

Summer 8-15-2016

Systematics and Evolution of the Arundinoideae and Micrairoideae (Poaceae)

Jordan Kinsley Teisher

Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Recommended Citation

Teisher, Jordan Kinsley, "Systematics and Evolution of the Arundinoideae and Micrairoideae (Poaceae)" (2016). *Arts & Sciences Electronic Theses and Dissertations*. 900.

https://openscholarship.wustl.edu/art_sci_etds/900

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Evolution, Ecology and Population Biology

Dissertation Examination Committee:

Barbara Schaal, Chair
Elizabeth Kellogg, Co-Chair
Garland Allen
Gerrit Davidse
Allan Larson
Peter Raven

Systematics and Evolution of the Arundinoideae and Micriarioideae (Poaceae)

by

Jordan K. Teisher

A dissertation presented to the
Graduate School of Arts & Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2016
St. Louis, Missouri

© 2016, Jordan K. Teisher

Table of Contents

List of Figures.....	vi
List of Tables.....	viii
Acknowledgments.....	ix
Abstract of the Dissertation.....	xiv
Introduction of the Dissertation.....	1
Study System: Arundinoideae and Micrairoideae.....	7
Outline of the Dissertation.....	16
Literature Cited.....	18
Chapter 1: "Awn" or Off – Evolution of Dispersal and Burial Traits in the PACMAD Grasses.....	26
1.1 Introduction.....	27
1.2 Materials and Methods.....	31
1.2.1 Taxon Sampling.....	31
1.2.2 DNA Isolation and Sequencing.....	32
1.2.3 Plastome Assembly and Phylogenetics.....	32
1.2.4 Morphological Character Coding.....	34
1.2.5 Ancestral State and Diversification Estimates.....	35
1.3 Results.....	37

1.3.1	Plastome Assembly and Alignment.....	37
1.3.2	Phylogenetic Analysis.....	38
1.3.3	Trait Evolution.....	39
1.3.4	Tree Dating and Diversificaiton Analysis.....	40
1.4	Discussion.....	41
1.4.1	Evolution of Dispersal/Burial Traits.....	42
1.4.2	Phylogenetic Position of Aristidoideae.....	44
1.4.3	Ancestral PACMAD Habitat.....	44
1.4.4	Divergence Date Estimates.....	45
1.4.5	Implications for Classification.....	47
1.4.6	Conclusion.....	49
	Literature Cited.....	51
	Figures and Tables.....	57
Chapter 2:	Phylogenetic Analysis of Polyploidy in Temperate Arundinoideae.....	75
2.1	Introduction.....	76
2.2	Materials and Methods.....	81
2.2.1	Taxon Sampling.....	81
2.2.2	RNA Extraction and Sequencing.....	81
2.2.3	Transcriptome Assembly.....	82
2.2.4	K _s Plots.....	83

2.2.5 Gene Clustering, Alignment, and Phylogenetic Analysis.....	84
2.3 Results.....	88
2.3.1 Transcriptome Assembly and Orthogroup Analysis.....	88
2.3.2 K _s Plots and PUG Analyses.....	89
2.4 Discussion.....	93
2.4.1 Evolution of Reedy Arundinoideae.....	94
2.4.2 PACMAD Genome Evolution.....	97
2.4.3 Approaches to Identifying WGDs.....	99
2.4.4 Conclusion.....	100
Literature Cited.....	101
Figures and Tables.....	107
Chapter 3: Evolution of C ₄ Photosynthesis in the Micrairoideae (Poaceae).....	117
3.1 Introduction.....	118
3.2 Materials and Methods.....	123
3.2.1 Collection of Material.....	123
3.2.2 DNA Isolation and Sequencing.....	123
3.2.3 Plastome Assembly.....	124
3.2.4 Alignment and Phylogenetic Analysis.....	125
3.2.5 Bioclimatic Data.....	125
3.2.6 Carbon Isotope Discrimination.....	126

3.2.7 Trait Evolution.....	127
3.3 Results.....	128
3.3.1 Plastome Assembly, Alignment and Phylogenetic Analysis...	128
3.3.2 Habitat Breadth in Micrairoideae.....	129
3.3.3 Carbon Isotope Ratios.....	129
3.3.4 Habitat Evolution in Eriachneae.....	130
3.4 Discussion.....	131
3.4.1 C ₄ Photosynthesis and Habitat Breadth in Micrairoideae.....	131
3.4.2 Evolution of Habitat Preference in Eriachneae.....	133
3.4.3 Notes on Classification.....	135
3.4.4 Conclusion.....	138
Literature Cited.....	139
Figures and Tables.....	144
Conclusion of the Dissertation.....	164
Literature Cited.....	171

List of Figures

Figure 1.1: SEM of <i>Alloeochaete gracillima</i> diaspore.....	58
Figure 1.2: ML phylogeny of grass plastomes.....	59
Figure 1.3: Parsimony ancestral state estimate of lemma awns.....	60
Figure 1.4: Parsimony ancestral state estimate of lemma callus.....	61
Figure 1.5: Parsimony ancestral state estimate of lemma hairs.....	62
Figure 1.6: BEAST dated phylogeny.....	63
Figure 1.7: Nine best credible shift sets from BAMM.....	64
Figure 1.8: ML phylogeny of PACMAD with branch lengths proportional to diversification rate shifts in posterior probability distribution.....	65
Figure 1.9: Best diversification rate shift set from BAMM.....	66
Figure 1.10: Cross-section of <i>Nematopoa longipes</i>	67
Figure 2.1: Hypothetical PUG example.....	108
Figure 2.2: Alternative topologies for Arundinoideae used in PUG.....	109
Figure 2.3: Distribution of contigs from arundinoid species in orthogroups.....	110
Figure 2.4: K_s plots for four temperate Arundinoideae.....	111
Figure 2.5: PUG results for chloroplast topology.....	112
Figure 2.6: PUG results for alternative topologies with <i>Arundo</i> as ingroup.....	113
Figure 2.7: PUG results for alternative topologies between <i>Phragmites</i> and relatives.....	114
Figure 3.1: Map of Australian collecting sites.....	145
Figure 3.2: ML phylogeny of Micrairoideae.....	146
Figure 3.3: Map of cleaned GBIF localities used in analysis	147
Figure 3.4: Principal components analysis of BioClim variables for GBIF localities	148
Figure 3.5: Disparity analysis of BioClim variables using <i>betadisper</i>	149

Figure 3.6: Carbon isotope ratios measured from dried leaf material.....	150
Figure 3.7: Subset of total Eriachneae climate space sampled in phylogeny.....	151
Figure 3.8: Mean annual temperature mapped on phylogeny of Micrairoideae.....	152
Figure 3.9: Mean annual precipitation mapped on phylogeny of Micrairoideae.....	153
Figure 3.10: Mean annual temperature mapped on phylogeny of Eriachneae.....	154
Figure 3.11: Mean annual precipitation mapped on phylogeny of Eriachneae.....	155
Figure 3.12: ML topology of Eriachneae colored according to Lazarides (1995) groups.....	156

List of Tables

Table 1.1: Alternative alignments used in plastome phylogeny.....	68
Table 1.2: Ages of select clades from BEAST analysis.....	69
Table S1.1: Samples used in phylogenetic analysis.....	70
Table S1.2: Species sampling fractions used in BAMM analysis.....	73
Table S1.3: Likelihood values for alternative awn ancestral trait estimations.....	74
Table 2.1: Sources for coding DNA sequences used in PUG.....	115
Table 2.2: Translation statistics for four temperate Arundinoideae transcriptomes.....	116
Table 3.1: Samples used in phylogenetic analysis.....	157
Table 3.2: Carbon isotope values of leaf material from Micrairoideae.....	159
Table 3.3: Variable loadings of BioClim variables on first two PCA axes.....	161
Table 3.4: Parameters from fitting BioClim variables under models of evolution.....	162

Acknowledgments

I would first like to express my sincere gratitude to my various funding sources, including the Division of Biology and Biomedical Sciences and the Mary Dell Chilton Distinguished Professor endowment at Washington University in St. Louis, the E. Desmond Lee and Family Endowment at the University of Missouri in St. Louis, the Graduate Assistance in Areas of National Need (GAANN) Program at the Department of Education, the Monsanto Graduate Excellence Fellowship, and the Missouri Botanical Garden Graduate Student Fund. I would also like to thank the Department of Biology in the College of Arts and Sciences at Washington University in St. Louis for the numerous teaching opportunities that supplemented my other funding during many semesters.

A great debt of gratitude is due to my committee members for their patience, encouragement, and guidance. My co-advisers, Barbara Schaal and Toby Kellogg, are two of the most inspiring and brilliant people I have ever met, while never failing to be equally kind and generous. Barbara has a dauntingly busy schedule, but always made time to talk with me about my research, the horrors of graduate school, politics, and anything else that popped into my head. Toby represents the consummate scientist to me. Her willingness to give her valuable time training students, her infectious enthusiasm for research, and her seemingly endless knowledge about grass genetics and evolution make her the best adviser I could imagine having. She is certainly better than I deserve, and I am incredibly thankful for the opportunity to learn from her. Peter Raven served as the chair during my committee meetings and was always encouraging, insightful, and very funny. I am honored to have had such a legend in the botanical community on my committee and have benefited tremendously from his advice. Allan Larson was an

inspiring teacher in my graduate courses, and his thoughtfulness and deep understanding of evolutionary concepts have helped to clarify my thinking over the past 6 years. Gerrit Davidse is an encyclopedia of grass systematic knowledge and was always patient with my endless questions about issues of classification and trait distributions. Gar Allen, a relatively late addition to the committee, has provided an invaluable non-specialist perspective on my dissertation work. He has also been a tremendous friend, mentor, and confidante, for which I am eternally grateful.

Having such a distinguished committee means scheduling meetings could be nightmarish. Fortunately, I had the help of the very capable Kasey Driscoll at Washington University and Donna Rogers at the Missouri Botanical Garden in navigating the impossibly busy schedules of Barbara and Peter, respectively.

The Schaal, Olsen, and Kellogg lab members have provided me with valuable technical expertise, comments on presentations and proposals, and general personal and professional support over the course of my Ph.D. In particular, Genevieve Croft introduced me to lab work in the Schaal Lab, Tonapha Pusadee helped me troubleshoot extractions and PCR of herbarium material, and Linda Small in the Olsen Lab very generously shared reagents as well as her extensive knowledge of all things related to laboratory techniques. Two undergraduates in the Schaal Lab, Gary Huang and Lawrence Lee, enthusiastically helped me with DNA extractions and PCR of my field-collected material. Jennifer Gruhn kept me company in the waning Schaal Lab, so I forgive her for defending her dissertation two weeks before me.

I am deeply indebted to the staff of the Biology Department and the Division of Biology and Biomedical Sciences at Washington University. I had the pleasure of working with three program coordinators: Melissa Torres, Andrea Krussel, and Stacia Burd. Betty Daues, Joshua Short, and Andrew Enders helped with room scheduling, copying, and faxing, and Judy Musick

was an invaluable contact for all things related to the department, not to mention a constantly friendly face on campus. Mike Dyer, Jeanne Cablish, Darlene Branson, and the temporary greenhouse staff at the Jeanette Goldfarb Plant Growth Facility were unfailingly helpful in growing plants for transcriptome extractions and developmental work. Gerry Rohde, Kelly March, Jan McDermott, and Barney Moss in the Biology Stockroom helped me track down equipment and reagents, often without the benefit of a complete name or accurate description from me. Mike Malolepszy was a wizard in dealing with printing and network issues.

The Donald Danforth Plant Science Center became a second home after the Schaal Lab closed, and much of the sequencing work was done there. Margaret Frank and Viktoriya Coneva in the Chitwood Lab showed me how to use the embedding station and microscope camera, and Mark Wilson of the Kellogg Lab and then Bart Lab grudgingly but good-humoredly suffered my computational ignorance.

The Missouri Botanical Garden is one of my favorite places on Earth and was a much-needed refuge when lab work became burdensome. I owe a special thank you to all of the staff there for making me feel so welcome, especially Carla Kostelac and Mary Merello, who were always available for my endless questions. Jim Solomon was of great help in securing loans, getting permissions to sample specimens, and finding cabinet space. Tom Croat was my first rotation adviser, and has been a constant source of encouragement through my dissertation. Roy Gereau helped organize a field expedition with colleagues in Tanzania that resulted in the successful collection of *Alloochaete ulugurensis*, which is included in the phylogeny in Chapter 1 for the first time. Mick Richardson and Peter Hoch were excellent graduate student coordinators for me and the other Garden students. Peter Stevens was always available for a chat and knows more about flowering plants than I will likely ever know about anything.

