PHYLOGENETIC ANALYSIS OF BREEDING-SYSTEM EVOLUTION
IN HETEROSTYLOUS MONOCOTYLEDONS

BY

SEAN W. GRAHAM

DEPARTMENT OF BOTANY

A THESIS SUBMITTED IN CONFORMITY WITH REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN THE
UNIVERSITY OF TORONTO

© SEAN W. GRAHAM 1997
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
PHYLOGENETIC ANALYSIS OF BREEDING-SYSTEM EVOLUTION
IN HETEROSTYLOUS MONOCOTYLEDONS

Sean W. Graham, Department of Botany, University of Toronto

1997

ABSTRACT

The evolutionary histories of various stylar conditions (heterostyly, enantiostyly, stylar dimorphism and monomorphism) were examined in two groups of monocotyledons, Pontederiaceae and Narcissus, using several sources of phylogenetic information and several different weighting schemes for reconstructions of character evolution.

The monophyly of Pontederiaceae is strongly supported by phylogenetic evidence from rbcL, but its local position in the monocotyledons is unclear. Several tests were used to assess the congruence of four phylogenetic data sets for the family: two chloroplast genes (ndhF and rbcL); a restriction-site survey of the plastome; and a revised morphology-based data set. Despite different “rules” of evolution for the three chloroplast sources of data, these data sets provided highly congruent, well resolved and well supported phylogenetic estimates of the family’s history. The morphological data provided a poorly supported phylogenetic estimate that showed both congruence and incongruence with the chloroplast evidence.

A combined analysis of ndhF and rbcL genes from taxa in five superorders of monocotyledons added support to the view that three families containing enantiostylous taxa, Commelinaceae, Haemodoraceae and Philydraceae, are the nearest living relatives of Pontederiaceae. Several lines of evidence indicate that despite the evolutionary distance between Pontederiaceae and these families, they contain useful historical information concerning the position of the family’s root.

Trees derived from the combined chloroplast evidence were used to reconstruct breeding-system evolution in Pontederiaceae. Moderately weighted optimization schemes
indicate that tristyly arose once only in the family, while predominantly-selfing, florally-monomorphic lineages of *Eichhornia* arose on multiple occasions. Self-incompatibility arose after the origin of the floral trimorphism, a scenario incompatible with some evolutionary models. Enantiostyly arose twice, with its occurrence in *Heteranthera* (*sensu lato*) possibly homologous with that in related families.

Reconstructions of breeding-system evolution in *Narcissus* are impeded by the partial lack of resolution of an *ndhF*-based tree of the genus and the probable role played by undetected hybridization events in the genus' phylogenetic history. A preliminary reconstruction of breeding-system evolution in *Narcissus* suggests that stigma-height dimorphism may have arisen on multiple occasions, and that tristyly in *Narcissus triandrus* may have evolved directly from floral monomorphism.
ACKNOWLEDGEMENTS

It is not very easy to sum up in a few short sentences the debt I owe to so many for making the last seven years a frequently wonderful time, and for making the less-than-wonderful times more than bearable... To Spencer, for taking me on and providing a wonderful problem and professional training, for trusting me enough to do this my own way, and for frequent jolts of psychic energy. Without Brian Morton this whole thing may not have finally got off the ground. Joshua Kohn shared in the prolonged birthing agony of the Evolution paper that I forced myself onto so naively, from which I learned so much about the art of phylogenetics and the science of diplomacy. I have been exceptionally lucky in the lab companions I have had over the years -- Elizabeth Bush, Domenica Manicacci, Majid Ghassemian, Taline Sarkissian, Holly Trewhitt, Mara Kerry, Brendon Larson, Linley Jesson, Peter Toppings, Brian Husband, Chris Eckert, Pam Diggle and Mitch Cruzan. In particular, Fanny Strumas, Bill Cole, Anne Worley, Angela Baker, Andrea Case, Brian Morton, and John Pannell provided friendship and support well above and beyond the call of duty. David McKnight provided crucial encouragement early in my academic training and Richard Abbott more recently.

Inside and outside the department, Tim Gray, Connie Soros, Sue and Emma Dexter, Kristjan Vitols, Rodger Evans, Petra Connelly, Zuzu Gadallah, Roberta Fulthorpe, Mara Kerry, Luba Skambara, Steven Schultz, Ben Teh, John Maxwell, Sandra Laronde, Kate Frego, Rolfe Vinebrooke, Christine Kampny, Fereshteh Hashemi, Usha Goel, Valda Zobens, Nick Provart, Elizabeth Tillier, Bob Latta, Rebecca Dotterer, Esther Levesque, Brenda Kostner and James Scott supplied endless talk, food, fun, and countless hours of procrastination. I thank Tamar Mamourian and Valerie Anderson for their caring professionalism, and Fernanda and the night-time crew for keeping me company into the wee hours.

Various faculty members provided support, help and advice along the way, but Jim Eckenwalder, Jim Anderson, Linda Kohn, Tim Dickinson, Tammy Sage and Nancy Dengler deserve particular thanks for this. Dick Olmstead and Mike Donoghue provided invaluable support and mentorship from afar. Thanks to Ken Sytsma and Linda Kohn for working so hard as my external and internal examiners. Mike Clegg and Jeff Palmer kindly permitted me to use their labs at critical points early on in my degree. Jennifer Richards, Mike Simpson and Charles Horn supplied useful advice on several papers. Many people wittingly or unwittingly supplied me with data and plant material over the years, but Linda Prince, John Blanchard, Mike Simpson, Tom Givnish, Alan Meerow and Mark Chase deserve particular thanks for going out of their way to do so.
Many individuals and institutions provided assistance on a less personal basis, but not less valuable for that. David Swofford, Wayne and David Maddison and Joe Felsenstein wrote the software without which my data would be a meaningless collection of letters and numbers. Gerard Zurawski and the DNAX Institute kindly provided the rbcL primers that allowed me to start. The taxpayers of Canada and Ontario provided the financial support that allowed me to do my work. The Connaught committee provided the financial support that allowed me to come to Canada. CUEW 2 and the University of Toronto provided excellent reasons for hypothermia on mid-February picket lines. I thank Alan Cross and the staff of CFNY for helping to keep me sane in an otherwise lonely molecular lab.

In Angela Baker, Jacquie Bede, Tracy Solomon and Andrea Case I have made wonderful friends. Thanks for all the larks and japes. I thank Christina Heidorn and Sue Dexter for "bean" great friends, for support in times of need, and for endless hours of hilarity. I thank Wendy Untereiner for her friendship, for all the chocolate, and for helping to make it go. I thank Lisa Law for her love and friendship, Rrrmmhh. I thank Philip Awadalla for his friendship, for his intellectual companionship, for sharing his epic love of life and most especially for laughing at my jokes. I thank Catherine Glass and Catriona Gordon for being the most complete and utter pair of bints I am ever likely to know. Thank you for showing me how. I thank Cary List and Peter Au for their friendship and for being the ones.

I thank my mum and dad, Joan and Jimmy Graham, and my sibs, Paula, Martyn and Peter, for their love and unwavering support across the great divide.

This thesis is dedicated to Elsie and Eric McDuff, and to my grandparents, Hannah and Sammy Nelson.

"...what ya thinking?  
What am I singing?  
A song of seeds,  
The food of love,  
Eat the music!"

# Table of Contents

<table>
<thead>
<tr>
<th>CONTENTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>xiv</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xvi</td>
</tr>
</tbody>
</table>

**CHAPTER 1: GENERAL INTRODUCTION**  
Models for the evolutionary origin and breakdown of heterostyly  
Objectives of the thesis  

**CHAPTER 2: PHYLOGENETIC SYSTEMATICS OF PONTEDERIACEAE:**  
IMPLICATIONS FOR BREEDING-SYSTEM EVOLUTION  
INTRODUCTION  
Suprafamilial Systematics  
A. Morphological evidence  
B. Molecular evidence  
Intrrafamilial Systematics  
A. Morphological evidence  
B. Molecular evidence  
Breeding-System Evolution in Pontederiaceae  
A. Origin and evolution of the tristyloous syndrome  
B. Effects of the selfing syndrome on phylogenetic reconstruction  
C. Evolution and adaptive significance of enantiostyly  
CONCLUSION  

vi
TABLE OF CONTENTS

CHAPTER 3: PHYLOGENETIC CONGRUENCE AND DISCORDANCE AMONG ONE MORPHOLOGICAL AND THREE MOLECULAR DATA SETS FROM PONTEDERIACEAE 60

INTRODUCTION 61

MATERIALS AND METHODS 64

Analyses 65

Measures of tree resolution and support 67
Measures of character incongruence 68
Templeton's test for significant differences among phylogenies 69
Evidence for a selfing syndrome in Eichhornia 70
Measures of taxonomic congruence 71

RESULTS 73

Tree resolution and support 73
Optimal and sub-optimal trees 96
Character incongruence 101
Templeton's test for overall incongruence among data sets and their shortest unrooted trees 101
Phylogenetic status of the selfing species of Eichhornia 102
Taxonomic congruence 105

DISCUSSION 112

Systematic implications of the chloroplast data 113
The morphological data is swamped by the molecular data 115
Conflict and agreement among molecules and morphology 115
Sources of incongruence among molecular and morphological trees 117
Summary 119
TABLE OF CONTENTS

CHAPTER 4: THE LOCAL POSITION OF PONTEDERIACEAE IN THE MONOCOTYLEDONS AND ITS ROOT LOCATION 121

INTRODUCTION 122

MATERIALS AND METHODS 124

Analyses 125

*Phylogenetic placement of Pontederiaceae in the monocotyledons* 126

*Are sub-optimal rootings of Pontederiaceae significantly different from the most-parsimonious root(s)?* 127

RESULTS 129

*Systematic relationships among the thirteen families* 130

*Congruence of the two chloroplast genes* 138

*The effect of a posteriori character re-weighting on phylogenetic estimation* 138

*The root of Pontederiaceae* 139

DISCUSSION 145

*Accuracy of the estimated phylogenies* 145

*Systematic relationships of the 13 families* 151

*Conflicts among trees estimated using the single and combined data sets* 151

*Summary of systematic conclusions* 153

*The root of Pontederiaceae* 155
# Table of Contents

**CHAPTER 5: ADAPTIVE RADIATION IN THE AQUATIC PLANT FAMILY PONTEDERIACEAE**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>158</td>
</tr>
<tr>
<td><strong>Taxonomy and Natural History</strong></td>
<td>159</td>
</tr>
<tr>
<td><strong>Taxonomy</strong></td>
<td>162</td>
</tr>
<tr>
<td><strong>Biogeography</strong></td>
<td>162</td>
</tr>
<tr>
<td><strong>Aquatic habitats and ecological differentiation</strong></td>
<td>163</td>
</tr>
<tr>
<td><strong>Aquatic life forms</strong></td>
<td>163</td>
</tr>
<tr>
<td><strong>Life-cycle duration</strong></td>
<td>164</td>
</tr>
<tr>
<td><strong>Clonality</strong></td>
<td>166</td>
</tr>
<tr>
<td><strong>Patterns of leaf development</strong></td>
<td>167</td>
</tr>
<tr>
<td><strong>Floral ecology, pollination and mating systems</strong></td>
<td>169</td>
</tr>
<tr>
<td><strong>Phylogenetic Systematics of Pontederaeace</strong></td>
<td>172</td>
</tr>
<tr>
<td><strong>Morphological and molecular evidence of phylogenetic relationships</strong></td>
<td>175</td>
</tr>
<tr>
<td><strong>Implications of the phylogenetic data and fossil evidence</strong> for the biogeography of the family</td>
<td>180</td>
</tr>
<tr>
<td><strong>Character Diversification and Adaptive Radiation in Vegetative and Reproductive Characters</strong></td>
<td>181</td>
</tr>
<tr>
<td><strong>Outgroups and their effect on character reconstruction in Pontederaeace</strong></td>
<td>181</td>
</tr>
<tr>
<td><strong>Character Codings</strong></td>
<td>183</td>
</tr>
<tr>
<td><strong>Reconstructions of Character Evolution</strong></td>
<td>186</td>
</tr>
<tr>
<td><strong>Aquatic habit</strong></td>
<td>186</td>
</tr>
<tr>
<td><strong>Life form</strong></td>
<td>186</td>
</tr>
<tr>
<td><strong>Duration</strong></td>
<td>190</td>
</tr>
<tr>
<td><strong>Clonality</strong></td>
<td>193</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

## CHAPTER 5: (contd.)

- Leaf developmental pathway 193
- Floral form and self-incompatibility 196

**SUMMARY** 203

## CHAPTER 6: A PHYLOGENY OF *NARCISSUS* L. (AMARYLLIDACEAE) BASED ON THE CHLOROPLAST GENE *NDHF*, AND ITS IMPLICATIONS FOR BREEDING-SYSTEM EVOLUTION IN THE GENUS 205

**INTRODUCTION** 206

- Taxonomy of *Narcissus* 207
- Phylogenetic position of *Narcissus* in *Amaryllidaceae* 208

**Variation in breeding-systems in *Narcissus* and related genera** 209

- Variation in *Narcissus* 209
- Variation in related taxa 210

**Biosystematic information from chromosomal variation in *Narcissus* and related genera** 211

- Variation in chromosome number 211
- Variation in chromosome shape 212

**Objectives of the study** 212

**MATERIALS AND METHODS** 213

**Analysis** 214

- Reconstruction of the evolution of stigma-height polymorphisms 214

**RESULTS** 215

- Phylogenetic evidence from *ndhF* 215
- Sectional relationships 218

- Reconstruction of the evolution of stylar polymorphisms in *Narcissus* 223
# TABLE OF CONTENTS

## CHAPTER 6: (contd.)

**DISCUSSION**

*Taxonomic Implications*  
231

*The monophyly of Narcissus*  
231

*Sub-genus Hermione*  
232

*Sub-genus Narcissus*  
232

*Evolution of chromosome number and form*  
234

*The Evolution of Stigma-Height Polymorphisms in Narcissus*  
236

*Tree resolution and character reconstruction*  
236

*Weighting schemes and phylogenetic reconstruction*  
237

## CHAPTER 7: GENERAL CONCLUSIONS

**Summary**  
249

## APPENDIX A: RECONSTRUCTION OF THE EVOLUTION OF REPRODUCTIVE CHARACTERS IN PONTEDERIACEAE USING PHYLOGENETIC EVIDENCE FROM CHLOROPLAST DNA RESTRICTION-SITE VARIATION

**Abstract**  
253

**INTRODUCTION**  
254

**MATERIALS AND METHODS**  
262

*Taxon sampling*  
262

*Molecular methods*  
263

*Tree reconstruction*  
263

*Character optimization*  
264

**RESULTS**  
265
APPENDIX A: (contd.)

Phylogenetic structure of the family 265
Reconstruction of reproductive character evolution 271

DISCUSSION 281

Systematics of Pontederiaceae 281

Implications of infrafamilial relationships for reproductive character evolution 282
Reconstruction of character evolution 283

Optimization schemes 284
Outgroup identity and coding 286
Uncertainty in tree topology 287
Origin of self-incompatibility 287
The evolution of selfing in Eichhornia 289

CONCLUSION 290

APPENDIX B: SOURCE AND LOCALITY OF SPECIMENS OF PONTEDERIACEAE 291

APPENDIX C: REVISION OF THE MORPHOLOGICAL DATA SET OF ECKENWALDER & BARRETT (1986) 295

APPENDIX D: TAXON PARTITIONS SUPPORTED IN 50% OR MORE BOOTSTRAP REPLICATES FROM THE UNROOTED ANALYSES OF THE COMBINED CHLOROPLAST DATA 302
# TABLE OF CONTENTS

**APPENDIX E: ACCESSION DETAILS FOR SPECIES ANALYSED IN CHAPTER 4**  
303

**APPENDIX F: ACCESSION DETAILS FOR SPECIES ANALYSED IN CHAPTER 6**  
305

**APPENDIX G: CHROMOSOME MORPHOLOGY IN NARCISSUS**  
313

**APPENDIX H: EXPERIMENTAL DETAILS FOR THE SEQUENCING STUDIES**  
316

**LITERATURE CITED:**  
320
LIST OF TABLES

Table 1.1  Major features of recent theoretical models for the evolution of heterostyly  12

Table 2.1  Shortest trees found in the analyses of 88 monocotyledon taxa using constrained and unconstrained heuristic searches  29

Table 2.2  Occurrence of heterostyly in the monocotyledons and general features of the syndrome  52

Table 2.3  Occurrence of enantiostyly in the monocotyledons and general features of the syndrome  54

Table 3.1  Oligonucleotides employed to amplify and sequence a 3'-portion of the chloroplast gene ndhF  66

Table 3.2  Summary statistics for phylogenetic trees of Pontederiaceae  74

Table 3.3  Tree topologies common among pairwise comparisons of the shortest unrooted trees from analyses of the individual and combined data-sets  97
Table 3.4  
Templeton's test applied to assess congruence between one fully bifurcated tree (of four shortest ones) from the analysis of the combined sequence data, and the ten shortest trees from the analysis of the restriction-site data, with respect to the two underlying data sets

Table 3.5  
Templeton's test applied to assess congruence between two fully bifurcated trees (of five shortest ones) from the analysis of the morphological data, and the four shortest trees from the analysis of the combined chloroplast data, with respect to the two underlying data sets

Table A.1  
Reconstructed number of shifts of floral form in Pontederiaceae. The range of frequencies of shifts is summarized for all most-parsimonious reconstructions of floral form on the ten shortest restriction-site based trees
**LIST OF FIGURES**

| Fig. 1.1 | The four major stylar conditions of hermaphroditic plants related to the breeding system that are examined in this thesis | 5 |
| Fig. 1.2 | A selected group of Pontederiaceae illustrating the diversity of aquatic life forms and range of leaf types | 8 |
| Fig. 2.1 | A portion of a strict consensus of the 64 shortest trees found in the unconstrained analysis of rbcL sequences from 88 monocotyledons | 31 |
| Fig. 2.2 | Phylogenetic reconstruction of breeding-system evolution in Pontederiaceae using the combined sequence evidence (ndhF and rbcL) | 40 |
| Fig. 3.1 | Results of parsimony analyses for three individual data sets from Pontederiaceae based on variation in the chloroplast genes ndhF and rbcL, and restriction-site variation in the chloroplast genome | 76 |
| Fig. 3.2 | Results of parsimony analyses of the three possible two-way combinations of the three chloroplast data sets from Pontederiaceae | 80 |
| Fig. 3.3 | Results of parsimony analyses for a combined chloroplast data set composed of the three individual chloroplast data sets from Pontederiaceae | 84 |
| Fig. 3.4 | Results of parsimony analyses for the morphological data and all current data from Pontederiaceae combined | 86 |
| Fig. 3.5 | Summary consensus trees of all shortest unrooted trees from the single and xvi |
combined analyses of the three chloroplast data sets (left-hand tree) and of all
the data sets (right-hand tree) 89

Fig. 3.6. Spectrum of bootstrap support for taxon partitions found in strict consensus
trees or at least 50% of replicates in bootstrap analyses of (A) the fully combined
chloroplast data set and each uncombined chloroplast data set from Pontederiaceae
(ndhF, rbcL and restriction site), (B) the fully combined chloroplast data, the
morphological data and all data combined 92

Fig. 3.7. Excess length to the shortest trees found with various data sets
from Pontederiaceae 98

Fig. 3.8. Neighbor-joining phenograms summarizing dissimilarity in tree-shape
as measured by the partition metric (number of symmetric differences) 106

Fig. 3.9. Distribution of tree-to-tree distances (measured as the number of symmetric
differences) for 22-taxon trees, estimated using random trees 110

Fig. 4.1. The phylogenetic position of Pontederiaceae in a local group of monocotyledons,
based on the chloroplast gene ndhF 131

Fig. 4.2. The phylogenetic position of Pontederiaceae in a local group of monocotyledons.
based on the chloroplast gene rbcL 133

Fig. 4.3. The phylogenetic position of Pontederiaceae in a local group of monocotyledons,
for a combined ndhF and rbcL data set 136
LIST OF FIGURES

Fig. 4.4. Optimal and sub-optimal rootings of a chloroplast-based tree of Pontederiaceae with closely related outgroups 141

Fig. 4.5. Frequencies of different rootings of a chloroplast-based tree of Pontederiaceae with the combined ndhF and rbcL data set for the family, for 100 random outgroup sequences 146

Fig. 4.6. The penalty in parsimony observed when real and random outgroups are attached to sub-optimal locations on a chloroplast-based tree of Pontederiaceae 148

Fig. 5.1. Estimate of phylogenetic history of Pontederiaceae using the combined evidence from the chloroplast genome. Biogeographic information is included 176

Fig. 5.2. Reconstruction of diversification in aquatic life-form in Pontederiaceae using the combined chloroplast evidence 188

Fig. 5.3. Reconstruction of diversification in life-cycle duration in Pontederiaceae using the combined chloroplast evidence 191

Fig. 5.4. Reconstruction of diversification in clonality in Pontederiaceae using the combined chloroplast evidence 194

Fig. 5.5. Reconstruction of diversification in leaf developmental pathway in Pontederiaceae using the combined chloroplast evidence 197

Fig. 5.6. Reconstruction of diversification in floral-form in Pontederiaceae using the combined chloroplast evidence 200

xviii
LIST OF FIGURES

Fig. 6.1. Step matrix describing the evolutionary difficulty of shifts in stilyar condition in Narcissus 216

Fig. 6.2. Results of parsimony analyses for an ndhF data set of Narcissus (majority-rule consensus tree) 219

Fig. 6.3. One of 50 shortest maximum-parsimony trees of Narcissus based on the chloroplast gene ndhF. Information on chromosome numbers is included 221

Fig. 6.4. Reconstructions of the evolution of stilyar class (monomorphism, dimorphism and tristyly) in Narcissus 224

Fig. 6.5. Estimated numbers of evolutionary shifts between stigma-height monomorphism and dimorphism in Narcissus and related genera for several different weighting schemes 228

Fig. A.1. Schematic representation of stamen and style configurations in tristylos, enantiostylos and monomorphic floral forms 256

Fig. A.2. Strict consensus of the most-parsimonious phylogenetic trees found in heuristic searches based on chloroplast DNA restriction site variation in Pontederiaceae 267
LIST OF FIGURES

Fig. A.3. Two of the ten shortest maximum parsimony trees (restriction-site data) from a heuristic search involving only taxa of Pontederiaceae 269

Fig. A.4. Example reconstructions of breeding-system evolution for selected maximum parsimony trees (restriction-site data) 274
CHAPTER 1

~ GENERAL INTRODUCTION ~
CHAPTER 1. GENERAL INTRODUCTION

The genes and phenotypic characters responsible for an organism's breeding system are among the most interesting to evolutionary biologists, because they direct the transmission of the entire genome from generation to generation. The breeding system is defined here to include those morphological and physiological aspects of reproduction in plants that are directly responsible for how male gametes are combined with female gametes during the mating process. These include whether individual organisms are hermaphroditic or unisexual, the spatial and temporal disposition of sexual organs among and within individuals, and the presence or absence of physiological barriers to self or same-morph reproduction. One particular breeding system, heterostyly, has fascinated biologists ever since the study of living organisms was revolutionized by Darwin's theory of evolution (Darwin 1859). The power of this theory lies in its extraordinary ability to provide testable adaptive hypotheses concerning the origin of complex organs like the eye, or of complex character syndromes like heterostyly, in which two or three "sexes" or hermaphroditic morphs co-exist within the same population. Heterostyly is a genetically determined polymorphism in which individuals possess reciprocal (and hence complementary) arrangements of male and female sex organs. The polymorphism functions to promote proficient cross-pollen transfer during mating (e.g., Kohn and Barrett 1992). The complexity of heterostyly greatly fascinated Darwin, and he devoted much time and effort to divining the meaning of this breeding system (see Darwin 1877).

Of the two major modes of investigation in evolutionary biology, manipulative and comparative, the latter attained a rigorous theoretical basis with the development of cladistic methodology by Hennig (1950, 1965, 1966) and subsequent workers. Hennig formalized a phylogenetic view of how organisms and their constituent characters evolve. A single species can be viewed as a coherently evolving set of organismal lineages. Speciation occurs when one set of lineages more or less permanently divides into two (or more) new
sets of lineages, and occasionally when previously divided lineages merge (hybrid speciation). The characters that different species possess may therefore be homologous, that is, similar or identical by descent. Characters evolve over time within and among different species. It is possible to deduce the level of character homology and the sequence of evolutionary change of characters by studying the patterns of descent of organisms and species. A phylogeny of species is a hierarchical record of their evolutionary history and of the history of their characters.

Phylogenies can be used as a powerful investigatory tool by evolutionary biologists (e.g., Donoghue 1989). They can provide useful historical frameworks for reconstructing the evolutionary diversification of individual characters, such as those involved in determining the breeding system. For example, recent models concerning the evolution of heterostyly (Charlesworth and Charlesworth 1979; Charlesworth 1979; Lloyd and Webb 1992a,b) differ in the sequence of assembly of the components of this syndrome, and should therefore be amenable to assessment using phylogenetic reconstruction (see below). However, there are several potential problems in using phylogenies to accurately dissect patterns of character evolution. One major set of complications concerns the reliability of the phylogenetic tree that is used to reconstruct patterns of character diversification. Factors that affect the reliability include whether there are a sufficient number of characters, whether some taxa (e.g., the outgroups) are particularly distant to others in the tree (Felsenstein 1978, Hendy and Penny 1989), and whether particular sub-sets of the characters have modes of evolution that result in distorted phylogenetic estimation (e.g., Huelsenbeck, Bull and Cunningham 1996). Robust phylogenies are essential for accurate reconstructions of character evolution (e.g., Brooks and McLennan 1994).

Given that the estimated phylogenetic history of a group of interest is indeed a good reflection of its actual history, a further concern is that there are many possible schemes for
performing reconstructions of character evolution, and it is not particularly obvious how to choose among them. Such schemes weight the "cost" or difficulty of change between different states in the character of interest, and are essential for mapping characters onto phylogenetic trees. By far the most commonly used weighting scheme is "Fitch optimization" (Fitch 1971; Hartigan 1973) in which all possible changes among discrete states of a character are equally weighted. However, this may not always be a biologically realistic model of the real cost of individual evolutionary shifts. This is a particular concern when reconstructing the evolution of complex characters (e.g., Farris 1977a,b), such as heterostyly. The appropriateness of particular weighting schemes for exploring phylogenetic reconstructions of morphological characters is a poorly explored issue in evolutionary biology, and one that I address in this thesis (Chapters 2, 5, 6 and Appendix A).

Substantial research over the last century on heterostyly has focused almost entirely on experimental (genetic, microevolutionary and developmental) aspects of the polymorphism (Barrett 1992a, 1993), with very little in the way of comparative studies. In this thesis I use a phylogenetic approach to address the evolutionary origin of heterostyly and associated stylar conditions (Fig. 1.1) in two unrelated groups of monocotyledons that possess tristyloous taxa, Pontederiaceae and Narcissus. The primary focus (Chapters 2 to 5) is on Pontederiaceae. Below I introduce these two groups and discuss the major features of recent models for the evolution of heterostyly that will be addressed using phylogenetic analysis of these groups. I then provide an overview of the structure of the thesis and outline specific objectives of individual chapters.
Fig. 1.1. Schematic representation of the four major stylar conditions related to the breeding system of hermaphroditic plants examined in this thesis. (A) In monomorphic species the stigma has the same spatial relationship to the anthers in all flowers within a population. (B) Dimorphism for stigma height occurs where the stigma is positioned beyond or below the anthers ("approach-" and "reverse-herkogamy") in different individuals in the same population. (C) In heterostylyous species a stigma-height polymorphism is coupled with a polymorphism in \textit{anther} height such that different individuals produce flowers with reciprocal positioning of male and female sex organs. The polymorphism is genetically controlled such that each individual produces flowers belonging to only one floral morph (see text). Tristylyous species have three floral morphs (long-, mid- and short-styled individuals). Distylyous species (not shown) possess only two floral morphs (pin and thrum). Different heterostylous species vary as to whether the floral heteromorphism is coupled with a self- and intra-morph incompatibility system. (D) Flowers of enantiostylyous species differ in their handedness, such that a bent stigma and a single bent anther point in opposite directions in outward facing flowers. In most enantiostylyous species (including all enantiostylous Pontederiaceae) flowers with left- and right-handedness are found on the same individual. In such cases this polymorphism is said to be "somatic" rather than "genetic," unlike the polymorphisms described above. Conditions A, C and D are found in Pontederiaceae, and conditions A, B and C are found in \textit{Narcissus}. Distylyous taxa are found in neither group.
A. Monomorphism

B. Stigma-height polymorphism

C. Tristyly

D. Enantiostyly
Pontederiaceae. -- This entirely aquatic family is composed of 6 to 9 genera and about 35 to 40 species, most of which are native to the New World tropics. All members of the family are hermaphroditic. They inhabit a broad array of freshwater habitats and display a remarkable diversity in their reproductive and life-history strategies (Fig. 1.1; 1.2; and see Chapters 2 and 5 for more details). Species are either predominantly outcrossing or predominantly selfing. Species with monomorphic flowers (Fig. 1.1A) are known in at least three of the four major genera, and in Eichhornia and Pontederia monomorphic taxa are predominantly selfing.

Two of the four main genera in the family, Eichhornia and Pontederia, include taxa that are tristylos (Fig. 1.1C) and predominantly outcrossing. Pollination in these species is performed by nectar-collecting bees and butterflies. All but two of the tristylos species possess an additional self- and intra-morph incompatibility system, which is controlled by a pair of genes (the S and M loci) that also determine floral morph (Barrett, Morgan and Husband 1989; Barrett unpubl. data). The genetic control of heterostyly in Pontederiaceae is a consequence of dominance relationships at each locus and epistatic interaction between them. Lewis and Jones (1992) provide a detailed review of the genetic control of heterostyly in this and other groups. Most taxa in the other two main genera, Heteranthera and Monochoria, are enantiostylos and self-compatible. In contrast to heterostyly, enantiostyly in Pontederiaceae is not a genetic polymorphism, since each individual may produce both left- and right-handed flowers (Fig. 1.1D). The functional significance of enantiostyly is unclear, but several features of these taxa indicate that they are serviced by pollen-collecting bees (see Chapter 2).
Fig. 1.2. A selected group of Pontederiaceae illustrating the diversity of aquatic life forms and range of leaf types (see Chapter 5 for a detailed discussion). A. *Pontederia cordata*; B. *Pontederia (Reussia) rotundifolia*; C. *Eichhornia crassipes*; D. *Monochoria vaginalis*; E. *Heteranthera zostericifolia*; F. *Heteranthera (Zosterella) dubia*; G. *Hydrothrix gardneri*; H. *Heteranthera (Eurystemon) mexicana*. All species except *H. mexicana* are represented in the phylogenetic reconstructions of Pontederiaceae (Chapters 2 to 5). Illustrations are from Cook (1990), with permission. All scale bars are 1 cm (except for 1.2C, which is 3 cm).
Narcissus. -- The daffodils are a European and North African genus of perennial, hermaphroditic geophytes consisting of some 30 to 60 species. Major pollinators include nectar- and pollen-collecting bees and hawkmoths. *Narcissus* contains a broad array of floral variation among its different species. There are several monomorphic species (Fig. 1.1A), numerous species with a stigma-height dimorphism (Fig. 1.1B) and a single tristylos species (Fig. 1.1C). Genetic evidence from *Narcissus tazetta* indicates that the stigma-height dimorphism is under the control of a one-locus, two-allele system (Dulberger 1964). Although a few species of *Narcissus* are known to be self-compatible, many species in the genus possess a self-sterility system. Several features of the breeding system are either rare or unknown in other heterostylos groups. Unlike almost all other heterostylos groups, the self-sterility system is not genetically coupled with styr polymorphisms (Bateman 1952, Dulberger 1964, Barrett, Lloyd and Arroyo 1996). Another unusual feature is how late-acting the self-sterility system is. It appears to operate around the time of ovule penetration or self-fertilization (Dulberger 1964; T. Sage, pers. comm.). The existence of stigma-height dimorphism in *Narcissus* is particularly intriguing to evolutionary biologists, because this rare polymorphism is thought by some recent workers to represent an intermediate stage in the evolution of heterostyly (see below).

*Models for the evolutionary origin and breakdown of heterostyly.* -- Charlesworth and Charlesworth (1979) and Lloyd and Webb (1992b) derived distinct models for the evolution of heterostyly. The two sets of models differ in the ancestral form of heterostylos groups and in the emphasis placed on the various selective forces responsible for the origin of the physiological and morphological components of heterostyly. The Charlesworths' models focus on how self-incompatibility (SI) may be selected to reduce self-fertilization in taxa experiencing high levels of inbreeding depression. Reciprocal herkogamy is then
subsequently selected to increase the efficiency of pollen transfer between incompatibility classes. In contrast, Lloyd and Webb’s model focuses on how reciprocal herkogamy may be selected to improve the proficiency of cross-pollen transfer, and hence directly improve male fitness. In their model, inbreeding depression may contribute to the subsequent origin of SI in heterostylos groups. Finally, the models differ critically in the evolutionary time-frame and historical sequence of assembly of these two major components of heterostyly. These models are discussed below from the point of view of these major differences (summarized in Table 1.1).

The ancestral form that is presumed to exist prior to the origin of heterostyly varies in different models for the evolution of heterostyly (Table 1.1). Charlesworth and Charlesworth (1979) explored the conditions favouring the evolution of distyly, and assumed that the ancestral form of distylos taxa is non-herkogamous, meaning that the stigma and anthers are positioned at the same height within each flower. In D. Charlesworth’s (1979) model for the evolution of tristyly, the ancestral form is instead herkogamous, with the stigma and anthers positioned at different levels in the same flower. Lloyd and Webb (1992a) postulated that the ancestral form is herkogamous in both distylos and tristylos groups.

The first step in Lloyd and Webb’s model is the invasion of a new stigma-height morph into a primitively herkogamous population, yielding a stigma-height dimorphism (Fig. 1.1B). This is selected because it can result in more proficient cross-pollen transfer than is possible between same-morph individuals (Table 1.1; Lloyd and Webb 1992b). Two major stylar classes are defined according to whether the stigma is extended beyond, or situated below the anthers, referred to as “approach” and “reverse” herkogamous
<table>
<thead>
<tr>
<th>Depression</th>
<th>Evolution of SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avoidance of depression</td>
<td>Selective force for herkogamy</td>
</tr>
<tr>
<td>Pollen-stylo coadaptation</td>
<td>Evolution of reciprocal reproduction</td>
</tr>
<tr>
<td>Progenial cross</td>
<td>Self-compatible (2 nuclear levels)</td>
</tr>
<tr>
<td>Homomorphic SI</td>
<td>Herkogamous</td>
</tr>
<tr>
<td>First, then intermediate homology</td>
<td>Non-herkogamous</td>
</tr>
<tr>
<td>Reciprocal herkogamy</td>
<td>Reciprocal herkogamy</td>
</tr>
<tr>
<td>Self-compatible (2 nuclear levels)</td>
<td>Self-compatible</td>
</tr>
<tr>
<td>Homokaryons</td>
<td>Homokaryons</td>
</tr>
<tr>
<td>Distyla</td>
<td>Distyla</td>
</tr>
</tbody>
</table>

**Table 1:** Major features of recent herkogamous models for the evolution of herkogamy.
<table>
<thead>
<tr>
<th>Uncoupled</th>
<th>Concurrently Reciprocal herkogamy evolve</th>
<th>Uncoupled</th>
<th>Microevolutionary</th>
<th>Time-frame Eventuality</th>
</tr>
</thead>
<tbody>
<tr>
<td>GC SL are evoluti onary</td>
<td>Reciprocal herkogamy</td>
<td>GC SL are evoluti onary</td>
<td>Reciprocal herkogamy</td>
<td>No special, but probably</td>
</tr>
<tr>
<td>Potentially derived</td>
<td>Helectrolysis stage</td>
<td>Potentially derived</td>
<td>Helectrolysis stage</td>
<td>Helectrolysis stage</td>
</tr>
<tr>
<td>Primitive and also</td>
<td>Incompatible</td>
<td>Primitive and also</td>
<td>Incompatible</td>
<td>Derived from self</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature</th>
<th>Distality</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1997)</td>
<td>Lloyd &amp; Webb</td>
</tr>
<tr>
<td>(1997)</td>
<td>Chadsworth &amp; Chadsworth</td>
</tr>
<tr>
<td>(1997)</td>
<td>Lloyd &amp; Webb</td>
</tr>
<tr>
<td>Chadsworth (1997)</td>
<td>Lloyd &amp; Webb</td>
</tr>
</tbody>
</table>

**TABLE 1.1 (cont’d)** Major features of recent theoretical models for the evolution of herkogamy.
mophs, respectively. Either moph may be the ancestral form, but comparative evidence indicates that approach-herkogamous species are much more common in the angiosperms, and are therefore more likely to be the ancestral form in most heterostyrous groups (Lloyd and Webb 1992a). Tristyly may evolve in some groups instead of distyly, if the ancestral herkogamous form had two distinct anther levels. This may permit the invasion of two floral mophs, rather than just one, through selection for proficient pollen transfer from each anther level in the ancestral form to a corresponding stigma level in either new moph.

Further selection to maximize the proficiency of cross-pollen transfer results in improvement in the reciprocity of anther and stigma positions, and eventually in the emergence of distyrous or tristyrous species from dimorphic ones (Lloyd and Webb 1992a,b; Barrett, Lloyd and Arroyo 1996). This morphological transition is assumed to be relatively rapid (Lloyd and Webb 1992a). However, the existence of a large number of species in Narcissus with a stigma-height dimorphism suggests that that this intermediate morphological stage in the evolution of heterostyly need not be short-lived or transitory, perhaps because there are special developmental circumstances in this genus that impede shifts to heterostyly (Barrett, Lloyd and Arroyo 1996). Alternatively, stigma-height dimorphisms may have arisen recently and on multiple occasions in Narcissus. Two goals of this thesis (Chapter 6) are to use phylogenetic evidence from the chloroplast genome to estimate the number of origins of the stigma-height dimorphism in the genus, and to assess whether the single tristyrous species in the genus (Narcissus triandrus) evolved from an intermediate stigma-height dimorphism.

Under the models of Charlesworth and Charlesworth (1979), and Charlesworth (1979), the major driving force for the evolution of heterostyly is the avoidance of inbreeding depression in selfed progeny (Table 1.1). For distyly, diallelic SI evolves in
order to reduce self-fertilization, and reciprocal herkogamy evolves subsequently to permit
segregated transfer of the two pollen incompatibility types, and hence reduce the frequency
of wasted pollen transfers to stigmas with the same compatibility reaction. An exception to
this pattern is that under some conditions distyly can evolve in the absence of SI. However
SI cannot subsequently evolve in such groups (Charlesworth and Charlesworth 1979, pp.
475), and so this model is not relevant to most heterostylos taxa. In D. Charlesworth’s
(1979) model for the evolution of tristyly and associated trimorphic SI systems, the
avoidance of inbreeding depression is again the major selective force driving the evolution
of the polymorphism. The first step in her model is the origin of a stigma or pollen
incompatibility reaction. When full reciprocal herkogamy and trimorphic SI subsequently
arise, they do so simultaneously, because pollen and stigma incompatibility reactions are
assumed to be a function of the sex-organ level.

Under Lloyd and Webb’s model (1992a,b), however, SI need not arise until long after
the origin of reciprocal herkogamy (Table 1.1). Reciprocal herkogamy evolves first through
selection for improved proficiency of cross-pollen transfer to and from the ancestral
herkogamous form (Lloyd and Webb 1992b). Self-incompatibility may then be selected
subsequently, either to reduce selfing, or as a passive consequence of pollen-style
coadaptation. Such coadaptation may arise due to proficient “legitimate” pollen transfer (see
Lloyd and Webb 1992a,b and Chapter 2). Legitimate pollinations are mating events
between the complementary sex-organ levels of different morphs.

The models from these two groups of workers thus make contrasting statements
concerning the historical sequence of the origin of the physiological or morphological
features of heterostyly (Table 1.1). This makes them potentially testable using phylogenetic
analysis. Self-incompatibility can arise after the evolution of reciprocal herkogamy in Lloyd
and Webb’s model, but in the Charlesworth’s models self-incompatibility evolves before reciprocal herkogamy (distyly), or concurrently with it (tristyly). Among heterostyrous taxa, self-compatibility is primitive, but potentially also derived (representing secondary losses of SI) under Lloyd and Webb’s model, but in the Charlesworths’ models self-compatible heterostyrous taxa can only represent secondary shifts from SI (Table 1.1). Baker (1966) proposed an historical sequence for the evolution of distyly in Plumbaginaceae in which SI arose prior to the origin of reciprocal herkogamy. This sequence of events seems to support the models of Charlesworth and Charlesworth (1979). However Baker did not use external evidence to determine the phylogenetic history of the family. He simply arranged the taxa to reflect what he considered to be the most probable sequence of events for the origin of distyly within Plumbaginaceae (Baker 1966, pp. 362).

A major objective of my thesis is to estimate the phylogenetic history of Pontederiaceae, and to use this to trace the historical pathway for the origin of heterostyly in the family and hence assess the competing hypotheses for the evolution of this polymorphism. Pontederiaceae includes both self-compatible and self-incompatible heterostyrous taxa, but unlike Plumbaginaceae, taxa lacking reciprocal herkogamy but possessing an SI system are unknown. Cladistic analysis of Pontederiaceae can not therefore indicate a sequence of events that refutes Lloyd and Webb’s model (i.e., diallelic or trimorphic SI arising prior to reciprocal herkogamy) unless this were assumed in advance using a sufficiently biased weighting scheme. However, it is possible for a reconstruction of the evolution of heterostyly in the family to be incompatible with the Charlesworth’s models, if it could be shown that self-incompatibility arose long after the evolution of reciprocal herkogamy.
CHAPTER 1. GENERAL INTRODUCTION

Structure of the thesis. -- Chapter 2 reviews previously available phylogenetic evidence for Pontederiaceae and uses new lines of evidence from the chloroplast genome to reconstruct the evolution of breeding systems in the family. A major theme addressed in the next two chapters concerns the general reliability of estimates of the family's phylogenetic history. Chapter 3 presents surveys of DNA sequence variation in the two chloroplast genes (ndhF and rbcL) employed in Chapter 2, and undertakes a detailed investigation of the degree of agreement among these and morphological (Eckenwalder and Barrett 1986) and restriction-site based (Appendix A) sources of phylogenetic information for the family. Chapter 4 attempts to clarify the position of the root of the family, and the local phylogenetic position of the family in the monocotyledons, using evidence from ndhF and rbcL. In Chapter 5, I use phylogenetic estimates from the three available molecular sources of data in Pontederiaceae for a final examination of the evolution of breeding systems and several other life-history traits in the family. In Chapter 6, I present a molecular phylogeny of the genus Narcissus and use it to provide a preliminary assessment of the number of origins of stigma-height dimorphism in the genus and to assess whether tristyly arose from a dimorphic taxon. Below I provide detailed objectives for each chapter.

Objectives of the thesis

Chapter 2. -- Patterns of breeding-system variation in Pontederiaceae and related families are reviewed and placed in a phylogenetic context. It is important to know what the nearest living relatives of Pontederiaceae are in order to polarize characters (i.e., to gauge the direction of change) within the family in reconstructions of breeding-system evolution. The robustness of previously available molecular evidence concerning the family's local position (a survey of variation in the chloroplast gene rbcL; Chase et al. 1993) is assessed using
CHAPTER 1. GENERAL INTRODUCTION

bootstrap analysis and an examination of sub-optimal trees. DNA sequence variation in two chloroplast genes (ndhF and rbcL; presented in more detail in Chapter 3) is used to reconstruct the phylogenetic history of the family. This phylogenetic framework is used to reconstruct evolutionary shifts in breeding-system characters in the family, in order to assess hypotheses (Charlesworth and Charlesworth 1979; Charlesworth 1979; Lloyd and Webb 1992a,b) concerning the evolution of heterostyly.

Chapter 3. -- The purpose of this chapter is to examine the degree of concordance in historical signal among four different available sources of phylogenetic evidence in Pontederiaceae (three chloroplast-based and one morphological) and thus determine whether there is any evidence of statistically detectable incongruence among them. Two of the chloroplast-based data sets are presented here (surveys of DNA sequence variation in two chloroplast genes, ndhF and rbcL), and the third is a survey of restriction-site variation in the chloroplast genome (see Appendix A). The fourth data set is a revision of previously available morphological evidence (Eckenwalder and Barrett 1986; Appendix C). It is useful to obtain as many pieces of phylogenetic evidence as possible when reconstructing phylogenetic history. However different sources of data should only be combined during phylogenetic estimation if it can be shown that they are not "statistically heterogeneous." Different subsets of the chloroplast genome are known to have different modes of evolutionary change (e.g., Wolfe, Li and Sharpe 1987), and it is probable that morphological characters have their own "rules" of evolution. Combinable sources of data do not result in significantly different estimates of phylogeny when analyzed separately, beyond those expected from sampling error (reviewed in Huelsenbeck, Bull and Cunningham 1996). The degree of congruence among the four available data sets is therefore assessed using several currently available tests for phylogenetic congruence and some new tools that are developed here. I also address an hypothesized distortion of a
Chapters 3 and 4 concern the local position of Pontederiaceae in the monocotyledons (i.e., to which taxa the family is most closely related) and the position of the root of the family. The objective of this chapter is to attempt to resolve these two major questions using combined evidence from two chloroplast genes (*ndhF* and *rbcL*) for a range of outgroup families for Pontederiaceae. Improved knowledge of these two facets of phylogeny (the identity of close relatives and the root position) is important for reconstructing character evolution, a theme taken up again in Chapter 5. A subsidiary objective of this chapter is to develop methods for assessing whether different sub-optimal root positions are significantly worse than the optimal root location.

Chapter 5. -- The focus of this chapter is to examine adaptive radiations in vegetative and reproductive characters in Pontederiaceae. The distribution of these characters among families related to Pontederiaceae (see Chapter 4) is assessed, and the ecological evidence concerning the adaptive significance of variation in these characters is reviewed. The diversification of these characters in Pontederiaceae is then examined using combined evidence from the three sources of molecular evidence from the chloroplast genome (see Chapter 3), and the single optimal rooting of the family indicated in Chapter 4.

Chapter 6. -- Finally, this chapter examines whether the chloroplast gene *ndhF* contains useful evidence concerning phylogenetic relationships in the genus *Narcissus* (Amaryllidaceae). This phylogenetic information is compared to detailed surveys of chromosome number and shape performed by Fernandes (for references see Chapter 6), and is used to provide a preliminary reconstruction of the evolution of stigma-height
polymorphisms (including tristyly) in the genus under a range of different weighting schemes.

These five chapters, and an additional published paper (Kohn et al. 1996) considering reconstructions of breeding-system evolution in Pontederiaceae using information derived from a restriction-site survey of the chloroplast genome, are written in research-paper format. A certain amount of repetition therefore exists among individual chapters. The additional paper is included as an appendix (Appendix A) rather than as a major chapter, since Chapter 5 repeats the analysis of breeding-system variation performed in that paper and in Chapter 2, but using the three chloroplast data sets combined.
CHAPTER 2

- PHYLOGENETIC SYSTEMATICS OF PONTEDERIACEAE:
  IMPLICATIONS FOR BREEDING-SYSTEM EVOLUTION -

Chapter 2. Breeding System Evolution in Pontederiaceae

Introduction

The family Pontederiaceae is composed of 6 to 9 genera and some 35 to 40 species of freshwater aquatics, the majority of which are native to the Neotropics. Members of the family are most readily distinguished by a sympodial growth pattern, herbaceous stems with sheathing leaf bases and petiolate leaves, often multi-flowered showy inflorescences subtended by a single bract, six petaloid tepals (blue, mauve, yellow or white) which are variously basally connate and in two series of three, variously dimorphic stamens which are adnate to the perianth, and superior ovaries with a single style. While plants are rarely misclassified as to family, there have been a variety of opinions concerning the local placement of Pontederiaceae within the monocotyledons. The family has been allied with a number of other families in a variety of combinations (reviewed in Dahlgren and Clifford 1982; Dahlgren, Clifford and Yeo 1985; Simpson 1987; Rosatti 1987; Goldberg 1989). Recent treatments suggest a close affinity of the family with Haemodoraceae and Philydraceae (e.g., Hamann 1966; Huber 1969, 1977; Simpson 1990; Thorne 1992a, 1992b).

Adaptive radiation to the multitude of ecological niches associated with aquatic environments has given rise to a diversity of life-histories and reproductive systems among members of Pontederiaceae (Barrett 1988a; Chapter 5; Fig. 1.2). Life-history variation is governed largely by the duration, predictability and depth of flooding. Annual life-histories are characteristic of ephemeral habitats, while perenniality is more commonly found associated with permanent water bodies. Annual species are largely self-pollinating, whereas perennial species are more frequently insect-pollinated and outbreeding. The reproductive ecology of populations is thus closely linked to their life-histories.

Of particular interest to evolutionary biologists is the occurrence of tristyly in the family. This sexual polymorphism has evolved on only a handful of occasions in the angiosperms
(Charlesworth 1979; Barrett 1993), and its origin and adaptive significance are still the subject of debate. Tristylos breeding systems appear to be particularly susceptible to evolutionary modification, giving rise to a range of derivative conditions, particularly involving autogamy (Ganders 1979; Weller 1992). The occurrence of variation in breeding systems and life-histories among members of Pontederiaceae provides opportunities for application of the comparative method for analysing character evolution and the origin of adaptations (Brooks and McLennan 1991; Harvey and Pagel 1991). Such approaches, however, are contingent upon the availability of sound phylogenetic information. A major objective of this review is therefore to evaluate current evidence provided by morphological and molecular data concerning the phylogenetic relationships of taxa within Pontederiaceae and its closest relatives.

This review has two major sections. Using available morphological and molecular evidence, I begin by evaluating contrasting schemes concerning the affinities of the family with other monocotyledonous taxa and provide a brief description of the systematic features of Pontederiaceae and its constituent genera. I then discuss molecular evidence concerning the relationships of taxa within the family and use this evidence to examine the major pathways of breeding-system evolution; in particular the evolutionary build-up and breakdown of tristyly. I also demonstrate how convergent floral evolution associated with multiple shifts from outbreeding to inbreeding can be difficult to detect when only morphological data are available for phylogenetic reconstruction. Finally, I discuss the evolutionary significance of the enantiostylos floral form in this and other families.

Suprafamilial Systematics.

A. Morphological Evidence. — Pontederales (sensu Dahlgren & Clifford 1982) is a monofamilial order of monocotyledons consisting of Pontederiaceae. Dahlgren and Rasmussen
(1983) used a cladistic approach in their morphologically-based study of suprafamilial systematics in the monocotyledons. Except for Zingiberales, they did not attempt intensive cladistic analyses. They presented what they felt were probable phylogenetic arrangements of taxonomic units. Dahlgren and Clifford (1982) and Dahlgren and Rasmussen (1983) discussed a range of morphological characters within the monocotyledons and provided argumentation concerning plesiomorphic versus apomorphic conditions. The degree of support for their phylogenetic groupings varied in terms of the number and quality of their proposed synapomorphies. Their study provides a useful framework for discussing morphological evidence concerning monocotyledon systematics. I discuss their phylogenetic groupings below with special reference to the local placement of Pontederiaceae. I employ the superordinal ending "-anae" throughout the discussion.

Dahlgren and Rasmussen (1983) proposed a major clade within the monocotyledons consisting of the members of Commelinanae, Zingiberanae and Bromelianae (and possibly Arecales) based on three proposed synapomorphies; UV-fluorescent cell walls, copiously starchy endosperm, and the *Strelitizia*-type of epicuticular wax. The *Strelitizia*-type of epicuticular wax has a scattered occurrence throughout the Commelinanae-Zingiberianae-Bromelianae complex and is also present in the Arecanae. Copiously starchy endosperm is present in other monocotyledons but its occurrence is concentrated and probably synapomorphic within the complex. The possession of UV-fluorescent cell walls is a highly consistent feature of this complex and is also present in the Arecanae. Their superorder Bromelianae includes Pontederiaceae, which although it lacks the *Strelitizia*-type of epicuticular wax, has UV-fluorescent cell walls (Harris and Hartley 1980) and a starchy endosperm (Dahlgren and Clifford 1982).

The further sub-clades within this complex that include Pontederiaceae were solely defined on the basis of single synapomorphies (Dahlgren and Rasmussen 1983). A number of exceptions and ambiguities weaken Dahlgren and Rasmussen's argumentation concerning
relationships among the orders. A Zingiberanae-Bromelianae complex was defined on the basis of a single synapomorphy; the possession of a showy petaloid perianth. The utility of this character in delimiting this group is somewhat dubious given the existence of showy petaloid tepals within Commelinanae. The superorder Bromelianae was further defined on the basis of a single synapomorphy; possession of helobial endosperm with a small, starch-free, and sometimes haustorial chalazal chamber. A further sub-clade within their Bromelianae consisting of the orders Pontederiales, Haemodorales, Philydraceae and Typhaceae was also defined by one synapomorphy; the possession of distichous leaves. However, this condition also has a widespread distribution in Commelinanae and Zingiberanae and has variable expression within Pontederiaceae, where the more broad-leaved taxa tend to have spiral phyllotaxy. Possession of an amoeboid tapetum was used to define a group consisting of Pontederiaceae, Haemodoraceae and Typhaceae (Philydraceae has a glandular-secretory type tapetum). The precise status of the tapetum in Pontederiaceae is, however, uncertain (Dahlgren, Clifford and Yeo 1985).

Earlier treatments of Pontederiaceae, Haemodoraceae and Philydraceae emphasised the liliaceous character of these families (Takhtajan 1969; Dahlgren 1975; Dahlgren and Clifford 1982; Cronquist 1988). Dahlgren and Clifford's study listed eleven (versus four) attributes that reflected the stronger liliaceous than commelinaceous character of Pontederiaceae. However, five of the liliaceous characters mentioned (sulcate pollen grains, presence of oxalate raphides, several to many ovules, axile placentation and dehiscent fruit) were later considered by Dahlgren, Clifford and Yeo (1985) to be primitive within the monocotyledons and therefore cannot be used as an indication of phylogenetic affinity.

The possession of oligosulcate pollen (disulcate pollen in Simpson 1987) and girdle-type endothecial thickening were considered by Dahlgren and Rasmussen to constitute synapomorphies of taxa in Pontederiaceae. Both are apomorphic conditions within the monocotyledons (Dahlgren and Rasmussen 1983). Other features of Pontederiaceae which may constitute apomorphies for the family include its aquatic habit (but note that Philydraceae also
inhabits semi-aquatic habitats), possession of petiolate leaves with stipule- or ligule-like structures, possession of hairs on the stamen filaments (Dahlgren and Clifford 1982), and possibly also their bifacial leaf anatomy, which Arber (1920) and Simpson (1990) suggested may be secondarily derived from a unifacial form. None of these features are unique to Pontederiaceae but may still represent synapomorphies of the family if their occurrence in other groups is found to be homoplasious.

