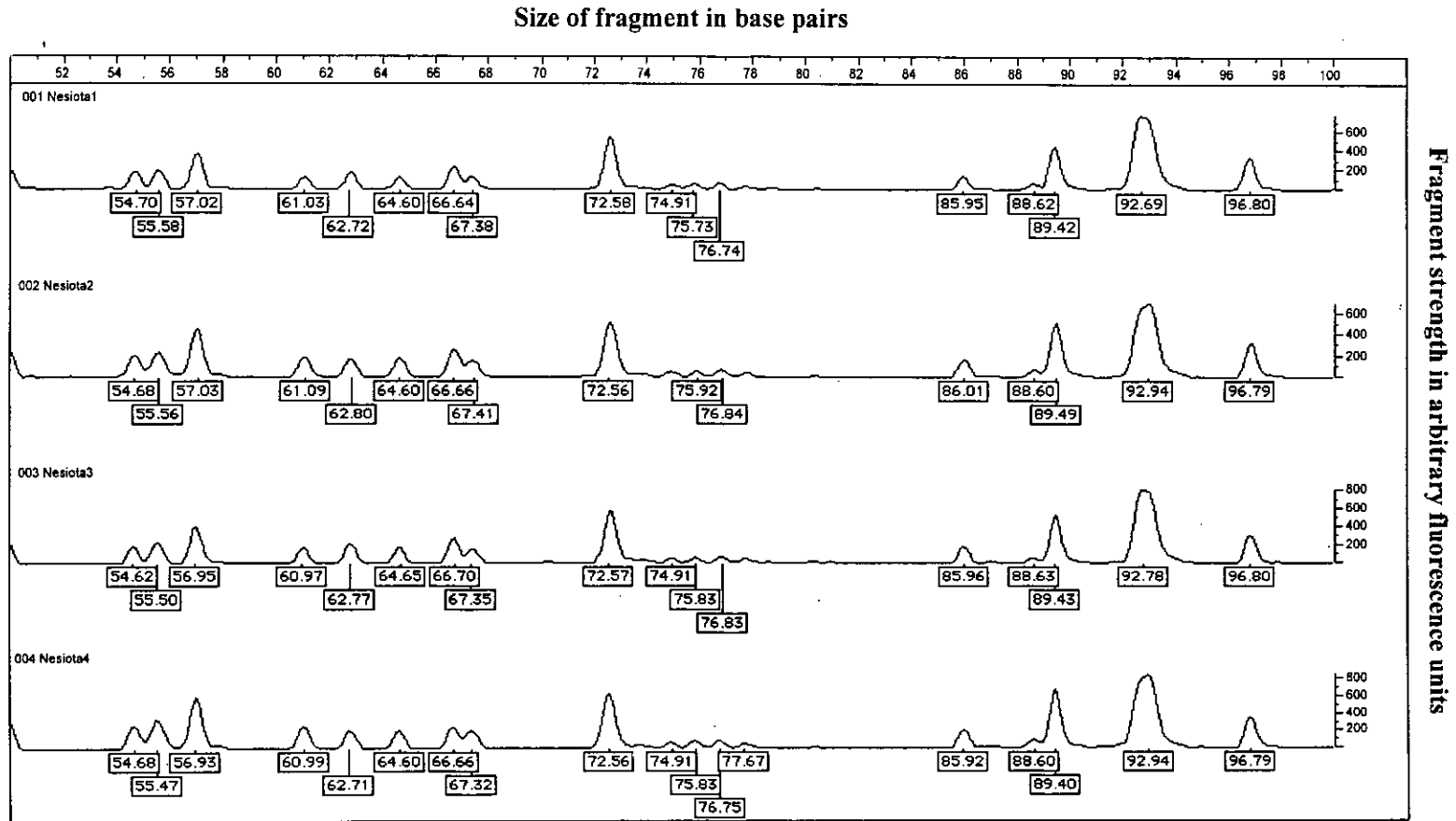


Figure 6.2. AFLP profiles of 50-100 bp fragments from *Nesiota elliptica*.

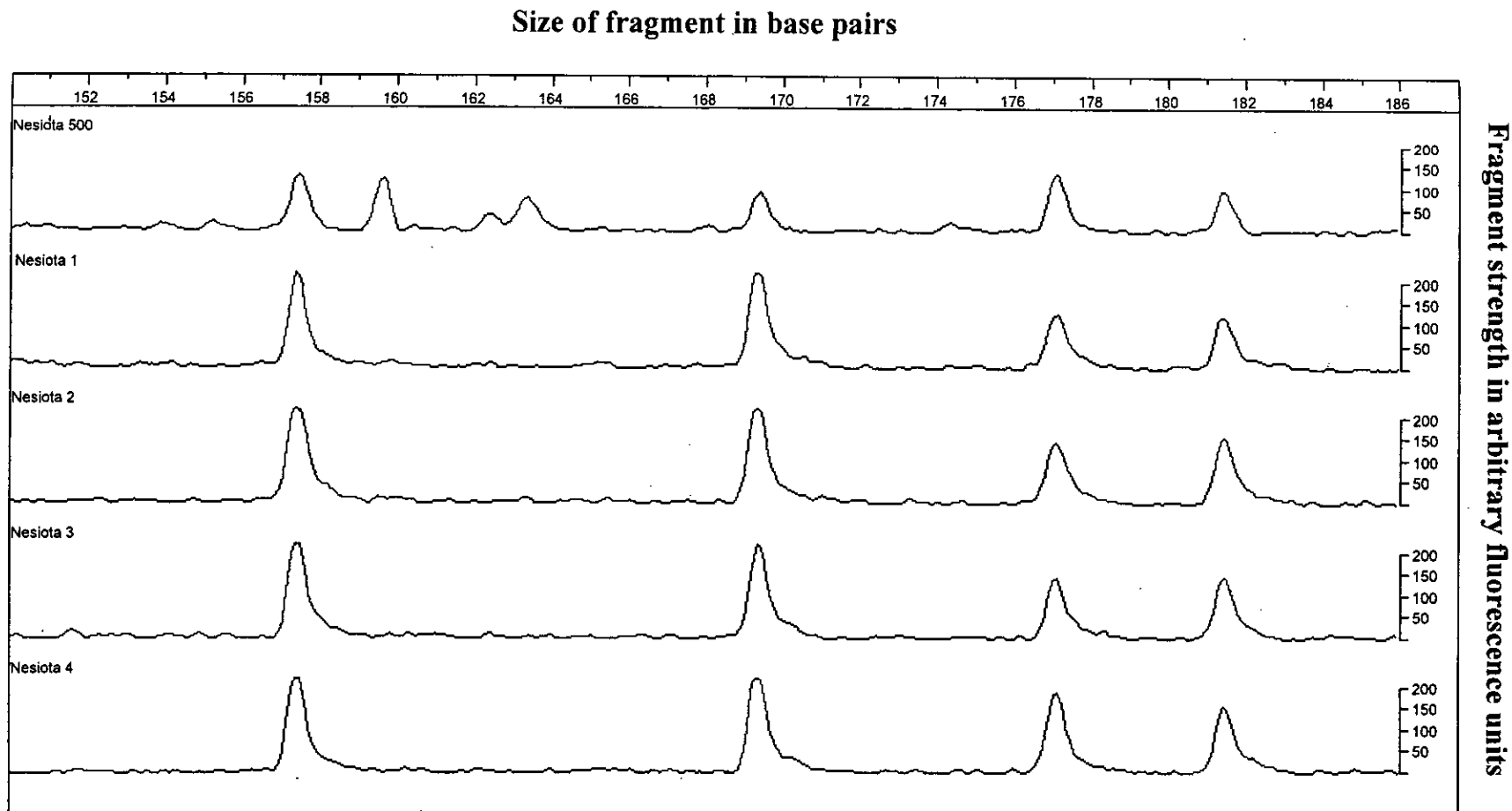


Figure 6.3. AFLP profiles of 150-180 bp fragments from *Nesiota elliptica* including profiles from the original surviving tree (*Nesiota* 500). The extra bands in the original tree are suspected to have been amplified from fungal contaminants.

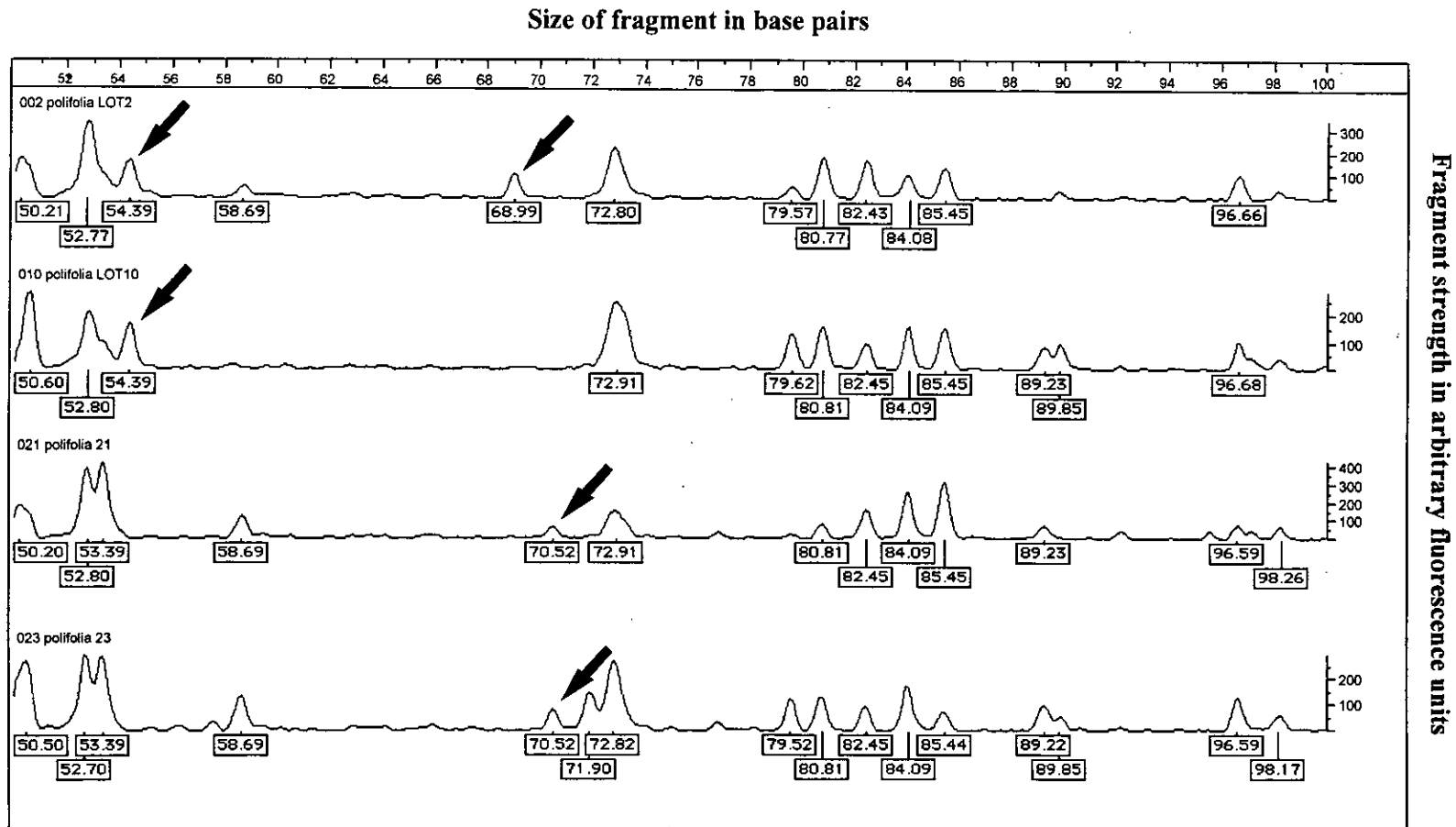


Figure 6.4. AFLP profiles of 50-100 bp fragments from two populations of *Phylica polifolia* (first two rows are the Lot population, the second two rows are the High Hill population). Arrows indicate polymorphisms.

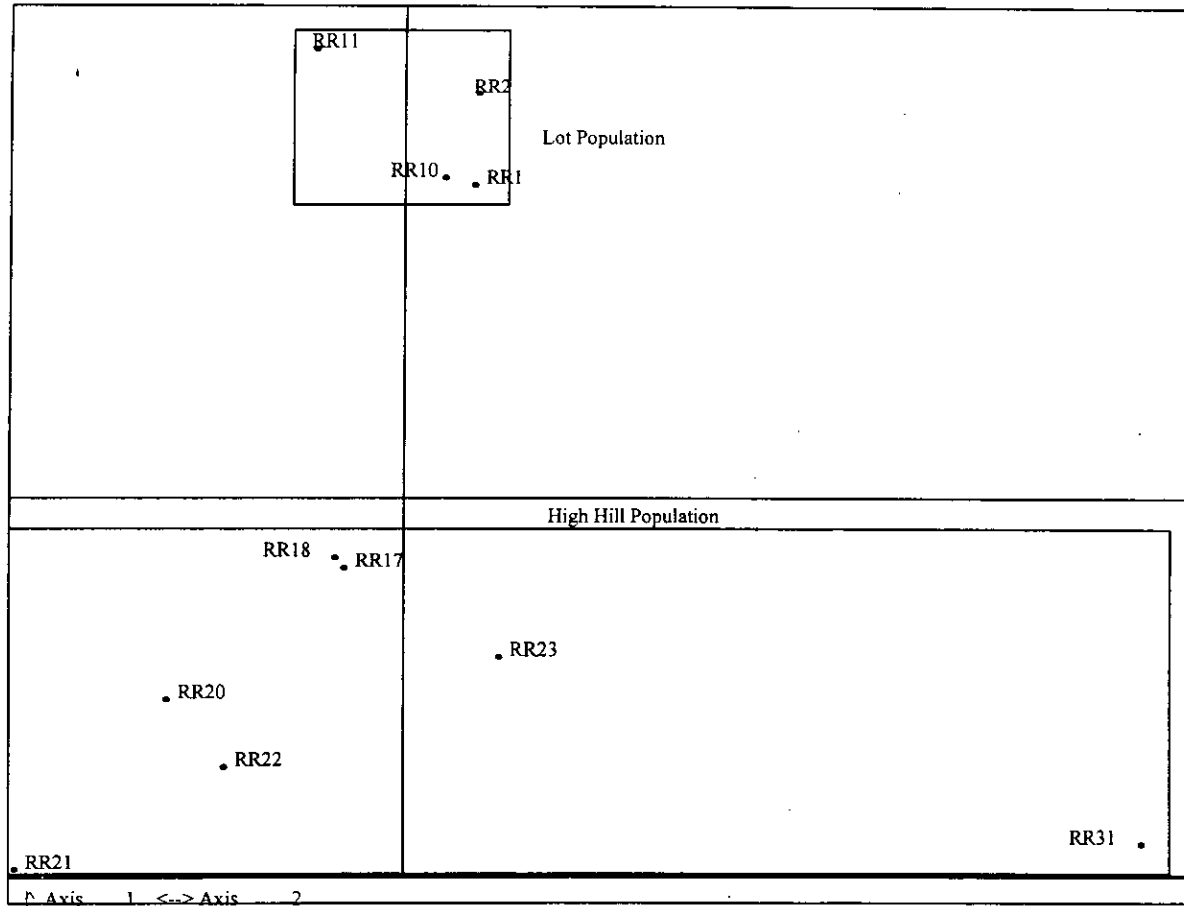


Figure 6.5. Principal co-ordinates analysis (using Jaccard's similarity co-efficient) of *P. polifolia*. Percentage variance of axis 1 = 43.8 and axis 2 = 24.8.

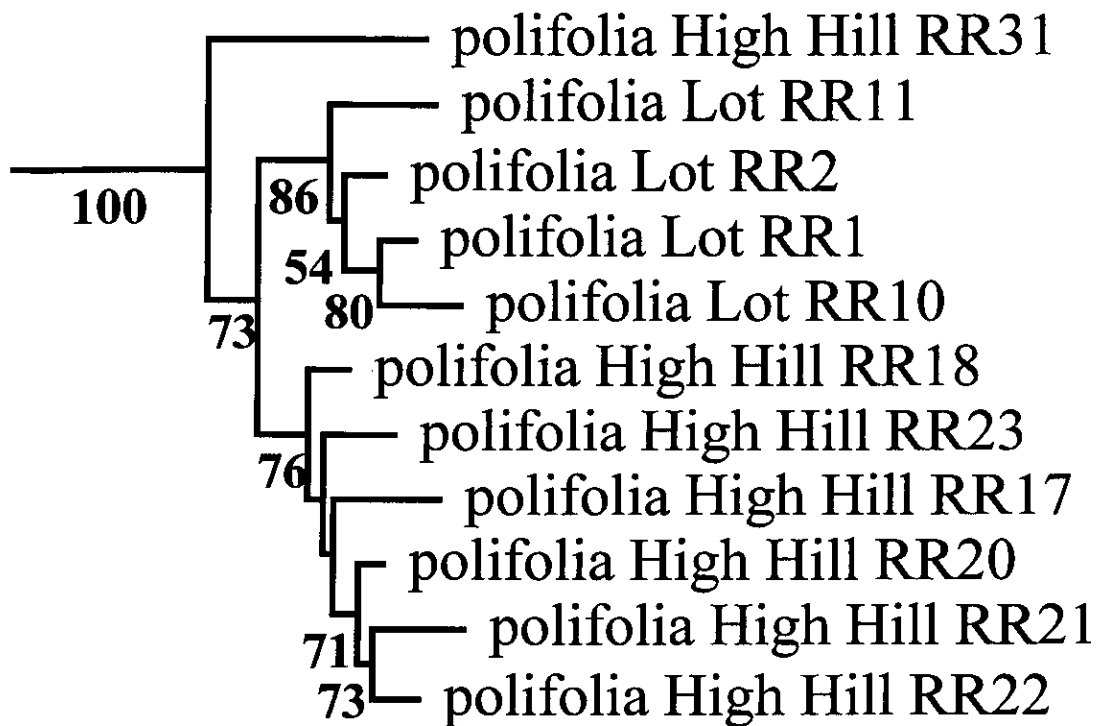


Figure 6.6. Tree for *P. polifolia* taken from the overall neighbour joining analysis on the island group presented in Chapter Five. Bootstrap percentages are shown below branches.

## 6.5. Discussion

The lack of variability shown in the *Nesiota* AFLP profiles does not necessarily mean that they actually are identical genotypes. However, studies on species or populations which are thought to be clonal (e.g. *Populus euphratica*; Fay *et al.*, in press, *Cosmos atrosanguineus*; Fay, pers. comm.) show AFLP profiles which are identical. Sample RRNes1 is a cutting from the last wild tree and RRNes2 to 4 are seedlings from this tree. Cuttings would normally be expected to be identical to the wild tree but seedlings should show some amount of variation due to segregation. Such variation was not detected by AFLPs. It is possible that the *N. elliptica* seedlings could have been formed from an unreduced gamete or by adventitious embryony which could explain their seemingly clonal AFLP profiles. Apomixis has not been recorded in Rhamnaceae although it has been recorded in the related



families Urticaceae and Rosaceae (Nygren, 1966; Asker and Jerling, 1992). The original tree (*Nesiota* 500) was infected with a number of species of fungi (Fay, 1989) and this may have resulted in the extra bands evident in the AFLP profile (Figure 6.4). The time between collection of the leaf sample and extraction of DNA meant that fungal growth could have occurred resulting in higher levels of contamination than there would have been in a fresh sample (Fay, pers. comm.).

The *N. elliptica* results may be contrasted with *P. polifolia* of which numbers and degree of genetic variability according to the AFLP results are greater. According to the PCO analysis *P. polifolia* is fairly clearly divided genetically into the two populations that exist on St Helena (with the exception of sample RR31) with the Lot population having strong bootstrap support in the NJ analysis. The geographic divisions are congruent with genotypic differences and because the two populations of *P. polifolia* at Lot and High Hill are distinct, I recommend that any seed orchards of these two populations that might be established be kept separate because mixing might disrupt the adaptation of these individuals to their particular habitats. The lower genetic diversity in the Lot population may be the result of its smaller size.

Although *P. nitida* on Mauritius is rare the degree of genetic variation between the limited number of samples in the study (Chapter Five) indicates that this population is also in a healthier state than *N. elliptica*. The sampling of *P. nitida* on Réunion is not sufficient to make any sound assessments regarding its conservation genetic status. The New Amsterdam population of *P. arborea*, which is also under threat, is also more variable than *N. elliptica* (Chapter Five). All of these rare or endangered species or populations may be contrasted with *P. arborea* on Tristan da Cunha (Chapter Five), which has an apparently healthy population both in terms of numbers and genetic diversity.

Because of the limited resources available for conservation it is necessary to identify taxa or areas which will maintain maximum diversity. Genetic diversity measures may indicate which taxa will have a better chance of long term survival. In this study the phylogenetic analysis (Chapter Four) identifies an endangered palaeoendemic taxon (*N. elliptica*) which is sister to a larger more recently derived group which also contains a number of endangered taxa. Vane-Wright *et al.* (1992) suggested that the taxa that are palaeoendemic or phylogenetically isolated should be

priorities for conservation. In other words *N. elliptica* should have a higher conservation priority than *P. polifolia*, *P. arborea* on New Amsterdam or *P. nitida* on Mauritius. However, the AFLP study shows that *N. elliptica* is in an extremely poor state in terms of levels of genetic diversity compared to the other more recently derived endangered taxa. Although *N. elliptica* has a higher conservation priority in terms of its phylogenetic position further factors regarding its long-term survival chances have to be taken into consideration before embarking on conservation programmes. The considerable efforts to increase the numbers of *Nesiota* individuals have so far proved relatively unsuccessful for reasons mentioned above. Even if propagation were successful, the long-term chances of survival of the species would be in doubt due to the lack of genetic variation detected. In terms of prioritising, it may therefore be more worthwhile to invest in a species such as *P. polifolia* for which chances of successful restoration are greater due to the greater levels of genetic variation that this taxon exhibits. However, because of its isolated phylogenetic position it is still better to persist with *N. elliptica* because it is more likely to contain novel genetic material than the recently derived *P. polifolia*. As mentioned in the Introduction there are cases in which species, which are not genetically diverse, survive perfectly well. It is therefore worth persisting with attempts to propagate and reintroduce *N. elliptica*.