Several people were of great help in planning and executing my two collecting trips in Australia. Ian Cowie at the Northern Territory Herbarium and Anne O'dea at Kakadu National Park helped me coordinate collections and make sure I had the necessary permissions. Terry McFarlane at the Western Australia Herbarium in Perth arranged for my field equipment in Broome and provided valuable guidance with regards to permits for Western Australia. Eliot Miller was an amazing field assistant on my second trip to Australia, making sure I did not fall down any holes and teaching me about the country's beautiful and diverse native birds.

I received leaf samples for DNA extractions from several collaborators, including Amanda Ingram at Wabash College, J. Travis Columbus at the Rancho Santa Ana Botanic Garden, Nigel Barker at the University of Pretoria, and the late Bryan Simon with the Queensland Government. Maria Vorontsova at the Royal Botanic Gardens – Kew provided me access to specimens that were the only sources of DNA for several species in Arundinoideae. Having a network of researchers willing to share material is vital for studies of global groups like the Arundinoideae and Micrairoideae, and the taxonomic breadth of this study would have been substantially reduced without the contributions of these individuals.

I cannot imagine graduate school without the support of friends and family. My graduate student cohort, the Dream Team of Matt Schuler, Emma Moran, and Katie Zelle, has been better than I could have imagined. Claudia Henriquez, my former office mate and ongoing dear friend, was there for me through all of the highs and lows (mostly lows) of lab work. Kate Waselkov is one of the best botanists and smartest people I know, and I learn something every time I talk to her. My friend and one-time roommate Jordan Terefencko gave me a place to live and hosted some excellent parties. My sister Jen and my brother-in-law Tim were gracious hosts to me in Roach, Missouri and Paris, France, and I have been lucky enough to be able to take time during

my graduate work to spend time with them and their children, my beautiful niece and nephews. My parents have been incredibly supportive my entire life, and I owe more to them than I could hope to fit in the acknowledgments section of a dissertation.

Lastly, I owe a tremendous debt of gratitude for the friendship and guidance of Michael McKain in the Kellogg Lab. I met Michael when my dissertation work was not proceeding well, to put it euphemistically, and it is through his tireless encouragement and skillful mentoring that I have a dissertation at all. Everything I know about high-throughput sequencing, genomic assembly, and bioinformatics has come from him, and our discussions about evolutionary biology have been some of the most engaging and entertaining I have had during my education. Michael's work ethic is inspiring, and his warmth and humor create an atmosphere that made me excited to go to lab every day. For all these reasons and more, I would like to dedicate this dissertation to him.

Jordan Teisher

Washington University in St. Louis

August 2016

ABSTRACT OF THE DISSERTATION

Systematics and Evolution of the Arundinoideae and Micrairoideae (Poaceae)

by

Jordan K. Teisher

Doctor of Philosophy in Biology and Biomedical Sciences

Evolution, Ecology and Population Biology

Washington University in St. Louis, 2016

Professor Barbara Schaal, Chair

Professor Elizabeth Kellogg, Co-Chair

The grass family, Poaceae, is one of the most ecologically and economically important plant groups on Earth. However, the large size (over 11,000 species) and geographic range of the family makes complete resolution of the evolutionary relationships in Poaceae challenging, with some significant groups remaining neglected. Two such understudied clades, subfamilies Arundinoideae and Micrairoideae, possess an incredible amount of morphological and ecological diversity for their sizes, making them a potentially rewarding system in which to study evolution of a wide range of grass features. In this dissertation, I resolved many of the long-standing systematic issues in Arundinoideae and Micrairoideae and used this improved phylogenetic framework to investigate evolutionary issues of broad importance to the grasses.

First, I conducted a molecular phylogenetic analysis of the grass family using high-throughput sequencing of chloroplast genomes, focusing sampling on the taxonomically problematic Arundinoideae. I then used this phylogeny along with observations of herbarium specimens to explore patterns in the evolution of lemma traits across the diverse PACMAD

clade, a group containing roughly half of all grass species. I found that the Arundinoideae are polyphyletic, with several genera belonging in other subfamilies of PACMAD. Possession of a straight awn near the apex of the lemma is found to be the ancestral state in PACMAD, with the evolution of a geniculate awn and loss of lemma awns each evolving multiple times across the clade. Possession of a hairy callus and hairs on the body of the lemma are strongly associated with presence of a lemma awn, supporting the existence of a dichotomous burial syndrome of either smooth and round diaspores or elongate, awned and hairy ones. However, burial syndromes were not associated with changes in diversification rate at this phylogenetic scale.

I explored the origin of the polyploid genomes in Arundinoideae using a phylogenetic analysis of transcriptomic sequence data for four species: *Arundo donax*, *Hakonechloa macra*, *Molinia caerulea*, and *Phragmites australis*. I found strong support for a shared whole genome duplication in the ancestor of the latter three species, with possible support for another such duplication shared by all four. However, limited sampling in Arundinoideae and closely-related subfamilies makes the placement of this second genome duplication equivocal.

Lastly, I tested whether the unique origin of C₄ photosynthesis in tribe Eriachneae of Micrairoideae meets some of the expectations of an adaptive radiation. I used carbon isotopes and plastome phylogenetics of 24 species of Eriachneae that I collected in northern and northwestern Australia to test the phylogenetic limits of the C₄ pathway. Eriachneae were found to all be C₄, with the rest of Micrairoideae using the C₃ pathway. An analysis of bioclimate data in Micrairoideae showed that the shift to C₄ is associated with a transition to hotter and drier climates. However, counter to expectations based on other instances of C₄ evolution in grasses, Eriachneae did not undergo rapid lineage accumulation or habitat diversification following this transition, suggesting that in this case C₄ photosynthesis did not facilitate an adaptive radiation.

INTRODUCTION OF THE DISSERTATION

The grass family, Poaceae, is one of the most diverse and ecologically important plant families, with over 11,000 species occupying virtually all habitats on all continents. Tropical and temperate grasslands make up roughly 40 percent of the earth's land cover (White *et al.*, 2000). Grasses are also arguably the most important plant family to humans, with cereal crops like rice, wheat, and maize contributing over 50 percent of the human race's caloric intake (Awika, 2011). Beer-lovers owe their favorite beverage to another grass, barley (subfamily Pooideae), and anyone with a sweet tooth is indebted to sugar cane, a member of the grass tribe Andropogoneae (subfamily Panicoideae). The desire to reduce humanity's dependency on nonrenewable resources for energy has stimulated interest in biofuel substitutes, and due to their fast growth rates and ability to grow on land that is unusable for agriculture, grasses constitute many of the most promising biofuel candidates (i.e. Porensky *et al.*, 2014; Laurent *et al.*, 2015).

Given the ecological and economic importance of the family, it is unsurprising that an enormous amount of research has been conducted on the genetics, physiology, ecology, and evolution of its members. This collection of literature, combined with the breadth of ecological adaptations in the family, presents almost limitless opportunities to explore fundamental issues in evolution in the grasses, from macroevolutionary trends to population-level dynamics. Vavilov (1922) drew heavily from data on cereal varieties to develop his concept of "homologous series of variation", and Arber (1943) outlined a detailed study of the grasses to understand questions raised by her broader studies in the monocots. More recently, Kellogg (2000) highlighted the utility of the family as a model for understanding the roles that heterotopic and heterochronic gene expression play in the evolution of anatomical and morphological features. Looking at smaller time scales, Glémin & Bataillon (2009) presented the Poaceae as an ideal system for conducting detailed comparative studies of the domestication process. The virtues of the grass

family as a model system are nearly endless, with the added benefit that basic discoveries in this system could enhance production in some of our most valuable crops in the future.

One feature that combines interesting evolutionary theory with potential practical value is the C₄ photosynthetic pathway. Most plants use C₃ photosynthesis, in which the enzyme Ribulose-1,6-bisphosphate carboxylase/oxygenase (RuBisCO) captures carbon dioxide (CO₂) in the mesophyll of the leaf. This enzyme evolved during a period of Earth's history when atmospheric CO₂ levels were much higher than today (Hayes, 1994). Thus, the fact that RuBisCO will also bind gaseous oxygen, producing potentially toxic phosphoglycolate that must be converted into useful metabolites, was of little consequence for most of the history of land plants (Sage, 1999). However, under higher concentrations of atmospheric oxygen, the energy-wasting oxygenation of RuBisCO and metabolism of phosphoglycolate, known collectively as photorespiration, become more significant selective forces (Sage *et al.*, 2012). In C₄ plants, a different enzyme, phosphoenolpyruvate carboxylase (PEPC), is used to capture CO₂ in the mesophyll. This enzyme is much more specific in its binding and does not accidentally fix oxygen instead of CO₂. The intermediate carbon molecule formed by PEPC and CO₂ is transported into specialized cells around the leaf vasculature called bundle sheath cells, where the CO₂ is released via the action of one of several different enzymes. Expression of RuBisCO in C₄ plants is restricted to the bundle sheath cells, in which CO₂ becomes highly concentrated compared to the mesophyll and the atmosphere (Kellogg, 2013). Thus, the efficiency of RuBisCO is maximized, and the energy that would be wasted in photorespiration can be diverted to other purposes.

By dividing carbon fixation into two steps, C₄ plants use CO₂ much more efficiently than did their C₃ progenitors (Sage, 2004; Kellogg, 2013). This efficiency grants C₄ plants an

advantage in conditions in which CO₂ is limited, as in arid environments in which stomata need to be closed in order to reduce water loss or in open habitats in which the availability of light overwhelms the supply of CO₂ (Taylor *et al.*, 2014). Several of the most important and productive crops are C₄, including maize, sorghum, and sugar cane (Brown, 1999), and considerable research has been directed towards engineering other major crops like rice, wheat, and soybean to use this pathway (i.e. Sage & Zhu, 2011; Slewinski, 2013; Wang *et al.*, 2014).

The success of the C₄ pathway is also shown by the fact that it has arisen over 60 times in the flowering plants (Sage *et al.*, 2011), with at least 22 independent origins in the grasses (GPWGII, 2011). These parallel transitions from C₃ to C₄ provide a rare example of historical replication, allowing evolutionary hypotheses about this trait to be tested with greater rigor. On average, grass lineages that use the C₄ pathway have greater diversification rates than those that use C₃, although the increase in rate often occurs after the transition to C₄ (Spriggs *et al.*, 2014). The transition to the C₄ pathway also generally coincides with a shift from shaded to open habitats, with subsequent exploitation of arid habitats more likely in C₄ than C₃ grasses (Edwards *et al.*, 2010). However, these general patterns are not without exceptions, and evolutionary history also plays a major role in the patterns seen in particular C₄ lineages (Edwards & Still, 2008). Additionally, the C₄ pathway is accomplished via a diversity of biochemical and anatomical modifications (Sinha & Kellogg, 1996; Liu & Osborne, 2015), and understanding the ways in which independent C₄ manifestations differ is critical to making generalizations about the role of this trait in shaping the evolution of the grasses.

Another major feature in plant evolution that is well-suited to study in Poaceae is polyploidy, or the possession by an organism of three or more complete sets of chromosomes (Ramsey & Schemske, 1998). Recent studies of plant genomes have shown that flowering plants

have experienced at least one ancient whole genome duplication (WGD) in their evolutionary history (Cui *et al.*, 2006; Jiao *et al.*, 2011). The history of the grass family is characterized by three recognizable WGD events: *tau* in the ancestor of most monocots (Jiao *et al.*, 2014), *sigma* in the early Poales (Tang *et al.*, 2010), and *rho* just prior to the origin of the Poaceae (Salse *et al.*, 2008; McKain *et al.*, 2016). Polyploidy can occur through duplication of the genome within a species (autopolyploidy), or through duplication of the genomes associated with a hybridization event between species (allopolyploidy) (Stebbins, 1947). Evidence for both kinds of polyploidy is abundant for many grass clades, including the bamboos (Triplett *et al.*, 2014), the tribe Andropogoneae (Estep *et al.*, 2012), the genus *Panicum* (Triplett *et al.*, 2012), and the subfamily Danthonioideae (Linder & Barker, 2014).