Several characters shared among Pontederiaceae, Haemodoraceae and Philydraceae may constitute synapomorphies for a clade consisting of these three families. Of the taxa he investigated palynologically, Simpson (1987) proposed that similarities in pollen exine sculpturing and architecture between Haemodoraceae and Pontederiaceae constitute synapomorphies of these two families. Dimorphic stamens are present in some members of Haemodoraceae and most Pontederiaceae. However, only a single stamen is found in Philydraceae. Possible synapomorphies of the three families include possession of placental sclereids, perianth tannin cells (Simpson 1990) and a form of herkogamy (the spatial separation of stigmas and anthers within a flower) known as enantiostyly, where flowers possess either right- or left-bending styles. Enantiostyly is present in all four genera of Philydraceae, in most genera of the tribe Haemodoreae of Haemodoraceae (Simpson 1990) and in two of the four main genera (Monochoria and Heteranthera) of Pontederiaceae (Eckenwalder and Barrett 1986). These latter characters in particular require further investigation in closely related families to determine if they represent evidence for monophylesis or are instead retained plesiomorphies. For example, it is unclear how widespread enantiostyly is in other families of monocotyledons -- the presence and type of herkogamy are not regularly recorded in taxonomical descriptions (Webb and Lloyd 1986). Enantiostyly is reported in Tecophilaeaceae (Dulberger and Ornduff 1980) and appears to be present in some species of Aneilema (Commelinaceae) (Faden 1991).
CHAPTER 2. BREEDING-SYSTEM EVOLUTION IN PONTEDERIACEAE

B. Molecular Evidence. -- Chase et al. (1993) used the chloroplast gene *rbcL* to investigate phylogenetic relationships within the seed plants. Their study included a broad range of monocotyledons (see their Figs. 5 & 6). Here I further analyse evidence from this gene concerning monocotyledon relationships and attempt to measure the degree of compatibility of this evidence with several recent suprafamilial taxonomical treatments.

Phylogenies were reconstructed using sequence data from 88 monocotyledon taxa -- 85 from the study of Chase et al., and an additional three from Pontederiaceae (see Chapter 3; experimental protocols for DNA extraction, amplification and sequencing are provided in Appendix H). All analyses were performed using PAUP version 3.1.1 (Swofford 1993). A two-tier heuristic search strategy was used during each analysis. NNI (nearest-neighbour interchange) branch swapping was used, with twenty five random-addition replicates employed to help uncover further islands of parsimony (Maddison 1991). The shortest trees found with these searches were then used as the starting point for a second round of searching using TBR (tree bisection-reconnection) branch-swapping. MULPARS and STEEPEST DESCENT options were activated in both tiers of searching. Analyses were performed both with and without topological constraints imposed on the search process. Topological constraints were defined using the treatments of Dahlgren, Clifford and Yeo (1985), Cronquist (1988), and Thorne (1992b), under the assumption that taxonomical units therein represent monophyletic groups. Sequences from the taxa were constrained both by superorder and order (by subclass and order for Cronquist 1988). Thorne (1992b) treated *Acorus* as a taxon of uncertain affinity. However, it was not possible to fully "unconstrain" the phylogenetic position of the *rbcL* sequence of this species for the Thorne analysis. Instead, the constraints employed allow it to freely associate with other taxa only at the superordinal level. Thorne's (1992b) Philydrales (i.e., Pontederiaceae, Haemodoraceae and Philydraceae) and Dahlgren and Rasmussen's (1983) Bromelianae (their Bromeliiflorae) were both used to delimit topological constraints for separate analyses involving these as the sole constraints. In the case of the Bromelianae constraint set,
additional topological structure was imposed based upon the cladistic arrangement of the orders within this superorder (i.e., Typhales, Velloziales, Bromeliales, Pontederiales, Haemodorales and Philydrales) proposed by Dahlgren and Rasmussen (1983). The monocotyledon portion of the cladogram presented from search 2 of the study of Chase et al. (see their Figs. 5B and 6B) was also reconstructed using MacClade version 3.0 (Maddison and Maddison 1992) in order to derive tree statistics for purposes of comparison with the searches performed here. The few taxa not shared between studies were cut from the tree of Chase et al. (1993), except that the three additional sequences from Pontederiaceae were added onto the terminal branch leading to *Pontederia sagittata* in the order found in the unconstrained analysis. A bootstrap analysis was also performed to determine the relative robustness of clades in the unconstrained analysis.

Table 2.1 lists tree statistics resulting from the various analyses. The degree of incongruence between molecular- and taxonomically-based treatments of monocotyledon affinities was taken as the increase in the number of steps and amount of homoplasy (as measured by CI and RI statistics; Table 2.1) of shortest trees found in analyses employing the constraint sets, as compared to those found in the unconstrained analysis. It is not possible to represent any non-explicit taxonomical concepts of "affinity" between or within groups, such as the relative location of groups depicted in "Dahlgrenograms". It should also be noted that this analysis does not determine which subsets of groups within constrained sets are relatively more incongruous with the historical signal present in molecular data, since it only compares the gross schemes.

The unconstrained analysis resulted in the shortest trees found overall (3194 steps). Imposing topological constraints on the tree-searching algorithm resulted in shortest trees of between 3 steps to 138 steps longer than this (Table 2.1). The greatest increase in tree length was observed using constraints based on Cronquist's (1988) scheme (4.32% more steps than the shortest unconstrained tree). Constraints based on the schemes of Dahlgren, Clifford and
<table>
<thead>
<tr>
<th>Phenelzine</th>
<th>120</th>
<th>0.574</th>
<th>0.285</th>
<th>3197</th>
<th>3198</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenelzine's</td>
<td>120</td>
<td>0.574</td>
<td>0.285</td>
<td>3197</td>
<td>3198</td>
</tr>
<tr>
<td>Imipramine (Fishman 1986)</td>
<td>199</td>
<td>0.578</td>
<td>0.278</td>
<td>3282</td>
<td>3288</td>
</tr>
<tr>
<td>Imipramine (Kupferman 1988)</td>
<td>199</td>
<td>0.578</td>
<td>0.278</td>
<td>3282</td>
<td>3288</td>
</tr>
<tr>
<td>Imipramine, Clinical and Xeo (1985)</td>
<td>199</td>
<td>0.578</td>
<td>0.278</td>
<td>3282</td>
<td>3288</td>
</tr>
<tr>
<td>Trimipramine (1986)</td>
<td>199</td>
<td>0.578</td>
<td>0.278</td>
<td>3282</td>
<td>3288</td>
</tr>
</tbody>
</table>

**Note:** Local

<table>
<thead>
<tr>
<th>Phenelzine</th>
<th>24</th>
<th>0.556</th>
<th>0.277</th>
<th>3332</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenelzine's</td>
<td>24</td>
<td>0.556</td>
<td>0.277</td>
<td>3332</td>
</tr>
<tr>
<td>Imipramine (Fishman 1986)</td>
<td>24</td>
<td>0.556</td>
<td>0.277</td>
<td>3332</td>
</tr>
<tr>
<td>Imipramine (Kupferman 1988)</td>
<td>24</td>
<td>0.556</td>
<td>0.277</td>
<td>3332</td>
</tr>
<tr>
<td>Imipramine, Clinical and Xeo (1985)</td>
<td>24</td>
<td>0.556</td>
<td>0.277</td>
<td>3332</td>
</tr>
<tr>
<td>Trimipramine (1986)</td>
<td>24</td>
<td>0.556</td>
<td>0.277</td>
<td>3332</td>
</tr>
</tbody>
</table>

**Note:** Local

<table>
<thead>
<tr>
<th>Phenelzine</th>
<th>192</th>
<th>0.547</th>
<th>0.273</th>
<th>3332</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenelzine's</td>
<td>192</td>
<td>0.547</td>
<td>0.273</td>
<td>3332</td>
</tr>
<tr>
<td>Imipramine (Fishman 1986)</td>
<td>192</td>
<td>0.547</td>
<td>0.273</td>
<td>3332</td>
</tr>
<tr>
<td>Imipramine (Kupferman 1988)</td>
<td>192</td>
<td>0.547</td>
<td>0.273</td>
<td>3332</td>
</tr>
<tr>
<td>Imipramine, Clinical and Xeo (1985)</td>
<td>192</td>
<td>0.547</td>
<td>0.273</td>
<td>3332</td>
</tr>
<tr>
<td>Trimipramine (1986)</td>
<td>192</td>
<td>0.547</td>
<td>0.273</td>
<td>3332</td>
</tr>
</tbody>
</table>

**Note:** Local

<table>
<thead>
<tr>
<th>Phenelzine</th>
<th>64</th>
<th>0.576</th>
<th>0.286</th>
<th>3194</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenelzine's</td>
<td>64</td>
<td>0.576</td>
<td>0.286</td>
<td>3194</td>
</tr>
<tr>
<td>Imipramine (Fishman 1986)</td>
<td>64</td>
<td>0.576</td>
<td>0.286</td>
<td>3194</td>
</tr>
<tr>
<td>Imipramine (Kupferman 1988)</td>
<td>64</td>
<td>0.576</td>
<td>0.286</td>
<td>3194</td>
</tr>
<tr>
<td>Imipramine, Clinical and Xeo (1985)</td>
<td>64</td>
<td>0.576</td>
<td>0.286</td>
<td>3194</td>
</tr>
<tr>
<td>Trimipramine (1986)</td>
<td>64</td>
<td>0.576</td>
<td>0.286</td>
<td>3194</td>
</tr>
</tbody>
</table>

**Note:** Local

<table>
<thead>
<tr>
<th>Phenelzine</th>
<th>2 (1)</th>
<th>0.572</th>
<th>0.285</th>
<th>3205</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenelzine's</td>
<td>2 (1)</td>
<td>0.572</td>
<td>0.285</td>
<td>3205</td>
</tr>
<tr>
<td>Imipramine (Fishman 1986)</td>
<td>2 (1)</td>
<td>0.572</td>
<td>0.285</td>
<td>3205</td>
</tr>
<tr>
<td>Imipramine (Kupferman 1988)</td>
<td>2 (1)</td>
<td>0.572</td>
<td>0.285</td>
<td>3205</td>
</tr>
<tr>
<td>Imipramine, Clinical and Xeo (1985)</td>
<td>2 (1)</td>
<td>0.572</td>
<td>0.285</td>
<td>3205</td>
</tr>
<tr>
<td>Trimipramine (1986)</td>
<td>2 (1)</td>
<td>0.572</td>
<td>0.285</td>
<td>3205</td>
</tr>
</tbody>
</table>

**Note:** Local

**Table 2.1:** Shown in the analyses of the monooxygenase level, consensual and unconsensual heuristic searches.
Yeo (1985), and Thorne (1992b) yielded trees longer than the shortest unconstrained tree by 2.94% and 2.63%, respectively. One of the greatest distinctions between Cronquist's scheme and the others is his Liliidae. This subclass contains a number of families (Haemodoraceae, Philydraceae, Pontederiaceae) treated quite differently by Thorne and Dahlgren and co-workers. Thorne (1992b) places these families in the order Philydrales of his Commelinanae, a superorder that strongly resembles the Commelinanae-Bromelianae-Zingiberanae complex of Dahlgren, Clifford and Yeo (1985) in terms of its constituent families. Constraints employed using only the Bromeliaceae (sensu Dahlgren and Rasmussen 1983) resulted in trees longer than the shortest unconstrained tree by 37 steps (1.16% longer), indicating that there is also some incompatibility between their morphology-based scheme in comparison to the arrangement implied by the rbcL data.

Figure 2.1 is a portion of the strict consensus tree computed from the 64 shortest trees found in the unconstrained analysis. No outgroup was defined in the analysis, but the tree is presented such that Acorus would be placed most basally (not included in the figure). It is notable that this large rbcL clade contains taxa only found in the Commelinanae-Bromelianae-Zingiberanae complex of Dahlgren, Clifford and Yeo, although one member of this complex (Velloziaceae) falls outside this section of the tree. Apart from Commelinales, all orders in this group are monophyletic, at least with respect to the taxa included. Neither Bromeliaceae nor Commelinanae (sensu Dahlgren and Rasmussen 1983) are monophyletic here. Many of the deeper branches in this section of the monocotyledon tree are unsupported or only moderately supported by the bootstrap analysis. In particular, relationships between the orders are not robust. The arrangement shown here is similar to that depicted in Figs. 6A, 6B of Chase et al. (1993).
Fig. 2.1. A portion of a strict consensus of the 64 shortest trees found in the unconstrained analysis of rbcL sequences from 88 monocotyledons. In terms of its constituent families, this clade strongly resembles Commelinanae sensu Thorne (1992a,b) and the Commelinanae-Bromelianae-Zingiberanae complex of Dahlgren, Clifford and Yeo (1985). Branch lengths, as projected onto the consensus tree (ACCTRAN optimisation), are indicated above branches. Bootstrap proportions (the percentage of bootstrap replicates supporting each branch from 110 replicates) are indicated in parentheses below branches. C = Commelinanae, B = Bromelianae, Z = Zingiberanae (after Dahlgren, Clifford and Yeo 1985). The arrow indicates an arrangement of Pontederiaceae-Commelinaceae observed in trees one step longer than the shortest trees (see text). Representatives of Pontederiaceae are: Eichhornia paniculata, Pontederia sagittata, Monochoria korsakovii and Heteranthera oblongifolia.
The monophyly of Pontederiaceae is strongly supported (bootstrap proportion, BP = 100%) based on the representative taxa from the four main genera employed in the unconstrained analysis, but the local position of the family within the complex is problematical. Commelinaceae is depicted as the immediate sister group of Pontederiaceae in all the shortest unconstrained trees found in the unconstrained analysis, but this association was not supported by the bootstrap analysis (BP < 50%). The shortest 192 trees from the NNI tier of the unconstrained search were one step longer than those found after TBR branch swapping. In all of these trees the clade consisting of Pontederiaceae and Commelinaceae was transposed relative to that illustrated in Fig. 2.1 such that it constituted the immediate sister-group of Anigozanthos (Haemodoraceae). Furthermore, the constraint set that prevented Commelinaceae from being the sister group of Pontederiaceae, by enforcing a monophyletic clade uniting Haemodoraceae, Philydraceae and Pontederiaceae (Philydrales sensu Thorne 1992b), resulted in shortest trees only 4 steps (0.13%) longer than the shortest unconstrained trees. Trees not uniting Commelinaceae and Pontederiaceae were thus not substantially longer than the shortest unconstrained trees. An additional constraint analysis was performed to examine the robustness of the membership of Pontederiaceae in the Commelinanae-Bromelianae-Zingiberanae complex. This analysis enforced a clade consisting of the taxa found in the portion of the monocotyledon tree shown in Fig. 2.1, corresponding to this complex, with the exception that species from Pontederiaceae were constrained to lie outside this clade (the "not local" constraint set in Table 2.1). The most parsimonious trees from this analysis were only 3 steps (0.09%) longer than those found in the unconstrained analysis, and showed Pontederiaceae as being sister to Velloziaceae. Since a very low penalty in the number of tree steps is needed to shift Pontederiaceae to a dispersed range of positions within the monocotyledons, it is apparent that the present evidence from the rbcL locus is not suitable for providing a strong indication of the local phylogenetic placement of Pontederiaceae within the monocotyledons.
A. Morphological Evidence. -- Six to nine genera have been recognised in Pontederiaceae, with the majority of species in only four -- Eichhornia, Pontederia, Monochoria and Heteranthera. Eckenwalder and Barrett (1986) treated 32 species and three varieties in their phylogenetic analysis of the family. Of the 42 morphological characters they examined, 35 were potentially informative within the family. Eckenwalder and Barrett's analysis did not fully resolve intergeneric relationships (see their Fig. 2), but they presented a range of character states considered to be synapomorphic for the genera and other clades in the family; the brief discussion below makes special reference to these character states. See also Lowden (1973), Horn (1985), Rosatti (1987), Barrett (1988a), and Cook (1989) for more complete discussions of individual genera.

The native distribution of Eichhornia is centred in the Neotropics (including the West Indies), with the African E. natans being the sole non-New World member of the genus. The genus is composed of 8 to 9 species and can be broadly subdivided into two groups on the basis of life-history, habit and chromosome number. This separation is reflected in a classification of the genus by Schwartz (1927). One group of species is clonal to various degrees, procumbent or free-floating and polyploid (E. azurea, E. crassipes, E. diversifolia, E. heterosperma and E. natans). The other has an erect and non-clonal habit, inhabits ephemeral aquatic habitats demanding a more amphibious existence and is diploid (E. meyeri, E. paniculata and E. paradoxa). Tristylos and non-tristylos taxa are found in both groups.

Eckenwalder and Barrett's (1986) cladistic analyses indicated a monophyletic group of non-tristylos Eichhornia species, but did not clearly resolve whether this clade was closely associated with tristylos species of Eichhornia. However, in the full cladogram presented by these authors (one of the shortest trees linking tristylos and non-tristylos species) a
paraphyletic *Eichhornia* was depicted, with *Pontederia* being derived from within *Eichhornia*. No characters were synapomorphic for *Eichhornia* alone. However, three characters were synapomorphic for a clade consisting of *Eichhornia* and *Pontederia*: a long-lived, perennial life-history, a geniculate infructescence attitude associated with submergent fruit maturation, and a zygomorphic perianth divided into upper and lower lips. Not all species of *Eichhornia* exhibit these character states. Only *E. azurea* and *E. crassipes* are long-lived perennials (the other taxa are short-lived perennials or annuals; these character states were treated as derived), and *E. meyeri*, *E. paniculata*, and *E. paradoxa* have an erect infructescence. Tristyly is probably uniquely derived within the family (see later) and serves as a further synapomorphy uniting *Eichhornia* and *Pontederia* (Eckenwalder and Barrett 1986), since homostylous species are usually interpreted as being derived from tristylos taxa (Barrett 1988a). However, it should be noted that this character was not employed in their analysis. The morphologically more advanced taxa of *Eichhornia*, as measured by advancement indices assigned in Eckenwalder and Barrett (1986), approach *Pontederia* in their overall morphology and perenniality.

*Pontederia* is composed of 6 species and has a primarily Neotropical distribution with extensions into cooler temperate regions (Canada and Argentina) at the boundaries of the range of the genus. All taxa are long-lived perennials with an erect to procumbent habit. Clonal propagation is achieved through trailing stems and rhizomes, with these organs additionally facilitating perennation under harsh conditions. The single-seeded fruits in *Pontederia* are utricles enclosed by a hardened residual perianth base. The utricle is light and the surrounding perianth-remainder is aeriferous. The consequent buoyancy of the fruit facilitates long-range dispersal (Lowden 1973). Lowden also noted that animal-mediated fruit dispersal may occur, especially in subgenus *Reussia*, where the fruit has prominently curved spines.

As measured by the five synapomorphies presented in Eckenwalder and Barrett's (1986) study, *Pontederia* is the best-supported genus in the family. Synapomorphies for *Pontederia* include the possession of a single fertile locule (the other two locules are non-fertile), a single
ovule per fruit (with terminal pendulous placentation), seed lengths exceeding 1.5 mm, an indehiscent fruit, and a non-smooth fruit wall. However, it is possible that the first three character states are evolutionarily correlated with each other. They could be interpreted as being part of a packaging strategy for single-seeded fruits, with large seed size a consequence of an energetic trade-off between seed size and number.

The perianth in *Pontederia* is strongly zygomorphic and consists of two lips. Various interpretations of perianth structure have been made in *Reussia* (Endlicher 1836; Solms-Laubach 1883; Lowden 1973), and this character has been used as a basis for generic segregation of *Reussia* from *Pontederia* (Endlicher 1836). Lowden (1973) concluded that earlier interpretations are somewhat confused and that this character represents a weak basis for generic segregation. He divided *Pontederia* into two subgenera. Subgenus *Pontederia* is supported by four synapomorphies; the possession of a pulvinus, a derived paniculate inflorescence type, more than 100 flowers per inflorescence, and a smooth-ribbed or tooth-ribbed fruit wall. Subgenus *Reussia* has two supporting synapomorphies; its procumbent habit and a spiny fruit wall.

*Monochoria* is composed of 7 to 8 species (Cook 1989) and is the only genus of Pontederiaceae restricted to the Old World. It has representatives in Australia and Africa and a centre of diversity in tropical Asia. Life-histories range from short-lived perennial to annual, with plants most commonly emergent in habit (i.e., with the plant body rooted in soil but growing above the water surface to varying degrees). The fruit is a capsule enclosed in a withered perianth and it is dispersed as a unit, later releasing the numerous seeds for water-mediated dispersal. Eckenwalder and Barrett (1986) found that two character states lend support for a monophyletic *Monochoria*; a geniculate infructescence and a poorly fused perianth (10-20% fusion) resulting in a nearly bowl-shaped flower. In contrast with the rest of the family, anther dehiscence in *Monochoria* is poricidal (Dahlgren and Clifford 1982; Cook 1989). This may represent a further synapomorphy of this genus.
Heteranthera and its allied genera comprise some 15 species and inhabit ephemeral aquatic habitats throughout the New World tropics, with two species in Africa (H. callifolia and Scholleropsis lutea). Most taxa are annuals or short-lived perennials, although Zosterella is a long-lived clonal perennial. Heteranthera in the broadest sense encompasses Zosterella, Eurystemon mexicanum, S. lutea and Hydrothrix gardneri (see the analysis of Eckenwalder and Barrett 1986). Synapomorphies grouping these taxa are: regular possession of cleistogamous flowers, 40-60% fusion of the perianth (10-25% in Hydrothrix), and possession of three or one stamen(s). Hydrothrix and Heteranthera species also have parietal placentation, in contrast to the rest of the family where placentation is axile (Dahlgren and Clifford 1982). This placentation type is rare in Bromeliaceae (Dahlgren, Clifford and Yeo 1985).

While genera of Pontederiaceae can be effectively divided into a clade of tristyloous and homostyloous taxa composed of Eichhornia and Pontederia, versus a largely enantiostyloous clade composed of Monochoria and Heteranthera (Fig. 1 in Eckenwalder and Barrett 1986), the morphological evidence supporting such a phylogenetic division is not strong (Table 5 and Fig. 2 in Eckenwalder and Barrett 1986). Only two synapomorphies (basifixed anthers and dimorphic stamens) support the clade composed of Monochoria and Heteranthera, although to these might also be added the enantiostyloous floral morphology. Four synapomorphies support a clade consisting of Eichhornia and Pontederia (see earlier). Uncertainty concerning phylogenetic relationships among genera, and the restricted numbers of morphological synapomorphies in Pontederiaceae, motivated me to obtain new sources of phylogenetic information to further clarify intergeneric relationships within the family.

B. Molecular Evidence. -- I reconstructed the phylogeny of Pontederiaceae with a combined data set based upon partial sequences from the chloroplast genes rbcL and ndhF. Twenty-five taxa (23 species, including three varieties of P. cordata) of Pontederiaceae were examined
Within the family 120 characters were potentially informative. Heuristic searches were performed as described earlier, except that 1000 random addition replicates were performed with TBR branch-swapping. *Philydrum lanuginosum* (Philydraceae) was used as an outgroup. A single tree with a length of 464 steps was found (CI = 0.552, excluding uninformative characters; RI = 0.775). The tree is shown in Fig. 2.2.

This chloroplast-based tree indicates that three of the four main taxonomic groups in the family are monophyletic; Monochoria, *Heteranthera* s.l. (including *Zosterella dubia* and *Hydrothrix gardneri*) and Pontederia s.l. (including one representative of subgenus *Reussia*, *P. rotundifolia*). *Heteranthera* is the sister group to the rest of the family. Two phylogenetically-distinct clades of *Eichhornia* each consist of a tristylos species (*E. paniculata* or *E. azurea*) together with two selfing species of *Eichhornia*. The clade consisting of *E. azurea*, *E. diversifolia* and *E. heterosperma* is sister to *Pontederia*. *Eichhornia crassipes* and *E. meyeri* are in neither of these groups of *Eichhornia*. The former is situated basally in a clade consisting of *Monochoria*, *Pontederia* and the *Eichhornia* group that includes *E. azurea*. The latter is basal to the clade consisting of *Pontederia, Monochoria* and all other species of *Eichhornia*.

Both morphological and molecular data sets thus support the monophyly of three of the four main genera of Pontederiaceae (*Pontederia, Monochoria* and *Heteranthera*). Significantly, however, the monophyly of *Eichhornia* is supported by neither the molecular nor morphological data sets. The unnaturalness of *Eichhornia* has taxonomic implications and raises the issue of whether the genus should be maintained as currently circumscribed. The non-monophyly of *Eichhornia* and the dispersed positions of the homostylos species of *Eichhornia* on the tree complicates phylogenetic reconstruction of the gain and loss of tristyly. However, as discussed below, it seems likely that this breeding system arose only once within the family.
A diversity of floral syndromes associated with the pollination biology and breeding systems of individual taxa are found within Pontederiaceae. Here I briefly review some of the major issues concerned with the evolution of breeding systems in Pontederiaceae, and focus particularly on the role of phylogenetic data and models in informing our understanding of the origin and evolutionary relationships of the three primary floral conditions (Fig. 1.1) that occur within the family -- tristyly, floral monomorphism and enantiostyly.

A. Origin and Evolution of the Tristylos Syndrome. -- Tristyly is a floral syndrome of animal-pollinated plants that functions to increase the proficiency of cross-pollen transfer (Darwin 1877; Lloyd and Webb 1992a, 1992b; Kohn and Barrett 1992). Three main components usually constitute the tristylos syndrome: reciprocal positioning of stigma and anther heights among the three floral morphs (reciprocal herkogamy), a self- and intramorph incompatibility system in which only pollen from the same level as the stigma is compatible, and a range of ancillary polymorphisms, primarily of pollen and stigmas. Each of the three floral morphs in tristylos species possesses a stereotypical combination of floral form and incompatibility type. Floral morph and self-incompatibility status is controlled by a simple genetic system involving two diallelic genes (the S and M loci) and is a consequence of dominance relationships at each locus and epistatic interaction between them (Barrett, Morgan and Husband 1989; Barrett unpubl. data).

In optimising the three floral conditions (tristyly, enantiostyly and homostyly) onto the chloroplast-based tree, I used a weighting scheme (see below) that favours the loss of tristyly or enantiostyly over their gain (Fig. 2.2A, 2.2B). An optimisation (not shown) that treats all shifts
Fig. 2.2. Phylogenetic reconstruction of breeding-system evolution in Pontederiaceae. The tree is the single shortest one found in an analysis based on sequence data from the chloroplast genes rbcL and ndhF (see text). Reconstruction of character evolution was performed using MacClade version 3 (Maddison & Maddison 1992). Philydmm lanuginosum was used to root the tree. A trichotomy involving E. azurea, E. heterosperma and E. diversifolia was arbitrarily resolved to permit character optimisation using user-defined character types in MacClade. Figs. 2.2A and B depict the evolution of tristyly, enantiostyly and homostyly in Pontederiaceae. The "Floral" user-defined character type employed in these reconstructions gives a slightly smaller weight to the loss of tristyly or enantiostyly (a shift to homostyly or floral monomorphism) than to the gain or interconversion between these two flower types. Weights employed: loss of tristyly or enantiostyly = 2 steps; gain of enantiostyly or tristyly or shift between them = 3 steps. Alternative resolutions of the trichotomy involving E. azurea, E. heterosperma and E. diversifolia lead to an optimisation with two, rather than one, loss(es) of tristyly in this clade (not shown). P. lanuginosum is enantiostylos, but other potential sister-groups to Pontederiaceae have some enantiostylos taxa (see text), or are florally monomorphic. Two different codings of the outgroup's floral state were therefore examined; the outgroup was coded as either monomorphic (2.2A) or enantiostylos (2.2B). Monochoria cyanea was coded as uncertain for floral form (i.e., enantiostylos or monomorphic).
Fig. 2.2(C). The evolution of heteromorphic self-incompatibility (SI) in Pontederiaceae. The outgroup was coded as self-compatible (SC) (see text). Transitions between SI and SC were equally weighted. Depending on the resolution of the trichotomy involving \textit{E. azurea}, \textit{E. heterosperma} and \textit{E. diversifolia}, SI arises either once or twice in the clade containing \textit{Pontederia} and \textit{E. azurea} and associated homostyles. With the resolution of \textit{E. azurea} and associated homostyles shown here, the origin of SI is equivocal (i.e., one or two origins of SI are possible). Regardless of how this equivocality is viewed, SI arises after the origin of tristyly in the family.

in floral form as equally weighted indicates an independent origin of enantiostyly in *Monochoria* (as is also the case with the optimisations presented in Fig. 2.2A and 2.2B), but is equivocal with regard to the origin of tristyly. This alternative optimisation requires between one and four independent origins of tristyly within the family; the number of origins of tristyly depends partly on how the trichotomy involving *E. azurea*, *E. heterosperma* and *E. diversifolia* is resolved. However, a range of microevolutionary and genetic evidence (described in the next section) indicates that the breakdown of tristyly to homostyly is a relatively simple process that occurs frequently. In addition, a number of lines of evidence suggest that the evolution of tristyly is likely to be a very infrequent event. The overall rarity of this breeding system argues against it arising twice or more within the same genus. Tristyly is known to have evolved within only four or five angiosperm families (Pontederiaceae, Lythraceae, Oxalidaceae, Amaryllidaceae and possibly Conneraceae; see Barrett 1993), only two of which are monocotyledons and all of which are phylogenetically distant from one another. Moreover, if the relative frequency of the two basic forms of heterostyly is any indication, tristyly appears to have much more difficulty evolving than distyly. The latter is believed to have evolved on at least 23 separate occasions in the flowering plants (Lloyd and Webb 1992a). Finally, striking differences in the developmental basis of the polymorphism exist among the tristyloous families (Richards and Barrett 1992). In contrast, there is a high degree of morphological consistency in the polymorphism within *Eichhornia* and *Pontederia* species, lending further support to the hypothesis that tristyly had a single origin within the family.

The weighted optimisation of floral conditions onto the tree indicates a single origin of tristyly in the family (Fig. 2.2A, 2.2B). There is a certain danger of circularity in using a weighting scheme that favours a single origin of tristyly and then using the resulting optimisation as further evidence of this fact. However, Maddison and Maddison (1992; Chapter 4) point out that workers should feel compelled to use the available biological evidence concerning a process when reconstructing the history of that process using local phylogenetic
Finally, I should point out that the weighting scheme employed in this analysis (a gain: loss weighting ratio of 3:2; see Fig. 2.2) only marginally favours the loss of tristyly or enantiostyly over their gain during the optimisation process.

In an evolutionary scheme proposed by Lloyd and Webb (1992a, 1992b), heterostyly evolves from a uniformly herkogamous taxon, i.e., an ancestor in which all individuals possess the same type of stigma-anther separation. Herkogamy serves to reduce self-interference during mating (Webb and Lloyd 1986). The reciprocal herkogamy that characterises heterostylosous plants represents a functional improvement upon the monomorphic condition because it acts to increase the efficiency of pollen transfer among individuals, by more precisely matching pollen dispatch-receipt points on the pollinator's body. A body of empirical evidence in Pontederiaceae lends support to this interpretation of the functional significance of heterostyly (Price and Barrett 1982; Barrett and Glover 1985; Wolfe and Barrett 1989; Lloyd and Webb 1992b).

Under the evolutionary model of Lloyd and Webb (1992a, 1992b), heteromorphic incompatibility, the class of self-incompatibility associated with the heterostylosous syndrome, arises after the floral heteromorphism. It arises either as a passive consequence of (co)adaptation of each class of pollen to the stylar morph to which it is most proficiently transferred, an hypothesis first suggested by Darwin (1877), or as an actively selected anti-selfing device. This hypothesis stands in opposition to the other major theoretical model for the evolution of heterostyly (Charlesworth and Charlesworth 1979) which presupposes that self-incompatibility arises as an anti-selfing device prior to the origin of reciprocal herkogamy, in the case of distyly, or concurrently with it, in the case of tristyly (see Chapter 2; Table 1.1).

The Lloyd and Webb model potentially permits different origins and evolutionary histories of self-incompatibility (SI) in each morph and among the different lineages of heterostylosous species. Differences in the site and strength of action of SI are well documented
in tristylos species of Pontederiaceae. For example, the different, illegitimate pollen classes fail at different, but characteristic points in the stylar tract and ovary of *P. cordata* (Anderson and Barrett 1986) and *P. sagittata* (Scribailo and Barrett 1991b). Self-incompatibility is stronger overall in tristylos *Pontederia* species than in *Eichhornia* species, where it is found only in *E. azurea*. In all tristylos *Pontederia* species examined, SI is much stronger in the long- and short-styled morphs than in the mid-styled morph, where illegitimate pollination results in abundant seed set (Barrett and Anderson 1985). In tristylos *Eichhornia* species only *E. azurea* has appreciable SI, although only data from the long-styled morph is available for this species. Even here, the strength of incompatibility differs among illegitimate anther levels (Barrett 1978). *Eichhornia paniculata* is completely self-compatible, in the traditional sense of this expression (i.e., full seed set upon application of self-pollen). However, differences in the prepotency of the pollen types of this species have been observed when different classes of pollen are forced to compete for access to ovules (Cruzan and Barrett 1993). Typically, legitimate pollen performs better than illegitimate pollen and the species can therefore be viewed as possessing a cryptic trimorphic incompatibility system.

Issues of homology make phylogenetic interpretation of the evolution of incompatibility systems difficult. Should cryptic SI in *E. paniculata* be viewed as homologous with full heteromorphic SI? Under the Lloyd and Webb evolutionary scheme (1992a, 1992b) both physiological systems can arise from the same evolutionary force, i.e., pollen-style adaptation. However, as traditionally defined, *E. paniculata* "functions" as a fully self-compatible species. If cryptic SI is homologous with full SI, then the optimisation of SI depicted in Fig. 2.2C would by this interpretation be misleading; its point of origin would indicate when full heteromorphic SI evolved, presumably from a version weak enough to masquerade as self-compatibility. It is also not clear if the trait I call heteromorphic SI is a single unified character, since under Lloyd and Webb's model, morph-specific pollen-style interactions leading to pollen failure can have different evolutionary and phylogenetic trajectories among different
heterostylos lineages.

An additional problem concerns the difficulty of accurately determining root position in phylogenetic reconstructions. The underlying structure of an ingroup phylogeny can be sturdy, but the precise location of the root of the tree can still remain unclear. Different root placements can lead to different optimisations of the origins of tristyly and SI onto the tree (Appendix A).

Figure 2.2C illustrates the optimisation of self-incompatibility and self-compatibility onto the chloroplast-based tree. Transitions between self-incompatibility and self-compatibility were equally weighted. The outgroup is coded as self-compatible in Fig. 2.2C, but coding it as self-incompatible does not produce a different reconstruction of the evolution of self-incompatibility within Pontederiaceae. Excluding Pontederiaceae, sporophytic self-incompatibility systems are unknown in the monocotyledons (Charlesworth 1985; Weller, Donoghue and Charlesworth 1995). Some taxa in Commelinaceae possess a gametophytic self-incompatibility system (reviewed in Owens 1981), but it seems highly unlikely that this system is homologous with the heteromorphic sporophytic system found in Pontederiaceae.

The optimisation of SI is equivocal and is also dependent on how the trichotomy involving *E. azurea* and its associated homostylos species is resolved. Depending on how these ambiguities in optimisation are disentangled, SI either originates at the base of the clade containing *Pontederia, E. azurea, E. heterosperma* and *E. diversifolia* (with one or two losses along the branch(es) leading to the two homostylos species), or it arises twice within this clade; once along the branch leading to *Pontederia*, and once along the branch leading to *E. azurea*. In either case, SI arises after the origin of tristyly. These conclusions concerning the evolutionary history of SI in Pontederiaceae also serve as a caution against assumptions that the existence of self-compatibility in heterostylos taxa (or morphs) always represent a degenerate condition (cf., Ornduff 1972; Weller 1992). In the optimisations presented here, self-compatibility can be the more primitive condition, a pattern consistent with Lloyd and Webb's
model (1992a,b) for the evolution of heterostyly.

**B. Effects of the Selfing Syndrome on Phylogenetic Reconstruction.** — In virtually every heterostylyous group, multiple shifts to predominant self-fertilisation have occurred via the evolution of homostyly. Homostyles possess anthers and stigmas at the same position within a flower, and as a result are largely self-fertilising. It has generally been assumed, following Darwin (1877), that homostyles are evolutionarily derived from heterostylyous ancestors. In many cases the evolution of homostyly in heterostylyous groups is closely associated with the development of reproductive isolation and speciation (Baker 1961).

Phylogenetic reconstruction is problematical in groups where such repeated transitions to predominant self-fertilisation (autogamy) have occurred, since this evolutionary shift is typically accompanied by multiple parallel changes in a broad range of floral characters (a "selfing syndrome"), as well as changes in life-history (Lloyd 1965; Ornduff 1969; Eckenwalder and Barrett 1986; Wyatt 1988; Morgan and Barrett 1989). The evolution of multiple, correlated morphological changes associated with shifts to autogamy violates the critical assumption of character independence that is implicit in phylogenetic reconstruction. Although floral characters represent some of the most important and numerous morphological data employed in phylogenetic reconstruction, and the shift to predominant self-fertilization from predominantly outcrossing breeding systems constitutes one of the most pervasive themes in floral evolution (Stebbins 1970; Jain 1976), the effect of this breeding-system shift on phylogenetic reconstruction is not well documented [although see the studies on *Leavenworthia* (Lloyd 1965), *Limnanthes* (Arroyo 1973; McNeill and Jain 1983), *Arenaria* (Wyatt 1988), and *Amsinckia* (Schoen 1993)].

Four of the seven taxa of *Eichhornia* and one of the eight taxa of *Pontederia* included in the morphology-based phylogenetic analysis of Pontederiaceae by Eckenwalder and Barrett
(1986) are homostylos. Their analysis indicated only two origins of homostyly within the family; one in *Pontederia* (*P. parviflora*) and the other in *Eichhornia*. They suggested that the finding of a single origin for homostyly in *Eichhornia* was a consequence of the distorting effects of the selfing syndrome on phylogenetic reconstruction.

Several lines of evidence indicate that selfing variants evolve readily in tristylos populations of *Eichhornia* (Barrett 1988a; Barrett, Morgan and Husband 1989). For example, relationships inferred among Brazilian populations of *E. paniculata* using genetic distance estimates from isozyme data (Husband and Barrett 1993) imply that populations possessing selfing variants have arisen repeatedly from outcrossing populations in different parts of the geographical range of the species. Although it is difficult to assess mutational versus migratory hypotheses for the origins of selfing in such populations from isozyme data alone, data on the genetic architecture of floral traits causing selfing are consistent with the multiple origin hypothesis (Fenster and Barrett 1994). Theoretical models and computer simulations (Eckert and Barrett 1992; Husband and Barrett 1992a) also demonstrate the inherent instability of the tristylos genetic polymorphism in the face of the kinds of levels of genetic drift observed in natural populations of species of *Eichhornia* (Husband and Barrett 1992b). These population-level studies indicate that the number of origins for homostyly may be considerably greater than can be revealed through phylogenetic analysis using species as OTUs. In the future, genealogical studies at the population level may enable more refined estimates of the number of evolutionary events that are occurring below the species level.

This range of microevolutionary and genetic evidence strongly suggests that evolutionary shifts to homostyly occur readily. The phylogenetic reconstruction based on molecular evidence from the chloroplast (Fig. 2.2A, 2.2B) indicates that tristyly evolved near the base of the family and was subsequently lost on at least three occasions, with at least two losses giving rise to homostylos species. One loss was associated with a shift to an enantiostylos floral form in *Monochoria*. *Pontederia* includes one species lacking tristyly (*P.*
parviflora). This was not available for the current molecular analysis, but probably represents another case of the loss of tristyly.

The conflict between molecular and morphological phylogenetic analyses concerning the evolution of selfing in Pontederiaceae suggests that either molecules or morphology (or both) are not telling the whole truth concerning phyletic descent in the family. A variety of phenomena can cause erroneous reconstruction of phylogenetic history when using molecular data based on single genetic linkage groups, e.g., lateral gene transfer, mistaken genetic orthology (Doyle 1992) and ancestral polymorphism (Pamilo and Nei 1988). The simplest interpretation of the systematic evidence from Pontederiaceae may be that the selfing syndrome has distorted phylogenetic reconstructions from morphological data (but see Chapter 3). This conflict among different data sets highlights the need to use a variety of sources of data in phylogenetic reconstruction.

C. Evolution and Adaptive Significance of Enantiostyly. — Outside Pontederiaceae, heterostyly does not occur in any putatively related family of monocotyledons (e.g., Commelinaceae, Haemodoraceae, Philydraceae). A report of heterostyly in Aneilema aequinoctiale (Commelinaceae) by Vogel (1955) is almost certainly a misinterpretation of the true nature of the polymorphism (Ornduff 1974; Faden 1991; S. Vogel, pers. comm.). Heterostyly has only reliably been reported from two other monocotyledonous taxa: Nivenia of Iridaceae (Mulcahy 1965; Goldblatt and Bernhardt 1990) and Narcissus triandrus of Amaryllidaceae (Fernandes 1935; Lloyd et al. 1990; Barrett et al. 1995). The distinctive nature of heterostyly in Nivenia and Narcissus (Table 2.2) compared with its expression in Pontederiaceae, and the fact that molecular evidence indicates that Iridaceae and Amaryllidaceae are distantly related to this family (Chase et al. 1993), lends strong support to the hypothesis that heterostyly is apomorphic within Pontederiaceae.
TABLE 2.2. Occurrence of heterostyly in the monocotyledons and general features of the syndrome.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Number of heterostylosus species (total number in genus)</th>
<th>Type of heterostyly</th>
<th>Incompatibility Type</th>
<th>Expression</th>
<th>Ancillary polymorphisms Stigmas</th>
<th>Pollen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pontederiaceae</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pontederia</em></td>
<td>5 (6)</td>
<td>Tristyly</td>
<td>TSI</td>
<td>Strong, with different sites of inhibition</td>
<td>Strong</td>
<td>Strong</td>
</tr>
<tr>
<td><em>Eichhornia</em></td>
<td>3 (9)</td>
<td>Tristyly</td>
<td>TSI</td>
<td>Variable, from absent or cryptic to weak</td>
<td>Weak</td>
<td>Weak</td>
</tr>
<tr>
<td><strong>Amaryllidaceae</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Narcissus</em></td>
<td>1 (30)</td>
<td>Tristyly</td>
<td>?</td>
<td>Late-acting SI, and/or inbreeding depression</td>
<td>Weak</td>
<td>Absent</td>
</tr>
<tr>
<td><strong>Iridaceae</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Nivenia</em></td>
<td>5 (9)</td>
<td>Distyly</td>
<td>Absent</td>
<td>Plants apparently fully self-compatible</td>
<td>Absent</td>
<td>Weak</td>
</tr>
</tbody>
</table>

In contrast, enantiostyly is reported from three families of monocotyledons with possible close affinities to Pontederiaceae: Philydraceae (Simpson 1990), Haemodoraceae (Wilson 1887; Ornduff and Dulberger 1978; Simpson 1990) and Commelinaceae (Faden 1991). This raises several issues concerning the evolution and phylogenetic origins of enantiostyly in these related families. What are the features of enantiostyly in these groups? Has the floral polymorphism originated independently in each family? What are the evolutionary relationships between heterostyly and enantiostyly in Pontederiaceae, and are the two conditions independent responses to pollinator-mediated selection for increased mating efficiency?

Several floral traits (e.g., heteranthery, zygomorphy, the absence of nectar secretion, and poricidal anther dehiscence) are commonly associated with enantiostyly in a variety of unrelated angiosperm taxa (Bowers 1975; Ornduff and Dulberger 1978; Dulberger and Ornduff 1980; Dulberger 1981; Buchmann 1983). These assemblages of floral characters are found to varying degrees in the enantiostylous monocotyledons (Table 2.3). These traits are discussed with respect to their distribution among enantiostylous taxa and the possible evolutionary significance of their associations with the enantiostylous floral form.

In the sense used here, enantiostyly is the possession of flowers with left- and right-bending styles, typically with a single stamen reflexed in a lateral position opposite the stigma. While this condition can apparently exist as a true genetic polymorphism [e.g., in Wachendorfia paniculata (Haemodoraceae) (Ornduff and Dulberger 1978)], with individual plants possessing either right- or left-bending styles, it more commonly occurs as a somatic polymorphism with both right and left-handed flowers occurring within the same individual. In Pontederiaceae the polymorphism is of this latter type and is usually associated with a clear stamen dimorphism. In Monochoria and Heteranthera, the reciprocally reflexed stamen tends to be larger than the other stamens and cryptically coloured. Such dimorphism is known as heteranthery when it represents a functional division of labour among the stamens into predominantly attractive
### TABLE 2.3. Occurrence of enantiostyly in the monocotyledons and general features of the syndrome.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Perianth Symmetry</th>
<th>Floral orientation</th>
<th>Stamen dimorphism</th>
<th>Anther dehiscence</th>
<th>Nectaries</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Zygo-</td>
<td>Outward</td>
<td>Upward</td>
<td>Present</td>
<td>Absent</td>
</tr>
<tr>
<td>Pontederiaceae</td>
<td>Actinomorphic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Heteranthera</em></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td><em>Monochoria</em></td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>Haemodoraceae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Wachendorfia</em></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td><em>Schiekia</em></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>Other genera 1</td>
<td>two species</td>
<td>most</td>
<td>species</td>
<td>✓</td>
<td>most</td>
</tr>
<tr>
<td>Philydraceae</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>(Single stamen only)</td>
</tr>
<tr>
<td>Tecophilaeaceae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Cyanella lutea, C. alba</em></td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Commelinaceae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Aneilema</em> 2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

"feeding" stamens and one or more cryptically-coloured "pollinating" stamens (Vogel 1978; Buchmann 1983; Lloyd 1992).

Enantiostyly and heteranthery are reported in a few phylogenetically unrelated angiosperm groups, and are commonly found associated together. Most taxa in *Monochoria* and *Heteranthera* are enantiostylosous and heterantherous. In *H. reniformis* and *M. vaginalis*, it is reported that pollinators ignore the single pollinating anther and are instead attracted to the feeding anthers (Müller 1883; Iyengar 1923). However, this division between attractive and fertilising functions is probably not absolute, since pollen from the feeding anthers is capable of fertilisation (S. C. H. Barrett, unpubl. data). *Zosterella dubia*, *Hydrothrix*, and *M. cyanea* all lack stamen dimorphisms (Eckenwalder and Barrett 1986) and at least the first two are also not enantiostylosous. Apart from Pontederiaceae and *Cyanella* (Dulberger and Ornduff 1980), it is not clear how many of the instances of stamen dimorphism indicated in Table 2.3 represent true heteranthery.

Faden (1991) was hesitant about calling the floral heteromorphism found in certain taxa of *Aneilema* true enantiostyly, since reciprocal deflection of a single stamen against the left- or right-bending style is not found. However, as is the case with the floral heteromorphism in *Cassia didymobotrya* (Caesalpiniaeae) (Dulberger 1981) flowers in these species possess two laterally-placed stamens such that whether the style is left- or right-bending, it is always reflexed against one of these stamens. I feel less hesitant about calling this floral heteromorphism enantiostyly.

Enantiostyly has most often been interpreted as an adaptation for increasing the proficiency of cross-pollination (Todd 1882; Iyengar 1923; Ornduff and Dulberger 1978; Webb and Lloyd 1986). In a manner analogous to heterostyly, the consistent spatial positioning on the pollinator's body of the sites of pollen dispatch and receipt may promote pollen transfer between individuals. One problem with this hypothesis is that this consistency in pollen
transfer may also actively promote geitonogamy (between-flower selfing) when regular visitation of opposite-form flowers takes place within the same plant. Promotion of geitonogamy may, however, be low if only one or a few flowers are open on a given day, or if the flight path of the pollinator is such that few flowers are visited per individual (Dulberger 1981). Geitonogamous matings will also be reduced if an SI system is present, as is the case with *Cyanella alba* and *C. lutea* (Tecophilaeaceae) (Dulberger and Ornduff 1980), but obviously SI by itself cannot act to increase the proficiency of cross-pollen transfer.

Dulberger (1981) suggested that the primary role of enantiostyly in species of *Cassia* is to provide pollinators with unobstructed access to feeding anthers, while protecting the gynoecium during vibrational collection of pollen (buzz pollination) from poricidal anthers. It is unclear if this explanation holds for enantiostylos species in general, since it is not certain that all enantiostylous species are buzz pollinated. For example, species of *Heteranthera* and most enantiostylous Haemodoraceae have longitudinal anther dehiscence and are therefore probably not buzz pollinated, since this pollen-collecting behaviour is strongly associated with poricidal anther dehiscence (Buchmann 1983).

Simpson (1990) suggested that in species with actinomorphic flowers, enantiostyly serves only to reduce the amount of self-pollination, by increasing stigma-anther separation. When the flower is zygomorphic, pollinators will be positioned consistently with respect to the pollinating anther and style. In actinomorphic flowers there may be no consistency in pollinator approach to the flower and hence in the sites of pollen dispatch and receipt on the pollinator's body. Thus, enantiostyly may not function to increase the proficiency of cross-pollen transfer in radially-symmetrical flowers. Among the enantiostylous taxa of Haemodoraceae, only *Wachendorfia, Schieckia, Xiphidium xanthorrhizon* and *Barberetta aurea* possess zygomorphic perianths (Simpson 1990; Fig. 10 in Simpson 1993). Species of Philydraceae have enantiostylos, strongly zygomorphic flowers with only a single stamen. Within Pontederiaceae, species of *Heteranthera* are mostly zygomorphic while species of *Monochoria*
are actinomorphic. Flowers borne on *Monochoria* inflorescences are outwardly facing, so that pollinators are likely to approach them in a consistent orientation. Because of this feature and their enantiostylous-heterantherous morphology, the flowers of species of *Monochoria* may be functionally zygomorphic. A parallel case is found in Tecophilaeaceae, where flowers of *Cyanella alba* and *C. lutea* are actinomorphic, but are enantiostylous, outwardly-facing and apparently heterantherous (Dulberger and Ornduff 1980).

Pollen from the feeding anthers of heterantherous species serves to attract pollinators in place of nectar. Secretion of nectar may not be present in *Monochoria* (Cook 1989) and nectaries are absent from *Heteranthera* (Van Heel 1988; Simpson 1990) and Philydraceae (Dahlgren and Clifford 1982). All species of *Monochoria* have poricidal anthers, and are therefore probably buzz-pollinated by pollen-collecting bees (Buchmann 1983). Although it is not a universal association, poricidal anthers are a well documented feature of enantiostylous taxa (*Solanum rostratum*: Bowers 1975; *Cyanella*: Dulberger and Ornduff 1980; *Cassia*: Dulberger 1981). Enantiostyly is also not always associated with an absence of nectar secretion. Apart from *Xiphidium*, all enantiostylous Haemodoraceae species have septal nectaries (Simpson 1993).

Repetition in elements of the enantiostylous syndrome among phylogenetically disjunct taxa is probably indicative of similar selective pressures operating on floral morphology. The regular association of enantiostyly with outwardly-facing, zygomorphic flowers (Table 2.3) suggests that consistent positioning of the pollinator is usually an important part of the functional operation of enantiostyly. Heteranthery, poricidal anther dehiscence and absence of nectaries are all likely to be associated with pollen-collection by pollinators, so their frequent co-occurrence may not be surprising. It would be particularly valuable to use a phylogenetic approach to determine if traits associated with pollen-collection by pollinators are truly more commonly associated with enantiostyly than might be expected by chance, and to perform experimental studies to examine the functional significance of the different components of the
enantiostylos syndrome, in much the same way as has been conducted for heterostylos plants (e.g., Ganders 1974; Kohn and Barrett 1992).

The precise evolutionary relationships of enantiostyly to heterostyly (if any) are unknown, but it is intriguing to note that both conditions involve forms of reciprocal herkogamy and dimorphic stamens. The primitive floral form in Pontederiaceae may be homologous with that found in Haemodoraceae and Philydraceae, if these are indeed the sister groups of the family. The reconstructions presented in Fig. 2.2A and 2.2B indicate an independent origin of enantiostyly in *Monochoria*. Depending on how the outgroup is coded, enantiostyly in *Heteranthera* represents either a second independent origin of the floral form in the family (Fig. 2.2A) or the primitive floral condition of the family (Fig. 2.2B). In both optimisations, enantiostyly is lost in two lineages within *Heteranthera* s.l.

**Conclusion**

Studies of floral evolution have largely been performed using contemporaneous, population-level evidence. However, in recent years the importance of adding an historical component to such studies has become widely appreciated (Donoghue 1989; Cox 1990; Sytsma, Smith and Berry 1991; Rieseberg, Hanson and Philbrick 1992; Weller, Donoghue and Charlesworth 1995). Phylogenetic systematics can provide this historical perspective. The addition of new phylogenetic data from a variety of different sources serves to strengthen our confidence in reconstructions of the evolutionary history of organisms and of their constituent character complexes. It can also function to highlight deficiencies in the capacity of any particular class of data to permit the accurate reconstruction of historical events. This paper brings together a range of phylogenetic evidence from morphological and molecular sources to examine the systematics of Pontederiaceae and the evolutionary history of polymorphic breeding
systems present in the family. I argue that tristyly probably evolved once in the family and that there have been multiple breakdowns of the syndrome to self-fertilisation via the evolution of homostyly. Given the diversity in form of enantiostyly in this and putatively related families, I suggest that these taxa provide excellent opportunities for phylogenetic, as well as functional, investigations of the evolutionary significance of this floral syndrome. Future systematic studies of Pontederiaceae should concentrate on providing more robust evidence concerning its local placement within the monocotyledons and on collecting phylogenetic data from a greater range of morphological and molecular sources.
CHAPTER 3

- Phylogenetic Congruence and Discordance Among One Morphological and Three Molecular Data Sets From Pontederiaceae -

60
CHAPTER 3. PHYLOGENETIC CONGRUENCE IN PONTEDERIACEAE

INTRODUCTION

When multiple sources of phylogenetic evidence are available for a given group of organisms, it seems desirable to combine all the available data (the "total evidence") (Kluge 1989; Barrett, Donoghue and Sober 1991). Unfortunately, different groups of characters may have experienced, and may therefore reflect, different evolutionary or genealogical processes. It is therefore potentially misleading to combine all the available sources of evidence, if the concerted effect of each process is to distort reconstruction of organismal phylogeny (Bull et al. 1994; Huelsenbeck et al. 1994; de Queiroz, Donoghue and Kim 1995). Concerted processes affecting sub-sets of the available data may involve co-ordinated adaptive processes or different functional constraints. Examples of such "process partitions" include adaptive syndromes affecting sub-sets of characters, such as the floral traits involved in wind- and self-pollination (Dahlgren, Clifford and Yeo 1985: pp. 346; Wyatt 1988); functional constraints associated with base-pairing requirements in ribosomal RNA (e.g., Tillier and Collins 1995); and differences in substitution rates among codon positions associated with the degeneracy of the genetic code (e.g., Albert, Chase and Mishler 1993). A different type of bias can occur when all or some of the characters used to reconstruct a phylogeny are genetically linked. Hybridization events and lineage sorting of ancestral polymorphisms can lead to incorrect inference of parts of "species trees" when linked characters, such as the entire chloroplast or mitochondrial genomes, are employed for phylogenetic reconstruction. Such genealogical biases can occur when groups of linked characters are simultaneously transferred among descendant lineages during sorting of ancestral polymorphisms, or between species during hybridization events (see Pamilo and Nei 1986; Doyle 1992).