## 6.6. Conclusions

Amplified fragment length polymorphisms proved useful in determining the conservation genetic status of island species in my studies. The lack of AFLP variation in *N. elliptica* can be contrasted with the levels of variation in *P. polifolia*, which can in turn be contrasted with the higher levels of variation found in *P. arborea*. One of the greatest advantages of AFLPs is that large numbers of markers can be produced more rapidly than with some other fingerprinting techniques such as RAPDs (AFLPs give 10-100 times more markers per primer than RAPDs) and they are therefore more suitable for detection of polymorphism between closely related individuals. The methods used for sizing and scoring bands are more reproducible and more accurate than for other fingerprinting methods. The chances of scoring

non-homologous bands as homologous bands are low. Disadvantages of AFLPs include the fact that they are dominant markers which means that the identity of homozygotes and heterozygotes cannot be reliably established. Levels of heterozygosity, which have been used as measures of fitness, can therefore not be determined. However, in the case of the individuals in this study it seems unlikely, particularly in the case of *N. elliptica*, that currently used co-dominant marker systems would detect polymorphisms. Further knowledge of the biology of these plants concerning breeding systems and pollinators is necessary to get a better idea of which strategies to employ in the conservation of these species, but AFLPs have provided a good basis from which to work.

### **6.7. Bibliography**

Asker, S.E. & L. Jehring. 1992. Apomixis in Plants. CRC Press, Boca Raton.

Bawa, K.A. & P.S. Ashton. 1991. Conservation of rare trees in tropical rainforests: a genetic perspective. *In* Falk, D.A. & K.E. Holsinger, Genetics and Conservation of Rare Plants, pp. 62-71. Oxford University Press, Oxford.

Beardmore, J.A. 1983. Extinction, survival and genetic variation. *In* Schonewald-Cox, C.M., S.M. Chambers, B. MacBryde & W.L. Thomas [eds.], A Reference Manual for Managing Wild Animal and Plant Populations, pp. 125-151. Benjamin/Cummings, Menlo Park, California.

Brown, A.H.D. & D.J. Schoen. 1992. Plant population genetic structure and biological conservation. *In* Sandlunds, O.J., K. Hindar & A.H.D. Brown [eds.], Conservation of Biodiversity for Sustainable Development, pp. 88-104. Scandinavian University Press, Oslo.

Clegg, M.T. & A.H.D. Brown. 1983. The founding of populations. *In* Schonewald-Cox, C.M., S.M. Chambers, B. MacBryde & W.L. Thomas [eds.], A Reference

Manual for Managing Wild Animal and Plant Populations, pp. 216-228. Benjamin/Cummings, Menlo Park, California.

Cronk, Q.C.B. 1984. The Historical and Evolutionary Development of the Plant Life of St Helena. Unpublished Ph.D thesis, Cambridge.

Cronk, Q.C.B. 1987. The plight of the St Helena Olive, *Nesiota elliptica*. Botanic Garden Conservation News 1: 30-32.

Cronk, Q.C.B. 1989. The past and present vegetation of St Helena. Journal of Biogeography 16: 47-64.

Falk, D.A. & P. Olwell. 1992. Scientific and policy considerations in restoration and reintroduction of endangered species. Rhodora 94: 287-315.

Fay, M.F. 1989. *Nesiota elliptica* - the St Helena Olive; new moves to safeguard its future. Botanic Gardens Conservation News 1: 7.

Fay, M.F., M.D. Lledó, M.M. Kornblum & M.B. Crespo. (in press). From the waters of Babylon? *Populus euphratica* in Spain is clonal and probably introduced. Biodiversity and Conservation.

Hill, M. H. Witsenboer, M. Zabeau, P. Vos, R. Kesseli & R. Michelmore. 1996. PCR-based fingerprinting using AFLPs as a tool for studying genetic relationships in *Lactuca* spp. Theoretical and Applied Genetics 93: 1202-1210.

Holsinger, K.E. & L.D. Gottlieb. 1991. Conservation of rare plants: principles and prospects. In Falk, D.A. & K.E. Holsinger, Genetics and Conservation of Rare Plants, pp. 195-208. Oxford University Press, Oxford.

Hoot, S.B., A.B. Kornkven & W.C. Taylor. 1998. The use of amplified fragment length polymorphisms (AFLP) in assessing genetic variation in the endangered species *Isoetes louisianensis*. *American Journal of Botany* 85: 95 [abstract]

International Union for the Conservation of Nature. 1994. IUCN Red List Categories. IUCN, Gland.

Jackson, A. 1991 Project Popeye - saving the St Helena Olive. Preliminary report to WWF (project no. 162/89). Royal Botanic Gardens, Kew.

Jackson, A. 1994. Project Popeye - saving the St Helena Olive. Final report to WWF (project no. 162/89). Royal Botanic Gardens, Kew.

Lande, R. & G.F. Barrowclough. 1987. Effective population size, genetic variation, and their use in population management. *In* Soulé, M.E. [ed.], *Viable Populations for Management*, pp. 87-124. Cambridge University Press, Cambridge.

Lucas, G. & H. Synge. 1978. *The IUCN Red Data Book*. IUCN, Gland.

Maunder, M., T. Upson, B. Spooner & T. Kendle. 1995. Saint Helena: Sustainable development and conservation of a highly degraded island ecosystem. *In* Vitousek *et al.* [eds.], *Ecological Studies* 115: 205- 217.

Melliss, J.C. 1875. *St Helena*. Reeve & Co., London.

Nygren, A. 1966. Apomixis in angiosperms. *In* Ruhland, W. [ed.], *Encyclopedia of plant physiology*. Volume 8. Sexuality, reproduction and alternation of generations, pp. 552-596. Springer-Verlag, Berlin.

Oldfield, S., C. Lusty & A. MacKinven [eds.]. 1998. *The World List of Threatened Trees*. World Conservation Monitoring Centre, World Conservation Press, Cambridge.

Pillans, N.S. 1942. The genus *Phylica* Linn. Journal of South African Botany 8: 1-164.

Qamaraz-Zaman, F., M.F. Fay, M.W. Chase & J.S. Parker. 1998. The use of AFLP fingerprinting in conservation genetics: a case study of *Orchis simia* (Orchidaceae). Lindleyana 13: 125-133.

Roxburgh, W. 1816. In Beatson, A. [ed.], Tracts relative to the island of Saint Helena; written during a residence of five years. Bulmer & Company, London.

Salwasser, H. 1990. Conserving biological diversity: a perspective on scope and approaches. Forest Ecology and Management 35: 79-90.

Schaeffer, M.L. 1981. Minimum population sizes for species conservation. Bioscience 31: 131-134.

Simberloff, D. 1988. The contribution of population and community biology to conservation science. Annual Review of Ecology and Systematics 19: 473-511.

Soulé, M.E. 1983. What do we really know about extinction? In Schonewald-Cox, C.M., S.M. Chambers, B. MacBryde & W.L. Thomas [eds.], A Reference Manual for Managing Wild Animal and Plant Populations, pp. 111-124. Benjamin/Cummings, Menlo Park, California.

Travis, S.E., J. Maschinski & P. Keim. 1996. An analysis of genetic variation in *Astragalus cremnocalyx* var. *cremnocalyx*, a critically endangered plant, using AFLP markers. Molecular Ecology 5: 735-745.

Vahl, M. 1794. Symbolae Botanicae 3: 41. Nicolaus Möller & Son.

Vane-Wright, R.I., C.J. Humphries & P.H. Williams. 1991. What to protect? – Systematics and the agony of choice. *Biological Conservation* 55: 235-254.

Walter, K.S. and Gillett, H.J. [eds.]. 1998. 1997 IUCN Red List of Threatened Plants. World Conservation Monitoring Centre, IUCN, Gland.

Winfield, M.O., G.M. Arnold, F. Cooper, M. Le Ray, J. White, A. Karp & K.J. Edwards. 1998. A study of genetic diversity in *Populus nigra* subsp. *betulifolia* in the Upper Severn area of the UK using AFLP markers. *Molecular Ecology* 7: 3-10.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97-159.

**CHAPTER SEVEN. CONCLUSIONS ON THE  
USE OF MOLECULAR DATA IN SOLVING  
SYSTEMATIC PROBLEMS AT DIFFERENT  
HIERARCHICAL LEVELS IN  
RHAMNACEAE**



## CHAPTER SEVEN. Conclusions On The Use Of Molecular Data In Solving Systematic Problems At Different Hierarchical Levels In Rhamnaceae

### 7.1. Rhamnaceae Study

The results of this study have led to a better understanding of relationships of genera within Rhamnaceae. Several new inter-generic relationships are uncovered. Based on *rbcL* and *trnL-F* nucleotide sequence data, Rhamnaceae are a strongly supported monophyletic group with their closest relatives being Dirachmaceae and Barbeyaceae. Three major strongly supported divisions within Rhamnaceae that were not apparent from assessments of morphological data alone are identified, and these “cryptic clades” are given informal names. Some tribes from Suessenguth’s (1953) and other systems are monophyletic, but the two large tribes Rhamneae and Zizipheae are paraphyletic. Eleven strongly supported tribes are recognised, three of which are new (Ampelozizipheae, Doerpfeldieae and Bathiorhamneae), the constitution of Rhamneae has been emended and the name of one tribe has been corrected (Zizipheae to Paliureae) and emended. Ventilagineae, Colletieae and Gouanieae are retained. Pomaderreae and Maesopsidae have been resurrected, as was Phylliceae which was also emended. The molecular trees permitted a better assessment of the biogeography of the family with two general patterns emerging. Informal sub-familial groupings have a wide predominantly Gondwanan distribution and clades within these groupings are usually restricted to individual plates.

The analysis of DNA sequences in this study resulted in more highly resolved trees than analysis of the morphological characters, but this is largely due to the larger number of characters available. Individual morphological characters do not perform badly in terms of their CI and RI values in comparison with many molecular characters; there are simply not enough of them. The fact that the morphological analysis of Rhamnaceae does not reveal the three major and well supported groups, evident in the molecular trees, indicates that the morphological characters used here are not useful in identifying deep clades in this group and that convergent morphological evolution subsequent to the formation of these clades may

obscure relationships. These results illustrate the difficulties involved in estimating phylogeny using only morphological characters in this group.

The molecular data indicate that many morphological character states have evolved in parallel, e.g. leaf venation patterns. Over-reliance on a few morphological characters can result in an incorrect estimate of phylogeny especially if these characters are homoplasious. A classification based on molecular data with the support of some morphological characters seems to be the best solution, and the molecular trees are used as the basis for recircumscribing tribes in Rhamnaceae.

Further studies should focus on finding morphological characters which might be used to define the “cryptic clades”, e.g. character states at various stages of floral apical development. The use of both molecular and morphological data will lead to a better understanding of the developmental and evolutionary biology of the group.

## 7.2. Phyliceae Study

The results of the Rhamnaceae study indicated that the genera *Nesiota* and *Noltea* formed a clade that is sister to *Phylica* and these genera were therefore included in a phylogenetic analysis of the tribe Phyliceae. Although Phyliceae are monophyletic, *Phylica* is polyphyletic with *P. stipularis* and *Nesiota elliptica* falling together in a clade that is sister to the rest of *Phylica*. *Phylica stipularis* is therefore placed in its own genus, *Trichocephalus*, a name that already exists for this taxon. The position of *N. elliptica* in the molecular trees indicates that it is a palaeoendemic taxon within the context of the tribe Phyliceae. All of the island species of *Phylica* form a well supported clade, the ‘island group’, with the southern African species *P. paniculata*, and this clade is derived from within the mainland group. Within the context of the ‘island group’, the Mascarene species *P. nitida* is palaeoendemic and the St Helenan, Tristan da Cunha and New Amsterdam species (*P. polifolia* and *P. arborea*) are recently derived neoendemic species.

The plesiomorphic morphology of the island species can be contrasted with that of their more derived mainland relatives. The fact that the island taxa are derived

from within the mainland taxa would seem to indicate that their plesiomorphic morphology arose due to reversals from more derived characteristics. However, the progenitor of this island group could have retained plesiomorphic morphological characteristics due to the fact that it was found in refugial areas (i.e. more mesic montane regions and along riverbanks). The retention of plesiomorphic, generalist morphological features meant that its capacity for dispersal was greater than that of more derived mainland species that are reliant on specific pollinators, soil types or climatic conditions. The retention of generalist morphology has therefore resulted in members of the island group having a greater chance of becoming established on dispersal to a variety of habitats and hence explains their current distribution on volcanic islands and montane regions in southern Africa.

In contrast to cases in which island taxa exhibit spectacular morphological specialisation (e.g. the Hawaiian silversword alliance), for *Phyllica* islands act as refugia for taxa that are highly restricted and likely to go extinct elsewhere in their range. The history of *Phyllica* on islands in the southern ocean indicates that island endemics are just as likely to be highly plesiomorphic as apomorphic in terms of their morphological characteristics.

Calibration of clocks based on degree of sequence divergence of closely related taxa is likely to be more accurate than estimates of divergence times based on comparisons between more phylogenetically isolated taxa because rates of change between the latter are likely to be more heterogeneous. The more distantly related the taxa, the more likely is an underestimate due to multiple undetected substitutions. The timings postulated here, assuming a molecular clock, seem to make sense from a biogeographic standpoint, given what is known about the history of southern Africa and the islands.

Analyses of both plastid and nuclear sequences indicated that *P. paniculata* is possibly paraphyletic, i.e. the island species evolved from different populations of *P. paniculata*. However, the putative paraphyly of this taxon could be due instead to low levels of divergence or lineage sorting of polymorphisms after divergence. Genes may diversify within a population prior to the diversification of the population itself, and organismal histories and gene histories can be partly independent. If these

polymorphisms persist through speciation events, the likelihood of gene and organismal trees having the same topology is low. Differential lineage sorting is more likely when time between nodes is short because newly acquired neutral mutations can take considerable time to become fixed, and the recent development of this group is compatible with such a scenario (1-0.5 mya, see Chapter Four). With these sequence data it is not possible to determine whether the island taxa are monophyletic or how many island species there are because these sequences were essentially invariant amongst these taxa. It was therefore necessary to look at a more variable source of data to try to answer these questions.

### **7.3. AFLP Study on the Island Species Of *Phyllica***

Amplified fragment length polymorphisms (AFLPs) are more variable than the sequence data used here and are therefore used to determine relationships between 'island group' taxa. The general consensus from the methods used to analyse the AFLP data is that each of the 'island group' species is monophyletic and that the possible paraphyly of *P. paniculata* (Chapter Four) is probably an artefact. Each species forms a unique group of genotypes indicating that gene flow between them ceased long ago. The results are consistent with the island species being the result of single introductions from a '*paniculata*-like' mainland ancestor with no subsequent gene flow. Increased sampling of *P. paniculata* may provide further evidence for the determination of its monophyly, but there is currently no reason to doubt its status. However, because of the subsequent period of isolation and continued gene flow among the continental populations, all evidence of which populations of *P. paniculata* were closer to the island species could have been removed. Continued interbreeding over a period in which new alleles arose and spread would make *P. paniculata* appear monophyletic. There may have been only slight divergence in *P. paniculata* prior to dispersal of the island taxa and considerable divergence since dispersal which would remove evidence of paraphyly. Therefore, even if *P. paniculata* were paraphyletic, proving it after one million years would be difficult.

Levels of polymorphism were high enough to allow within-species genotypic relationships to be revealed and to indicate the possible origins of some island populations. Genotypic distinctness could be assessed, and the extent of current and previous levels of gene flow could be estimated. Some island populations are shown to be distinct from other island populations indicating a period of isolation or separate introductions from different genetic stocks, e.g. Gough/New Amsterdam genotypes are distinct from Tristan da Cunha/Nightingale populations. Better assessments of these phenomena could be achieved by increasing the level of sampling.

Phylogenetic reconstruction breaks down if there is gene flow between populations. In this study a lack of gene flow was detected between certain isolated populations or species. This is not unreasonable given the large geographical distances between some of the species populations. For example little gene flow would be expected between the Tristan da Cunha Group and New Amsterdam since they are 6000 kilometres apart. The New Amsterdam individuals of *P. arborea* are a subset of the variation found within individuals from Gough Island which is consistent with the New Amsterdam individuals being derived from dispersal from Gough. The fact that there is strong bootstrap support in the parsimony and neighbour joining analyses for a New Amsterdam cluster lends support to the idea that they were derived from a single founder event. There is no strong bootstrap support for any other inter-population relationships within *P. arborea* and relationships break down in the strict consensus tree indicating that there has been recent gene flow between these other populations.

Although the direction and timing of founder events estimated here was consistent in part with the age of the islands, the AFLP data did not conclusively prove the origins of island populations. This again was partly due to the low sampling levels, but may also be due to the fact that there have been too many subsequent changes within populations or species to be able to detect the patterns at the time of divergence. Whether any other markers can identify the origins of island populations or prove that taxa such as *P. paniculata* are paraphyletic is an open question. The AFLP study has provided a focus for further studies which could

include the evaluation of cytoplasmic markers (e.g. plastid microsatellites or RFLPs) which may provide better evidence of origins. For example, if the Gough/New Amsterdam populations have a distinct cytotype, then this may be expected to be found on Tristan in the "hybrid" plants.

#### 7.4. Conservation genetics

One of the aims of conservation of endangered species or populations is to maintain the maximum amount of diversity, for which there are a variety of ways to produce estimates. These measures will indicate which taxa or areas should have priority. Given that the resources available for conservation are limited, it is necessary to identify taxa or areas that will maintain maximum diversity. According to Vane-Wright *et al.* (1992) the taxa which should be prioritised for conservation should be those which are palaeoendemic or phylogenetically isolated. In other words *N. elliptica* should have a higher conservation priority than *P. polifolia* or *P. arborea* on New Amsterdam or *P. nitida* on Mauritius because it is found on a long branch within the tree as a sister to a more derived group. However, the AFLP study shows that *N. elliptica* is in an extremely poor state in terms of levels of genetic diversity compared to the other more recently derived endangered taxa. Genetic variation may be lost from small isolated populations because of genetic drift, and deleterious alleles may become fixed through inbreeding. Genetic variation is necessary to maintain adaptive potential and populations lacking genetic variability are less likely to respond to changing environmental conditions and are therefore more likely to become extinct. What should be prioritised in the case of endangered island species in the tribe Phylliceae? Because of the isolated phylogenetic position of *N. elliptica* it is more likely to contain unique genetic material and should consequently be considered more valuable than the recently derived *P. polifolia* which has several close relatives. A lack of genetic variation does not necessarily mean that a species is unsuccessful and because of its uniqueness it is worth persisting with attempts to propagate and reintroduce *N. elliptica*.

## 7.5. General Conclusions

The Rhamnaceae molecular trees are used to produce a more natural supra-generic classification, show that the family is monophyletic, delimit several strongly supported groups that were not identified from morphological studies alone and indicate that some previously delimited tribes were paraphyletic. The trees provide the basis for further more critical studies of the evolutionary biology of the family.

The molecular study of Phylloceae is highly significant because it is the first phylogenetic analysis that reveals that derived taxa have retained plesiomorphic morphology in island and mainland species of the same group. Previous molecular phylogenetic studies revealed that island taxa that are morphologically derived are also phylogenetically derived (e.g. Baldwin, 1990, 1992; Hawaiian silverswords) or island taxa that are morphologically primitive are sister groups to phylogenetically derived groups (e.g. Fay *et al.*, 1997; Medusagynaceae). The molecular phylogenetic study of *Phyllica* indicated that taxa on islands and some taxa on mainland southern Africa with plesiomorphic morphological features are phylogenetically derived. The retention of plesiomorphic morphology in these species is due to their distribution in refugial areas such as on islands or in mesic montane or riverside localities on the mainland.

The AFLP study revealed that taxa in the 'island group' each form a distinct set of genotypes that is consistent with them being monophyletic. Of the species included in the study, one species is found on St Helena, one species on the Tristan da Cunha Group and New Amsterdam and another on Mauritius and Réunion. The AFLP study is a significant first step towards linking phylogenetics and population genetics. It revealed sufficient polymorphism to be able to distinguish between populations and to reveal the distribution of genotypes. The study indicates a lack of gene flow due to geographical isolation between some island species and populations. This information can be used to undertake a more directed study of how variation is partitioned using other markers. As discussed in Chapter One the molecular markers currently available for use in the study of population genetics and phylogenetics each have a number of advantages and disadvantages which when used together may

complement each other. A more complete picture of patterns and processes among closely related species linking the separate disciplines of population genetics and phylogenetics will be determined by using more than one type of molecular marker. Plastid data has been used to elucidate progenitor-derivative relationships in a number of crop species, the origin of both polyploids and diploids, introgression and genetic differentiation both among and within populations (reviewed in Soltis *et al.*, 1992). Because of the maternal, non-recombining mode of inheritance of plastid DNA it could provide cytotypes that might be ordered within species to yield gene genealogies which could determine infra-specific phylogeography. Therefore plastid DNA (RFLPs or microsatellites) could potentially be used to provide further evidence for the origin of island species or populations of *Phyllica* which have been hypothesised using AFLP data.