The effects of polyploidy on subsequent evolution in a lineage remain equivocal. Genome doubling is associated with many physiological and developmental changes in plants (Levin, 1983; Otto & Whitton, 2000). Polyploid species have also been identified as being more prevalent in the arctic (Brochmann *et al.*, 2004), and increased ploidy may have been associated with success during the Cretaceous-Tertiary mass extinction (Lohaus & Van de Peer, 2016). Polyploid plants have also been shown to be more likely to be invasive than their diploid progenitors (Pandit *et al.*, 2011). Ancient polyploidy is associated with increased diversification rates, though often only after a substantial lag period (Tank *et al.*, 2015). On the other hand, most recently-formed polyploid lineages diversify more slowly than their diploid relatives (Mayrose *et al.*, 2011), and several authors have considered polyploidy to be an evolutionary dead end (Stebbins, 1971; Arrigo & Barker, 2012). Consensus as to how auto- and allopolyploidy affect evolution in plants is slowly growing, however, as more ancient and recent WGD events are identified and modern techniques are used to characterize them (Madlung, 2013; Kellogg, 2016).

A robust phylogenetic systematic framework is needed to test evolutionary hypotheses. In this regard again the Poaceae is exceptional as a result of centuries of taxonomic study. Phylogenetic analyses of the family have identified two large sister clades, named BOP and PACMAD after their constituent subfamilies, that each contain roughly half of the species diversity in Poaceae (Clark *et al.*, 1995; GPWG, 2001; GPWGII, 2011). Rice, wheat, bamboos, and most of the cool-season grasses are included in the BOP clade (Bambusoideae, Oryzoideae, and Pooideae), while maize, sorghum, sugar cane, and the tropical savannah grasses are included in PACMAD (Panicoideae, Aristidoideae, Chloridoideae, Micrairoideae, Arundinoideae, and Danthonioideae). The deepest phylogenetic splits in Poaceae separate three relatively species-poor subfamilies from each other and from the BOP+PACMAD clade: Anomochlooideae - four species in two genera, Pharoideae - twelve species in three genera, and Puelioideae – eleven species in two genera (Kellogg, 2015; Soreng *et al.*, 2015).

Analyses of relationships between the PACMAD subfamilies have generally treated the Aristidoideae as sister to the remaining subfamilies, with Panicoideae sister to a clade consisting of the pairs Chloridoideae+Danthonioideae and Arundinoideae+Micrairoideae (Clark *et al.*, 1995; GPWGII, 2011). However, Cotton *et al.* (2015) found support using whole-chloroplast genome sequence data for an alternative topology in which the Panicoideae is sister to the rest of PACMAD, though the contrasting topology cannot be rejected with their data. Aside from the position of the Aristidoideae, the relationships between the PACMAD subfamilies appear to be robust to additional sampling of taxa and molecular markers. Relationships have also been identified for many of the major clades within the four largest subfamilies: Panicoideae (Aliscioni *et al.*, 2003; Doust *et al.*, 2007; Sanchez-Ken & Clark, 2007), Chloridoideae (Duvall

et al., 2016), Aristidoideae (Cerros-Tlatilpa *et al.*, 2011), and Danthonioideae (Barker *et al.*, 2007).

Subfamilies Arundinoideae and Micrairoideae are the two smallest and least well-studied subfamilies in the PACMAD clade. Molecular phylogenetic analyses identify a clade consisting of these two subfamilies that is sister to the Chloridoideae+Danthonioideae (GPWG II, 2011). The clade formed by the Arundinoideae and Micrairoideae possesses a remarkable amount of morphological and ecological diversity given the relatively small number (<200) of species it contains. This diversity among a manageable number of species presents an opportunity to investigate many evolutionary phenomena that are of interest in this clade, among the rest of the PACMAD grasses, and among grasses and flowering plants in general. In many ways the Arundinoideae and Micrairoideae form a snapshot of grass diversity in miniature, including a unique origin of the C₄ photosynthetic pathway, transitions to cold climates from tropical ones, adaptation to both dry and aquatic habitats, development of woody culms, evolution of multiple ploidy levels, and one of only two genera in Poaceae to have spiral phyllotaxis.

Study System: Arundinoideae and Micrairoideae

Since its description by Beilschmied in 1833, the subfamily Arundinoideae has been used as a holding place for taxa that did not fit well elsewhere in the classification of the grasses. As a result, the generic composition of the subfamily has varied widely between treatments. Tateoka (1957) included seventeen tribes in subfamily Arundoideae, including members from across the currently recognized grass phylogeny. Renvoize (1981) examined leaf blade anatomy in 72 genera that could not be assigned to one of four anatomically distinct subfamilies: Bambusoideae, Pooideae, Chloridoideae, and Panicoideae. Using a multivariate analysis of 65

coded anatomical characters, he identified a "core Arundinoideae" of 43 genera that possess a unique set of characters separating them from the other four subfamilies. Aside from a few wildly misplaced taxa (i.e. *Lygeum*, a member of the Pooideae in the BOP clade), this assemblage contains what would later become the current Danthoioideae and Arundinoideae.

Following the analysis of Renvoize (1981), Clayton & Renvoize (1986) circumscribed the Arundinoideae to include four tribes: one tribe containing the "core Arundinoideae" and three others to accommodate the taxa identified by Renvoize (1981) as "peripheral". The authors consider this subfamily to contain the ancestors of the tropical savannah grasses (i.e. subfamilies Panicoideae and Chloridoideae), noting the geographically fragmented distribution of the Arundinoideae as evidence of a declining group. This position was supported by Conert (1986), who described the Arundinoideae as a "very old group", also citing the scattered geographic ranges and small numbers of species among the arundinoid genera. A phylogenetic analysis of structural characters by Kellogg & Campbell (1987) revealed that the Arundinoideae was polyphyletic.

Watson & Dallwitz (1992), using a phenetic approach, recognized eleven tribes within subfamily Arundinoideae, including *Micraira* and the Eriachneae and with members of modern Danthoioideae forming a separate tribe. The heterogeneous nature of the Arundinoideae was recognized by these authors, who referred to it as "...an unsatisfactory assemblage of convenience, which is not amenable to anything approaching a diagnostic description, and is probably polyphyletic" (Watson & Dallwitz, 1992 p. 47). A common theme throughout studies of the Arundinoideae is that even with different techniques, structural character sets, and classification schemes, the relationships among the heterogeneous taxa in this subfamily resist

clarification, leading to an unnatural group that obscures the evolutionary history of the grass family.

Molecular phylogenies have been critical in identifying and resolving the chronic polyphyly of the Arundinoideae. Barker *et al.* (1995) and Clark *et al.* (1995) were able to remove several taxa from the Arundinoideae on the basis of chloroplast *rbcL* and *ndhF* sequence data, respectively. These studies also identified a core Arundinoideae consisting of *Arundo*, *Phragmites*, *Molinia*, and, in the case of Barker *et al.*, *Hakonechloa* and *Monachather*. Barker (1997) and Linder *et al.* (1997) added *Amphipogon*, *Elytrophorus*, and *Styppeiochloa* to this core Arundinoideae on the basis of *rbcL* sequence data. Using the *rpoC2* chloroplast insert, Barker *et al.* (1999) found a strongly supported relationship between the South African genus *Dregeochloa* and *Phragmites*, but were unable to resolve the placement of *Arundo* and placed *Amphipogon* outside the restricted Arundinoideae. Two phylogenetic analyses using ribosomal ITS sequences Hsiao *et al.* (1998; 1999) found support for a broader monophyletic Arundinoideae that includes taxa falling out in the Danthonioideae, Aristidoideae, and Panicoideae in chloroplast analyses. However, the combined analysis of multiple data sets including ITS and chloroplast sequences as well as structural characters from across the grass family, only a reduced "arundinoid core" excluding the Danthonieae and *Aristida* was found to form a clade (GPWG, 2001). This result was supported by the largest phylogenetic analysis of grass species using three chloroplast markers (GPWG II, 2011) and by a recent study using whole-chloroplast genomes (Cotton *et al.*, 2015).

In the only phylogenetic study focused on the reduced set of arundinoid taxa, Linder *et al.* (1997) found weak morphological support for a clade consisting of several other putative members of Arundinoideae, including *Crinipes*, *Leptagrostis*, *Piptophyllum*, *Nematopoa*,

Zenkeria, and *Styppeiochloa*. These genera were moved to the Arundinoideae from what is now the Chloridoideae by De Winter (1961), Jacques-Félix (1962), and Hubbard (1967) on the basis of leaf anatomy. The clade formed by these taxa is resolved in the Linder *et al.* (1997) phylogeny as sister to a clade consisting of *Phragmites*, *Molinia*, *Arundo*, *Hakonechloa*, and *Dichaetaria*. In this analysis, *Arundo* is more closely related to *Phragmites* than the latter is to *Hakonechloa*, which is at odds with earlier chloroplast topologies. The African genus *Alloeochaete* is considered by Linder *et al.* to be a member of the Danthonioideae, though they acknowledge that difficulties in rooting the morphological tree make resolving relationships between the broader clades in the analysis difficult. Since this genus has not been included in any phylogenetic analysis in the Danthonioideae, and its placement in this subfamily is equivocal, it is treated here as a putative member of the Arundinoideae.

Two other poorly-studied monotypic genera have been placed in the Danthonieae by Watson & Dallwitz (1992) and are treated here as putative arundinoids. They are the Ethiopian genus *Phaenanthoecium* and the Indian genus *Danthonidium*. The spikelet of *Phaenanthoecium* is similar to that found in the Danthonioideae (Kabuye & Renvoize, 1975), but the same general spikelet morphology is found in the "crinipoid group" in Arundinoideae. *Danthonidium* was described by Linder *et al.* (1997) as being part of a group of genera "probably misplaced in the Arundineae", but that group included *Amphipogon*, which Barker (1997) showed is closely related to *Arundo*.

Because the Arundinoideae have no known synapomorphies, it is also possible that there are taxa currently classified in other subfamilies that belong in this group. An example of such a discovery using a phylogenetic analysis of chloroplast and nuclear markers is presented in Ingram *et al.* (2011). The authors of this study sought to resolve the paradox of *Eragrostis*

walteri, the only known C₃ species in an otherwise entirely C₄ genus, and found that this taxon is not a member of *Eragrostis* at all, but rather most likely belongs in the Arundinoideae.

The nineteen genera (including *Eragrostis walteri*) remaining in the more limited Arundinoideae are still heterogeneous morphologically and ecologically, with no clear geographic center of diversity. Nine genera occur across tropical East Africa, two are endemic to Australia, four occur only in east Asia, and a few, like *Phragmites*, *Molinia*, *Elytrophorus*, and *Arundo*, have very broad distributions across multiple continents. Eight genera are monotypic, and the most species-rich genus, the Australian endemic *Amphipogon*, has only eight members. As mentioned above, the lack of species numbers in this subfamily is at odds with the very high morphological diversity and disparity among its members. Under more inclusive delimitations, some authors considered the possibility that Arundinoideae is paraphyletic and made up of the relatively unsuccessful ancestors of the savannah grasses in the highly speciose Panicoideae and Chloridoideae (Clayton & Renvoize, 1986; Conert, 1987). This explanation made sense in the context of the polyphyletic former taxonomic treatments of the Arundinoideae, but has become less likely given more recent phylogenies. The sampled Arundinoideae form a relatively young clade within PACMAD and still do not possess any identifiable synapomorphies (GPWG II, 2011). Still, most of the genera in the subfamily have not been included in any molecular analysis, so the possibility that some of those taxa belong in other subfamilies, thus explaining at least part of the morphological disparity of the group, cannot be discarded. If the Arundinoideae as currently circumscribed is monophyletic, it represents a remarkable phylogenetic clustering of morphological evolution in the absence of substantial species accumulation. If it is still polyphyletic, then identifying the proper placement of the genera in this subfamily becomes important for inferring character evolution across the PACMAD grasses.