Characters governed by different evolutionary processes or genetic histories may still be capable of converging on the correct gene or organismal tree, if these biases can be accounted
CHAPTER 3. PHYLOGENETIC CONGRUENCE IN PONTEDERIACEAE

for during phylogenetic reconstruction through the application of appropriate weighting schemes (e.g., Albert, Mishler and Chase 1992; Doyle 1992; Chippindale and Wiens 1994; Huelsenbeck et al. 1994; Tillier and Collins 1995), but some processes may operate idiosyncratically within particular lineages or groups of characters (see Ritland and Eckenwalder 1993). There also seems to be no consensus on how to convert empirical evidence concerning such process partitions into appropriate weighting schemes (see Maddison and Maddison 1992: pp. 60-62, 295; de Queiroz, Donoghue and Kim 1995).

However, if there is good correspondance or "congruence" among phylogenies derived from different data sets, this can be taken as independent corroboration that each data set is correctly inferring the "true" gene or organismal phylogeny (Penny, Foulds and Hendy 1982). Conversely, discordance among data sets may serve to highlight cases where different processes are acting concertedly on groups of characters to bias phylogenetic inference. Discordance may also be a function of inadequate character sampling in one or more data sets. Two major approaches for assessing the degree of corroboration among data sets consider "character congruence," how well each data set fits the available phylogenetic hypotheses, and "taxonomic congruence," the degree of topological agreement among trees from different data sets (reviewed in Swofford 1991).

Four phylogenetic data sets for the monocotyledonous family Pontederiaceae were assessed for phylogenetic congruence. One of the data sets is based on morphological evidence obtained by Eckenwalder and Barrett (1986), who suggested that a "selfing syndrome" involving concerted evolutionary shifts in a subset of these characters acted to subvert phylogenetic inference in their study. This motivated us to investigate new sources of phylogenetic evidence for the family. A study of chloroplast DNA restriction-site variation has recently been published (Appendix A), and new surveys of nucleotide variation in portions of two chloroplast genes, ndhF and rbcL, are presented here. Evidence from rbcL provides strong support for the monophyly of the family (Chapter 2). More than one source of chloroplast data
was considered to address the concern that a single chloroplast data set would not provide a sufficient sampling of characters for inferring a sturdy phylogeny.

Because all three chloroplast data sets (ndhF, rbcL and the restriction-site survey) are part of the same genetic linkage group, they can not give independent estimates of phylogenetic history during any periods of introgression or lineage sorting of ancestral polymorphisms. However, this does not mean that different sources of data from the chloroplast genome will necessarily converge upon a single (chloroplast) phylogeny. With respect to the restriction-site data, it is known that different regions of this genome evolve at different rates. The inverted-repeat region, for example, evolves at a slower rate than the rest of the chloroplast genome (Jansen and Palmer 1987; Wolfe, Li and Sharpe 1987). Rates of restriction-site change are also affected by among-codon variation in rates of nucleotide substitution and differences in transition and transversion rates (Albert, Mishler and Chase 1992), and these rates can vary from gene to gene. With regard to the sequence data sets, for example, the ratio of synonymous to non-synonymous substitutions, and of transitions to transversions, are known to differ within and between rbcL and ndhF (see Albert, Chase and Mishler 1993; Kim and Jansen 1995). Since different subsets of the chloroplast genome have different rules of evolutionary change, they may therefore give biased estimates of the phylogenetic history of the chloroplast genome in the family.

Although different rules of evolution occur among and within different chloroplast-based data sets, it is not clear how to translate this information into appropriate weighting schemes (although see Albert, Mishler and Chase 1992; Albert, Chase and Mishler 1993), and of course, not all evolutionary processes within the chloroplast genome are either known or well characterized. With a known phylogeny appropriate weightings would be easier to derive (Huelsenbeck et al. 1994), but usually the phylogeny of a group is not known in advance. Because of these difficulties, I assumed equal rates of change for all character-state transformations during phylogenetic inference. Obtaining congruent phylogenetic
reconstructions under this null model thus requires that the phylogenetic information in different data sets is congruent, and that this simple model of character evolution does not unduly interfere with the congruency in the phylogenetic signal, or less likely, that it compensates for any incongruency in the phylogenetic signal. I first examined whether estimates of the reconstructed phylogenetic history of the family from the three chloroplast data sets disagreed in any significant manner, prior to combining these three different data sets (Bull et al. 1994). This combined estimate of phylogenetic history was then compared to one obtained from the morphological data, to provide an appraisal of agreement between morphological and chloroplast sources of phylogenetic information. In examining the effect of these "classes" of data (ndhF, rbcL and restriction site; chloroplast-based and morphology-based) I recognize that these are only several of many possible partitionings of the data that may be relevant to their combinability (Chippindale and Weins 1994; de Queiroz, Donoghue and Sober 1995).
However, if these classes are found to be concordant, this suggests they are combinable, and if they are not, this provides an impetus for further investigation of the source of their discordance.

MATERIALS AND METHODS

Twenty four taxa of Pontederiaceae (22 species and three varieties of Pontederia cordata) and an outgroup species (Philydrum lanuginosum from Philydaceae) were considered in this study. A substantial proportion of the 35-40 described species in the family were represented. All genera were represented, apart from the rare monotypic Scholleropsis. These taxa were examined for DNA sequence variation in two chloroplast genes, rbcL and ndhF. The rbcL locus codes for the large subunit of ribulose-1,5-bisphosphate carboxylase oxygenase (EC 4.1.1.39) and ndhF for subunit F of a chloroplast NADH oxidoreductase of unidentified
function (Gruissem and Tonkyn 1993). These data sets were examined in concert with a revised version of the morphological data set of Eckenwalder and Barrett (1986) and a survey of restriction-site variation in the chloroplast genome (Appendix A).

Minor revisions and additions to the morphological data are presented in Appendix C. The portion of \textit{rbcL} sequenced represents the majority of this gene (Appendix H). The portion of \textit{ndhF} examined is a major portion of the highly variable 3'-end region of this gene (Appendix H). This region of \textit{ndhF} was sequenced using primers I designed based on \textit{ndhF} sequences from one monocotyledon and two dicotyledons (see Appendix H). The oligonucleotides used as primers are listed in Table 3.1. Experimental protocols for DNA extraction, amplification and sequencing are provided in Appendix H. Experimental protocols for inference of restriction-site variation are given in Appendix A. All data have been submitted to GenBank and have accession numbers U41573-U41597 for \textit{rbcL} partial sequences and U41598-U41622 for \textit{ndhF} partial sequences. Source information for each taxon is given in Appendix B.

\textbf{Analyses}

Maximum parsimony analyses were conducted using PAUP version 3.1.1 (Swofford 1993). Length differences were not encountered in the region of \textit{rbcL} examined, but two short (6 bp) indels were observed in \textit{ndhF}, one of which was potentially informative. These indels were coded as an additional character state using the GAPMODE option in PAUP, and the single informative gap was coded as an extra character, an approach that is valid because no other sequence variation was observed in taxa lacking the gap (Eernisse and Kluge 1993). Seven restriction sites involving the enzymes used to produce the restriction-site data were located in \textit{rbcL} (one was potentially informative and three were autapomorphic) and four in \textit{ndhF} (two were potentially informative). The latter two potentially informative sites in \textit{ndhF} may correspond to sites in the restriction-site data and were therefore excluded from \textit{ndhF} data.
### TABLE 3.1. Oligonucleotides employed to amplify and sequence a 3’-portion of the chloroplast gene *ndhF*.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
<th>Base pair in <em>ndhFa</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>ndh2F</td>
<td>5’-actcatgcttacgaaagc</td>
<td>1042-1061</td>
</tr>
<tr>
<td>ndh3F&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5’-tatccaatatgttatggg</td>
<td>1420-1439</td>
</tr>
<tr>
<td>ndh4F&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5’-cttattcattggatcaatggaat</td>
<td>1655-1679</td>
</tr>
<tr>
<td>ndh4R&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5’-gagtttaaccctttgataata</td>
<td>1712-1732</td>
</tr>
<tr>
<td>ndh2R&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5’-ctataaaccgctttatgacc</td>
<td>1961-1984</td>
</tr>
<tr>
<td>ndh1.6R</td>
<td>5’-ctctacctgtgtaattcctt</td>
<td>2066-2087</td>
</tr>
<tr>
<td>ndh1R</td>
<td>5’-aataataaagagcaatggcc</td>
<td>2134-2156</td>
</tr>
</tbody>
</table>

<sup>F</sup> = forward strand, <sup>R</sup> = reverse strand.

<sup>a</sup> Reference sequence = *Oryza sativa.*  <sup>b</sup> Used as sequencing primer.
in the combinations involving these two data sets. The three chloroplast data sets were each analyzed individually, in three pair-wise combinations and in one three-way combination. The morphological data were analyzed individually and in a four-way combination with the chloroplast data. These heuristic searches were performed using the TBR (tree bisection-reconnection) branch-swapping option in PAUP, with MULPARS and "Steepest descent" options activated, and with all character and character-state changes equally weighted. All character-state transformations were unordered. One hundred random addition replicates were employed in all searches to increase the efficacy of the heuristic algorithm (Maddison 1992). The analyses were repeated both with and without the outgroup included.

Measures of tree resolution and support -- The degree of resolution of most-parsimonious trees resulting from each unconstrained analysis was measured as the number of fully resolved nodes and the number of internal taxon partitions (non-terminal branches) retained in strict consensuses of these trees. The strict consensus summarizes taxon partitions common to all shortest trees. Non-parametric bootstrap analyses (Felsenstein 1985a) were performed to estimate the robustness of taxon partitions. Tree support was taken as the average bootstrap support of taxon partitions retained in the strict consensus of the shortest trees from each maximum parsimony analysis (Olmstead and Sweere 1994). Although there have been a number of statistical (and philosophical) objections to the use of the bootstrap for assessing the reliability of phylogenies, reviewed in Sanderson (1995), bootstrapping remains the most widely-used approach for doing this. Hillis and Bull (1993) suggest that bootstrap analysis provides fairly conservative estimates of how well the reconstructed phylogeny infers actual trees, and provide some rough guidelines for what constitutes good bootstrap support. A few other methods have been used to assess tree support. One "non-statistical" method is to derive "decay indices" for branches that describe their stability (persistence) in trees less optimal than the shortest ones (Bremer 1988). Empirical comparisons of bootstrapping versus decay indices
suggest that these measures are addressing the same parameter (Olmstead and Sweere 1994), and so I employed only one of them in this study.

*Measures of character incongruence.* -- One way to assess the congruency of phylogenetic signals among different data sets is to measure the size of the penalty in parsimony required to find sub-optimal trees with one data set that are optimal for another data set (Swofford 1991). This is of interest because trees with topologies identical to those derived from another data set may be found by slightly relaxing the parsimony criterion, in which case the data sets are highly congruent, or may require substantial decreases in optimality, in which case they are not. These sub-optimal trees may be found by a search procedure described by Swofford (1991). However, a simpler way to "find" them is to directly optimize each data set onto the most-parsimonious trees from alternative data sets and determine the resulting increases in tree lengths beyond the shortest "native" trees for the data set. I use the term "native" to refer to the set of trees found by analysis of a given data set (or to that data set), and "non-native" to refer to sets of trees found using other data sets (or to those data sets). I determined the range and average increase in tree length for several comparisons of individual and combined data sets. These excess lengths were normalized by expressing them as a percentage of the shortest native trees.

Two incongruence indices, $I_{MF}$ of Mickevich and Farris (1981), and $I_M$ of M. M. Miyamoto (in Kluge 1989), were calculated to further assess character congruence between the morphological data and the combined chloroplast data. These incongruence measures are derived from counts of the number of "extra steps" for a given data set (on a particular tree or trees) beyond the minimum possible number of synapomorphies for that data set. This may be partitioned into within and between data-set incongruence, and each index is determined by expressing the between data-set incongruence as a proportion of the total incongruence. For $I_{MF}$ the total incongruence is defined with respect to a combined data set and the trees derived from that data set, whereas for $I_M$ it is the sum of the incongruence of each data set with respect
to trees derived from the other data set. Thus, these incongruence indices actually assess "character by tree" incongruence, since they are made with reference to the fit of one or more data sets to the available phylogenetic hypotheses. Since a range of values of extra steps is usually obtained upon optimizing one data set onto trees obtained from another data set, $I_M$ can be based on the lowest number of extra steps for each of the two sets of optimizations (Swofford 1991). It could also be based on the mean number of extra steps in each case, but since this made little difference to the final value of $I_M$ I obtained, I report only the former value here.

A novel measure of congruence considered here is the average bootstrap support for branches common to all shortest trees derived from one data set, based on independent bootstrap analysis of the other data set. This can be calculated because taxon partitions seen in the strict consensus of shortest trees from one data set, but not in the strict consensus of another, may still be found among the replicates of a bootstrap analysis of that other data set. Support for all partitions observed among the bootstrap replicates can be obtained using PAUP, and it is relatively straightforward to determine which taxon partitions correspond to those in the non-native strict consensus tree and use this information to calculate the average bootstrap support by a given data set of the non-native strict consensus tree. Those non-native taxon partitions that are not supported by any bootstrap replicates can be assigned a value of zero in this calculation. This value can then be compared to the average bootstrap support by each data set for its native trees.

Templeton's test for significant differences among phylogenies -- An explicitly statistical approach for measuring character incongruence among different phylogenies can be obtained using a test due to Templeton (1983; see also Felsenstein 1985b) that is closely related to maximum likelihood and maximum parsimony approaches described by Kishino and Hasegawa (1989). Templeton's test is implemented in PHYLIP (Felsenstein 1995). The test assesses
whether differences between trees can be explained by chance alone, by considering the mean and variance of step differences between different phylogenies, taken across the characters under consideration. Trees are deemed to be significantly different if the mean of one tree is more than 1.96 standard deviations different from another. When the trees being compared are derived from different data sets, Templeton's test is a measure of the congruence between these data sets (Paterson, Wallis and Gray 1995). All tests were performed with the DNAPARS algorithm of PHYLIP version 3.5. Felsenstein's current implementation of MIX does not handle multistate unordered characters and so the morphological data set was re-coded as nucleotide data. Re-coding is valid because most characters in the morphological data set have four or less unordered character-states, and all character-state transformations are unordered. Characters 16 and 26 in the morphological data set each have five character-states (Eckenwalder and Barrett 1986; Appendix C) but re-coding was possible after dropping one autapomorphic character-state in each case. Two taxa that were coded as polymorphic for character 16 (Appendix C) were re-coded as "missing data" for PHYLIP. The restriction-site data set and the two indels in the *ndhF* data set were also re-coded as nucleotide data. Templeton's test detects overall incongruence among trees, given a particular data set. When the trees being compared are derived from two different data sets, it seems appropriate to repeat the test for both data sets (cf., Paterson, Wallis and Gray 1995). I take the existence of non-native trees that are not significantly different from at least some native trees, in tests made with reference to each data set, as an indication that the two data sets are not significantly incongruent. If there is no such "overlap" between native and non-native trees, then the data sets are significantly incongruent.

*Evidence for a selfing syndrome in* Eichhornia. -- Templeton's test can be used to detect process partitions, if the predicted effect of the process partition can be described in terms of a topological constraint, such as the linkage of a particular group of taxa. Eckenwalder and
Barrett (1986) suggested that during the evolution of selfing species of *Eichhornia*, an adaptive syndrome occurred involving convergent adaptive shifts in a whole suite of floral characters. These multiple parallel shifts may be mistaken for synapomorphies, in which case they would (falsely) tend to link selfing taxa during phylogenetic reconstruction. This selfing syndrome is a type of process partition that would be expected to affect morphology-based phylogenies, but not those based on the chloroplast genome (see Appendix A). The phylogenetic linkage of predominantly selfing species of *Eichhornia* was enforced in a separate set of heuristic searches using the "Topological constraints" option in PAUP, to examine the plausibility of this phylogenetic arrangement (and hence the existence of a selfing syndrome) for the revised morphological data and for the combined molecular data. Two different sets of constraints were considered. In one of these, all of the selfing species of *Eichhornia* formed one constrained group (i.e., *Eichhornia diversifolia*, *E. heterosperma*, *E. meyeri*, *E. paradoxa* and a closely related but undescribed species referred to as *Eichhornia* sp. here). In the other, *E. meyeri* was excluded from this constrained group (see Appendix A). In a novel application of Templeton's test, the trees found using the topologically constrained searches that united selfing taxa of *Eichhornia* were compared to the shortest unconstrained trees, for the chloroplast data and for the morphological data. If the trees from the constrained analysis are found to exhibit significant increases in tree length, the selfing syndrome can be considered to be an implausible evolutionary scenario. Because the constraint analyses I performed did not include the outgroup, it should be noted that these analyses did not constrain monophyly *per se*, only particular taxon partitionings (monophyly statements require that the root is known).

*Measures of taxonomic congruence* -- In addition to computing strict consensus trees, I assessed the similarity in the shape of trees from different analyses. A variety of metrics are available for assessing the topological distinctness of different trees to each other (see Penny and Hendy 1985), and I used one of these, the partition metric of Robinson and Foulds (1979,
1981). The metric is simply the total number of unique partitionings of the taxa observed in pair-wise comparisons of trees. The distribution of this metric has been assessed for up to 16 taxa and is highly skewed towards a maximum distance of $2n-6$ symmetric differences for $n$ taxa (Steel and Penny 1993), so random pairs of trees are extremely unlikely to have a low partition distance. The distribution was estimated for 24-taxon trees from 999 random trees produced using the random-tree generator in MacClade version 3 (Maddison and Maddison 1992). Three methods of random tree generation were available: equiprobable trees, random joining and random partitioning. These methods can result in somewhat different distributions of tree shape (see Maddison and Maddison 1992). Each method was used to produce one third of the random trees. If trees derived from different data sets are closer to each other than random trees are to each other, this indicates that the different data sets are estimating, at least in part, the same phylogenetic history (Penny, Foulds and Hendy 1982).

Tree-to-tree distances were determined in PAUP for all of the unrooted trees from the parsimony analyses. To summarize the overall similarity among these trees, the resulting distance matrix was converted into a "tree of trees" using the neighbor-joining algorithm in PHYLIP version 3.5c (Felsenstein 1995). The shared similarity among all the trees from a given data set is accounted for on this summary phenogram, and so the overall distance among sets of trees from two or more data sets can be assessed. The within tree-set correction for assessing taxonomic incongruence among sets of trees is only possible when trees from particular data sets occur as distinct clusters in the summary phenogram. The idea of treating trees as distinct objects whose dissimilarities can be defined and used to construct a summary tree has been suggested by other workers (see Podani and Dickinson 1984), but to my knowledge has not been used to address phylogenetic congruence among different data sets. The validity of using a hierarchical (phenogram-based) summary of taxonomic incongruence among trees was assessed by measuring how well correlated the raw matrix of tree-to-tree distances was to a secondary matrix derived from tree-to-tree distances on the summary
phenogram (Sneath and Sokal 1973: pp. 278-280). Cophenetic correlation coefficients (rcs) between these matrices were calculated using NYTSYS version 1.80 (Rohlf 1993).

RESULTS

Basic tree statistics for unrooted trees from the analyses of the individual and combined data sets are presented in Table 3.2. Both chloroplast genes yielded approximately the same number of potentially informative characters, despite the fact that almost three times the length of sequence was examined for variation in rbcL as compared to ndhF. The restriction-site data set had approximately two-thirds more potentially informative characters than either sequence data set alone. Consistency indices (CI; excluding autapomorphies) for the molecular data ranged from 0.454 for the restriction-site data to 0.628 for the ndhF-sequence data. The CI for the morphological data (0.474) was slightly higher than for the restriction-site data. Among the chloroplast data sets, the three individual (uncombined) data sets showed the greatest range of CIs, and the five examined combinations of the chloroplast and morphological data sets had CIs in the middle of this range. Retention indices (RI) were tightly correlated with the two estimates of CI across the nine data sets and data set combinations. The correlation between RI and CI (excluding autapomorphies) was 0.977; correlation between RI and CI (including autapomorphies) was 0.939.

Tree resolution and support -- With the molecular trees, there was a general trend for increase in tree resolution and support with increasing size of the uncombined or combined data set (Table 3.2). From 12 to 20 fully resolved nodes (from the rbcL data set to the combined rbcL and
TABLE 3.2. Summary statistics for phylogenetic trees of Pontederiaceae (ingroup taxa only) derived from individual and combined analyses of the chloroplast and morphological data-sets.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>No. potentially informative characters</th>
<th>Length&lt;sup&gt;a&lt;/sup&gt;</th>
<th>No. trees</th>
<th>No. partitions (No. fully resolved nodes)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>CI&lt;sup&gt;c&lt;/sup&gt;</th>
<th>CI&lt;sup&gt;d&lt;/sup&gt;</th>
<th>RI</th>
<th>Mean bootstrap value&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>ndhF partial sequence data&lt;sup&gt;e&lt;/sup&gt;</td>
<td>56</td>
<td>144</td>
<td>9</td>
<td>17 (14)</td>
<td>0.628</td>
<td>0.708</td>
<td>0.833</td>
<td>78</td>
</tr>
<tr>
<td>rbcL partial sequence data</td>
<td>61</td>
<td>168</td>
<td>48</td>
<td>15 (12)</td>
<td>0.576</td>
<td>0.685</td>
<td>0.820</td>
<td>77</td>
</tr>
<tr>
<td>cp restriction-site (R.S.) data</td>
<td>104</td>
<td>299</td>
<td>10</td>
<td>18 (16)</td>
<td>0.454</td>
<td>0.582</td>
<td>0.718</td>
<td>80</td>
</tr>
<tr>
<td>Combined ndhF &amp; rbcL data</td>
<td>117</td>
<td>313</td>
<td>4</td>
<td>19 (18)</td>
<td>0.598</td>
<td>0.693</td>
<td>0.824</td>
<td>84</td>
</tr>
<tr>
<td>Combined ndhF &amp; R.S. data</td>
<td>158</td>
<td>440</td>
<td>8</td>
<td>18 (16)</td>
<td>0.507</td>
<td>0.620</td>
<td>0.757</td>
<td>88</td>
</tr>
<tr>
<td>Combined rbcL &amp; R.S. data</td>
<td>165</td>
<td>468</td>
<td>2</td>
<td>20 (20)</td>
<td>0.496</td>
<td>0.618</td>
<td>0.757</td>
<td>88</td>
</tr>
<tr>
<td>All chloroplast data combined</td>
<td>219</td>
<td>609</td>
<td>4</td>
<td>19 (18)</td>
<td>0.525</td>
<td>0.637</td>
<td>0.775</td>
<td>90</td>
</tr>
<tr>
<td>Morphological data</td>
<td>33</td>
<td>121</td>
<td>5</td>
<td>19 (19)</td>
<td>0.474</td>
<td>0.496</td>
<td>0.715</td>
<td>52</td>
</tr>
<tr>
<td>All current evidence combined</td>
<td>252</td>
<td>748</td>
<td>12</td>
<td>17 (14)</td>
<td>0.499</td>
<td>0.599</td>
<td>0.749</td>
<td>94</td>
</tr>
</tbody>
</table>

<sup>a</sup> Including autapomorphies.  
<sup>b</sup> In strict consensus trees, of a maximum of 21 internal partitions (branches) and 22 fully resolved nodes for 24 taxa.  
<sup>c</sup> Excluding autapomorphies.  
<sup>d</sup> Based on partitions observed in strict consensus trees (including those with less than 50% bootstrap support).  
<sup>e</sup> Including one indel.
restriction-site data set) were retained in the strict consensus trees. There were typically a few more taxon partitions supported in all strict consensus trees than there were fully resolved nodes. The mean bootstrap support ranged from 77% for the \( rbcL \) data alone, to around 90% for several of the two-way and the fully combined chloroplast data. Despite the small number of characters in the morphological data set, its strict consensus tree was highly resolved (Table 3.2; Fig. 3.4). However, its average bootstrap support (52%) was much lower than for any of the molecular data sets. When all of the available data were combined there was somewhat poorer resolution than in the fully combined molecular tree, with 14 fully resolved nodes in the former case compared to 18 in the latter, but the average bootstrap support was marginally higher at 94% (compared to 90%).

The strict consensus trees from the various (unrooted) individual and combined analyses are presented in Figs. 3.1-3.4 (left-hand side trees). All trees were rooted at the same taxon partition, which is also the root found in the combined sequence analysis (Chapter 2, 4 and see Fig. 3.2A), to emphasise the high degree of topological similarity apparent among all trees found in analyses of the molecular data. Bootstrap support for each taxon partition is reported on the corresponding branches of strict consensus trees. Single representative shortest trees from among each set of most-parsimonious trees are included in Fig. 3.1, 3.3-3.4 (right-hand side trees), to illustrate branch lengths. ACCTRAN optimization was employed to compute branch lengths with respect to the rooting depicted. Constructing strict consensuses of trees from all the analyses results in a vast reduction in the amount of information about relationships among taxa. The degree of resolution of strict consensuses of all chloroplast trees, or of all trees (including those derived from morphology) is poor (Fig. 3.5), but worse for the strict consensus of all trees. There were six fully resolved nodes and ten taxon partitions in the former case, \textit{versus} one fully resolved node and five internal partitions in the latter case.

All trees presented in the figures and tables include ingroup taxa only. Inclusion of the outgroup did not resolve where the root of the family is. In several cases no single most-
Fig. 3.1. Results of parsimony analyses for three individual data sets from Pontederiaceae based on variation in the chloroplast genes *ndhF* and *rbcL* (Fig. 3.1A and 3.1B respectively) and restriction-site variation in the chloroplast genome (Fig. 3.1C). Only taxa in Pontederiaceae were considered, but an "R" indicates the root position observed when the outgroup was included in analyses. For the *ndhF* trees, no single most parsimonious root position was found in analyses that included the outgroup. Trees on the left-hand side are strict consensus trees. Bootstrap values are indicated above branches (outgroup excluded) and below branches (outgroup included); "(--)" indicates less than 50% bootstrap support for a branch. Inclusion of the outgroup adds an extra branch to trees, and so two (rather than one) bootstrap values are reported below the branch intersected by the root. Each tree on the right-hand side is a single topology taken from the set of most-parsimonious trees for each data set, used to demonstrate branch lengths.
A. nzyf sequence data alone
B. ricol: sequence data alone
C. Restriction-site data alone

Pontederia cordata v. ovalis
  P. sagittata
  P. cordata v. cordata
  P. cordata v. lancifolia
  P. rotundifolia
Eichhornia azurea
E. heterosperma
E. diversifolia
Monochoria korsakovii
  M. vaginalis
  M. hastata
  M. cyanea
Eichhornia crassipes
  E. paniculata
  Eichhornia sp.
  E. paradoxa
  E. meyeri
Heteranthera limosa
  H. rotundifolia
  H. oblongifolia
  H. seubertiana
  H. dubia
  H. zostericofolia
Hydrothrix gardneri
Fig. 3.2. Results of parsimony analyses of the three possible two-way combinations of the three chloroplast data sets from Pontederiaceae. (A) Strict consensus tree of the most-parsimonious trees from combined analysis of ndhF & rbcL data sets. (B) Strict consensus tree of the most-parsimonious trees from combined analysis of ndhF & chloroplast restriction-site data. (C) Strict consensus tree of the most-parsimonious trees from combined analysis of rbcL & chloroplast restriction-site data. An "R" indicates the root position observed when the outgroup was included in analyses (Fig 3.2A; Fig 3.2B). Bootstrap values are indicated above branches (outgroup excluded) and below branches (outgroup included). "(--)" indicates less than 50% bootstrap support for a taxon partition. No single most parsimonious root position was found in the combined analysis of the ndhF & chloroplast restriction-site data (Fig. 3.2A) when the outgroup was included in the analysis, but a single root position (indicated by an asterisk) was supported in more than 50% of bootstrap replicates.
A. \textit{ndhF} \& \textit{rbcL} sequence data combined
B. ndhF & restriction-site data combined

```
Pontederia cordata v. ovalis
  P. sagittata
  P. cordata v. cordata
  P. cordata v. lancifolia
  P. rotundifolia
  Eichhornia azurea
  E. heterosperma
  E. diversifolia
  Monochoria korsakovii
  M. vaginalis
  M. hastata
  M. cyanea
  Eichhornia crassipes
  E. paniculata
  Eichhornia sp.
  E. paradoxa
  E. meyeri
  Heteranthera limosa
  H. rotundifolia
  H. oblongifolia
  H. seubertiana
  H. dubia
  H. zosterifolia
  Hydrothrix gardneri
```

Pontederia

Monochoria

Heteranthera s.l.
C. \textit{rbcL} & restriction-site data combined

\begin{itemize}
\item \textit{Pontederia} cordata v. ovalis
\item \textit{P.} sagittata
\item \textit{P.} cordata v. cordata
\item \textit{P.} cordata v. lancifolia
\item \textit{P.} rotundifolia
\item \textit{Eichhornia} azurea
\item \textit{E.} heterosperma
\item \textit{E.} diversifolia
\item \textit{Monochoria} korsakovi
\item \textit{M.} vaginalis
\item \textit{M.} hastata
\item \textit{M.} cyanea
\item \textit{Eichhornia} crassipes
\item \textit{E.} paniculata
\item \textit{Eichhornia} sp.
\item \textit{E.} paradoxa
\item \textit{E.} meyeri
\item \textit{Heteranthera} limosa
\item \textit{H.} rotundifolia
\item \textit{H.} oblongifolia
\item \textit{H.} seubertiana
\item \textit{H.} dubia
\item \textit{H.} zostericola
\item \textit{Hydrothrix} gardneri
\end{itemize}
Fig. 3.3. Results of parsimony analyses for a combined chloroplast data set composed of the three individual chloroplast data sets from Pontederiaceae. An "R" indicates the root position observed when the outgroup was included in analyses. The tree on the left-hand side is a strict consensus tree. Bootstrap values are indicated above branches (outgroup excluded) and below branches (outgroup included). "(--)" indicates less than 50% bootstrap support for a taxon partition. Branches experiencing a substantial drop (less than 20%) in bootstrap support with the inclusion of the outgroup taxon are circled. The tree on the right-hand side is a single topology taken from the set of most-parsimonious trees, used to demonstrate branch lengths.
All chloroplast data combined
Fig. 3.4. Results of parsimony analyses for (A) the morphological data and (B) all current data from Pontederiaceae combined. An "R" indicates the root position observed when the outgroup was included in analyses. For the morphological trees, no single most parsimonious root position was found in analyses that included the outgroup. Bootstrap values are indicated above branches (outgroup excluded) and below branches (outgroup included). Less than 50% bootstrap support for a branch is indicated with "(--)". Trees on the right-hand side are individual most-parsimonious trees found for each data set that are included to demonstrate branch lengths.
A. Morphological data alone

- Pontederia cordata v. ovalis
- P. cordata v. lancifolia
- P. cordata v. cordata
- P. sagittata
- P. rotundifolia
- Eichhornia crassipes
- E. paniculata
- E. meyeri
- Monochoria korsakovii
- M. vaginalis
- M. hastata
- M. cyanea
- Eichhornia azurea
- E. heterosperma
- E. diversifolia
- Eichhornia sp.
- E. paradoxa
- Heteranthera limosa
- H. rotundifolia
- H. oblongifolia
- H. zostericola
- H. dubia
- Hydrothrix gardneri
- Heteranthera seubertiana
Fig. 3.5. *Summary* consensus trees of all shortest unrooted trees from the single and combined analyses of the *three* chloroplast data sets (left-hand tree) and of *all the data sets* (right hand tree).
Strict consensus trees: All cp trees & all trees
parsimonious root position was found (i.e., with the \textit{ndhF} data by itself, the \textit{ndhF} and restriction-site data combined, and the morphological data alone). For the other data sets there was a single most-parsimonious root position, but this position differed among data sets (compare Figs. 3.1-3.4). In all analyses where there was a single agreed-upon root position among shortest trees, there was also poor bootstrap support for the rooting (typically less than 50\%). Moreover, a 20\% or greater drop in bootstrap support was always observed for one or more branches neighbouring the root (see the circled values in Fig. 3.3, for example).

Inclusion of the outgroup taxon also resulted in loss of resolution, as measured by the number of taxon partitions or fully resolved nodes in the strict consensus trees (S.W. Graham, unpubl. results). The branch leading to the outgroup taxon accounts for 7-9\% of total tree length for the morphological data, and 22-31\% of tree length for the other cases.

Bootstrap support for branches seen in analyses of the three chloroplast data sets increased upon combination of the data sets, and there was little evidence of serious conflict among well-supported taxon partitions. With regard to the fully combined chloroplast data, for example, bootstrap support was typically greater for the combined data set than for any of the uncombined data sets (Fig. 3.6A). Most taxon partitions (20 out of 25) were seen in at least two of the four analyses, and none of the "shared" partitions common to two of the four analyses contradicted other such partitions in terms of their constituent taxa. There were only five unique taxon partitions from the analyses of these data sets with greater than 50\% bootstrap support, and although they were all contradictory to some aspect of phylogenetic structure among the shared partitions, they each had only moderately high bootstrap support (around 50\% to 70\% of bootstrap replicates; Hillis and Bull 1993). Eleven of the thirteen most highly ranked taxon partitions seen in the fully combined data set had greater than 50\% bootstrap support from all three uncombined data sets. Interestingly, three shared taxon partitions not found in some of the strict consensus trees had moderate support from the corresponding bootstrap analyses.
Fig. 3.6. **Spectrum of bootstrap support for taxon partitions** found in strict consensus trees and at least 50% of replicates in bootstrap analyses of (A) the fully combined chloroplast data set and each uncombined chloroplast data set from Pontederiaceae (ndhF, rbcL and restriction site), (B) the fully combined chloroplast data, the morphological data and all data combined. Partitions shared among at least two of the data sets being compared are ranked according to their support in the three-way combined chloroplast analysis. Those partitions unique to particular analyses are unranked and separated from the other data by a dashed line. Partitions supported by bootstrap analysis that are not found in the corresponding strict consensus tree are included and indicated with a bold arrow. Taxon partitions are illustrated in Appendix D.
The strict consensus tree from the analysis of the morphological data was the only unrooted tree in this study with any branches with less than 50% bootstrap support (10 of 19 branches; Fig. 3.4A). Of the nine branches with greater than 50% bootstrap in this tree, six were also found on trees from the fully combined chloroplast data (Fig. 3.6B). The level of bootstrap support for these six branches was in all cases less than with the molecular data. Bootstrap support for most branches in the molecular trees changed very little upon combination with the morphological data (Fig. 3.6B), a further reflection of the swamping of the morphological data by the molecular data. A few major changes in bootstrap support were seen upon combination of the morphological and the molecular data, relative to the analysis of the combined molecular data. One branch from the molecular analysis experienced 37% more support in the analysis of the combined morphological and molecular data (partition "l"; Fig. 3.6B, Appendix D). One branch that was well-supported in the morphology-based trees and the combined analysis but not the analysis of the molecular data, was a resolution of a polyclotomy seen at the base of Pontederia in the latter case (partition "w"; Fig. 3.6B, Appendix D). Several branches in Heteranthera that were supported by more than 50% of bootstrap replicates in the analysis of the molecular data had less than 50% bootstrap support in the analysis of the combined molecular and morphological data. In terms of their constituent taxa, some of these branches were also contradictory to three branches from the morphological or combined morphological and molecular analyses that had 50% or greater bootstrap support; partitions "s," and "p" from the molecular analysis conflicted with partitions "z," "aa," and "ab" from the other two analyses (Fig. 3.6B, Appendix D). With the exception of partition "s," all of these branches had only moderate (50% to 70%) bootstrap support, and two of these five partitions were supported by the bootstrap analysis, but were not retained in the corresponding strict consensus trees (partitions "p" and "z").
Optimal and sub-optimal trees -- When the three uncombined data sets were mapped onto other chloroplast-based trees, the resulting excess tree length was generally less when the sub-optimal trees were originally derived from the larger or more combined data sets (Fig. 3.7A). There was only one case (when the restriction-site data set was mapped onto the rbcL-based trees) where the mean excess tree length was greater than 5% of native tree length. When the restriction-site data set or the rbcL data set were mapped onto trees derived from the various combined chloroplast data sets, there were several cases where these non-native trees were as short as the native trees (Fig. 3.7A; see also Table 3.3). Two specific examples of this are when either data set was mapped onto trees derived from the fully combined chloroplast data. When the ndhF data set was mapped onto this same set of trees, the most suboptimal tree was only around 2% longer, and the least sub-optimal tree was less than 1% longer than the native trees for this data set (Fig. 3.7A).

When the combined morphological and molecular data set was mapped onto trees derived from the combined chloroplast data, the trees were almost as parsimonious as the native trees of the former data sets (Fig. 3.7B). This was also the case when the fully combined chloroplast data set was optimized onto trees derived from the combined morphological and molecular data. The penalty in parsimony paid when optimizing either combined data set onto trees derived from the uncombined morphological data, or the morphological data onto trees derived from either combined data set, was large. In all four comparisons, these sub-optimal trees were around 15-20% longer than the native trees (Fig. 3.7B).

By examining the intersection of all pairs of sets of unrooted trees from the maximum parsimony analyses, an option available when inputting trees in PAUP, I found that several of the data sets produced some shortest trees that were topologically identical (Table 3.3, and see Fig. 3.8). One of these shortest trees was employed to illustrate branch lengths in Fig. 3.1C and 3.3. This tree is actually the single common tree in the intersection of all of the following sets of most-parsimonious trees: trees from the restriction-site data, from the combined ndhF
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(9)</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[1+1] 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>[2+1] 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data-set (no. of data)</th>
<th>(12) data</th>
<th>(5) data</th>
<th>(4) data</th>
<th>(8) data</th>
<th>(9) data</th>
<th>(10) data</th>
<th>(48) data</th>
<th>(48) data</th>
</tr>
</thead>
<tbody>
<tr>
<td>All data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restriction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fully</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 3.2: The topologies common among pairwise comparisons of the shortest unordered trees from analyses of the individual and combined data-sets.
Fig. 3.7. Excess length to the shortest trees found with various data sets from Pontederiaceae. The plot shows the mean excess number of steps required to map each individual data set onto unrooted topologies found from parsimony analyses of other data sets. The excess length is expressed as a percentage of the length of the native trees of the former data sets. Error bars represent the range of excess tree lengths for that data set/tree-set combination. The arrows indicate the native trees for that data set. (A) Excess length for the three uncombined chloroplast data sets (i.e., the ndhF data alone, the rbcL data alone and the restriction-site data alone) when mapped onto seven sets of chloroplast-based trees. (B) Excess length for the fully combined chloroplast data set, the morphological data set and all the data combined when mapped onto trees derived from each of these three data sets.
Excess length to sub-optimal trees

Data sets

% Excess length to sub-optimal trees

Restriction site data alone
Tel data alone
mCherry data alone

Tree sets

All data

A
Excess length to sub-optimal trees

Data-Set:

- All data combined
- Morphological data alone
- All chlorophyll data combined

Data sets:

- Shortest tree sets
- Longest data sets

% Excess length to sub-optimal trees
and restriction-site data set, from the combined \textit{rbcL} and restriction-site data set, and from the fully combined chloroplast data set. Another single topology is common to trees derived from three combined data sets (the fully combined chloroplast data set, the combined \textit{ndhF} and restriction-site data set, and the combined molecular and morphological data set), and was used to illustrate branch lengths in Fig. 3.4B.

When mapping each of the nine data sets and data-set combinations onto tree-sets obtained from the other eight cases, several instances were observed where additional non-native tree topologies were as short as the native trees (Table 3.3; values in square brackets). For the uncombined \textit{rbcL} data, a few non-native topologies derived from several of the combined chloroplast data sets were as parsimonious as its native trees. For the combined \textit{ndhF} and \textit{rbcL} data set, one tree derived from the combined \textit{rbcL} and restriction-site data set was as parsimonious as its four native trees. However, all of these "additional" topologies were found to contain some branches with a maximum possible length of zero. When such branches were collapsed using the "Condense trees" option in PAUP, the resulting topologies were identical to trees from the native set of trees.

\textit{Character incongruence} -- Mean bootstrap support of the molecular strict consensus by the morphological data was 37\%, quite close to the bootstrap support by this data set for its own strict consensus tree (52\%). In contrast, the mean support of the morphology-based strict consensus tree by the molecular data was 29\%, considerably less than for its own strict consensus (90\%). Incongruence indices for the molecular and morphological data were 6\% for I_{MF} and 34\% for I_{M}. With respect to the two definitions of total character incongruence, these fractions of the total incongruence were thus due to between data-set character incongruence.

\textit{Templeton's test for overall incongruence among data sets and their shortest unrooted trees} -- Only fully bifurcated trees can be compared for significant incongruence using Templeton's test
in the current implementation of PHYLIP (Felsenstein 1995). Since all \textit{rbcL}-derived trees and \textit{ndhF}-derived trees contained one to three minor polychotomies, trees from these data sets could not be compared to other trees using this test. However, one of the four shortest trees found with the combined \textit{ndhF} and \textit{rbcL} data set was fully bifurcated (the other three trees contained either one or two trichotomies). This tree was compared to the ten shortest trees found with the restriction-site data (Table 3.4) and was not found to be significantly different from them, at least with respect to the restriction-site data. Using the combined sequence data, eight of ten restriction-site trees were significantly longer than the combined-sequence tree. The other two restriction-site trees were not significantly longer (Table 3.4). Templeton's test indicated that all trees derived from the uncombined morphological data were significantly different to trees derived from the fully combined chloroplast data, with respect to either data set (Table 3.5).

\textit{Phylogenetic status of the selfing species of Eichhornia} -- The plausibility of the hypothesized monophyly of selfing species of \textit{Eichhornia} (Eckenwalder and Barrett 1986) was assessed using Templeton's test. Constrained trees were compared to unconstrained trees with regard to the morphological data and to the combined molecular data, respectively. Only fully-bifurcated trees could be assessed in these tests, but the partially bifurcated trees were identical in length to the bifurcated trees considered below. The first constraint-set unites four of the selfing species of \textit{Eichhornia} - \textit{E. diversifolia}, \textit{E. heterosperma}, \textit{E. paradoxa}, \textit{Eichhornia} sp. Five of the 17 shortest constrained trees found with the morphological data were fully bifurcated, and two of the five shortest unconstrained trees were fully bifurcated. The constrained trees compared in Templeton's test were 3.33\% longer than the corresponding unconstrained trees for this data set (i.e., four steps longer) and the SD for each constrained tree was 3.50 steps. This moderate increase in tree length was not significant. For the molecular data, two shortest constrained trees were found. Both they and the four shortest unconstrained trees were fully bifurcated.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>6.98</td>
<td>3.47</td>
<td>Yes</td>
<td>3.18</td>
<td>Yes</td>
<td>3.17</td>
<td>Yes</td>
<td>3.16</td>
<td>Yes</td>
<td>3.15</td>
<td>Yes</td>
<td>3.14</td>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
<td>6.90</td>
<td>2.97</td>
<td>No</td>
<td>3.12</td>
<td>No</td>
<td>3.17</td>
<td>No</td>
<td>3.16</td>
<td>No</td>
<td>3.15</td>
<td>No</td>
<td>3.14</td>
<td>No</td>
</tr>
</tbody>
</table>

**Resolution still-nee** refers to the number of clear neonates, and the number of clear neonates from the

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>6.98</td>
<td>3.47</td>
<td>Yes</td>
<td>3.18</td>
<td>Yes</td>
<td>3.17</td>
<td>Yes</td>
<td>3.16</td>
<td>Yes</td>
<td>3.15</td>
<td>Yes</td>
<td>3.14</td>
</tr>
<tr>
<td>No</td>
<td>6.90</td>
<td>2.97</td>
<td>No</td>
<td>3.12</td>
<td>No</td>
<td>3.17</td>
<td>No</td>
<td>3.16</td>
<td>No</td>
<td>3.15</td>
<td>No</td>
<td>3.14</td>
</tr>
</tbody>
</table>

**Resolution still-nee** refers to the number of clear neonates, and the number of clear neonates from the

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>6.98</td>
<td>3.47</td>
<td>Yes</td>
<td>3.18</td>
<td>Yes</td>
<td>3.17</td>
<td>Yes</td>
<td>3.16</td>
<td>Yes</td>
<td>3.15</td>
<td>Yes</td>
<td>3.14</td>
</tr>
<tr>
<td>No</td>
<td>6.90</td>
<td>2.97</td>
<td>No</td>
<td>3.12</td>
<td>No</td>
<td>3.17</td>
<td>No</td>
<td>3.16</td>
<td>No</td>
<td>3.15</td>
<td>No</td>
<td>3.14</td>
</tr>
</tbody>
</table>

**Resolution still-nee** refers to the number of clear neonates, and the number of clear neonates from the

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>6.98</td>
<td>3.47</td>
<td>Yes</td>
<td>3.18</td>
<td>Yes</td>
<td>3.17</td>
<td>Yes</td>
<td>3.16</td>
<td>Yes</td>
<td>3.15</td>
<td>Yes</td>
<td>3.14</td>
</tr>
<tr>
<td>No</td>
<td>6.90</td>
<td>2.97</td>
<td>No</td>
<td>3.12</td>
<td>No</td>
<td>3.17</td>
<td>No</td>
<td>3.16</td>
<td>No</td>
<td>3.15</td>
<td>No</td>
<td>3.14</td>
</tr>
</tbody>
</table>

**Resolution still-nee** refers to the number of clear neonates, and the number of clear neonates from the

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>6.98</td>
<td>3.47</td>
<td>Yes</td>
<td>3.18</td>
<td>Yes</td>
<td>3.17</td>
<td>Yes</td>
<td>3.16</td>
<td>Yes</td>
<td>3.15</td>
<td>Yes</td>
<td>3.14</td>
</tr>
<tr>
<td>No</td>
<td>6.90</td>
<td>2.97</td>
<td>No</td>
<td>3.12</td>
<td>No</td>
<td>3.17</td>
<td>No</td>
<td>3.16</td>
<td>No</td>
<td>3.15</td>
<td>No</td>
<td>3.14</td>
</tr>
</tbody>
</table>

**Resolution still-nee** refers to the number of clear neonates, and the number of clear neonates from the

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>6.98</td>
<td>3.47</td>
<td>Yes</td>
<td>3.18</td>
<td>Yes</td>
<td>3.17</td>
<td>Yes</td>
<td>3.16</td>
<td>Yes</td>
<td>3.15</td>
<td>Yes</td>
<td>3.14</td>
</tr>
<tr>
<td>No</td>
<td>6.90</td>
<td>2.97</td>
<td>No</td>
<td>3.12</td>
<td>No</td>
<td>3.17</td>
<td>No</td>
<td>3.16</td>
<td>No</td>
<td>3.15</td>
<td>No</td>
<td>3.14</td>
</tr>
</tbody>
</table>

**Resolution still-nee** refers to the number of clear neonates, and the number of clear neonates from the

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>6.98</td>
<td>3.47</td>
<td>Yes</td>
<td>3.18</td>
<td>Yes</td>
<td>3.17</td>
<td>Yes</td>
<td>3.16</td>
<td>Yes</td>
<td>3.15</td>
<td>Yes</td>
<td>3.14</td>
</tr>
<tr>
<td>No</td>
<td>6.90</td>
<td>2.97</td>
<td>No</td>
<td>3.12</td>
<td>No</td>
<td>3.17</td>
<td>No</td>
<td>3.16</td>
<td>No</td>
<td>3.15</td>
<td>No</td>
<td>3.14</td>
</tr>
</tbody>
</table>

**Resolution still-nee** refers to the number of clear neonates, and the number of clear neonates from the

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>6.98</td>
<td>3.47</td>
<td>Yes</td>
<td>3.18</td>
<td>Yes</td>
<td>3.17</td>
<td>Yes</td>
<td>3.16</td>
<td>Yes</td>
<td>3.15</td>
<td>Yes</td>
<td>3.14</td>
</tr>
<tr>
<td>No</td>
<td>6.90</td>
<td>2.97</td>
<td>No</td>
<td>3.12</td>
<td>No</td>
<td>3.17</td>
<td>No</td>
<td>3.16</td>
<td>No</td>
<td>3.15</td>
<td>No</td>
<td>3.14</td>
</tr>
</tbody>
</table>

**Resolution still-nee** refers to the number of clear neonates, and the number of clear neonates from the
### Table 3.5: Templeton's test (Templeton 1983; Relethford 1988, 1995) applied to assess confidence between two fully biometrically

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully combined or fec-4</td>
<td>6.92</td>
<td>7.00</td>
<td>7.14</td>
<td>138</td>
</tr>
<tr>
<td>Fully combined or fec-3</td>
<td>6.9</td>
<td>7.21</td>
<td>8.9</td>
<td>111</td>
</tr>
<tr>
<td>Fully combined or fec-2</td>
<td>6.9</td>
<td>7.21</td>
<td>8.9</td>
<td>111</td>
</tr>
<tr>
<td>Fully combined or fec-1</td>
<td>6.9</td>
<td>7.21</td>
<td>8.9</td>
<td>111</td>
</tr>
</tbody>
</table>

Chloroplasts (cp) data with respect to the two underlying data sets:

1. Two unphosphorized characters in the morphological data were coded as missing data for this test (sec text).
The constrained trees were 10.1% longer than the unconstrained trees (i.e., 61 steps longer) and the SDs for the constrained tree were 8.55 and 8.78 steps respectively. This increase in tree length was significant.

The second set of searches were more highly constrained in that they united all of the selfing species of *Eichhornia*. For the morphological data, 100 shortest constrained trees were found, 40 of which were fully bifurcated. The constrained trees compared in Templeton's test were 6.6% longer than the corresponding unconstrained trees for this data set (i.e., eight steps longer) and the SDs for the constrained trees ranged from 4.29 to 5.16 steps. The difference in tree length was again not significant. For the molecular data, two shortest constrained trees were found, both of which were fully bifurcated. The constrained trees were 11.8% longer than the unconstrained trees (i.e., 71 steps longer) and the SD for each constrained tree was 9.11 steps. This increase in tree length was significant for the molecular data.

**Taxonomic congruence** -- A neighbor-joining tree summarizing the similarity among all most-parsimonious unrooted chloroplast-based trees is presented in Fig. 3.8A. The scale-bar in Fig. 3.8A is the minimum (uncorrected) partition distance between two fully bifurcated, non-identical trees (i.e., 2 symmetric differences). The summary tree demonstrates the high degree of topological similarity among all of the chloroplast-based trees. Individual trees found with the *rbcL* data alone, the *ndhF* data alone and these two data sets combined, formed three distinct clusters on the summary tree, whereas for the other cases there were often topologies from other data sets that were more similar than other shortest trees from the same data set, so that all of these trees became intermingled on the summary tree. This phenogram also indicates some cases where different data sets yielded the same most-parsimonious topologies. In one instance four different data sets produced the same single topology (see above). Interestingly, one of the two restriction-site trees found not to be significantly different from the tree derived from the
Fig. 3.8. Neighbor-joining phenograms summarizing dissimilarity in tree-shape as measured by the partition metric (number of symmetric differences). The phenograms are midpoint rooted for compactness. (A) A neighbor-joining tree of all shortest unrooted trees from analyses involving each single, and all possible two-way and three-way combinations of the three chloroplast data sets. *ndhF*- and *rbcL*-derived trees group into distinct clusters and are highlighted in labeled boxes. Other most-parsimonious trees are individually labeled according to the data set or data set combination they were derived from. The open arrow indicates the single fully bifurcated tree from the combined *ndhF* and *rbcL* data that was employed in Table 3.4, the shaded and black arrows indicate the closest and furthest restriction-site trees to this tree, respectively. (B) A neighbor-joining tree of all shortest unrooted trees from analyses involving the combined chloroplast data, the morphological data and all the data combined. Individual trees are labeled according to the data set they were derived from.
A

Symmetric Differences

Restriction-site data

ndhF & CL data combined

ndhF & RS. data combined

dicL & RS. data combined

All cp data combined

2 Symmetric Differences

ndhF alone

rbcL alone

Restriction-site data

ndhF & rbcL data combined

ndhF & R.S. data combined

rbcL & R.S. data combined

All cp data combined
B

- × All cp data combined
- □ All data combined
- ● Morphological data

5 Symmetric differences
combined *ndhF* and *rbcL* data set (see above) is this highly converged-upon tree, a tree that is also topologically the closest of the ten restriction-site trees to the combined-sequence tree (Table 3.4; Fig. 3.8A).

Trees derived from the combined molecular data set and the combined chloroplast and morphological data are all very similar to each other, and in fact one tree topology is common to both data sets (Fig. 3.8B, and see Table 3.3, Fig. 3.7B). The morphology-based trees are the most topologically distinctive entities found in any analyses, and a large distance (just under 22 symmetric differences) separates them from the other trees considered in Fig. 3.8B. The morphological trees are still much more similar to the molecular trees than would be expected by chance (Fig. 3.9). The distance between the observed trees and random trees in this figure is probably somewhat exaggerated, since all of the random trees were fully bifurcated and some of the observed ones had one to three minor polychotomies. However, the unrooted trees from all the individual and combined data sets considered in this study were still well into the tail of the distribution of the possible partition distances for 24-taxon trees.

The summary phenogram portrays relationships among different tree topologies in a hierarchical manner. The validity of using a hierarchical approach for summarising tree dissimilarity was assessed by deriving a cophenetic correlation coefficient ($r_{CS}$) between the raw tree-to-tree distances and the distances implied by each phenogram. A high correlation coefficient provides evidence that these distances are indeed hierarchical (see Rohlf and Fisher 1968). For the phenogram in Fig. 3.8A, $r_{CS}$ was 0.940 ($p < 0.01$; approximate Mantel t-test). For the phenogram in Fig. 3.8B, $r_{CS}$ was 0.993 ($p < 0.01$). A large degree of the hierarchy present in this phenogram results from a single very long branch separating the morphological trees from the other trees, since with regard to the morphology-based trees alone, $r_{CS}$ was 0.394 ($p > 0.10$). However, the low degree of hierarchy among the morphological trees in the phenogram is also suggested by the phenogram itself, since several branches in this part of the
Fig. 3.9. Distribution of tree-to-tree distances (measured as the number of symmetric differences) for 22-taxon trees, estimated using random trees generated in MacClade version 3.0 (Maddison & Maddison, 1992). The portions of the distribution occupied by tree-to-tree distances used to derive the neighbor-joining trees in Fig. 3.8A and 3.8B are bracketed. The length of the branch in the neighbor-joining tree in Fig. 3.8B that separates morphological trees from the other trees is indicated with an asterisk.
Trees compared in Fig. 3.8A:

Trees compared in Fig. 3.8B:

1,000 Random trees

Tree-tree distances
(no. of symmetric differences)
tree have nearly zero length. For the remaining trees in Fig. 3.8B, \( r_{CS} \) was 0.809 (\( p < 0.01 \)). These lines of evidence therefore suggest that it is valid to use this phenetic approach for deriving summaries of taxonomic (in)congruence among trees.

**DISCUSSION**

These analyses indicate that there is considerable agreement among the three chloroplast-based data sets from Pontederiaceae. The striking similarity of the individual and combined chloroplast-based trees (Figs. 3.1-3.4, 3.8A), the high level of resolution and support of these trees (Table 3.2, Fig. 3.6A), the relative ease with which sub-optimal trees corresponding to topologies derived from other chloroplast data sets could be found (Fig. 3.7A), and the proximity and intermingling of the shortest trees from different chloroplast data sets on the summary phenogram (Fig. 3.8A) strongly suggest that the chloroplast data are converging towards the true chloroplast tree of Pontederiaceae. That topological disagreements among chloroplast data sets occurred with respect to only moderately supported taxon partitions (Fig. 3.6A) is also consistent with this view.