## 7.6. Bibliography

Baldwin, B.G., Kyhos, D.W. & Dvorak, J. 1990. Chloroplast DNA evolution and adaptive radiation in the Hawaiian Silversword alliance (Madiinae, Asteraceae), *Annals of the Missouri Botanic Garden* 77: 96-109.

Baldwin, B.G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. *Molecular Phylogenetics and Evolution* 1: 3-16.

Fay, M.F., S.M. Swensen & M.W. Chase. 1997. Taxonomic affinities of *Medusagyna oppositifolia* (Medusagynaceae). *Kew Bulletin* 52: 111-120.

Soltis, D.E., P.S. Soltis & B.G. Milligan. 1992. Intraspecific chloroplast DNA variation: systematic and phylogenetic implications. *In* Soltis, P.S., D.E. Soltis & J.J. Doyle, *Molecular Systematics of Plants*. Chapman & Hall, New York.



Suessenguth, K. 1953. Rhamnaceae, Vitaceae, Leeaceae. *In* Engler, A. & K. Prantl [eds.], *Die Natürlichen Pflanzenfamilien* 2, Aufl., 20d. Duncker & Humblot, Berlin.

Vane-Wright, R.I., C.J. Humphries & P.H. Williams. 1991. What to protect? – Systematics and the agony of choice. *Biological Conservation* 55: 235-254.

## Appendix 1. Matrix of *trnL-F* sequences with insertions and deletion characters indicated by asterices.

Sageretia thea	AAAAAAATTA	TAAAAA---T	TATTGGAT--	GAGCCTTGGT	AT-GGAAAC-	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Rhamnus lycioides	ATAAAAATTA	TAAAAA---T	AATTGGAT--	GAGCCTTGGT	AT-GGAAACC	TACC-AGGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Frangula alnus	ATAAAAATTA	TAAAAA----	----GGAT--	GAGC-TTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Rhamnella franguloides	AAAAAAATTA	TAAAAA---T	TATTGGAT--	GAGCCTTGGT	AT-GGAA-CC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Krugiodendron ferreum	AAAAAAATTA	TAAAAA--TT	AATTGGAT--	GAGCCTTGGT	AT-GGAA-CC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Rhamnidium cfelaeo	AAAAAAATTA	TAAAT?--TT	AATTGGAT-G	CAGCCTTGGT	AT-GGAA-CC	TACC-GAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Karwinskia humboldtiana	AAAAAAATTA	TAAAT?--TT	AATTGGATT-	GTACCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Condalia microphylla	AAAAAAATTA	TAAAT?--T	AATTGGATT-	GAGCCTTGGT	ATTGGAAACT	TACC-GGGTG	ATTACTTTCA	AATTCAGAGA	AACCCTGGAA
Scutia buxifolia	AAAAAAATTA	TAAAAA---T	GATTGGATT-	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Berchemia discolor	AAAAAAATTA	TAAAAA--TT	AATTGGATT-	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Maesopsis eminii	AAAAAAATTA	TAAAAA--TT	AATTGGAT--	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Ventilago viminalis	AAAAAAATTA	TTATAA--TT	AATT-AGT-G	GAGC-TTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Ventilago leiocarpa	AAAAAAATTA	TTATAA--TT	AATT-AGT-G	GAGC-TTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Reynosia uncinata	AAAAAAATTA	TAAAAA--TT	AAT-GAGT--	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Bathiorhamnus cryptophorous	AAAAAAATAA	TATAT?--TT	AATTGG--G	CAGC-TTGGT	AT-GGAAACC	TACC-GGCTG	ATAACTTTCA	AATTCAGAGA	AACCCTAGAA
Ampeloziziphus amazonicus	AAAAAAATAA	TATAT?----	--TTGGAT--	GAGC-TTGGT	AT-GGAA-CC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTAGAA
Doerpfeldia cubensis	AAAAAAATAA	TATAT?--TT	AATTGGATT-	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTAGAA
Hovenia dulcis	AAAAAA?A?A	AAAAAA--TT	AATTGGAT--	GAGC-TTGGT	AT-GGAAAC-	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Ceanothus	AAAAAA?A?A	AAAAAA--TT	AATTGGAG--	GAGC-TTGGT	AT-GGAAA-C	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Gouania mauritiana	TATAAA?A?T	AAAAAA--TT	AATTG-AT--	GGGT-TTGGT	AT-GGAAAC-	TACC-GAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Reissekia smilacina	TATAAA?A?T	AAAAAA--TT	AATTGAGTTC	GAGCCTAGGT	AT-GGAAACC	TATC-GAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Crumenaria erecta	TATAAA?A?T	AAAAAA--TT	AATTG-AT-G	GGGC-TTGGT	AT-GGAAACC	TACC-GAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Helinus integrifolius	TATAAA?A?A	AAAAAA--TT	AATTGGAT--	GCGC-TTGGT	AT-GGAAACC	TACC-GAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Pleuranthodes hillebrandii	TATAAA?A?T	AAAAAA--TT	AATT-GAT-G	GGGC-TTGGT	AT-GGAAACC	TACC-GAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Schistocarpaea johnsonii	AAAAAA?A?A	AAAAAA--TT	AATTGGATTG	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Colubrina asiatica	AAAATT?A?A	AAAAA----	--ATGGA--G	GAGC-TTGGT	AT-GGAAAC-	TACT-GAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Emmenosperma alphitonioides	AAAAAA?A?A	AAAAAA--TT	AATTGGATT-	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Alphitonia excelsa	AAAAAA?A?A	AAAAAAATT	AATTGGATT-	GAGCCTTGGT	AT-GGAAACT	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Lasiodiscus mildbraedii	AAAAAA?A?A	AAAAAA--TT	AATTGGAT-G	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Paliurus spinachristi	AAAAAA?A?A	AAAAAA--TT	AATTGGAT-G	GAGC-TTGGT	AT-GGAAACC	CGCC-AGGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Ziziphus glabra	AAAAAA?A?A	AAAAAA--TT	AATTGGAT--	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Ziziphus ornata	AAAAAA?A?A	AAAAAA--TT	AATTGGAT--	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Phylica pubescens	AA?TAA?A?A	AAAAA----	-----C	GAGC-TTGGT	AT-GGAAAC-	TAC--AA-TG	ATAACTTTCA	AATTCAGAGA	AACC-TGGAA
Phylica polifolia	AA?TAA?A?A	AAAAAAATT	AATTGGATT-	GAGCCTTGGT	AT-GGAAACC-	TACC-AAGTG	ATAGCTTTCA	AATTCAGAGA	AACCCTGGAA
Phylica arborea nitida	AA?TAA?A?A	AAAAA----	-----	GAGC-TTGGT	AT-GGAAAC-	TACC-AAGTG	ATAGCTTTCA	AATTCAGAGA	AACCCTGGAA
Nesiota elliptica	AA?TAA?A?A	AAAAA--TT	-ATTGG-TT-	GAGC-TTGGT	AT-GGAAAC-	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACC-TGGAA
Noltea africana	AA?TAA?A?A	AAAAA----	-----	-----GGC	C--GGAAACC	TACC-AAGTC	ATAACTTTCA	AATTCAGAGA	AACC-TGGAA
Discaria chacaye	AAAAAA?A?A	AAAAA----	AATTGGAT--	GAGC-TTGGT	AT-GGAA-CC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Spyridium spl	AAAAAA?A?A	AAAAA----	-----	-----GT	ATGGGAAAC-	TAC--AGGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Spyridium globulosum	AAAAAA?A?A	AAAAA--TT	AATTGGATT-	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA

Spyridium sp2	AAAAAA?A?A	AAAAAA--TT	AATTGGATT-	GAGCCTTGGT	ATGGAAACC-	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Cryptandra sp	AAAAAA?A?A	AAAAAA--TT	AATTGGATT-	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Trymalium sp1	AAAAAA?A?A	AAAAAA--TT	AATTGGATT-	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Trymalium sp2	AAAAAA?A?A	AAAAAA--TT	AATTGGATT-	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Pomaderris rugosa	AAAAAA?A?A	AAAAAA----	-ATTGGAT--	GAGC-TTGGT	AT-GGAA--C	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Siegfriedia darwinioides	AAAAAA?A?A	AAAAAA--TT	AATTGGAT--	GAGC-TTGGT	AT-GGAAA-C	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Colletia ulicina	AAAAAA?A?A	?AAAAA	AATTGGGAT-	GAGC-TTGGT	AT-GGAA-CC	TACCCAAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Barbeya oleoides	AAAAAATTAA	TAAAAA	AATTGGATT-	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	AGAACTTTCA	AATTCAGAGA	AACCCTGGAA
Dirachma socotrana	AAAAAATTAA	TAAAAA--TT	AATTGATT--	GAC--TTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Dorstenia psilurus	AAAAAATTAA	TAAAAA--TT	AATTGGATT-	GAGCCTTGGT	AT-GG-AAACC	TACC-AAGAG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Ficus	AAAAAATTAA	TAAAAA	GATTGAATTC	GAGCGTTGGT	AT-GG-AAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Artocarpus heterophyllus	AAAAAATTAA	TAAAAA--TT	AATT-GAT--	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Cannabis sativa	AAAAAATTAA	TAAAAA----	--TTGGATT-	GAGCCTTGGT	AT-GGAAACC	TACC-AAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Shepherdia argentea	?????ATTTA	TAAAAA????	??????????	??????????	??????????	??????????	??????????	??????????	??????????
Hippophae rhamnoides	AAAAAATTAA	TAAAAA--TT	AATTGGAT--	GAGC-TTGGT	AT-GGAAAC-	TACC-AAGTG	A-GAATTTCA	AATTCAGAGA	AACCCTGGAA
Elaeagnus	AAAAAATTAA	?AAAAA--TT	AATTGGAT--	GACC-TTGGT	AT-GGAAACC	TACC-AAGTG	A-GAATTTCA	AATTCAGAGA	AACCCTGGAA
Dryas drummondii	AAAAAATTAA	TAAA?T--TT	AATTGGATT-	GATC-TTGGT	AT-GG--CCG	TACCAGTGTG	AGAACTTTCA	AATTCAGAGA	AACCCTGGAA
Spiraea	AAAAAATTAA	TAAA?T--TT	AATTGGA--G	GAGC-TTGGT	AT-GGAACC-	TACC-AAGTG	AAAACTTTCA	AATTCAGAGA	AACCCTGGAA
Pyrus	AAAA?ATTTA	TAAA?T--TT	AATTGGAT--	GAGC-TTGGT	AT-GGAACC-	TACC-AAGTG	AGAACTTTCA	AATTCAGAGA	AACCCTGGAA
Colubrina reclinata	AAAATT?A?A	AAAA?T--TT	AATTGGAT-G	GAGCCTTGGT	AT-GGAAACC	TGCC-GAGTG	ATAACTTTCA	AATTCAGAGA	AACCCTGGAA
Gironniera	AAAAAATTAA	TAAA?T----	-----	-----	A--G---AC-	TACG-GAGTG	ATAACTTTCA	AATTCATAGA	AACCCTGGAA
Boehmeria	AAAAAA?TTA	TAAA?T--TT	AAT--GATT-	GACC-TTGGT	AT-GGACAC-	TACC-GAGTG	ATAACTTTCA	AATTCAGAGA	AACCCGGGAA

Sageretia thea	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	AAAGG-----	TT-CAGAAAG	C-----GA
Rhamnus lycioides	TT--AAAAAC	GGGCAATCCT	GAGCCAAATC	CCG-----TT	TTCTG-----	-AAAACAAAC	AAAGG-----	TT-CAGAAAG	C-----GA
Frangula alnus	TT--AAAAAC	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	AAAGGAAAGG	TT-CAGAAAG	C-----GA
Rhamnella franguloides	TT--AAAAAC	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	AAAAG-----	TT-CAGAAAG	C-----GA
Krugiodendron ferreum	TT--AAAAAC	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	AAAGG-----	TT-CAGAAAG	C-----GA
Rhamnidium cfelaeo	TT--AAAAAC	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	AAAGG-----	TT-CAGAAAG	C-----GA
Karwinskia humboldtiana	TT--AAAAAC	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	AAAGG-----	TT-CAGAAAG	C-----GA
Condalia microphylla	TT--AAAAAC	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	AAAGG-----	TT-CAGAAAG	C-----GA
Scutia buxifolia	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAA	AAAGG-----	TT-CAGAAAG	C-----GA
Berchemia discolor	TT--AAAAAC	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	AAAGG-----	TT-CAGAAAG	C-----GA
Maesopsis eminii	TT--AAAAAC	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	AAAGG-----	TT-CAGAAAG	C-----GA
Ventilago viminalis	TT--AAAAAC	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	AAAGG-----	TT-CAGAAAG	C-----GA
Ventilago leiocarpa	TT--AAAAAC	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	AAAGG-----	TT-CAGAAAG	C-----GA
Reynosa uncinata	TT--AAAAAC	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	AAAGG-----	TT-CAGAAAG	C-----GA
Bathiorhamnus cryptophorous	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	GAAGG-----	TT-CAGAAAG	C-----GA
Ampeloziziphus amazonicus	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CCG-----TT	TTCTG-----	-AAAACAAAC	GAAGG-----	TT-CGGAAAG	C-----GA
Doerpfeldia cubensis	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CCG-----TT	TTCTG-----	-AAAACAAAC	GAAGG-----	TT-CGGAGAT	C-----GA
Hovenia dulcis	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG-----TT	TTCTG-----	-AAAACAAAC	AAGGA-----	TT-CAGAAAG	T-----GA
Ceanothus	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTC-----TT	TTCTG-----	-AAAACAAAC	AAGGA-----	TT-CAGAAAG	T-----GA

Gouania mauritiana	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CGG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Roissekia smilacina	TT--ACAAAT	GGGCCATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-GATAATGG
Crumenaria erecta	TT--AGAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Helinus integrifolius	TT--ACAAAC	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Pleuranthodes hillebrandii	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAT	AAGGA----	TT-CAGAAAAG	T-----GA
Schistocarpaea johnsonii	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	C-----GA
Colubrina asiatica	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Emmenosperma alphonoioides	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Alphitonia excelsa	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Lasiodiscus mildbraedii	TT--ACAAAC	GGGCAATCCT	GAGCCAAATC	CTG----	TG	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Paliurus spinachristi	TT--ATAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAAGA----	TT-CAGAAAAG	T-----GA
Ziziphus glabra	TT--ATAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Ziziphus ornata	TT--ATAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Phylica pubescens	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAT	AAGGA----	TT-CAGAAAAG	T-----GA
Phylica polifolia	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Phylica arborea nitida	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAA	AAGGA----	TT-CAGAAAAG	T-----GA
Nesiota elliptica	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTT----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Noltea africana	TT--ACAAAC	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Discaria chacaye	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Spyridium spl	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAGAC	AAGTA----	TT-CAGAAAAG	T-----GA
Spyridium globulosum	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGTA----	TT-CAGAAAAG	T-----GA
Spyridium sp2	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAGAC	AAGTA----	TT-CAGAAAAG	T-----GA
Cryptandra sp	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGTA----	TT-CAGAAAAG	T-----GA
Trymalium spl	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGTA----	TT-CAGAAAAG	T-----GA
Trymalium sp2	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGTA----	TT-CAGAAAAG	-----GA
Pomaderris rugosa	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGTA----	TT-CAGAAAAG	T-----GA
Siegfriedia darwinioides	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGTA----	TT-CAGAAAAG	T-----GA
Colletia ulicina	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTT----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	T-----GA
Barbeya oleoides	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTATG----	-AAAA--AAC	AAAGG----	TT-CAGAAAAG	C-----GA
Dirachma socotrana	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CCG----	TT	TTCTG----	-AAAACAAAG	AAAGG----	TT-AAGAATT	CAGAAAACGA
Dorstenia psilurus	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CGG----	TT	TTCTG----	-AAAACAAAG	AAGGG----	TT-CAGAAGG	C-----CA
Ficus	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CGG----	TT	TTCTG----	-AAAACAAAC	AAGGG----	TT-CAGAAGG	C-----GA
Artocarpus heterophyllus	TT--AAAAAT	GGGTAATCCT	GAGCCAAATC	CGG----	TT	TTCTG----	-AAAACAAAC	AAGGG----	TT-CAGAAGG	C-----GA
Cannabis sativa	TTCAAAAAA	GGGCAATCCT	GAGCCAAATC	CGG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAGG	C-----AA
Shepherdia argentea	?????????	?????????	?????????	?????????	?????????	?????????	?????????	?????????	?????????	?????????
Hippophae rhamnoides	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CAG----	TT	TTCTGACTCT	TAAAACAAAC	AAGGG----	TT-CAGAAAAG	C-----
Elaeagnus	TT--AATAAT	GGGCAATCCT	GAGCCAAATC	CAG----	TT	TTCTGACTCT	GAAAACAAAC	AAGGG----	TT-CAGAAAAG	C-----
Dryas drummondii	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTATG----	-AAAACAAAC	AAGGG----	TTTCAGAAAAG	C-----GC
Spiraea	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTATG----	-AAAACAAGC	AAGGG----	TTTCATAAAC	TCATAAACGA
Pyrus	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTATG----	-AAAATAAAC	AAGGG----	TTTCATAAAC	C-----GA
Colubrina reclinata	TT--ACAAAT	GGGCAATCCT	GAGCCAAATC	CTG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAGG	T-----GA
Gironniera	TTAAAAAAT	GGGCAATCCT	GAGCCAAATC	CGG----	TT	TTCTG----	-AAAACAAAC	AAGGA----	TT-CAGAAAAG	C-----GA
Boehmeria	TT--AAAAAT	GGGCAATCCT	GAGCCAAATC	CGT----	GT	TTATG----	-----	AAAGG----	TT-CGAAAAG	T-----GG

Sageretia thea	TAATAAAAA- ----GGGAT AGATA--GGG ATAGG-----	TGCAGAGACT CAATGGAAGT TGTTCTAACA AA----TGGA GTTGGCCACG
Rhamnus lycioides	TAATAAAAAA ----- -GG ATAGG-----	TGCAGAGACT CAATGGAAGT TGTTCTAACA AA----TGGA GTTGGCCACG
Frangula alnus	TAATAAAAAA -----GGAT AGAGA--GGG ATAGG-----	TGCAGAGACT CAATGGAAGT TGTTCTAACA AA----TGGA GTTGGCCACG
Rhamnella franguloides	TAATAAAAAA -----GGAT AGATA--GGG ATAGG-----	TGCAGAGACT CAATGGAAGT TGTTCTAACA AA----TGGA GTTGGCCACG
Krugiodendron ferreum	TAATAAAAAA -----GGAT AGATA--GGG ATAGG-----	TGCAGAGACT CAATGGAAGT TGTTCTAACA AA----TGGA GTTGGCCACG
Rhamnidium cfelaeo	TAATAAAAAA -----GGAT AGATA--GGG ATAGG-----	TGCAGAGACT CAATGGAAGT TGTTCTAACA AA----TGGA GTTGGCCACG
Karwinskia humboldtiana	TAATAAAAAA -----GGAT AGATA--GGG ATAGG-----	TGCAGAGACT CAATGGAAGT TGTTCTAACA AA----TGGA GTTGGCCACG
Condalia microphylla	TAATAAAAAA -----GGAT AGATA--GGG ATAGG-----	TGCAGAGACT CAATGGAAGT TGTTCTAACA AA----TGGA GTTGGCCACG
Scutia buxifolia	TAATAAAAAA A-----GAT AGATA--GGG ATAGG-----	TGCAGAGACT CAATGGAAGT TGTTCTAACA AA----TGGA GTTGGCCACG
Berchemia discolor	TAATAAAAAA -----GGAT AGATA--GGG A-----	TGCAGAGACT CAATGGAAGT TGTTCTAACA AA----TGGA GTTGGCCACG
Maesopsis eminii	TAATAAAAAA A-----GGAT AGATA--TGG ATAGG-----	TGCAGAGACT CAATGGAAGT TGTTCTAACA AA----TGGA GTTGGCCACG
Ventilago viminalis	TGA-AAAAAA AAAA--GGAT -GAAT-GGGG ATGGG-----	TGCAAAAACCT CCTTGAAGT TGTTCTAACA AA----TGGA GTTGGCCGCG
Ventilago leiocarpa	TAAAAA AAAA--GGAT -AAAT-GGGG ATGGG-----	TGCAAAAACCT CCATGGAAGT TGTTCTAACA AA----TGGA GTTGGCCGCG
Reynosia uncinata	TAATAAAAAA -----GGAT AGATA--GGG ATAGG-----	TGCAGAGACT CAATGGAAGT TGTTCTAACA AA----TGGA GTTGGCCACG
Bathiorhamnus cryptophorus	TAATAAAAAA -----GGAC AGATA--GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA TA----GGGA GTTACCTACT
Ampeloziziphus amazonicus	TAATAAAAAA -----GGAC AGATA--GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----GGGA GTTACCTACG
Doerpfeldia cubensis	TAATAAAAAA -----GGAA AGATA--GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----GGGA GTTACCTACG
Hovenia dulcis	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Ceanothus	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Gouania mauritiana	TAATAAAAAA A-----GGG ATAGG-----	TGCAGAGACT CAACGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Reissekia smilacina	TAATAAAAAA AAAA-----GGGG ATGGG-----	TGCAAAAACCT CAACGGAAGC TGTTCTAACA AA----TGGA GTTGGCCGCG
Crumenaria erecta	TAATAAAAAA A-----GGG--GGGG ATAGG-----	TGCAGAGACT CAACGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Helinus integrifolius	TAATAAAAC-----GGG ATAGG-----	TGCAGAGACT CAACGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Pleuranthodes hillebrandii	TAATAAAAAA AA---GGG--GGGG ATAGG-----	TGCAGAGACT CAACGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Schistocarpaea johnsonii	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----CGGA GTTGGCTGCG
Colubrina asiatica	TAATAAAAA-----GGG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Emmenosperma alphitonioides	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Alphitonia excelsa	TAATAAAA-----GGG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AACAAATGGA GTTGGCTGCG
Lasiodiscus mildbraedii	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Paliurus spinachristi	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Ziziphus glabra	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Ziziphus ornata	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TCTTCTAACA AA----TGGA GTTGGCTGCG
Phylica pubescens	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Phylica polifolia	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Phylica arborea nitida	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Nesiota elliptica	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Noltea africana	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAACGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Discaria chacaye	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Spyridium sp1	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Spyridium globulosum	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG
Spyridium sp2	TAATAAAAAA -----GG ATAGG-----	TGCAGAGACT CAATGGAAGC TGTTCTAACA AA----TGGA GTTGGCTGCG