As an example, the members of Arundinoideae possess a wide range of spikelet characters involved in dispersal and burial of the seed and its protective structures, known collectively as the diaspore (van der Pijl, 1982). In particular, one of the two large bracts wrapped around the grass fruit is called the lemma and ranges from being relatively simple and associated with a rounded spikelet to being adorned with hairs, a pointed base, and a needle-like projection near its apex, known as an awn. These structures have been shown to facilitate guided and active burial, in which the awn changes configuration in response to changes in humidity to propel the diaspore across the ground and into suitable burial sites (Peart, 1979; 1981; Elbaum *et al.*, 2007). Humphreys *et al.* (2010) analyzed the evolution of lemma traits in the Danthonioideae and found that lemmas tended to be awned, pointed, and hairy, or unawned, rounded, and smooth, with relatively few intermediate species possessing a mixture of these traits. The authors describe these two suites of character states as being opposite poles of a "burial syndrome" that represent alternative adaptations to different habitats. They found that possession of hygroscopic (water-sensitive) lemma awns is the ancestral condition in subfamily Danthonioideae, and that loss of these awns corresponds with changes in life history and a statistically insignificant decrease in diversification rates. However, the applicability of these results to other subfamilies of grasses or to the PACMAD grasses as a whole is unknown. Resolving the phylogenetic relationships in Arundinoideae would address this problem in two ways. First, the subfamily possesses both extremes of the burial syndrome among a small number of species, so the differences in species numbers in this group are unlikely to be attributable to differences in burial syndrome. Second, if Arundinoideae is polyphyletic, the misplaced taxa could substantially change estimates of ancestral states depending on where they belong in the phylogeny. The same

logic applies for a large host of other morphological and anatomical traits, making systematic studies in the Arundinoideae especially appealing.

The Arundinoideae are also noteworthy for genomic evolution and ecological invasion of new habitats. Most taxa in the subfamily reside in the Old World Tropics, but two separate invasions of the temperate zone can be inferred from chloroplast phylogenies. One such transition seems to have occurred in the ancestor of *Phragmites*, *Molinia*, and *Hakonechloa*, while the other occurred in the ancestor of the species of *Arundo*. *Phragmites* consists of four species, with *P. australis*, or common reed, possessing a nearly global distribution. *Molinia* and *Hakonechloa* have two and one species, respectively. *Molinia* has a distribution extending across Europe and in China, while *Hakonechloa* is restricted in its native range to the main island of Japan. *Arundo* has five species with a center of diversity in Eurasia (Hardion *et al.*, 2012). One species, *A. donax* (giant reed), is similar to *P. australis* in being a large-statured invasive reed (Lambert *et al.*, 2010) and in having a cosmopolitan distribution, although *A. donax* tends to avoid the cold more than *P. australis*. These four genera are all polyploids, with ploidy levels up to 12x in *Molinia* (Dančák *et al.*, 2012) and *Phragmites* (Clevering & Lissner, 1999) and up to 10x in *Arundo* (Bucci *et al.*, 2013). A great deal of physiological and genetic research has been conducted on *P. australis* and *A. donax* due to their invasiveness (Saltonstall, 2002) and potential use as biofuels (Laurent *et al.*, 2015), and *Molinia caerulea* is a major component of European heathland (Taylor *et al.*, 2001), but the source of the duplicated genomes in Arundinoideae is unknown. Elucidation of the history of the genomes in these taxa could provide valuable insights into the evolution of cold tolerance and invasiveness as well as potentially explain the convergence in morphology and ecology between *Phragmites* and *Arundo*.

Relationships within the Micrairoideae have also been difficult to resolve until relatively recently. Bentham (1878; 1881) placed the genera *Isachne*, *Eriachne*, and *Micraira* in the same tribe, but subsequent authors placed *Micraira* in various tribes and subfamilies, including the Aveneae in Pooideae (Bentham and Hooker, 1883), the Bambusoideae (Tateoka, 1957), Arundinoideae (Clayton & Renvoize, 1986; Watson & Dallwitz, 1992), and Eragrostoideae (Clifford, 1964). Pilger (1954) and Lazarides (1979) placed the genus in its own subfamily, Micrairoideae because of the difficulty in assigning it to any existing taxonomic group within the Poaceae. The taxonomic history of *Eriachne* has been similarly inconsistent, with different authors placing the genus in the Aveneae (Hubbard, 1973), the Danthonieae (Watson & Clifford, 1976), and the Aristideae (Brown, 1977). Eck-Borsboom (1980) ruled out these placements on the basis of morphology and anatomy and erected a tribe, Eriachneae, that includes *Eriachne* and its close relative *Pheidochloa*. The Isachneae, including the genera *Isachne*, *Coelachne*, *Heteranthoecia*, *Limnopoa*, and *Sphaerocaryum*, was recognized as a natural group on the basis of leaf anatomy by Hubbard (1943) and supported by subsequent studies by Potztl (1952) and Prakash & Jain (1987). Later, the rare Indian genus *Hubbardia* was described by Bor (1950) and added to the tribe. The Isachneae were placed in the Paniceae by most authors (i.e. Pilger, 1954; Jacques-Felix, 1962; Clayton & Renvoize, 1986; Watson & Dallwitz, 1992; GPWG, 2001), although some authors thought it was better suited to the Pooideae (Bor, 1960) or in its own subfamily (Prakash & Jain, 1984).

Once again, molecular phylogenies fundamentally changed the classification of the Micrairoideae. The classification of the grass family by GPWG (2001) left *Micraira* and *Eriachne* as *incertae sedis* due to a lack of support for their placement in the phylogeny. However, Duvall *et al.* (2003) found that *Isachne*, *Eriachne*, and *Pheidochloa* form a clade that

is separate from the Panicoideae in a phylogenetic analysis of that subfamily using *rpoC2* and *ndhF* chloroplast markers. A subsequent study by Duvall *et al.* (2007), aimed at filling in some of the taxonomic holes in the GPWG (2001) study, found support for a clade consisting of *Isachne*, *Eriachne*, and *Micraira* that is sister to the Arundinoideae. This result was supported by Sánchez-Ken & Clark (2007) using *ndhF* and *rpl16* intron sequences as well as structural data, which led Sánchez-Ken *et al.* (2007) to reinstate the subfamily Micrairoideae on the basis of expanded sampling within the group. Thus was the "M" put in "PACMAD".

The Micrairoideae in its current circumscription consists of approximately 188 species divided between the three tribes Micraireae, Isachneae, and Eriachneae. Micraireae has only one genus, *Micraira*, which contains fifteen species occupying a disjunct distribution featuring one species in eastern Queensland and the other fourteen species in the Northern Territory (Lazarides, 1979; Lazarides *et al.*, 2005). This odd genus possesses a moss-like growth habit, with spiral phyllotaxis (Philipson, 1935) and forming dense mats on seasonally moist, rocky sites. The Micraireae form the sister clade to a clade containing the Isachneae and Eriachneae. Tribe Isachneae contains roughly 120 species in the six genera mentioned above, with the vast majority of species belonging in the widespread genus *Isachne*. Members of this tribe occupy primarily moist habitats in the tropics and subtropics of Australia, Southeast Asia, Central Africa, and South America, with a greater diversity of species in the eastern hemisphere and particularly in Indo-Malaysia (Prakash & Jain, 1987a). The Eriachneae consists of two genera, *Eriachne* and *Pheidochloa*, with 48 and 2 species, respectively. Members of this tribe are mostly endemic to Australia, occupying dryer, open habitats across the continent.

From an evolutionary standpoint, the Micrairoideae represents a unique opportunity to investigate the impact of the acquisition of C₄ photosynthesis on diversification and ecology in a

clade that appears to oppose general trends for grasses. As mentioned above, the C₄ pathway is on average associated with higher diversification rates compared with C₃ sister clades in grasses (Spriggs *et al.*, 2014). Part of the explanation for this phenomenon is that the C₄ pathway allows a plant access to habitats that would otherwise be too harsh in terms of solar intensity, heat, and aridity (Taylor *et al.*, 2014). In the Micrairoideae, however, both of these general trends are reversed, in that the C₃ grasses in Isachneae occupy a broader geographic range and include more species than their C₄ sister tribe Eriachneae. Additionally, members of the tribe Isachneae appear to occupy a wider range of habitat types, with those of Eriachneae and Micraireae preferring open or shady and rocky or sandy sites in mostly arid habitats. If this is taken as the ancestral habitat for the Micrairoideae, then C₃ Isachneae experienced a diversification following a transition into shadier, wetter habitats, while its C₄ sister taxon Eriachneae remained in habitats more similar to the ancestral condition and diversified more slowly. This pattern is inconsistent with the general trend found in Hawaii for Paniceae (Christin & Osborne, 2014) and for most other PACMAD lineages (Edwards *et al.*, 2010). However, only a few species in the Micrairoideae have been studied phylogenetically, and the taxonomic limits of the C₄ pathway within Eriachneae have not been broadly tested. Additionally, the tendency for Eriachneae and Micraireae to occupy similar habitats as compared to Isachneae has not been evaluated quantitatively.

Outline of the Dissertation

Three evolutionary phenomena in the Arundinoideae and Micrairoideae are the subjects of this dissertation. In Chapter 1, I explore the evolution of morphological characters associated with seed dispersal and burial in the PACMAD grasses using a whole-chloroplast genome

phylogeny of the grasses with a focus on sampling members of the Arundinoideae. This subfamily has the potential to affect ancestral state estimates across the PACMAD grasses because it possesses a wide range of character states and is probably polyphyletic. As one example, I use the plastome phylogeny to analyze the evolution of three lemma traits that have been shown to be important in grass dispersal and burial. Since this is also the largest sampling of whole-chloroplast genomes in a phylogeny of the grass family, I also estimate divergence dates for the family and use those dates to test for significant shifts in diversification rate within the PACMAD clade.

The focus of Chapter 2 is on the evolution of polyploidy among the temperate members of Arundinoideae. I combine novel transcriptome data from *Arundo donax*, *Phragmites australis*, *Molinia caerulea*, and *Hakonechloa macra* with existing data from the subfamilies Panicoideae, Oryzoideae, Bambusoideae, and Anomochlooideae in a phylogenetic framework to identify potential whole genome duplications in the Arundinoideae and PACMAD clades. I also discuss the possible implications of polyploidy on the ecology and evolutionary history of *A. donax* and *P. australis*, which are both large invasive reeds with cosmopolitan distributions.

In Chapter 3, I address the role that the acquisition of C₄ photosynthesis has played in the Micrairoideae. I first examine the phylogenetic distribution of the C₄ pathway in this poorly-studied subfamily using carbon-isotope ratios from herbarium specimens. I also compare climatic distributions of C₄ and C₃ taxa using a principal components analysis of nineteen BioClim variables from the WorldClim database. Finally, I construct a whole-chloroplast phylogeny of the C₄ tribe Eriachneae using field collections I gathered in the Northern Territory and Western Australia to test the hypothesis that the pathway has facilitated an adaptive radiation into new habitats.

LITERATURE CITED

- Aliscioni, S. S., L. M. Giussani, F. O. Zuloaga, and E. A. Kellogg. 2003. A molecular phylogeny of *Panicum* (Poaceae: Paniceae): tests of monophyly and phylogenetic placement within the Panicoideae. *American Journal of Botany* 90: 796–821.
- Arber, A. 1943. *The Gramineae; a Study of Cereal, Bamboo, and Grass*. New York: MacMillan.
- Arrigo, N., and M. S. Barker. 2012. Rarely successful polyploids and their legacy in plant genomes. *Current Opinion in Plant Biology* 15: 140–146.
- Awika, J. M. 2011. Major cereal grains production and use around the world. Pp. 1-13 in *Advances in Cereal Science: Implications to Food Processing and Health Promotion*, eds. J. M. Awika, V. Piironen, and S. Bean. Washington, DC: American Chemical Society.
- Barker, N. P. 1997. The relationships of *Amphipogon*, *Elytrophorus* and *Cyperochloa* (Poaceae) as suggested by *rbcL* sequence data. *Telopea* 7(3): 205–213.
- Barker, N. P., Linder, H. P., and E. H. Harley. 1995. Polyphyly of Arundinoideae (Poaceae): evidence from *rbcL* sequence data. *Systematic Botany* 20(4): 423–435.
- Barker, N. P., H. P. Linder, and E. H. Harley. 1999. Sequences of the grass-specific insert in the chloroplast *rpoC2* gene elucidate generic relationships of the Arundinoideae (Poaceae). *Systematic Botany* 23: 327-350.
- Barker, N. P., C. Galley, G. A. Verboom, P. Mafa, M. Gilbert, and H. P. Linder. 2007. The phylogeny of the austral grass subfamily Danthonioideae: evidence from multiple data sets. *Plant Systematics and Evolution* 264: 135–156.
- Bentham, G. 1863-1878. *Flora Australiensis*. London: L. Reeve.
- Bentham, G. 1881. Notes on Gramineae. *Journal of the Linnean Society, London, Botany* 19: 93-94.
- Bentham, G. and J. D. Hooker. 1862-1883. *Genera plantarum*. London: L. Reeve.
- Bor, N. L. 1950. A new genus of Indian grasses. *Kew Bulletin* 5:385-388.
- Bor, N. L. 1960. *The Grasses of Burma, Ceylon, India, and Pakistan*. London: Oxford.
- Brochmann, C., A. K. Brysting, I. G. Alsos, L. Borgen, H. H. Grundt, A.-C. Scheen, and R. Elven. 2004. Polyploidy in arctic plants. *Biological Journal of the Linnean Society* 82: 521–536.
- Brown, R. H. 1999. Agronomic implications of C₄ photosynthesis. Pp. 473-508 in *C₄ Plant*