When there are trees derived from one data set that are not significantly different from one or more trees derived from a different data set, this is evidence for a lack of incongruence between the two data sets. When I used Templeton's test as a measure of congruence, I repeated it with respect to both data sets. The trees derived from the restriction-site data were not significantly different from the tree derived from the combined sequence data set, at least given the restriction-site data (Table 3.4). When the combined sequence data (i.e., data from the \( ndhF \) and \( rbcL \) genes combined) was considered in this test, there were two of ten restriction-site trees found not to be significantly different from this combined-sequence tree. All of the restriction-site based trees were on the border of being significantly different from this
sequence-based tree. These tests therefore provided overall support for the chloroplast data sets being combinable with each other, but also suggested that there may be weak elements of incongruence between the restriction-site data and the combined sequence data.

Unfortunately, trees obtained from the two individual sequence data sets could not be compared to other trees using Templeton's test because none of them were fully bifurcated. However, there were several trees among those derived from the various combinations of the chloroplast data sets, upon which the \textit{rbcL} data could be mapped with no decrease in parsimony (Table 3.3; Fig. 3.7A). These non-native trees are not "statistically different" (in the sense of Templeton's test) from the native \textit{rbcL} trees, at least given the \textit{rbcL} data. The most topologically distinctive trees produced by an individual chloroplast data set were those derived from the \textit{ndhF} locus (Fig. 3.8B), but it was in general as easy to map this data set onto other chloroplast-derived trees, as it was to map the \textit{rbcL} data onto non-native trees (Fig. 3.7A). Thus, while there may be limited elements of incongruence among the three individual chloroplast data sets, these lines of evidence suggest that they are at most quite weak.

The restriction-site data set had the highest amount of homoplasy of any data set examined in this study (Table 3.2). Trees inferred from this data set are well supported and congruent with those from the other chloroplast data sets, so the relatively low homoplasy indices associated with this data set are not associated with a poorer ability to reconstruct phylogenetic history. More homoplasious classes of characters do not necessarily infer phylogeny more poorly; the amount of homoplasy in a data set is not necessarily a measure of the quality of the data (Sanderson and Donoghue 1989).

Systematic implications of the chloroplast data -- These have been summarized elsewhere with regard to the restriction-site evidence (see Appendix A). Apart from general strengthening in taxon support, the systematic conclusions in that study apply to the fully combined chloroplast data. It seems unlikely that the chloroplast-based phylogeny will change substantially with
addition of further taxa, since all of the currently missing species belong taxonomically within the three well-supported genera (*Monochoria*, *Heteranthera* and *Pontederia*). A few elements of phylogenetic relationship remain poorly resolved with regard to the fully combined chloroplast evidence (Fig. 3.4). The most basal portions of *Pontederia* and *Heteranthera s.l.* are still unresolved. The branching order within the clade consisting of *Monochoria hastata*, *Monochoria korsakovii* and *Monochoria vaginalis* is also ambiguous. It seems probable that the distorting effects of sampling errors (or weak process partitions) become more important in the more poorly supported regions of trees, since most of the conflicting or ambiguous regions among the chloroplast-based trees are in regions with relatively short internal branches (Figs. 3.1-3.4). Presumably, speciations occurred in fairly rapid succession in these parts of the trees. The position of *Eichhornia crassipes* within the family has important implications for reconstructions of breeding-system evolution in the family (see Appendix A). Two classes of topology were found in the restriction-site study (Appendix A) with distinct positions of this taxon relative to the position of *Monochoria* in trees derived from the restriction-site data. Their "topology class A" is the sole topological class found with the fully combined chloroplast evidence, and bootstrap support for this topological configuration reached 75% in this analysis (the branch corresponding to partition "j" in Fig. 3.6A; see also Fig. 3.3 and Appendix D).

The single major unresolved issue from the chloroplast evidence is the location of the root of the family. When Philydraceae was included in the searches as an outgroup, there was a loss of support and resolution for phylogenetic structure around the root of Pontederiaceae (e.g., Fig. 3.3). The drop in bootstrap support around the root suggests that the root position is subject to problems of long-branch attraction (Trueman 1995). The outgroup branch is by far the longest branch in the rooted trees. Uncertainty in root placement because of long outgroup branches may be a more widespread problem in phylogenetic studies than has been generally appreciated. This issue is explored in more detail in Chapter 4.
The morphological data is swamped by the molecular data — The morphology-based trees are well resolved, but are not well supported by bootstrap analysis. Furthermore, the trees resulting from the combination of all the morphological and molecular data are virtually (and in one case actually) identical to those derived from the combined molecular evidence alone (Table 3.3, Figs. 3, 4, 8b), they have very similar levels of resolution and support (Table 3.2, Fig. 3.6B) and it is extremely easy to find the sub-optimal trees from the combined molecular data that correspond to optimal trees from the combined molecular and morphological data, and vice versa (Fig. 3.7B). When the morphological and molecular sources of data are combined, the former is thus overwhelmed by the latter. This finding contrasts with other workers' experiences in combining molecular and morphological data (e.g., DeSalle et al. 1992; Donoghue and Sanderson 1992; Eernisse and Kluge 1993, Wheeler, Cartwright and Hayashi 1993; Chavarria and Carpenter 1994), where such swamping was not considered to be a major problem. This "swamping" is a further indication that there is insufficient evidence for deriving robust phylogenies of the family using the currently available morphological data. An examination of Table 3.2 illustrates a potential source of the problem: there are far fewer informative morphological characters per taxon than molecular characters. Indeed, there was a significant correlation \( r = 0.823 \) between the mean bootstrap support and the number of informative characters for a data set, across the nine data sets and data set combinations considered in this study.

Conflict and agreement among molecules and morphology — The various incongruence indices provide an estimate of the degree of character conflict between molecules and morphology. From 6% \( (I_{MF}) \) to 34% \( (I_M) \) of total incongruence can be ascribed to between-data-set character incongruence, depending on which measure of total character incongruence is employed. These values are comparable to those found in other studies, although \( I_{MF} \) is at the upper end of values recorded from other groups (Omland 1994). For both measures, the sum of within data-set
incongruence was greater than that due to the sum of incongruence between the molecular and morphological data sets. It was difficult to find sub-optimal trees with the molecular data that corresponded topologically to those found by the morphological data (Fig. 3.7B), and it was proportionately just as hard to optimize the morphological data onto the molecular trees (Fig. 3.7B), despite the poor support by the morphological data for its own shortest trees. This may in part be an artifact of normalizing the excess tree length by the total evolution accounted for by the data set on its native trees, since shortest tree length is proportionately much less for the morphological trees as a consequence of the small number of characters in this data set.

A new measure of incongruence lends some credence to this hypothesis. The mean bootstrap support for taxon partitions in the molecular strict consensus by the morphological data was only 15% less than that observed for its own trees (52%). The morphological data thus provides poor support for most of the topological structure seen in the trees derived from it, but only moderately poorer support for structure in non-native trees derived from the molecular evidence. The overall lack of bootstrap support by the molecular data for the topological structure in the morphological trees was much poorer. The molecular data strongly supported its own trees, and strongly disagreed with phylogenetic relationships suggested by the morphological evidence. The mean bootstrap support by the combined molecular data for the morphological strict consensus tree was 61% less than for its own strict consensus (90%). With regard to elements of the morphology-based tree with more than 50% bootstrap support, there was actually more agreement than disagreement with the chloroplast-based trees (Fig. 3.6B). These points of agreement include several of the taxonomically more interesting taxon partitions in the family, namely those linking taxa within three of the four major genera in the family (Monochoria, partition "i"; Pontederia s.l., partition "c"; and Heteranthera s.l., partition "o"; see Appendix D). The bootstrap analysis also suggested several elements of conflict between the morphological and molecular data, but these involved branches that were not particularly well supported by either data set (see Results).
Sources of incongruence among molecular and morphological trees -- Despite the fact that the morphological data is overwhelmed by the molecular data when combined with it, and provides poor support for its own trees (and the molecular trees), there was support from Templeton’s test for significant incongruence between the morphological and molecular trees, given either set of data (Table 3.5). While trees from these two major sources of evidence share substantial components of structure (see above), this indicates that there is significant disagreement between them over one or more aspects of phylogenetic history. Templeton’s test gives an overall measure of whether or not two data sets are incongruent, but unless an a priori prediction can be made concerning the effect of particular process partitions on tree topology, it can not readily indicate with which characters or taxa the incongruence is associated. Another statistical test that assesses whether data sets are incongruent was described by Farris et al. (1994), but it also does not pinpoint where incongruence lies.

Rodrigo et al. (1993) describe a test for identifying sources of incongruence that are limited to certain clades. Their procedure involves sequential pruning of apparently conflicting taxa from each data set, followed by comparisons of the shape of the trees resulting from each pruned data-matrix to null distributions of tree-to-tree distances derived from bootstrapped trees from each data source. This test is most practical when tree conflicts are limited. A simpler method for locating taxonomic sources of incongruence is to identify well-supported clades (or taxon partitions) in trees derived from one data set that conflict with well-supported clades in trees from another data set. However, none of the taxon-partitions in the morphological data that conflict with partitions in the molecular data are very well supported by the morphological data (see above). The current morphology-based estimates are therefore not sturdy enough to determine the source(s) of conflict between the two major types of available evidence. An expanded morphological data-base for this family is needed to address this issue.
CHAPTER 3. PHYLOGENETIC CONGRUENCE IN PONTEDERIACEAE

The significant incongruence among the morphological and chloroplast data indicates that discordance among these sources of evidence is not simply because there were too few characters in the former data set to allow accurate phylogenetic reconstruction. One possible source of this incongruence is the action of various genealogical phenomena on the linked chloroplast characters. Whether or not the chloroplast tree corresponds exactly to the true "species tree" of Pontederiaceae depends largely on the extent to which lineage-sorting or hybridization events have played a role in the family's history. Although there is little evidence of hybridization among extant members of Pontederiaceae (see Appendix A), some taxa in the family are polyploids (reviewed in Eckenwalder and Barrett 1986). It is not known whether these represent auto- or allopolyploids, but their existence raises the possibility that hybridization events may have played a role in the family's history. However, it seems unlikely that lineage sorting of ancestral polymorphisms has led to major distortions in the inference of the phylogenetic history of the family. This is because fixation times of novel mutations in the chloroplast genome are likely to be short relative to the time between successive speciation events, a consequence of the small effective population sizes of these haploid genomes (Moore 1995). There may be other, non-genealogical sources of conflict among the chloroplast and morphological data. Different evolutionary rules operating in either data set could also lead to such conflict. Although some process partitions are known to operate within and between the chloroplast data sets, these did not interfere with the combinability of various chloroplast sources of data, at least with each other.

The credibility of one process partition in the morphological data set, that of a selfing syndrome, was assessed by constraining selfing taxa within the family to form a taxon partition during phylogenetic inference. A monophyletic group of selfing taxa of Eichhornia was found in some most-parsimonious cladograms in a morphology-based analysis of this family by Eckenwalder and Barrett (1986). This led these workers to suggest that repeated homoplasious shifts in multiple morphological characters could corrupt phylogenetic inference (see also
Chapter 3. Phylogenetic Congruence in Pontederiaceae

Chapter 1, Appendix A). Unlike Eckenwalder and Barrett's study, the shortest morphology-based trees did not link the selfing taxa of *Eichhornia*. In part this may be because my analysis considered a different subset of taxa within the family (including two extra selfing taxa of *Eichhornia*). The difference may also be a consequence of revisions to the morphological data set (see Appendix C), or because I treated all character-state transformations as unordered, whereas quantitative characters were ordered in Eckenwalder and Barrett's study. In addition, Eckenwalder and Barrett used several auxilliary criteria to choose among most-parsimonious trees that I do not consider here. In any case, some of the most-parsimonious trees that Eckenwalder and Barrett found also did not portray monophyly of the selfing taxa of *Eichhornia* (Eckenwalder and Barrett, Fig. 2). In the current study, trees supporting the linkage of these selfing taxa were 3% to 7% longer than the shortest morphology-based trees and 10% to 12% longer than the shortest trees derived from the fully combined chloroplast data, depending on the stringency of the constraint criterion employed. The morphological evidence thus does not support this scenario and the molecular evidence strongly contradicts it.

Summary — The three chloroplast data sets are substantially in agreement. Upon combination of these data sets, bootstrap support increased for most branches. This increase in agreement may occur because phylogenetic “noise” associated with sampling error is reduced with more data points (de Queiroz, Donoghue and Sober 1995), but combination may also override the distorting effects on phylogenetic reconstruction of weak process partitions in individual data sets. While there are known differences in the patterns of change among subsets of the chloroplast characters, these did not appear to substantially interfere with the convergence of trees derived from these data sets, even in the absence of differential weighting schemes for these processes (cf., Miyamoto et al. 1994). However, it is still possible that these same process partitions may distort the reconstruction of phylogenetic relationships among less closely related organisms than those examined in this study.
CHAPTER 3. PHYLOGENETIC CONGRUENCE IN PONTEDERIACEAE

The morphological data set is swamped when combined with the molecular data, and is too small to derive well-supported phylogenies of its own. Nonetheless, the morphology-based evidence has significantly detectable elements of agreement and disagreement with the molecular data. This partial discordance is probably not simply a consequence of sampling error associated with the small size of the morphological data, which indicates that there may have been different evolutionary processes or genealogical histories associated with the chloroplast and morphological sources of data. It is clear that new morphological data, and perhaps molecular evidence from the nuclear genome, are needed to address this possibility. However, an hypothesized process partition involving the morphological data, a selfing syndrome, is incompatible with all available sources of data for the family.
CHAPTER 4

- The Local Position of Pontederiaceae in the Monocotyledons and its Root Location -
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

INTRODUCTION

The root of a monophyletic group is the ancestral species from which all other species in that group are descended. Of all the aspects of a group’s phylogenetic history, it is perhaps the most important, for without it, it is not usually possible to polarize reconstructions of the evolutionary events occurring within the group. The assumption that the root of a group is known with any degree of certainty can be a precarious one to make. Phylogenetic trees from studies involving molecular data are typically rooted by including outgroups in the phylogenetic analysis. Following phylogenetic reconstruction, the root is the point of attachment of the outgroup on the ingroup sub-tree. Unfortunately, ample opportunities for artifactual rootings may exist in many phylogenetic studies because of the relatively long branches that tend to connect ingroups to their outgroup taxa. These long branches often accumulate multiple homoplasious events, which collectively can be highly misleading concerning phylogenetic relationships (Felsenstein 1978; Hendy and Penny 1989) and consequently the root’s location (Miyamoto and Boyle 1989; Wheeler 1990).

The outgroup used to root an ingroup need not be its sister group (e.g., Nixon and Carpenter 1993). However, there should be fewer problems of spurious attraction of long outgroup branches to relatively long branches in the ingroup, the more closely related the outgroup is to the ingroup (Wheeler 1990; Maddison, Ruvolo and Swofford 1992; Swofford et al. 1996). Careful choice of the outgroup taxon used to root a tree is therefore important. Smith (1994) pointed out that a good sampling of taxa in the sister-group will help break up long outgroup branches better than increased sampling of less closely related outgroups (see also Hendy and Penny 1989). However, it is not always obvious what group is the sister-group, and even where this is known with some certainty, the ingroup may still be quite distantly related to its extant sister-group.
Chloroplast-based evidence provides a clear picture of most aspects of the phylogenetic structure of Pontederiaceae, except for the position of its root (Chapter 3). While the monophyly of the family is strongly supported by chloroplast-based evidence, it is not clear to which groups it is most closely related. Current studies based on the chloroplast gene rbcL (e.g., Chase et al. 1993) and restriction-site variation in the chloroplast genome (Davis 1995) provide only weak indications of the local position of Pontederiaceae in the monocotyledons (see Chapter 1 and Duvall et al. 1993). The goal of this study is to clarify which taxa are most closely related to Pontederiaceae, and to use these taxa as outgroups to determine whether a definitive rooting can be obtained for the family.

The uncertain placement of Pontederiaceae may be partly a function of insufficient data. It may also be a function of the preponderance of relatively long terminal branches compared to short branches connecting the different families to each other (see e.g., Chase et al. 1993). I address the first problem by approximately doubling the number of phylogenetically informative characters, by employing a new source of data from the chloroplast genome (the ndhF gene) in tandem with the available rbcL evidence. I attempt to side-step the second issue by using various weighting schemes to correct for the "multiple hits" (homoplasious events) that, if undetected during phylogenetic estimation, can lead to long-branch attraction. The branches connecting Pontederiaceae to individual candidate outgroup taxa are known to be long (Chapter 3), and so I use two tests for assessing whether sub-optimal roots are significantly different from the best root. One of these is a novel application of Templeton's (1983) test for whether different phylogenetic trees are significantly different, and the other is an extension of Miyamoto and Boyle's (1989) method for comparing the performance of real and random outgroup sequences during tree rooting.
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICAE AND ITS ROOT

MATERIALS AND METHODS

DNA sequences were obtained from portions of two chloroplast genes (*ndhF* and *rbcL*) for representatives of families from five monocotyledon superorders (Arecales, Bromeliaceae, Commelinaceae, Liliaceae and Zingiberaceae *sensu* Dahlgren, Clifford and Yeo 1985; see Appendix E). These taxa were chosen to provide a good spread of families in the complex of superorders consisting of Arecales, Bromeliaceae, Commelinaceae, Zingiberaceae and several taxa outside this complex. There is some morphological and molecular evidence that Pontederiaceae is related to several taxa within this group of superorders, particularly Commelinaceae, Haemodoraceae and Philydraceae (reviewed in Chapter 2). 1343 bp of sequence were obtained from *rbcL*, representing a major portion of that gene, and 490 bp of DNA sequence were obtained from around the 3'-end of *ndhF*.

The *ndhF* and *rbcL* sequences for 24 taxa in Pontederiaceae were obtained for the analyses in Chapter 3. Most of the other *ndhF* sequences were obtained for this study, using primers and methods detailed in Appendix H. The remaining sequences were obtained directly from GenBank or were supplied by other workers. Collection details and a list of the sources of the remaining sequences are provided in Appendix E. In all cases, I attempted to obtain the two different chloroplast sequences from the same species, but in three cases (*Anigozanthos*, *Canna* and *Musa*) the *ndhF* and *rbcL* sequences were potentially from different species in the same genus, and in one case the two sequences were from different genera in the same family (*Barbienia* and *Vellozia* in Velloziaceae). The *Anigozanthos* and *Canna* specimens used to obtain *ndhF* sequences are cultivars of undetermined hybrid parentage. The *Musa* specimen used to obtain an *ndhF* sequence was from a pre-flowering individual of uncertain specific affinity.
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

Analyses

Length differences were not found in the region of rbcL examined, but several short indels (one-two codons long) were encountered in ndhF. One indel was observed in each of Canna sp. (Cannaceae), Caryota mitis (Arecaceae) andTradescantia zebrina (Commelinaceae), and two other indels were previously noted in Eichhornia meyeri and two varieties of Pontederia cordata (Pontederiaceae) (see Chapter 3). Nine nucleotides neighbouring the indel site in the ndhF sequence of Caryota mitis were scored as missing data because of uncertain homology to other sequences. In each case, the entire indel was coded as a single new character using the GAPMODE option in PAUP. However, apart from the indel shared by two varieties of Pontederia cordata, they were all uninformative (autapomorphic) in the cladistic sense.

Unless otherwise stated, all maximum-parsimony analyses (basic searches and bootstrap analyses) were performed using PAUP version 3.1.1 (Swofford 1993). A pre-release version of PAUP 4 (version 4.0d47) was used to test for incongruence between the two chloroplast genes for the twelve outgroup families and two members of Pontederiaceae using a test described by Farris et al. (1994), and to perform parsimony analyses using characters that were re-weighted according to their estimated site-likelihoods on particular trees (see below). The congruence test of Farris et al. (1994) is analogous to the Mann-Whitney U test (Lindgren 1962) for assessing whether two populations differ in location. It works by randomly re-partitioning data from each data set into new data-sets of equal size to the original ones. The shortest trees are found for each re-partitioned data-set, and if the sum of tree lengths for the two re-partitioned data-sets is greater than the tree length for the combined data in more than 95% of random re-partitionings, the two sources of data are deemed to be significantly incongruent. I used the non-parametric bootstrap (Felsenstein 1985a) as an index of the robustness of particular branches. Whether highly robust branches are good reflections of true phylogenetic structure depends on a number of factors (see Discussion).
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

Phylogenetic placement of Pontederiaceae in the monocotyledons. -- Three sets of analyses were performed to determine the position of Pontederiaceae among the twelve additional families considered in this study. One set of analyses involved only rbcL, one involved only ndhF, and a third involved a combined data set consisting of both genes. For each set of analyses, heuristic searches were performed to find the most parsimonious tree(s), with all character-state changes unordered and thus equally weighted (i.e., "Fitch optimization"; Fitch 1971), with all characters equally weighted, and using TBR (tree bisection-reconnection) branch swapping, with MULPARS and "Steepest descent" options activated, and with 100 random-addition replicates. Bootstrap analyses (100 replicates, one random addition sequence per replicate) were performed using the same heuristic search conditions under this weighting scheme.

Two types of a posteriori weighting were undertaken in an attempt to correct for multiple hits at each variable site. In the first of these I re-weighted each character according to its rescaled consistency index (RC, a product of the consistency index and the retention index; Farris 1989a,b) on the tree(s) found with all characters equally weighted, with maximum RC values used when there was more than one such tree (Swofford 1993). Multiple rounds of re-weighting can be performed, using new weights in each round that are determined using the previous most-parsimonious tree(s). This process can be repeated until the same tree(s) or character weighting scheme is converged upon (Farris 1969; Swofford 1993). For all re-weighted data-sets, bootstrap support for tree structure was also assessed under that weighting scheme.

I performed a second set of analyses using a posteriori character weights determined from individual site-likelihoods that were estimated using the most-parsimonious tree obtained with all characters equally weighted. Where there were tied trees under this optimization scheme, the one with the lowest -log_e likelihood score (as determined using the default
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

likelihood settings in PAUP 4) was used for determining site-likelihoods. Two likelihood models were used to determine site-likelihoods: the default maximum likelihood model in PAUP 4 and a more parameter-heavy likelihood model. This second model (estimated for the combined data set only) took account of several parameters which were all estimated from the data: the transition to transversion ratio, the number of invariant sites, and variation among sites in the substitution rate. The rate variation was estimated using a discrete approximation to the gamma distribution, with the gamma shape parameter estimated from the data, and with six different rate categories permitted, each category being represented by a mean rate. For the re-weighted data-sets, bootstrap support for tree structure was again assessed under each weighting scheme. The indels were excluded from likelihood assessments and from all searches that were performed using the site-likelihood estimates of character weights. A single round of re-weighting was performed for all three data sets.

Are sub-optimal rootings of Pontederiaceae significantly different from the most-parsimonious root(s)? -- The three families that were found to be most closely related to Pontederiaceae (see Results) were used to determine the most-parsimonious rooting of the family, and whether there were additional root placements that did not require significant increases in tree length. The phylogeny of Pontederiaceae assumed for these tests was one of four shortest trees found from an unrooted combined analysis of three highly congruent chloroplast-based data-sets (two sequence data sets and a restriction-site data set of the family), a single tree that was also found from three other single and combined analyses of these three data sets (Chapter 3).

This unrooted tree of Pontederiaceae was re-rooted with individual or composite outgroups at all 45 internal and terminal branches using MacClade 3.0 (Maddison and Maddison 1992), with some additional editing of NEXUS-format tree files in PAUP to prune particular outgroups. The penalty in parsimony for the sub-optimal rootings was determined with all
characters equally weighted, for different outgroups that were attached in turn on each branch of the unrooted tree (see Lundberg 1972; Miyamoto and Boyle 1989). This method of rooting is preferred when it is suspected (as here; Chapter 3) that the long branch between the ingroup and outgroup nodes may otherwise distort estimated phylogenetic relationships within the ingroup (Swafford et al. 1996).

Rooted tree lengths were determined in PAUP and also in DNAPARS from the PHYLIP 3.5c package of Felsenstein (1995) via its user-tree option. DNAPARS was also used to determine which trees (and hence root locations) were significantly different from the best tree (= root location) for the combined sequence data, using a test derived from Templeton (1983). Templeton’s original formulation of the test is a Wilcoxon ranked-sums test (Seigel 1956) that compares the scores of each character on the best tree and an alternative tree (see also Felsenstein 1985b; Larson 1994). Felsenstein’s (1995) version of the test is analogous to a maximum-likelihood test due to Kishino and Hasegawa (1989) that examines loge likelihood differences between trees. It uses the among-character variance of the difference between two trees to assess whether their tree lengths are significantly different. Sub-optimal root locations were also assessed using the restriction-site data-set of Pontederiaceae (Appendix A), for two outgroups (Philydrum and Tradescantia) which that study shares with this study. Restriction-site presence/absence character-states were re-coded as DNA nucleotides for assessment in DNAPARS (see Chapter 3).

I also examined the propensity for random outgroups to be “attracted” to particular ingroup branches in this tree of Pontederiaceae for the combined sequence data-set. One hundred random outgroups were constructed in MacClade, using base frequencies determined from the outgroup taxa shown in Fig. 4.1-4.3 for the combined sequence data set. No significant heterogeneity in base frequencies was observed across the fourteen taxa, according to a test available in PAUP 4 [Mean base composition = 29% A, 18% G, 22% C and 32% T;
The length of these random sequences was set at 397 nucleotides. This is the number of variable nucleotides observed between three real outgroup taxa (representatives from Haemodoraceae, Philydraceae and Commelinaceae) and all of the ingroup taxa. The ingroup was reduced to these 397 characters and tree files were constructed with each random outgroup sequence attached to the 45 ingroup branches. The optimal rooting for each random-outgroup sequence was determined by examining tree lengths for each possible rooting. Worse performance of real outgroups than random outgroups for the next-best rooting was used as evidence of historical signal in real outgroups by Miyamoto and Boyle (1989). I extend their method by looking at rootings less optimal than the next-best-to-optimal rooting. For each outgroup, the increase in parsimony beyond the most-parsimonious rooting was determined (as before) for all 45 root locations in PAUP, and the mean and standard deviation in this penalty were determined across ranked penalties. In cases where there were tied penalties for individual outgroup sequences), these were arbitrarily assigned ranks that spanned the number of ties (see Fig. 4.6).

RESULTS

The two genes provided similar numbers of variable characters that were potentially informative in the cladistic sense. For the taxa shown in Figs. 4.1-4.3, the rbcL locus provided 151 such characters, while the ndhF locus provided 140 of these characters. For conciseness, I refer to the Areconae-Bromelinae-Commelininae-Zingiberanae and Bromelinae-Commelininae-Zingiberanae groups of taxa as the "ABCZ" and "BCZ" complexes, respectively. I refer to the shortest trees found with all characters equally weighted as "equally weighted trees," those found after one round of re-weighting characters according to their maximum rescaled consistency indices on the equally weighted trees as "RC re-weighted trees,"
and those found after one round of re-weighting characters according to their site-likelihoods on the equally weighted trees as "site-likelihood re-weighted trees" (see also Materials and Methods). For all three data sets, only a single round of re-weighting using the rescaled consistency indices was required for convergence. Length and fit measures for the equally weighted trees are given on Figs. 4.1-4.3. Individual taxa are generally referred to by the families they belong to, and trees are arranged with the representatives from two Lilianean families in the order Asparagales sensu Dahlgren, Clifford and Yeo (1985) (Cyanastraceae and Amaryllidaceae) placed most basally.

Systematic relationships among the thirteen families. -- For the ndhF locus, one equally weighted tree was found (Fig. 4.1), which had the same topology as the RC re-weighted tree. The single site-likelihood re-weighted tree found for the ndhF data set (not shown) differed from the equally weighted tree only in the relative branching order within the clade consisting of taxa from Flagellariaceae, Bromeliaceae and Typhaceae. For rbcL, two new shortest equally weighted trees were found, one of which (Fig. 4.2) was also the single shortest RC re-weighted tree for this data set. The two trees had quite different topologies. The one shown in Fig. 4.2 contains a clade consisting of the taxa from Pontederiaceae, Commelinaceae, Philydraceae and Haemodoraceae. The second equally weighted tree for the rbcL data (not shown) differed from the first in that Velloziaceae was the sister-group of Pontederiaceae, and Commelinaceae was depicted as being the sister-group of Flagellariaceae. In this tree, a clade consisting of Haemodoraceae and Philydraceae was the sister-group of a clade containing Velloziaceae, Pontederiaceae and the Zingiberalean taxa. The single site-likelihood re-weighted tree for the rbcL data had a third distinctive topology (not shown), in which Commelinaceae was the sister-group of Pontederiaceae, Haemodoraceae was the sister group of both families, and Velloziaceae was the sister-group of Philydraceae. All three rbcL topologies thus had Velloziaceae disjunct from the other taxa in Lilianae and embedded within the BCZ complex.
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

Fig. 4.1. The phylogenetic position of Pontederiaceae in a local group of monocotyledons, based on the chloroplast gene ndhF. The tree is the single most-parsimonious one found with all characters equally weighted and was also the shortest tree found after characters were re-weighted according to their rescaled consistency indices (see text). It has a length of 561 steps, CI = 0.654, CI (excluding autapomorphies) = 0.526, and RI = 0.428. The root has been placed (a posteriori) such that Cyanastrum cordifolium and Narcissus elegans are the outgroups to the other taxa. The reconstructed number of mutations is indicated above each branch (ACCTRAN optimization). The bootstrap support for different branches was estimated using various weighting schemes and is indicated below each branch. Each value in square brackets is the bootstrap support under equally weighted parsimony. The two values in curly brackets are from analyses where characters were re-weighted according to their rescaled consistency indices (first value) or their site-likelihoods under a simple likelihood model (second value) using the equally weighted tree. The superorders that individual taxa belong to are labelled as follows: A = Arecales, B = Bromeliales, C = Commelinales, L = Liliales, Z = Zingiberales.
Flagellaria indica C
Ananas comosus B
Typha latifolia B
Anigozanthos sp. B
Tradescantia zebrina C
Caryota mitis A
Philydrum lanuginosum B
Pontederia sagittata B
Hydrothrix gardneri B
Canna sp. Z
Musa sp. Z
Barbacenia elegans B
Cyanastrum cordifolium L
Narcissus elegans L
Fig. 4.2. The phylogenetic position of Pontederiaceae in a local group of monocotyledons, based on the chloroplast gene \textit{rbcL}. The tree is one of two most-parsimonious ones found with all characters equally weighted and was also the shortest tree found after characters were re-weighted according to their rescaled consistency indices (see text). The two trees have a length of 558 steps, CI = 0.634, CI (excluding autapomorphies) = 0.496, and RI = 0.424. The root has been placed (\textit{a posteriori}) such that \textit{Cyanostromum cordifolium} and \textit{Narcissus elegans} are the outgroups to the other taxa. The reconstructed number of mutations is indicated above each branch (ACCTRAN optimization). The bootstrap support for different branches was estimated using various weighting schemes and is indicated below each branch. Each value in square brackets is the bootstrap support under equally weighted parsimony. The two values in curly brackets are from analyses where characters were re-weighted according to their rescaled consistency indices (first value) or their site-likelihoods under a simple likelihood model (second value) using the equally weighted tree with the lowest -log$_{e}$ likelihood. The superorders that particular taxa belong to are labelled as follows: A = Arecanae, B = Bromelianae, C = Commelinanae, L = Lilianae, Z = Zingiberanae.
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

For the combined data set, three new shortest equally weighted trees were found. These trees differed from each other only in the relative positions of Flagellariaceae, Bromeliaceae and Typhaceae, within a clade consisting of these three families (all three possible variants were seen). One of these had the same topology as the single RC re-weighted tree for the combined data (Fig. 4.3), and another equally weighted tree had the same topology (with Bromeliaceae sister to Typhaceae; not shown) as the only site-likelihood re-weighted tree found with the combined data.

The major differences between the ndhF-based trees and those found with the combined data set lay in whether or not Arecanae was the sister-group to the BCZ complex of taxa, and in what constituted the sister-group of Pontederiaceae (Haemodoraceae-Commelinaceae in the former case, Philydraceae in the latter case). The relatively low bootstrap values (under equally weighted parsimony and the re-weighted schemes) for several clades in the ndhF tree containing Arecanae (Fig. 4.1) suggest a lack of support by this data-set for this arrangement. Furthermore, this taxon is the sister-group of the BCZ complex in trees derived from the combined data, although the moderate bootstrap support for the latter clade under three of the four weighting schemes (Fig. 4.3) suggests that more sequence data (and taxa) are needed to conclusively rule out a position of Arecanae within the latter complex.

Some major differences between the rbcL-based trees and the trees found with the combined data set lay in where Velloziaceae was positioned, and in what constituted the sister-group of Pontederiaceae (Commelinaceae or Velloziaceae for the former data-set, Commelinaceae-Haemodoraceae for the latter). The various rbcL-based trees were the most distinctive entities found in all the analyses, both from each other and from the trees found from the ndhF data set and the combined data set. In most cases, phylogenetic structure in trees
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

Fig. 4.3. The phylogenetic position of Pontederiaceae in a local group of monocotyledons, for a combined ndhF and rbcL data set. The tree is one of three most-parsimonious ones found with all characters equally weighted and was also the shortest tree found after characters were re-weighted according to their rescaled consistency indices (see text). The three trees have a length of 1130 steps, CI = 0.638, CI (excluding autapomorphies) = 0.504, and RI = 0.410. The root has been placed (a posteriori) such that Cyanastrum cordifolium and Narcissus elegans are the outgroups to the other taxa. The reconstructed number of mutations is indicated above each branch (ACCTRAN optimization). The bootstrap support for different branches was estimated using various weighting schemes and is indicated below each branch. Each value in square brackets is the bootstrap support under equally weighted parsimony. The three values in curly brackets are from analyses where characters were re-weighted according to their maximum rescaled consistency indices (first value) on the three shortest equally weighted trees, or their site-likelihoods under simple or parameter-heavy likelihood models (second and third values respectively; the two likelihood models are described in the text) using the equally weighted tree with the lowest -loge likelihood. The superorders that particular taxa belong to are labelled as follows: A = Arecales, B = Bromeliales, C = Commelinales, L = Liliales, Z = Zingiberales.
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

derived from this data-set (Fig. 4.2) had lower bootstrap support than for the other two data sets, particularly in instances where the rbcL trees had major conflicts with trees derived from the combined data set. This was the case even after re-weighting the rbcL data (compare Fig. 4.1 and 4.3, for example).

*Congruence of the two chloroplast genes.* -- Although the combinability analysis indicated no significant incongruence between the two chloroplast genes for the taxa shown in Figs. 4.1-4.3, there was a trend for this ($P = 0.17$), which raises the possibility of marginal incongruity between the two genes at this taxonomic level. Excluding the four families (Cannaceae, Haemodoraceae, Musaceae and Velloziaceae) for which there were no exact correspondents between the data sets resulted in a reduction in the number of re-partitioned replicates where the combined length of the re-partitioned data sets exceeded shortest tree length for the combined data sets [$P = 0.56$ with all four families excluded; $P = 0.44$ with *Anigozanthos* (Haemodoraceae) excluded; $P = 0.26$ with *Canna* (Cannaceae) and *Musa* (Musaceae) excluded; $P = 0.19$ with *Barbacenia/Vellozia* (Velloziaceae) excluded]. This suggests that lack of exact identity in the taxa employed for different data sets can be a source of phylogenetic incongruence, although not a major one in this study.

*The effect of a posteriori character re-weighting on phylogenetic estimation.* -- The character re-weighting schemes tended to result in convergence on one of the shortest trees found under equally weighted parsimony for that data set, and in increased bootstrap support compared to that observed with equally weighted parsimony. Those aspects of the topological structure of trees from the single-gene data sets which conflicted among data sets generally had poor bootstrap support under equally weighted parsimony and one of the two re-weighting schemes (character weights derived using rescaled consistency indices or site-likelihoods under a simple
maximum likelihood model), although particular conflicting branches sometimes had improved bootstrap support under the remaining re-weighting scheme (compare Figs. 4.1 and 4.2). Analysis of the combined sequence data using weights based on equally weighted parsimony and all three re-weighting schemes (character weights derived using rescaled consistency indices or site-likelihoods under a simple or complex maximum likelihood model) yielded identical or nearly identical trees, generally with strong bootstrap support. Areas of these trees that were more poorly supported under equally weighted parsimony, generally were strongly supported under all three re-weighting schemes examined for the combined sequence data (Fig. 4.3).

The root of Pontederiaceae. -- Including several related families as outgroups substantially broke up the long branch connecting the family to these taxa. This branch length is the distance between the ingroup node (the ancestor of all extant members of Pontederiaceae) and the first outgroup node. (Where a single taxon was used as an outgroup, the outgroup node is the outgroup itself.) When Commelinaceae, Haemodoraceae and Philydraceae were individually employed as outgroups with the combined data, this branch varied between 185 steps (with Commelinaceae as outgroup; ACCTRAN optimization) to 134 steps (with Philydraceae as outgroup; DELTRAN optimization) for the optimal rooting of Pontederiaceae shown in Fig. 4.4C. When the sister-group of the family (i.e., Commelinaceae-Haemodoraceae) was used as the outgroup, the branch connecting the ingroup node and the first outgroup node was 77-80 steps long (DELTRAN-ACCTRAN optimization). Attachment of Philydraceae to this branch reduced the distance between ingroup and the new first outgroup node to 71-72 steps (DELTRAN-ACCTRAN optimization). By comparison, the longest branch in the ingroup was just under a third as long as this (Fig. 4.5).

For each tree defining a different possible rooting of Pontederiaceae, I used Felsenstein's (1995) version of Templeton's (1983) test to examine whether or not excess tree
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

lengths of each sub-optimal root position (compared to the optimal root) were larger than 1.96 standard deviations in the step differences between this rooting and the best one (see Fig. 4.4C for example values). For all the Lundberg-rooting experiments that used sequences from real outgroup taxa (Fig. 4.4A-C), there were multiple sub-optimal root locations that were not significantly different from the most-parsimonious root location. The number of sub-optimal root positions that were not significantly different from the most optimal one was greater for each individual outgroup (19 for Commelinaceae, 9 for Philydraceae, 11 for Haemodoraceae; Fig. 4.4B), than for the composite outgroup composed of these three families (eight; Fig. 4.4A). Most possible rootings were not significantly different from the best root for the restriction-site data (34 of 44 sub-optimal branches when Commelinaceae and Philydraceae were used to root the tree using the restriction-site data, versus 13 when the combined sequence data were used for these same outgroups; Fig. 4.4C).

The same most parsimonious root position (the branch separating Heteranthera s.l. from the other taxa in the family) was found for the combined sequence data for almost all combinations of the closely related taxa using equally weighted parsimony. These include the use of Haemodoraceae or Philydraceae as outgroups by themselves (Fig. 4.4B); the three possible pair-wise combinations of outgroups involving Commelinaceae, Haemodoraceae and Philydraceae [including the one used in Fig. 4.4C and the sister group by itself (not shown)]; and the sister-group together with the next most closely related taxon [i.e., ((Commelinaceae-Haemodoraceae), Philydraceae) together; Fig. 4.4A]. This latter sub-tree constitutes the sister group of the family plus the next most closely related taxon, as determined using the combined sequence data (Fig. 4.3). This composite outgroup taxon should provide the best estimate of the root of Pontederiaceae, because it involves the two most closely related clades and has the shortest branch connecting the ingroup node to the first outgroup node.

I repeated the rooting experiments using this latter outgroup for the three weighting schemes that were derived from the twelve outgroup taxa for the combined sequence data (two
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

Fig. 4.4. Optimal and sub-optimal rootings of a chloroplast-based tree of Pontederiaceae with closely related outgroups. Arrows indicate the most-parsimonious roots of the family as determined by each outgroup. Templeton's (1983) test (as implemented in DNAPARS of PHYLIP 3.5c; Felsenstein 1995) was used to assess whether the excess lengths for each root were significantly different from the best root (with all characters equally weighted). To illustrate the test, excess lengths and one standard deviation in the step differences between each root and the best root (in parentheses) are included for each branch in 4A (indels excluded). All sub-optimal rootings that did not yield a significant increase in tree length are indicated with symbols on the relevant branches -- for each outgroup, placing the root of the family on those branches lacking symbols yielded significantly longer tree lengths than the most-parsimonious rooting. The chloroplast-based tree of Pontederiaceae is one of four most-parsimonious unrooted trees found from a combined analysis of the three chloroplast data-sets (see text). (A) Rooting with a composite outgroup composed of the sister-group of the family indicated in Fig. 4.3 [i.e., (Haemodoraceae, Commelinaceae)], together with the next most closely related outgroup (Philydraceae), for the combined sequence data. (B) Rooting with three individual outgroups (Philydraceae, Haemodoraceae or Commelinaceae respectively), for the combined sequence data. (C) Rooting with a composite outgroup composed of Philydraceae and Commelinaceae for a restriction-site data-set (Appendix A) and for the combined sequence data set.
A. Data-set = $[ndhF \text{ plus } rbcL]$  
Outgroup = $(Philydrum, (Anigozanthos, Tradescantia))$
B.

Data-set = [ndhF plus rbcL]
Outgroup = Anigozanthos (△), Philydrum (○), or Tradescantia (□)
C. Data-set = \([ndhF \text{ plus } rbcL}\) (●) or Restriction Site (□)
Outgroup = (\textit{Philydrum}, \textit{Tradescantia})
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

based on site-likelihoods and one based on the rescaled consistency index; see Materials and Methods). The same most-parsimonious root position was found under all three additional weighting schemes, although the length difference between optimal and next-most suboptimal root position (which was on the same branch as the one of the two next most sub-optimal ones found under equally weighted parsimony; Fig. 4.4C) was sometimes very small.

Random outgroup sequences attached to the ingroup tree in a non-random manner (Fig. 4.5; left-hand tree). The optimal rootings for the majority of these sequences were onto the three longest ingroup branches in the unrooted tree of Pontederiaceae (Fig. 4.5; right-hand tree). However, several quite long branches, including the branch that most of the real outgroups that I examined attach to, did not attract any of the random sequences (Fig. 4.5; compare left- and right-hand tree). Each random sequence was not equally attracted to the different branches in the ingroup. There was a definite ranking in preferred points of attachment to the ingroup (Fig. 4.6). The estimated mean and variance for the 100 random sequences in the increase in tree length over the optimal rooting, for consecutively worse-ranked root positions, is shown in Fig. 4.6. When all possible root positions for the real composite outgroup were ranked in the same manner, a similar pattern of decreasing parsimony in real and random outgroups was observed. The real outgroup had a greater penalty than the mean random outgroup for all ranks. However, it was only after the 21st worse-ranked case rooting that a significantly greater penalty was observed for the real outgroup than for random outgroups (Fig. 4.6).

DISCUSSION

Accuracy of the estimated phylogenies. -- Bootstrap analysis can provide good but somewhat biased estimates of "accuracy," the probability that a specified branch is contained on the true tree, if the data are not "inconsistent" under a particular optimality criterion such as maximum
Fig. 4.5. Frequencies of different rootings of a chloroplast-based tree of Pontederiaceae with the combined \textit{ndhF} and \textit{rbcL} data set for the family, for 100 random outgroup sequences. For each random outgroup sequence, tree length was determined for all 45 possible root locations on the tree (with all characters equally weighted), with the best root(s) being the shortest tree or trees in each case. The frequencies (indicated on each branch of the left-hand tree) sum to greater than 100, as some outgroups have more than one most-parsimonious root position. Branches on the left-hand tree that lack values were not found to be most-parsimonious root locations for any of the random outgroups. This chloroplast-based tree of the family is one of four most-parsimonious unrooted trees found from a combined analysis of three chloroplast data-sets (see text). Branch lengths for the ingroup for the combined \textit{ndhF} and \textit{rbcL} data set are shown on the right-hand tree (values above each branch). Note that one branch has a maximum length of zero with the \textit{ndhF} and \textit{rbcL} data. Where DELTRAN estimates of branch length differed from ACCTRAN estimates, the former is indicated in parentheses.
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

Fig. 4.6. The penalty in parsimony observed when real and random outgroups are attached to sub-optimal locations on a chloroplast-based tree of Pontederiaceae. The tree is one of four most-parsimonious unrooted trees found from a combined analysis of three chloroplast data-sets (see text). The real outgroup (circles) is composed of the sister-group of the family indicated in Fig. 4.3 [i.e., (Haemodoraceae, Commelinaceae)], together with the next most closely related outgroup (Philydraceae). The penalty is the number of extra steps (with all characters equally weighted) for the second-best, third-best (and so on) root locations compared to the most-parsimonious root location for each outgroup sequence. In cases where there were tied penalties (see the real outgroup for examples), these were arbitrarily assigned ranks spanning the number of ties. The rank labelled "0" on the y-axis corresponds to the most-parsimonious root locations. For the 100 random outgroups examined, the mean penalty (dashes) is shown for each consecutively worse-ranked root locations, plus or minus two standard deviations.
parsimony (reviewed in Swofford et al. 1996). ["Inconsistency" means that there are so many undetectable homoplasious evolutionary events on particular branches that the data set tends to converge on an incorrect tree, even with increasing number of characters (Felsenstein 1978).] This bootstrap bias is such that high bootstrap values give underestimates of accuracy and low ones give overestimates of it (Hillis and Bull 1993). The degree of bias is affected by a number of factors, including taxon and character number (Hillis and Bull 1993; Zharkikh and Li 1995; Li and Zharkikh 1995).

It is consequently possible that some of the phylogenetic structure of the trees found with the combined data is a consequence of long-branch attraction rather than a true record of phylogenetic history (see Felsenstein 1978; Hendy and Penny 1989). I attempted to correct for the possible distorting effects of multiple hits along these long branches using several character re-weighting schemes. Improved taxon sampling in groups related to Pontederiaceae (see below) may further help decrease the possible distorting effects of long-branches on phylogenetic estimation (Smith 1994). Tree and bootstrap estimates from analyses that successively weight using consistency indices may experience another kind of bias that results from the starting tree used to estimate weights. This additional bias may result in convergence to a local optimum rather than towards the true tree (Maddison and Maddison 1992; Swofford et al. 1996), and may be avoided with extreme weightings based on each character's consistency index (J. McGuire and J. Huelsenbeck, pers. comm. in Swofford et al. 1996). Swofford et al. (1996) suggested that maximum-parsimony analysis performed using ML-based weighting schemes may provide improved handling of potentially inconsistent data over equally weighted parsimony, with the attendant tractability of maximum-parsimony analysis. They also suggested that the relative frequencies of various changes used to derive such weights are not greatly biased by the starting tree.

Weights based on site-likelihoods should therefore provide a more accurate correction for multiple hits in DNA sequences than weights derived from the rescaled consistency index.
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

However, for the single gene data sets, re-weighting using site-likelihoods did not consistently reduce bootstrap support for topological structure that conflicted among the individual data sets. For the combined sequence data, re-weighting schemes based on site-likelihoods (under two different maximum-likelihood models) and rescaled consistency indices produced identical or nearly identical trees to those found using equally weighted parsimony, and provided very similar levels of bootstrap support for individual branches — most of these branches were well supported under all three re-weighting schemes (Fig. 4.3). I treat the trees found with the combined sequence data as the current best estimate of the local position of Pontederiaceae in the monocotyledons.

Systematic relationships of the 13 families

Conflicts among trees estimated using the single and combined data sets. — The sister group of Pontederiaceae varied among the data sets. The various shortest rbcL trees found under the different weighting schemes had conflicting and mostly poorly supported arrangements of taxa associated with Pontederiaceae (see above). Two of the shortest trees depicted Commelinaceae as the sister-group (Fig. 4.2) of Pontederiaceae, one depicted Velloziaceae as the sister-group. One of the two shortest equally weighted trees found using the rbcL data also depicted a clade containing Pontederiaceae and three other families (Commelinaceae, Haemodoraceae, and Philydraceae). This clade (with variable arrangement of the four families) was seen in some more taxon-dense studies of rbcL (Chase et al. 1993; Duvall et al. 1993; Chase et al. 1995) but not in others (Chapter 2). Support for this clade in these studies (as measured by bootstrap and decay analyses) was poor. Bootstrap support for this clade for the rbcL data was also poor in this study (<50% for all weighting schemes; Fig. 4.2), but was somewhat higher with the different re-weighted analyses than under equally weighted parsimony. Except for the weighting scheme based on site likelihoods, the various weighting schemes did not provide
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

good support for any particular taxon being the sister-group of Pontederiaceae, for the $rbcL$ data set (Fig. 4.2).

Various other $rbcL$-based studies (e.g., Chase et al. 1993; Chase et al. 1995) and a study based on chloroplast DNA restriction-site variation indicated a placement of Velloziaceae outside the ABC2 complex and depicted it as the sister-group of a clade consisting of the superorders Pandananae and Cyclanthanae sensu Dahlgren, Clifford and Yeo (1995). However, the analyses based on $rbcL$ data here (Fig. 4.2) and in Chapter 2 sometimes indicated a close relationship of Velloziaceae to other taxa in Bromelianae sensu Dahlgren, Clifford and Yeo (1985). The Pontederiaceae-Velloziaceae clade was observed in one of the $rbcL$ equally weighted trees here and in a slightly sub-optimal $rbcL$-based tree in Chapter 2. Bootstrap support for local clades containing Velloziaceae was poor in this study and in that of Duvall et al. (1993). There was little bootstrap support by the $rbcL$ data for the Pontederiaceae-Velloziaceae clade (34% under equally weighted parsimony, 33% for weights based on rescaled consistency indices and <5% for weights based on site-likelihoods). This suggests that the uncertain positions of Velloziaceae in the $rbcL$ trees here and in previous studies contributed to the low bootstrap support for local clades containing this family.

The most-parsimonious tree found with the $ndhF$ data set depicted a clade containing Pontederiaceae, Haemodoraceae and Philydraceae, with Arecaceae embedded within it. However, the bootstrap support for the membership of Arecaceae in this and other local clades was poor for the $ndhF$-based trees under most weighting schemes (Fig. 4.1). Philydraceae was depicted as the sister-group of Pontederiaceae for this data set, an arrangement with poor to strong bootstrap support under the various weighting schemes (Fig. 4.1). This arrangement was also seen in the shortest equally weighted trees from a phylogenetic study that employed restriction-site variation in a slowly evolving region (the inverted repeat) of the chloroplast genome (Davis 1995), although there was also very poor bootstrap support (38%) for the Philydraceae-Pontederiaceae clade in that study. Bootstrap support for a clade consisting only
of Commelinaceae, Haemodoraceae, Philydraceae and Pontederiaceae was very poor for the *ndhF* data set, varying between 14% under equally weighted parsimony, 17% for weights based on rescaled consistency indices and 24% for weights based on site-likelihoods.

*Summary of systematic conclusions.* -- The combined analysis of the two chloroplast genes yielded shortest trees with substantial bootstrap support for most branches, both with and without character re-weighting. The monophyly of several clades (Pontederiaceae, Zingiberanae and to a more limited extent Flagellariaceae-Bromeliaceae-Typhaceae) was strongly supported in most analyses, as was the placement of the two Asparagalean taxa (Cyanasteraceae and Amaryllidaceae) apart from the other taxa (Fig. 4.1-4.3). Neither Bromelianae nor Commelinanae (*sensu* Dahlgren, Clifford and Yeo 1985) were monophyletic for any data set, in line with several previous studies of *rbcL* (Chase et al. 1993; Duvall et al. 1993; Chase et al. 1995; Chapter 2) and chloroplast DNA restriction-site data (Davis 1995). In the current study, trees obtained from the combined sequence data concurred over those features that were consistent between the various single gene trees (i.e., the monophyly of Zingiberanae and Pontederiaceae; the existence of a Flagellariaceae-Typhaceae-Bromeliaceae clade and an Amaryllidaceae-Cyanasteraceae clade). The combined data also provided strong support for the isolated position of Velloziaceae from other members of Bromelianae and its exclusion from the ABCZ complex (see the weighted analyses of the combined data set; Fig. 4.3). Arecanae was placed as the sister-group of the BCZ complex, but a firm exclusion of this super-order from this BCZ complex was not provided by any bootstrap analysis of the combined data (Fig. 4.3). A denser taxon sampling would be valuable to further clarify the positions of Velloziaceae and Arecanae.

Several local monocot clades containing Pontederiaceae that were not seen on either single-gene tree were seen in all shortest trees for the combined data, and generally had robust support from the bootstrap analyses. A clade consisting of Commelinaceae, Haemodoraceae,
Philydraceae and Pontederiaceae was seen in all on the shortest trees with the combined data. While it was poorly supported under equally weighted parsimony (by only 56% of bootstrap replicates), it was strongly supported in bootstrap analyses using all three character re-weightings (Fig. 4.3). The clade consisting of Commelinaceae and Haemodoraceae on these trees (Fig. 4.3) had moderate to very strong support under all four weighting schemes, and was depicted as being the sister-group of Pontederiaceae in all most-parsimonious trees for the combined data. Bootstrap support for this latter relationship was very poor under equally weighted parsimony (only 27%), but enjoyed moderate (67%) to strong (91%) bootstrap support with all three re-weighted analyses (Fig. 4.3).

These conclusions concerning relationships among the families examined in this study should be tempered with the realization that a denser taxonomic sampling of the ABCZ complex may overturn some of them. For this reason, it would be particularly valuable to obtain ndhF sequences from Hanguanaceae, Calectasiaceae and Dasypogonaceae, several families that Dahlgren, Clifford and Yeo (1985) placed in Lilianae, but which the rbcL-based analysis of Chase et al. (1995) indicate may belong within an expanded ABCZ clade. A representative of Cyperales s.l. (Commelinanae) is also missing here. I was unable to amplify and sequence the ndhF locus for Cyperus alternifolius. However, there is no indication from the current molecular and morphological evidence that any of these taxa are very closely related to Pontederiaceae. It would also be profitable to increase taxonomic sampling within Commelinaceae, Haemodoraceae and Philydraceae for these two genes, as this may further stabilize the phylogenetic relationship of Pontederiaceae to these taxa (cf. Smith 1994). The best estimate of the local position of Pontederiaceae in the monocotyledons, however, is that its sister group is Commelinaceae-Haemodoraceae and that the next most closely related taxon is Philydraceae.
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

The root of Pontederiaceae

If enough “multiple hits” go undetected on the branches connecting the ingroup to the outgroup taxa, the wrong root may be estimated. The use of multiple, but closely related, outgroup taxa to root Pontederiaceae substantially broke up the length of the first branch connecting this family to other taxa. Splitting such branches is potentially very important, because it should decrease the possibility that undetected homoplasious events (“multiple hits”) will lead to spurious rooting (Hendy and Penny 1989; Smith 1994). Such rooting artifacts occur because these undetected homoplasies can be “misinformative” when estimating phylogenies using maximum parsimony (Swofford et al. 1996). I used two taxa from the sister group of the family (one each from Commelinaceae and Haemodoraceae) and one from Philydraceae (the next most closely related taxon) to obtain a best current estimate of the most-parsimonious root of Pontederiaceae. Using these taxa together to root the family helped to break up the long branch below the ingroup node, but even at its shortest, this branch was still around three times longer than any branch in the ingroup (see Results). This suggested that there is still some potential for long-branch attraction leading to a rooting of Pontederiaceae that is an artifact of the data, rather than a reflection of true historical signal in the outgroup sequences.