Cryptandra sp	TAATAAAAAA	-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGGCTGCG
Trymalium sp1	TAATAAAAAA	-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGGCTGCG
Trymalium sp2	-----	-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGGCTGCG
Pomaderris rugosa	TAATAAAAAA	-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGGCTGCG
Siegfriedia darwinioides	TAATCAAAAA	-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGG	GTTGGCTGCG
Colletia ulicina	TAATAAAAAA	-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TATTTAACA	AA----	TGGA	GTTGGCTGCG
Barbeya oleoides	TAATAAAAAA	-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	AGGA	GTTGGCTGCG
Dirachma socotrana	TA-TAAGAAA	AAA-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGGTTGCG
Dorstenia psilurus	TAATAAAAAA	A-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	CGGA	GTTGGCTGAA
Ficus	TAATAAAAAA	-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGACTGCG
Artocarpus heterophyllus	TAATAAAAAA	-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGGCTGCA
Cannabis sativa	TAATAAAAAA	-----	-----GA	ATAGGATAGG	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGGCTGCG
Shepherdia argentea	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????
Hippophae rhamnoides	-AATACAAA-	-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGGCTGCG
Elaeagnus	-AATAAAA-	-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGGCTGCG
Dryas drummondii	GAATAAAAA-	-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGGCTGCA
Spiraea	GAATAAAAG-	-----	-----AGG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGACTGCA
Pyrus	AAATAAAAA-	-----	-----AGG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGGCTGCA
Colubrina reclinata	TAATAAAAAA	-----	-----GG	ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGCCTGCG
Gironniera	TAATAAAAAA	-----	GAAT A-----	GG ATAGG-----	TGCAGAGACT	CAATGGAAGC	TGTTCTAACA	AA----	TGGA	GTTGGTTGCG
Boehmeria	TAAAAAAA	-----	T -AAA-----	GG ATAGG-----	TGCAGAGACT	CAATGGAAGT	TGTTCTAACA	AA----	TGGA	GTTGGCTACT

Sageretiá thea	ATGCGTTAGT	A-----	AA	AGAC-----	T CCTTCCATCG	AAACTCCAGA	AAGTAT----	-GAAGAATAA	ACGTATA--	T ATACG-TATA
Rhamnus lycioides	ATGCGTTAGT	A-----	AA	GGAC-----	T CCTTCCATCG	AAACTCCAGA	AAGTAT----	-GAAGAATAA	ACGTATA--	T ATACG-----
Frangula alnus	ATGCGTTAGT	A-----	AA	GGAC-----	T CCTTCCATCG	AAACGCCAGA	AAGTAT----	-GAAGAATAA	ACGTATA--	T ATACG-TATA
Rhamnella franguloides	ATGCGTTAGT	A-----	AA	GGAC-----	T CCTTCCATCG	AAACTCCAGA	AAGTAT----	-GAAGAATAA	ACGTATA--	T ATACG-----
Krugiodendron ferreum	ATGCGTTAGT	A-----	AA	GGAC-----	T CCTTCCATCG	AAACTCCAGA	AAGTAT----	-GAAGAATAA	ACGTATA--	T ATACG-----
Rhamnidium cfelaeo	ATGTGTTAGT	A-----	AA	GGAC-----	T CCTTCCATCG	AAACTCCAGA	AAGTAT----	-GAAGAATAA	ACGTATA--	T ATACG-----
Karwinskia humboldtiana	ATGTGTTAGT	A-----	AA	GGAC-----	T CCTTCCATCG	AAACTCCAGA	AAGTAT----	-GAAGAATAA	ACGTATA--	T ATACG-----
Condalia microphylla	ATGCGTTAGT	A-----	AA	GGAC-----	T CCTTCCATCG	AAACTCCAGA	AAGTAT----	-GAAGAATAA	ACGTATA--	T ATACG-----
Scutia buxifolia	ATGCGTTAGT	A-----	AA	GGAC-----	T CCTTCCATCG	AAACTCCGA	AAGTAT----	-GAAGAATAA	ACGTATA--	T ATACG-----
Berchemia discolor	ATGCGTTAGT	A-----	AA	GGAC-----	T CCTTCCATCG	AAACTCCAGA	AAGTAT----	-GAAGAATAA	ACGTATA--	T ATACG-----
Maesopsis eminii	ATGCGTTAGT	A-----	AA	GGAC-----	T CCTTCCATCG	AAACTCCAGA	AAGGAT----	-GAAGAATAA	ACCTATA--	T ATACG,-----
Ventilago viminalis	ATCCATTATT	A-----	AA	GGAC-----	T CCTTCCATCG	AAACTCCAA	AAGGAT----	-GAAAAATAA	ACCTATA--	T TTCCT-----
Ventilago leiocarpa	ATGCTTTAGT	A-----	AA	GGAC-----	T CCTTCCATCG	AAACTCCAA	AAGGAT----	-GAAAAATAA	CCCTATA--	T ATACT-----
Reynosa uncinata	AT-----	GT A-----	AA	GGAC-----	T CCTTCCATCG	AAATTCCAGA	AAGTAT----	-GAAGAATAA	ACGTATA--	T ATACG-----
Bathiorhamnus cryptophorous	ATGCGTTAGT	A-----	AA	GGAA-----	T ACTTACATCG	AAACGCCAGA	AAGGAT----	-GAAGAATAA	ACCTATA--	T ATACG-----
Ampeloziziphus amazonicus	ATGCGTTAGT	AA--	AGTAAA	GAAA-----	T CCTTCCATCG	AAACGCCAGA	AAGGAC----	-GAAGAATAA	ACCTATA--	T ATACG-TATA
Doerpfeldia cubensis	ATGCGTTAGT	A-----	AA	GAAATAAAAT	CCTTACATCG	AAACGCCAGA	AAGGAC----	-GAAGAATAA	ACCTATA--	T ATACG-TATA
Hovenia dulcis	CTGCGTTAGT	A-----	AA	GGAA-----	T CCTTCCATCG	AAACTCCAGA	AAGGAT----	-GAAGAATAA	ACCTATA--	T ATACG-TATA
Ceanothus	ATGCGTTAGT	A-----	AA	GGAA-----	T CCTTCCATCG	AAACTCCAGA	AAGGAT----	-GAAGAATAA	ACCTATA--	T ATACG-TATA
Gouania mauritiana	GTGCGTTAGT	A-----	AA	GGAA-----	T CCTTCCATCG	AAACTCCAGA	AAGGAT----	-GAAGAATAA	ACCTATA--	T CTACG-TATA

Reissekia smilacina	GCCCGTTAGT	A-----AA	GGAA-----T	CCTTCTATCG	AAACCCCAA	AAGGAT----	-GAAAATAA	ACCTATT--T	CAACG-TATA
Crumenaria erecta	GCGCGTTAGT	A-----AA	GGAA-----T	CCTTCTATCG	AAACTCCAGA	AAGGGT----	-GAAGAATAA	ACCTATA--T	CTACG-TATA
Helinus integrifolius	GTGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAGA	AAGGAT----	-GAAGAATAA	ACCTATA--T	CTACG-TATA
Pleuranthodes hillebrandii	GCGCGTTAGT	A-----AA	GGAA-----T	CCTTCTATCG	AAACTCCAGA	AAGGAT----	-GAAGAATAA	ACCTATA--T	CTACG-TATA
Schistocarpaea johnsonii	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAGA	AAAGAT----	-GAAGAATAA	ACCTATA--T	ATACG-TATA
Colubrina asiatica	ATGTGTTAGT	A-----AC	GGAA-----T	TCTTCCGTCG	AAACTACAGA	AAGGAT----	-GAAGAATAA	ACGTATA---	-----TATA
Emmenosperma alphonoioides	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAGA	AAGGAT----	-GAAGAATAA	ACCTATA--T	ATACG-TATA
Alphitonia excelsa	GTGCGTTAGT	AAA-AGTAAA	GGAA-----T	CCTTCCATCG	AAACTCCAGA	AAGGAT----	-GAAGAATAA	GCCTATG--T	ATACG-TATA
Lasiodiscus mildbraedii	AAGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAGA	AAGGAT----	-GAAGAATAA	ACCTATA--T	ATACG-TATA
Paliurus spinachristi	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAGA	AAGGAT----	-GAAGAATAA	ACCTATA--T	ATACG-TATA
Ziziphus glabra	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAGA	AAGGAT----	-GAAGAATAA	ACCTATA--T	ATACG-TATA
Ziziphus ornata	ATGCGTTAGT	A-----AA	GGAA-----T	-----CG	AAACTCCAGA	AAGGAT----	-GAAGAATAA	ACCTATA--T	ATACG-TATA
Phylica pubescens	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAATTCAG	AAGGAG----	-GAAGGATAA	ACGTATA--T	ATACG-TATA
Phylica polifolia	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAATTCAG	AAGGAG----	-GAAGGATAA	ACGTATA--T	ATACG-TATA
Phylica arborea nitida	ATGCGTGAAT	A-----AA	GGAA-----T	CCTTCCATCG	AAATTCAG	AAGGAG----	-GAAGGATAA	ACGTATA--T	ATACG-TATA
Nesiota elliptica	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAATTCAG	AAGGAT----	-GAAGGATAA	ACCTATA--T	ATACG-TATA
Noltea africana	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAATTCAG	AAGGAT----	-GAAGGATAA	ACCTATA--T	ATAGG-TATA
Discaria chacaye	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAG	AAGGAT----	-GAAGAATAA	ACCTATA--T	-----TATA
Spyridium sp1	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAG	AAGGAT----	-GAAGAATAA	ACCTATA--T	-TACG-TATA
Spyridium globulosum	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAG	AAGGAT----	-GAAGAATAA	ACCTATA--T	ATACG-TATA
Spyridium sp2	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAG	AAGGAT----	-GAAGAATAA	ACCTATA--T	ATACG-TATA
Cryptandra sp	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAG	AAGGAT----	-GAAGAATAA	ACCTATA--T	ATACG-TGTA
Trymalium sp1	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAG	AAGGAT----	-GAAGAATAA	ACCTATA--T	-TACG-TATA
Trymalium sp2	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAG	AAGGAT----	-GAAGAATAA	ACCTATA--T	-TACG-TATA
Pomaderris rugosa	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAG	AAGGAT----	-GAAGAATAA	ACCTATA--T	ATACG-TATA
Siegfriedia darwinioides	ATGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAG	AAGGAT----	-GAAGAATAA	ACCTATA--T	ATACG-TATA
Colletia ulicina	ATGCGTTAGT	A-----AA	GGAA-----T	-----TCG	AAACTCCAG	AAGGAT----	-GAAGAATAA	ACCTATA--T	-----TATA
Barbeya oleoides	TTGCGTTAGT	A-----AA	GGAA-----T	CCTTCCAGTG	AAACTCCAG	AAGGAT----	-GAAGAATAA	ACCTATA--T	ATACG-----
Dirachma socotrana	TTGAGTTAGT	A-----AC	GGAA-----T	CCTTCCCTCA	AAAGTCCATA	AAGTAT----	-GAAGAATAA	ACCTATA--T	ATACG-----
Dorstenia psilurus	GTGCG-----	-----AAA	GGAA-----T	CACTCCA--A	AAA-----A	AAGGAT----	-GAAGAAT--	-----ATA--T	ATACG-TATA
Ficus	GTGCGTTAGT	A-----AA	GGAA-----T	CACTCCA--	-----GA	AAGGAT----	-GAAGAATAA	ACCTATA--T	-TACG-TATA
Artocarpus heterophyllus	GTGCATTAGT	A-----AA	GGAA-----T	CACTTCA--	-----GA	AAGGAT----	-GAAGAATAA	ACGTATA--T	-TACG-TATA
Cannabis sativa	TT-----	-----	-----	-----	-----	-----	-----	-----	-----
Shepherdia argentea	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????
Hippophae rhamnoides	TTGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTTACGA	AAGGAT----	-GAAGAATAC	-CCTATA--T	ATACGATATA
Elaeagnus	TTGTGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTTCCGA	AAGGATGAGA	TGAAGAAGAC	-CCTATA--T	ATACGATATA
Dryas drummondii	TTGTGTTAGT	A-----AA	GGAA-----T	CCTTACATCG	AAACTTCCGA	AAGGAT----	-GAAGGATAA	ACGTATATAC	ATACG-TATA
Spiraea	TTGTGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTTCCGA	AAGTAT----	-GAAGGATAA	ACTTAAAGAC	ATACA-TATA
Pyrus	TTGTGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTTCCGA	AAGGAT----	-GAAGGATAA	ACGTATATAC	ATACG-TATA
Colubrina reclinata	ATGTGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCAG	AAGGAT----	-GAAGAATAA	ACGTATATAT	ATACG-TATA
Gironniera	TTGCGTTAGT	A-----AA	GGAA-----T	CCTTCCATCG	AAACTCCATA	AAGGAT----	-GAAGAATAA	ATCTATA--T	-TACG-TATA
Boehmeria	TTGCGTTAGT	AGTTAGTAAA	GGAA-----T	CCTTCCATTG	AAACTCCAG	AAGGAT----	-GAAGAATAA	ATGTATA--T	-TAGG-----

Sageretia thea	CGTACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	AAATC----	-----	TTTTTTTTAT	TTATAT----
Rhamnus lycioides	--TACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATAT----
Frangula alnus	CGTACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATAT----
Rhamnella franguloides	--TACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATAT----
Krugiodendron ferreum	--TACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATAT----
Rhamnidium cfelaeo	--TACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATAT----
Karwinskia humboldtiana	--TACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATAT----
Condalia microphylla	--TACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATAT----
Scutia buxifolia	--TACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATAT----
Berchemia discolor	--TACTGAAA TACTATCTCC	AACTCCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	ATATAT----
Maesopsis eminii	--TACTGAAA TACTATCT--	----CCAAAC	CGAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTACAA----
Ventilago viminalis	--TTCTGAAA TAATATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTTTTT--
Ventilago leiocarpa	--TTCTGAAA TAATATCT--	----CCAAAC	CCAATCATTA	TTGAC-ACCC	CAATC--C--	-----	TTTTTTTTTT	TTTTATAT--
Reynosia uncinata	--TACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATAT----
Bathiorhamnus cryptophorus	--TATTGAAA TACTATCT--	----TCAAAC	CAAATGATTA	ATGACGACCC	AAATC----	-----	TTTTTTTTTT	TTATATG-TT
Ampeloziziphus amazonicus	CATAGTGAAA TACTATCT--	----TCAAAC	CAAATGATTA	ATGAAGACCC	CAATC----	-----	TTTTTTTTTT	TTATATG---
Doerpfeldia cubensis	CATATTGCAA TACTATCT--	----TCAAAC	CAAATGATTA	ATGAAGACCC	CAATC----	-----	TTTTTTTTTT	TTATATG---
Hovenia dulcis	CGTACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATATG-TT
Ceanothus	CGTACTGAAA TACTATCG--	----CCAAAC	CAAATGATTA	AT-----	-----	-----	-----	---TATG-TT
Gouania mauritiana	CGTACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATATG-TT
Reissekia smilacina	CGTACTAAAA TACTGTCT--	----CCAAAC	CCAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTA-ATA-TT
Crumenaria erecta	CGTACTAAAA TAGTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATATG-T
Helinus integrifolius	CGTACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATATG-TT
Pleuranthodes hillebrandii	CGTACTAAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGGCCC	GACTC----	-----	TTTTATTTAT	TTATATG-TT
Schistocarpaea johnsonii	CGTACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATATG-TT
Colubrina asiatica	CGTACTGAAA TCCTATCT--	----CCAAAC	CAAATGATTA	ATGACGACTC	GAATC----	-----	TTTTTT-TAT	TTCTATG-TT
Emmenosperma alphonoides	CGTACTGAAA TACTATCT--	----CCAAAC	CAAACGATTA	ATGACGACCC	GAATC----	-----	TTTTTTATAT	TCATATG-TT
Alphitonia excelsa	CGTGCTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTATAT	TTATATG-TT
Lasiodiscus mildbraedii	CGTACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTATAT	TTATATG-TT
Paliurus spinachristi	CGTACTGAAA TACTATCT--	----CCAAGC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTATAT	TTATATG-TT
Ziziphus glabra	CGTACTGAAA TACTATCT--	----CCAAGC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATATG-TT
Ziziphus ornata	CGTACTGAAA TACTATCT--	----CCAAGC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATATGGTT
Phylica pubescens	CGTACTGAAA TACTGTCT--	----CCAAAC	AAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATATG-TT
Phylica polifolia	CGTACGGAAA TACTGTCT--	----CCCAAC	AAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTATATG-TT
Phylica arborea nitida	CGTACGGAAA TACTGTCT--	----CCAAAC	AAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTAAATG-TT
Nesiota elliptica	CGTACTGAAA TACTGTCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTTTTT	TTA----TT
Noltea africana	CGTACTGAAA TACTGTCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATC----	-----	TTTTTTATAT	TTATA-GGTT
Discaria chacaye	CGTACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	AAGACGACCC	GAATC----	-----	TTTTTTATAT	TTATATG-TT
Spyridium sp1	CGTACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATA--TAT	ATATATA---	---TTTTTTT	TTTATATG-
Spyridium globulosum	CGTACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATA----	-----	---TTTTTTT	TTATATG-TT
Spyridium sp2	CGTACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATA----	-----	---TTTTTTT	TGATA-GATC
Cryptandra sp	CGTACTGAAA TACTATCT--	----CCAAAC	CAAATGATTA	ATGACGACCC	GAATA--TAT	AT-----	---TTTTTTT	ATTTATATG-





Crumenaria erecta	TAGAT-----	-----	---GAAAA	CGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTCAAA-	----GAAAA-
Helinus integrifolius	TATAT-----	-----	---GAAACA	GGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAA-	----GAAAA-
Pleuranthodes hillebrandii	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAA	----GAAAA
Schistocarpaea johnsonii	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AGTAAAAAAA
Colubrina asiatica	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTCAAAA	AAAAAAAAAA
Emmenosperma alphitonioides	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AA-----
Alphitonia excelsa	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAAAAA---
Lasiodiscus mildbraedii	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAAA-----
Paliurus spinachristi	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAA-----
Ziziphus glabra	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAA-----
Ziziphus ornata	TCTAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAAA-----
Phylica pubescens	TATATATATG	-----TTA	ATATGA-AAA	TGAAA---A	-ATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAA-----
Phylica polifolia	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAAA-----
Phylica arborea nitida	TATAT-----	-----	---GAAAA	TGAAA---C	ACTTGTGTG	AATCGATTG	-----	CAAGTAAAAA	AAAA-----
Nesiota elliptica	TATAT-----	---GCTTA-	TATGAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAAA-----
Noltea africana	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATAGATTG	-----	CAAGTAAAAA	AAAA-----
Discaria chacaye	TATATATATG	TTTATATAAA	AAATAAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAAAAA---
Spyridium sp1	TTTATAT---	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAAAAA---
Spyridium globulosum	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAAAAA---
Spyridium sp2	TTT-T-----	-----	---GAAAA	TGAAA---G	-----	-----	-----	-----	-----
Cryptandra sp	TTTATAT---	-----	---GAAAA	CGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAAAAA---
Trymalium sp1	-----	-----	-----	-----	-----	-----	-----	-----	-----
Trymalium sp2	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAAAAA---
Pomaderris rugosa	TATAT-----	-----	---GAAAA	TAAAA---A	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAAAAA---
Siegfriedia darwinioides	-----	-----	-----	-----	-----G	AATCGATTG	-----	CAAGTAAAAA	AAAA-----
Colletia ulicina	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CAAGTAAAAA	AAAA-----
Barbeya oleoides	GAT-----	-----	---AAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	CTACTTGAAA	AAAAGAA---
Dirachma socotrana	TATAGGACAA	ATCAAATATA	G---GACAAA	TGAAA---G	AATTGTTGTG	AAGCGATTG	TCAATCAAGT	CAAGTTGAAA	AAA-GAA---
Dorstenia psilurus	-----	-----	-----	-----	-----	-----	-----	-----	-----
Ficus	CAT-----	-----	---GAAAA	TTAAA---G	AATTATTGTA	AATCAATTA	-----T	TAAGTTGAAA	AAA-GAA---
Artocarpus heterophyllus	CAT-----	-----	---TAAAA	TTAAA---G	AATTGTTGTG	AATCAATTA	-----	TAAGTTGAAA	AAA-GAA---
Cannabis sativa	CAT-----	-----	---GAAAA	TCAAA---G	AATTGTTGTG	AATCAATTC	-----	TAAGTTGAAA	AAT-GAA---
Shepherdia argentea	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????
Hippophae rhamnoides	-----	-----	-----	-----	---G	AATTGATATG	AA-----	---AATAGAA-	-----
Elaeagnus	T-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	ACATTGAAA	AAAAAA---
Dryas drummondii	-----	-----	---GAAAA	TAAAA---G	AATTGTTGTG	AATCGATTG	-----	TAAATTGAAA	AAA-GAA---
Spiraea	-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCGATTG	-----	AAAATTGAAA	AAA-GAA---
Pyrus	-----	-----	---GAAAA	TGAAA---G	ACTTGTGTG	AATCGATTA	-----	AAAATTGAAA	AAA-GAA---
Colubrina reclinata	TATAT-----	-----	---GAAAA	TGAAA---G	AATTGGTGTG	AATCGATTG	-----	CACGTAAAAA	AAAA-----
Gironniera	-----	-----	---GAAAA	TGAAA---G	AATTGTTGTG	AATCAATTC	-----	TAAATTGAAA	AAC-GAA---
Boehmeria	-----	-----	---GAAAA	TAAAA---G	AATTCCTGTG	AATAAATTC	-----	TAAGTTGAAA	AAA-GATA---