- Biology*, eds. R. F. Sage and R. K. Monson. San Diego, CA: Academic Press.
- Brown, W. V. 1977. The Kranz syndrome and its subtypes in grass systematics. *Memoirs of the Torrey Botanical Club* 23: 63–85.
- Bucci, A., E. Cassani, M. Landoni, E. Cantaluppi, and R. Pilu. 2013. Analysis of chromosome number and speculations on the origin of *Arundo donax* L. (Giant Reed). *Cytology and Genetics* 47: 237–241.
- Cerros-Tlatilpa, R., J. T. Columbus, and N. P. Barker. 2011. Phylogenetic relationships of *Aristida* and relatives (Poaceae, Aristidoideae) based on noncoding chloroplast (trnL-F, rpl16) and nuclear (ITS) DNA sequences. *American Journal of Botany* 98: 1868–1886.
- Christin, P.-A., and C. P. Osborne. 2014. The evolutionary ecology of C₄ plants. *New Phytologist* 204(4): 765–81.
- Clark, L. G., W. Zhang, and J. F. Wendel. 1995. A phylogeny of the grass family (Poaceae) based on *ndhF* sequence data. *Systematic Botany* 20: 436–460.
- Clayton, W. D. and S. A. Renvoize. 1986. *Genera Graminum*. London: Her Majesty's Stationery Office.
- Clevering, O. A., and J. Lissner. 1999. Taxonomy, chromosome numbers, clonal diversity and population dynamics of *Phragmites australis*. *Aquatic Botany* 64: 185–208.
- Clifford, H. T. 1964. The systematic position of the grass genus *Micraira* F. Muell. *University of Queensland Papers* 4: 87–94.
- Conert, H. J. 1987. Current concepts in the systematics of the *Arundinoideae*. Pp. 239–250 in *Grass Systematics and Evolution*, eds. T. R. Soderstrom, K. W. Hilu, C. S. Campbell, and M. E. Barkworth. Washington: Smithsonian Institution Press.
- Cotton, J. L., W. P. Wysocki, L. G. Clark, S. A. Kelchner, J. C. Pires, P. P. Edger, D. Mayfield-Jones, and M. R. Duvall. 2015. Resolving deep relationships of PACMAD grasses: a phylogenomic approach. *BMC Plant Biology* 15(1): 1–11.
- Cui, L., P. K. Wall, J. H. Leebens-Mack, B. G. Lindsay, D. E. Soltis, J. J. Doyle, P. S. Soltis, et al. 2006. Widespread genome duplications throughout the history of flowering plants. *Genome Research* 16: 738–749.
- Dančák, M., M. Duchoslav, and B. Trávníček. 2012. Taxonomy and cytogeography of the *Molinia caerulea* complex in central Europe. *Preslia* 84: 351–374.
- De Winter, B. 1961. *Styppeiochloa* De Winter, gen. nov. *Bothalia* 9:134–137.
- Doust, A. N., A. M. Penly, S. W. L. Jacobs, and E. A. Kellogg. 2007. Congruence, conflict, and

- polyploidization shown by nuclear and chloroplast markers in the monophyletic “Bristle Clade” (Paniceae, Panicoideae, Poaceae). *Systematic Botany* 32: 531–544.
- Duvall, M. R., D. E. Saar, W. S. Gaybum, and G. P. Holbrook. 2003. Complex transitions between C₃ and C₄ photosynthesis during the evolution of Paniceae: a phylogenetic case study emphasizing the position of *Steinchisma hians* (Poaceae) a C₃-C₄ intermediate. *International Journal of Plant Science* 164: 949-958.
- Duvall, M. R., J. I. Davis, L. G. Clark, J. D. Noll, and J. G. Sanchez-Ken. 2007. Phylogeny of the grasses (Poaceae) revisited. In *Monocots: Comparative Biology and Evolution, 2 Volumes*, eds. J. T. Columbus, E. A. Friar, C. W. Hamilton, J. M. Porter, L. M. Prince, and M. G. Simpson. California: Rancho Santa Ana Botanic Garden.
- Duvall, M. R., A. E. Fisher, J. T. Columbus, A. L. Ingram, W. P. Wysocki, S. V. Burke, L. G. Clark, and S. A. Kelchner. 2016. Phylogenomics and plastome evolution of the chloridoid grasses (Chloridoideae: Poaceae). *International Journal of Plant Sciences* 177: 235–246.
- Eck-Borsboom, M. H. J., Van. 1980. Revision of *Eriachne* R. Br. (Gramineae) in Asia and Malesia. *Blumea* 26: 127-138.
- Edwards, E. J., and C. J. Still. 2008. Climate, phylogeny and the ecological distribution of C₄ grasses. *Ecology Letters* 11(3): 266–76.
- Edwards, E. J., S. A. Smith, and M. J. Donoghue. 2010. Phylogenetic analyses reveal the shady history of C₄ grasses. *Proceedings of the National Academy of Sciences of the United States of America* 107(6): 2532–2537.
- Elbaum, R., L. Zaltzman, I. Burgert, P. Fratzl, L. Zaltzman, and I. Burgert. 2007. The role of wheat awns in the seed dispersal unit. *Science* 316(5826): 884–886.
- Estep, M. C., M. R. McKain, D. Vela Diaz, J. Zhong, J. G. Hodge, T. R. Hodkinson, D. J. Layton, *et al.* 2014. Allopolyploidy, diversification, and the Miocene grassland expansion. *Proceedings of the National Academy of Sciences of the United States of America* 111: 15149–15154.
- Glémin, S., and T. Bataillon. 2009. A comparative view of the evolution of grasses under domestication. *New Phytologist* 183: 273–290.
- Grass Phylogeny Working Group [GPWG]. 2001. Phylogeny and subfamilial classification of the grasses (Poaceae). *Annals of the Missouri Botanical Garden* 88: 373-457.
- Grass Phylogeny Working Group II. 2011. New grass phylogeny resolves deep evolutionary relationships and discovers C₄ origins. *New Phytologist* 193(2): 304–312.
- Hardion, L., R. Verlaque, A. Baumel, M. Juin, and B. Vila. 2012. Revised systematics of

- Mediterranean *Arundo* (Poaceae) based on AFLP fingerprints and morphology. *Taxon* 61: 1217–1226.
- Hayes, J. M. 1994. Global methanotrophy at the Archean-Proterozoic transition. Pp. 220-236 in *Early Life on Earth*, ed. S. Bengtson. New York: Columbia University Press.
- Hsiao, C., S. W. L. Jacobs, N. P. Barker, and N. J. Chatterton. 1998. A molecular phylogeny of the subfamily Arundinoideae (Poaceae) based on sequences of rDNA. *Australian Systematic Botany* 11: 41-52.
- Hsiao, C., S. W. L. Jacobs, N. J. Chatterton, and K. H. Asay. 1999. A molecular phylogeny of the grass family (Poaceae) based on the sequences of nuclear ribosomal DNA (ITS). *Australian Systematic Botany* 11: 667-688.
- Hubbard, C. E. 1943. *Limnopoia meeboldii* (C.E.C. Fischer) C.E. Hubbard, in Hooker, *Icon. Pl.*: T. 3432, pp. 1-4.
- Hubbard, C. E. 1967. *Habrochloa bullockii* C. E. Hubb. *Gramineae*. Tribus *Danthonieae* (*sensu lato*). *Icones. Pl.* 37(19) t 3645: 1–5.
- Hubbard, C. E. 1973. In *The Families of Flowering Plants, Arranged According to a New System Based on Their Probably Phylogeny*, J. Hutchinson, ed. Oxford: Oxford University Press.
- Humphreys, A. M., A. Antonelli, M. D. Pirie, and H. P. Linder. 2010. Ecology and evolution of the diaspore burial syndrome. *Evolution* 65(4): 1163–1180.
- Ingram, A. L., P. Christin, and C. P. Osborne. 2011. Molecular phylogenies disprove a hypothesized C₄ reversion in *Eragrostis walteri* (Poaceae). *Annals of Botany* 321–325.
- Jacques-Félix, H. 1962. *Les graminées (Poaceae) d'Afrique Tropicale*. Paris: Institute de Recherches Agronomiques Tropicales et des Cultures Vivrières.
- Jiao, Y., J. Li, H. Tang, and A. H. Paterson. 2014. Integrated syntenic and phylogenomic analyses reveal an ancient genome duplication in monocots. *The Plant Cell* 26: 2792–2802.
- Jiao, Y., N. J. Wickett, S. Ayyampalayam, A. S. Chanderbali, L. Landherr, P. E. Ralph, L. P. Tomsho, *et al.* 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473: 97–100.
- Kabuye, C.H.S. and S.A. Renvoize. 1975. The genus *Alloeochoaete*, tribe Danthonieae (Gramineae). *Kew Bulletin* 30(3): 569-577.
- Kellogg, E. A. 2000. The grasses: a case study in macroevolution. *Annual Review of Ecology & Systematics* 31: 217–238.

- Kellogg, E. A. 2013. C4 photosynthesis. *Current Biology* 23: R594–R599.
- Kellogg, E. A. 2015. Flowering Plants. Vol. 13. Monocots: Poaceae. In *The Families and Genera of Vascular Plants Series*, ed. K. Kubitzki. Heidelberg, Germany: Springer.
- Kellogg, E. A. 2016. Has the connection between polyploidy and diversification actually been tested? *Current Opinion in Plant Biology* 30: 25–32.
- Kellogg, E. A. and C. S. Campbell. 1987. Phylogenetic analyses of the Gramineae. Pp. 310-322 in *Grass Systematics and Evolution*, eds. T. R. Soderstrom, K. W. Hilu, C. S. Campbell and M. E. Barkworth. Washington: Smithsonian Institution Press.
- Lambert, A. M., T. L. Dudley, and K. Saltonstall. 2010. Ecology and impacts of the large-statured invasive grasses *Arundo donax* and *Phragmites australis* in North America. *Invasive Plant Science and Management* 3: 489–494.
- Laurent, A., E. Pelzer, C. Loyce, and D. Makowski. 2015. Ranking yields of energy crops: a meta-analysis using direct and indirect comparisons. *Renewable and Sustainable Energy Reviews* 46: 41–50.
- Lazarides, M. 1979. *Micraira* F. Muell. (Poaceae, Micrairoideae). *Brunonia* 2: 67-84.
- Lazarides, M., A. Van Den Borre, and C. M. Weiller. 2005. Micraireae. Pp. 132-176 in *Flora of Australia. 44B. Poaceae 3*, ed. K. Mallet. Australia: CSIRO Publishing.
- Levin, D. A. 1983. Polyploidy and novelty in flowering plants. *The American Naturalist* 122: 1–25.
- Linder, H. P., and N. P. Barker. 2014. Does polyploidy facilitate long-distance dispersal? *Annals of Botany* 113: 1175–1183.
- Linder, H.P., G. A. Verboom, and N.P. Barker. 1997. Phylogeny and evolution in the *Crinipes* group of grasses (Arundinoideae: Poaceae). *Kew Bulletin* 52(1): 91-110.
- Liu, H., and C. P. Osborne. 2015. Water relations traits of C₄ grasses depend on phylogenetic lineage, photosynthetic pathway, and habitat water availability. *Journal of Experimental Botany* 66(3): 761–73.
- Lohaus, R., and Y. Van de Peer. 2016. Of dups and dinos: evolution at the K/Pg boundary. *Current Opinion in Plant Biology* 30: 62–69.
- Madlung, A. 2013. Polyploidy and its effect on evolutionary success: old questions revisited with new tools. *Heredity* 110: 99–104.
- Mayrose, I., S. H. Zhan, C. J. Rothfels, K. Magnuson-Ford, M. S. Barker, L. H. Rieseberg, and S. P. Otto. 2011. Recently formed polyploid plants diversify at lower rates. *Science* 333:

1257–1257.