The number of sub-optimal root positions that were not significantly different from the most-parsimonious one was smaller when the composite outgroup was used to root the family, than when any of the three individual outgroups were used (cf. Fig. 4.4A-B). This suggests that using multiple outgroups to break up the long branches connecting them to the ingroup can result in greater discriminating power for locating root position than using the outgroups individually. More than three-quarters of all possible root positions were not significantly different from the most optimal root position for the restriction-site data set (Fig. 4.4C). This indicates that this data set has a much poorer ability than the combined sequence data to
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

discriminate where the true root of the family lies. This may be because the branches connecting the ingroup tree to the outgroups are effectively saturated with multiple hits for the restriction-site data set.

Wheeler (1990) described a test for whether individual outgroup taxa are likely to contain any historical signal. However, he ignored the potential for multiple outgroup taxa to break up the long branches on which the historical signal can become randomized. I used several alternative means for determining whether any historical signal in a given outgroup has been effectively randomized by multiple hits. Random outgroup sequences do not attach randomly to unrooted phylogenetic trees of ingroups. Long-branch attraction means that random sequences (or real sequences in which the phylogenetic signal has been effectively randomized) have definite preferred points of attachment on trees in which there are branches with unequal length. Most random sequences attach to one of the three longest branches on the chloroplast-based tree of Pontederiaceae (Fig. 4.5). The rooting found with the real outgroups for the combined sequence data did not correspond to that found by any random outgroup sequences I examined (cf. Fig. 4.4, 4.5).

I also compared the length increases for consecutively worse sub-optimal roots by a real outgroup, to the length increases for same-ranked rootings by random outgroups. This was first suggested by Miyamoto and Boyle (1989), although these workers examined only the best and next-best root position to assess whether real outgroups possess useful historical signal concerning root location. However, the absence of a significant difference between the best and next best rooting does not mean that an outgroup is no better at determining root positions than random sequences. It may be necessary, as here (Fig. 4.6), to examine substantially less parsimonious root positions to demonstrate that this is not the case. There was an apparent mismatch between this analysis and the one based on Templeton’s (1983) test in terms of how many sub-optimal outgroup locations are significantly worse than the most-parsimonious one (just over 50% in the former case and over 80% in the latter case; cf. Figs. 4.4C, 4.6).
CHAPTER 4. THE LOCAL POSITION OF PONTEDERICEAE AND ITS ROOT

Templeton’s test may be a more sensitive test, because variances in the other test are calculated for equivalently ranked preferences of the random sequences in the other method, rather than for each actual root position. The calculation was done that way because random sequences have different preferred patterns of attraction to ingroup branches, so the actual locations of the tenth worst root location for one random sequence will usually not correspond to the tenth worst one for another, for example. The variances calculated using Templeton’s test are instead for each actual rooting by the outgroup.

How far can we trust the rootings of the family indicated by the sequence data from the two chloroplast genes? One answer is that we must treat the entire cluster of sub-optimal roots that are not significantly worse than the optimal root (Fig. 4.4A) as the best estimate of root position. However, several lines of evidence lend weight to the rooting of the family between *Heteranthera s.l.* and the other genera being the true root of the chloroplast tree. First, despite being only half as long as the longest branch in the ingroup, this was a branch to which none of the random sequences were attracted (Fig. 4.5). Outgroups that have had their historical signal completely randomized with multiple hits are only rarely attracted to this long branch. Second, this root was the most-parsimonious one under equally weighted parsimony, when different combinations of closely related taxa were used as outgroups. Several of these outgroups constitute observations that are completely independent, apart from any historical signal that they may share. Even if (as seems probable) some randomization of phylogenetic signal has occurred along the branches leading to these outgroup taxa, it is unlikely that this would bias independent outgroups to attach in the same place, particularly to a branch to which random outgroups are not particularly attracted. Finally, this was also the most-parsimonious rooting under several different weighting schemes. These character re-weighting schemes should at least partially correct for the differential ability of different sites in the DNA sequences to record phylogenetic history.
CHAPTER 5

ADAPTIVE RADIATION IN THE AQUATIC PLANT

FAMILY PONTEDERIACEAE
The invasion of aquatic environments from land has occurred repeatedly during the evolutionary history of the flowering plants. The precise number of transitions from land to water is not known with certainty, although Cook (1990) recently estimated that it may have taken place a minimum of 50 but possibly many times more. Approximately 33 diverse families of monocotyledons and dicotyledons are exclusively aquatic and numerous aquatic genera are found in predominantly terrestrial plant families. Aquatic plants constitute only 1-2% of angiosperms but they have received considerable attention from botanists and ecologists, primarily because of the high degree of ecological specialization that they exhibit. Adaptation to life in water has demanded the evolution of a distinctive array of morphological, anatomical, physiological and biochemical attributes that have developed on multiple occasions among the lineages that have invaded aquatic habitats.

Depending on the degree to which the life cycle of an aquatic plant is spent in water, individual taxa show increasing divergence from their terrestrial ancestors. At one extreme are species that spend their entire lives submerged below the water surface and are most distinct from land plants, to amphibious taxa that are equally at home on land or in water and that closely resemble their strictly terrestrial relatives. Aquatic groups often display considerable evolutionary diversification resulting from adaptation to the wide range of ecological conditions that occur in wetland habitats. This diversity offers excellent opportunities for relating form to function (reviewed in Arber 1920; Sculthorpe 1967; Crawford 1987; Barrett, Eckert and Husband 1993).

While the concept of adaptive radiation is central to evolutionary theory, there is a wide range of viewpoints as to what it entails. Futuyma (1986) states that adaptive radiation is simply diversification into different ecological niches by species derived from a common
ancestor. According to Simpson (1953), however, such diversification is a direct response to a novel ecological or geographic circumstance experienced by the common ancestor of species involved in the radiation. More recently, the idea that increased species richness may (or must) be associated with adaptive radiation has become prevalent among phylogenetic systematists (e.g., Brooks and McLennan 1991). In this paper however, I follow Simpson's perspective that adaptive radiation occurs through character diversification among different lineages in response to a novel set of ecological circumstances or a key innovation. This process may involve an increase in speciation rate, no change in speciation rate, or even a reduction in speciation rate. In this view a radiation in slow motion is still a radiation -- the number of lineages arising from an adaptive radiation is of secondary importance to the patterns of character diversification among lineages.

Aquatic plant groups have rarely been investigated from the perspective of adaptive radiation. While in part this is undoubtedly associated with the paucity of phylogenetic data available for most angiosperm families, it may also have been because of a widespread belief that aquatic environments are relatively homogeneous compared with those on land, thus providing less opportunity for evolutionary diversification. Indeed, such arguments have frequently been used to explain the apparently conservative macroevolutionary patterns found in certain aquatic groups (Sculthorpe 1967; Hutchinson 1975; Les 1988; Cook 1990). Of the 33 strictly aquatic families, 30 include fewer than 10 genera, 17 contain only one genus and three consist of a single species (Sculthorpe 1967). Increasing commitment to an exclusively aquatic existence appears to be associated with reduced taxonomic differentiation, as groups containing primarily amphibious and emergent aquatics show little evidence of reduced species diversity. The suggestion that some aquatic radiations are associated with reduced species richness (via decreased speciation rates or increased extinction rates) is intriguing, but requires confirmation with detailed phylogenetic analysis of the sort suggested by Sanderson and Donoghue (1994) and Nee and Harvey (1994).
CHAPTER 5. ADAPTIVE RADIATION IN PONTEDERIACEAE

The wide spectrum of life forms and diversity of reproductive strategies found in aquatic plants suggests that extensive character diversification has occurred in response to the novel ecological opportunities afforded by the aquatic habit. The breadth of adaptations implies that the habitats occupied by aquatic plants are far from ecologically uniform, as is often supposed. Because of their many specialized features, aquatic plant groups can provide outstanding opportunities for studies of adaptive radiation and character evolution, as has been undertaken in many animal groups that are restricted to aquatic environments.

Pontederiaceae is a small monocotyledonous family of exclusively freshwater aquatics composed of approximately six to nine genera and 35 to 40 species, most of which are native to the New World tropics (Barrett 1978a). Members of the family display a remarkable diversity of life-history and reproductive strategies ranging from highly clonal, long-lived taxa that inhabit permanent marshes and river systems to exclusively sexual species that are annual and occur in ephemeral pools, ditches and rice fields. Linking these extremes are species with various combinations of sexual and asexual reproduction and a variety of different pollination and mating systems. Evolutionary studies of the family over the past two decades have focused primarily on the floral biology and sexual systems of selected taxa (reviewed in Barrett 1988a, 1993; Barrett, Kohn and Cruzan 1992). More recently, phylogenetic reconstructions using both morphological (Eckenwalder and Barrett 1986) and molecular data (Chapters 2, 3, 4 and Appendix A) have been employed to investigate character evolution and the systematic relationships of taxa within the family and its close relatives.

The diversity of life-history traits in Pontederiaceae suggests that this family might provide a valuable opportunity for investigating the processes of adaptive radiation in an aquatic plant group. Carson (1985) and Johnson (1996) suggested two major modes of adaptive radiation in plants: growth-environment-driven and pollinator-driven. Below I review the patterns of character variation and ecological differentiation in vegetative traits (with a particular focus on traits important for growth under aquatic conditions) and reproductive characters in
Pontederiaceae. These lines of evidence suggest that selection acting on reproductive and vegetative characters has contributed to the radiation of taxa in this family.

I begin by providing a brief review of the taxonomy and natural history of Pontederiaceae, emphasizing the diversity of life-history and morphological adaptations to life in and out of water that occur in the family. I then perform phylogenetic reconstructions to examine the origins and evolution of a range of life-history and reproductive attributes, including aquatic life form, life-cycle duration, patterns of leaf development, types of clonality, floral form and self-incompatibility system. Throughout this chapter two particular issues form the basis of much of the discussion: (1) What is the ecological evidence that the various morphological characters considered are adaptations in response to an aquatic life-style? (2) What is the phylogenetic pattern of diversification in these characters, and which traits have evolved on multiple occasions within the family?

**Taxonomy and Natural History**

**Taxonomy** -- Pontederiaceae is composed of four main genera; *Eichhornia* (8-9 spp.), *Pontederia* (6 spp.), *Heteranthera* (10-12 spp.) and *Monochoria* (7-8 spp.) and several smaller, segregate genera [*Eurystemon* (1 sp.), *Hydrothrix* (1 sp.), *Scholleropsis* (1 sp.) and *Zosterella* (1-2 spp.)] allied with *Heteranthera*, and one genus [*Reussia* (2-3 spp.)] allied with *Pontederia*. It is not clear to which groups of monocotyledons Pontederiaceae is most closely related (see Dahlgren and Clifford 1982; Dahlgren, Clifford and Yeo 1985; Simpson 1987; Rosatti 1987; Goldberg 1989) although recent treatments suggest a close affinity to Philydraceae and Haemodoraceae (Hamann 1966; Huber 1977; Simpson 1990; Thorne 1992a, 1992b).
Biogeography -- Members of Pontederiaceae are largely tropical in distribution with the primary concentration of species occurring in the Neotropics, particularly lowland South America and especially Brazil. Several taxa occur in North America, with some reaching as far north as Canada [e.g., Pontederia cordata (Fig. 1.2A) and Heteranthera (Zosterella) dubia (Fig. 1.2F)]. In common with many other aquatic plants (see Ridley 1930; Sculthorpe 1967; Cook 1987), most species of Pontederiaceae have widespread distributions, often involving strikingly disjunct areas (e.g., Eichhornia paradoxa, Eichhornia paniculata; Barrett 1988a). All members of the family occur in freshwater habitats frequented by waterfowl and wading birds that are capable of mediating long-distance dispersal. Aside from Pontederia and Reussia which are relatively large-seeded, all species have small diaspores that are likely to adhere to mud and be easily transported on the feet of waterfowl. In some cases long-distance dispersal can be achieved by stem fragments acting as floating vegetative propagules. The occurrence of Eichhornia crassipes (Fig. 1.2C) and Eichhornia azurea throughout the large river systems of South America and also on some Caribbean islands may have been largely the result of dispersal by vegetative means (Barrett 1978b; Barrett and Forno 1982). The natural distributions of a handful of Pontederiaceae have been extended over the past century due to human influences. Several New World Heteranthera species (Heteranthera limosa, Heteranthera rotundifolia, Heteranthera reniformis) occur as weeds of rice in Europe and Asia (C. Horn, pers. comm.; S.W. Graham, pers. observ.). Monochoria vaginalis (Fig. 1.2D), a noxious weed of Asian rice, has also been introduced to Californian rice fields, probably as a seed contaminant (Barrett and Seaman 1980). The most widespread and economically important member of the family is the notorious clonal weed water hyacinth (E. crassipes). Originally native to lowland South America, vast floating mats of this species now infest lakes, rivers, reservoirs and drainage canals in many parts of the warmer regions of the world (Barrett 1989c).

Aquatic habitats and ecological differentiation. -- Members of Pontederiaceae can be found in a
wide variety of natural and man-made habitats provided by lakes, rivers, streams, permanent marshlands, bogs and fens, seasonal pools, drainage ditches, low-lying pastures, and rice-fields, indicating a wide range of habitat preferences within the family. Aquatic habitats can be exceptionally diverse and therefore provide considerable opportunities for ecological differentiation by aquatic plants. Extensive field observations of the family over the past two decades by S.C.H. Barrett indicate that the most significant features of aquatic environments that determine whether a particular species of Pontederiaceae can persist relate to the permanency of the habitat, depth of water, extent of water-level fluctuations, amount of nutrient loading, and the degree of interspecific competition from other aquatic plants. Of particular importance in determining the aquatic life form, duration of the life cycle and degree of clonality of individual species are the overall depth of water, the predictability of the habitat and the degree of interspecific competition.

I outline below the variation in growth forms, life histories and reproductive strategies found in members of the Pontederiaceae and discuss their likely ecological and evolutionary bases. In order for phylogenetic reconstructions of character evolution to be conducted it is necessary to classify the range of observed variation in traits. This exercise can be difficult, particularly where apparently continuous variation occurs or where detailed morphological and developmental information on the homology of different structures is not available. Nevertheless such classifications are attempted below in order to begin to explore the patterns of adaptive radiation in life-history traits in the family.

Aquatic life forms. -- Life form classifications of aquatic plants are many and varied (reviewed in Raunkiaer 1934; Den Hartog and Segal 1964; Sculthorpe 1967; Hutchinson 1975). Here I adopt the classification of Sculthorpe (1967) which involves distinguishing four main classes of aquatic life form; emergent, floating-leaved, free-floating and submergent. All occur in
Pontederiaceae (Fig. 1.2) and are closely associated with the habitat preferences of individual taxa. The emergent life form, in which the plant body is rooted in soil but grows above the water surface to varying degrees, is the most common aquatic form and occurs in the majority of taxa in the family. Populations with this habit can be found over a broad range of aquatic conditions from temporary pools to more permanent wetland habitats. Leaf-blades of emergent taxa are held by self-supporting petioles. This allows them to overshadow and outcompete floating-leaved and submerged species in shallow waters. In deeper waters such petioles would be too costly to produce (Givnish 1995). The emergent life form is therefore restricted to shallow locations at the edges of ponds, lakes or rivers. I distinguish two subclasses of emergents here depending on whether the growth form is largely erect or procumbent. This dichotomy is somewhat artificial since a few species occur that link the two extremes (e.g., Monochoria vaginalis, Pontederia (Reussia) subovata) and considerable plasticity in the degree of erectness is evident depending on water depth (e.g., Heteranthera seubertiana; Horn 1988) and stand density. Nevertheless, this is a useful distinction because most taxa are distinguished by whether or not internodal elongation is extensive, producing plants that have either a creeping stem or a compact, erect rosette (Fig. 1.2).

The only members of the family exhibiting the floating-leaved life form in mature plants are Eichhornia diversifolia, Eichhornia natans and Scholleropsis lutea. In these species plants are rooted to the substrate, with the stems and leaves floating on the water surface. I distinguish this growth form from the procumbent class of emergents by the predominance of truly floating leaves possessed by species in this category. While species such as Pontederia (Reussia) rotundifolia (Fig. 1.2B), E. azurea and H. reniformis frequently grow out from land over the surface of water, the majority of leaves that they produce are held erect as a result of upturned petioles and laminas. Many species of Pontederiaceae with emergent life forms produce a small number of floating leaves as they emerge from below the water surface, following seed germination or perennation (e.g., Horn 1988). However, these leaves can be viewed as
transitional, since the majority of mature leaves produced by these forms are adapted for terrestrial rather than aquatic conditions.

_Eichhornia crassipes_ is the only species in the family that is truly free-floating (Fig. 1.2C). The free-floating life form is characterized by only a brief dependence on solid substrate to enable seed germination and establishment. Once this is over, young seedlings sever their connection with the sediments in which they have established and float to the water surface. Floating is accomplished via swollen, aerenchymous petioles. Subsequent growth, clonal propagation and dispersal occurs entirely independently of land. While many taxa of Pontederiaceae with procumbent or floating-leaved growth forms can form floating mats, these are incapable of extensive growth and regeneration unless rooted to the substrate.

The final aquatic life form in Pontederiaceae is the submersed life form, represented by _Heteranthera zosterifolia_ (Fig. 1.2E), _H. dubia_ (Fig. 1.2F), and _Hydrothrix gardneri_ (Fig. 1.2G). In these species the entire plant body is submersed below the water surface, except during flowering when reproductive parts may be elevated just above the water (Wylie 1917; Rutishauser 1983). _Heteranthera dubia_ and _H. zosterifolia_ can tolerate partial emergence (e.g., mud-flat ecotypes of _H. dubia_; Horn 1983), but apart from occasional flowers above water, _Hydrothrix gardneri_ is obligately submersed. The four main aquatic life forms therefore appear to show different degrees of adaptation to the aquatic environment and could be thought of as involving an evolutionary transition from a terrestrial ancestor through an amphibious existence to a fully aquatic habit. Phylogenetic reconstruction may assist in evaluating this hypothesis by determining the direction and sequence of evolutionary change within Pontederiaceae and the ecological basis of such variation.

*Life-cycle duration.*-- The adaptive basis of life-cycle duration in flowering plants has been the subject of much discussion, with a variety of ecological and demographic factors invoked as
important selective agents (reviewed in Harper 1977; Grime 1979). Members of Pontederiaceae display a spectrum of life histories that are frequently associated with the permanency of the aquatic habitat occupied. These range from annual species that occur in ephemeral aquatic habitats such as seasonal pools, ditches, and ricefields (e.g., *E. diversifolia* and *Heteranthera* spp.) to very long-lived taxa that are largely restricted to permanent marshlands (*Pontederia* spp.) or large river and lake systems (*E. azurea*) such as those found in Amazonia and the Pantanal region of South America.

I distinguish three categories of life-cycle duration: annual, short-lived perennial and long-lived perennial. Annual species are those in which the majority of populations of a species complete their life cycle within a year. Short-lived perennials may persist for up to five years and long-lived perennials often live for considerably longer time periods. Once again these categories are not rigid, since altered ecological conditions may modify patterns of longevity in any species. For example, several of the species that I classify as annuals (e.g., *E. paniculata*) because they usually cannot persist vegetatively in their native habitats from season to season as a consequence of severe desiccation, can continue growing almost indefinitely in the greenhouse if provided with suitable conditions. In contrast, populations of some annual species (e.g., *Eichhornia meyeri, H. limosa*) display obligate annualness, undergoing programmed senescence regardless of growing conditions. Among several of the species I classify as short-lived perennials are populations that appear to be annual when grown under glasshouse conditions (e.g., *M. vaginalis* from Californian ricefields).

**Clonality.** A considerable literature has been devoted to addressing questions concerned with the ecology and evolution of clonal versus sexual reproduction (Williams 1975; Maynard Smith 1978; Bell 1982). Valuable perspectives on the adaptive basis of clonality in plants have been provided by Abrahamson (1980), Leakey (1981) and Cook (1985). Aquatic plants are of
particular interest in these discussions because of their heavy reliance on asexual methods of propagation (Arber 1920; Hutchinson 1975) and it has often been suggested that cloning may be favored in aquatic environments where regular seed reproduction is difficult in deep or turbulent water (Sculthorpe 1967). However, in a recent review Grace (1993) drew attention to the variety of different clonal strategies found in aquatic plants and argued that at least six major selective forces may be involved in the evolution of clonal growth in aquatics: numerical increase, dispersal, resource acquisition, storage, protection, and anchorage.

Clonality in members of the Pontederiaceae appears to be closely linked with the life form and longevity of individual taxa. Propagation in annual species is typically entirely sexual but most perennials in this family possess some form of clonal growth. This includes local colony expansion through rhizome growth in erect, emergent taxa such as Pontederia sagittata and Monochoria hastata, fragmentation of creeping stems in procumbent taxa with extensive internodal elongation [e.g., P. rotundifolia (Fig. 1.2B) and E. azurea], fragmentation of stems in submersed taxa (e.g., H. dubia) and the formation of slender stolons with daughter rosettes in the free-floating E. crassipes. As in many other perennial aquatics (see Eckert and Barrett 1993) the balance between sexual and asexual reproduction in Pontederiaceae can vary with habitat conditions and the combination of growth form and clonality that occurs (Richards 1982). Seed reproduction is common in most emergent taxa with rhizomatous growth or stem fragmentation because they usually occupy habitats suitable for seed germination and seedling establishment. In contrast, in submersed and free-floating taxa sexual recruitment probably occurs less often, despite seed formation, because of deep water conditions that restrict seedling establishment, and in some taxa there are populations that regenerate exclusively through clonal propagation (e.g., E. crassipes; Barrett 1980a, 1980b)

The various clonal strategies displayed by members of the Pontederiaceae serve different functions. One of these is numerical increase (i.e., reproduction via ramet formation), which is most obvious in taxa with regular stem fragmentation or stolon production. Dispersal of these
vegetative structures by water currents also enables exploitation of new environments, with the free-floating daughter rosettes in *E. crassipes* representing the most specialized adaptation for vegetative dispersal in the family. For species that experience long periods with unfavorable growing conditions, such as during winter in eastern North America, rhizomes and stem fragments are also used as perennating structures (e.g., *P. cordata* and *H. dubia*; Lowden 1973, Horn 1983). However, these structures are also capable of withstanding considerable desiccation and in tropical habitats prone to drought can contribute to persistence during dry periods. Finally, in taxa of Pontederiaceae with creeping stems or stolons, the structures involved in clonal growth are also photosynthetic and produce roots. Thus they are highly effective in resource acquisition, the exploitation of suitable habitat patches and in competition with co-existing aquatic species.

*Patterns of leaf development.* -- Aquatic plants display striking foliar plasticity involving continuous variation in leaf shape to the formation of discrete leaf-types with very distinct morphologies on a single individual. The latter condition has been referred to as heterophyll and a considerable literature exists on the proximate ecological, physiological and developmental mechanisms that control changes in leaf shape in heterophyllous species (Arber 1919; Sinnott 1960; Sculthorpe 1967; Lee and Richards 1991). Less attention has been paid to the genetic and evolutionary basis of such patterns (although see Bradshaw 1965; Cook and Johnston 1968). It is usually assumed that in aquatic plants the formation of flaccid, ribbon-shaped (Fig. 1.2E,F) or highly dissected leaves or leaf whorls (Fig. 1.2G) represent adaptive responses to submersed conditions. Heterophyll has often been considered to be a manifestation of heteroblastic leaf development, that is, the ontogenetic sequence in which early-formed 'juvenile' leaves are markedly different in appearance from later 'adult' ones. However, because the distinction between leaf types is often not clear cut and so-called 'juvenile' leaves can often be retained throughout the life cycle by neotony (Sculthorpe 1967, and see below), it is important to realize
that considerable diversity exists in the patterns of leaf development found in aquatic plants and that any attempt at classification is likely to be somewhat artificial.

With the exception of a detailed investigation of leaf ontogeny in *E. crassipes* (Richards 1983) and descriptions of heterophylly in *Heteranthera* (Horn 1988) there has been little work on the developmental basis of leaf-shape variation in Pontederiaceae. For the purpose of this study I tentatively recognize five basic classes (referred to as patterns A-E) that differ primarily in the duration of the 'juvenile' phase. In the first and most common type (pattern A) plants first produce a small number (one to four) of juvenile, linear, strap-shaped leaves, the width and size of which varies with species and degree of submersion, before producing 'adult' aerial leaves (Fig. 1.2A-D) with distinct petioles and laminae. In *Heteranthera* the juvenile leaves can be very narrow, whereas in *Pontederia* they can be up to several centimeters in width. Species with this type of leaf development are usually amphibious with seedlings commonly developing in shallow water or on wet mud. The important feature of this leaf development strategy is a rapid transition to the formation of aerial leaves, the characteristics of which vary according to species.

In the second class that I recognize (pattern B), this transition is much slower and a greater number (more than twenty) of submerged ribbon-like leaves are produced before the transition to aerial or floating leaves. This pattern of leaf development is quite restricted in the family and occurs only in *E. azurea*, *E. diversifolia*, *Eichhornia heterosperma* and *E. natans*, all species that commonly germinate in deep water and experience extended periods of seedling development under water. While ribbon-shaped leaves are always the first leaves to be produced by species in class B, damage to the shoot apex through herbivory or disease in adult plants can result in a temporary reversion to the 'juvenile' phase, indicating that heteroblastic development is not necessarily developmentally fixed.

The next two categories involve taxa in which all foliar leaves retained throughout the life-
cycle are sessile, linear and ribbon-shaped and are similar in appearance to the 'juvenile' leaves initially produced by taxa in the first two categories. I chose to recognize two separate classes of these (presumably) 'retained juvenile' or paedomorphic forms, because it seems probable that they have different developmental origins (neotenic versus progenetic; J. H. Richards pers. comm.) and ecological significance. Petiolate leaves are never produced in these taxa, the lower spathe in inflorescences of *H. zostericola* being the sole exception.

Pattern C occurs in the perennial, submersed aquatics *H. zostericola* (Fig. 1.2E) and *H. dubia* (Fig. 1.2F), and presumably reflects their predominantly submersed existence. The pattern appears to have arisen through neoteny (i.e., through slower somatic development relative to the onset of reproductive maturity), analogous to but much more accentuated than in pattern B. Such a neotenic shift may be a consequence of direct selection for retention of the juvenile leaf-form throughout the life-cycle. Ribbon-shaped leaves have been interpreted as a mechanical (anti-drag) adaptation to moving water (e.g., Sculthorpe 1967), and may play a role in counteracting diffusive limitations on photosynthesis underwater (see below). Pattern D is found in two annual taxa, *H. seubertiana* and *Heteranthera (Eurystemon) mexicana* (Fig. 1.2H). These two species are primarily emergent and are found in ephemeral pools. This pattern could have arisen through an earlier onset of reproduction (i.e., progenesis; see Alberch et al. 1979). Such precocious reproduction may be an adaptive response to the ephemeral nature of some aquatic environments (Arber 1920; Van Steenis 1957; Sculthorpe 1967). Emerged leaves from the four species exhibiting these two patterns are not flaccid but are stiff and erect and clearly adapted for terrestrial conditions.

The final class of leaf development (pattern E) is restricted to the submersed species *Hydrothrix gardneri* (Fig. 1.2G). This monotypic genus is characterized by whorls of small, thread-like leaves whose developmental origin is highly unusual (Rutishauser 1983). The leaf whorl is analogous to whorled leaves in other aquatic plant groups. This general leaf-form (whorled leaves or leaf whorls) is referred to as the 'Hippuris syndrome' by Cook (1978). The
occurrence of this syndrome in *Hydrothrix* is virtually unique within the monocotyledons (Cook 1978).

The functional significance of the *Hippuris* syndrome may be to reduce interference by upper leaves of light capture by the lower leaves in deep water (Arbor 1920). It has also been interpreted as a means of reducing drag in flowing water (e.g., Arber 1920, but see Sculthorpe 1967) or for increasing total assimilating area (e.g., Sculthorpe 1967, but see Arber 1920). A major function of ribbon- or thread-like leaves in submersed aquatics may be to decrease the impedance of leaf-surrounding boundary layers on CO₂ diffusion, which is much slower in water than in air (Givnish 1987). The more effectively 'narrow' the leaf or leaf-part is, the smaller this boundary layer will be, and the less it will act to impede diffusion and thus limit photosynthesis (Givnish and Vermeij 1977; Givnish 1987). More slowly-flowing waters should therefore favor narrower leaves. This relationship is supported by a set of comparative data compiled by Madsen (1986; see Givnish 1987).

*Floral ecology, pollination and mating systems.* — Angiosperms display a spectacular array of floral diversity associated with the pollination biology and mating systems of individual species. Indeed, reproductive adaptations associated with different pollen vectors are among the few plant traits that have been considered explicitly in the context of adaptive radiation (Grant and Grant 1965; Stebbins 1970). In addition to descriptive studies of pollination syndromes, recent functional interpretations of floral radiation have emphasized the importance of individual selection for fitness gain through female and male reproductive function (e.g., Bell 1985; Campbell 1989; Devlin and Ellstrand 1990). Using this approach it is important to distinguish aspects of floral display and design that reduce the incidence of self-pollination and inbreeding depression (Charlesworth and Charlesworth 1987) from those that promote outcrossed siring success through more effective pollen dispersal (Harder and Barrett 1996). This is of particular
importance in species that possess physiological self-incompatibility systems (Lloyd and Webb 1986; Bertin and Newman 1993; Harder and Barrett 1995).

The flowers of Pontederiaceae are showy and blue, mauve, yellow or white and can be solitary or displayed in inflorescences. They exhibit a broad range of morphological specializations associated with a variety of pollination mechanisms and mating systems. Despite being an entirely aquatic family, the floral syndromes of Pontederiaceae involve either animal pollination or self-pollination, with no evidence of the kinds of adaptations towards hydrophily that occur in many other exclusively aquatic taxa (Cox 1988). Flowers in the family are largely pollinated by bees, and to a lesser extent by butterflies. In Pontederia and Eichhornia the flowers are tubular and moderately zygomorphic and pollination is largely achieved through the services of long-tongued bees that feed primarily on nectar (Wolfe and Barrett 1988; Husband and Barrett 1992c). In Heteranthera and Monochoria, floral visitors are mostly pollen-collecting bees (Iyengar 1923; Wang, Miura and Kusanagi 1995; S. C. H. Barrett pers. observ.). In most species of Heteranthera, Hydrothrix gardneri, E. diversifolia, E. natans and M. vaginalis, flowers are also produced which develop underwater and hence are completely self-fertilized. This phenomenon is known as pseudo-cleistogamy and is commonly reported in other aquatic groups (Sculthorpe 1967). Amphibious flowers with underwater ovaries but aerial pollination organs are found in H. dubia (Wylie 1917; Horn 1985). Floral tubes in this species are very variable in length (Horn 1985) and can reach over 11 cm long (R. Rutishauser, pers. comm. in Endress 1995). Amphibious flowers and pseudo-cleistogamy represent the only obvious examples of shifts in reproductive characters in response to the aquatic habit of the family.

Two conspicuous floral polymorphisms (tristyly and enantiostyly) occur in Pontederiaceae (Fig. 1.1). The genetic polymorphism tristyly occurs in all species of Pontederia s.l. except Pontederia parviflora, and in three species of Eichhornia (E. azurea, E. crassipes and E. paniculata). Tristylious species possess a reciprocal arrangement of style and anther heights (Fig. 1.1) and an associated syndrome of ancillary characters exhibiting
polymorphisms of pollen and stigmas. The remaining taxa in both genera are small-flowered and monomorphic for style and stamen length. A self- and intramorph-incompatibility system accompanies the floral heteromorphism in all tristylious species except for *E. crassipes* and *E. paniculata*, which are tristylious but highly self-fertile (Barrett 1979; Barrett 1988a; Barrett and Anderson 1985). Experimental studies of *Pontederia* and *Eichhornia* support Darwin’s (1877) original hypothesis that the tristylious polymorphism functions to promote proficient cross-pollination among plants through the reciprocal arrangement of male and female sex organs (Barrett and Glover 1985; Kohn and Barrett 1992; Lloyd and Webb 1992a,b).

The second floral polymorphism, enantiostyly (Fig. 1.1), occurs in species of *Heteranthera s.l.* and *Monochoria* (Iyengar 1923; Eckenwalder and Barrett 1986; Wang, Miura and Kusanagi 1995). The outward-facing flowers possess either left- or right-bending styles with a single stamen reflexed in a lateral position opposite the stigma. This condition rarely occurs as a true genetic polymorphism with populations composed of plants with either right- or left-handed flowers, but not both (e.g. in *Wachendorfia*; Ornduff and Dulberger 1978). More commonly however, it exists as a somatic polymorphism with individual flowers possessing both flower types. All enantiostylious members of Pontederiaceae have the somatic form of the polymorphism. They also display a striking anther dimorphism with the reflexed stamen cryptically-colored and larger than the remaining stamens. Such dimorphism is termed heteranthery and represents a functional division of labor into attractive ‘feeding’ anthers and a cryptically colored ‘pollinating’ anther (Vogel 1978; Buchmann 1983; Lloyd 1992; Fenster 1995). Enantiostyly has most often been interpreted as an adaptation to increase the effectiveness of cross-pollination, in a manner analogous to heterostyly. However, there is little empirical evidence to support this hypothesis and where the polymorphism is somatic other factors must be involved in its origin and maintenance (Chapter 2).

While mating patterns have not been quantified in the majority of Pontederiaceae, some inferences can be drawn from information on floral biology and experimental studies conducted
by S.C.H. Barrett and colleagues over the past two decades. All species of Monochoria and Heteranthera s.l. are highly self-compatible. Undisturbed flowers of species in these two groups can usually achieve full seed-set through autonomous self-pollination. Although pollen-collecting bees visit flowers, it seems likely that populations of these taxa experience considerable self-pollination, particularly in taxa with pseudo-cleistogamous flowers. In contrast, tristylos species of Pontederia and Eichhornia with self-incompatibility must be largely outcrossing because of their physiological barrier to self-fertilization. Even where tristyly is associated with self-compatibility, marker-gene studies indicate that populations can exhibit high outcrossing rates (Barrett, Kohn and Cruzan 1993). Among non-heterostylos species of Pontederia and Eichhornia, self-fertilization is likely to predominate, since these taxa are self-compatible and homostylos, with anthers and stigmas close together within each flower (Fig. 1.1; and see Barrett 1988a).

Phylogenetic Systematics of Pontederiaceae

Morphological and molecular evidence of phylogenetic relationships. -- Molecular evidence from the chloroplast gene rbcL strongly supports the monophyly of Pontederiaceae (Chapter 2). Three highly congruent data sets derived from the chloroplast genome (based on restriction-site variation and sequence data from rbcL and ndhF) yield robust and well-resolved estimates of the phylogenetic history of the family (Fig. 5.1; Chapter 3). These estimates indicate that Monochoria, Pontederia s.l., and Heteranthera s.l. are monophyletic, but that Eichhornia is not. Morphological evidence concerning the phylogenetic history of Pontederiaceae also rejects the monophyly of Eichhornia (Eckenwalder and Barrett 1986), but is largely insufficient for
Fig. 5.1. Reconstructed phylogenetic history of Pontederiaceae using combined evidence from DNA sequence variation in the chloroplast genes *rbcL* and *ndhF*, and restriction-site variation in the chloroplast genome (Chapter 3). The tree is a strict consensus of four shortest unrooted trees (Chapter 3). Branch lengths were determined using ACCTRAN optimization (for one of the four shortest trees) and are indicated above each branch. Bootstrap values (from 100 bootstrap replicates) are indicated in bold below branches. The root indicated (with an ‘R’) is that found by a variety of closely related taxa using the two chloroplast DNA sequence-based data sets (see text and Chapter 4). For the four most-parsimonious trees found with the combined chloroplast data, tree length (including autapomorphies) = 609 steps, CI (excluding autapomorphies) = 0.525, CI (including autapomorphies) = 0.637, RI = 0.775 (Chapter 3). The bars indicate whether each taxon is found in the Old World (Asia, Australia or Africa), or the New World. Question marks indicate possible long-distance dispersal event involving several taxa missing from this analysis (see text).
CHAPTER 5. ADAPTIVE RADIATION IN PONTEDERIACEAE

estimating a robust phylogenetic history of the family (Chapter 3). In combined analyses with
the molecular evidence, the morphological evidence has almost no impact on phylogenetic
reconstructions (Chapter 3).

Despite being swamped by the molecular evidence, there was statistically detectable
incongruence between the molecular and morphological data (Chapter 3). Eckenwalder and
Barrett (1986) hypothesized that a 'selfing syndrome' in the family (involving multiple parallel
shifts in reproductive characters during the origin of predominantly self-fertilizing species)
could result in incorrect phylogenetic reconstructions using the morphological data. However,
as this hypothesis is not actually supported by the morphological evidence and is contradicted
by the molecular data, it can not be the source of the incongruence between these two major
lines of evidence (Chapter 3). There is little evidence of hybridization among modern species of
Pontederiaceae, but it is possible that undetected ancient hybridization events have contributed to
the observed incongruence between the morphological and chloroplast data (see Doyle 1992).
An improved morphological data-base (or new molecular evidence from the nuclear genome) is
needed to pinpoint the precise source and extent of the incongruence between these two sources
of data. In the meantime, the chloroplast evidence is the only substantial source of evidence
concerning the phylogenetic history of Pontederiaceae, and I use it here to investigate the
radiation of morphological characters in the family.

Given the trees found with the morphological data set of Eckenwalder and Barrett (1986)
as modified in Appendix C, the morphological data have levels of homoplasy very close to that
expected for this number of taxa. The observed CI (excluding autapomorphies) for the 24
ingroup taxa in Fig. 5.1 is 0.474. The expected CI for this number of taxa is 0.495, based on
Sanderson and Donoghue's (1989) survey of 60 data sets. However, the trees derived from the
molecular data indicate that the amount of homoplasy in the morphological data is substantially
higher than this. The CI (excluding autapomorphies) of the morphological data on the four
shortest trees found with the combined chloroplast evidence ranges from 0.404 to 0.410. This
lower value may partly be a reflection of the incongruence between the morphological and molecular sources of data, and given the low number of informative morphological characters (only 33), it is possible that these estimates of homoplasy in the morphological data are a quite imprecise reflection of the real levels of homoplasy in morphological characters. The patterns of homoplasy among subsets of the morphological characters are nonetheless intriguing. On the molecular trees, the CI for 17 informative floral characters (excluding autapomorphies and post-anthesis characters) ranges from 0.446 to 0.453; for 11 informative vegetative characters, the CI (excluding autapomorphies) ranges from 0.310 to 0.316. The trend towards high levels of homoplasy relative to that expected for this number of taxa (Sanderson and Donoghue 1989) in both classes of data suggests that they have been subject to elevated levels of character diversification in this family, at least compared to unrelated groups.

The location of the family's root was one of the few elements of the phylogenetic structure of Pontederiaceae left unresolved by the molecular studies in Chapter 3 and Appendix A. The local position and membership of Pontederiaceae within a cluster of superorders consisting of Arecales, Bromeliaceae, Commelinales and Zingiberales (sensu Dahlgren, Clifford and Yeo 1985) was also not well supported by evidence from rbcL (Chapter 2; Duvall et al. 1993) or surveys of chloroplast restriction-site variation (Davis 1995). A range of taxa in this complex of superorders were surveyed for variation in the chloroplast genes ndhF and evidence from this gene was employed in tandem with the rbcL evidence to determine which taxa are most closely related to the family, and to establish the position of the root of the family (Chapter 4). Using equal-weighting of all characters and a variety of different unequal weighting schemes to correct for among-site rate variation, combined analyses of these two genes indicated that the sister-group of the family is a clade consisting of Commelinaceae and Haemodoraceae, and that the next most closely related clade is Philydraceae (Chapter 4). These three most closely related families (Commelinaceae, Haemodoraceae and Philydraceae) also converged upon a single most-parsimonious root location of Pontederiaceae (Chapter 4; Fig. 4.4A), which is the a
posteriori rooting employed in character reconstructions here (Figs. 5.2-5.6).

Implications of the phylogenetic data and fossil evidence for the biogeography of the family. --
As discussed earlier, the modes of dispersal of taxa in Pontederiaceae provide substantial opportunities for long-range dispersal. With a few isolated exceptions, the basal clades in the family are all currently limited to the New World, and the exclusively Old World genus Monochoria is located quite far from the base of the tree, suggesting an ancient inter-continental dispersal (Fig. 5.1). The fossil record of the family reaches back into the Eocene (c. 50 mya) when Africa, Australia and South America were no longer in direct contact (Briggs 1987). Monochoria is the only genus in Pontederiaceae currently restricted to the Old World (Fig. 5.1). Fossilized seeds and leaf material similar to modern Eichhornia, and seeds like those of modern Monochoria, are known from the upper Eocene onwards in Europe (see Collinson, Boulter and Holmes 1993). The presence of Eichhornia-like fossils in Europe raises the intriguing possibility that the current limitation of this genus to the New World may be a consequence of ancient extinctions in the Old World. Fossilized root and stem fragments of Pontederiaceae are known from the Eocene in India (Patil and Singh 1978). However, the uncertain generic affinity of this material (Eckenwalder and Barrett 1986) means that it is not clear whether this represents a lineage that arose before or after the divergence of the extant members of the family. Other fossils ascribed to the family (Bureau 1892; Knowlton 1922; Fritel 1928) are from North American and European sites, but as these reports are based solely on leaf material, they must also be treated as being of unclear affinity to modern genera. Philydraceae is currently limited to Australia and Asia (Dahlgren, Clifford and Yeo 1985; Adams 1987), but the two families constituting the sister-group (Commelinaceae and Haemodoraceae) are distributed throughout the Old and New World (Dahlgren, Clifford and Yeo 1985; Simpson 1990). The only other species in Pontederiaceae with Old World distributions are Heteranthera callifolia, S. lutea, and E. natans. These three species, together with two species of Monochoria, are limited
to Africa. All of these taxa are missing from the current study, but their taxonomic affinities and positions in the morphology-based analysis of Eckenwalder and Barrett (1986) suggest that their current distributions are a consequence of several long-range dispersal events (Fig. 5.1). However, the possibility that some of these species are relicts from a more cosmopolitan distribution of these genera can not be ruled out, although this is unlikely for *E. natans* given its probable close relationship to *E. diversifolia* (see below). Multiple long-range dispersal events probably also contributed to the modern distributions of the various New World taxa. North American taxa in general seem to be more morphologically apomorphic than those in tropical South America, suggesting that they may have migrated north after intercontinental contact in the Miocene (Eckenwalder and Barrett 1986).

*Character Diversification and Adaptive Radiation in*

*Vegetative and Reproductive Characters*

*Outgroups and their effect on character reconstruction in Pontederiaceae.* -- The inclusion of outgroup taxa in phylogenetic analysis serves two major purposes: locating the position of the root of the ingroup, and polarizing character-state transformations within this group. Although these analytical goals are often addressed simultaneously, they need not be if the characters used to reconstruct a phylogeny (as in this study) are not the ones in whose evolutionary transformation we are interested (e.g., Brooks and McLennan 1991; Maddison and Maddison 1992). The sister-group plays a major role in polarizing character reconstructions, but other less closely related outgroups also play an important role in this (Maddison, Donoghue and Maddison 1984; Nixon and Carpenter 1993).

I employed the three most closely related families to Pontederiaceae to provide information concerning the polarity of the character transformations discussed below: Commelinaceae and
Haemodoraceae (which together constitute the sister-group), and Philydraceae (the next most closely related taxon). In cases where there was character-state variation among the constituent taxa of individual outgroups, knowledge of the internal phylogenetic structure of each outgroup would be valuable for obtaining ‘globally parsimonious’ reconstructions of the evolution of such characters (Maddison, Donoghue and Maddison 1984; Maddison and Maddison 1992, pp. 47). Simpson (1990) provided a phylogeny of Haemodoraceae based on morphological data, but there are no published phylogenies for Commelinaceae or Philydraceae. It was consequently often necessary in this study to code individual families as polymorphic for character-states for which there was known to be variation among different species within each family. Using polymorphic coding to account for this variation is a less-than-ideal solution to lack of knowledge concerning the phylogenetic structure within individual outgroup families (Maddison and Maddison 1992; Nixon and Davis 1991). However, the reconstructions obtained here, while conditional on increased knowledge of the phylogenetic structure of these groups and improved knowledge of character distributions in them, are nonetheless the most-parsimonious ones given our current state of knowledge (Maddison and Maddison 1992, pp. 47).

The reconstructions of character diversification were performed using MacClade version 3.0 (Maddison and Maddison 1992) and employed the four most parsimonious unrooted trees of the family found in the combined analysis of the three chloroplast sources of evidence (Chapter 3; Table 3.2), with the rooting determined using combined evidence from rbcL and ndhF (Chapter 4; Fig. 4.4A). Only minor differences existed among the four trees concerning the placements of P. rotundifolia and Hydrothrix gardneri in Pontederia s.l. and Heteranthera s.l., respectively. All reconstructions were performed using MacClade version 3.0 (Maddison and Maddison 1992) and used ‘Fitch optimization’ (Fitch 1971), that is with all character-state changes treated as equally likely events (i.e., unordered or equally weighted), apart from an analysis of reproductive characters in which ‘relaxed Dollo’ schemes of character evolution
(Swofford and Olsen 1990) were also assessed (see below). With the aid of the ‘equivocal cycling’ tool in MacClade, I obtained counts of the number of gains of each character-state for each character within Pontederiaceae, and determined the primitive state of the family, for all most-parsimonious reconstructions of each character on the four trees. The results are indicated in boxes on each figure. The character reconstructions in Figs. 5.2-5.6 exemplify much of the diversification observed for each character: the tree used in these figures is one of the four most-parsimonious ones, and is the most highly converged-upon tree found in analyses of several different combinations of the available chloroplast evidence (Chapter 3; Fig. 3.1C, 3.3).

**Character Codings**

A total of 24 species were surveyed, representing approximately two-thirds of the family and including all major taxonomic groups (for source and voucher information see Appendix B). Except for leaf developmental pathway, the character codings for the taxa of Pontederiaceae considered here are derived from Eckenwalder and Barrett (1986), Chapter 2 and Appendix A. Codings for the three outgroup taxa I included are presented here. Of the 50 genera in Commelinaceae, a few are found in wet places, but only *Murdannia* possesses aquatic species (at least two of 50 species; Cook 1990). Cook (1990) lists only one species of Philyraceae as being helophytic (*Philydrum lanuginosum*), but all six species in the family are found in marshes and wet rainforest habitats (Adams 1987). Species of Haemodoraceae are almost all xeric (M.G. Simpson, per. comm.), although *Tribonanthes* is found in similar habitats (low, winter-wet flats) to *Philydrella* (Philyraceae) (Simpson 1990). Commelinaceae and Haemodoraceae are therefore almost exclusively terrestrial groups. The three families most closely approach the ‘emergent’ condition in Pontederiaceae. Procumbent and erect life-forms are known in Commelinaceae (Faden 1988), and so this outgroup is coded as polymorphic for
these two forms. Haemodoraceae and Philydraceae contain only erect taxa (M.G. Simpson pers. comm.; Adams 1987; Cook 1990) and these families are coded accordingly for life-form.

For life-cycle duration, Commelinaceae and Philydraceae are both coded as polymorphic for annuality and short and long-lived perenniality (Dahlgren, Clifford and Yeo 1985; Faden 1988; Cook 1990). A ‘long-lived perennial’ coding is appropriate for Haemodoraceae (Simpson, pers. comm.). Non-clonal species are found in Commelinaceae (Faden 1988), and some species in this family express clonality via rhizomes, stolons or spreading-stem fragmentation (Faden 1988). Commelinaceae is coded as polymorphic for ‘non-clonality,’ and for these three kinds of clonality (‘via rhizomes,’ ‘via stolons,’ and ‘via stem fragmentation’). There is no direct evidence of clonality in Haemodoraceae, but extensive underground rhizome/stolon systems are known in this family (M. G. Simpson, pers. comm.). Haemodoraceae is provisionally coded as polymorphic for ‘non-clonality,’ ‘clonality via rhizomes,’ and ‘clonality via stolons.’ Species of Philydraceae are rhizomatous or cormous (Adams 1987), but it is not known if these structures are involved in regeneration. Philydraceae is provisionally coded as polymorphic for ‘non-clonality’ and ‘clonality via rhizomes.’

Information on the timing of the transition to adult leaves, and on the homology of such pathways among the outgroup families and Pontederiaceae is mostly lacking. However, Tillich (1994, 1995) noted that the seedlings of Pontederiaceae and Philydraceae are very alike and that the primary leaves in both families are ribbon-like. The homology of adult leaf-form among these families is uncertain. Members of Commelinaceae possess bifacial leaves, but Haemodoraceae and Philydraceae possess unifacial, ensiform leaves (Dahlgren and Rasmussen 1983). Anatomical data suggest that the bifacial leaves typical of taxa in Pontederiaceae have a unifacial origin (see Arber 1925), and Simpson (1990) hypothesized that an origin of the bifacial leaf was associated with the shift to an aquatic environment in Pontederiaceae. I provisionally coded the three families as ‘unknown’ for leaf developmental pathway.
Showy, insect-pollinated flowers are typical of species in Commelinaceae, Haemodoraceae and Philydraceae. Species lacking either somatic or genetic polymorphisms in styal class (referred to here as 'monomorphic flowers') predominate in Commelinaceae and Haemodoraceae and in most monocotyledons. Tristyly is only known in Pontederiaceae, but enantiostyly is found in some species of Commelinaceae (Faden 1991) and Haemodoraceae (Simpson 1990) and all species of Philydraceae (Simpson 1990). I did not count as enantiostylosous those species that possess flowers with bent styles but that lack a 'handedness,' such as Hydrothrix gardneri (Rutishauser 1983) and H. dubia in Pontederiaceae. Species in Barberetta, Schiekia and Wachendorfia and Xiphidium in Haemodoraceae possess flowers with bent styles that have true left- versus right-handedness, i.e., with zygomorphic or outward-facing flowers (Simpson 1990). The non-basal position of these taxa in morphology-based cladograms of Haemodoraceae (Simpson 1990) suggests that monomorphic flowers were ancestral in this family. Haemodoraceae was therefore coded as 'monomorphic' for floral form, Philydraceae as 'enantiostylous,' and Commelinaceae as polymorphic for these two conditions.

Owens (1981) reported self-compatible and self-incompatible species in Commelinaceae. The form of self-incompatibility in Commelinaceae is gametophytic and non-heteromorphic and is consequently highly unlikely to be homologous to that in Pontederiaceae (Chapter 2). This family was therefore coded as polymorphic for gametophytic SI and self-compatibility. The self-incompatibility status of species in the two other outgroup families is largely unknown. Philydrum lanuginosum is fully autogamous (S. C. H. Barrett, per. obs.). Hamann (1966) also reported autogamy in the family, but no explicit surveys for self-incompatibility have been performed. Philydraceae is therefore coded as 'unknown' for this character. There is a single report of a weakly developed incompatibility system in Wachendorfia paniculata that appears to be associated with true genetic enantiostyly, in a manner analogous to distylos or tristylos self-incompatibility systems (Ornduff and Dulberger 1978; see also Wilson 1887). However, I am aware of no other data concerning the self-incompatibility status of species in
HAEMODORACEAE, AND SO CODE IT AS 'UNKNOWN' FOR THIS CHARACTER.

Reconstructions of Character Evolution

Aquatic habit.-- Is the capacity to thrive in an aquatic habitat a synapomorphy for the taxa in Pontederiaceae? Of the taxa that are closely related to Pontederiaceae, the two families that constitute its sister-group (Haemodoraceae and Commelinaceae; see above) are almost exclusively terrestrial. For each of these families, it is almost certain that the primitive forms were adapted to a completely terrestrial existence (see above). However, the next most closely related clade (Philydraceae) is semi-aquatic. Whether the aquatic habit is homologous in Philydraceae and Pontederiaceae depends partly on the distribution of aquatic versus terrestrial taxa in clades more distant to Pontederiaceae than its sister-group and Philydraceae. If it is assumed that these more distant taxa are exclusively terrestrial, then the parsimony criterion indicates that the aquatic habit either arose independently in the two families, or it arose prior to the origin of Commelinaceae, Haemodoraceae, Philydraceae and Pontederiaceae, with a subsequent loss prior to the split of Commelinaceae and Haemodoraceae. If more distant clades than Philydraceae are aquatic, then homology of the aquatic habit between the two families is more parsimonious than non-homology. In either case, homology of the aquatic habit between Philydraceae and Pontederiaceae is a definite possibility. However, aquatic adaptations in Pontederiaceae are much more complete and diverse than in Philydraceae, indicating that the majority of the character diversification constituting this radiation has taken place in this family alone.

Life form. -- An emergent, erect habit is reconstructed as the primitive condition in Pontederiaceae for all four shortest trees (see the branch in Fig. 5.2 that connects the outgroup
taxa to Pontederiaceae). The free-floating form typical of *E. crassipes* arose directly from the emergent, erect habit, and the floating-leaved form arose from an emergent, erect or an emergent, procumbent form. In *Eichhornia* and *Pontederia* the emergent, procumbent form arose from an emergent, erect form on one or two occasions, although an origin of the former condition from a floating-leaved form in *Eichhornia* was also possible. Surprisingly, there was no phylogenetic record of any transitional forms between the emergent, erect and submersed life forms in *Heteranthera*. There were up to three independent origins of the latter from the former in this genus. There were no parsimonious reconstructions for any of the trees in which the emergent, procumbent habit was homologous between *Heteranthera* versus *Pontederia* and *Eichhornia* (Fig. 5.2). There were two independent origins of this life form in *Heteranthera* s.l. under all reconstructions, and up to two independent origins of this form in species of *Pontederia* and *Eichhornia*. Under some of the most-parsimonious reconstructions, the emergent, erect life form within *Heteranthera* and *Pontederia* represented a reversion from a submersed or emergent, procumbent form. The emergent, erect habit in *Heteranthera* is thus also potentially not homologous with the occurrence of this life form outside the genus (Fig. 5.2).

A number of taxa that are currently missing from the phylogenetic estimate of Pontederiaceae are likely to make an impact on future reconstructions of this character. These include several emergent, procumbent taxa in *Heteranthera* (*H. reniformis* and allies) and two emergent, erect species (*Heteranthera spicata* and *H. mexicana*). One missing species with a floating-leaved life form is *S. lutea*. This species probably belongs in *Heteranthera* s.l. (Eckenwalder and Barrett 1986) and consequently may well represent an additional origin of the floating-leaved form in the family. Inclusion of these taxa may also indicate that some or all of the submersed taxa in *Heteranthera* did not arise directly from emergent, erect forms, a finding that in any case seems to us to be biologically implausible. The other missing floating-leaved
Fig. 5.2. Reconstruction of diversification in aquatic life-form in Pontederiaceae and its closest relatives (Commelinaceae, Haemodoraceae and Philydraceae). The ingroup tree is one of four most-parsimonious trees found using the combined chloroplast evidence from *rbcL*, *ndhF* and a survey of restriction-site variation. The root of the ingroup tree is that indicated by a variety of closely related taxa using the two chloroplast DNA sequence-based data sets. See text for a description of the character-states and codings. Commelinaceae was coded as polymorphic for this character (see text). The reconstructed numbers of gains in each character-state are indicated in the box. They refer to changes within Pontederiaceae (not the whole tree). These values were the same across all four shortest trees.
form is *E. natans*, but as this species appears to be very closely related to *E. diversifolia* (Verdcourt 1968; Eckenwalder and Barrett 1986), it probably does not represent a novel origin of this life form.