\*\*1

Sageretia thea --TCGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAGAAT  
Rhamnus lycioides --TGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TTTGATAGAT AGATCTTTT- ---GAAGAAT  
Frangula alnus --TCGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGATAGAT AGATCTTTT- ---GAAGAAT  
Rhamnella franguloides --TCGAAT-A -TTCATTAAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAGAAT  
Krugiodendron ferreum --TCGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAGAAT  
Rhamnidium cfelaeo --TCGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAGAAT  
Karwinskia humboldtiana --TCGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATC-TTTG ---GAAGAAT  
Condalia microphylla --TCGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AAATCTTTT- ---GAAGAAT  
Scutia buxifolia --TCGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAGAAT  
Berchemia discolor --TCGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAGAAT  
Maesopsis eminii --TCGAAT-A -TT----- --AA-TC--- ----ATTTA CTCTATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAGAAT  
Ventilago viminalis --TCGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T ?GATCTTTT- ---GAAGAAT  
Ventilago leiocarpa --TCGAAT-A -ATCATTGAT CAAA-TC--- ----ATTTA ATCCATCG-- -----TAA TCTGA----T ?GATCTTTT- ---GAAGAAT  
Reynosia uncinata --TCGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T CGATCTTTT- ---GAAGAAT  
Bathiorhamnus cryptophorous --TCGAAT-A -TTCATTGAT CAAA-TT--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T CGATC--TTG ---GAAAAAT  
Ampeloziziphus amazonicus --TCGAAT-A -TTCATTGAT CAAA-TT--- ----ATTTA CTCCGTCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Doerpfeldia cubensis --TCGAAT-A -TTCATTGAT CAAA-TT--- ----ATTTA CTCCGTCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Hovenia dulcis --TGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Ceanothus AATGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Gouania mauritiana --TGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTC CTCCATCG-- -----TAA TCTGA----T AGATCTTTTG TTTGAAAAAT  
Reissekia smilacina --TGGAAT-A -TTCATTGAT CAAA-TCATT ACTCCATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTTG ---GAAAAAT  
Crumenaria erecta --TGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Helinus integrifolius --TGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAACAAT  
Pleuranthodes hillebrandii --CGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Schistocarpaea johnsonii AATGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Colubrina asiatica A-TGGAAT-A -TTCATTGGT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Emmenosperma alphonitonioides --TGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Alphonitonia excelsa --CGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Lasiodiscus mildbraedii --TGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Paliurus spinachristi --TGGAAT-A -TTCATTGAT CAAA-TA--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Ziziphus glabra --TGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTTG ---GAAAAAT  
Ziziphus ornata --TGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Phylica pubescens --TGGAAT-A -TTCGTTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T ATATCTTTT- ---GAAAAAT  
Phylica polifolia --TGGAAT-A -TTCGTTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Phylica arborea nitida --TGGAAT-A -TTCGTTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Nesiotia elliptica --TGGAAT-A -TTCGTTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Noltea africana --TGGAAT-A -TTAGTTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Discaria chacaye --TGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCACCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Spyridium sp1 --TGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT  
Spyridium globulosum --TGGAAT-A -TTCATTGAT CAAA-TC--- ----ATTTA CTCCATCG-- -----TAA TCTGA----T AGATCTTTT- ---GAAAAAT

\*\*\*\*2

Spyridium sp2	-----	-----	-----	-----	-----	-----	-----	-----	-----			
Cryptandra sp	--TGAAT-A	-TTCATTGAT	CAAA-TC--	-----	ATTTA	CTCCATCG--	-----	TAA	TCTGA----T	AGATCTTTT-	---GAAAAAT	
Trymalium sp1	-----	-T-A	---CTATCTC	CAAA-TC--	-----	ATTTA	CTCCATCG--	-----	TAA	TCTGA----T	AGATCTTTT-	---GAAAAAT
Trymalium sp2	--TGAAT-A	-TTCATTGAT	CAAA-TC--	-----	ATTTA	CTCCATCG--	-----	TAA	TCTGA----T	AGATCTTTT-	---GAAAAAT	
Pomaderris rugosa	--TAGAAT-A	-TTCATTGAT	CAAA-TC--	-----	ATTTA	CTCCATCG--	-----	TAA	TCTGA----T	AAATCTTTT-	---GAAAAAT	
Siegfriedia darwinioides	-----	-T-A	-TTCATTGAT	CAAA-TC--	-----	ATTTA	CTCCATCG--	-----	TAA	TCTGA----T	AGATCTTTT-	---GAAAAAT
Colletia ulicina	--TGAAT-A	-TTCATTGAT	CAAA-TC--	-----	ATTTA	CTCCACCG--	-----	TAA	TCTGA----T	AGATCTTTT-	---GAAAAAT	
Barbeya oleoides	--TCGAAT-A	-TTCATTGAT	CAAA-TC--	-----	ATTTA	CTCCATCA--	-----	TAA	TCTGA----T	AGATCTTTT-	---GAAGAAT	
Dirachma socotrana	--TCGAAT-A	-TTCATTAAT	CAAA-TC--	-----	ATTTA	CTCCATCA--	-----	TAA	TCTGA----T	AGATCTTTT-	---GAAGAAT	
Dorstenia psilurus	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	
Ficus	--TTAAAT-A	-TTCATTAAT	CAAA-TC--	-----	ATTTA	CTCCATCA--	-----	AAA	TCTGA----T	AGATCTTTT-	---GAAGAAT	
Artocarpus heterophyllus	--TCAAAT-A	-TTCATTGAT	CAAA-TC--	-----	ATTTA	CTCCATCA--	-----	AAA	CCTGA----T	AGATCTTTT-	---GAAGAAT	
Cannabis sativa	--TTGAAT-A	-TTCATTAAT	CAAA-TC--	-----	ATTTA	CTCCATCA--	-----	AAA	TCTGA----T	AGATTTTTT-	---GAAGACT	
Shepherdia argentea	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	
Hippophae rhamnoides	--TCGAAT-A	-TTCATTGAT	CAAA-TC--	-----	ATTTA	CTCCATCA--	-----	TAA	TCTGA----T	AGATCTTTT-	---GAAGAAT	
Elaeagnus	-----	-T-A	-TTGATTGAT	CAAA-TC--	-----	ATTTA	CTCCATCTAC	TCCATCATAA	TCTGA----C	AGATCTTTT-	---GAAGAAT	
Dryas drummondii	--TCGAAT-A	-TTCATTGAT	CAAA-TC--	-----	ATTCA	CTCCACCA--	-----	TAG	TCTGA----T	AGATCTTTT-	---TAAGAAT	
Spiraea	--TCGACT-A	-TTCATTGAT	CAAA-TC--	-----	ATTCA	CTCCATCA--	-----	TAG	TCTGA----T	AGGTCCTTT-	---TAAGAAT	
Pyrus	--TCGAAT-A	-TTCATTGAT	CAAA-CC--	-----	ATTCA	CTCCACCG--	-----	TAG	TCTGA----T	AGATCTTTT-	---TAATAAT	
Colubrina reclinata	---GGAAT-A	-TTCATTGGT	CAAA-TC--	-----	ATTTA	CTCCATCG--	-----	TAA	TCTGA----T	AGATCTTTT-	---GAAAAAG	
Gironniera	--TCGAAT-A	-TTCATTGAT	CAAA-TC--	-----	ATTTA	CTCCATCA--	-----	AAA	TTTGA----T	AGATCTTTT-	---CAAGACT	
Boehmeria	--TCAAATCA	CTTCTCCAT	CAAAATC--	-----	-----	-----	-----	-----	--TGA----T	AGATTTTTT-	---GAAGAAG	

Sageretia thea	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG
Rhamnus lycioides	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG
Frangula alnus	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG
Rhamnella franguloides	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG
Krugiodendron ferreum	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG
Rhamnidium cfelaeo	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG
Karwinskia humboldtiana	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG
Condalia microphylla	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG
Scutia buxifolia	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG
Berchemia discolor	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG
Maesopsis eminii	CGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG-----	--TCAATATC	-----	-----	-----	GACAATAATG
Ventilago viminalis	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTTT	ACATG-----	--TTAATATC	-----	-----	-----	GACAACAATG
Ventilago leiocarpa	TGATTAATCG	G-ACGAGTAA	TAAAGATAGA	GTCCCATTTT	ACATG-----	--TTAATATC	-----	-----	-----	GACAACAATG
Reynosa uncinata	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG-----	--TCAATATC	G-----A	TGTCAATATC	-----	GACAACAATG
Bathiorhamnus cryptophorous	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTTT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG
Ampeloziziphus amazonicus	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTTT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG
Doerpfeldia cubensis	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTTT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG
Hovenia dulcis	TTATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTTT	ACATG-----	--TCAATATC	-----	-----	-----	GGCAACAATG
Ceanothus	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTTT	ACATG-----	--TCAATATC	-----	-----	-----	GACAACAATG

Gouania mauritiana	TGACTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATGCTACA	TGTCAATATT	-----	-----	GGCAACAATG
Reissekia smilacina	TGACTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GGCAACAATG
Crumenaria erecta	TGACTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GGCAACAATG
Helinus integrifolius	TGACTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATT	-----	-----	GGCAACAATG
Pleuranthodes hillebrandii	TGACTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GGCAACAATG
Schistocarpaea johnsonii	GGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Colubrina asiatica	TGATTACTCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Emmenosperma alphitonioides	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Alphitonia excelsa	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Lasiodiscus mildbraedii	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Paliurus spinachristi	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Ziziphus glabra	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Ziziphus ornata	GGATTAATCC	GGACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Phylica pubescens	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Phylica polifolia	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Phylica arborea nitida	GGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Nesiota elliptica	GGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Noltea africana	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Discaria chacaye	GGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Spyridium sp1	GGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Spyridium globulosum	GGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Spyridium sp2	--ATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Cryptandra sp	GGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Trymalium sp1	GGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Trymalium sp2	GGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Pomaderris rugosa	GGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACGTG----	--TCAATATC	-----	-----	GACAACAATG
Siegfriedia darwinioides	GGATTAATCC	GGACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Colletia ulicina	GGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Barbeya oleoides	TGATAAATCC	GGACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Dirachma socotrana	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Dorstenia psilurus	----TAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCAGTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Ficus	GGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Artocarpus heterophyllus	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	GACAACAATA	TGTCAATATC	GACAACAATG
Cannabis sativa	TGATAAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Shepherdia argentea	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????
Hippophae rhamnoides	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Elaeagnus	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Dryas drummondii	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Spiraea	TAATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Pyrus	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Colubrina reclinata	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCGTTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG
Gironniera	TGATTAATCG	G-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	GATG----	--TCAATATC	-----	-----	GACAACAATG
Boehmeria	TGATTAATCG	T-ACGAG-AA	TAAAGATAGA	GTCCCATTCT	ACATG----	--TCAATATC	-----	-----	GACAACAATG

\*\*\*\*\*3

\*\*\*\*\*4

Sageretia thea	CAATTTATAG	TAAAAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	CA-----
Rhamnus lycioides	CAATTTATAG	TAAAAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	CA-----
Frangula alnus	CAATTTATAG	TAAAAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	CAAAAAGTACC
Rhamnella franguloides	CAATTTATAG	TAAAAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	CA-----
Krugiodendron ferreum	CAATTTATAG	TAAAAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	CA-----
Rhamnidium cfelaeo	CAATTTATAG	TAAAAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	CA-----
Karwinskia humboldtiana	CAATTTATAG	TAAAAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	CA-----
Condalia microphylla	CAATTTATTG	TAAAAGGAGG	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	CT-----
Scutia buxifolia	CAATTTATAG	TAAAAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	CA-----
Berchemia discolor	CAATTTATAG	TAAAAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	CA-----
Maesopsis eminii	CAATTTATAG	TAAAAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	CA-----
Ventilago viminalis	ATATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-ATAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	TA-----
Ventilago leiocarpa	ATATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-ATAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	CA-----
Reynosia uncinata	CAATTTATAG	TAAAAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGTCC	CG-----
Bathiorhamnus cryptophorous	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CA-----
Ampeloziziphus amazonicus	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CA-----
Doerpfeldia cubensis	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CA-----
Hovenia dulcis	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CT-----
Ceanothus	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CA-----
Gouania mauritiana	CAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-ATAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCCAAC	CAAAAAGGC-	CA-----
Reissekia smilacina	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-ATAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCCAAC	CAAAAAGGCC	CA-----
Crumenaria erecta	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-ATAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCCAAC	CAAAAAGGCC	CA-----
Helinus integrifolius	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCCAAC	CCAAAAGGCC	CG-----
Pleuranthodes hillebrandii	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-ATAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCCAAC	CAATAGGGCC	CG-----
Schistocarpea johnsonii	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CT-----
Colubrina asiatica	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CA-----
Emmenosperma alphonoioides	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CA-----
Alphitonia excelsa	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CA-----
Lasiodiscus mildbraedii	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAGAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CA-----
Paliurus spinachristi	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CA-----
Ziziphus glabra	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CA-----
Ziziphus ornata	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CA-----
Phylica pubescens	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-----ACC	CA-----
Phylica polifolia	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-----ACC	CA-----
Phylica arborea nitida	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-----ACC	CA-----
Nesiota elliptica	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-----ACC	CA-----
Noltea africana	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-----ACC	CA-----
Discaria chacaye	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC---	-AAAAAGGCC	CA-----

Spyridium sp1	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	CA-----
Spyridium globulosum	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	CA-----
Spyridium sp2	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	CA-----
Cryptandra sp	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	CA-----
Trymalium sp1	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	CA-----
Trymalium sp2	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	TGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	CA-----
Pomaderris rugosa	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	CA-----
Siegfriedia darwinioides	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	CA-----
Colletia ulicina	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	CA-----
Barbeya oleoides	AAAATTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	TA-----
Dirachma socotrana	AAATTTATAG	TAAAGGAAA	ATCCGTCGAC	TTTTAGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAGAG-	-----
Rorstenia psilurus	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AAAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAGGGCC	CA-----
Ficus	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AAAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAGGGCC	CA-----
Artocarpus heterophyllus	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AAAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGTC	CA-----
Cannabis sativa	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AAAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	AA-----
Shepherdia argentea	??????????	??????????	??????????	??????????	??????????	??????????	??????????	?	AAAGTCC?T?	TATCCCC
Hippophae rhamnoides	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	TA-----
Elaeagnus	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAGGGCC	TA-----
Dryas drummondii	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGACC	TG-----
Spiraea	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAACGACC	TG-----
Pyrus	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAACGACC	CG-----
Colubrina reclinata	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AGAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	CA-----
Gironniera	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AAAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGGCC	AA-----
Boehmeria	AAATTTATAG	TAAGAGGAAA	ATCCGTCGAC	TTT-AAAAAT	CGTGAGGGTT	CAAGTCCCTC	TATCCCC	---	-AAAAAGTCC	CA-----

\*\*\*\*\*5

Sageretia thea	-TTGGATTCC	CTAATTATTT	----ATCCTA	TGCTCTCATT	TCGTT-----	AACGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Rhamnus lycioides	-TTGGATTCC	CTAATTATTT	----ATCCTA	TGCTCTCATT	TCGTT-----	AACGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Frangula alnus	ATTGGATTCC	CTAATTATTT	----ATCCTA	TGCTCTCATT	TCGTT-----	AACGGTTCAA	AA-----TTC	CTTATGTTTC	T-----CA
Rhamnella franguloides	-TTGGATTCC	CTAATTATTT	----ATCCTA	TGCTCTCATT	TCGTT-----	AACGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Krugiodendron ferreum	-TTGGATTCC	CTAATTATTT	----ATCCTA	TGCTCTCATT	TCGTT-----	AACGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Rhamnidium cfelaeo	-TTGGATTCC	CTAATTATTT	----ATCCTA	TGCTCTCATT	TCGTT-----	AACGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Karwinskia humboldtiana	-TTGGATTCC	CTAATTATTT	----ATCCTA	TGCTCTCATT	TCGTT-----	AACGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Condalia microphylla	ATTGGATTCC	CTAATTATTT	----ATCCTA	TGCTCTCAGT	TCGTC-----	AACGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Scutia buxifolia	-TTGGATTCC	CTAATTATTT	----ATCCTA	TGCTCTCATT	TCGTT-----	AACAGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Berchemia discolor	-TTGGATTCC	CTAATTATTT	----ATCCTA	TGCTCTCATT	TCGTT-----	AACGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Maesopsis eminii	-GTGGATTCC	CTAATTATTT	----ATCCTA	TGCTCTCATT	TCGTT-----	AACGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Ventilago viminalis	-TAAGATTCC	CTAATTATTT	----ATCCTA	TGCTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Ventilago leiocarpa	-TAAGATTCC	CTAATTATTT	----ATCCTA	TGCTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Reynosia uncinata	-TTGGATTCC	CGAATTATTT	----ATCCTA	TGCTCTCATT	TCGTT-----	AACGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA

\*\*\*\*\*6

Bathiorhamnus cryptophorous	-TTGATGCC	CTAATTATTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGGTTC	G-----CA
Ampeloziziphus amazonicus	-TTGGATGCC	CTAATTATTT	----ATCCTA	TACTCTCATT	TTGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGGTTC	G-----CA
Doerpfeldia cubensis	-TTGTATGCC	CTAATTATTT	----ATCCTA	TACTCTCATT	TTGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGGTTC	G-----CA
Hovenia dulcis	-TTTGATGCC	CTAATT-TTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGAGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Ceanothus	-TTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Gouania mauritiana	-TTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TCC	GTTATGTTTC	T-----CA
Reissekia smilacina	-TTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AC-----TCT	GTTATGTTTC	T-----CA
Crumenaria erecta	-TTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AC-----TCT	GTTATGTTTC	T-----CA
Helinus integrifolius	-TTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TCC	GTTATGTTTC	T-----CA
Pleuranthodes hillebrandii	-TTGATGCC	GTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AC-----TCT	GTAATGTTTC	T-----CC
Schistocarpaea johnsonii	-TTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Colubrina asiatica	-TTTGATGCC	CTAATTTTTT	----ATGCTA	TACTCTCATT	TCATTTTCGTT	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	TCGTTTCTCA
Emmenosperma alphonoides	-TTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	CCGCT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	C-----CA
Alphitonia excelsa	-TTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	G-----CA
Lasiodiscus mildbraedii	-TTTGATGCC	CTAGTATTTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Paliurus spinachristi	-TTTGATGCC	CTAATT-TTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Ziziphus glabra	-TTTGATGCC	CTAATT-TTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Ziziphus ornata	-TTTGATGCC	CTAATT-GTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Phylica pubescens	-TTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Phylica polifolia	-TATGATGCC	TTAATTTTTT	----ATCATA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Phylica arborea nitida	-TTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Nesiotia elliptica	-TTTGATGCC	CCTAATTTTT	T---ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Noltea africana	-TTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Discaria chacaye	-TTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGCGGTTCAA	AA-----TTC	GTTAGGTTTC	T-----CA
Spyridium sp1	-GTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCATT-----	ATCGGTTCAA	AA-----TTC	ATTATGTTTC	T-----CA
Spyridium globulosum	-GTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	ATCGGTTCAA	AG-----TTC	ATTATGTTTC	T-----CA
Spyridium sp2	-GTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	ATCGGTTCAA	AA-----TTC	ATTATGTTTC	T-----CA
Cryptandra sp	-GTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	ATCGGTTCAA	AA-----TTC	ATTATGTTTC	T-----CA
Trymalium sp1	-GTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	ATCGGTTCAA	AA-----TTC	ATTATCTTTC	T-----CA
Trymalium sp2	-GTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	ATCGGTTCAA	AA-----TTC	ATTATCTTTC	T-----CA
Pomaderris rugosa	-GTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	ATCGGTTCAA	AA-----TTC	ATTATCTTTC	T-----CA
Siegfriedia darwinioides	-GTTGATGCC	CTAATTTTTT	----ATCCTA	TACTCTCATT	TCGTT-----	ATCGGTTCAA	AA-----TTC	ATTATCTTTC	T-----CA
Colletia ulicina	-TTTGATGCC	CTCATTTTTT	TTTTATCCTA	TACTCTCATT	TCATT-----	AGTGGTTCAA	AA-----TTC	GTTAGGTTTC	T-----CA
Barbeya oleoides	-TTTGATGCC	CTAATTATTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGGATTCAA	AA-----TTC	GCTATGTTTC	T-----CA
Dirachma socotrana	----GATGCC	CTAATTATTT	----ATCCTA	TACTCTCATT	TCATT-----	AGGATTCAA	AA-----TTC	GCTATGTTTC	T-----CA
Dorstenia psilurus	-TTTGATGCC	CTAATTATTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGGATTCAA	AA-----TTC	GCTATGTTTC	T-----CA
Ficus	-TTTGATGCC	CTAATTATTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGGATTCAA	AA-----TTC	GCTATGTTTC	T-----CA
Artocarpus heterophyllus	-TCTGATGCC	CTAATTATTT	----ATCCTA	TACTCTCATT	TCGTT-----	AGGATTCAA	AA-----TTC	GCTATGTTTC	T-----CA
Cannabis sativa	-ATG-ATGCC	CTAATTATTT	----ATCC--	TACTCTCATT	CCGTT-----	AGTGGTTCT	AA-----TTC	GTTATGTTTC	T-----CG
Shepherdia argentea	-TTTGATGCC	?AATTATTT	----ATCCCC	TACTCTCATT	TCGTT-----	AGGATTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Hippophae rhamnoides	-TTTGATGCC	CTAATTATTT	----ATCCCC	TACTCTCATT	TCGTT-----	AGGATTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Elaeagnus	-TTTGATGCC	C-----	-----C	TACTCTCATT	TCGTT-----	AGTGGTTCT	-----TTC	GTTATGTTTC	T-----CA
Dryas drummondii	---CGACTCA	TTAATTATTT	----ATTTTC	TACTCTCATT	TTGTT-----	AGGATTCAA	AA-----TTC	GTTATGTTTC	T-----CA