- McKain, M. R., H. Tang, J. R. McNeal, S. Ayyampalayam, J. I. Davis, C. W. dePamphilis, T. J. Givnish, *et al.* 2016. A phylogenomic assessment of ancient polyploidy and genome evolution across the Poales. *Genome Biology and Evolution* 8(4): 1150–1164.
- Otto, S. P., and J. Whitton. 2000. Polyploid incidence and evolution. *Annual Review of Genetics* 34: 401–437.
- Pandit, M. K., M. J. O. Pocock, and W. E. Kunin. 2011. Ploidy influences rarity and invasiveness in plants. *Journal of Ecology* 99: 1108–1115.
- Peart, M. H. 1979. Experiments on the biological significance of the morphology of seed dispersal units in grasses. *Journal of Ecology* 67(3): 843–863.
- Peart, M.H. 1981. Further experiments on the biological significance of the morphology of seed-dispersal units in grasses. *Journal of Ecology* 69(2): 425–436.
- Philipson, W.R. (1935). A grass with spiral phyllotaxis: *Micraira subulifolia*. *Bulletin of Miscellaneous Information (Royal Botanic Gardens, Kew)* 1935(5): 324–326.
- Pilger, R. 1954. Das System der Gramineae (excl. Bambusoideae). *Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie* 76: 365-366.
- Porensky, L. M., J. Davidson, E. A. Leger, W. W. Miller, E. M. Goergen, E. K. Espeland, and E. M. Carroll-Moore. 2014. Grasses for biofuels: A low water-use alternative for cold desert agriculture? *Biomass & Bioenergy* 66: 133–142.
- Potzta, E. 1952. Über die Blattanatomie der Isachneae. *Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie* 75: 551-569.
- Prakash, V. and S. K. Jain. 1984. Poaceae: Tribe – *Isachneae*. *Fascicles of Flora of India* 14:7-37.
- Prakash, V. and S. K. Jain. 1987. Anatomy of the grass tribe *Isachneae* (Fam. Poaceae) and its taxonomic significance-I. *Journal of the Indian Botanical Society* 66: 89-94.
- Prakash, V. and S. K. Jain. 1987a. On the phytogeography of the tribe *Isachneae* (Poaceae). *Journal of the Indian Botanical Society* 66: 107-115.
- Ramsey, J. and D. W. Schemske. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology & Systematics* 29: 467–501.
- Renvoize, S. 1981. The subfamily *Arundinoideae* and its position in relation to a general classification of the *Gramineae*. *Kew Bulletin* 36:85-102.

- Sage R. F. 1999. Why C₄ photosynthesis? Pp. 3-16 in *C₄ Plant Biology*, eds. R. F. Sage, R. K. Monson. San Diego, CA: Academic Press.
- Sage, R. F. 2004. The evolution of C₄ photosynthesis. *New Phytologist* 161(2): 341–370.
- Sage, R. F., P.-A. Christin, and E. J. Edwards. 2011. The C₄ plant lineages of planet earth. *Journal of Experimental Botany* 62(11): 3155–69.
- Sage, R. F., T. L. Sage, and F. Kocacinar. 2012. Photorespiration and the evolution of C₄ photosynthesis. *Annual Review of Plant Biology*. 63: 19–47.
- Sage, R. F., and X.-G. Zhu. 2011. Exploiting the engine of C₄ photosynthesis. *Journal of Experimental Botany* 62(11): 2989–3000.
- Salse, J., S. Bolot, M. Throude, V. Jouffe, B. Piegu, U. M. Quraishi, T. Calcagno, *et al.* 2008. Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution. *The Plant Cell* 20: 11–24.
- Saltonstall, K. 2002. Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. *Proceedings of the National Academy of Sciences of the United States of America* 99: 2445–2449.
- Sánchez-Ken, J. G. and L. G. Clark. 2007. Phylogenetic relationships within the clade Centothecoideae + Panicoideae (Poaceae), based on *ndhF* and *rpl16* intron sequences and morphological data. In *Monocots: Comparative Biology and Evolution, 2 Volumes*, eds. J. T. Columbus, E. A. Friar, C. W. Hamilton, J. M. Porter, L. M. Prince, and M. G. Simpson. Claremont: Rancho Santa Ana Botanic Garden.
- Sánchez-Ken, J. G., L. G. Clark, E. A. Kellogg and E. E. Kay. 2007. Reinstatement and emendation of subfamily Micrairoideae (Poaceae). *Systematic Botany* 32(1): 71-80.
- Sinha, N. R., and E. A. Kellogg. 1996. Parallelism and diversity in multiple origins of C₄ photosynthesis in the grass family. *American Journal of Botany* 83(11): 1458–1470.
- Slewinski, T. L. 2013. Using evolution as a guide to engineer Kranz-type C₄ photosynthesis. *Frontiers in Plant Science* 4: 1–13.
- Soreng, R. J., P. M. Peterson, K. Romaschenko, G. Davidse, F. O. Zuloaga, E. J. Judziewicz, T. S. Filgueiras, *et al.* 2015. A worldwide phylogenetic classification of the Poaceae (Gramineae): phylogenetic classification of the grasses. *Journal of Systematics and Evolution* 53(2): 117–37.
- Spriggs, E. L., P.-A. Christin, and E. J. Edwards. 2014. C₄ photosynthesis promoted species diversification during the Miocene grassland expansion. *PLoS ONE* 9(5): 1–10.
- Stebbins, G. L. 1947. Types of polyploids: their classification and significance. *Advances in*

Genetics 1: 403–429.

- Stebbins, G. L. 1971. *Chromosomal Evolution in Higher Plants*. London: Edward Arnold.
- Tang, H., J. E. Bowers, X. Wang, and A. H. Paterson. 2010. Angiosperm genome comparisons reveal early polyploidy in the monocot lineage. *Proceedings of the National Academy of Sciences of the United States of America* 107: 472–477.
- Tank, D. C., J. M. Eastman, M. W. Pennell, P. S. Soltis, D. E. Soltis, C. E. Hinchliff, J. W. Brown, *et al.* 2015. Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytologist* 207: 454–467.
- Tateoka, T. 1957. Miscellaneous papers on the phylogeny of the Poaceae (10). Proposition of a new phylogenetic system of Poaceae (10). *Japanese Journal of Botany* 32: 275–287.
- Taylor, K., A. P. Rowland, and H. E. Jones. 2001. *Molinia caerulea* (L.) Moench. *Journal of Ecology* 89: 126–144.
- Taylor, S. H., B. S. Ripley, T. Martin, L.-A. De-Wet, F. I. Woodward, and C. P. Osborne. 2014. Physiological advantages of C₄ grasses in the field: a comparative experiment demonstrating the importance of drought. *Global Change Biology* 20(6): 1992–2003.
- Triplett, J. K., L. G. Clark, A. E. Fisher, and J. Wen. 2014. Independent allopolyploidization events preceded speciation in the temperate and tropical woody bamboos. *New Phytologist* 204: 66–73.
- Triplett, J. K., Y. Wang, J. Zhong, and E. A. Kellogg. 2012. Five nuclear loci resolve the polyploid history of switchgrass (*Panicum virgatum* L.) and relatives. *PLoS ONE* 7: 1–19.
- Van der Pijl, L. 1982. *Principles of Dispersal in Higher Plants*. Berlin: Springer-Verlag.
- Vavilov, N. I. 1922. The law of homologous series in variation. *Journal of Genetics* 12: 47–89.
- Wang, L., R. A. Mertz, L. K. Dedow, D. W. Bryant, S. Weissmann, A. Studer, Y. Shao, *et al.* 2014. Comparative analyses of C₄ and C₃ photosynthesis in developing leaves of maize and rice. *Nature Biotechnology* 32(11): 1158–65.
- Watson, L. and M. J. Dallwitz. 1992. *The Grass Genera of the World*. Wallingford: CAB International.
- White, R. P., S. Murray and M. Rohweder. 2000. *Pilot Analysis of Global Ecosystems: Grassland Ecosystems*. Washington, DC: World Resources Institute.

CHAPTER 1

"Awn" or Off: Evolution of Dispersal and
Burial Traits in the PACMAD Grasses

1.1 INTRODUCTION

Plant seeds are confronted with the joint challenges of moving away from the parent and getting to an appropriate place for germination. Dispersal can be accomplished through a variety of forces, including wind, water, gravity, and animal movement (Van der Pijl, 1982; Cousens *et al.*, 2008). Modifications of seeds and their accessory dispersal structures, together forming the dispersal unit or "diaspore", facilitate movement via these forces, such as the fur-catching burrs of *Geum* (Sorensen, 1986; Kiviniemi, 1996) or the wind-riding samaras of maple trees (Green, 1980). Additionally, plant structures can generate considerable mechanical force to propel seeds, as in the well-known case of explosive seed pods in the touch-me-not, genus *Impatiens* (Hayashi *et al.*, 2009). As can be seen from these examples, dispersal operates over very different spatial scales for different taxa. Furthermore, there is no guarantee that the habitat reached by the diaspore will be conducive to germination. This is especially problematic for small seeds, as heterogeneity in soil microhabitats can present significant challenges to germination in otherwise suitable broader habitats (Harper *et al.*, 1965; Hamrick & Lee, 1987). The orientation of the seed in or on the soil can also affect germination rates, in large part through water loss from exposed attachment scars (Sheldon, 1974). To address these finer-scale challenges, many species have evolved moisture-sensitive bristles, hairs, or other tissue projections that push or pull the diaspore across the soil surface and help orient the seed in the soil (examples in Van der Pijl, 1982 and below). Diaspore structures thus have broad importance for plant evolution, being involved in both long distance dispersal and facilitation of seed germination in new habitats.

A wide variety of diaspore modifications can be found in the grass family, Poaceae, corresponding to the range of habitats occupied by its members (Davidse, 1987; Kellogg, 2015). At least part of this widespread success may be attributable to the diversity of narrow tissue

extensions called awns that stick out of the grass diaspore, most often occurring at or near the apex of one of the protective floral bracts called the lemma (Figure 1.1). These awns can facilitate long-distance dispersal by sticking to animal fur or human clothing (e.g. Ansong & Pickering, 2013). They may also help orient the diaspore in soil microsites by providing passive structural support (Sheldon, 1974; Peart, 1981), actively moving the diaspore short distances (Peart, 1979), and/or pushing the base of the diaspore into the soil (Garnier & Dajoz, 2001; Schöning *et al.*, 2004; Elbaum *et al.*, 2007; Johnson & Baruch, 2014). The latter two functions occur through moisture-sensitive coiling and uncoiling of the awn, translating environmental variation into unidirectional movement (Kulić *et al.*, 2009; Wolgemuth, 2009). Such hygroscopic (water-sensitive) awns are often associated with stiff hairs at the base of the lemma (called a callus), which prevents the diaspore from being pushed out of the soil from the force of the emerging radicle (Peart, 1979). Hairs on the body of the lemma may similarly serve to anchor the dispersal unit in the soil. These hairs may also serve other purposes in aiding dispersal or deterring herbivory. Differences in these traits have significant effects on germination rates and are thus potentially under great selection pressures (Peart, 1984; Peart & Clifford, 1987). Additionally, active burial via hygroscopic awns may play a role in protecting seeds from fire (Garnier & Dajoz, 2001) and ant predation (Schöning *et al.*, 2004).

Lemma awns are common and diverse in the PACMAD clade, a group containing a little over half of the family's over 11,000 species (Kellogg, 2015). Members of this enormously successful clade have diversified into virtually all habitats around the world and include several dominant prairie and savannah grasses as well as agricultural giants like corn, sugar cane and sorghum. Humphreys *et al.* (2010) explored the evolution of awns in one PACMAD subfamily, Danthonioideae. Their analysis identified a strong association between an apical awn, a hairy

callus, and indumentum on the lemma body in this subfamily. Together these traits form one end of a burial syndrome dichotomy in which species tend to be adapted either for active or passive burial. Humphreys *et al.* argue that presence of awns is ancestral in Danthonioideae and that their loss occurs less frequently than would be expected by chance. The authors cite the tendency for awnless species to lack lemma hairs and hairy calli as possible evidence for selective pressure towards the passive burial syndrome.