**Duration.** — Despite uncertainty at the base of the tree, with a variety of possible most-parsimonious reconstructions of shifts in life-history among and within the outgroup taxa, an annual life history was reconstructed as the primitive condition of Pontederiaceae for all shortest trees (Fig. 5.3). Under all most-parsimonious reconstructions on these trees, the annual life-history was homologous for all species exhibiting this life-history. It should be borne in mind, however, that many of the species that I coded as annual are capable of growing as short-lived perennials and are therefore 'facultative' annuals. In contrast, species coded as short-lived perennials are incapable of an annual existence.

Most workers assume that in herbaceous groups annuals are derived from perennials (Stebbins 1974), a shift in life-history that has normally been invoked in the context of adaptive radiations from mesic to arid environments. However the reverse shifts may well have occurred in Pontederiaceae during invasion of permanent aquatic habitats. Such environments require specialized aquatic adaptations and may have represented relatively unsaturated niches. Under these circumstances, lack of species competition could have aided an evolutionary transition that for terrestrial groups occurs less frequently, a situation that may be analogous to the evolution of perenniality in island floras (see Carlquist 1974).

Long-lived perennials arose two or three times in the family, and long-lived perenniality in *H. dubia* was not homologous with other species in the family for any most-parsimonious reconstruction. Instances of long- and short-lived perenniality in *Monochoria* were homologous with the occurrences of these forms in *Eichhornia* and *Pontederia* in some reconstructions on
Fig. 5.3. Reconstruction of diversification in life-cycle duration in Pontederiaceae and its closest relatives (Commelinaceae, Haemodoraceae and Philydraceae). The ingroup tree is one of four most-parsimonious trees found using the combined chloroplast evidence from *rbcL*, *ndhF* and a survey of restriction-site variation. The root of the ingroup tree is that indicated by a variety of closely related taxa using the two chloroplast DNA sequence-based data sets. See text for a description of the character-states and codings. Commelinaceae and Philydraceae were coded as polymorphic for this character (see text). The reconstructed numbers of gains in each character-state are indicated in the box. They refer to changes within Pontederiaceae (not the whole tree). These values were the same across all four shortest trees.
Philydracaeae
Haemodoraceae
Commelinaceae

P. cordata var. ovalis
P. cordata var. wrdata
P. cordata var. landolia

Outgroups (unordered)
the shortest trees but not in others (e.g. Fig. 5.3). Short-lived perenniality in *E. diversifolia* and *H. zosterifolia* arose uniquely in the terminal lineages leading to these species. Missing perennial taxa from this study include *H. reniformis* and its allies in *Heteranthera*, and several species of *Monochoria* and *Pontederia s.l.*

**Clonality.** A non-clonal form is reconstructed as the most primitive condition in *Pontederiaceae*. Vegetative reproduction via stolons (typical only of *E. crassipes*) arose directly from this primitive form, as did the instance of clonal reproduction via rhizomes in *M. hastata*, which consequently must have arisen independently from the other instances of this clonal form in *Eichhornia* and *Pontederia*. Species of *Pontederia* and *Eichhornia* that express clonality via stem fragmentation can also clone via rhizomes (I scored them using the former coding only), so the question of the number of origins or interconversions between these two clonality modes in the local part of the tree containing these taxa is somewhat moot. Examination of the most-parsimonious reconstructions on the four shortest trees indicates that clonality via stem-fragmentation arose independently in *Heteranthera* versus *Pontederia* and *Eichhornia*, and may have arisen from one to three (Fig. 5.4) times in the former genus. Several missing taxa of *Pontederiaceae* in this study that are clonal include *H. reniformis* and its allies (which all express clonality via stem fragmentation) and several species in *Pontederia s.l.* that can reproduce via rhizomes or stem fragmentation. The only taxa to reproduce via fragmentation of creeping stems are those with a growth form that could broadly be described as procumbent (emergent, submersed or floating-leaved).

**Leaf developmental pathway.** Pattern A, with a rapid transition to adult leaves, is the primitive form in the family, although it is also possible that it also arose once by reversion from pattern
Fig. 5.4. Reconstruction of diversification in clonality in Pontederiaceae and its closest relatives (Commelinaceae, Haemodoraceae and Philydraceae). The ingroup tree is one of four most-parsimonious trees found using the combined chloroplast evidence from \textit{rbcL}, \textit{ndhF} and a survey of restriction-site variation. The root of the ingroup tree is that indicated by a variety of closely related taxa using the two chloroplast DNA sequence-based data sets. See text for a description of the character-states and codings. The three outgroup taxa were coded as polymorphic for this character (see text). The reconstructed numbers of gains in each character-state are indicated in the box. They refer to changes within Pontederiaceae (not the whole tree). Where numbers of gains differ across the four shortest trees, this is indicated in square brackets.
C or D in the terminal lineage leading to *Heteranthera oblongifolia* (Fig. 5.5). Pattern B, in which the production of petiolate leaves is more delayed than in pattern A, evolved once from pattern A in *Eichhornia*, supporting the idea that this represents a neotenic shift. The suggestion that the other two patterns (C and D, which possess adult leaves resembling the juvenile leaves of pattern A) arose by various paedomorphic processes was supported by only some of the reconstructions on the shortest trees. A variety of shifts between patterns A, C, D and E were seen in different most-parsimonious reconstructions on the four shortest trees: pattern D arose from patterns A or C in different reconstructions, pattern C arose from patterns A, D or E (with either one or two origins), and the leaf form unique to *Hydrothrix gardneri* (pattern E, with whorls of highly reduced leaves) could have arisen from patterns A or C. To my knowledge, all of the taxa missing from the current study exhibit pattern A, apart from *H. mexicana* (pattern D) and *E. natans* (pattern B; probably homologous with the other three instances of this pattern). It is likely that an increased sampling of taxa within *Heteranthera s.l.* would provide a less equivocal reconstruction of the evolution of patterns C to E, since much of the variation in leaf developmental pathway is in this genus.

*Floral form and self-incompatibility.* -- I considered two types of evolutionary scheme for the shifts in floral form and self-incompatibility status: Fitch optimization, in which all character-state shifts are equiprobable and there is no implied order of change; and ‘relaxed Dollo’ schemes that weight against the origin of self-incompatibility and the two polymorphic floral forms in the family, enantiostyly and tristyly. A large body of comparative and microevolutionary evidence indicates that evolutionary gains of tristyly are much more difficult than their loss (Chapter 2, Appendix A), and it seems likely that the same is also true for self-incompatibility systems. Schemes that disfavor the origin of such complex characters are probably more biologically and historically accurate than those that weight all character-state shifts equally.
Fig. 5.5. Reconstruction of diversification in leaf developmental pathway in Pontederiaceae and its closest relatives (Commelinaceae, Haemodoraceae and Philydraceae). The ingroup tree is one of four most-parsimonious trees found using the combined chloroplast evidence from *rbcL*, *ndhF* and a survey of restriction-site variation. See text for a description of the character-states and codings. The root of the ingroup tree is that indicated by a variety of closely related taxa using the two chloroplast DNA sequence-based data sets. The three outgroup taxa were coded as 'unknown' (missing data) for this character (see text). The reconstructed numbers of gains in each character-state are indicated in the box. They refer to changes within Pontederiaceae (not the whole tree). These values were the same across all four shortest trees.
LEAF DEVELOPMENTAL PATHWAY
(Unordered)

- Pattern A: 0-1 gain
- Pattern B: 1 gain
- Pattern C: 1-2 gains
- Pattern D: 1 gain
- Pattern E: 1 gain
- Equivocal reconstruction
CHAPTER 5. ADAPTIVE RADIATION IN PONTEDERIACEAE

The choice of weighting scheme can have a profound influence on reconstructions of character evolution (compare Fig. 5.6A and 5.6B). The scheme that weights all character shifts equally (Fitch optimization) indicates a wide possible range of gains in each floral form, including up to four origins of tristyly (Fig. 5.6B). Not surprisingly, the scheme that weights gains of the floral polymorphisms more heavily that their reversion to monomorphism indicates a single origin of tristyly and up to two origins of enantiostyly, but multiple (four to six) origins of floral monomorphism in the family (Fig. 5.6B). Enantiostylosous flowers in *Monochoria* and *Heteranthera* were not homologous under either scheme, for any of the most-parsimonious reconstructions on the shortest trees. Floral monomorphism and enantiostyly were the primitive floral forms in the family in different most-parsimonious reconstructions (using the Fitch or relaxed-Dollo optimization schemes). Of course, the precise weights that correspond to the actual probabilities of change in these floral forms are unknown, but only very small weighting biases (around 3:2 to 2:1; Chapter 2, Appendix A) were required to reconstruct a single origin of tristyly in the family.

The reconstructed shifts in self-incompatibility (SI) status indicates that it arose at most twice in the clade consisting of *Pontederia* s.l. and the species of *Eichhornia* associated with *E. azurea*. One most-parsimonious reconstruction is shown in Fig. 5.6B under a small weighting bias. Under equal-weighting (Fitch optimization), or unequal-weighting (2:1 bias against the origin of SI with DELTRAN optimization), two origins of SI are indicated on the tree (see Appendix A). Under unequal-weighting (2:1 bias with ACCTRAN optimization), a single origin and two losses of SI are implied (Fig. 5.6B, and see Appendix A). In either case, SI must have arisen subsequent to the origin of tristyly, given reconstructions of floral form that indicate a single origin of tristyly.

This evolutionary sequence casts doubt on models of the evolutionary origin of heterostyly (Charlesworth and Charlesworth 1979; Charlesworth 1979) in which an origin of
Fig. 5.6. Reconstruction of diversification in floral-form in Pontederiaceae and its closest relatives (Commelinaceae, Haemodoraceae and Philydraceae). The ingroup tree is one of four most-parsimonious trees found using the combined chloroplast evidence from \textit{rbcL}, \textit{ndhF} and a survey of restriction-site variation. The root of the ingroup tree is that indicated by a variety of closely related taxa using the two chloroplast DNA sequence-based data sets. \textit{Monochoria cyanea} was coded as 'enantiostylous or monomorphic,' and Commelinaceae was coded as polymorphic for this character (see text for descriptions of the character-states and codings). The reconstructed numbers of gains in each character-state are indicated in the box. They refer to changes within Pontederiaceae (not the whole tree). Where numbers of gains differ across the four shortest trees, this is indicated in square brackets. A. Reconstructions of shifts in floral form, performed using Fitch optimization. B. Reconstructions of shifts in floral form and self-incompatibility (SI) status performed using a 'relaxed Dollo' weighting scheme described in the step-matrix beside the tree. Gains of either polymorphic form (tristyly or enantiostyly) were coded more heavily (3:2 weighting, see step matrix) than reversions to floral monomorphism. Gains of SI were coded more heavily (2:1, ACCTRAN optimization, see step matrix) than reversions to self-compatibility. Abbreviations: GSI = gametophytic self-incompatibility; Het SI = heteromorphic, sporophytic self-incompatibility.
<table>
<thead>
<tr>
<th>Philydraceae</th>
<th>Haemocarpaceae</th>
<th>Commenaceae</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. corda ta</em></td>
<td><em>P. sagittata</em></td>
<td><em>P. cordata</em></td>
</tr>
<tr>
<td><em>P. cordata v. ovalis</em></td>
<td><em>P. cordata v. ovata</em></td>
<td><em>P. cordata v. lancifolia</em></td>
</tr>
</tbody>
</table>

OUTGROUPS

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

FLORAL FORM (Unguiculate Wedge)
self-incompatibility is required prior to the origin of floral heteromorphisms (Table 1.1; see also Chapter 2 and Appendix A). Floral shifts between the three floral types are undoubtedly associated with shifts in pollination mode. Tristylos species of *Eichhornia* and *Pontederia* are predominantly pollinated by nectar-collecting bees, enantiostylous species of *Monochoria* and *Heteranthera* are pollinated by pollen-collecting bees, and monomorphic species throughout the family are predominantly self-pollinating (see above). The reconstructions also indicate that at least some predominantly selfing lineages of *Eichhornia* have existed for substantial evolutionary periods and were even capable of speciation (see also Appendix A), a finding at odds with Stebbins' (1957) view that selfing species are evolutionary dead-ends. Missing taxa that could influence reconstructions of character evolution include *E. natans* and *P. parviflora* (both monomorphic and presumably self-compatible).

**Summary**

There is substantial ecological evidence that a number of life-history traits in Pontederiaceae are involved in or affected by an aquatic existence, and that a variety of reproductive characters have undergone diversification in response to shifts in their mode of pollination. I used currently available molecular evidence of phylogenetic relationships within this family to reconstruct patterns of character diversification associated with adaptive radiations in vegetative and floral characters. Shifts in pollination mode, particularly those resulting in predominant self-pollination, occurred on multiple occasions in the family. Vegetative characters in the family are particularly prone to convergence (see also Eckenwalder and Barrett 1986). The extensive homoplasy in vegetative characters suggests that aquatic habitats are far from ecologically uniform, as has often been supposed. As with all phylogenetic analyses, these findings are liable to new interpretations when more taxa are sampled inside and outside
the family, and when new sources of phylogenetic evidence become available. More detailed developmental and ecological work is also needed to determine how plastic some of these character classes are in different taxa. However, several of the reconstructions of character diversification in Pontederiaceae challenge widely held views on the course of plant evolution. These include the shift from the annual to the perennial habit, the evolutionary longevity of some predominantly selfing lineages and the sequence in which morphological and physiological traits became associated in the heterostylos syndrome. In challenging these orthodox views, I hope that these analyses may help to provoke future research on these topics in Pontederiaceae and other aquatic plant families.
CHAPTER 6

A Phylogeny of Narcissus L. (Amaryllidaceae) Based on the Chloroplast Gene ndhF, and Its Implications for Breeding-System Evolution in the Genus
CHAPTER 6. PHYLOGENY OF NARCISSUS

INTRODUCTION

*Narcissus* includes all the species known as daffodils (Blanchard 1990). All members of the genus are perennial geophytes. They are concentrated geographically in the Iberian peninsula and Morocco, but the genus extends around the Mediterranean sea and its islands, out along the Atlantic coasts of Europe and North Africa, and across to the Canary Islands (reviewed in Fernandes 1951; Blanchard 1990). Several taxa reach northwards to Great Britain and eastwards into Asia, as far as Japan, but the pre-modern status of these populations is doubtful (Webb 1980, Blanchard 1990). Various species of *Narcissus* have been in cultivation for at least four centuries. *Narcissus cyclamineus*, for example, was recorded in cultivation in the early seventeenth century (Vallet 1608, cited in Blanchard 1990), and other species were probably valued in earlier centuries for their aesthetic and medicinal properties (Jefferson-Brown 1991).

This purpose of this study is to provide a new phylogenetic context for breeding-system variation in the genus. *Narcissus* contains a rich array of floral variation within and among different species. However, despite at least a century of breeding effort, knowledge of the reproductive biology of most taxa in this genus is still fairly meager (Barrett, Lloyd and Arroyo 1995). Considerable information is available concerning variation in chromosome number and shape in the genus. Data on chromosomal information is interesting from a biosystematic point of view, and may provide insights into phylogenetic relationships in the genus. In the following sections I review the taxonomy of the genus and what is known about its phylogenetic position in Amaryllidaceae. I then review what is currently known concerning chromosomal variation, and variation in the breeding system in the genus. The remainder of the chapter presents phylogenetic evidence of relationships in the family based on the chloroplast
gene ndhF. I use this preliminary phylogeny to reconstruct the evolution of breeding-system variation in the genus.

**Taxonomy of Narcissus.** -- Estimates of the number of species in the genus have varied widely and the taxonomy of the genus is still quite unsettled. Early workers recognised up to 160 species in six to sixteen genera (Haworth 1831; Herbert 1837: cited in Blanchard 1990) or as few as sixteen species (Baker 1888). More recent treatments divide the genus into nine or ten sections. Fernandes (1968a) recognised two sub-genera containing a total of ten sections and 61 species between them, while Webb (1980) recognised 26 species in the sections found in Europe. Much of the difference between Webb’s and Fernandes’ treatments lies in the rank they gave to a number of closely related taxa, particularly those allied to *Narcissus pseudonarcissus*, *Narcissus papyraceus* and *Narcissus tazetta*. Webb reduced many of these taxa below specific rank. I follow the example of Blanchard (1990) and largely employ Fernandes’ (1968a) basic scheme in this paper.

Species delimitation remains problematical in several cases, but ascribing taxa to particular sections is usually straightforward. Of the sections containing more than one species, *Bulbocodii* (the hoop petticoats), *Jonquillae* and *Apodanthae* (the jonquils), *Narcissus* (the poets), *Pseudonarcissi* (the wild trumpet daffodils), and *Tazettae* (which includes the paper whites) are fairly well circumscribed. However, there are several cases where, on closer inspection, sectional boundaries appear less well-defined. For example, the multiflowered members of *Apodanthae* approach *Tazettae* (Blanchard 1990), and *Apodanthae* may itself deserve to be collapsed within *Jonquillae* (Webb 1980). Many important morphological features used to define particular sections are notoriously variable within sections, and in many cases even within individual species (Fernandes 1968a; Blanchard 1990). These include flower number and colour, flowering season, corona and floral-tube length and shape, extent of scape
CHAPTER 6. PHYLOGENY OF NARCISSUS

compression, and leaf width and degree of glaucousness. A thorough examination of morphological and anatomical variation in the genus is currently lacking, and little effort has been made to place the known variation in the context of related genera.

Phylogenetic position of Narcissus in Amaryllidaceae. -- Dahlgren, Clifford and Yeo (1995) indicated that tribe Galantheae (the snowdrops and snowflakes, Galanthus and Leucojum) might be united with tribe Narcisseae (consisting of Narcissus, Tapeinanthus and perhaps Sternbergia), although the former tribe may be distinguished by anther dehiscence via apical pores rather than slits. Relationships of Narcisseae with other tribes in Amaryllidaceae have not been proposed (Meerow 1995). Traub (1969) transferred the monotypic Tapeinanthus to Narcissus [under the name Narcissus humilis (Cav.) Traub (= N. cavanillesii A. Barra & G. López)], apparently on the basis of Fernandes and Fernandes' (1945) chromosomal evidence. Chromosomal evidence also supports a possible transfer of Sternbergia to Narcisseae (Flagg and Flory 1962). Sternbergia and section Hermione of Narcissus have similar base chromosome numbers [x = 9-11 (see Webb 1980) in the former case, possibly representing a duplication plus aneuploid shifts from a base of x = 5, which is the base number in the latter case (Fernandes 1967)]. They also both have "cephalobrachial" (short acrocentric to telocentric) elements (Fernandes 1967; Flagg and Flory 1962; Bedalov and Susnik 1970), and similar sized chromosomes (Flagg and Flory 1962). Preliminary analyses of rbcL data from Amaryllidaceae support a close relationship between Narcissus and Sternbergia of tribe Narcisseae, and between that tribe and Galanthus and Leucojum of tribe Galantheae (A. Meerow, pers. comm.).
CHAPTER 6. PHYLOGENY OF NARCISSUS

Variation in breeding-systems in Narcissus and related genera

Variation in Narcissus. -- The flowers of Narcissus are long-lasting and hermaphroditic, and dichogamous to varying degrees (both protogynous and protandrous taxa are known; Barrett, Lloyd and Arroyo 1996). They are typically herkogamous (i.e., with stigma and anthers at different heights), although a few species are known in which the stigma is at the same level as the anthers (e.g., Narcissus viridiflorus and Narcissus elegans). Fernandes and earlier workers recognised the existence of tristyly in the genus (reviewed in Lloyd, Webb and Dulberger 1990; Barrett, Lloyd and Arroyo 1996). However, heterostyly is now recognised only in N. triandrus, because only this species possesses sufficiently reciprocal positioning of the anthers and stamens of different morphs (Barrett, Lloyd and Arroyo 1996).

Many species in sections Apodanthae, Jonquillae and Tazetae are dimorphic for stigma height. Populations of these taxa contain a mix of individuals whose flowers are either "approach" or "reverse" herkogamous, i.e., the stigma is respectively beyond or below the level of the anthers (see Webb and Lloyd 1986). Although there are some indications in dimorphic taxa of shifts in anther-height position towards reciprocity of sex-organ position, the distances involved are relatively small (Barrett, Lloyd and Arroyo 1996), so that these species can not be viewed as being distylous. Records of authentic non-heterostylous stigma-height polymorphisms are extremely rare in the angiosperms. When they do arise, they are thought to typically represent short-lived, intermediate stages on the road to complete reciprocal herkogamy (Lloyd and Webb 1992a). In Narcissus, the genetic control of the stigma-height dimorphism is apparently quite simple. In N. tazetta the short-styled morph is the result of a single dominant allele, with the long-styled morph a homozygous recessive genotype (Dulberger 1964).

Self-sterility appears to be the typical condition in Narcissus, suggesting that it may indeed be primitive within the genus. Self-sterility is reported from taxa in the following
sections: *Apodanthae, Bulbocodi, Ganymedes, Narcissus, Pseudonarcissi* and *Tazettae* (Bateman 1954; Dulberger 1964; Barrett et al. 1996). To date, fully self-compatible species have been reported from only two sections: *Pseudonarcissi* and *Serotini* (Herrera 1995; Barrett, Lloyd and Arroyo 1996). However, not all species in the genus have been surveyed for their self-sterility status, and it is not clear whether the mechanism of self-sterility is homologous across all species in the genus. In the dimorphic and tristylos species, different individuals with the same style morph are cross-compatible, and so floral heteromorphisms in *Narcissus* are not associated with the dialellic or trimorphic incompatibility systems typical of most heterostylos groups (Bateman 1952; Dulberger 1964; Barrett, Lloyd and Arroyo 1996). Self and outcross pollen tubes grow at the same rate in styles of *Narcissus* (Bateman 1954). Self-sterility in the genus appears instead to be a consequence of events that occur later in the reproductive process, as demonstrated by cytological studies of fertilization in *N. tazetta* (Dulberger 1964) and *N. triandrus* (T. Sage, pers. comm.). One consequence of the apparent late-acting self-sterility is that self-pollination in *Narcissus* can result in the costly loss of ovules, possibly due to their penetration by self-pollen tubes (T. Sage, pers. comm.). It has been proposed that avoiding wastage of selfed ovules was an important force in the origin of stigma-height polymorphisms in the genus (Barrett, Lloyd and Arroyo 1996).

**Variation in related taxa.** -- Information concerning the reproductive biology of *Sternbergia, Leucojum* and *Galanthus* is limited. Species in *Leucojum* and *Galanthus* are approach herkogamous (Stern 1956), as is *Sternbergia clusiana* (Dafni and Werker 1982). The self-compatibility or sterility status of related genera is more incomplete. In *Sternbergia clusiana* self-pollination yields degenerate, few-seeded capsules (Dafni and Werker 1982). However, the mechanism of this apparent self-sterility is not known. Taxa from *Galanthus* and *Leucojum* are reported as being self-compatible (Clapham, Tutin and Warburg 1962), although there are no experimental data to back up this claim.
Chapter 6. Phylogeny of Narcissus

Biosystematic information from chromosomal variation in Narcissus and related genera.

Variation in chromosome number. -- Fernandes documented chromosome numbers over the course of more than four decades of work (summarized in Fernandes 1967, 1968a, 1968b, 1975; Appendix G). The two sub-genera of Narcissus are defined by taxa that possess quite distinct chromosome shape and number (Fernandes 1975, Brandham and Kirton 1987). Most taxa in sub-genus Hermione have chromosome counts of $2n = 20$ or 22, whereas the most common chromosome count in sub-genus Narcissus is $2n = 14$, rising to an octaploid level of $2n = 8x = 56$. In his earlier work Fernandes (1951) postulated that the base chromosome number in the genus was $x = 7$ and that the common chromosome number of $2n = 20$ or 22 in N. tazetta and its relatives was derived by diploidization of a $3n = 21$ triploid. However, the important discovery of a form of N. serotinus (sub-genus Hermione, section Serotini) with $2n = 10$ (Fernandes 1967), indicated instead that chromosome numbers of $2n = 20$, 22, 30 and 44 were derived by polyploidy from a base number of $x = 5$ or 6 in this sub-genus (Fernandes 1975).

Fernandes (1975) left partially unresolved what the base chromosome number of Narcissus must be, but favoured a base number of $x = 6$, from which $n = 5$ and 7 were both derived. The most common chromosome number in the family is $2n = 22$ (Traub 1963; Goldblatt 1976; Flory 1977). Meerow (1995) suggested that the base number in Amaryllidaceae as a whole is $x = 11$. However, several tribes in Amaryllidaceae, including Galantheae, have some taxa with base numbers of $x = 6$ or 7 and Traub (1963) concluded that the base chromosome number of the entire family was $x = 6$. Reductions in chromosome number require one evolutionary event for each single-chromosome reduction, whereas a doubling in chromosome number can be achieved through a single polyploidization event. It consequently seems more likely that the base chromosome number in the Amaryllidaceae is towards the lower of these two proposals, since if the reverse were true, this would require
CHAPTER 6. PHYLOGENY OF NARCISSUS

multiple gradual reductions in chromosome number from \( n = 11 \) in the family. If the base chromosome number of the immediate ancestor of Narcissus was \( x = 11 \), for example, this would require numerous aneuploid reductions in the genus to generate \( x = 7 \) in sub-genus Narcissus and the \( x = 5 \) in sub-genus Hermione. However, a definitive test of these hypotheses awaits a robust phylogeny of Amaryllidaceae.

_Variation in chromosome shape._ -- Fernandes also reported extensively on the gross shapes of chromosomes in different taxa (summarized in Appendix G; and see below). This karyotypic data constitutes another possible source of systematic information for the genus. However, in the absence of data concerning homology of chromosomes or chromosome segments among different taxa, this limits the use of this data in assessing chromosomal evolution or phylogenetic relationships in the genus. Nonetheless, in a general sense taxa with identical or nearly identical karyotypes seem more likely to be closely related than those with more dissimilar karyotypes.

,Objectives of the study._ -- Apart from Fernandes' chromosomal work, there is virtually no evidence concerning how the different taxa in Narcissus are related to each other. The purpose of this study is to produce new evidence concerning systematic relationships within Narcissus, and to employ this phylogenetic information to assess several hypotheses concerning floral evolution in the genus. Fernandes (1975) hypothesized that stigma-height monomorphism was the basal condition of Narcissus. Barrett, Lloyd and Arroyo (1996) suggested that tristyly in _N. triandrus_ was derived from a stigma-height dimorphism, rather than directly from a taxon monomorphic for style length. Because it has been hypothesized that stigma-height dimorphisms are in general a short-lived transition during the evolution of heterostyly (Lloyd and Webb 1992a), it would also be useful to know whether all instances of this polymorphism
CHAPTER 6. PHYLOGENY OF NARCISUS

In Narcissus are homologous. If this were so, it would imply that the polymorphism need not be short-lived. The chloroplast gene \textit{ndhF} can provide phylogenetic resolution at the generic level and below (e.g., Chapter 3). I present the results of a phylogenetic survey of this locus in \textit{Narcissus} and compare this to the biosystematic information provided by Fernandes' karyotypic surveys. I then use the chloroplast evidence to investigate which stylar condition was ancestral, whether tristyly evolved via an intermediate dimorphic stage, and to estimate the number of origins of stigma-height dimorphism in the genus.

MATERIALS AND METHODS

Twenty-nine species from the ten sections of \textit{Narcissus} recognised by Fernandes (1968a, 1975) were examined for DNA sequence variation in a portion of the chloroplast gene \textit{ndhF}. Collection details are provided in Appendix F. Species from the following sections were represented: five species from \textit{Apodanthe}, three from \textit{Bulbocodium}, five from \textit{Jonquillae}, two from \textit{Narcissus}, five from \textit{Pseudonarcissus} and five from \textit{Tazettae}. Species in the four monotypic sections of the genus (\textit{Aurelia}, \textit{Ganymedes}, \textit{Serotina} and \textit{Tapeinanthes}) were each represented by two or three collections. Sequences from \textit{ndhF} were also obtained from outgroup species representing three genera (\textit{Sternbergia}, \textit{Galanthus} and \textit{Leucojum}) in Amaryllidaceae related to \textit{Narcissus}. Experimental protocols for DNA extraction, amplification and sequencing are provided in Appendix H. In two cases (position 484 in \textit{N. cuatrecasasii} and position 163 in the Mangualde population of \textit{N. triandrus}), the forward and reverse strands had different nucleotides (C and A in the former case, C and T in the latter), indicating within-population polymorphisms. Both polymorphisms involved unique (autapomorphic) nucleotides. The polymorphisms were not re-confirmed by new sequences from the same
individuals, and so these positions were coded as unknown. All sequences have been submitted to GenBank and have accession numbers U79188-U79224.

*Analysis.* -- Alignment of the 37 sequences included in this study was straightforward. No indels were observed. A maximum-parsimony analysis of these data was conducted using PAUP version 3.1.1 (Swofford 1993). Two sets of heuristic searches were performed, with the taxa of known hybrid origin (*Narcissus dubius* and *Narcissus tortifolius*; see below) either included or excluded. Each search employed 100 random addition replicates with TBR (tree bisection-reconnection) branch-swapping, and with MULPARS and “Steepest descent” options activated. All character and character-state changes were equally weighted. Support for branches in the maximum-parsimony trees was evaluated from heuristic searches of 100 bootstrap replicates of the data, performed using the same conditions described above.

*Reconstruction of the evolution of stigma-height polymorphisms.* -- Reconstructions were performed on all the shortest maximum-parsimony trees from the analysis in which *N. dubius* and *N. tortifolius* were excluded. The root of the entire tree was placed between tribes *Narcisseae* and *Galantheae* (Fig. 6.4). The outgroups were coded as monomorphic. Taxa were scored according to whether they were monomorphic (i.e., either approach-herkogamous or non-herkogamous), dimorphic or tristylos (Barrett, Lloyd and Arroyo 1996; S.C.H. Barrett pers. obs.). Although dimorphic populations are known in tristylos *N. triandrus*, and monomorphic or near-monomorphic populations are known in several dimorphic taxa (Barrett, Lloyd and Arroyo 1996), I interpreted these as being recently derived intraspecific phenomena and so did not score these taxa as having among-population polymorphisms for stigma-height polymorphisms. Based on Fernandes’ (1940) review of 19th century descriptions of *N.*.
broussonetii, it is possible that this species also possesses a stigma-height dimorphism. Until this claim can be critically assessed, this taxon is provisionally coded as “unknown.”

Polychotomies in the shortest trees were interpreted as “soft,” that is, they result from a lack of power to resolve the speciation events rather than simultaneous multiple speciations. To estimate the frequency of shifts in stylar class, 100 arbitrary resolutions of each shortest tree were examined. All shortest trees were scored simultaneously for the various possible evolutionary transitions using the “State Changes and Stasis” option in MacClade version 3.0 (Maddison and Maddison 1992). The weighting scheme used to evaluate character-state transitions is an underexplored but often critical component of phylogenetic reconstruction of evolutionary shifts (Chapter 2, 5, Appendix A). I therefore examined four different weighting schemes for the character-state transitions among monomorphism, dimorphism and trimorphism. These differed only in the relative weights assigned to the gain versus loss of the stigma-height polymorphisms. Gains of a stigma-height polymorphism (dimorphism or trimorphism) were weighted equally, two times, three times or four times more heavily than shifts back to monomorphism. The step-matrix in Fig. 6.1 illustrates the basic form of these weighting schemes.

RESULTS

Phylogenetic evidence from ndhF. -- There were 79 variable nucleotides among the taxa examined, of which 42 were potentially phylogenetically informative for maximum parsimony analysis. An heuristic search that included all taxa yielded 50 shortest trees of length 88 steps, with a consistency index of 0.886 (0.818 excluding uninformative characters) and a retention index of 0.946. Fourteen branches (“taxon partitions”) were seen across the 50 shortest trees,
Fig. 6.1. Step matrix describing the evolutionary difficulty of shifts in styal condition in *Narcissus*. Shifts to the styal polymorphism (stigma-height dimorphism and tristyly) were weighted $n$-times as heavily as shifts back to monomorphism. Equal-weights: $n = 1$; Unequal weights (2:1, 3:1 or 4:1): $n = 2, 3$ and 4, respectively.
and branches supporting an additional six clades were observed in at least half of the shortest trees (Fig. 6.2). Despite a lack of variation among some taxa (Fig. 6.3), there was ample support for several important clades in the family, including some of the taxonomically more interesting groups. All branches observed in the majority-rule consensus of the shortest trees were represented in at least 50% of bootstrap replicates, and more than half had around 70% or more bootstrap support (Fig. 6.3). Several of these clades correspond to sections, but the monophyly of several sections was robustly contradicted by the chloroplast evidence (Fig. 6.3). Most of the topological differences among the 50 trees lay in the degree of resolution of several groups, rather than being a consequence of conflicting arrangements of taxa among different shortest trees.

Sectional relationships. — Variation among the shortest trees in the phylogenetic structure of sub-genus Narcissus (Fig. 6.2) represented cases where single branches supporting particular clades dissolved into polychotomies in some trees but not others (i.e., the nodes marked with unfilled circles in Fig. 6.4). None of these represented cases where conflicting clades were present among different shortest trees. This was also true for sub-genus Hermione (Fig. 6.2), with the exception that different trees conflicted over what constituted the basal taxon of this clade. Four clades corresponded to individual sections (Apodanthae, Bulbocodii, Ganymedes and Tapeinanthus), two of which represent multiple populations of single species (Fig. 6.2, 6.3). The population pair representing the monotypic section Aurelia was supported as a monophyletic group in some trees but not in others. The population pair representing the monotypic section Serotini was part of a polychotomy that included two to four other taxa in sub-genus Hermione. Sections Jonquillae and Pseudonarcissus were not monophyletic.

Two taxa in Jonquillae (N. assoanus and N. gaditanus, corresponding to Fernandes'
Fig. 6.2. The phylogenetic history of *Narcissus* based on the chloroplast gene *ndhF*. This is a majority-rule consensus of 50 shortest maximum-parsimony trees. The numbers on each branch are the percentage of shortest trees with this branch. Sections whose monophyly is supported in all shortest trees are marked with bold brackets. $J_1$ and $J_2$ correspond to sub-sections *Juncifoliiae* and *Jonquillae* of section *Jonquillae*, respectively.
Fig. 6.3. One of 50 shortest maximum-parsimony trees of *Narcissus* based on the chloroplast gene *ndhF*. The reconstructed number of mutations is indicated above each branch. Where DELTRAN estimates differed from ACCTRAN estimates, the former is indicated in parentheses. Bootstrap support is indicated below each branch in square brackets.

Chromosome numbers (Fernandes 1967, 1975; Romero, Sánchez Castillo, and Ruiz Rejón 1983) are given where known; polyploid series are known in several species, including *Narcissus bulbodium*, where all possible polyploids from triploids to octaploids are known. The base chromosome number is \( x = 7 \) in sub-genus *Narcissus* and \( x = 5 \) or \( 6 \) in sub-genus *Hermione*. Chromosome counts and possible base chromosome numbers in the three outgroup genera (*Sternbergia*, *Leucojum* and *Galanthus*) are from Webb (1980).
sub-section *Juncifoliae*) were well isolated from other taxa in this section, and were located in a poorly resolved clade containing taxa from *Ganymedes, Pseudonarcissus, Narcissus* and *Tapeinanthus* (Fig. 6.2). In the most resolved shortest trees (e.g., Fig. 6.3, 6.4A), sub-section *Jonquillae* was the most basal clade of sub-genus *Narcissus, Apodantheae* was the next most basal clade, followed by two sister-clades consisting of sections *Bulbocodium* and the group containing the remaining sections of this sub-genus.

In the heuristic search that included the hybrid taxa *N. dubius* and *N. tortifolius*, these taxa belonged to a clade that corresponds to sub-section *Jonquillae* (Fig. 6.2). The search that excluded these taxa yielded 30 shortest trees of length 87 steps, with a consistency index of 0.897 (0.833 excluding uninformative characters) and a retention index of 0.947. Pruning these two taxa from the 50 shortest trees obtained from the previous search yielded the same 30 trees. These trees were used in the reconstructions of the evolution of stigma-height polymorphism.

*Reconstruction of the evolution of stylar polymorphisms in Narcissus.* -- Lack of resolution in several areas of the phylogeny impeded reconstructions of the evolution of stylar polymorphisms (see Fig. 6.4A). Some clades were found only in sub-sets of these trees (the corresponding nodes are indicated with unfilled circles in Fig. 6.4A-C). Some unresolved areas were seen among all 30 shortest trees (nodes marked with filled circles in Fig. 6.4B,C correspond to arbitrary resolutions of these polytomies). Reconstructions were performed on 100 randomly resolved replicates of each shortest tree to estimate the range of evolutionary transitions in the genus. For any given weighting scheme, there was typically a large degree of uncertainty in the reconstructed number of shifts between stigma-height dimorphism and monomorphism, but the weighting scheme also had a strong influence on the reconstructed numbers of these two classes of event (Fig. 6.5).
Fig. 6.4. Reconstructions of the evolution of stylar class (monomorphism, dimorphism and tristyly) in *Narcissus* using one of the 30 shortest maximum-parsimony trees from an analysis in which the two known hybrid taxa *Narcissus dubius* and *Narcissus tortifolius* were excluded. The tree used (Fig. 6.4A) was arbitrarily resolved in Fig. 6.4B and 6.4C. The root of the entire tree was placed between tribes *Narcisseae* (composed of *Sternbergia* and *Narcissus*) and *Galantheae* (composed of *Galanthus* and *Leucojum*). Unfilled circles indicate nodes that are not resolved in all shortest trees; filled circles indicate arbitrarily resolved nodes in this shortest tree. An equal-weighting scheme for transitions among herkogamy classes was employed for Fig. 6.4A and 6.4B. An unequal-weighting scheme (2:1 weighting) was employed for Fig. 6.4C. See the text and Fig. 6.1 for a description of the weighting schemes.
Gaianthus nivalis
Leucojum aestivum
Stembergia lutea
N. atlanticus
N. cuatrecasasii
N. rupicola
N. acaberulus
N. calcicola
N. bulbocodium
N. cantabricus
N. hydraeanthus
N. triandrus concolor
N. triandrus cenuus
N. triandrus triandrus
N. bicolor
N. asturiensis
N. pseudonarcissus
N. hispanicus bujei
N. poeticus hellenicus
N. radiiflorus poetarum
N. assoanus
N. gaditanus
N. longispatus
N. cavanillesii
N. cavanillesii
N. fernandesii
N. jonquilla
N. viridiflorus
N. broussetii
N. broussetii
N. elegans
N. serotinus
N. serotinus
N. tazetta
N. papyraceus
Fig. 6.5. Estimated numbers of evolutionary shifts between stigma-height monomorphism and dimorphism in *Narcissus* and related genera for several different weighting schemes that determine the relative difficulty of transitions between stylar monomorphism and polymorphisms (see text and Fig. 6.1). Bars indicate the range of character-state transitions (minimum-maximum) observed across 100 arbitrary polychotomy resolutions of 30 shortest maximum-parsimony trees (*N. dubius* and *N. tortifolius* were excluded). Estimated numbers of shifts from monomorphism to dimorphism are indicated by bold bars and the single cross. Shifts from dimorphism to monomorphism are indicated by grey bars.
No. of shifts between monomorphism and dimorphism

Character-state weighting scheme

Equal weighting

2:1 weighting

3:1 weighting

4:1 weighting
When all evolutionary transitions were equally weighted, as many as six (Fig. 6.4B) and as few as zero origins of stigma-height dimorphism were indicated (zero gains indicate that the character-state was primitive for the whole tree). There was a similar possible range in the number losses of the dimorphism. When origins of the stigma-height polymorphisms were more heavily weighted (corresponding to a lower probability of these types of change) there were, not surprisingly, fewer origins of stigma-height dimorphism indicated (e.g., Fig. 6.4C). With a 4:1 weighting scheme, a single (plesiomorphic) origin of stigma-height dimorphism was indicated in all of the reconstructions I examined (Fig. 6.5).

There was a single origin of tristyly in the genus for all trees and weighting schemes examined. This is not surprising, considering that the tristylos section Ganymedes is a monophyletic group in all shortest trees and that the three intraspecific taxa in this section all possess a character state (tristyly) not seen anywhere else on the tree. Whether tristyly originated from a stigma-height dimorphism or directly from monomorphism varied among the resolutions of the different trees for any weighting scheme examined, but most resolutions indicated an origin of tristyly directly from monomorphism. However, for unequal weighting schemes, when nodes that give rise to N. assoanus or N. gaditanaus were situated one node away from the node giving rise to section Ganymedes (e.g., Fig. 6.4C), it was equivocal whether tristyly arose from monomorphism or stigma-height dimorphism. Such reconstructions also permitted a single loss of tristyly on the tree (Fig 6.4C).

Given the degree of equivocality in the number of origins of stigma-height dimorphism, it is not surprising that the primitive breeding-system of the genus is itself equivocal. Inspection of arbitrary resolutions of the 30 shortest trees revealed numerous instances where this was the case under the equal (e.g., Fig. 6.4A) and 2:1 weighting schemes. For heavier weighting schemes, all arbitrary resolutions I inspected had stigma-height dimorphism as the primitive
state of the genus. The effect of these heavier weighting schemes was to push the origin of dimorphism to the base of the entire tree.

DISCUSSION

I treat the *ndhF*-based phylogeny of *Narcissus* as a provisional statement of the phylogenetic history of the genus. It is not yet known how extensive the contribution of ancient (or recent) hybridization has been to the phylogenetic history of the genus, and the current chloroplast evidence lacks resolution in several key areas. Resolution will be improved as more chloroplast sequence data become generated. In the meantime however, our uncertainty can be bounded by careful examination of the available phylogenetic evidence. Below I discuss how well the *ndhF* trees tally with biosystematic evidence from the karyotype in *Narcissus*, and the implications of the chloroplast evidence for breeding-system evolution in the genus.

*Taxonomic Implications*

The monophyly of *Narcissus*. — *Narcissus* is closely related to *Sternbergia lutea* and the species from the two genera in *Galantheae* (Fig. 6.3). There was quite low bootstrap support (only 55%; Fig. 6.3) for the monophyly of *Narcissus* from the chloroplast-based evidence. In 32% of bootstrap replicates *Sternbergia lutea* was actually positioned within sub-genus *Hermione* of *Narcissus*. These findings are likely a consequence of the relatively short distance (only two or three steps) separating the ingroup node of *Narcissus* from the first outgroup node (Fig. 6.3). A greater taxonomic sampling of outgroup taxa, particularly other species of *Sternbergia*, would be valuable in future phylogenetic studies of *Narcissus*. The root of the genus divides it neatly
CHAPTER 6. PHYLOGENY OF NARCISSUS

into two well supported clades that correspond to the sub-genera (*Narcissus* and *Hermione*) accepted by Fernandes (1967).

*Sub-genus* Hermione. -- The three sections of sub-genus *Hermione* (*Aurelia*, *Serotini* and *Tazetae*) were intermingled on the chloroplast tree. *Narcissus serotinus* (section *Serotini*) has the lowest known base chromosome in the genus (*x* = 5). However, whether or not it is the most basal entity in sub-genus *Hermione* was unresolved. The *ndhF* sequence of this species is identical to that of *N. tazetta* and *N. elegans* (Fig. 6.3). A clade consisting of a polyclotomy involving these three species was positioned as a basal component of the sub-genus in some of the shortest trees but not others (e.g., Fig. 6.3). It is possible, therefore, that more than one polyploid speciation event was involved in the origin of higher chromosome numbers in this sub-genus.

*Sub-genus* Narcissus. -- The basal components of sub-genus *Narcissus* were also not resolved, although a group consisting of sub-section *Jonquillae* (and two taxa of hybrid origin between *Jonquillae* and *Tazetae*) was the sister-group of the rest of sub-genus *Narcissus* in eighty percent of the shortest trees. *Apodanthae*, the section that Fernandes (1966) erected from part of section *Jonquillae* [an elevation that Webb (1980) did not recognise], was a well-supported clade that was distinct from any members of section *Jonquillae*. Two other clades are candidates for having basal placements in sub-genus *Narcissus*: section *Bulbocodii* (a well-supported, monophyletic clade in the sub-genus), and a clade containing a group of taxa from sections *Ganymedes*, *Narcissus*, *Pseudonarcissus*, *Tapeinanthus*, and sub-section *Juncifoliae* of *Jonquillae* (Fig. 6.2)

The occurrence of this latter clade is one of the most intriguing results of this study, as it includes some morphologically and taxonomically distinct taxa. The *ndhF* sequences of
CHAPTER 6. PHYLOGENY OF NARCISSUS

Narcissus longispathus (section Pseudonarcissus), N. cavanillesii (section Tapeinanthus), N. assoanus and N. gaditanus (sub-section Juncifoliae) are nearly identical to each other and form part of a basal polychotomy in this clade (Fig. 6.2, 6.3). Collectively these taxa span much of the morphological variation within the genus, ranging from the diminutive N. cavanillesii to the much sturdier N. longispathus. The isolated position of N. assoanus and N. gaditanus from the rest of section Jonquillae was another unexpected feature of the phylogeny. Intriguingly, Fernandes and Fernandes (1945) found that the karyotypes of N. cavanillesii (section Tapeinanthus) constitutes a tetraploid version of the chromosomal complement found in N. gaditanus (sub-section Juncifoliae).

Fernandes (1975) also noted a resemblance between the karyotypes of sub-section Juncifoliae and section Ganymedes (Fernandes 1975). His karyotypic formulae for Narcissus triandrus and Narcissus bulbocodium are nearly identical (Fernandes 1949, 1967, 1968b; Appendix G). However, the chloroplast tree depicts Ganymedes as having a position quite isolated from section Bulbocodium, a position embedded within the clade containing taxa from Narcissus and Pseudonarcissus and other sections (Fig. 6.2, 6.3). All members of sections Narcissus and Pseudonarcissi that Fernandes examined had identical karyotypes (Fernandes 1967, 1968b; Appendix G). The closeness of members of section Pseudonarcissus and Narcissus on the chloroplast tree is thus in line with Fernandes' karyotypic work.

Tapeinanthus is the most recent genus to be combined with Narcissus (Traub 1969) and it has not been universally accepted as a member of the genus (Blanchard 1990). This may partly be because of its unusual combination of features. Narcissus cavanillesii (section Tapeinanthus) has a very short floral tube, a rudimentary corona, and is autumn-flowering. Only Narcissus cyclamineus (section Pseudonarcissi) has a floral tube that approaches that of N. cavanillesii (Blanchard 1990), but the former species has a long corona. Several taxa have very short coronas, but only in N. broussonetii and N. cavanillesii is the corona virtually absent.
CHAPTER 6. PHYLOGENY OF NARCISSUS

(see Webb, 1980; Blanchard 1990). Only *N. elegans* (in *Tazetiae*), *N. serotinus* (in *Serotini*), and *N. viridiflorus* (in *Jonquillae*) are also autumn-flowering in *Narcissus*. No other species in *Narcissus* has this combination of rare features. The chloroplast-based phylogeny strongly supports the inclusion of *Tapeinanthus* in *Narcissus*, and its position rather deep in the tree suggests that its reduced stature and unique floral morphology are derived features.

*Evolution of chromosome number and form.* -- Fernandes (1967) designed schemes showing how the different karyotypes found in sub-genus *Hermione* may have evolved, and he also commented (Fernandes 1967, 1975) on the similarities among karyotypes of the various members of sub-genus *Narcissus*. Section *Jonquillae* can be sub-divided on the basis of two related karyotypes, one of which corresponds to his sub-section *Juncifoliae* (*Narcissus assoanus* and *Narcissus gaditanus*), and the other to sub-section *Jonquillae* (the remaining taxa in *Jonquillae*) (Fernandes 1975).

Fernandes (1937) determined that *Narcissus dubius* (*2n = 50*) is a hybrid allopolyploid whose genome contains 28 chromosomes derived from *N. assoanus* (*2n = 14*) and 22 from *N. papyraceus* (*2n = 22*) or one of its allies. Fernandes (1967) hypothesized that the hybridization involved a diploid gamete of *N. assoanus* (from an unreduced diploid gamete or a gamete derived from a tetraploid form of this species), and a haploid gamete from *N. papyraceus*, with subsequent doubling-up of the hybrid genome [(14+11) x 2 = 50], presumably to restore fertility. Another intersectional hybridisation event between sections *Tazetiae* and *Jonquillae* is thought to have yielded *Narcissus tortifolius* (*2n = 36*), with *N. gaditanus* acting as one of the parents (A. Fernandes, pers. comm. in Romero, Sánchez Castillo, and Ruiz Rejón 1983). Fernandes’ cytological studies indicated that the parents of the allopolyploids *N. dubius* and *N. tortifolius* came (in both cases) from sections *Jonquillae* and *Tazetiae*. *Narcissus dubius* and *N. tortifolius* are both stable, fertile allopolyploids (Blanchard 1990).
Because the ndhF-based phylogeny is a single-gene tree, the hybrid taxa should be positioned with one of the two parental taxa on it. *Narcissus dubius* and *N. tortifolius* are located within sub-section *Jonquillae* on this tree, but are not found associated with *N. assoanus* or *N. gaditanus*. This could indicate that *N. assoanus* and *N. gaditanus* were both involved in palaeo-hybridization events (with an unknown third taxon) after the hybrid origins of *N. dubius* and *N. tortifolius*. This would explain the current disjunct status of sub-section *Juncifoliae* on the chloroplast tree, but as it requires a total of four independent hybridization events it otherwise seems highly improbable. An alternative possibility is that a member of sub-section *Jonquillae*, rather than sub-section *Juncifoliae* was one of the parents for both hybrid taxa, or at least that *N. assoanus* (not *N. gaditanus*) was one of the parents of *N. tortifolius*. It is interesting in this respect that Blanchard (1990) questioned the conclusion of Fernandes (not Fernández Casas as he states) that *N. gaditanus* gave rise to *N. tortifolius*.

Other possible conflicts between the chloroplast gene tree and Fernandes’ biosystematic work (see above) may be explicable if some of the taxa in the genus underwent introgression events at some point in their phylogenetic history, or were themselves of hybrid origin. Allopolyploid speciation has probably been a relatively rare occurrence in sub-genus *Narcissus*, because most species in the sub-genus have base chromosome numbers of x = 7 (Fig. 6.3). The higher base numbers typical of species in sub-genus *Hermione* (Fig. 6.3) may be a result of at least one allo- or autopolyploidization event. However, numerous instances of inter-sectional hybridization have been recorded in the wild (reviewed in Blanchard 1990), and hybrid plants are sometimes partially fertile (e.g., Jefferson-Brown 1991). The possible historical occurrence in *Narcissus* of introgression of linkage groups (such as the chloroplast genome) between species and across sectional boundaries, or of “recombinational speciation” between taxa possessing chromosomal rearrangements to yield fertile diploid hybrids (Stebbins 1957; Grant 1958), will only be detected when phylogenetic markers from other linkage groups become available.
CHAPTER 6. PHYLOGENY OF *NARCISSUS*

It is not known whether any diploid taxa in *Narcissus* originated through recombinational speciation, but the existence of self-sterility in *Narcissus* should reduce the likelihood of such taxa arising (Grant 1971). Nonetheless, the possible historical occurrence of undetected recombinational speciation or introgression means that parts of the chloroplast tree presented in this study, while amenable to better resolution through the addition of more chloroplast data, may in part be a poor reflection of species relationships. Sequence information from loci in the nuclear genome are consequently needed to assess how well the chloroplast has tracked species phylogeny in this genus. However, until such data become available, I treat the *ndhF*-based phylogeny as a first approximation of the phylogeny of species of *Narcissus*, and use it to begin to think about the evolution of characters of interest.

*The Evolution of Stigma-Height Polymorphisms in Narcissus*

*Tree resolution and character reconstruction.* -- The lack of resolution of certain regions of the *ndhF*-based phylogeny of *Narcissus* presented problems for the reconstruction of character evolution. Estimates of the range of evolutionary events can nonetheless be obtained by sampling arbitrary resolutions of these unresolved areas. For a variety of weighting schemes, these resolutions revealed a large degree of uncertainty in the reconstructed number of origins and losses of stylar polymorphisms in *Narcissus* (Fig. 6.5) and in whether or not stigma-height monomorphism was primitive in the genus. Tristyly evolved once in the genus, and most reconstructions indicated that it arose directly from monomorphic stigma height. If there were an intermediate, short-lived stigma-height polymorphism, the phylogenetic tree does not record this fact.
**CHAPTER 6. PHYLOGENY OF *NARCISSUS***

*Weighting schemes and phylogenetic reconstruction.* – A major unresolved issue is the appropriateness of particular weighting schemes for reconstructing the evolution of reproductive (and other) characters. Weighting schemes are necessary for reconstructing character evolution on a given phylogenetic tree. They describe the evolutionary difficulty of transitions between different character states. Most published accounts of reconstructions of character evolution assume that all possible evolutionary shifts are evolutionarily equally likely (e.g., Brooks and McLennan 1991), but this is as strong an assumption as any other weighting scheme (Swofford et al. 1996). I examined a range of schemes that weighted the origin of stigma-height polymorphisms (dimorphism or tristyly) more heavily than their conversion back to monomorphism (Fig. 6.5). With a 4:1 weighting scheme, all the resolutions of the shortest tree showed a single origin of the dimorphism. Weighting schemes are commonly derived from logarithmic transformations of character-transition probabilities (Swofford et al. 1996), so this particular scheme probably reflects a substantially greater than four-fold difficulty of evolutionary shifts than the equal weighting scheme.

A 4:1 weighting scheme is in fact sufficiently biased to indicate a single origin of stigma-height dimorphism for most *possible* trees constrained to have the observed terminal states in *Narciissus*. For 10,000 equiprobable random trees generated using MacClade (Maddison and Maddison 1992) with the observed terminal character states, the average number of origins of stigma-height dimorphism from monomorphism was 0.205 (range 0-4) across all trees and equivocal resolutions. The 4:1 weight is a heavier weighting scheme than that required in Chapter 3 (and Appendix A) to indicate a single origin of tristyly inPontederiaceae. Single origins of the dimorphism in *Narciissus* were found for some resolutions of the shortest trees, even when all character-state changes were equally weighted.

Transformation probabilities are unknown in advance for most characters, so it seems advisable to examine a range of different weighting schemes. The stigma-height dimorphism is
developmentally quite simple. It requires a simple shift in style height and is the result of a single dominant allele (at least in *N. tazetta*; Dulberger 1964). Tristyly is more complex than stigma-height dimorphisms or stylar monomorphism, as it involves three distinct, but reciprocal morphologies. It also has a more complex genetic basis (where this is known; see Lewis and Jones 1992). While this greater complexity means that it is probably "harder" or less likely to evolve, whether the origin of tristyly should be weighted differently from the origin of dimorphism (cf. Fig. 6.1) is a moot point in this study, because of the limitation of tristyly to a single species of *Narcissus*.

The scarcity of non-heterostyloous stigma-height dimorphisms in the angiosperms as a whole may not be an accurate indication of the evolutionary difficulty of evolving this system. Lloyd and Webb (1992a) suggested that after this initial dimorphism evolves it may quickly shift to heterostyly. If this is so, then the relatively large number of origins of distyly in the angiosperms (perhaps more than 23; Lloyd and Webb 1992a) indicates that stigma-height polymorphisms are not exceptionally difficult systems to evolve. Because herkogamy class is rarely recorded in taxonomic descriptions (Webb and Lloyd 1986), it is probable that the known number of stigma-height polymorphisms in the angiosperms is an underestimate of the true number.