Spiraea	-GTTGACTCC	CTAATTATT	----ACTTTA	TCA-----TT	TGTT-----	AGGGATTCAA	AA-----TTC	GTTATGTTTC	T-----CA
Pyrus	-GTTGACTCC	CTAATTATT	----ATTTTC	--ATTTTATC	ATTTT---GT	AGCGATTCAA	ATAAAAATTC	GTTATATTTA	T-----CA
Colubrina reclinata	-TTTGATTCC	CTAATTTTT	----ATCCTA	TACTCGCATT	TCATTTTCGTT	AGCGGTTCAA	AA-----TCC	GTTATGTTTC	TCGTTTCTCA
Gironniera	-ATGAATTCC	CTAATTATT	----ATCCTA	TCCTTCTCATT	CCATT-----	AGCGGTTCAA	AA-----TTC	GTTATCTTTC	T-----CG
Boehmeria	-TTTGATTCC	CTAATTTTT	----ATCCTA	TCTTCTCAGT	TAATT-----	AGCAGTTCAA	AA-----TTC	GTCATGTTTC	T-----CG

\*\*\*\*\*7

\*\*\*\*\*8

\*\*\*\*\*9

Sageretia thea	TTCATT----	--CTACTCTT	TTACTTT-AC	AAATGG----	-----TCT	GAGCGGAAAT	TT-TTTTCTT	T-TCACAA--	-----GCCTT
Rhamnus lycioides	TTCATT----	--CTACTCTT	TTAGTTT-AC	AAATGG----	-TAATGGTCT	GAGCGGAAAT	TTT-TTTATT	T-TCACAA--	-----GCCTT
Frangula alnus	TTCAGT----	--CTACTCTT	TTCCTTT-AC	AAATGG----	-TAATGGTCT	GAGCGGAAAT	T-CTTTTCTT	T-TCACAA--	-----GCCTT
Rhamnella franguloides	TTCATT----	--CGACTCTT	TTACTTT-AC	AAATGG----	-----TCT	GAGCGGAAAT	TAGTTT-CTT	T-TCACAA--	-----GCCTT
Krugiodendron ferreum	TTCATT----	--CTACTCTT	TTACTTT-AC	AAATGG----	-----TCT	GAGCGGAAAT	TT-TTTTCTT	T-TCACAA--	-----GCCTT
Rhamnidium cfelaeo	TTCATT----	--CTACTCTT	TTACTTT-AT	AAATGG----	-----TCT	GAGCGGAAAT	TT-TTTTCTT	T-TCACAA--	-----GCCTT
Karwinskia humboldtiana	TTCATT----	--CTACTCTT	TTACTTT-AT	AAATGG----	-----TAT	GAGCGGAAAT	TT-TTTTCTT	T-TCACAA--	-----GCCTT
Condalia microphylla	TTCATT----	--CTACTCTT	TTACTTT-AC	AAATGG----	-----TCT	GAGCGGAAAT	TTGTTT-CTT	T-TCACAA--	-----GCCTT
Scutia buxifolia	GTCATT----	--CTACTCTT	TTACTTT-AC	AAATGG----	-----TCT	GAGCGGAAAT	TT-TTTTCTT	T-TCACAA-T	CACAAGCCTT
Berchemia discolor	TTCATT----	--CTACTCCT	TTACTTT-AC	AAATGG----	-----TCT	GAGCGGAAAT	T-GTTTTCTT	T-TCACAA--	-----GCCTT
Maesopsis eminii	TTCATT----	--CTACTCTT	TTACTTT-AC	AAAGGG----	-----TCT	GAGCGGAAAT	CTTTTT-CGT	-ATCACAA--	-----GCCTT
Ventilago viminalis	TTCATT----	--TTACTTTT	TTACTTT-AC	AAATGG----	-----TTT	GAGCGGAAAT	TT-TTTTCTT	T-TCACAA--	-----GCCTT
Ventilago leiocarpa	TTCATT----	--TTACTTTT	TTACTTT-AC	AAATGG----	-----TTT	GAGCGGAAAT	TT-TTTTCTT	T-TCACAA--	-----GCCTT
Reynosia uncinata	TTCATT----	--CTACTCTT	TTACTTT-AC	AAATGG----	-----TCT	GAGCGGAAAT	TT-TTTTCTT	T-TCACAA--	-----GCCTT
Bathiorhamnus cryptophorous	TTCATT----	--CTACTCTT	TT-----AC	AAATGGA----	-----TCT	GAGCGTAAA-	-----TTCTT	T-TCACAA--	-----GCCTT
Ampeloziziphus amazonicus	TTCATT----	--CTACTCTT	TT-----AC	AAATGAA----	-----TCT	AATCGTAAA-	-----TTCTT	T-TCACAA--	-----GCCTT
Doerpfeldia cubensis	TTCATT----	--CTACTCTT	TT-----AC	AAATGTA----	-----TCT	AATCGTAAA-	-----TTCTT	T-TCACAA--	-----GCCTT
Hovenia dulcis	TTCATT----	-----	-----	-----	-----	-----AAT	TT-TTTTCTT	T-TCACAA--	-----GCCTT
Ceanothus	TTCATT----	-----	-----	-----	-----	-----AAT	TT-TTTTCTT	T-TCACAA--	-----GCCTT
Gouania mauritiana	TCCATT----	-----	-----	-----	-----	-----AAT	TTCT-----	---CACAA--	-----GCCTT
Reissekia smilacina	TTCATT----	-----	-----	-----	-----	-----AAT	TTTT-----	---CACAA--	-----GCCTT
Crumenaria erecta	TTCATT----	-----	-----	-----	-----	-----AAT	TTTT-----	---CACAA--	-----GCCTT
Helinus integrifolius	TTCATT----	-----	-----	-----	-----	-----AAG	TTTTTTTCGT	T-TCACAA--	-----GCCTT
Pleuranthodes hillebrandii	GTCATT----	-----	-----	-----	-----	-----AAT	TTTT-----	---CACAA--	-----GCCTT
Schistocarpaea johnsonii	TTCATT----	-----	-----	-----	-----	-----AAT	TT-TTTTCTT	T-TCACAA--	-----GCCTT
Colubrina asiatica	TTCATT----	-----	-----	-----	-----	-----AAT	TT-TTTTCTT	T-TCACAA--	-----GCCTT
Emmenosperma alphonoides	TTCATT----	-----	-----	-----	-----	-----AAT	TT-TTTTCTT	T-TCACAA--	-----GCCTT
Alphitonia excelsa	TTCATT----	-----	-----	-----	-----	-----AAT	CTTTTT-CTT	T-TCACAA--	-----GCCTT
Lasiodiscus mildbraedii	TTCATT----	-----	-----	-----	-----	-----AAT	TTGTTT-CTT	T-TCACAA--	-----GCCTT
Paliurus spinachristi	TTCATT----	-----	-----	-----	-----	-----AAT	TT-TTTTCTT	T-TCACAA--	-----GCCTT

Ziziphus glabra	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TT-TTTTCTT T-TCACAA--	-----	GCCTT
Ziziphus ornata	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TTTTTT-CTT T-TCACAA--	-----	ACCTT
Phylica pubescens	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TTTTTG-ATT T-TCACAA--	-----	GCCTT
Phylica polifolia	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TTTTTG-ATT T-TCACAA--	-----	GCCTT
Phylica arborea nitida	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TTTTTG-ATT T-TCACAA--	-----	GCCTT
Nesiotia elliptica	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	CAT TTTTTT-ATT T-TCACAA--	-----	GCCTT
Noltea africana	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TTTTTT-ATT T-TCACAA--	-----	GCCTT
Discaria chacaye	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	CAT TTTTTT-CTT T-TCACAA--	-----	GCCTT
Spyridium sp1	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TTTTTT-CTT T-TCACAA--	-----	GCCTT
Spyridium globulosum	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	CAT TTTTTT-CTT T-TCACAA--	-----	GCCTT
Spyridium sp2	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TTTTTT-CTT T-TCACAA--	-----	GCCTT
Cryptandra sp	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TTTTTT-CTT T-TCACAA--	-----	GCCTT
Trymalium sp1	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TTTTTT-CTT T-TCACAA--	-----	GCCTT
Trymalium sp2	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TTTTTT-CTT T-TCACAA--	-----	GCCTT
Pomaderris rugosa	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TTTTTT-CTT T-TCACAA--	-----	GCCTT
Siegfriedia darwinioides	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TTTTTT-CTT T-TCACAA--	-----	GCCTT
Colletia ulicina	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	CAT TTTTTT-CTT T-TCACAA--	-----	GCCTT
Barbeya oleoides	TTCATTGAA	TTTTACTCTT	TC-----	AC	AAACGAGTCT	GAACGGGTCT	GGGCGGAAAT	TTG-TTTATT	-ATCACAAAT	C-----	TT	
Dirachma socotrana	CTCATTCTAA	TTCTACTTTT	TC-----	AC	AAAC---TCA	CAAACGG---	-AGGAAAAT	TT-TTTTCTT	-ATCACATAT	CATAT--	AT	
Dorstenia psilurus	TTCATTTTAA	TTCT-----	-----	AC	AAACGTA---	-----	TTT	GATCAAAAAT	TT-TTTTCTT	-ATCACAA--	-----	GCCTT
Ficus	TTCATTCTAA	TTTT-----	-----	AC	AAACGTA---	-----	TCT	GAGCGAAAAT	CT-TTTTCTT	TATCACAA--	-----	GCCTT
Artocarpus heterophyllus	TTCATTCTAA	TTCT-----	-----	AC	AAACGTA---	-----	TCT	GAGCGAAAAT	TT-TTTTCTT	-ATCACAA--	-----	GCCTT
Cannabis sativa	TTCATTCTAA	CTTT-----	-----	AC	AAACGGGA---	-----	CCT	GAATGACCTT	TT-TTTTATT	-ATCACAA--	-----	GCCTT
Shepherdia argentea	TTCATTCTAA	CTATAATCTT	TCGCATTAC	AAATGTA---	-----	TTT	GATCAGAAAAT	TT-TTTTCTT	-ATCACAA--	-----	GCCTT	
Hippophae rhamnoides	TTCATTCTAA	CTATAATCTT	TCGCATTAC	AAATGTA---	-----	TTT	GTCGCATAT	CATATG---	T	-ATCACAG--	-----	GCCTT
Elaeagnus	TTCATTCTAA	CTATAATCTT	TCGCATTAC	AAATGTA---	-----	TTT	TCATAAAT	TT-TTTTCTT	-ATCACAA--	-----	GCCTT	
Dryas drummondii	GTCATTATCA	TTCTACTCTT	TC-----	AC	AAGCGGA---	-----	TCT	GAGCGTAAAT	TTATTTTCTT	-ATCACAA--	-----	GCCTT
Spiraea	TTCATTCTCA	TTATACTCTT	TC-----	AC	AAACGTA---	-----	TCT	GAGCGTAAAT	TT-TTTTCT-	-ATCACAA--	-----	GCCTT
Pyrus	TTCATTCTCA	TTATACTCTT	TT-CTTTCAC	AAATGGA---	-----	TCT	GAGCGAAAAT	TT-TTTTCTT	-ATCAAAA--	-----	GACTT	
Colubrina reclinata	TTCATT----	-----	-----	-----	-----	-----	-----	-----	-----	AAT TT-TTTTCTT T-TTACAA--	-----	GCCTT
Gironniera	TTCATTATAA	TTCT-----	-----	AC	AAACGGGA---	-----	TCT	GAGCGGAAAAT	TT-TTTTCTT	-ATCACTTAT	CACAGGCCTT	
Boehmeria	GTCATT----	-----	-----	-----	-----	-----	CT	GAACGGAAAAT	-T-TTTTCTT	-ATCAAAA--	-----	GATTT

\*\*

Sageretia thea	GTGATA----	-----	--TATAGGAT	AC-----	AC	GTACAAATGA	ACATCGTTGG	GC----	ACGT	AACCCCGATT	GT-AAATT--	
Rhamnus lycioides	GTGATA--TT	GT-----	GA	TATATAGGAT	AC-----	AC	GTACAAATGA	ACATCGTTGG	GC----	ACGT	AACCCGTGATT	GT-AAATT--
Frangula alnus	GTGATA----	-----	--TATAGGAT	AC-----	AC	GTACAAATGA	ACATCGTTGG	GC----	ACGT	AACCCCGATC	GT-AAATT--	
Rhamnella franguloides	GTGATATCTT	GT-----	GA	TATATAGGAT	AC-----	AC	GTACAAATGA	ACATCGTTGG	GC----	ACGT	AACCCCGATT	GT-AAATT--
Krugiodendron ferreum	GTGATA----	-----	--TATAGGAT	AC-----	AC	GTACAAATGA	ACATCGTTGG	GC----	ACGT	AACCCCGATT	GT-AAATT--	
Rhamnidium cfelaeo	GTGATA----	-----	--TATAGGAT	AC-----	AC	GTACAAATGA	ACATCGTTGG	GC----	ACGT	AACCCCGATT	GT-AAATT--	
Karwinskia humboldtiana	GTGATA----	-----	--TATAGGAT	AC-----	AC	GTACAAATGA	ACATCGTTGG	GC----	ACGT	AACCCCGATT	GT-AAATT--	
Condalia microphylla	GTGATA--TT	GTGATATTGA	TATATAGGAT	AC-----	AC	GTACAAATGA	ACATCGTTGG	GC----	ACGT	AACCCCGATT	GT-AAATT--	
Scutia buxifolia	GTGATA----	-----	--TATAGGAT	AC-----	AC	GTACAAATGA	GCATCGTTGG	GC----	ACGT	AACCCCGATT	GT-AAATT--	

Berchemia discolor	GTGGTA----	-----	--TATAGGAT	AC-----AC	GTACAAATGA	ACATCGTTGG	GC----GCGT	AACCCCGATT	GT-AAATT--	
Maesopsis eminii	GTGATA----	-----	--TATAGGCT	AC-----AC	GTACAAATGA	ACATCGTTGG	GC----ACGT	AACCCCGATT	GT-AAATT--	
Ventilago viminalis	GTGATA----	-----	--TATAGGAT	AC-----AC	GTACAAATAA	ACATCGTTGA	GC----ACGT	AACAACGATT	G-AAAATT--	
Ventilago leiocarpa	GTGATA----	-----	--TATAGGAT	AC-----AC	GTACAAATGA	ACATCGTTGA	GC----ACGT	AACAACGATT	G-AAAATT--	
Reynosia uncinata	GTGATA----	-----	--TATAGGAT	AC-----AC	GTACAAATGA	ACATCGTTGG	GC----ACGT	AACCCCGATT	GT-AAATT--	
Bathiorhamnus cryptophorous	TTGATA----	-----	--TATACGAT	AC-----AC	GTACAAATGA	ACATCGTTGA	GC----AAGT	AATCCCGATT	GT-AAATT--	
Ampeloziziphus amazonicus	TTGATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCGTTGA	GC----AAGT	AATCCCGATT	GT-AAATT--	
Doerpfeldia cubensis	TTTTTT-----	-GA	TATATATGAT	AC-----AC	GTACAAATGA	ACATCGTTGA	CC-----	-ACTCCA--	-----TT--	
Hovenia dulcis	GTGATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Ceanothus	GTGATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Gouania mauritiana	GT-ATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTGA	GC----AAGT	AACCTCGATT	GG-AAATTGA	
Reissekia smilacina	GT-ATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTGA	GC----AAGT	AACTCCGATT	GT-AAATT--	
Crumenaria erecta	GT-ATA----	-----	--TATATGAT	AC-----AC	GTACAAACGA	ACATCATTGA	GC----AAGT	AACTCCGATT	GG-AAATT--	
Helinus integrifolius	GTGATA----	-----	--TATATGAT	AC-----CC	GTACAAATGG	AATCATTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Pleuranthodes hillebrandii	GT-ATA----	-----	--TATATGAT	AC-----AC	GTACAAGTGG	ACATCATTGA	GC----AAGT	AACTCCGATT	GG-AAATT--	
Schistocarpaea johnsonii	GTGATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-AAACT--	
Colubrina asiatica	GTGATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-ATATT--	
Emmenosperma alphitonioides	GTGATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTGA	GC----AAGT	AACCCCAATT	GT-AAATT--	
Alphitonia excelsa	GTGGTA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCGTTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Lasiodiscus mildbraedii	GTG-TA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATCGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Paliurus spinachristi	GTGATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTGA	GC----AAGT	AACCTTGATT	GT-AAATT--	
Ziziphus glabra	TTGATA----	-----	--TATACGAT	AC-----AC	GTACAAATGA	ACATCATTGA	GC----AAGT	AA'CCCGATT	GT-AAATT--	
Ziziphus ornata	GTGGTA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Phylica pubescens	GTGATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTTA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Phylica polifolia	GTGATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTTA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Phylica arborea nitida	GTAATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTTA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Nesiota elliptica	GTGATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Noltea africana	GTGATA----	-----	--TATATGAT	AC-----AC	G?ACAAATGA	ACATCATTGA	GC----ACGT	AACCCCGATT	GT-AAATT--	
Discaria chacaye	GTGATA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Spyridium sp1	GTGGTA----	-----	--TATATGAT	AC-----AC	ATACAAATGA	ACATCATTGA	GC----AAGT	AAACCCGATT	GT-AAATT--	
Spyridium globulosum	GTGGTA----	-----	--TATATGAT	AC-----AC	ATACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Spyridium sp2	GTGGTA----	-----	--TATATGAT	AC-----AC	ATACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Cryptandra sp	GTGGTA----	-----	--TATATGAT	AC-----AC	ATACAAATGA	ACATCGTTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Trymalium sp1	GTGGTA----	-----	--TATATGAT	AC-----AC	ATACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Trymalium sp2	A-----	-----	--TATATGAT	AC-----AC	ATACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Pomaderris rugosa	GTGGTA----	-----	--TATATGAT	AC-----AC	ATACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Siegfriedia darwinioides	GTGGTA----	-----	--TATATGAT	AC-----AC	ATACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-AAATT--	
Colletia ulicina	GTGGTA----	-----	--TATATGAT	AC-----AC	GTACAAATGA	ACATCATTGA	GC----AAGT	AACCCCGATT	GT-AAAT---	
Barbeya oleoides	GGGATA----	-----	--TATATGAT	TATGATATAC	GTACAAATGA	ACATCTTTGA	GC----AAGC	AATCCCGATT	TC-AAATT--	
Dirachma socotrana	ATGATA----	-----	-CACG	TACAAATGAT	AC-----AC	GTACAAATGA	ACATCTTTGA	GCAAACAAGT	AATCCCGATT	TT-AAAA--
Dorstenia psilurus	GTGATT----	-----	--CATATGAA	AC-----AC	GTACAAATGA	ACATCTTTGA	GA----AGGG	AACCC-ACG	TT-AAATC--	
Ficus	GTGATC----	-----	--TATATGAA	AG-----AC	GTACAAATGA	ACATCTTTGA	GA----AAGG	AATCCCAATG	TT-AAATT--	
Artocarpus heterophyllus	GTGATA----	-----	--GAGGTGAA	AC-----AC	GTACAAATGA	ACATCTTTGA	GA----AAGG	AATCCCAATG	TT-AAATT--	

Cannabis sativa	GTGATA----	-----	--TATATGAA	AG-----	AC	CTACAAATGA	ACAT-----	-----	AAGG	AATCCCAATG	TG-CAATT--	
Shepherdia argentea	ATGATA----	-----	--CATATGAT	AT-----	GC	GGACAAATGA	ATATCTTTGA	GC----	AAGT	CATTCCAATP	TC-AAATP--	
Hippophae rhamnoides	GTGATA----	-----	--CATATGAT	AT-----	GC	GGACAAATCA	ATATCTTTGA	GCAA--	AAGT	AATTCCCATT	TC-AAATP--	
Elaeagnus	GCGATA----	-----	---CATATG	ATATTATCAT	AT-----	GC	GGACAAATGA	ATATCTTTGA	AC----	AAGT	AATTCCAATP	TG-AAA----
Dryas drummondii	GCGTGT----	-----	-----GATA	TATATATGAT	AT-----	GC	GTATAAATGA	ACATCTTTGA	GT----	AAGG	AATCCCCATT	T-AAAATP--
Spiraea	GTGTAT----	-----	---GA	TATATATGAT	AC-----	AC	GTACAAATGA	ACAGCGTTGA	GA----	AAGG	AATCCCCATT	TTAAAATP--
Pyrus	GTGTGT----	-----	---GA	TATATATGAT	AC-----	GC	GTACAGTACA	AATGATTTGA	GC----	AAGG	AATCC--ATT	---AAATP--
Colubrina reclinata	GTGATA----	-----	--TATATGAT	AC-----	AC	GTACAAATGA	ACATTATTGA	GC----	AAGT	AACCCCGATT	TT-ATATP--	
Gironniera	GTAGTA----	-----	--TATATGAA	AT-----	AC	TTACAAATGA	ACAT-----	-----	AAGT	AATCCCAATA	TT-AAATP--	
Boehmeria	GTGATA----	-----	--TATATGAA	AA-----	AC	GTACAAATGA	ACATGTTTGA	GA----	AAGG	AATCCTAATA	TT-AAATA--	