Trait combinations from across the grass burial spectrum can be found among the rest of the PACMAD clade, including lemmas with non-geniculate (straight) and geniculate (typically hygroscopic, as in the Danthonioideae) awns. Both types of awn have been shown to guide the orientation of the diaspore to ensure proper burial (e.g. Peart, 1979; 1984). Geniculate awns are also found in the Panicoideae and Arundinoideae in addition to the Danthonioideae, suggesting convergent evolution based on our current understanding of relationships between these subfamilies. Understanding the broader evolutionary trends in lemma traits associated with dispersal and burial could help clarify the role that these processes have had in the diversification of major grass clades. However, such an analysis requires a well-resolved phylogenetic framework.

Subfamily Arundinoideae represents a significant obstacle to inferring character evolution across the PACMAD clade because it currently contains a heterogeneous group of species of uncertain phylogenetic placement. Some of these species possess diaspore characters similar to those found in other distantly-related subfamilies, so that their misplacement within Arundinoideae would artificially increase estimates of how many times such characters have evolved independently. With 50 species divided among 18 genera, Arundinoideae is the smallest subfamily in PACMAD. However, a tremendous amount of morphological and ecological

diversity is contained among these species, indicating that the subfamily may be polyphyletic. This subfamily has a long history of including heterogeneous and unrelated taxa (e.g. Renvoize, 1981; Conert, 1987; Watson & Dallwitz, 1992). Molecular phylogenetic studies were crucial in removing some of these taxa from the polyphyletic Arundinoideae (Barker *et al.*, 1995; 1998; GPWGII, 2011). However, these studies also revealed close relationships between several traditionally arundinoid genera, supporting the recognition of the subfamily. Still, many genera currently included in the subfamily have never been included in a molecular phylogeny, due largely to the difficulty of acquiring field-collected material of these narrowly-distributed and remotely-located species. Herbarium specimens are a tremendously valuable resource and are the only source of morphological, anatomical, and genetic information for many species in Arundinoideae. However, DNA extracted from these specimens is often highly degraded, making PCR amplification and Sanger sequencing of plastome regions difficult or impossible (i.e. Särkinen *et al.*, 2012). In such cases, genome survey sequencing (GSS) and comparison with reference sequences could be a powerful tool (i.e. Besnard *et al.*, 2014). The small size of fragments used in this type of sequencing (500 base pairs or less) is potentially well-suited to the degraded DNA found in herbarium specimens, and the enormous amount of sequence data generated means that rigorous quality control can be used to remove any contaminants or poor-quality fragments.

In this chapter, I explore evolutionary patterns of dispersal- and burial-associated traits across the PACMAD clade. I use a new phylogeny of Poaceae based on full chloroplast genomes and with a focus on Arundinoideae in its current taxonomic sense (called Arundinoideae *sensu lato* in this paper). Most sequences were taken from herbarium samples, including six genera not part of any previous molecular phylogenetic analysis. I included

published plastomes from all other subfamilies in Poaceae to test polyphyly of the Arundinoideae *s.l.* This phylogenetic framework represents the largest whole-chloroplast phylogeny of the grass family published thus far.

1.2 MATERIALS AND METHODS

1.2.1 Taxon Sampling

Fifteen of the nineteen genera currently assigned to Arundinoideae were sampled for DNA, including multiple species within a genus wherever possible (Table S1.1). To test for polyphyly of the subfamily, I also included a broad sample of published plastomes from all other PACMAD subfamilies. I considered the possibility that some "arundinoid" taxa might actually be more closely related to other subfamilies. To test this rigorously, I deliberately included samples of taxa previously identified as phylogenetically outside a group comprising the remaining taxa of each subfamily so I could be confident that placement was not an artifact of limited sampling. Published plastomes for 23 BOP clade taxa as well as samples from the early-diverging grass lineages Anomochlooideae, Pharoideae and Puelioideae were included to test congruence of this larger sample with previously published phylogenies of the family. All other plastomes were taken from GenBank, with the exception of several from subfamily Chloridoideae provided by M. Duvall at Northern Illinois University, *Danthoniopsis dinteri* from Washburn *et al.* (2015), and *Chasmanthium laxum*, which was assembled from genome sequences (Kellogg Lab, unpublished data). In total, 88 full plastomes representing all subfamilies in Poaceae were included in the phylogenetic analysis.

1.2.2 DNA Isolation and Sequencing

Plant material was obtained either from field-dried specimens or from herbarium specimens and ground by hand using a mortar and pestle with sterilized sand. Total DNA was extracted using either the QIAGEN EasyDNA Plant Mini Kit, a modified CTAB protocol (Cota-Sánchez *et al.*, 2006), or a combination of the two in which QIAGEN columns were used to clean and isolate the extracted DNA. Sample DNA was sheared using a Covaris S220 sonicator with peak power of 175 and duty factor of 5.0 for 200 cycles for 30 seconds. Libraries were prepared using the NEBNext Ultra DNA Library Prep Kit for Illumina (New England BioLabs, Inc.) according to the manufacturer's instructions. Fragments were size selected to 400-500bp and purified using AMPure XP Beads (Beckman Coulter, Inc.) and sequenced using an Illumina 2x250 paired-end HiSeq run at the University of Illinois at Urbana-Champaign Roy J. Carver Biotechnology Center.

1.2.3 Plastome Assembly and Phylogenetics

All sequence assemblies and analyses were run on the Apollo Cluster at the Donald Danforth Plant Science Center, the CIPRES Gateway (Miller *et al.*, 2010), or Google Cloud. Raw Illumina paired-end reads were cleaned using Trimmomatic version 0.32 for TruSeq3-PE adapters, using a sliding window of 10 basepairs (bp) with a minimum phred score of 20 and keeping fragments with minimum length 40 (Bolger *et al.*, 2014). Trimmed fragments were assembled initially with SPAdes version 3.1.0 with k values of 55, 87 and 121 (Bankevich *et al.*, 2012). SPAdes output for each sample was assembled with the full trimmed read data set to create longer contigs using **afin** (bitbucket.org/benine/afin) with parameters as follows: a stop extension value of 0.1, an initial trim of 100 bp from contigs, a maximum extension of 100 bp

per loop and 50 search loops. This program trims ends from input contigs and extends their length iteratively using matching trimmed reads, ultimately attempting to connect the resulting extended sequences. Contigs generated by **afin** were assembled by hand into complete plastomes in Sequencher version 5.3 (Gene Codes Corporation) by identifying Inverted Repeat (IR) boundaries and, where necessary, manually searching trimmed reads to connect any remaining fragments. Gaps in the final alignment were filled with N's. Some variation was found between the IR regions in some samples, but read lengths were not long enough to phase SNPs; therefore the Inverted Repeat B (IRB) region was duplicated and inverted to serve as IRA. A coverage analysis (<https://github.com/mrmckain/Chloroplast-Genome-Assembly>) was done on completed plastomes to check assemblies for accuracy, with further modifications to the assemblies made as necessary. Plastome sequences were oriented to start at the beginning of the Large Single Copy (LSC) and end with IRA and will be deposited in GenBank. Annotations and Circos graphs of finished plastomes were done in Verdant (verdant.iplantcollaborative.org)(McKain *et al.*, submitted).

Finished plastomes were divided into three regions for alignment: IRB, SSC and LSC. Each region was aligned using MAFFT version 7.029b with default parameters (Katoh, 2013). The three alignments were concatenated into a single alignment, which was then trimmed with Gblocks version 0.91b (Castresana, 2000). Three options regarding treatment of gaps in Gblocks were used to create edited alignments: 1) all sites with gaps excluded (no gaps), 2) all sites with gaps in less than half of the sampled taxa included (less than half gaps), and 3) all sites included (all gaps). All four alignments – untrimmed, no gaps, less than half gaps, and all gaps – were analyzed using maximum likelihood with RAxML version 8.0.22 with 500 bootstrap replicates (Stamatakis, 2014). Trees with different outgroups were also constructed to test the robustness of

the results. The subfamilies outside BOP-PACMAD – Anomochlooideae, Pharoideae, and Puelioideae – were used as outgroups, as was *Avena sativa* (Pooideae) in a reduced phylogeny of only the PACMAD taxa. Alternative topologies were tested using the Shimodaira-Hasegawa test (Shimodaira & Hasegawa, 1999) in RAxML. The no gaps alignment with *Anomochloa* as an outgroup was also analyzed using Bayesian phylogeny inference in MrBayes version 3.2.6 (Ronquist *et al.*, 2012). Trees were visualized and edited using FigTree version 1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>) and with the *plot.phylo* function in R package **ape** 3.0 (Popescu *et al.*, 2012).

1.2.4 Morphological Character Coding

Observations of morphology were made on herbarium specimens for all genera in Arundinoideae and compared with data taken from the literature (Clayton & Renvoize, 1999; Watson & Dallwitz, 1992 onwards). Three characters associated with seed burial (Peart, 1981; 1984; Humphreys *et al.*, 2010) were coded for all PACMAD taxa in the phylogenetic analysis, using the condition found in the genus as a whole for each species in the phylogeny. The first character, presence and type of awn on the lemma, was coded as either unordered multistate – absent (0), straight (1) or geniculate (2) – or as binary – absent (0), present (1). A hairy callus and indumentum on the lemma were each scored as either absent (0) or present (1). Taxa displaying both character states were scored as polymorphic unless one of the states is rare, in which case the more common state was chosen. The outgroup, *Avena sativa*, was artificially treated as either missing data for all characters or as lacking awns, a hairy callus and lemma indumentum to provide a conservative approach to testing whether or not these traits are ancestral in PACMAD.

1.2.5 Ancestral State and Diversification Estimates

Duplicate species were reduced to a single sample in the phylogeny prior to trait analyses to avoid artificially inflating the influence of those taxa. Character histories were analyzed with parsimony using Mesquite version 3.04 (Maddison & Maddison, 2015), with maximum likelihood using the function *rayDISC* in R package **corHMM** (Beaulieu *et al.*, 2013), and using stochastic character mapping with function *make.simmap* in R package **phytools** (Revell, 2012). Additional Panicoideae genera were added by hand to the phylogeny in Mesquite using Estep *et al.* (2014) and GPWGII (2011) as guides for placement to create an expanded cladogram with more representative sampling of awn types in that subfamily. Three different models of trait evolution were used for likelihood and stochastic analyses. The first model, ER, assumes equal rates of change between all character states. The symmetric model, SYM, assigns different rates to transitions between each pair of character states with equal rates for forward and reverse transitions. The final option, all rates different (ARD), assigns a different rate to each transition, including reversals. In the case of binary characters, the ER and SYM models are identical. Akaike's Information Criterion (AIC) was calculated for these three character models for each character set using the *AIC* function in R's basic stats package (R Core Team, 2014).

BEAST v. 1.8.3 (Drummond *et al.*, 2012) was used to estimate a dated, ultrametric tree to test the effect of increased sampling with whole plastomes on divergence dates within the PACMAD clade and as a basis for analyses of diversification rates. BEAUti v. 1.8.3 was used to set parameters for the analysis. Ten separate identical runs of 100 million generations each were run on the CIPRES Gateway, starting with a random tree and sampling trees every 1,000 generations using an uncorrelated relaxed clock model with a lognormal relaxed distribution and

with a Yule process model of speciation used as a tree prior. A GTR+Gamma+I nucleotide substitution model with four gamma categories was used with base-pair frequencies being estimated from the plastome alignment. Four fossil calibrations were specified as lognormal distributions with mean of zero, standard deviation of one, an offset from zero equal to the estimated age of the fossil minus one, and an initial value of the fossil age. These fossils were assigned positions in the phylogeny according to Vincentini *et al.* (2008) as follows: 7 mya for the node connecting *Setaria* and *Panicum* (Elias, 1942); 19 mya for stem Chloridoideae (Strömberg, 2005); 35 mya for the ancestor of BOP+PACMAD (Strömberg, 2005); and 55 mya for all grass subfamilies excluding Anomochlooideae (Crepet & Feldman, 1991). These groups were also constrained to be monophyletic in the dating analysis to reduce computational effort slightly. LogCombiner v. 1.8.3, distributed with the BEAST package, was used to combine the last 1,000 trees taken from each of the ten BEAST runs, and the concatenated tree file was annotated in TreeAnnotator v. 1.8.3. The annotated tree was examined with FigTree v. 1.4.2.