This study raises the distinct possibility that different instances of stigma-height dimorphism in *Narcissus* may not be homologous with each other. If stigma-height dimorphism did arise only once in the genus, this contradicts Lloyd and Webb's general statement concerning the short-lived nature of stigma-height dimorphisms. Barrett, Lloyd and Arroyo (1996) postulated that unusual developmental constraints may exist in the genus that make it difficult for the stigma-height dimorphism to evolve into heterostyly. The nature of such constraints is obscure, and it would therefore be interesting to perform a comparative study of stigma-height dimorphisms in *Narcissus* and the few other known instances of this
polymorphism in the angiosperms (see Barrett, Lloyd and Arroyo 1996) to see if any common developmental themes emerge that are suggestive of such constraints. However, if stigma-height dimorphisms did arise on multiple (and hence recent) occasions in the genus, it is relevant to ask what it is about *Narcissus* that has made it prone to repeatedly evolve stigma-height dimorphisms, while none of its close relatives in Amaryllidaceae have done so. The existence of this polymorphism in *Narcissus* thus leads to some interesting questions that will only be answered with a more resolved picture of its phylogenetic history and an improved understanding of floral development and pollination ecology in the genus.
Chapter 7. General Conclusions

The major objective of this thesis was to investigate the phylogenetic history of two monocotyledonous groups possessing tristylos taxa (Pontederiaceae and Narcissus), and to use these estimates of phylogenetic history to reconstruct the evolution of breeding-systems in these groups. Several new and revised data sets were used to estimate the phylogenetic history of the two groups. These include surveys of DNA sequence variation in two chloroplast genes (ndhF and rbcL) in Pontederiaceae (Chapters 2, 3); a survey of restriction-site variation in this family obtained in collaboration with Dr. Joshua Kohn (University of California, San Diego; Appendix A); a morphological data set of the family (Eckenwalder and Barrett 1986) revised here in collaboration with Dr. James Eckenwalder (Chapter 3, Appendix C); a survey of DNA sequence variation in ndhF for several families related to Pontederiaceae (Chapter 4) and a survey of DNA sequence variation in ndhF for the genus Narcissus. These surveys were the starting point for the various studies investigating the evolution of the breeding-system and other characters in these two groups. I summarize here the main conclusions from each of the five research chapters and suggest future lines of inquiry that might modify or strengthen the major results.

Chapter 2. -- Previously available data from the chloroplast gene rbcL suggested affinities of Pontederiaceae with other families within and outside several related superorders (Commelinanae, Bromelianae and Zingiberanae), including Commelinaceae, Haemodoraceae and Philydraceae (Chase et al. 1993). I re-examined this data using a bootstrap analysis and determined the phylogenetic position of Pontederiaceae in some marginally sub-optimal trees. These analyses indicated that the rbcL evidence by itself is incapable of a robust determination of the taxa to which the family is most closely related. Duvall et al. (1993) independently obtained the same result. However the same analysis (using new molecular data for the family for the chloroplast gene rbcL) strongly supported the family’s monophyly. New rbcL-based
evidence for the family was used in tandem with new evidence from the chloroplast gene \textit{ndhF} to investigate the phylogenetic relationships of taxa within Pontederiaceae. Of the four main generic groupings in the family, the molecular evidence supports the monophyly of \textit{Monochoria} and two other broadly defined generic groupings (\textit{Pontederia} and \textit{Heteranthera} and their respective allied genera) and indicates that \textit{Eichhornia} is an unnatural genus as currently circumscribed. These findings are in line with the morphological analysis of Eckenwalder and Barrett (1986).

Reconstructions of breeding-system evolution in the family using the molecular evidence indicated multiple losses of the sexual polymorphism tristyly within \textit{Eichhornia}. A weighting scheme that is only marginally biased against multiple origins of enantiostyly and tristyly indicates a single origin of tristyly in the family, two non-homologous instances of enantiostyly in the family, and multiple reversions of these polymorphic floral forms to floral monomorphism. Since tristylos species are predominantly outcrossing and monomorphic lineages derived from tristylos species are predominantly self-fertilizing, this reconstruction also indicates multiple origins of self-fertilization in the family. Self-incompatibility in the family arose on one or two occasions \textit{after} the origin of tristyly (assuming a single origin of tristyly in the family), a temporal sequence compatible with only one theory concerning the origin of heterostyly, that of Lloyd and Webb (1992a,b; Table 1.1; and see Darwin 1877).

The significance of breeding-system variation and evolution in tristylos, homostylos and enantiostylos taxa of Pontederiaceae was assessed in a broader evolutionary and phylogenetic context. Tristyly in the family is highly unlikely to be homologous with instances of heterostyly in two other monocotyledonous groups [distyly in \textit{Nivenia} (Iridaceae) and tristyly in \textit{Narcissus} (Amaryllidaceae)] because these taxa are only distantly related to Pontederiaceae. Enantiostyly in the family may well be homologous with its occurrence in other families (Commelinaceae, Haemodoraceae and Philydraceae). I therefore reviewed the general
morphological features of the enantiostyly syndrome in the monocotyledons. This floral syndrome is often associated with a variety of reproductive traits associated with pollen-collecting bees (heteranthery, poricidal anther dehiscence, the absence of nectaries), an observation of unknown significance. The functional significance of enantiostyly has been interpreted variously. It has most often been portrayed as a breeding system that increases the proficiency of cross-pollination, an adaptive hypothesis for which there is currently little experimental evidence.

Chapter 3. -- Several approaches were used to assess statistical support for different aspects of phylogenetic congruence among the four data sets. One data set is a revision of Eckenwalder and Barrett's (1986) morphological data set (Appendix C), a second is a restriction-site survey of the chloroplast genome (Appendix A), and the third and fourth are two new data sets based on DNA sequence variation in the chloroplast genes rbcL and ndhF. The degree of phylogenetic congruence and conflict in separate and combined analyses of these data sets was assessed using a variety of methods, including a test for inferring shared descent among different sources of data (Penny, Foulds and Hendy 1989), a character-based statistical test of incongruence between phylogenies (Templeton 1983; Felsenstein 1985b, 1995), and direct comparisons of bootstrap support for observed taxon partitions by different data sets.

A tree-based representation of the topological dissimilarity among trees from the same and different data sets was used to assess the degree of "taxonomic congruence" (Mickevich 1978) among data sets. Taxonomic congruence is the similarity in taxonomic groupings implied by trees derived from different data sets. When trees from different data sets were examined in this manner, significant cophenetic correlation coefficients were obtained between raw tree-to-tree distances and the distances implied by the phenetic representations of tree dissimilarity.
This indicates that the raw tree-to-tree distances are hierarchical and that this phenetic approach to assessing taxonomic congruence is valid, at least for these data sets.

Trees derived from the various chloroplast-based sources of evidence were well supported and well resolved, and the support and resolution of their phylogenetic structure increased with data-set size and combination. Unrooted chloroplast-based trees were largely congruent with each other, and in combination these data appear to be converging towards the true chloroplast tree of the family. The trees were all derived from analyses in which all characters and all possible character-state changes were equally weighted. Despite the fact that different evolutionary rules are known to operate among sub-sets of the chloroplast data, there was no statistically detectable heterogeneity among these three data sets, and they can therefore safely be combined for unweighted phylogenetic analysis (Heuelsenbeck, Bull and Cunningham 1996). The single major unresolved aspect of these molecular analyses concerned the position of the root of the family (see Chapter 4).

In contrast, the morphological data produced trees that were highly distinctive topologically, compared with trees based on the chloroplast evidence. These trees were poorly supported by bootstrap analysis, indicating that there is insufficient morphological evidence to reconstruct the phylogeny of the family. This result is corroborated by the fact that the morphological data set was almost completely overpowered by the molecular evidence when all four data sets were combined. Nonetheless, a comparison of the topological distances between morphological, molecular and random trees indicated that trees from the morphological and molecular sources of evidence do agree on substantial elements of descent. However, despite the small size of the morphological data set, there was still statistically detectable character incongruence between the morphological and molecular classes of data, beyond that expected by sampling error in either data set.
An hypothesized distortion due to a "selfing syndrome" on phylogenetic reconstructions using the morphological data (Eckenwalder and Barrett 1986) was also examined using Templeton's (1983) test. This hypothesis predicted that selfing species of *Eichhornia* erroneously become grouped during phylogenetic analysis using morphological evidence, as a consequence of multiple parallel shifts in reproductive and other life-history characters. Neither the morphological nor the molecular data supported this hypothesis, as trees uniting selfing species of *Eichhornia* were substantially longer than the shortest trees in the former case, and significantly longer in the latter case.

Chapter 4. -- I obtained sequence data from the chloroplast gene *ndhF* for twelve families in five different superorders (Arecales, Bromeliaceae, Commelinaceae, Liliales and Zingiberales) and combined this evidence with data from *rbcL*, to attempt to clarify the phylogenetic position of Pontederiaceae in the monocotyledons. When the taxa in a phylogenetic analysis are all distantly related to each other, it is important to try to correct for "multiple hits" (repeated convergences and reversals) at individual nucleotide positions in the sequence data. If these repeated hits are not corrected for, there is a serious danger that long branches in such a phylogenetic tree will "attract" each other during phylogenetic reconstruction, yielding erroneous estimates of phylogeny (Felsenstein 1978; Penny and Hendy 1989). I assessed two different classes of re-weighting schemes, one based on the relative likelihoods of change at different sites, the other based on the reconstructed amount of homoplasy at each site, as determined on a preliminary tree produced by an unweighted analysis. Weighted and unweighted bootstrap analyses of the combined sequence data yielded identical or nearly identical trees. Bootstrap analyses also indicated that most branches in these trees were robust. The closest taxa to Pontederiaceae from these analyses were Commelinaceae and
Haemodoraceae (which together constitute the sister-group of the family) and Philydraceae (which is the next most closely related group).

A major finding of the preceding chapter was that the position of the root of the family is uncertain. "Lundberg rooting experiments" (Lundberg 1972) using the three closely related families, were performed to assess optimal and sub-optimal root locations in combination with tests due to Templeton (1983) and Miyamoto and Boyle (1989). Using the combined ndhF and rbcL data and one of four shortest unrooted trees of Pontederiaceae found using the combined chloroplast data, I successively attached the three outgroup families (singly and in combination) to each branch on this unrooted tree, and determined the penalty in parsimony required to place the roots in sub-optimal positions, for unweighted versions of the combined sequence data. I repeated the same experiment using 100 random outgroup sequences.

For the real outgroups, Templeton’s (1983) test was used to assess whether sub-optimal root positions were significantly different from the optimal root. This was found to be the case for most but not all sub-optimal root positions. Of all the outgroups I examined, the outgroup combination involving all three families, with two of them (Commelinaceae and Haemodoraceae) constrained to form a clade sister to Pontederiaceae, was found to have the best ability to discriminate whether different sub-optimal roots were significantly different from the optimal one (about 20% of sub-optimal root positions were not significantly different from the optimal one). When this experiment was performed using the restriction-site data, the majority of root positions were not significantly different from the optimal root position, suggesting that historical signal in the outgroup branches in the restriction-site data has been almost completely randomized by multiple hits.

A second, less sensitive test [an extension of one due to Miyamoto and Boyle (1989)], was used to determine the decrease in parsimony required to root each random outgroup sequence in sub-optimal positions. This was compared to the penalty required to root a real
composite outgroup (consisting of the three outgroup families) in its suboptimal positions. The real outgroup performed significantly worse than the random outgroup sequences for around half of the sub-optimal root positions. This test thus also indicates that despite the length of the branches leading to the outgroup families, the historical signal on these branches has not been completely randomized.

This ingroup branch where the root attached was moderately long, raising the possibility that the outgroup branch was erroneously attracted to it, rather than to the true root. Several lines of evidence indicate that this is not the case. For almost all of the real outgroups that I examined individually or collectively, the most-parsimonious root position was found on this branch. The same branch was determined to contain the optimal root when the data were weighted under three different weighting schemes to correct for multiple hits. Of 100 random outgroup sequences, the optimal root position was found to be on this branch for none of them. Outgroups that have had their historical signal “overwritten” with multiple hits therefore do not preferentially attach to this long branch.

Chapter 5. The ecological evidence that a number of life-history traits in Pontederiaceae are involved in or affected by an aquatic existence were reviewed, and reconstructions of evolution of vegetative and reproductive characters were performed, but this time (cf., Chapter 2; Appendix A) using the trees obtained from the three combined chloroplast data sets (Chapter 3), with the single optimal rooting indicated by the most closely related families to Pontederiaceae (Chapter 4). These trees were also used to reconstruct patterns of character diversification associated with adaptive radiations in vegetative characters involved with the aquatic habit in the family. Vegetative characters in the family are particularly prone to convergence (see also Eckenwalder and Barrett 1986). The extensive homoplasy in vegetative characters suggests that aquatic habitats are far from ecologically uniform, as has often been supposed.
Reconstructions of the most primitive Pontederiaceae indicate that it was annual, non-clonal, emergent and erect, and that it possessed a pattern of leaf development in which the seedling rapidly switches to the production of adult, petiolate leaves. There were consequently multiple origins in the family of perenniality, the various types of clonality, several types of life form, and various developmental patterns of leaf development. Reconstructions of the evolution of stalar polymorphisms and self-incompatibility in the family confirm those in Chapter 2, which is not surprising considering that tree structure for the two different combinations of the chloroplast data are for the most part very similar (Chapter 3).

Chapter 6. -- The molecular evidence from ndhF yielded an incompletely resolved phylogeny of Narcissus. These phylogenetic estimates support the division of the genus into two natural groups (sub-genera Hermione and Narcissus) that correspond to a basic division in chromosome number in the genus (a probable base number of n = 5 in the former case and n = 7 in the latter). Only a few of the sub-sectional divisions of the genus corresponded to monophyletic groups (Apodantheae and Bulbocodii and Ganymedes, all with strong bootstrap support). Bootstrap analysis robustly contradicted the monophyly of several sections (Pseudonarcissi and Jonquillae). In agreement with evidence from chromosomal morphology, this study supported the close relationship of several sections (Narcissus with Pseudonarcissi; Aurelia and Serotini with Tazettae). Traub's (1969) inclusion of the distinctive monotypic genus Tapeinanthus within Narcissus, in a phylogenetic position (close to N. gaditanus) predicted by Fernandes and Fernandes (1945), and the taxonomic distinctness of sub-section Juncifoliae and section Apodantheae from the rest of section Jonquillae. The two subsections of section Jonquillae were located in very different parts of the tree, perhaps a consequence of ancient, undetected hybridization events. Hybridization among extant taxa in the genus is common, so it is a distinct possibility that such ancient hybridization played a role in the origin of some taxa in
the genus. Two taxa of known hybrid origin (*Narcissus dubius* and *Narcissus tortifolius*) were included in one set of analyses. Their position on the tree was only partly in agreement with the parental taxa suggested by Fernandes (1967 and pers. comm. in Romero, Sánchez Castillo, and Ruiz Rejón 1983).

If (as seems probable) ancient hybridization events played an important role in the evolutionary history of *Narcissus*, this reduces the value of the chloroplast-based phylogenetic tree for performing reconstructions of breeding-system evolution in the genus. Nonetheless, there is heuristic value in using this as a provisional estimate of the phylogenetic history of *Narcissus*. The ndhF-based phylogeny was not resolved enough to permit very conclusive reconstructions of the evolution of stigma-height polymorphisms in the genus, but bounds on the numbers of evolutionary transitions were estimated under several different character-state weighting schemes that bias against the origin of the polymorphic forms (dimorphic and trimorphic). Under most reconstructions, tristyly does not appear to arise from stigma-height dimorphism in the genus.

**Summary**

A phylogenetic systematist's endpoint (the tree) is the starting point for other biologists interested in the evolutionary history of their group of interest (Swofford 1993). The determination of a robust phylogeny is consequently a prerequisite for performing useful reconstructions of the evolution of breeding systems and other characters. In this thesis I have produced several such "endpoints," and have employed them to think about the evolution of breeding-systems in two groups of heterostylovous monocotyledons. One chapter considers the thorny problem of how to root phylogenetic trees when the extant outgroup taxa are distantly related. The available chloroplast data seem to be converging on a single rooting of
CHAPTER 7. GENERAL CONCLUSIONS

Pontederiaceae (Chapter 4), but more chloroplast or nuclear data would be valuable for increasing our confidence in this result.

Of the four available data sets for Pontederiaceae, the three chloroplast-based ones provide robust, congruent estimates of phylogenetic history. The fourth data set, based on morphology, does not provide a sturdy estimate of phylogeny. Nonetheless, statistically detectable incongruence between the morphological and molecular data was encountered (Chapter 3). Further studies involving increased numbers of morphological characters and molecular data sets from the nuclear genome will be required to resolve the source of this congruence, which may or may not turn out to be trivial with respect to the findings of the three studies (Chapter 2, 5, Appendix A) in which I map breeding-system and other morphological and life-history characters onto chloroplast-based estimates of the phylogenetic history of Pontederiaceae. One major source of incongruence between these two classes of data may have been the occurrence of ancient hybridization events in the family, although the lack of hybridization among extant members of the family suggests this may not have played a major role (Chapter 3, Appendix A).

In contrast, extensive evidence of current and historical hybridization in Narcissus (for references see Chapter 6) indicates that my single-gene estimate of the phylogenetic history of the genus must be treated with caution. This study should therefore be viewed as a preliminary step on the road to determining the number and type of shifts in the breeding-system of this genus. Future phylogenetic studies of Narcissus should expand the chloroplast data base, by looking at more chloroplast genes (to improve the robustness of the chloroplast-based tree), and more samples per taxon (to estimate the current extent of introgression among taxa). Multiple chromosomal or sequence-based markers will be necessary to come to grips with the possible role played by undetected ancient hybridization events in the history of the genus. The congruence of the ndhF-based tree with some karyotypic evidence in the genus is, however,
encouraging. More basic experimental studies on the form and function of floral and physiological traits associated with the breeding system among different species of Narcissus and related genera would also be valuable.

The only other known heterostylosus group in the monocotyledons is Nivenia (Iridaceae), which possesses a mixture of distylosus and monomorphic species. When I wrote the preceding chapters I had not yet seen a morphology-based treatment of the genus and its close relatives by Goldblatt (1993), in which he examined the evolution of distyly in the genus. Two equally parsimonious mappings of distyly are possible in that genus (compare his Fig. 3 and 4). A single gain with two subsequent losses, or three separate gains of distyly, are both equally parsimonious. Both scenarios require three evolutionary steps. However, Goldblatt (1993, pp. 41) considered that the relative rarity of heterostyly in the angiosperms in general, and the monocotyledons in particular, is evidence that the former scenario is more likely than the latter. This is one of the lines of reasoning I employed in the reconstructions of breeding-system evolution in Pontederiaceae and Narcissus. Character-states that are rare in a broad group of organisms may be a priori less likely to evolve on multiple occasions within a given group, particularly if the rare forms of the trait are developmentally or genetically more complex. For complex morphological characters, this argues strongly for an examination of optimization schemes in which complex character shifts are not treated as equally likely events.
APPENDIX A

- RECONSTRUCTION OF THE EVOLUTION OF REPRODUCTIVE CHARACTERS IN PONTEDERIACEAE USING PHYLOGENETIC EVIDENCE FROM CHLOROPLAST DNA RESTRICTION-SITE VARIATION -

**APPENDIX A. RESTRICTION-SITE BASED PHYLOGENY OF PONTEDERIACEAE**

*Abstract.*—We reconstructed phylogenetic history in Pontederiaceae using chloroplast DNA restriction-site variation from approximately two-thirds of the species in this family of aquatic monocotyledons. The molecular phylogeny was used to evaluate hypotheses concerning the evolution of reproductive characters associated with the breeding system. The family has four main genera, two of which (*Eichhornia* and *Pontederia*) have tristylos, predominantly outcrossing species, while two (*Monochoria* and *Heteranthera*) have enantiostylos taxa. Self-incomptability is restricted to some but not all tristylos species. In *Eichhornia* and *Pontederia*, predominantly selfing species with small monomorphic flowers (homostyly) have been hypothesized to result from the multiple breakdown of tristyly. Restriction-site variation provided a well supported phylogeny of ingroup taxa, enabling the mapping of reproductive characters onto trees. Two contrasting optimization schemes were assessed, differing in the relative weights assigned to shifts in character states. The reconstructed sequence of floral character-state change was used to assess competing hypotheses concerning the origin and breakdown of tristyly, and the relationships between tristylos and enantiostylos syndromes. Our results indicate that the class of optimization scheme used was the most critical factor in reconstructing character evolution. Despite some topological uncertainties and difficulty in reconstructing the primitive floral form in the family, several broad conclusions were possible when an unordered, unequally-weighted optimization scheme was used: (1) tristyly originated either once or twice, while the occurrence of enantiostyly in *Monochoria* and *Heteranthera* was always found to have independent origins, (2) tristyly has repeatedly broken down leading to selfing, homostylous taxa, and (3) self-incompatibility probably arose after the origin of floral trimorphism, a sequence of events that conflicts with some evolutionary models.
The integration of phylogenetic analyses with studies of character evolution has blossomed over the past decade following the acceptance of cladistic techniques and the introduction of molecular data to systematics. Recognition of the utility of historical reconstruction as a tool for testing evolutionary hypotheses has awakened an interest in phylogenetics among evolutionary ecologists. Several authors have recently pointed out that when competing hypotheses differ in their proposed order of character-state change, phylogenetic analysis can provide a tool for hypothesis evaluation (e.g. Donoghue 1989; Brooks and McLennan 1991; Harvey and Pagel 1991). For at least three reasons, plant breeding-system evolution is well suited for analysis by historical reconstruction: (1) Closely related groups often show abundant quantitative and qualitative variation in floral traits influencing mating patterns (Jain 1976; Wyatt 1988; Barrett 1989b), (2) Theories of breeding-system evolution provide detailed hypotheses which often conflict with respect to the proposed order of character state change as well as the selective forces involved (e.g. Charlesworth and Charlesworth 1978; Thomson and Brunet 1990; Barrett 1990; Lloyd and Webb 1992a, 1992b; Uyenoyama, Holsinger and Waller 1993), (3) There exists, at least in some plant groups, a wealth of microevolutionary studies of mating-system variation providing a useful context for macroevolutionary inquiry (Schoen 1982; McNeill and Jain 1983; Weller et al. 1990; Barrett, Kohn and Cruzan 1992; Fenster and Ritland 1992). A major goal of evolutionary studies should be to find ways to integrate findings from micro- and macroevolutionary inquiry. Studies of plant breeding-system evolution are likely to provide a useful model in this regard.

In this study we perform phylogenetic analysis on chloroplast DNA (cpDNA) restriction-site variation in the small monocotyledonous family Pontederiaceae to evaluate
hypotheses concerning the evolution of reproductive characters associated with the breeding system. Included within this family of some 35 species are four major genera: *Eichhornia* (8-9 spp.), *Pontederia* (6 spp.), *Heteranthera* (10-12 spp.) and *Monochoria* (7-8 spp.) and several smaller, segregate genera containing 1-3 species each, *Eurystemon*, *Hydrothrix*, *Scholleropsis*, *Reussia* and *Zosterella*. Three of the smaller genera are included in this study: *Reussia*, allied to or combined with *Pontederia* (Lowden 1973); *Zosterella*, allied to or combined with *Heteranthera* (Horn 1985; Rosatti 1987); and *Hydrothrix*, a monotypic genus placed within the family by Hooker (1887) which may also be associated with *Heteranthera* (Rutishauer 1983; Eckenwalder and Barrett 1986). All species are freshwater aquatics occurring primarily in tropical regions, with the center of diversity for the family located in lowland South America. Three distinct floral conditions (Fig. A.1) involving the relative positions of male and female sex organs are present in the family: **tristyly**, a genetic polymorphism found in *Eichhornia* and *Pontederia*; **enantiostyly**, a floral polymorphism that predominates among species of *Heteranthera* and *Monochoria*; and **floral monomorphism**, found in all 4 major genera.

Tristyly is a genetic polymorphism in which populations contain three floral morphs that differ from one another in style length and stamen height (Fig. A.1a). The reciprocal arrangement of anther and stigma heights (reciprocal herkogamy) in the three morphs is a mechanism that promotes animal-mediated cross-pollination between morphs (Darwin 1877; Barrett and Glover 1985; Kohn and Barrett 1992; Lloyd and Webb 1992b). In most tristyloous species a physiological self-incompatibility system (heteromorphic incompatibility) prohibits both self and intramorph matings by preventing fertilization unless pollen comes from the same level as the stigma. Tristyly is usually associated with a suite of ancillary polymorphisms in which pollen size, exine sculpturing and stigma papillae length vary among stamen and style levels (Dulberger 1992). In Pontederiaceae,
Fig. A.1. Schematic representation of stamen and style configurations in tristylos, enantiostylous and monomorphic floral forms: a) Tristyly. A genetic polymorphism controls whether individuals produce long-, mid- or short-styled flowers. Within each flower, stamens are positioned at the two levels not occupied by the stigma. b) Monomorphism. Populations are composed of a single floral form with either one or two sets of anthers at the same level as the stigma. Homostylous taxa of Pontederiaceae typically have one set of anthers at the stigma, and this most commonly involves the mid-styled morph (see Barrett, 1988a). A homostylous mid-styled flower is shown, with short-level anthers adjacent to the stigma. c) Enantiostyly. Flowers have either left- or right-bending styles. A cryptically colored stamen (one of six stamens in Monochoria, one of three in Heteranthera) bends in the opposite direction. In contrast with tristyly, individuals can produce both flower types simultaneously.
A. Tertiary
- Short
- Mid
- Long

B. Monomorphism

C. Enantiostrategy
- Right
- Left

(Homostyly)
tristyly occurs in three species of *Eichhornia* (*E. azurea*, *E. crassipes*, *E. paniculata*) and in all but one species of *Pontederia*. All tristyloous taxa are insect-pollinated and primarily outcrossing. Among tristyloous species of Pontederiaceae there is considerable variation both in the strength of the incompatibility system and the expression of ancillary polymorphisms (Barrett 1988a, 1993). This variation ranges from species such as *E. paniculata*, which are highly self-compatible, setting equivalent numbers of seed following self and intermorph pollinations (Barrett 1985), to species of *Pontederia* with strongly developed self-incompatibility (Barrett and Anderson 1985). The strength of self-incompatibility is correlated with the expression of pollen and stigma heteromorphisms. Self-compatible tristyloous species have weak pollen and stigma polymorphisms (Barrett 1988a), whereas self-incompatible taxa exhibit pronounced differences in pollen size among stamen levels and well developed stigma polymorphisms (Scribailo and Barrett 1991a).

Enantiostyly is a floral polymorphism in which flowers possess either left- or right-bending styles (Fig. A.1c). It is most commonly associated with a stamen dimorphism in which one anther is positioned opposite the stigma and the remaining anthers are located together elsewhere in the flower. Such stamen dimorphism is referred to as heteranthery where, as in Pontederiaceae and several other enantiostyloous taxa, it reflects a functional division of labor into predominantly attractive, contrastingly-colored "feeding" anthers and a cryptically-colored "pollinating" anther (Müller 1883; Iyengar 1923; Vogel 1978; Buchmann 1983; Lloyd 1992; Chapter 2). There have been few attempts to investigate the ecological significance of enantiostyly (Bowers 1975, Fenster 1995) and while its function as a mechanism promoting inter-flower and, perhaps, inter-plant pollen dispersal seems likely, experimental evidence in support of this hypothesis is lacking. While the enantiostyloous floral syndrome can apparently exist as a true genetic polymorphism (Ornduff and Dulberger 1978), it usually occurs as a somatic polymorphism with right- and
left-handed flowers occurring within the same individual, as is the case in enantiostylous species of *Monochoria* and *Heteranthera*.

Among all four major genera of Pontederiaceae are species that lack either the tristylos or enantiostylos floral syndromes. Flowers are uniform with respect to style and stamen characters and we therefore refer to them as monomorphic. In *Eichhornia*, monomorphic species are small-flowered, self-compatible and largely self-pollinating, with one set of anthers positioned close to the stigma (Fig. A.1b). This floral condition is referred to as homostyly following Darwin (1877) and other workers on heterostylos groups (e.g. Ornduff 1972; Ganders 1979), who have assumed that homostylos species are derived from heterostylos ancestors through loss of floral morphs and selection for self-fertilization. Evidence in support of this hypothesis in *Eichhornia* has come from microevolutionary studies in which selfing, homostylos populations have been documented at the margins of the range in each of the three tristylos species (Barrett 1978b, 1979, 1985). In *E. paniculata*, the evolutionary pathway from trimorphism and outcrossing to monomorphism and predominant self-fertilization is associated with major changes in floral architecture and has been the focus of intensive microevolutionary investigation over the past decade (Glover and Barrett 1986; Barrett, Morgan and Husband 1989; Barrett and Husband 1990; Barrett, Kohn and Cruzan 1992; Fenster and Barrett 1994; Kohn and Barrett 1994).

Variation in floral syndromes and mating-systems, a wealth of microevolutionary information, and the relatively small size of the family make Pontederiaceae particularly appropriate for phylogenetic analysis of the evolution of reproductive characters. A major objective in this study was to use molecular phylogenetic evidence to reconstruct the history of evolutionary transitions in floral characters. We posed three questions about the evolution and breakdown of floral syndromes, paying particular attention to character-state changes associated with the buildup and breakdown of tristyly: (1) What is the evolutionary
relationship between tristyly and enantiostyly and were there single or multiple origins of each polymorphism? (2) What is the order of character-state changes leading to the buildup of the tristylos floral syndrome? In particular, is self-compatibility the ancestral or derived condition among tristylos species? (3) Is homostyly among selfing taxa of *Eichhornia* homologous or were there multiple transitions from tristylos, outcrossing taxa to homostylos, selfing ones? We next develop the rationale for each of these questions.

Tristyly has been well documented in only four unrelated angiosperm families (Lythraceae, Oxalidaceae, Pontederiaceae, Amaryllidaceae; reviewed in Barrett 1993) and therefore its origin is likely to have been an infrequent event. The rarity of tristyly is undoubtedly associated with its developmental and genetic complexity (Charlesworth 1979; Lewis and Jones 1992). Little is known of the floral characteristics of the immediate ancestors of tristylos species. Based on their developmental studies of various tristylos groups, Richards and Barrett (1992) proposed that a within-flower stamen dimorphism was probably a prerequisite for the evolution of tristyly, raising the possibility of an evolutionary linkage to enantiostyly. Several monocotyledonous families (e.g. Commelinaceae, Tecophilaeaceae, Haemodoraceae, Philydraceae), some of which are closely allied to Pontederiaceae (see Chapter 2), also contain enantiostylos taxa and it is therefore possible that enantiostyly in Pontederiaceae is homologous to that found in related families.

An important issue concerned with the evolution of heterostyly is the order of establishment of the morphological and physiological components of the syndrome (Barrett 1992b). For example, models of the evolution of distyly conflict as to whether heteromorphic incompatibility evolves before reciprocal herkogamy (Charlesworth and Charlesworth 1979) or after it (Lloyd and Webb 1992a,b). Charlesworth (1979) proposed that a distylos condition (with two levels of anthers in each flower) serves as an intermediate stage in the transition from monomorphism to tristyly. In her model, modifiers
that increase the strength of heteromorphic incompatibility spread, though perhaps not to fixation, before floral trimorphism becomes established. Where self-compatibility occurs in tristylos taxae, it has usually been assumed to represent a secondary loss from a self-incompatible ancestor (Ornduff 1972; Weller 1992). Where heterostylous taxae within a clade differ with regard to the presence or absence of incompatibility, phylogenetic reconstruction of character-state evolution may allow us to determine the sequence of evolutionary events.

Microevolutionary studies of the dissolution of tristyly in *Eichhornia* suggest that the polymorphism has broken down repeatedly, giving rise to selfing, homostylous species (e.g. Barrett 1988a; Husband and Barrett 1993). However a phylogenetic study of Pontederiaceae involving 34 taxa and 42 morphological characters, 27 of which were reproductive traits, suggested instead that the shift from tristyly to homostyly occurred once, resulting in a monophyletic group of selfing species (Eckenwalder and Barrett 1986). This result was not in accord with previous population-level studies and Eckenwalder and Barrett (1986) proposed that the apparent monophyly of homostylous *Eichhornia* species may have resulted from the repeated evolution of traits associated with a selfing syndrome (see also Wyatt 1988). In the present study, we use restriction-site variation of the chloroplast genome because this source of data is likely to be relatively free of multiple convergences caused by similar selective forces acting on reproductive characters in different lineages.

Reconstruction of character evolution commonly involves mapping traits onto a phylogenetic tree. This is usually achieved by treating all shifts in characters as equally likely events (e.g. Brooks and McLennan 1991). Although details of the optimization method are not always explicitly stated, this has generally been the approach employed in phylogenetic studies concerned with the evolution of reproductive traits in flowering plants (Hart 1985; Eckenwalder and Barrett 1986; Donoghue 1989; Olmstead 1989; Cox 1990,
Sytsma, Smith and Berry 1991; Rieseberg, Hanson and Philbrick 1992; Armbruster 1993; Weller, Donoghue and Charlesworth 1995; Bruneau 1993). Equal weighting is methodologically simpler and is used to avoid subjectivity in weighting various classes of evolutionary change. However, there is abundant evidence that certain evolutionary transitions are more likely than others, such as shifts from diploidy to polyploidy in plants (Stebbins 1950) or the evolution of flightlessness in island-inhabiting birds (Carlquist 1974), compared to the reverse changes. Where external evidence bearing on the likelihood of particular character-state changes is available, it would seem wise to take account of this information in reconstructing the course of evolution (Maddison and Maddison 1992). Hence, an additional goal of our study was to explore the importance of different optimization schemes on the interpretation of the evolution of reproductive characters in Pontederiaceae.

MATERIALS AND METHODS

*Taxon sampling.* — Localities for the 24 accessions from Pontederiaceae and three outgroup taxa used in this study are given in Appendix B. The 24 taxa from Pontederiaceae included eight species of *Eichhornia*, four taxa of *Pontederia*, five species of *Heteranthera*, four species of *Monochoria*, the monotypic *Hydrothrix*, and one species each from *Reussia* and *Zosterella*. Outgroup taxa representing Commelinaceae, Philydraceae and Liliaceae were used. Eckenwalder and Barrett (1986) used Liliaceae to root their morphological phylogeny of Pontederiaceae. Based on morphological evidence, Commelinaceae and especially Philydraceae have been closely allied with Pontederiaceae by several workers (Hamman 1966; Takhtajan 1969; Dahlgren, Clifford and Yeo 1985; Thorne 1992b); evidence from the chloroplast gene *rbcL* provides only limited support for these alliances.
APPENDIX A. RESTRICTION-SITE BASED PHYLOGENY OF PONTEDERIACEAE

(Chase et al. 1993; Clark et al. 1993; Chapter 2). An undescribed species of Eichhornia (referred to here as Eichhornia sp.) was originally identified in Eckenwalder and Barrett (1986) as E. paradoxa (Mart.) Solms-Laub. Differences in a range of vegetative and reproductive traits between this accession and E. paradoxa (S.C.H. Barrett, unpubl. data) led to a study of differentiation between these two taxa. F₁ hybrids proved vigorous but sterile and a study of isozyme differentiation between the taxa showed levels of divergence normally found among congeneric species (Cole and Barrett 1989).

Molecular methods. — Total genomic DNA was isolated by the method of Doyle and Doyle (1987) using approximately 1g of leaf tissue from single individuals ground initially under liquid nitrogen. For Monochoria cyanea, DNA was extracted from pooled seedlings instead of an individual plant because we were unable to maintain plants beyond the seedling stage. DNA (approximately 1-3µg) was cut with one of ten restriction enzymes (BamHI, BglII, DraI, EcoRI, EcoRV, HindIII, KpnI, PstI, SalI, XhoI) and then separated on 0.8% agarose gels and transferred onto nylon filters (Sambrook, Fritsch, and Maniatis 1989). Filters were then sequentially hybridized to 42 clones from the Nicotiana chloroplast genome (Sugiura et al. 1986; Olmstead and Palmer 1992). Clones were radiolabeled with 32P-labeled dATP using the random hexanucleotide priming method (Feinberg and Vogelstein 1983) and bands were visualized using autoradiography (see Doyle et al. 1990, for hybridization methods). Restriction sites were inferred by examination of autoradiogram banding patterns from adjacent cpDNA clones.

Tree reconstruction. — Taxa were coded 0 for absence and 1 for presence of inferred restriction sites and analyzed using PAUP version 3.1.1 (Swofford 1993). The data matrix is available on request from the senior author. Sites were scored wherever unambiguous but some sections of the chloroplast genome for some enzyme-taxon combinations were not mapped due to poor hybridization or difficulty in site reconstruction among divergent taxa. Heuristic searches were performed using Fitch optimization with tree bisection-
reconnection (TBR) branch-swapping and the “MULPARS” and “Steepest descent” options in effect. To evaluate the effectiveness of the heuristic search, 100 random addition replicates were performed (Maddison 1991). These found no additional trees of equal or shorter length. The ingroup was monophyletic in all trees. Separate analyses were performed in which outgroup taxa were excluded or included. To assess tree robustness, two bootstrap analyses were performed, one with the three outgroup taxa and one without, using the same heuristic search settings as before. The addition sequence was random for each of the 200 bootstrap replicates performed.

Character optimization. — Reproductive characters were optimized onto trees in two independent sets of analyses, one involving floral form (tristyly, enantiostyly, and monomorphism) and the other the self-incompatibility status of taxa in Pontederiaceae. For both floral conditions and self-incompatibility status, two different types of optimization scheme were assessed. In one, all character state transitions were treated as unordered and equally-likely (equally-weighted optimization scheme, Table A.1a, Fig. A.4b). The second type of scheme (hereafter “2:1” weighting, Table A.1b, Fig. A.4a,c,d) was also unordered but applied weights in the following manner: for floral form, the loss of tristyly or enantiostyly was favored by a two-fold margin over their gain or interconversion; for self-incompatibility, the loss of self-incompatibility was favored by a two-fold margin over its gain. We discuss more fully below our rationale for favoring reconstructions which use these unequally-weighted schemes. Reconstructions were performed using the “State Changes and Stasis” option of MacClade version 3.0 (Maddison and Maddison 1992).

Two \textit{a posteriori} methods were used to root trees resulting from the phylogenetic analysis of ingroup taxa: (1) The most parsimonious rooting (the branch to which the outgroups join to the family when these taxa are included in the searches), and (2) Midpoint rooting (the midpoint of the longest path connecting any pair of taxa). The instability of the root position found using outgroup analysis (see results) motivated the examination of the
APPENDIX A. RESTRICTION-SITE BASED PHYLOGENY OF PONTEDERIACEAE

effect of alternative root position on character reconstruction and midpoint rooting was considered as one example of a slightly sub-optimal root position. Although Pontederiaceae is a monophyletic group within the monocotyledons (Chapter 2), the identity of the sister-group to Pontederiaceae is unclear with respect to current molecular and morphological evidence (Chapter 2; Clark et al. 1993). A hypothetical outgroup was therefore placed at the base of trees for the root positions discussed above, and was coded either as monomorphic, enantiostylous or tristylious to examine the effect of different outgroup codings on the reconstruction of shifts in floral condition within the family. Heteromorphic sporophytic self-incompatibility systems are unknown in the monocotyledons outside Pontederiaceae (Charlesworth 1985; Weller, Donoghue and Charlesworth 1995). The homomorphic gametophytic self-incompatibility systems found in some monocotyledonous groups (e.g. in some taxa of Commelinaceae) are unlikely to be homologous with heteromorphic sporophytic incompatibility in Pontederiaceae. The hypothetical outgroup was therefore coded as self-compatible.

RESULTS

Phylogenetic structure of the family. — We scored 356 restriction sites of which 292 were variable and 104 were potentially informative within Pontederiaceae. Twenty most-parsimonious trees of length 492 steps and with a consistency index (CI) of 0.593 (0.464 excluding uninformative characters), retention index (RI) of 0.692 and rescaled consistency index (RC) of 0.411 were found in the analysis of the entire data set. When included, the outgroup taxa joined Pontederiaceae along the terminal branch subtending Eichhornia crassipes in all most-parsimonious trees. Ten shortest unrooted trees of length 299 steps, CI = 0.582 (0.454 excluding uninformative characters), RI = 0.718 and RC = 0.418 were
found when outgroup taxa were excluded. Strict consensus trees of the trees from analyses in which outgroup taxa were either included or excluded are presented in Fig. A.2. Bootstrap values are indicated on branches seen in 50% or more of replicates. When outgroup taxa were excluded, the phylogenetic structure of the family is quite well resolved. Most branches have bootstrap support of 70% or higher. Simulation studies by Hillis and Bull (1993) suggest that bootstrap values of about 70% or more represent well supported branches. The inclusion of outgroup taxa in the analysis led to a collapse of support for several of the branches at the base of the rooted tree, indicating uncertainty in the position of the root of the family. In particular, the rooting seen in all most parsimonious rooted trees, with *E. crassipes* sister to the rest of Pontederiaceae (Fig A.3), was supported by only 27% of bootstrap replicates.

In all unrooted trees, three of the four main taxonomic groups, *Pontederia* s.l., *Monochoria* and *Heteranthera* s.l., of Pontederiaceae are monophyletic. Two separate clades of *Eichhornia* occur that are each composed of a tristyloous species together with two monomorphic species of *Eichhornia*. Tristyloous *E. paniculata* is sister to the clade consisting of monomorphic *Eichhornia* sp. and *E. paradoxa* while tristyloous *E. azurea* forms a clade with monomorphic *E. heterosperma* and *E. diversifolia*. *Eichhornia crassipes* and *E. meyeri* were not part of either clade. The two enantiostyloous genera (*Monochoria* and *Heteranthera* s.l.) did not form a monophyletic group, but instead were well separated on the trees (Figs. A.2, A.3).

Relatively minor differences in topology were observed among the ten shortest unrooted trees. These varied in three ways (Fig. A.3): Within *Pontederia*, *P. sagittata* was seen in the two different positions shown. Within *Monochoria*, a clade consisting of *M. korsakovii*, *M. vaginalis* and *M. hastata* varied in their relative branching order such that all three possible branching orders of these taxa were found (two of the three are shown). Two major topological classes are defined by the relative positions of *E. crassipes* and the
Fig. A.2. Strict consensus of the most-parsimonious phylogenetic trees found in heuristic searches based on chloroplast DNA restriction site variation found in Pontederiaceae (see text). Outgroup taxa were either included (left side) or excluded (right side) from the analysis. Bootstrap support for phylogenetic structure in the consensus trees is indicated above each branch. Branches with less than 50% bootstrap support are indicated by "(--)". Indicated below each branch is the range of lengths across the shortest trees (computed using ACCTRAN optimization). Tree lengths and summary statistics are provided in the text. The symbols beside each taxon indicate its floral form and self-incompatibility status.
APPENDIX A. RESTRICTION-SITE BASED PHYLOGENY OF PONTEDERIACEAE

Fig. A.3. Two of the ten shortest maximum parsimony trees from a heuristic search involving only taxa of Pontederiaceae. The two trees summarize almost all of the topological variation seen among the ten shortest trees and illustrate the two classes (A and B) of topological variation important for reconstructing the evolution of reproductive characters (see text and Fig. A.4). Each class is defined with respect to two possible positions of *Eichhornia crassipes* relative to *Monochoria*. The three arrows indicate topological shifts which would convert one tree into the other. The branch containing the tree midpoint is indicated with an asterisk. ACCTRAN optimization was used to compute all branch lengths.
APPENDIX A. RESTRICTION-SITE BASED PHYLOGENY OF PONTEDERIACEAE

genus *Monochoria* within the family. With respect to the *a posteriori* rooting indicated in Fig. A.3, *Monochoria* was either the sister-group of the clade consisting of *Pontederia s. l.* and the clade of *Eichhornia* species that includes *E. azurea* (hereafter referred to as "class A topology"), or it was the sister-group of all taxa in the family excluding *E. crassipes*, ("class B topology"). The ten most-parsimonious unrooted trees all have topological equivalents among the twenty most-parsimonious rooted trees.

*Reconstruction of reproductive character evolution.* — Reproductive characters were optimized onto these ten shortest trees in the separate analyses for floral form and self-incompatibility status. Apart from the two possible positions of *E. crassipes* relative to *Monochoria* shown in Fig. A.3, other topological variations among the shortest trees had no effect on character reconstructions since all taxa in the relevant clades possess the same character state.

Table A.1a,b summarizes shifts in floral condition reconstructed across the shortest trees for the different outgroup codings and rootings using the two different optimization schemes. Under the first scheme all floral shifts were treated as equally-weighted, unordered character-state transitions, and up to four reconstructed gains of tristyly are possible (Table A.1a, Fig. A.4b). However, several lines of evidence indicate that this optimization scheme is unlikely to reflect the relative likelihood of the gain versus loss of a complex polymorphism such as tristyly (see below). The second optimization scheme was also unordered, but the weights used favored the loss of tristylos or enantiostylous floral syndromes over their gain by a two-fold margin (Table A.1b, Fig. A.4a,c,d). For trees with class B topology, it is equivocal whether tristyly arose once or twice, if outgroup rooting is employed and the outgroup is coded as enantiostylos (see Fig. A.4d). Under such conditions, tristyly arises either a single time prior to the origin of the extant lineages of Pontederiaceae, or it arises twice, once within the *E. crassipes* lineage, and once prior to
Table A.1. Reconstructed number of shifts of floral form in Pontederiaceae. The range of frequencies of shifts is summarized for all most-parsimonious reconstructions of floral form on the ten shortest chloroplast-based trees. (a) Equally weighted optimization scheme; (b) “2:1” optimization scheme. Outgroup and midpoint rootings were assessed in conjunction with the three possible codings of the outgroup floral form. The primitive floral form in the family is listed as equivocal if ambiguous for at least some of the trees.

(a).

<table>
<thead>
<tr>
<th>Outgroup Coding</th>
<th>Monomorphic</th>
<th>Enantiostylys</th>
<th>Tristylys</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Floral Shift</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equally-Weighted Optimization Scheme, Outgroup Rooting:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tristyly Gained</td>
<td>4</td>
<td>0 - 4</td>
<td>0 - 3</td>
</tr>
<tr>
<td>Tristyly Lost</td>
<td>0</td>
<td>0 - 6</td>
<td>1 - 5</td>
</tr>
<tr>
<td>Enantiostyly Gained</td>
<td>2</td>
<td>1 - 3</td>
<td>2</td>
</tr>
<tr>
<td>Enantiostyly Lost</td>
<td>1</td>
<td>1 - 5</td>
<td>1 - 2</td>
</tr>
<tr>
<td>Total Events and Steps</td>
<td>7</td>
<td>7 - 8</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primitive Floral Form</th>
<th>Monomorphic</th>
<th>Equivocal</th>
<th>Tristylys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equally-Weighted Optimization Scheme, Midpoint Rooting:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tristyly Gained</td>
<td>1 - 4</td>
<td>0 - 4</td>
<td>0 - 1</td>
</tr>
<tr>
<td>Tristyly Lost</td>
<td>0 - 4</td>
<td>0 - 6</td>
<td>4 - 5</td>
</tr>
<tr>
<td>Enantiostyly Gained</td>
<td>2</td>
<td>1 - 3</td>
<td>2</td>
</tr>
<tr>
<td>Enantiostyly Lost</td>
<td>1 - 2</td>
<td>1 - 4</td>
<td>1</td>
</tr>
<tr>
<td>Total Events and Steps</td>
<td>7</td>
<td>8</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primitive Floral Form</th>
<th>Monomorphic</th>
<th>Equivocal</th>
<th>Tristylys</th>
</tr>
</thead>
</table>
Outgroup Coding

<table>
<thead>
<tr>
<th>Floral Shift</th>
<th>Monomorphic</th>
<th>Enantiostyly</th>
<th>Tristyly</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;2:1&quot; Optimization Scheme, Outgroup Rooting:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tristyly Gained</td>
<td>0</td>
<td>0 - 2</td>
<td>0</td>
</tr>
<tr>
<td>Tristyly Lost</td>
<td>6</td>
<td>4 - 6</td>
<td>5</td>
</tr>
<tr>
<td>Enantiostyly Gained</td>
<td>2</td>
<td>1 - 3</td>
<td>2</td>
</tr>
<tr>
<td>Enantiostyly Lost</td>
<td>1</td>
<td>1 - 3</td>
<td>1</td>
</tr>
<tr>
<td>Tristyly to Homostyly</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Tristyly-Enantiostyly Interconversion</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Total Events (Steps)</td>
<td>8 (10)</td>
<td>8 (11)</td>
<td>7 (9)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primitive Floral Form</th>
<th>Monomorphic</th>
<th>Equivocal</th>
<th>Tristyly</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;2:1&quot; Optimization Scheme; Midpoint Rooting:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tristyly Gained</td>
<td>0 - 1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Tristyly Lost</td>
<td>4 - 6</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Enantiostyly Gained</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Enantiostyly Lost</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Tristyly to Homostyly</td>
<td>3 - 5</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Tristyly-Enantiostyly Interconversion</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Total Events (Steps)</td>
<td>7 - 8 (10)</td>
<td>8 (10)</td>
<td>7 (9)</td>
</tr>
<tr>
<td>Primitive Floral Form</td>
<td>Equivocal</td>
<td>Enantiostyly</td>
<td>Tristyly</td>
</tr>
</tbody>
</table>
Fig. A.4. Example reconstructions of breeding-system evolution for selected maximum-parsimony trees. Reconstruction of character evolution was performed using MacClade version 3 (Maddison & Maddison, 1992). *Monochoria cyanea* was coded as uncertain for floral form (i.e., as monomorphic or enantiostyious). Floral form and self-incompatibility were each assessed independently. Step matrices employed for reconstructions are depicted in each figure. All optimization schemes were unordered. The outgroup was coded as self-compatible for all reconstructions of shifts in compatibility status. a) Character state reconstruction using unequal weighting schemes, class A tree topology (see Fig. A.3), the outgroup coded as monomorphic for floral form, and the root location found by outgroup analysis (see Fig. A.2). The reconstruction of self-incompatibility was equivocal under "2:1" weighting and was resolved here using ACCTRAN optimization (see text). b) Same conditions as Fig. A.4a, except that all character transformations were weighted equally. c) Same conditions as Fig. A.4a, except that midpoint rooting was employed. d) Unequal weighting schemes, Class B topology (see Fig. A.3), outgroup rooting, and the outgroup coded as enantiostyious. This reconstruction illustrates a scenario where the unequal weighting scheme for floral form can lead to two origins of tristyly.
Floral Form (8 events)  Self-Incompatibility (3 events)

"2:1" optimization scheme

FROM:  TO:

Outgroup
P. cordata v. ovalis
P. cordata v. cordata
P. cordata v. lancifolia
P. sagittata
P. rotundifolia
E. azurea
E. heterosperma
E. diversifolia
E. paniculata
Eichhornia sp.
E. paradoxa
E. meyeri
H. limosa
H. rotundifolia
H. oblongifolia
H. seubertiana
H. dubia
H. zosterifolia
Hydrothrix gardneri
M. korsakovii
M. vaginalis
M. hastata
M. cyanea
E. crassipes

Class B Topology, Outgroup Rooting

Ponederia s.l.
Heteranthera s.l.
Monochoria
the origin of the other tristylosous taxa. Under all other combinations of root-placement, ingroup topology, and outgroup coding using the "2:1" weighting scheme, all instances of tristyly are inferred to be homologous within the family.

Using MacClade, the number of gains of particular character states does not necessarily give an accurate account of the total number of homoplastic (i.e. non-homologous) occurrences of each state. This is because a character state can have \( x \) separate homoplastic instances across a tree, but have only \( x-1 \) gains reported. The "missing" gain is a symplesiomorphic (shared primitive) instance of the character retained from the base of the tree, hence cases with zero gains of tristyly in Table A.1a and A.1b. In some cases the number of gains of enantiostyly indicated in Table A.1b under-reports the number of homoplastic incidences. We examined all cases where a single gain of tristyly occurred (Table A.1b) for the "2:1" weighting scheme and confirmed that these corresponded to a single origin of tristyly on the tree. In all cases with the "2:1" weighting scheme, there were at least two homoplastic cases of enantiostyly on the tree. In some trees with the outgroup coded as enantiostylous, there was an additional homoplastic reconstruction of its occurrence on the branch leading to the outgroup, hence cases with three gains of enantiostyly (Table A.1b). Finally, when "2:1" weighting is used, from three to five shifts from tristyly to homostyly are inferred under all combinations of outgroup coding, root placement, and ingroup topology. A total of seven to eight events (each event representing a single shift in floral form) occurred across all classes of rooting, outgroup coding and optimization scheme.

Using the "Enforce topological constraints" option of PAUP we examined the hypothesis that monomorphic species form a monophyletic clade, as was found in several of the morphological cladograms of Eckenwalder and Barrett (1986, Fig. 1.2). Shortest trees in our search were 26 steps (5.28%) longer than the most parsimonious unconstrained trees. Constraining only \( E. \) diversifolia, \( E. \) heterosperma, \( E. \) paradoxa and \( Eichhornia \) sp.,
three of the four selfing taxa used by Eckenwalder and Barrett (1986), to be monophyletic resulted in shortest trees 23 steps (4.67%) longer than the most-parsimonious trees. Both scenarios therefore require a substantial decrease in parsimony.

Reconstructions of the evolution of heteromorphic incompatibility also differ according to the weighting scheme used to model its gain and loss. With an equal weighting scheme, incompatibility is seen to arise on two separate occasions in the family; once along the terminal branch leading to *E. azurea*, and once in the lineage leading to *Pontederia s.l.* (shown in Fig. A.4b; shifts in incompatibility are represented by bars ).

With a "2:1" weighting scheme for the gain versus loss of incompatibility, however, it is equivocal whether it arises once or twice. When delayed transformation (DELTRAN) is employed to resolve this ambiguity, two origins of incompatibility are inferred on these two branches. With accelerated transformation (ACCTRAN), a single gain of heteromorphic incompatibility is inferred along the branch supporting the clade consisting of *E. azurea*, *E. heterosperma*, *E. diversifolia* and *Pontederia s.l.* (as shown in Fig. A.4a,c,d). This single gain is followed by reversions to self-compatibility along the terminal branches leading to the two monomorphic species in this clade. In fact, all weighting schemes favoring the gain over the loss of SI by a greater than two-fold margin result in a single origin and two losses of incompatibility in the family. Regardless of which scenario is inferred, incompatibility always arises several branch segments after the origin of tristyly, when the "2:1" weighting scheme is used to reconstruct shifts in floral form, for all topologies and rootings examined here.
APPENDIX A. RESTRICTION-SITE BASED PHYLOGENY OF PONTEDERIACEAE

DISCUSSION

Analysis of chloroplast DNA restriction-site variation produced a highly resolved phylogenetic hypothesis for Pontederiaceae with many groupings within the family strongly supported by the bootstrap analysis. However, topological uncertainties involving the root placement and the relative positions of *Eichhornia crassipes* and *Monochoria* in the family posed complications for reconstruction of the evolution of reproductive characters. In this discussion we first take up the systematic implications of the molecular data and present the implications of infrafamilial relationships for understanding shifts in floral characters. Next we discuss the reconstruction of floral character state transitions using different optimization schemes and alternative topologies. We show that the class of optimization scheme employed is a critical issue in interpreting character evolution, particularly where biological evidence dictates that equal-weighting of evolutionary transitions is of dubious validity.