\*\*\*\*\*10

\*\*\*\*\*11

\*\*\*\*\*12

Sageretia thea	-----	---	GTAATG	ATTAACAATA	CATATTATT-	---	ACT---	-----	-TGTACTG--	---	AAACTTA	CAAAGTCTTC
Rhamnus lycioides	-----	---	GTAATG	ATTAACAATA	CATATTCTT-	---	ACT---	-----	-TGTATGTA	CTGAAACGTA	CAAAGTCTTC	
Frangula alnus	-----	---	GTAATA	ATTAACAATA	CATATTATT-	---	ACT---	-----	-TGTACTGTA	CTGAAACGTA	CAAAGTCTTC	
Rhamnella franguloides	-----	---	GTAATG	ATTAACAATA	CATATTATT-	---	ACT---	-----	-TGTACTGTA	CTGAAACTTA	CAAAGTCTTC	
Krugiodendron ferreum	-----	---	GTAATG	ATTAACAATA	CATATTATT-	---	ACT---	-----	-TGGACTGTA	CTGAAACTTA	CAAAGTCTTC	
Rhamnidium cfelaeo	-----	---	GTAATG	ATTAACAATA	CATATTATT-	---	ACT---	-----	-TGTACTGTA	CTGAAACTTA	CAAAGTCTTC	
Karwinskia humboldtiana	-----	---	GTAATG	ATTAACAATA	CATATTATT-	---	ACT---	-----	-TGTACTGTA	CTGAAACTTA	CAAAGTCTTC	
Condalia microphylla	-----	---	GTAATG	ATTAACAATA	CATATTATT-	---	ACT---	-----	-TGTACTGTA	CTGAAACTTA	CAAAGTCTTC	
Scutia buxifolia	-----	---	GTAATG	ATTAACAATA	CATATTATT-	---	ACT---	-----	-TGTACTGTA	CTGAAACTTA	CAAAGTCTTC	
Berchemia discolor	-----	---	GTAATG	ATTAACAATA	CATATTATT-	---	AGT---	-----	-TGTACTGTA	CTGAAACTTA	CAAAGTCTTC	
Maesopsis eminii	-----	---	GTAATG	ATTAACAATA	CATATTATT-	---	ACT---	-----	-CGTAGTGA	CTGAAACTTA	CAAAGTCTTG	
Ventilago viminalis	-----	---	GTAATG	ATTAACAATA	CATATTATTT	TATTACT-	---	-----	-CGTACTGTA	CTAAAACCTTA	CAAAGTCTTC	
Ventilago leiocarpa	-----	---	GTAATG	ATTAACAATA	CATATTATTT	TATTACT-	---	-----	-CGTACTGTA	CTAAAACCTTA	CAAAGTCTTC	
Reynosia uncinata	-----	---	GTAAG	ATTAACAATA	CATATTATT-	---	ACT---	-----	-TGTACTGTA	CTGAAACGTA	CAAAGTCTTC	
Bathiorhamnus cryptophorous	-----	---	GGAATG	ATTAACAATA	CATATTATT-	---	ACT---	-----	-CGGACTGTA	CTAAATCTTA	CAAAGTCTTC	
Ampeloziziphus amazonicus	-----	---	GGAATG	ATTAACAATA	CATATTATT-	---	ACT---	-----	-CGGACTGTA	CTGAATCTTA	CAAAGTCTTC	
Doerpfeldia cubensis	-----	---	GGAATG	ATTAACAATA	CATATTATT-	---	ACTACA	TAATTATTAC	TCGGGCTGTG	CTGAATCTTA	AAAAGTCTTC	
Hovenia dulcis	-----	---	GGAATG	AT-----	CATATTATC-	---	GCT---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTTC	
Ceanothus	-----	---	GGAATG	AT-----	CATATTATC-	---	GCT---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTTC	
Gouania mauritiana	AATTATTGTA	AATTGGAATA	AT-----	CATATTATC-	---	ACT---	-----	-----	-CGTACTGTA	CCGAAACTCA	CAAAGTCTTC	
Reissekia smilacina	---ATTGTA	AATTGGAATG	AT-----	CATATTATC-	---	ACT---	-----	-----	-CGTACTGTA	CTGAAACTCA	CAAAGTCTTC	
Crumenaria erecta	---ATTGTA	AATTGGAATG	AT-----	CATATTATC-	---	ACT---	-----	-----	-CGTACTGTA	CTGAAACTCA	CAAAGTCTTC	
Helinus integrifolius	-----	---	GAAATG	AT-----	CATATTATC-	---	CCT---	-----	-CGTACTGTA	CTGAAACTCA	CAAAGT-IT-	
Pleuranthodes hillebrandii	---GTTGTA	AATTGGAATG	AT-----	CATATTATC-	---	GCT---	-----	-----	-CGTACTGTA	CTGAAACTCA	CAAAGTCTTC	
Schistocarpha johnsonii	-----	---	GGAATG	AT-----	CATATTATC-	---	GCT---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTTC	
Colubrina asiatica	-----	---	GGAATG	AT-----	CATATTATC-	---	GCT---	-----	-CGTACTGTA	CTGAAACTTG	CAAAGTCTTC	

Emmenosperma alphitonioides	-----	----GGAATG AT-----	CATATTATC-	----GCT---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTTA
Alphitonia excelsa	-----	----GGAATA AT-----	CATATTATT-	----GCT---	-----	-CGTACTG--	---AAACTTA	CAAAGTCTTC
Lasiodiscus mildbraedii	-----	----GGAATG AT-----	CATATTATC-	----GCT---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTTC
Paliurus spinachristi	-----	----GGAATG AT-----	CATATTATC-	----GCT---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTTC
Ziziphus glabra	-----	----GGAATG AT-----	CATATTATC-	----GCT---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTTC
Ziziphus ornata	-----	----GGAATG AT-----	CATATTATC-	----GCT---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTTC
Phylica pubescens	-----	----GGAATG GT-----	CATATTATC-	----GCT---	-----	-CGTACTGTA	TTGAAACTTA	CAAAGTCTTC
Phylica polifolia	-----	----GGAATG AT-----	CATATTACC-	----GCT---	-----	-CGTACTGTA	TTGAAACTTA	CAAAGTCTTC
Phylica arborea nitida	-----	----GGAATG AT-----	CATATTATC-	----TCT---	-----	-CGTACTGTA	TTGAAACTTA	CAAAGTCTTC
Nesiota elliptica	-----	----GGAATG AT-----	CATATTATC-	----GCC---	-----	-CGTACTGTA	TTGAAACTTA	CAAAGTCTTC
Noltea africana	-----	----GGAATG AT-----	CATATTATC-	----GCT---	-----	-CGTACTGTA	TTGAAACTTA	CAAAGT---
Discaria chacaye	-----	----G-AATG AT-----	CAGATTATC-	----GCT---	-----	-CGTACTGTA	TTGAAACTTA	CAAAGTCTTC
Spyridium sp1	-----	----GGAATG AT-----	CATATTATC-	----GCC---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTATT-
Spyridium globulosum	-----	----GGAATG AT-----	CATATTATC-	----GCC---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTT-
Spyridium sp2	-----	----GGAATG AT-----	TATATTATC-	----GCC---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTATT-
Cryptandra sp	-----	----GGAATG AT-----	CATATTATC-	----GCC---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTT-
Trymalium sp1	-----	----GGAATG AT-----	CATATTATC-	----GCC---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTT-
Trymalium sp2	-----	----GGAATG AT-----	CATATTATC-	----GCC---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTT-
Pomaderris rugosa	-----	----GGAATG AT-----	CATATTATC-	----GCC---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTT-
Siegfriedia darwinioides	-----	----GGAATG AT-----	CATATTATC-	----GCC---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTT-
Colletia ulicina	-----	-----	CAGATCATC-	----GCT---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTATTTC
Barbeya oleoides	-----	----TAAATG ATTAACAATA	GATATCATT-	----AGT---	-----	-CGTACTGTA	CTGAAAGTTA	GAAAGTCTT-
Dirachma socotrana	-----	----GGAATG ATTAACAATA	CATATCATT-	----ACC---	-----	-CGTACTGTA	CTGAAAGTTA	CAAAGTCTTC
Dorstenia psilurus	-----	----TGAATA ATTAATAAATT	CTTTTTT--	-----	-----	-CTTACTGTA	CTGAAACTTA	CAAAGTATT-
Ficus	-----	----TGAATA ATTAATAAATT	CATTTTATT-	----ACT---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTT-
Artocarpus heterophyllus	-----	----TGAATA ATTAATAAATT	CATTTTATT-	----ACT---	-----	-CGTACTGTA	CTGAAACTTA	CAAAGTCTT-
Cannabis sativa	-----	----GGAATA ATTAACAA--	-----	-----	-----	-----	-----	-----
Shepherdia argentea	-----	----GGAATG ATTTACAATA	CATATTATT-	----ACT---	-----	-CATACGTTA	ATGAAACTTT	GGAAGTTTCA
Hippophae rhamnoides	-----	----GGAATG ATTTACAATA	CATACCATT-	----ACT---	-----	-CATACGTTA	ATGAAACTTT	GGAAGTTT--
Elaeagnus	-----	-----	-----	-----	-----	-----	-----	-----
Dryas drummondii	-----	----TGAATA ATTAACAATA	CATACCATT-	----ACT---	-----	-TGTACTGTA	CTGAAACTTA	CAAAGTTTTTC
Spiraea	-----	----TGAATA ATTAACAATA	CATACCATT-	----ACT---	-----	-TGTACT---	--GAAACTTA	GAAATAAA--
Pyrus	-----	----TGAATA ATTAACAATA	CATATCATT-	----ACT---	-----	-TGTACTGTA	CTGAAACTTT	GAAATTTA-
Colubrina reclinata	-----	----GGAATG AT-----	CATATTATC-	----GCT---	-----	-CGTACTGTA	CTGAAACTTG	CAAAGTCTTC
Gironniera	-----	----TGAATA ATTAACAATA	ATGTAATCC-	----CTT---	-----	-----	-----	-----
Boehmeria	-----	----TGAATA ATTAATAAATT	CACTTAATT-	----ATT---	-----	-CATATTTTA	CTGAAGTAAT	-----

\*\*\*\*13

\*\*\*\*14

Sageretia thea	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAACC	TGGA----	TA	AGGCTTTGTA	AT----	CCCC	TTTTCGT---
Rhamnus lycioides	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGACC	TGGA----	TA	AGGCTTTGTA	AT----	CCCC	TTTTCGT---
Frangula alnus	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGACC	TAGA----	TA	AGGCTTTGTA	AT----	CTCC	TTTTCGT---
Rhamnella franguloides	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGACC	TGGA----	TA	AGGCTTTGTA	AT----	CCCC	TTTTCGT---
Krugiodendron ferreum	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGACC	TGGA----	TA	AGGCTTTGTA	AT----	CCCC	TTTTCGT---
Rhamnidium cfelaeo	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGACT	CGGA----	TA	AGGTTTGTGTA	AT----	CCCC	TTTTCGT---
Karwinskia humboldtiana	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGACC	CGGA----	TA	AGGTTTGTGTA	AT----	CCCC	TTTTCGT---
Condalia microphylla	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGACC	CGGA----	TA	AGGTTTGTGTA	AT----	CCCC	TTTTCGT---
Scutia buxifolia	-----	-TTTTG---	-AAGATCCAA	---TAAATTC	CACCAAGACC	TGGA----	TA	AGGCTTTGTA	AT----	CCCC	TTTTCGT---
Berchemia discolor	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGACC	TGGA----	TA	AGGCTTTGTA	AT----	CCCC	TTTTCGT---
Maesopsis eminii	TCTTC----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGACC	AGGA----	TA	AGGCTTTGTA	AT----	CCCC	TTTTCGT---
Ventilago viminalis	-----	-TTTTG---	-AAGATCCA	A---GAAATTC	CACCAAGGCC	CGGA----	TA	AGACTTTGTA	ATCGCCCC	TTTTCGT---	
Ventilago leiocarpa	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGGCC	CGGA----	TA	AGACTTTGTA	ATCTCCCC	TTTTCGT---	
Reynosia uncinata	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGACC	TGGA----	TA	AGGCTTTGTA	AT----	CCCC	TTTTCGT---
Bathiorhamnus cryptophorous	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGGCC	CGGCCGATA	AGACTTTATA	AT----	TCCC	TTTTCAT---	
Ampeloziziphus amazonicus	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGGCC	TGACTGGATA	AGACTTTGTA	AT----	TCCC	TTTTTGT---	
Doerpfeldia cubensis	-----	-TTTTG---	-AAGATCCAA	---GAAATTC	CACCAAGGCC	TGACTGGATA	AGACTTTGTA	AT----	TCCC	TTTTTGT---	
Hovenia dulcis	-----	TT TTTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AC----	CCCC	TTTTCGT---
Ceanothus	-----	-TTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCC	TTTTCGT---
Gouania mauritiana	-TTTTTCTT	TTTTTTG---	-ACGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCT	TTTTCGT---
Reissekia smilacina	-----	TT TTTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	CGGA----	TA	AGACTTTGTA	AT----	CCCT	TTATCGT---
Crumenaria erecta	-----	TT TTTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCC	TTTTCGT---
Helinus integrifolius	--TTTTTTT	TTTTTTG---	-CAGATCTAC	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCT	TTTTCGT---
Pleuranthodes hillebrandii	-----	TT TTTTTG---	-AAGATCTAA	---GAAATTC	CACCGAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCTT	TGTTTGT---
Schistocarpea johnsonii	-----	-TTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCC	TTTTCGT---
Colubrina asiatica	-----	-TTTTG---	-AAGATCTAA	---GAAATTC	TAACAAGGTC	TGGC----	TA	AGACTTTGTA	AT----	CCCC	TTTTCGC---
Emmenosperma alphonoioides	-----	-TTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCC	TTTTCGT---
Alphitonia excelsa	-----	-TTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCC	TTTTCGT---
Lasiodiscus mildbraedii	-----	-TTTTG---	-AAGATCTAA	GAAGAAATTC	CGCCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCC	TTTTCGT---
Paliurus spinachristi	-----	TTTT TTTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	CT----	CCCC	TTTTCGT---
Ziziphus glabra	-----	TTTT TTTTTG---	-AAGAGCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	CT----	CCCC	TTTTCGT---
Ziziphus ornata	-----	TTT TTTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	CT----	CCCC	TTTTCGT---
Phylica pubescens	-----	-TTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCC	TTTTCGT---
Phylica polifolia	-----	-TTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCC	TTTTCGT---
Phylica arborea nitida	-----	-TTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCC	TTTTCGT---
Nesiota elliptica	-----	-TTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCC	TTTTCGT---
Noltea africana	-----	--TTTTG---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCC	TTTTCGT---
Discaria chacaye	-----	-TTGTTG---	-AAGATCTAA	---CAAATTC	CACCAAGGCC	TGGA----	TA	AGACTTTGTA	AT----	CCCC	TTTTCGT---
Spyridium spl	-----	T TTTTTT---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	ACACTTTGTA	AT----	CCCC	TTTTCGT---
Spyridium globulosum	-----	T TTTTTT---	-AAGATCTAA	---GAAATTC	CACCAAGGCC	TGGA----	TA	ACACTTTGTA	AT----	CCCC	TTTTCGT---

Spyridium sp2	-----T TTTTT--- -AAGATCTAA ---AAAATTC CACCAAGGCC TGGA----TA ACACTTTGTA AT----CCCC TTTCGT---
Cryptandra sp	-----T TTTTT-AAT TAAGATCTAA ---GAAATTC CACCAAGGCC TGGA----TA ACACTTTGTA AC----CCCC TTTCGT---
Trymalium sp1	-----T TTTTT--- -AAGATCTAA ---GAAATTC CACCAAGGCC TGGA----TA ACACTTTGTA AT----CCC- TTTCGT---
Trymalium sp2	-----TT TTTTT--- -AAGATCTAA ---GAAATTC CACCAAGGCC TGGA----TA ACACTTTGTA AT----CCC- TTTCGT---
Pomaderris rugosa	-----TTT TTTTT--- -AAGATCTAA ---GAAATTC CACCAAGGCC TGGA----TA ACACTTTGTA AT----CCC- TTTCGT---
Siegfriedia darwinioides	-----T TTTTT--- -AAGATCTAA ---GAAATTC CACCAAGGCC TGGA----TA ACACTTTGTA AT----CCC- TTTCGT---
Colletia ulicina	----- -TTGTTG--- -AAGATCTAA ---GAAATTC CACCACGGCC TGGA----TA AGACTTTGTA AT----CCCC TTTCGT---
Barbeya oleoides	-----T TTTTTG--- -AAGATCTAA ---GAAATTC CTCCAGGGCC CGAA----TA AGACTTTGTA AT----CCAC TTTCAT---
Dirachma socotrana	----- -GTTTTG--- -AAGATCTAA ---GAAATTC CATCAGGGAC TGGA----TA ATACTTTGTA AT----- ---TCGT---
Dorstenia psilurus	-----T TTTTTG--- -AAGATCTAA ---GAAATTC CAGCAAGACA TGGA----TA AGAATTTGCG AT----CCTC CTTTCGT---
Ficus	----- -TTTTG--- -AAGATCTAA ---GAAATTC CAACAAGGCC TGGA----TA AGAATTTGAA AT----TCCC CTTTCGT---
Artocarpus heterophyllus	----- -TTTTG--- -AAGATCTAA ---GAAATTC CACCAGGGCT TGGA----TA AGACTTTCCA AC----TCCC CTTTCGT---
Cannabis sativa	----- -TTTTG--- -AAGATCTAA ---GAAATTC CA-TGAGGCT TGGA----TA AACTTTAGA AT----CCCT TTTC-T---
Shepherdia argentea	-----T ATTTTG--- -AAGATCTAA ---GAAATTC CATGAGGC-T GGA----TA AACTTTGGA AT----CCCT TTTC-T---
Hippophae rhamnoides	----- -TTTTG--- -AAGATCTAA ---GAAATTC TA--AGGGTC TGGA----TA AACTTTCTA AT----CCCT TTTC-T---
Elaeagnus	----- -TTTTG--- -AAGATCTAA ---GAAATTC TATCAGGGAC TGTA----TA ATACTTTGTA AT----ACTT TTTCAT---
Dryas drummondii	----- -TTTTG--- -AAGATCTAA ---GAAATTC TATTAGATCC TGTA----TA ATACTTTGTA AT----ACT- TTTCGT---
Spiraea	----- -TTTTG--- -AAGATCTAA ---GAAATTC CACCACGGCC TGGA----TA AGACTTTGTA AT----CCCC TTTCGT---
Pyrus	----- -TTTTG--- -AAGATCTAA ---GAAATTC CACCACGGCC TGGA----TA AGACTTTGTA AT----CCCC TTTCGT---
Colubrina reclinata	----- -TTTTG--- -AAGATCTAA ---GAAATTC CACCACGGCC TGGA----TA AGACTTTGTA AT----CCCC TTTCGT---
Gironniera	----- -TTTTG--- -AAGATCTAA ---GAAATTC CACCACGGCC TGGA----TA AGACTTTGTA AT----CCCC TTTCGT---
Boehmeria	----- -TTTTG--- -AAGATCTAA ---GAAATTC CACCACGGCC TGGA----TA AGACTTTGTA AT----CCCC TTTCGT---

\*\*\*\*\*15

\*\*\*\*\*16

Sageretia thea	-CTTTTCAT TGACATAGAA CCAAGTCC-- ---TCTATTA CAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Rhamnus lycioides	-CTTTTCAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Frangula alnus	-CTTTTCAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Rhamnella franguloides	-CTTTTCAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Krugi dendron ferreum	-CTTTTCAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Rhamnidium cf. laeo	-CTTTTCAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Karwinskia humboldtiana	-CTTTTCAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Condalia microphylla	-CTTTTCAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Scutia buxifolia	-CTTTTCAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Berchemia discolor	-CTTTTCAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Maesopsis eminii	-CTTTTCAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Ventilago viminalis	-CTTTTAAAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Ventilago leiocarpa	-CTTTTAAAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Reynosa uncinata	-CTTTTAAAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Bathiorhamnus cryptophorus	-CTTTTAAAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA
Ampeloziziphus amazonicus	-CTTTTAAAT TGACATAGAA CCAAGTCC-- ---TCTATTA AAAT----- GAGGATGGTG CGTCGTGAAT -----GGTCG GGATAGCTCA

Doerpfeldia cubensis	-ATTTTAAAT	TGACGAAGAC	CCAAGCCA--	---TCTATGA	AAATTACAAT	CAGGATGATG	CGCCGTGAAT	-----GGTCG	GGATAGCTCA
Hovenia dulcis	-CTTTTAAAT	TGACATAGAC	CCAAGCCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Ceanothus	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGGTGTG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Gouania mauritiana	-CTTTTAAAT	TGACATAGAC	CCAAGTCC--	---TATATTA	AAAT-----	GAGGATGGTG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Reissekia smilacina	-CTTTGTATT	TGACATAGAC	CCACGTCA--	---TATATTT	AAAT-----	GAGGCTGATG	CGCCGCAAAT	-----GGTCG	GGATAGCTCA
Crumenaria erecta	-CTTTTAAAT	TGACAAAGAC	CCACGTCA--	---TCTATTA	AAAT-----	GGGGTGTG	CGTCGTCAAT	-----GGTCG	GGATAGCTCA
Helinus integrifolius	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Pleuranthodes hillebrandii	-CTTTTAAAT	TGACATAGAC	CCACGTCA--	---TCTATTA	AAAT-----	GGGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Schistocarpaea johnsonii	-CTTTTAAAT	TGACATAGAC	CCAAGCCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Colubrina asiatica	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CAGCGTGAAT	-----GGTCG	GGATAGCTCA
Emmenosperma alphonoides	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGGTGTG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Alphitonia excelsa	-CTTTTAAAT	TGACAGAG-C	CCAAGTCA--	---TCTGTTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Lasiodiscus mildbraedii	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Paliurus spinachristi	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Ziziphus glabra	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Ziziphus ornata	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Phylica pubescens	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGACT	-----GGTCG	GGATAGCTCA
Phylica polifolia	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Phylica arborea nitida	-CTTTTAAAT	AGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Nesiota elliptica	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Noltea africana	-CTTTTAAAT	AGACATAGAA	CCAAGTCA--	---TCTATTC	AAAT-----	GAGGATGATG	CGTCGAGAA-	-----GGTCG	GGATAGCTCA
Discaria chacaye	-TTTTTAAAT	TGACATAGAC	CCAAGTCC--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Spyridium sp1	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Spyridium globulosum	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Spyridium sp2	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Cryptandra sp	-CTTTTAAAT	TGACAGAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Trymalium sp1	-CTTTTAAAT	TGACATAGAC	CCAAGTAA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Trymalium sp2	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Pomaderris rugosa	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Siegfriedia darwinioides	-CTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGCCGTGAAT	-----GGTCG	GGATAGCTCA
Colletia ulicina	-TTTTTAAAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Barbeya oleoides	-CTTTT-AAT	TGACATAGAC	CCAAGTCA--	---TCTATTA	AAAT-----	GAGGATGATG	CGTCGTGAAT	TGAATGGTCG	GGATAGCTCA
Dirachma socotrana	-CTTTTAAAT	TGACATAGAC	CCCAGTCTTT	CA-TATATTA	AAAT-----	GAAGATGATT	TGTCGTGAAT	-----GGTCG	GG-TAGCTCA
Dorstenia psilurus	-CTTTTAAAT	TGACATAGAC	CCAAGTACTC	TACTCTATTA	AAAA-----	GAGGATGATG	CATAAGGGAT	-----GGTCG	GGATAGCTCA
Ficus	-CTTTTAAAT	TGACATAGAC	CCAAGTCC--	---TATATTA	AAAT-----	GAGGATG-TG	CGTAAGGGAT	-----GGTCG	GGATAG-TCA
Artocarpus heterophyllus	-CTTTTAAAT	TGACATAGGC	CCAAGTCC--	---TCTATTA	AAAT-----	GAGGATGATG	CGTAAGGAAT	-----GGTCG	GGATAGCCCA
Cannabis sativa	-CTTTTAAAT	TGACATAGTC	CCAAGTCC--	---TCTAGTA	AAAT-----	GATGATGATG	TATCATGAAT	-----GGTCG	GG-TAGCTCA
Shepherdia argentea	-CTTTTAAAT	TGACATAGAC	CCCAGCTA--	---TCTATTA	GAAT-----	AAGGATGGTG	CGTTG-GAAT	-----GGTCG	GGATAG-TCA
Hippophae rhamnoides	-CTTTTAAAT	TGACATAGAC	CCAAGCTA--	---TCTATTA	GAAT-----	AAGGATGGTG	CGTTGTGAAT	-----GGTCG	GG-TAGCTCA
Elaeagnus	-CTTTTAAAT	TGACATAGAA	CCCAGCCA--	---TCTATTA	GAAT-----	AAGGATGGTG	CGTTGTGAAT	-----GGTCG	GGATAGCTCA
Dryas drummondii	-TTTTTAAAT	TGACATAGAC	CTAAGTCC--	---TATATTA	AAAT-AAAAAT	GAGGCTGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Spiraea	-TTTTGTAAT	?GACATAGAT	CCAAGTCC--	---CTATATTA	AAAT-AAAAAT	TAGGATGATG	CGTCGTGAAT	-----GGTCG	GGATAGCTCA
Pyrus	-TTTTCTAAT	TGACATAGAC	CCAAGTCC--	---TATATTA	AAAT-AAAAAT	GAGGATGATG	CGTCATGG-T	-----GGTCG	GGATAG-TCA



Colubrina reclinata	-CTTTTAAAT TGACATAGAC CCAAGTCA-- ---TCTATTA AAAT-AAAAT GAGGATGATG CAGCGTGAAT -----GGTCG GGATAGCTCA
Gironniera	-CTTTTAAAT TGACATAG-C CCAAGTCG-- ---TCTATTA AAAT-AAAAT GATGATGATG CATCATGAAT -----GGTCG GGATAGCTCA
Boehmeria	TCTTTTAAAT TGACATAGAC CCAAGTCT-- ---TCTATTA AAAT-AAAAT GAAGATGGGA CTTTCATCAG- -----GGGCC GGATAGCTCA

Sageretia thea	GCTGGTAGAG -----
Rhamnus lycioides	GCTGGTAGAG C-AGAG-ACT GAA-----
Frangula alnus	GCTGGTAGAG C-AGA---CT GA-----
Rhamnella franguloides	GCTGGTAGAG C-AGAGGA-T G-----
Krugiodendron ferreum	GCTGGTAGAG C-AGAG-ACT GAATA----
Rhamnidium cfelaeo	GCTGGTAGAG CGAGAGGACT GAATA----
Karwinskia humboldtiana	GCTGGTAGAG G-AGAGGACT GAATA----
Condalia microphylla	GCTGGTAGAG G-AGAGGACT GAATA----
Scutia buxifolia	GCTGGTAGAG C-AGAGG-CT GTAAGAGGA
Berchemia discolor	GCTGGTAGAG C-AGAGGACT GA-----
Maesopsis eminii	GCTGGTAGAG C-AGAG-ACT GAA-----
Ventilago viminalis	G-TG-TAGAG C-AG---ACT GAA-----
Ventilago leiocarpa	GCTGGTAGAG C-AGAGGACT GAA-----
Reynosia uncinata	GCTGGTAGAG C-AGA-CACT GA-----
Bathiorhamnus cryptophorus	GCTGGTAGAG CGAGAGGACT GAATA----
Ampeloziziphus amazonicus	GCTG-TAGAG C-AGAG-ACT GA-----
Doerpfeldia cubensis	GCTGGTAGAG C-AGAG-AC- -----
Hovenia dulcis	GCTG-TAGAG G-AGAG-ACT GA-----
Ceanothus	GCTG-TAGAG C-AGAG-ACT GAATA----
Gouania mauritiana	GCTGGTAGAG C-AGGG-ACT G-----
Reissekia smilacina	GCTGGTAGAG C-AGCG-ACT GAA-A----
Crumenaria erecta	GCTGGTAAAG C-AGAG-ACT GAA-----
Helinus integrifolius	GCTGGTAGAG C-AGAG-A-T GA-----
Pleuranthodes hillebrandii	GCTGGTAGAG C-AGAG---- -AA-----
Schistocarpha johnsonii	GCTGGTAGAG C-A?AGGACT GAAAATCCT
Colubrina asiatica	GCTGGTAGAG G-AGAG-A-T GAA-----
Emmenosperma alphonoioides	GC-----A- -----
Alphitonia excelsa	GCTGGTAGAG C-AGAGGACT GAAAATCCT
Lasiodiscus mildbraedii	GCTG--AAGG C-AGAGGACT GAATAA----
Paliurus spinachristi	GCTGGTAGAG C-AGAG-ACT GAA-----
Ziziphus glabra	GCTGGTAGAG C-AGAGGACT GAAAT----
Ziziphus ornata	GCTGGTAGAG C-AGAGGACT G-----
Phyllica pubescens	GCTGGTAGAA --AGAG--CT GAACAG---
Phyllica polifolia	GCTGGTAGAG C-AGAGGACT GA-----
Phyllica arborea nitida	GCTG-TAGA- -----
Nesiota elliptica	GCTGGTAGAG CGAGAG-ACT GAATAA----
Noltea africana	-----
Discaria chacaye	GCTGGTAGAG C-AGAGGACT GAAT-----

Spyridium sp1	GCTGGTAGAG C-AGAG-A-T GAATAA----
Spyridium globulosum	GCTGGTAGA- -----
Spyridium sp2	GCTGGTAGAG C-A-----
Cryptandra sp	GCTGGTAGAG C-AGGG-ACT GAA-----
Trymalium sp1	GCTGGTAGAG C-AGGG-A-T GAA-----
Trymalium sp2	GCTGGTAGAG C-AGAG-ACT G-----
Pomaderris rugosa	-----
Siegfriedia darwinioides	GCTGGTAGAG C-AGAGGACT GAATA----
Colletia ulicina	GCTGGTAGAG C-AGAGGACT GAAAATCCT
Barbeya oleoides	GCTGGTAGAG C-AGAGGACT GAA-----
Dirachma socotrana	GCTGGTAGAG CA-----
Dorstenia psilurus	GCTG-AAGAG C-AGAGGACT GAA-----
Ficus	GCTG-AAGA- -----
Artocarpus heterophyllus	GCTTGAAGAG C-AGAGGACT GAA-----
Cannabis sativa	G-TTG-----
Shepherdia argentea	G-T-?TAGAG C-AGAG-ACT GAAT----
Hippophae rhamnoides	GTTGGTAGAG C-AGGG-ACT GAA-----
Elaeagnus	GTG-TAGAG C-AGAGGACT GAA-----
Dryas drummondii	GCTGGTAGAG C-AGAGGACT GA-----
Spiraea	GCC-GTA--- -----
Pyrus	G-T-GTAG-- -----
Colubrina reclinata	GCTGGTAGAG C-AGAGGACT GA-----
Gironniera	GCT-----
Boehmeria	GCC-----















*P. arborea* JER67 010010000000100010000010000000000000100110  
0010000010011100000010000000000011000000100001010000000  
0010000010000010110000010000001000010000000000001100000  
0000001100100000000100001000000001010010000001100000000  
0001000000000000000000000100010000100001010100000101100  
0000000000000000100000000000000000000000000000000000000  
0000000000010000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0010100111001011000100000011000010110010100010000000000  
0000000100000110001000000001000000000000000000010100000000  
00000000010000001001100000000000000000011100000000000000  
0000000000000000000000001000000001000000000000000100000000  
0000000000000000000000001000000000000000000000000000000

*P. arborea* JER71 1101101000001000000000000000100000100100110  
0010011010011100000010100000000111000000100001010000000  
00000000110000101100000100000011000010000000000001100000  
0000101100100000000100001000000000010010000001100000001  
0001000000000000000000000000000000000000000000000000000  
0000000000000000100000000000000000000000000000000000000  
0000010000010000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0000100111001011000000000000000000000000000000000000000  
00000001000001100010000000000000000000000000000010100000000  
0000000001000000100010000000000000000000011001110000000000  
00000000000000000000000010000000010000000000000000000000  
0000000000000000000000000100000000000000000000000000000

*P. arborea* JER72 1101101000001000000000000000100000000100110  
00000010000111000000000000000000111000000100001010000000  
0000000010000010110000010000001101010000000000000000000  
0000001000000000000000000000000000000000000000000000000  
0001000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0000100111001011000100000001000010110010100010000000000  
00000000000001100010000000000000000000000000000001010000000  
000000000100000010001000000000000000000001100000000000000  
0000000000000000000000001000000001000000000000000000000  
0000000000000000000000000100000000000000000000000000000

*P. arborea* JER74 110110100000100010000010000100000100100110  
0010001010011100000010000000000111000000100001010010000  
00100000110000101100000100000011000100000000000001100000  
0000101100100000000100001000000000000010000000101000000  
0001000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0000010000010000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000  
0000100110001011000000000011000010110010100000000000000  
0000000000000110001000000000000000000000000000000000000  
1000000010000001000000000000000000000000000000000000000  
0000100000000000000000000000000000000000000000000000000  
0000000000000000000000000100000100000000000000000000000

*P. arborea* JER75 110000100000100000000000000000000000100100010  
000000001001110000000000000000000111000000100001000000000



0001000000000110000000000100010000000001010000100100100  
000001000100100100000001100001000000000000000000100100000  
000001000001001000000000000001001010000001001000110000000  
00000000000001000000000001000000000000000000000000000000  
00000000000000000000000000000001000000000000000010001000  
0010100111001011000100000001000010110010100010010000000  
00000000000000110001010000000000000100010000010100000000  
00000000010000001000000000000000000000000100110000000000  
00001000100000000001000000001000000000000000000000000001000  
0000000000000000000000000100100000000000000000000000000

*P. arborea* JER94 110110100100101010100010001100000100100111  
00100110100111100000000101001000111000000100001011000001  
0000000010000010110000010000001101011000000000001110010  
00000001000000000001001010000010010000100000000000000001  
0001000000000110110000010000010010010011010000100101100  
000000000000000100000001010010000000000000000000100100000  
00001000000110000000000000000001010000001000000000000010  
10000000000011000000000000000000000000000000000000000000  
00000000000000100000000000000000010000000000000010001000  
0010100111001111000000000001000010110010100010000010000  
0000000000000110001000000000000000000000000000010100000000  
0000001001000000100000000000000000000000001001110000000000  
0000100010000000000100000000100000000000000000001000000100  
0000000000000000000000000100000000000000000000000000000

*P. arborea* JER96 11011010000010101010000000000000000100100110  
0010101000001100000010000000000111000000100001010000000  
0010000011000010010000110000001101011000000000000000000  
0000101100000000000100101000000001000010000001100000000  
0001000000000000000000000000000100000000000011010000000101100  
0000000000000001000000010100000000000000100000000000000  
000001000001000000000000000000000000000000000000100100010000000  
0000000000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000000  
0010100111101011000000000001000010110010100010000000000  
00000000000001100010000000000000000000000000000010100000000  
000000000100000010000000000000000000000000100000000000000  
000000000000000000000000000000000000000000000000000000000  
00000000000000000000000000000000000000000000000000000000

*P. arborea* JER97 110110101000100010000001000100000000100111  
0010011010001100000110101001000011000000100001010000000  
00000000100000101100001100001011000110000000000101100010  
0000101100000000000101001000001000010010000001000000001  
1001000000000110000000011100010010000011010100100101100  
0000000001001001000000000010000000000001000000100100000  
0011010000010000010010000000001010000001001000100000000  
10000000000010000000000000000000000000000000000000000000  
0000000000000010000000100000000000000000000000000000000  
0010100111001011000000000001000010110010100010000000000  
00000000000001100010000000000000000000000000000010100000000  
000000000100000010000000000000000000000000100000000000000  
0000000010000000000100000000100000000000000000000000000  
00000000000000000000000000000000000000000000000000000000

*P. arborea* JER100 1101101000001000000000000000010000010000011  
0001000101001110000001010000000011100000010000101000000  
00010000010000010110000010000000001000110000000000000000  
00000101100000000001000010000000001000010000000010000000  
0000100000000000000000000110000000001000100010000010110  
0000000000000000000000000000000000000000000000000000000









0100010000011001001000010101001001000101010000010110000  
01101000000000000000110100100000000010000000000000000  
0000010001100000000000010000101000000100100001000000001  
0001000000100000000000100000000000000000000000000000000  
00000000000000000000000000000000000000000000000000000000  
1010011000101100000000000100001010001010001000000000000  
0000000000011000110000000000000100000000010100000000000  
0000010100000010001000000000000000000000000000000000000  
0000001000000000010000010010000000000000000000000000000  
0000000000000000000100000000000000000000000000000000000

*P. arborea* JR5 11011010010010101100100100010100010010001100  
1101001000110000011010100100010100100010000101100000000  
1001001100001011000001000000110000100100000100110000000  
0010110000000000010100100000100001000000000100000000000  
00000000000100010010000100000000000000000001010010010010000  
0110100000000000000000001010000000011010100001010000000000  
000100000110000000000000000000000000000000000000000000000  
010000000000100000000000000000000000000000000000000000000  
00000000000000000000000000000000000000000000000000000000  
1010011101111100000000000100000011001010001000000000000  
0000000000011000100000000000000000000000000000000000000  
00000000000000000000000000000000000000000000000000000000  
0000001000000000010000010010000000100000000100000000000  
00000000000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000

*P. arborea* JR6 11011010100010100000000100110000110010011100  
1001101000110000001010100100001100100010000101000010000  
10000010000010110000110000101000000000000000000010111000000  
00101100000000000010100100000100001000001010010000000000  
0100010000010001000000110001000001000101010010010110000  
0110100000000000000000000000000000000000000000000000000  
00010000010000000000000000000000000000000000000000000000  
00000000000000000000000000000000000000000000000000000000  
00000000000000000000000000000000000000000000000000000000  
10100111000011000000000001000010100010100010000000000000  
00000000000110001000000000000000000000000000000000000000  
00000001000000100000000001000000000100000000101000000  
0000001000000000000000001001000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000

*P. arborea* JR7 11011010010010001010001010010000010010011100  
1001101001111000011010100100001100110010000101100100000  
0000001000001011000011000000110000101000000000110000000  
00101000000000000010111100000100001000001001110000000100  
0100000000011001000000110000000001000101010010010110000  
0100100000000000000000000000000000000000000000000000000  
00010000011000000000000000000000000000000000000000000001  
000100000010000000000000000000000000000000000000000000001  
00000000000000000000000000000000000000000000000000000000  
1010011100111100001000100100001010001010001000000001000  
00000000000110001000000000000000000000000000000000000000  
00001000000000100010000000100000000001001100000100000000  
0000001000000000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000

*P. arborea* JR9 01011010010010001000000000010000010010011000  
100110100011000001010100100011100000010000101000000000  
10000010000010110000010000001000000000000000000000000000  
00101100000000000000101101000001000010000000000000000000  
0100000000010001000000110001000001000101010010010110000  
0110100000001000000011000000000000000000000000000000000







000000100010100000000100000000000000000000010100000000000  
0000000000000000000000000000000000000000000000000000000000010  
0000000000000000000000000000000000000000000000000000000000001  
000010000000000000000000000000000000000000000000000000000000

*P. nitida* YM1 1110010000001001000000000000000000000000000011001  
0001110000010000000100000000001110000000010010010000000  
0000001000000001000000000100001000000000000000000000000010001  
00001100000000001000101000100000000000000000000000000000000010000001000  
01000000000000000100000011010000000000000010000000001000010  
00000000000000001011000000000000000000000000000000000000000000000000  
0001000000000000000000000100000010000000000000000000000000000000000000  
0000000000000000100000000010000000000000000000000000000000000000000000  
0000000000000000000000000000000000000000000000000000000000000000000000  
????????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????????

*P. nitida* YM2 1110010000001001000001000000000100000001101100  
0111000001000000010010000000111000000001000101000000000  
000010000000010000000000000001000000000000000000000000000000001000100  
0011000000000001000101000100000000000000000000000000000000000000100001  
0000000000000001000000100100000000000000000000000000000000000000000000  
0000000000000001011000000000000000000000000000000000000000000000000000  
0100001001000000000000000000000000000000000000000000000000000000000000  
0000000000000100010000010001001101000000000000000000000000000000000000  
00000000000000000000000000000000000000000000000000000000000000000000??  
????????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????????

*P. nitida* YM3 11000100000010010000000000000000000000000000101000000  
00000101000100001000001000000010000001000100010000000000  
0100101001000101000100000001000000010010101000000010001  
01000000001001000000000010000010001000000000000000000000000000000000  
01000000000000000000000000000000000000000000000000000000000000000000  
00000000000000000000000000000000000000000000000000000000000000000000  
0001000000010100000100000001000000000000000000000000000000000000000000  
00000000000100000100010100000000000000000000000000000000000000000000  
001000010000000100100000010010000001000000100000100001000010000  
0000010100101100000010000100010001001010000001000000100000  
0011001010001000000000101000000000000000000000000000000000000000000000  
00000000000000000000000000000000000000000000000000000000000000000000  
01000101000000001000100000000101100000000000000000000000000000000000  
001000000010000000010000001110000000101000

*P. nitida* YM4 111001000000100100000000010000100000000011011  
0001110000010000000100101000001110000000010001010000000  
0000001000010001000000000000000000000000000000000000000000000000000000  
00001100000010001000101000100000000000000000000000000000000000000000  
01000000000000000000000000000000000000000000000000000000000000000000  
1000000000000000010110000000000000000000000000000000000000000000000000  
0001000000010000000000000000000000000000000000000000000000000000000000  
00000110000000100000000010001001101000000000000000000000000000000000  
00000000000000000000000000000000000000000000000000000000000000000000  
0000010000001000000000000000000000000000000000000000000000000000000000  
001100101000100000010010100000000100000000011000000000000  
00000011000000000000000000000000000000000000000000000000000000000000

0000000000000000101001000000000000000010000000000000100  
00100000001010000000000001000000000100000  
*P. nitida* YM5 11100100000010010000000000000000000000000000011000  
0001110100010100100100000000001110001001110001000001001  
000000100000000101010010010000000000000000000000000000010001  
0000110000000000001010100110000000101000000110000011001  
011000010000000000000001101010000010000000000000000000000  
0000000000000001011000000000001100101000000000000000000000  
00010000000000000101000000011000000000000000000000000000  
1000000000000001100001011000000000000000000000000000000000  
0000000100100000000000000100100000110000000000000000000000  
000001010010110010000000010000001000001000000010100000  
00000000000010000001000000000000000000000000000110000000000  
000000000100001000000000000000000000000000000000000000010  
000000000000000000100000000000000000000000000000000000001  
00000000000000000001110000000000000000000000000000000000  
*P. nitida* YM6 ??????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
????????????????????????????????????????????????????????????  
1000000100000010000100110010100010000001001010000100000  
00010001100100010000000000000000000000000000000000000001010  
0100000001000000001001010001000000000110101000000001000  
00010000000000000011100100000001010000001000000110000100  
001000