*Systematics of Pontederiaceae*

Many of the relationships among taxa within Pontederiaceae were well resolved, allowing several conclusions to be drawn from the phylogenetic analysis of chloroplast characters. The degree of divergence of the chloroplast genomes of *Eichhornia* sp. and *E. paradoxa* (Fig. A.3) is consistent with the proposed species-level status of each taxon. Based on the root placement indicated by outgroup analysis, *Monochoria, Pontederia s.l.* and *Heteranthera s.l.* are all monophyletic. However, bootstrap support for this root placement is well below 50%. Bootstrap support for several other branches, including the one supporting the monophyly of *Monochoria*, falls below 50% when outgroups are
APPENDIX A. RESTRICTION-SITE BASED PHYLOGENY OF PONTEDERIACEAE

included in the analysis (Fig. A.2). We interpret this collapse in support as being a function of the uncertainty in root placement, since these branches are well supported when outgroups are excluded from the analysis. Finally, the genus *Eichhornia* is an "unnatural" or polyphyletic group, regardless of the placement of *E. crassipes*.

The apparent polyphyly of *Eichhornia* could have a number of causes. It is always possible that the chloroplast "gene" tree is not an accurate reflection of the "species" tree of the organisms in which the chloroplasts reside, due to phenomena such as ancient introgression of chloroplast genomes (reviewed in Doyle 1992). While there is little evidence for hybridization among extant species in Pontederiaceae (S.C.H. Barrett, unpubl. data), historical introgression of chloroplasts could have led, for example, to the capture by the ancestor of *Monochoria* of the chloroplast genome of an early species of *Eichhornia*. Phylogenetic evidence from several unlinked nuclear genes may help clarify whether such events have occurred at this or any other point on the tree. However, it is worth noting that phylogenetic evidence based on morphology also indicates that *Eichhornia* is at least paraphyletic, if not polyphyletic (Eckenwalder and Barrett 1986). Rather than appealing to phenomena such as ancient hybridization to explain the dispersed placement of *Eichhornia* species across Pontederiaceae, it may be safer to assume, in the absence of evidence to the contrary, that the unnaturalness of the genus *Eichhornia* is real.

Implications of infrafamilial relationships for reproductive character evolution. — Despite the uncertainty of root placement, at least two conclusions are warranted on the basis of ingroup topology alone. The naïve expectation that species sharing the same complex floral syndrome would be grouped together is not supported by molecular data. In fact, because of the manner in which tristyrous and enantiostyrous groups are interspersed on the tree, at least one of these complex polymorphisms must have multiple origins in the family regardless of root placement or ancestral floral form. Second, monomorphic species of *Eichhornia* are not monophyletic. Both the bootstrap analysis and use of the "Constraints"
option in PAUP confirm that trees which unite the monomorphic species of *Eichhornia* are highly unlikely.

Eckenwalder and Barrett (1986) suspected that monomorphic species of *Eichhornia* grouped together in their morphological phylogenetic reconstruction because of convergent evolution of characters associated with selection for a selfing syndrome. When morphological and molecular phylogenetic data conflict, it sometimes may be difficult to determine which form of data is more reliable. However, there is a widespread belief that the evolution of the selfing syndrome involves multiple morphological convergences (e.g. Ray and Chisaki 1957; Lloyd 1965; Ornduff 1969; Wyatt 1984; Ritland and Ritland 1989). The possibility that apparently synapomorphic morphological characters associated with selfing are actually homoplastic suggests that classes of characters not known to be involved in this syndrome, such as the cpDNA characters used here, are particularly useful when testing phylogenetic hypotheses concerning the evolution of autogamy. Based on their intuition and limited cytogenetic evidence, Eckenwalder and Barrett (1986) predicted particular associations between tristyloous and homostyloous species of *Eichhornia*: *E. paradoxa* with *E. paniculata*, *E. heterosperma* with *E. azurea*, and *E. diversifolia* with *E. crassipes*. Our data lend support to some of these suggestions. *Eichhornia paradoxa* and *Eichhornia* sp. represent a clade sister to *E. paniculata* while *E. heterosperma* is allied with *E. azurea*. However, *E. diversifolia* is sister to *E. azurea* and *E. heterosperma*, rather than having its suspected close relationship to *E. crassipes*.

Reconstruction of character evolution

Conclusions concerning the number and direction of character-state transitions require careful reconstruction of reproductive characters on the phylogenetic trees. Progress
APPENDIX A. RESTRICTION-SITE BASED PHYLOGENY OF PONTEDERIACEAE

has been made in reconstructing the evolutionary history of character-state transitions despite uncertainties in outgroup identity and character state, root placement, and ingroup structure. These difficulties make it hard to determine the precise points of particular evolutionary shifts, but the following broad conclusions can be drawn: (1) tristyly broke down to homostyly on multiple occasions within the family, (2) instances of enantiostyly in Monochoria and Heteranthera are not homologous, (3) there was probably one, but perhaps also a second, origin of tristyly, and (4) heteromorphic incompatibility arose after the origin of floral trimorphism.

Optimization Schemes. — The conclusions stated above are sensitive to whether character-state transformations are uniformly or non-uniformly weighted (Table A.1a,b; Figure A.4a-d). They largely depend upon acceptance of the unequally-weighted optimization scheme for floral form. How can we justify choosing between different optimization scenarios?

In the absence of external evidence concerning the relative difficulty of particular evolutionary shifts, it may be operationally simpler to assume an equally-weighted optimization scheme. However, equally-weighted optimization schemes are themselves strong assumptions (Swofford and Olsen 1990). Several lines of evidence indicate that the origin of tristyly is an improbable occurrence relative to its evolutionary dissolution. Tristyly is perhaps the most complex breeding system in flowering plants. Its relatively elaborate developmental basis (Richards and Barrett, 1992) and great rarity within the angiosperms (Barrett 1993) certainly argue that it is difficult to evolve. A range of microevolutionary evidence also demonstrates that, at the population level at least, tristyly readily breaks down to yield predominantly self-fertilizing floral variants both in Pontederiaceae (Barrett, Morgan and Husband 1989; Husband and Barrett 1992, 1993; Fenster and Barrett 1994) and other tristyous families (e.g., Stout 1925; Mayura Devi and Hashim 1966; Ornduff 1972; Eckert and Barrett 1994). Where such "external" evidence exists concerning the relative difficulty of particular evolutionary events, it would be
unwise to ignore it (Maddison and Maddison 1992; p.171). Given that all possible shifts in floral form are not equally probable, it would consequently be inappropriate to give all possible character-state transformations equal weighting during character reconstruction.

Dollo parsimony (Le Quesne 1975; Farris 1977a, 1977b) is an extreme example of unequal-weighting which traditionally has been used as a criterion for modeling the evolution of complex characters, or for using those characters in the reconstruction of phylogenies. An infinitely large cost is used to reject multiple gains of a complex character, but multiple losses are possible under the logic that complexity is much easier to lose than originate. Dollo weighting has been criticized on the grounds that the infinite penalty against homoplasy in complex structures is unrealistic (Swofford and Olsen 1990; Albert, Mishler and Chase 1992). Our "2:1" weighting scheme is an example of a "relaxed" Dollo criterion (Swofford and Olsen 1990) whereby single-gain, multiple-loss scenarios are preferred but not demanded during character reconstruction.

Linear, quadratic or logarithmic transformations have been used to translate the probabilities of particular evolutionary events into weighting schemes (Felsenstein 1981; Williams and Fitch 1989, 1990; Wheeler 1990). The two-fold weighting imbalance we employed may correspond to greater than two-fold differences in the relative probabilities of transformation. Weighting imbalances similar to the one we used have been suggested for use when reconstructing phylogenies with molecular data (e.g. Albert, Mishler and Chase 1992). There is empirical evidence that different classes of change at individual nucleotides or restriction sites occur at different rates. The origin of a floral polymorphism such as tristyly, or a physiological mechanism such as self-incompatibility, is assuredly a much more complex evolutionary event than a substitution event in a DNA sequence, but it is less obvious how to rate the evolutionary difficulty of the origin or dissolution of such characters.
With equal-weighting, tristyly may arise up to four times in the family (Table A.1a; Fig. A.4b). Under the "2:1" scheme for shifts in floral form, a dramatic shift occurs in the pattern of reconstructed events; Tristyly is gained only once or perhaps twice, but lost on multiple occasions (Table A.1b; Fig. A.4a,c,d). A greater than two-fold imbalance does not change this conclusion (S.W. Graham et al., unpubl. results), and less than two-fold biases can still suffice to reject multiple origins of tristyly in Pontederiaceae (Chapter 2, 5). The boundary weighting imbalance needed to yield this shift depends on details of tree structure, root position and outgroup coding (S.W. Graham et al., unpubl. results), but clearly the bias required to favor the scenario with a single origin of tristyly over multiple origins is not very large.

Enantiostyly is also a relatively elaborate floral form. We treated both floral polymorphisms (tristyly and enantiostyly) equivalently in the "2:1" weighting scheme, i.e., two steps for their gain, and one step for a shift to floral monomorphism. Interconversion between them was also assigned two steps. This approach is conservative in that it does not downwardly weight gains of enantiostyly, compared to gains of tristyly, even though enantiostyly is arguably a less complex floral polymorphism than is tristyly. Although the same pattern of weights was assigned to shifts involving the two polymorphic floral forms, in all reconstructions enantiostyly is homoplastic, whereas in most cases instances of tristyly are homologous across the tree. At least one shift between tristyly and enantiostyly is necessary in all reconstructions examined (Table A.1b).

Outgroup Identity and Coding. — We assessed reconstructions with the outgroup coded, in turn, as each of the three floral forms seen in the family. However, tristyly is an improbable candidate character state for the outgroup. Apart from an isolated case in the genus Narcissus (Amaryllidaceae) (Fernandes 1935; Lloyd, Webb and Dulberger 1990; Barrett, Lloyd and Arroyo 1996), it has not been reported in other monocotyledonous families. Several potential outgroup families are enantiostyloous or include enantiostyloous
taxa (for example, Commelinaceae, Haemodoraceae, Philydraceae; see Chapter 2), and floral monomorphism of one form or another is the predominant condition within the monocotyledons. Employing different outgroup codings had a marginal effect on the broad pattern of reconstructed shifts in the family, compared with changing the optimization scheme (Table A.1a,b). One obvious effect is on the reconstructed primitive floral condition in the family. This was strongly influenced by the character state assigned to the outgroup, but was not always the same as the outgroup state (Table A.1b). The issue of the primitive floral condition will remain unresolved until we have a clearer idea of the phylogenetic placement of Pontederiaceae within the monocotyledons.

Uncertainty in tree topology. — We performed character reconstructions using the two most-parsimonious positions for Monochoria relative to E. crassipes, and two alternative root positions; the most-parsimonious rooting and the midpoint of ingroup taxa. As with outgroup coding, the topological differences analyzed had little effect on broad patterns of character evolution, in comparison to the effect of the different optimization schemes. Under only one combination of topology and outgroup coding was tristyly found to be potentially homoplastic, while enantiostyly was homoplastic under all optimizations using "2:1" weighting. In addition, multiple transitions from tristyly to homostyly were always inferred (Table A.1b).

Origin of Self-Incompatibility. — Whether its gain is equally or unequally weighted, self-incompatibility is shown to arise after the one or two origins of floral trimorphism inferred when shifts in floral form are unequally weighted. Models that posit the origin of incompatibility before floral heteromorphism (Charlesworth and Charlesworth 1979; Charlesworth 1979) thus fail to find support, since self-compatibility can be a primitive condition among tristyloous species. This finding contradicts the common assumption that self-compatibility in heterostyloous taxa must necessarily represent a secondary loss from a
APPENDIX A. RESTRICTION-SITE BASED PHYLOGENY OF PONTEDERIACEAE

self-incompatible ancestor, although reversions to self-compatibility are still required in scenarios where a single gain of incompatibility is observed (Fig. A.4).

However, there are at least two potential problems in accepting this reconstruction of events. First, bootstrap analysis indicates that the position of the root is uncertain, and many root positions are possible with only slight increases in tree length. For the present study we examined only two rootings in detail. Attaching the outgroup taxa at the midpoint requires only three extra steps (495 steps or 0.6% longer) compared to the most parsimonious root position (492 steps), but other nearly-optimal root positions are possible (S.W. Graham et al., unpubl. data). Not all root positions necessarily lead to the conclusion found here concerning the order of origin of floral trimorphism and heteromorphic self-incompatibility. This issue will be explored in detail in a future paper.

Weller, Donoghue and Charlesworth (1995) treat self-incompatibility as a qualitative character during the reconstruction of its evolutionary history, as we do here. However, there is a growing body of evidence indicating that the expression of self-incompatibility can vary quantitatively, and in these cases can be difficult to distinguish from inbreeding depression (Seavey and Bawa 1986; Charlesworth 1985; Barrett 1988b). Where incompatibility is quantitative, some judgment is required in classifying species as self-incompatible or self-compatible (reviewed in Rigney et al. 1993). In our study, the self-compatibility status of taxa was based upon the criterion of full or near full (≥75%) seed set upon selfing. We coded *E. paniculata* and *E. crassipes* as self-compatible because they satisfy this criterion. However, more detailed studies of pollen-pistil interactions in *E. paniculata* indicate that this species exhibits cryptic self-incompatibility (*sensu* Bateman 1956). In all three style morphs, outcross pollen from the same level as the stigma has higher siring success than self pollen or outcross pollen from other anther levels. This pattern may represent a weaker but homologous form of the strong incompatibility found in species of *Pontederia* (Cruzan and Barrett 1993). These considerations suggest that
incompatibility might better be treated as a quantitative rather than a qualitative character in phylogenetic analysis. Quantitative measures of the strength of the incompatibility reaction in all tristyrous taxa would be needed to utilize comparative methods developed for quantitative traits (e.g. Harvey and Pagel, 1991, and references therein). Such measures, particularly competitive abilities of different pollen types when placed in mixtures on stigmas, are unavailable for most of the taxa under study. However, even if quantitative data were available for each species, it seems unlikely that the weak incompatibility found in *E. paniculata* and perhaps *E. crassipes*, would be derived relative to the strong incompatibility expressed in *E. azurea* and *Pontederia s.l.*, given the topologies examined in this study.

*The Evolution of Selfing in Eichhornia.* — All reconstructions using "2:1" weighting indicate that at least 3 transitions from tristyly to homostyly have occurred. Stebbins (1957) suggested that selfing species, because of low genetic variability, were often evolutionary dead ends and rarely gave rise to new phyletic lines. If true, phylogenetic reconstructions generally should show relatively short branch lengths between selfing taxa and their outcrossing relatives and speciation should be rare within selfing lineages. Recent work by Schoen et al. (unpubl. MS) on the phylogeny of *Amsinckia* is consistent with this view. Their work suggests that there were multiple shifts from distyly to homostyly in the genus, and they found little evidence of either great longevity of selfing lineages or speciation within these lineages. Our data, in contrast, are not entirely consistent with this view.

Homostyly is probably a shared derived condition for *E. paradoxa* and *Eichhornia* sp. (e.g. Fig. A.4a,c,d), indicating that speciation occurred within this lineage subsequent to the origin of self-fertilization. Under some reconstructions (e.g. Fig. A.4b,c) the floral monomorphism found in *Eichhornia meyeri* may be a retained primitive condition, whereas in others, it arises as a consequence of the loss of tristyly (e.g. Fig. A.4a,d). Unlike other selfing species of *Eichhornia*, all six stamens are at the level of the stigma in *E. meyeri* and
it does not possess "residual" pollen heteromorphism (Barrett 1988a). These distinctive features could either be a consequence of a substantial time depth since the loss of tristyly, or may simply reflect a retained primitive monomorphic condition for this lineage. Regardless of which interpretation is correct, it is clear from the lengths of the branches following the origin of monomorphism (Figs. A.2, A.4) that monomorphism is of ancient origin in this lineage. Both speciation within one monomorphic lineage and the ancient origin of monomorphism in another indicate that not all selfing species are phylogenetically evanescent (see also Armbruster 1993).

CONCLUSION

Uncertainties in a variety of factors can impede the reconstruction of character evolution. Our study investigated the effect of a number of these factors on the reconstruction of historical shifts in floral syndromes in Pontederiaceae. The choice of optimization scheme is perhaps the most critical issue when mapping evolutionary shifts, more important in our study than the phylogenetic uncertainties we encountered. Just as microevolutionary studies should not be performed without reference to phyletic history, phylogenetic investigations of character-state changes need to be integrated with other lines of biological evidence concerning the likelihood of occurrence of particular evolutionary transitions. Considerable genetic, developmental and ecological information is available for many heterostylos groups, and hence these taxa and their reproductive adaptations provide useful model systems for the analysis of character evolution in a phylogenetic context.
Appendix B. Source and locality of specimens of Pontederiaceae used in the analysis of chloroplast DNA restriction-site variation (Appendix A) and in Chapter 2, 3, 4 and 5 (for Pontederiaceae and Philydraceae only).

<table>
<thead>
<tr>
<th>Taxon</th>
<th>¹Source</th>
<th>Locality (Collector)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Swartz) Kunth</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>E. heterosperma</em> (Alex.)</td>
<td>Barrett 1400</td>
<td>Quixadá, Ceará, Brazil (S.C.H. Barrett)</td>
</tr>
<tr>
<td><em>E. meyeri</em> Schulz</td>
<td>Barrett 1409</td>
<td>Nuevo Asunción, Paraguay (Billiet &amp; Jadin)</td>
</tr>
<tr>
<td><em>E. paniculata</em> (Spreng.) Solms-Laub.</td>
<td>Barrett 1401</td>
<td>Population B46, Quixadá, Ceará, Brazil (S.C.H. Barrett)</td>
</tr>
<tr>
<td><em>E. paradoxa</em> (Mart.) Solms-Laub.</td>
<td>Barrett 1402</td>
<td>Pato, Paraíba, Brazil (S.C.H. Barrett)</td>
</tr>
<tr>
<td><em>Eichhornia</em> sp.</td>
<td>Barrett &amp; Shore 1399</td>
<td>Propriá, Sergipe, Brazil (S.C.H. Barrett &amp; J.S. Shore)</td>
</tr>
</tbody>
</table>

1 Each voucher is a representative individual of material under cultivation at Toronto. Source populations are listed in the locality column. One to several individuals from each population were used for DNA extractions.

2 An undescribed species of *Eichhornia* (referred to here as *Eichhornia* sp.) was identified in Eckenwalder and Barrett (1986) as *E. paradoxa*. 
<table>
<thead>
<tr>
<th>Taxon</th>
<th>Locality (Collector)</th>
<th>Source</th>
<th>barracks</th>
<th>state abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazilian (J. Doerfler)</td>
<td>Quechan University, N. Sun, Chaffee's loca.</td>
<td>Battalion 1411</td>
<td>H.</td>
<td>Nevada</td>
</tr>
<tr>
<td>Brazilian (J. Doerfler)</td>
<td>Quechan University, N. Sun, Chaffee's loca.</td>
<td>Battalion 1412</td>
<td>H.</td>
<td>Nevada</td>
</tr>
<tr>
<td>Perpermuco, Brazil (S.C.H. Battalion)</td>
<td>Pic Real do Colégio, Alagoas, Brazil</td>
<td>Battalion 1411</td>
<td>H.</td>
<td>Florida</td>
</tr>
<tr>
<td>Perpermuco, Brazil (S.C.H. Battalion)</td>
<td>Pic Real do Colégio, Alagoas, Brazil</td>
<td>Battalion 1402</td>
<td>H.</td>
<td>Florida</td>
</tr>
<tr>
<td>Brazilian (J. Doerfler)</td>
<td>Quechan University, N. Sun, Chaffee's loca.</td>
<td>Battalion 1414</td>
<td>H.</td>
<td>Nevada</td>
</tr>
<tr>
<td>Brazilian (J. Doerfler)</td>
<td>Quechan University, N. Sun, Chaffee's loca.</td>
<td>Battalion 1413</td>
<td>H.</td>
<td>Nevada</td>
</tr>
<tr>
<td>Brazilian (J. Doerfler)</td>
<td>Quechan University, N. Sun, Chaffee's loca.</td>
<td>Battalion 1406</td>
<td>H.</td>
<td>Nevada</td>
</tr>
<tr>
<td>Taxon</td>
<td>Source</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. Reussii</td>
<td>Rondevalia L.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. sagittata</td>
<td>Pretl.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. condama var. mon Mitch.</td>
<td>P. condama var. mon Mitch.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. condama var. longiflora</td>
<td>P. condama var. longiflora</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. condama var. condama L.</td>
<td>P. condama var. condama L.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. baniem</td>
<td>Bumil (I.) Pretl.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. sargentii</td>
<td>M. sargentii</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. korshakoffi</td>
<td>M. korshakoffi</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. hassman</td>
<td>M. hassman</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(S. Macintyre)</td>
<td>(S. Macintyre)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Monochaeta crassa</td>
<td>3 Monochaeta crassa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I. oceania (Collector)</td>
<td>I. oceania (Collector)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appendix B (cond.)</td>
<td>Appendix B (cond.)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plant</th>
<th>Locality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bireli (S.C. H. Batei)</td>
<td>&quot;P&quot; de Jiri, Peru, Lower Amazon Basin, Brazil</td>
</tr>
<tr>
<td>Veracruz, Vera Cruz, Mexico (D.E. Clovet)</td>
<td></td>
</tr>
</tbody>
</table>
Appendix B. (contd.)

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Source</th>
<th>Locality (Collector)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Chlorophyllum</em> sp. (Liliaceae)</td>
<td>-na-</td>
<td>In cultivation at Botany Greenhouse, University of Toronto, Toronto, Canada (J.R. Kohn)</td>
</tr>
<tr>
<td><em>Philydrum lanuginosum</em> * Banks &amp; Solander (Philydraceae)</td>
<td>Graham &amp; Barrett 1</td>
<td>In cultivation at Botany Greenhouse, University of Toronto, Toronto, Canada (S.W. Graham)</td>
</tr>
<tr>
<td><em>Tradescantia</em> sp. (Commeinaceae)</td>
<td>-na-</td>
<td>In cultivation at Botany Greenhouse, University of Toronto, Toronto, Canada (J.R. Kohn)</td>
</tr>
</tbody>
</table>
APPENDIX C

Appendix C. Revision of the morphological data set of Eckenwalder & Barrett (1986).

Revised character states are given for affected taxa, with sources. One new character (number 43) is added to the matrix, and some problematical characters and codings are discussed. Several character states have new names; the original name is given in parentheses. One character is deleted (number 19), but the numbering system used in Eckenwalder and Barrett (1986) is retained for consistency. Additional character states are added in several instances.

Character 1. Duration. States: 0 = "long-lived perennial"; 1 = "short-lived perennial"; 2 = "annual." Revised codings: E. paniculata (2)a; E. heterosperma (0)a; M. hastata (0)a; M. korsakovii (1)a.

Character 4. Habit. States: 0 = "erect"; 1 = "procumbent." Revised coding: M. hastata (0)b.

Character 7. Leaf whorls ("Axillary dwarf shoots"). All but the first of the members of each leaf whorl in H. gardneri may arise from one or more short shoots with annular insertion on the long shootc,d, or alternatively, they may be intercalary leaves produced from a meristematic ring below the shoot apexd. By either interpretation, the leaf whorls of this species are a unique feature of this taxon and the coding employed in Eckenwalder and Barrett (1986) is retained.

Character 8. Stipules or ligules. States: 0 = "present"; 1 = "absent." Revised codings: H. dubia (0); H. seubertiana (0); H. gardneri (0). In Pontederiaceae, the structures variously described as stipules or ligules are outgrowths of the lower leaf zone above the insertion of the upper leaf zone. Confusion over whether to call these structures stipules or ligules appears to have arisen from an historical precedent set by de Candolle that monocotyledons do not have stipules. In Richard's view, "stipule" is a broader term that can include the sort of structures called "ligules" in grasses (the tongue-like outgrowths between sheath and blade): although less elaborate, grass ligules are
developmentally homologous with the stipule in Pontederiaceae. *Heteranthera dubia* and *H. seubertiana* possess ligule-like structures, and the leaf whorl members of *H. gardneri* possess stipules.

**Note:** Characters 11-15 (leaf blade characters) refer here to the blades of adult, petiolate leaves, which in Pontederiaceae includes the lower (and sometimes upper) of the (one-) two "spathes" or bracts subtending each inflorescence. In taxa where petioles are never produced (as in *H. dubia* and *H. seubertiana* and *P. lanuginosum* of Philydraceae) these characters refer to the blades of mature, sessile leaves. Since the filiform leaf of *H. gardneri* has no structures that are obviously homologous to a petiole or leaf-blade, it is not clear whether the characters involving these structures are applicable. These characters (9 and 11-15) are therefore treated as missing data ("?"") for this taxon.

**Character 9. Petiole.** States: 0 = "normal"; 1 = "inflated"; 2 = "absent." Note: 2 is a new character state. Revised coding: *H. dubia* (2); *H. seubertiana* (2); *H. gardneri* (?); *P. lanuginosum* (2).

**Character 11. Width/length ratio of lamina** ("Leaf width/ length"). States: 0 = "0.3-0.5"; 1 = "0.1-0.3"; 2 = "< 0.1"; 1' = "> 0.5." Revised coding: *H. dubia* (2); *H. gardneri* (?); *M. vaginalis* (0); *P. lanuginosum* (2).

**Character 12. Maximum length of lamina** ("Maximum leaf size"). States: 0 = "5-10 cm"; 1 = "> 10 cm"; 1' = "< 5 cm." Note: These are modifications of the ranges given in Eckenwalder and Barrett (1986), where state 0 was "4-10 cm" and state 1' was "< 4 cm." Revised codings: *H. dubia* (0); *H. limosa* (1); *H. oblongifolia* (0); *H. seubertiana* (1); *H. zostericlydia* (1); *H. gardneri* (?); *M. cyannea* (1); *M. hastata* (1); *M. korsakovii* (1); *M. vaginalis* (1); *P. lanuginosum* (1).
APPENDIX C

Character 13. **Lamina base shape** ("Leaf base shape"). States: 0 = "cuneate"; 1 = "cordate"; 2 = "sagittate." Revised coding: *H. gardneri* (?); *M. cyanea* (0)b.

Character 14. **Brodest point of blade.** States: 0 = "below middle"; 1 = "at middle"; 2 = "above middle." Revised coding: *H. gardneri* (?).

Character 15. **Lamina apex shape** ("Leaf apex shape"). States: 0 = "acute"; 1 = "obtuse." Revised codings: *H. limosa* (0)e; *H. oblongifolia* (1)e; *H. gardneri* (?).

Character 16. **Inflorescence type.** States: 0 = "panicle"; 1 = "derived panicle"; 2 = "spike - raceme"; 3 = "umbel - sub-umbel"; 4 = "pseudanthium." Note: 3 and 4 are new character states. Character state 2 refers to a class of panicle that has highly contracted side-branches and is presumably derived from a regular panicle. Revised codings: *H. gardneri* (4); *M. hastata* (2 and 3); *M. korsakovii* (0); *M. vaginalis* (2 and 3). The "flower" in *H. gardneri* is a two-flowered pseudanthiumd. *M. korsakovii* has a paniculate inflorescenceb. The inflorescences of *M. hastata* and *M. vaginalis* are sub-umbellate to shortly racemoseb and are therefore coded as polymorphic for these two forms.

Character 20. **Peduncle pubescence.** States: 0 = "glabrous"; 1 = "hairy." Note: Taxa are coded as "hairy" if any part of the inflorescence axis is pubescent. Revised codings: *E. azurea* (1)g; *E. crassipes* (1)g; *E. diversifolia* (1)b; *E. heterosperma* (1)i; *P. cordata* var. *cordata* (1)g; *P. cordata* var. *lancifolia* (1)i; *P. sagittata* (1)k; *P. rotundifolia* (1)g. The inflorescence axis of *E. diversifolia* is sparsely pubescentb.

Character 21. **Flower attachment.** States: 0 = "pedicellate"; 1 "sessile." Revised coding: *M. cyanea* (0)b. Note: The "pedicellate" coding is employed here if at least some of the flowers in the inflorescence have pedicels.
Character 22. Flower number. States: 0 = "50-100"; 1 = "10-50"; 2 = "2-10"; 3 = "1"; 1' = "> 100." Revised codings: *P. lanuginosum* (1)\(^b\); *H. zostericola* (2)\(^c\). Note: The flower count for *P. lanuginosum* refers to the entire panicle, not each spike.

Character 23. Flower symmetry. States: 0 = "actinomorphic"; 1 = "zygomorphic." Revised codings: *H. gardneri* (1)\(^d\); *H. limosa* (1)\(^a\); *H. oblongifolia* (1)\(^a\); *H. zosterifolia* (1)\(^a\).

Character 24. Perianth length ("Flower size"). State: 0 = "2-4.4 cm"; 1 = "> 4.5 cm"; 1' = "< 2 cm." Revised codings: *P. lanuginosum* (1)\(^h\); *M. cyanea* (1)\(^b\); *M. hastata* (1)\(^b\); *M. korsakovii* (1)\(^b\); *M. vaginalis* (1)\(^b\).

Character 26. % Perianth fusion. State: 0 = "0"; 1 = "10-25"; 2 = "25-40"; 3 = "40-60"; 4 = "60-70"; 5 = "70-80." Revised coding: *H. gardneri* (2)\(^h\).

Character 33. Stamen diversity. State: 0 = "monomorphic stamens"; 1 = "dimorphic stamens." Revised codings: *H. gardneri* (?); *P. lanuginosum* (?). Note: *H. gardneri* has a single fertile stamen and two staminodes\(^d\). *P. lanuginosum* has only a single stamen. This character is therefore inapplicable for the latter taxon and may be inapplicable for the former, so they are coded as missing data ("?").

Character 34. Filament inflation. State: 0 = "none"; 1 = "some inflated." Revised coding: *H. zostericola* (1)\(^c\).

Character 36. Anther attachment. State: 0 = "dorsifixed"; 1 = "basifixed." Revised coding: *P. lanuginosum* (0)\(^l\).
Character 38. **Fertile carpel number** ("Ovary locule number"). State: 0 = "3"; 1 = "2"; 2 = "1."
Revised codings: *P. lanuginosum* (0)^m; *H. gardneri* (0)^d.

Character 39. **Ovule number per flower** ("Ovule number"). State: 0 = "> 50"; 1 = "2-50"; 2 = "1."
Revised codings: *E. azurea* (0)^a; *E. heterosperma* (0)^n.

Character 42. **Seed length.** State: 0 = "< 1 mm"; 1 = "1-1.5 mm"; 2 = "> 1.5 mm." Revised codings: *M. cyanea* (0)^b; *M. korsakovi* (1)^b.

Character 43. **Anther dehiscence.** Note: this is an additional character. State: 0 = "regular"; 1 = "poricidal." Coding: All species of *Monochoria* (1); *P. lanuginosum* and other taxa in Pontederiaceae (0).
Complete character state listings are given below for two additional taxa (*Eichhornia paradoxa* and *E. meyeri*) and are repeated for an unnamed species of *Eichhornia* identified in Eckenwalder & Barrett (1986) as *E. paradoxa*, and called *Eichhornia* sp. here. Characters and character-states are as listed in Eckenwalder & Barrett (1986), except where modified above. The boundary between petiole and lamina is not sharp in *Eichhornia* sp., but this did not unduly interfere with the assignment of character states to those characters involving the leaf-blade (i.e., characters 11-15). In several instances (marked by an asterisk) the character states for *Eichhornia* sp. have been revised from those listed in Eckenwalder & Barrett (1986). Additions: *E. meyeri*\(^{a,g}\), *E. paradoxa*\(^{b}\); changes: *Eichhornia* sp.\(^{b}\).

### Character

<table>
<thead>
<tr>
<th>Taxon</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Eichhornia paradoxa</em></td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>Eichhornia meyeri</em></td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>Eichhornia</em> sp.</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>20</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Eichhornia paradoxa</em></td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>Eichhornia meyeri</em></td>
<td>1'</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><em>Eichhornia</em> sp.</td>
<td><em>2</em></td>
<td><em>0</em></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td><em>3</em></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Eichhornia paradoxa</em></td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>Eichhornia meyeri</em></td>
<td>1</td>
<td>1</td>
<td>1'</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>Eichhornia</em> sp.</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### APPENDIX C

<table>
<thead>
<tr>
<th>Character</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
<th>41</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Eichhornia paradoxa</em></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>Eichhornia meyeri</em></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>Eichhornia sp.</em></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>42</th>
<th>43</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Eichhornia paradoxa</em></td>
<td>0</td>
</tr>
<tr>
<td><em>Eichhornia meyeri</em></td>
<td>0</td>
</tr>
<tr>
<td><em>Eichhornia sp.</em></td>
<td>0</td>
</tr>
</tbody>
</table>

**Sources:** (a) S. C. H. Barrett (pers. comm.); (b) Cook (1989); (c) Goebel (1913); (d) Rutishauser (1983); (e) Horn (1985); (f) J. E. Eckenwalder (pers. comm.); (g) Schulz (1942); (h) S. W. Graham (pers. obs.); (i) Velasquez, J. (1972); (j) Castellanos (1959); (k) Presl (in Castellanos, 1959); (l) Dahlgren & Clifford. (1982); (m) Dahlgren, Clifford and Yeo (1985); (n) Barrett (1988); (o) Lowden. (1973); (p) Richards (1980).
Appendix D. Taxon partitions supported in 50% or more bootstrap replicates from the unrooted analyses of the combined chloroplast data, illustrated with the strict consensus tree from Fig. 3.3. Labels used for these internal partitions are arbitrary.

Other partitions seen in strict consensus trees (Fig. 3.1 - 3.5). Following Penny and Hendy (1985) the smaller of the two subsets of taxa that could be used to describe each partition is reported.

p = (Heteranthera limosa, H. rotundifolia, H. oblongifolia, H. zosterifolia, H. seubertiana, H. dubia)
u = (Pontederia rotundifolia, P. sagittata, P. cordata v. ovalis)
v = (Hydrothrix gardneri, Heteranthera oblongifolia, H. zosterifolia, H. seubertiana, H. dubia)
w = (Pontederia sagittata, P. cordata v. cordata, P. cordata v. lancifolia, P. cordata v. ovalis)
x = (Hydrothrix gardneri, Heteranthera limosa, H. oblongifolia, H. zosterifolia, H. seubertiana, H. dubia)
y = (Heteranthera limosa, H. oblongifolia, H. zosterifolia, H. seubertiana, H. dubia)
z = (Heteranthera zosterifolia, H. seubertiana, H. dubia)
aa = (Heteranthera rotundifolia, H. oblongifolia)
ab = (Hydrothrix gardneri, Heteranthera zosterifolia, H. seubertiana, H. dubia)
### APPENDIX E. Accession details for species analysed in Chapter 4, arranged alphabetically by family.

<table>
<thead>
<tr>
<th>SPECIES</th>
<th>FAMILY</th>
<th>GENE</th>
<th>VOUCHER/SOURCE</th>
<th>CITATION</th>
<th>GENBANK ACCESSION</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Caryota mitis</em> Lour.</td>
<td>Arecales</td>
<td>ndhf</td>
<td>Graham 1020</td>
<td>This paper</td>
<td>U79227</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rbcl</td>
<td>None</td>
<td>Gaut et al. 1992</td>
<td>M81811</td>
</tr>
<tr>
<td><em>Narcissus elegans</em> (Haw.) Spach</td>
<td>Amaryllidaceae</td>
<td>ndhf</td>
<td>Barrett 1434</td>
<td>Graham &amp; Barrett MS</td>
<td>U79216</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rbcl</td>
<td>Unknown</td>
<td>Fay et al 1995</td>
<td>n/a</td>
</tr>
<tr>
<td><em>Ananas comosus</em> (L.) Merrill</td>
<td>Bromeliaceae</td>
<td>ndhf</td>
<td>Graham 1000</td>
<td>This paper</td>
<td>U79225</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rbcl</td>
<td>Unknown</td>
<td>Duvall et al. 1993</td>
<td>L19977</td>
</tr>
<tr>
<td><em>Canna sp.</em></td>
<td>Cannaceae</td>
<td>ndhf</td>
<td>Graham 1010</td>
<td>This paper</td>
<td></td>
</tr>
<tr>
<td><em>Canna indica</em> L.</td>
<td>Cannaceae</td>
<td>rbcl</td>
<td>R.K. Godfrey 57928</td>
<td>Prince et al. 1995</td>
<td>n/a</td>
</tr>
<tr>
<td><em>Tradescantia zebrina</em> Bosse</td>
<td>Commelinaceae</td>
<td>ndhf</td>
<td>Graham 1030</td>
<td>This paper</td>
<td>U79229</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rbcl</td>
<td>J. Cortesi 7324, GMUF</td>
<td>Duvall et al. 1993</td>
<td>L05042</td>
</tr>
<tr>
<td><em>Cyanastrum cordifolium</em> Oliv.</td>
<td>Cyanastraceae</td>
<td>ndhf</td>
<td>Graham &amp; Barrett 2, TRT</td>
<td>This paper</td>
<td>U79228</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rbcl</td>
<td>Graham &amp; Barrett 2, TRT</td>
<td>Chase et al. 1993</td>
<td>U41572</td>
</tr>
</tbody>
</table>

1 Details for taxa of Pontederiaceae considered in Chapter 4 are provided in Appendix B (and see Appendix A, Chapter 3). Superorder: A = Arecales, B = Bromeliales, C = Commeliniales, L = Lituanias, Z = Zingiberanae (all *sensu* Dahlgren, Clifford and Yeo 1985).
Source/Locality

The source is the voucher (or living collection if marked by an asterisk) from A. Femendas and J. Joaquim. The source is the voucher (or living collection if marked by an asterisk) from A. Femendas and J. Joaquim. One sub-section of section Narcissus, shown in the legend is the total of section Juncifolia DC., into sub-section Juncifolia A. Species, the other sub-sections in Narcissus, section Amaryllis (Amaryllis, Baker, Section, Petalum, and Tassil DC, belong to sub-genus Hernione (Haw, and related taxa used in the survey of sequence variation in the chloroplast gene.!
(Webb & Heider) Colmeno
N. hederaeana

N. canadensis DC.

N. bulbocodium L.

Section BULBOCODIUM DC.

Collection (2) (Narrow-leaved variety)

*Blanchard, ex Michael Salmon stock
Blanchard 9130, Western Atlas, Morocco

Collection (1) (Broad-leaved variety)

N. pronssionii Lagesse

Section Aurelia (L. Cary) Baker

TAXON

SOURCE

APPENDIX E - SOURCE LIST FOR NARCISSUS
*Blanchard, ex Rosswame

N. multiflorus Salisb.

(Pueblay) A. Fermades

N. pectoris L. var. helleucus

Section NARCISSUS L.

N. vittiliorum Schw.: Schouboe

N. jonquilla L.

Section JONQUILLAE DC. (compl.)
N. pseudonarcissus L.

N. longiflora Pusley

var. pulchra Bernhauer Casas

N. hispanica Cano

N. austriaca (Jordan) Pusley

N. bicolor L.

Section Pseudonarcissi DC.

Bauern 1432, Sierra de Cazorla, Jaen, Spain

Bauern 1431, Sierra de Cazorla, Cabra, Cordoba, Spain

Branchard 940, Torre Serra da Estrela, Casela Bravo, Portugal

Branchard 9211, Spanish Pyrenees

Bauern 9401, Auvergne, France
APPENDIX F - SOURCE LIST FOR *NARCISSUS*

## TAXON

<table>
<thead>
<tr>
<th>SOURCE/LOCALITY</th>
</tr>
</thead>
</table>

### Section SEROTINI Parlatore

*N. serotinus* L.

| Collection (1) | Barrett 1433A (ex J. Arroyo collection), La Puebla del Río, Sevilla, Spain |
| Collection (2) | Barrett 1433B (ex J. Arroyo collection), La Pueblo del Río, Sevilla, Spain |

### Section TAPEINANTHUS (Herbert) Traub

*N. cavanillesii* A. Barra & G. López

| Collection (1) | *Barrett* (ex J. Arroyo collection), Venta del Cruce, Sevilla, Spain |
| Collection (2) | Blanchard 9131, Western High Atlas, Morocco |
## APPENDIX F - SOURCE LIST FOR NARCISSUS

<table>
<thead>
<tr>
<th>TAXON</th>
<th>SOURCE/LOCALITY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Section TAZETTAEE DC.</strong></td>
<td></td>
</tr>
<tr>
<td><em>N. dubius</em> Gouan</td>
<td>Barrett 1433, Montagne d'Hortus, Valflaunes, Montpellier, France</td>
</tr>
<tr>
<td><em>N. elegans</em> (Haw.) Spach</td>
<td>Barrett 1434 (ex J. Arroyo collection), Brieux, Morocco</td>
</tr>
<tr>
<td><em>N. papyraceus</em> Ker-Gawler</td>
<td>Barrett 1435, Aznalcazar, Sevilla, Spain</td>
</tr>
<tr>
<td><em>N. tazetta</em> L.</td>
<td>Barrett 1436, Pérols, Montpellier, France</td>
</tr>
<tr>
<td><em>N. tortifolius</em> Fernández Casas</td>
<td>*Blanchard, Almeria, S.E. Spain</td>
</tr>
<tr>
<td>Taxon</td>
<td>Source/Locality</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td><em>Leucocinum aestivum</em> L.</td>
<td><em>Braeurol 1438</em>, Commercial Source</td>
</tr>
<tr>
<td><em>Cochliopus nigricans</em> L.</td>
<td><em>Braeurol 1437</em>, Commercial Source</td>
</tr>
<tr>
<td><em>Stemergeria linear (L.) Ker-Gawler</em></td>
<td><em>Braeurol 1439</em>, Commercial Source</td>
</tr>
</tbody>
</table>
Appendix G. Karyotypic formulae (where known) of the various species of *Narcissus* examined in the current study (taken from Fernandes 1967, 1968b, Fernandes and Fernandes 1945). The terminology is that of Fernandes (1931a,b, 1934). Each letter refers to one arm of the chromosome, with the sizes (so far as I can make out) arranged as follows, in descending order: L, l, m, i, P, p. Telocentric to short-acrocentric ("céphalobrachiaux") chromosomes are denoted with a period (e.g., "P."). Arms with satellites (nucleolar organizers) are indicated with apostrophes.

---

**Sub-genus Hermione (x = 5)**

**Sect. Aurelia** *(N. broussonetii)*

\[ x = 11 = 1 \text{ LL} + 1 \text{ Lp} + 2 \text{ L.} + 1 \text{ l.} + 1 \text{ P.'} + 4 \text{ P.} + 1 \text{ pp} \]

**Sect. Serotini** *(N. serotinus)*

\[ x = 5 = 1 \text{ LL} + 1 \text{ 'Ll} + 1 \text{ LP} + 1 \text{ l.} + 1 \text{ li} \]

\[ x = 15 = 2 \text{ LL} + 2 \text{ LP} + 3 \text{ Lp} + 1 \text{ l.} + 2 \text{ li} + \text{ P.'} + 4 \text{ P.} \]

**Sect. Tazettae** *(N. elegans)*

\[ x = 10 = 2 \text{ Lp} + 2 \text{ L.} + 1 \text{ li} + 1 \text{ lp} + 1 \text{ P.'} + 3 \text{ P.} \]

\[ x = 10 = 3 \text{ Lp} + 2 \text{ L.} + 1 \text{ P.'} + 3 \text{ P.} + 1 \text{ pp} \]

(N. *papyraceus s.l.)*

\[ x = 11 = 1 \text{ Lp} + 2 \text{ L.} + 1 \text{ li} + 1 \text{ l.} + 2 \text{ P.'} + 3 \text{ P.} + 1 \text{ pp} \]

(N. *tazetta s.l.)*

\[ x = 10 = 2 \text{ Lp} + 2 \text{ L.} + 1 \text{ li} + 1 \text{ lp} + 1 \text{ P.'} + 3 \text{ P.} \]

\[ x = 11 = 1 \text{ Lp} + 2 \text{ L.} + 1 \text{ li} + 1 \text{ l.} + 2 \text{ P.'} + 3 \text{ P.} + 1 \text{ pp} \]
Sub-genus *Narcissus* \((x = 7)\)

**Sect. Apodanthae**  
\((N.\ \text{calcicola}, \ N.\ \text{rupicola}, \ N.\ \text{scaberulus})\)

\[ x = 7 = 1 \text{ LL} + 2 \text{ Ll} + 1 \text{ Lp} + 1 \text{ li} + 1 \text{ lp} + 1 \text{ Pp}' \]

**Sect. Bulbocodium**  
\((N.\ \text{bulbocodium})\)

\[ x = 7 = 3 \text{ Lp} + 1 \text{ li} + 2 \text{ PP} + 1 \text{ Pp}' \]

\((N.\ \text{cantabricus})\)

\[ x = 7 = 2 \text{ Lp} + 1 \text{ li} + 1 \text{ lp}' + 3 \text{ PP} \]

**Sect. Ganymedes**  
\((N.\ \text{triandrus})\)

\[ x = 7 = 3 \text{ Lp} + 1 \text{ li} + 2 \text{ PP} + 1 \text{ Pp}' \]

**Sect. Jonquilla**

**Sub-sect. Jonquillae**

\((N.\ \text{fernandesii}, \ N.\ \text{jonquilla}, \ N.\ \text{viridiflorus})\)

\[ x = 7 = 2 \text{ Ll} + 1 \text{ Lm} + 1 \text{ Lp} + 1 \text{ li} + 1 \text{ lp} + 1 \text{ lp}' \]

**Sub-sect. Juncifoliae**

\((N.\ \text{assoanus})\)

\[ x = 7 = 2 \text{ Ll} + 2 \text{ Lp} + 1 \text{ li} + 1 \text{ Pp} + 1 \text{ Pp}' \]

\((N.\ \text{gaditanus})\)

\[ x = 7 = 2 \text{ Ll} + 1 \text{ Lp} + 1 \text{ lp}' + 1 \text{ li} + 1 \text{ Pp} + 1 \text{ P} \].
Sub-genus *Narcissus*  \((x = 7)\) (contd.)

**Sect. Narcissus**  
\((N. poeticus, N. radiiflorus)\)
  
  \(x = 7 = 1 L_l + 1 L_P + 1 L_p + 2 l_i + 1 l_p + 1 P_p'\)

**Sect. Pseudonarcissus**  
\((N. asturiensis, N. bicolor, N. hispanicus, N. longispatus)\)
  
  \(x = 7 = 1 L_l + 1 L_P + 1 L_p + 2 l_i + 1 l_p + 1 P_p'\)

**Sect. Tapeinanthus**  
\((N. cavenilesii)\)
  
  \(x = 14 = 2 L_m + 2 'L_p + 2 L_P + 2 L_p' + 2 l_i + 2 P_p + 2 P.\)
Appendix H. Experimental details for the sequencing studies.

DNA isolation. -- Total genomic DNAs were isolated using the method of Doyle and Doyle (1987) from approximately 0.1 - 2 g of fresh green tissue (whole or partial leaves, flowering stems (Narcissus), or entire stems (Hydrothrix)), typically from single individuals (ramets). Two extracts were made for each collection, which were from the same or different ramets. Tissue was frequently stored at -80 °C prior to extraction. Yields were highly variable but typically in the range 10 to 200 μg. For Narcissus, as little as 0.1 g of one to four day old tissue provided good yields. DNAs were resuspended in 0.1 - 1 ml of TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) and stored at -20 °C.

Polymerase Chain Reaction (PCR) amplification and DNA sequencing. -- Primers used for sequencing and amplifying rbcL were designed by Zurawski, Clegg and Brown (1984). Gerard Zurawski (DNAX institute, Palo Alto, CA, USA) kindly provided aliquots of these primers for use as sequencing primers. For rbcL, oligonucleotides Z1F and Z1375R were employed to amplify 1375 bp of the coding region of this gene. The coding sequence of rbcL is 1,431 bp long in Zea mays. Clegg (1993) noted that minor length differences are known in the extreme 3'-end of the gene, but none were observed in Pontederiaceae. For Hydrothrix gardneri, Z1375R did not permit amplification and so primer Z1204R was used as the 3'-end amplification primer. An 1343 bp internal region at the 5'-end of rbcL was sequenced for all taxa, except for Hydrothrix, for which an 1169 bp region was sequenced.

The coding sequence of ndhF is 2,223 bp long in tobacco (Nicotiana tabacum). Amplification and sequencing primers for an internal 3'-region of ndhF were designed using an alignment of published sequences from rice, tobacco and broad-bean. These oligonucleotides are listed in Table 3.1. Amplification was achieved using primer pair ndh2F and ndh1R or ndh1.6R. A 490 bp segment of these amplification products was sequenced using primers...
APPENDIX H - EXPERIMENTAL DETAILS

ndh3F and ndh4F (for the forward strand) and ndh2R and ndh4R (for the reverse strand). I focused on this region of \textit{ndhF} because I noticed its high substitution rate relative to \textit{rbcL} and the rest of the \textit{ndhF} locus, in comparisons of previously published rice, bean and tobacco sequences for these genes. This finding has been commented and elaborated upon by other workers (Olmstead and Sweere 1994; Kim and Jansen 1995).

Each double-stranded PCR product was used as a template for generating forward and reverse single-stranded DNA products, in separate asymmetric PCR reactions using only one of the original primers. These single-stranded DNAs were then used for sequencing with internally-situated forward or reverse sequencing primers. In general, only those primary PCR reactions that yielded a single amplification product went on to provide useful sequence. Secondary asymmetric PCR reactions often yielded a complex mixture of fragment sizes (sometimes a smear with no single identifiable most-common fragment size), but this bore no relation to the final quality of the sequence. Cross-contamination of PCR products was minimized by including a negative-control PCR reaction (i.e., with all components but template DNA) for primary amplifications, and through the use of dedicated pipetmen for genomic DNA extraction and PCR set-up, dedicated laboratory bench space, and pre-sterilized tips, tubes and reaction components. Since all taxa were sequenced twice (once from each DNA extract), this enabled rapid detection of cross-contamination or sample-switching, of which several instances were noted.

All PCR reactions were performed in a 100 \( \mu \)l total reaction volume using standard PCR reaction components: 1x PCR buffer (10 mM Tris-HCl, 50mM KCl, pH 8.3); 1.5 mM Mg; 0.2mM dGTP, dATP, dTTP and dCTP; and 0.5 \( \mu \)M of each primer. 1 \( \mu \)l of genomic DNA extract was used as the template DNA for the primary amplifications, and 5 \( \mu \)l of double-stranded PCR product for the secondary amplifications. 2.5 U of Taq DNA Polymerase (various manufacturers) was added after denaturing the template DNA at 94 \(^\circ\)C for 120 seconds.
Each polymerase chain reaction consisted of 30 cycles of template DNA denaturation, primer annealing and polymerase-mediated extension of new strands. Denaturation was performed at 94 °C for 90 seconds, annealing at 42-48 °C for 120 seconds, and extension at 72 °C for 180 seconds. A final 15 minute extension at 72 °C was performed after the completed chain reaction. For rbcL, the annealing step was performed at 48 °C for the double-stranded reaction and 42 °C for the single-stranded reaction. For ndhF, annealing was performed at 47 °C for both series of reactions. Single-stranded DNAs were purified by precipitation with 60 μl of 20% PEG 8000, 2.5 M NaCl after a minimum two hours incubation at 4 °C. Pellets were obtained by centrifugation at 13K rpm for 10-15 minutes. After two washes with 70% ethanol, pellets were air-dried overnight, resuspended in 20 μl of TE, and stored at -20 °C.

One third of the purified single-stranded DNA from a 95 μl reaction volume was employed in each sequencing reaction using the dideoxy chain-termination method (Sanger, Niklen and Coulson 1977), with the modified T7 DNA polymerase provided with the Sequenase version 2.0 DNA sequencing kit (United States Biochemical Corporation, product number 70770/70775). Reaction conditions were those indicated in the Sequenase manual (with the suggested addition of inorganic pyrophosphatase), except that DTT was added to the annealing reaction when this had cooled to 34 °C (rather than room temperature), labelling and termination reactions were permitted to run for 8 minutes (to allow 8 reactions to be run in parallel), and the termination reactions were performed between 42 and 45 °C.

Sequencing plates were pretreated to bind the gel to the short plate. Repulsive silane was applied to the long plate (two polishings with 2 ml of Sigmacote, product number SL-2, Sigma Chemical Company), and attractive silane was applied to the short plate (2.5 ml of 0.003% ethanoic acid, 0.003% γ-methacyrloxypropyltrimethoxysilane in ethanol, followed by removal of excess silane with a final polishing with ethanol). ³⁵S-radiolabeled sequencing ladders were fractionated on 6% polyacrylamide sequencing gels. Long runs and short runs (about 4 hours and 1.5 hours, respectively) were performed on separate gels at around 1600V,
APPENDIX H - EXPERIMENTAL DETAILS

60-80 W, 45 - 50 °C. This saved time with the short runs. After fixing the gel in 10% methanol, 10% ethanoic acid, and washing in running distilled water for 5 minutes, it was dried horizontally for 30 - 60 minutes using a hairdryer in a fumehood. Labelled bands were visualized by overnight exposure to autoradiography film. Longer exposures were sometimes used when the signal was weak. Bound gels were soaked in water prior to being removed from the plate with a fresh razor blade.

For each taxon, one strands-worth of sequence was obtained from each extract. Typically both forward and reverse strands were sequenced (i.e., one strand from each extract), but for rbcL duplicate sequences were sometimes obtained from same-sense strands. Duplicate sequences for each gene and taxon were thus always derived from independent extracts (and hence potentially from different individuals). Several potentially polymorphic sites were observed in ndhF and rbcL sequences from Pontederiaceae and Narcissus. Because they were not re-confirmed by sequencing alternative strands from each duplicate extract, they were coded as missing characters for the relevant taxa in analyses with PAUP and PHYLIP. The putatively polymorphic characters all involved an autapomorphic and an invariant state, and so they would in any case have no influence on phylogenetic inference using parsimony.


LITERATURE CITED

Lovett Doust, eds.). Oxford University Press, New York, USA.


LITERATURE CITED


LITERATURE CITED


LITERATURE CITED


LITERATURE CITED


Darwin, C. 1877. The different forms of flowers on plants of the same species. Murray, London, UK.


LITERATURE CITED


Fernandes, A. 1931b. Estudos nos cromosomas das Liliáceas e Amarilidáceas. Boletim Sociedada Broteriana (sér. 2) 7:3-110.


LITERATURE CITED


LITERATURE CITED


LITERATURE CITED


LITERATURE CITED


LITERATURE CITED


LITERATURE CITED


Ridley, H. N. 1930. The dispersal of plants throughout the world. Reeve, Ashford, UK.


LITERATURE CITED


LITERATURE CITED


LITERATURE CITED


LITERATURE CITED

Cambridge, UK.


Swofford, D. L. 1993. PAUP: Phylogenetic analysis using parsimony, version 3.1.1

Computer program and documentation distributed by the Illinois Natural History Survey, Champaign, Illinois, USA.


Traub, H. P. 1963. Genera of the Amaryllidaceae. American Plant Life Society, La Jolla, California, USA.


LITERATURE CITED


Wylie, R. B. 1917. Cleistogamy in *Heteranthera dubia*. Bulletin from the Laboratories of
Natural History. University of Iowa 7:48-58.


divergence between barley and maize chloroplast DNA. Genetics 106:735-749.