Bayesian Analysis of Macroevolutionary Mixtures (BAMM) version 2.5.0 (Rabosky, 2014) was used to test for significant shifts in diversification rate across the PACMAD clade, with priors set using the function *setBAMMpriors* and results visualized using R package **BAMMtools** (Rabosky *et al.*, 2014). BAMM is potentially well-suited to the current study because it allows for substantial numbers of missing taxa, provided some information about the placement of those taxa is known. Such is the case in the current phylogeny, as virtually all species in the PACMAD clade can be assigned to a subfamily, and generally to a tribe or other smaller clade within the subfamily. Thus, diversification rate shifts can theoretically be identified, at least at the subfamily or tribal levels. Some other analyses often associated with studies of trait evolution, like testing for significantly asymmetrical character transition

probabilities or phylogenetic-independent correlations between character states (i.e. Pagel, 1994), are inappropriate for the current study due to strongly biased sampling in the phylogeny. Additionally, Maddison & Fitzjohn (2014) argue that phylogenetically-controlled correlation tests for discrete characters suffer from serious flaws that make their use in testing hypotheses of evolution and adaptation ill-advised. In any case, the current study aims to explore macroevolutionary patterns of diaspora evolution across the PACMAD grasses to identify clades of interest and to generate testable hypotheses for future work. For this purpose, we ran BAMM for 1,000,000 generations, sampling every 1,000 generations, on the dated, ultrametric tree produced by BEAST, which was trimmed to include only members of PACMAD. A sampling fraction was applied to each PACMAD tribe in our tree using Kellogg (2015) as a guide for species numbers (Supplementary Table S1.2).

Alignments, trees, BEAST and BAMM control and output files, the sampling fractions file, and morphological states will be stored in the Dryad Digital Repository (www.datadryad.org).

1.3 RESULTS

1.3.1 Plastome Assembly and Alignment

Average single-copy coverage and total plastome length for each of the 29 samples generated by this study are reported in Table S1.1. Average coverage for the single-copy regions ranged from 31x to 452x, with total plastome lengths of 133,327 to 139,395 bp. Lengths of the unedited and Gblocks-trimmed alignments can be found in Table 1.1. They range from just under 80,000 bp when all gaps are excluded to almost 157,000 bp without any trimming, demonstrating the considerable extent of gaps in the full alignment. Part of the reason for this is the inclusion of

Anomochloa, which lacks some characteristic features of grass plastome structure, such as the absence of an *rpoC1* intron and a 39-bp subrepeat in the *rpoC2* insert instead of the 21-bp subrepeat found in the rest of the grasses (Morris & Duvall, 2010). Use of *Pharus* as an outgroup reduces the number of ambiguous regions in the alignment, but as discussed below does not significantly affect inferred phylogenetic relationships.

1.3.2 Phylogenetic Analysis

The ML tree (Figure 1.2) was the result of analysis of the full unedited alignment from MAFFT using *Anomochloa marantoidea* as an outgroup. The tree topology was robust to outgroup sampling and alignment trimming except that the placement of Aristidoideae changed among three different positions (Figure 1.2 insert). Bootstrap support for the placement of this subfamily ranged from 52 to 77% with no topology showing a consistently higher support value across analyses. None of the three alternative topologies could be rejected by a Shimodaira-Hasegawa (SH) test (Shimodaira & Hasegawa, 1999), with log likelihood scores as follows for the unedited alignment: best tree, -930564.68; Panicoideae sister to rest of PACMAD, -930577.20; Aristidoideae+Panicoideae sister to rest of PACMAD, -930565.38.

Monophyly of Arundinoideae *s.l.* was strongly rejected by a SH test (log likelihood -945857.72), with four genera falling into other subfamilies. The Zimbabwean monotypic genus *Nematopoa* groups with members of the Chloridoideae, the Ethiopian monotypic genus *Phaenanthoecium* groups with the Danthonioideae, and *Dichaetaria* and *Alloeochaete* form the sister group to the remainder of the Panicoideae. These placements all have 100% bootstrap support, as does the monophyly of the remaining Arundinoideae. This clade, referred to hereafter as Arundinoideae *s.s.*, includes: the cosmopolitan reeds *Arundo* and *Phragmites*; the temperate

genera *Molinia* and *Hakonechloa*; the African genera *Crinipes*, *Styppeiochloa*, *Dregeochloa* and the misnamed "*Eragrostis*" *walteri*; the Australian genera *Amphipogon* and *Monachather*; and the African-Australian-Asian genus *Elytrophorus*. Relationships among genera in this subfamily are strongly supported, with all but two nodes found in 100% of bootstrap trees. The nodes that are less well supported describe the relationships between *Arundo*, *Amphipogon* and *Dregeochloa*+*Monachather* and are recovered in the maximum likelihood phylogeny of the unedited alignment in 83-87% of bootstrap trees.

1.3.3 Trait Evolution

Parsimony optimizations for presence and form of awns, calli and lemma hairs in PACMAD are given in Figures 1.3-5, respectively. The ancestor of this clade is inferred as having a straight (i.e. non-geniculate) awn (Figure 1.3); the same result occurs when Panicoideae is treated as the sister taxon to the rest of PACMAD and when a clade made up of Aristidoideae and Panicoideae is in this position. Members of Panicoideae, Danthonioideae and Arundinoideae have evolved geniculate awns independently at least five times collectively, and all subfamilies except Aristidoideae have experienced complete loss of awns in one or more taxa (awnless taxa in Danthonioideae were not included in the current phylogeny). Maximum likelihood and stochastic character mapping yielded results similar to parsimony analysis under three models of trait evolution: equal rates (1 parameter), symmetric rates (3 parameters), and all rates different (6 parameters). In both of these sets of analyses, no model was significantly more likely than the others according to likelihood ratio tests. The log-likelihoods for each model and analysis are presented in Table S1.3.

Presence of a hairy callus is also estimated to be the ancestral condition in PACMAD (Figure 1.4), while the presence of hairs on the lemma in this ancestor is unknowable based on current sampling (Figure 1.5). Hairy calli have been lost at least six times, with absence of awns a strong predictor of absence of a hairy callus. Of the 22 taxa unequivocally lacking awns in our analysis, only *Molinia* possesses a consistently hairy callus. Similarly, four out of the 22 taxa lacking awns also lack hairs on the body of the lemma. Among the 18 taxa with predominantly geniculate lemma awns, only 2 lack a hairy callus, while 3 possess mostly hairless lemmas. Straight awns showed associations comparable to geniculate ones, with only 2 taxa out of 19 straight-awned taxa lacking a hairy callus and 6 out of 18 taxa with geniculate awns unequivocally lacking hairs on the lemma.

1.3.4 Tree Dating and Diversification Analysis

BEAST recovered an optimal tree with identical topology and very similar support values as the maximum likelihood tree (Figure 1.6). The placement of Aristidoideae as the sister taxon to the rest of PACMAD is recovered with a posterior probability of 0.81. Relationships between *Amphipogon*, *Dregeochloa*, and *Monachather* are slightly better supported in the BEAST tree, with only the position of *Amphipogon* recovered with less than a posterior probability of 1 (value of 0.93 in BEAST tree as compared to bootstrap value of 87% in RAxML tree). Ages of select clades are given in Table 1.2 with their corresponding 95% highest probability density (HPD) intervals.

BAMM identified two or three shifts in diversification rate across the PACMAD tree as having the highest posterior probability. These shifts were associated with 33 credible shift sets collectively accounting for 95% of the posterior probability from the MCMC analysis; the first

nine of these shift sets are depicted in Figure 1.7. These shift sets most often occur in the core Panicoideae or both the Panicoideae and Chloridoideae as shown by the tree in Figure 1.8 in which branch lengths are proportional to the frequency of inferred shifts in diversification rate occurring on that branch out of the total posterior distribution. As an example, the single best shift set, accounting for 16% of the posterior probability, is shown in Figure 1.9. This shift set contains a rate increase in the common ancestor of the core Panicoideae, which includes the tribes Paniceae, Andropogoneae, and Paspaleae, and another smaller increase in crown Chloridoideae. In many other shift sets, the rate increase in the Panicoideae occurs on the branch leading to the divergence of *Lecomtella* from the core Panicoideae, followed by a rate decrease on the branch leading to *Lecomtella*. The inability of BAMM to distinguish between these alternative scenarios is due in part to the fact that *Lecomtella* is on a long branch in the BEAST tree, so that the prior probability of a rate shift occurring on that branch is fairly high. Decreasing the prior on the number of rate shifts would tend to favor the rate shift after the divergence of *Lecomtella*, while increasing the same prior would favor the scenario with two rate shifts: an increase followed by a decrease in *Lecomtella*.

1.4 DISCUSSION

Polyphyly of Arundinoideae *s.l.* was confirmed by our phylogenetic analysis and has significant consequences for evolutionary inferences across the PACMAD clade. In particular, the placement of *Dichaetaria* and *Alloeochoete* in a small clade that is the sister group to the rest of Panicoideae complicates existing interpretations of early habitat evolution in PACMAD and contributes strongly to estimations of ancestral character states and evolutionary transitions for burial and dispersal characters. These results are discussed below, as are the implications of the

phylogenetic analysis on issues of classification, including placement of former "arundinoid" taxa in other subfamilies and the resulting Arundinoideae *s.s.*

1.4.1 Evolution of Dispersal/Burial Traits

The possession of a straight awn and hairy callus as the ancestral state in PACMAD has several interesting implications. As Humphreys *et al.* (2010) reported in subfamily Danthonioideae, the passive burial syndrome – corresponding to absence of awns and typically hairless lemma body and callus – has evolved multiple times independently across PACMAD lineages. Genuate lemma awns, identified by Humphreys *et al.* as the ancestral state in the Danthonioideae, have originated several times, with no obvious phylogenetic clustering in these origins. Similar to the results of Humphreys *et al.*, possession or lack of awns shows strong associations with the presence/absence of hairy calli and lemma body hairs, supporting the existence of those authors' "burial syndrome" across the PACMAD clade. However, straight awns and genuate awns show similar patterns in my analysis, suggesting that active burial and passive-and-guided burial make use of similar supportive structures.

While testing hypotheses regarding the influence of burial syndrome on diversification rates requires more detailed sampling, there are noteworthy trends in the current broad-scale analysis. In particular, the awnless and smooth-callused condition predominates in such large genera as *Panicum*, *Setaria*, *Eragrostis* and *Sporobolus*. These results are partially compatible with the BAMM analysis in that two out of the three major awnless clades are associated with increased diversification rates. However, the rate shifts occur prior to the inferred loss of awns, suggesting that absence of this character is not the driving force behind increased diversification. Another case of greater species richness in awnless clades occurs in the Micrairoideae, where the

passively buried tribe Isachneae contains over twice as many species as its awned sister tribe, Eriachneae. According to BAMM, the Isachneae are not associated with a higher diversification rate. Thus, at least at the tribal level, increased diversification rate is not strongly associated with absence of lemma awns. It is possible that absence of awns is a better general adaptive strategy, so that genera that occupy a wide range of habitats, and thus tend to accumulate species, tend to lose their awns. The Isachneae occupy a broader geographic and ecological range than the Eriachneae (see Chapter 3 of this dissertation), though whether this has anything to do with awn loss is difficult to establish.

Interestingly, a geniculate awn does not occur in many species-diverse clades. The Danthonioideae, members of which commonly have this trait, is one of the smaller subfamilies, and within Arundinoideae the two geniculate-awned genera contain a total of three species. Similarly, *Alloochaete* and the members of the Tristachyidae in subfamily Panicoideae have low to moderate numbers of species. However, the largely geniculate-awned tribe Andropogoneae is highly diverse, both in numbers of species and in kinds of habitats. This tribe also possesses straight-awned and unawned taxa, making it an ideal candidate group for future studies on this important but understudied trait. Another interesting group in terms of awn structure is the arundinoid clade recovered in this study containing *Arundo*, *Amphipogon*, *Monachather* and *Dregeochloa*. Awnless, straight-awned and geniculate-awned lemmas occur in this very small group (~15 species total), and these genera vary greatly in their geographic extent, from the exclusively South African *Dregeochloa* to the highly cosmopolitan *Arundo*. The large discrepancies in dispersal and burial strategies and successes among so few species make this clade of great interest for understanding grass biogeography and evolution.

1.4.2 Phylogenetic Position of Aristidoideae

Aristidoideae has been recovered as the sister taxon to the remainder of the PACMAD clade (Clark *et al.*, 1995; GPWGII, 2011), or only to the CMAD group (Cotton *et al.* 2015). In this study, I find weak support for an additional relationship, with Aristidoideae and Panicoideae forming a clade that is the sister taxon to the remaining PACMAD. I find no statistical support for any of the three topologies over the others. Likewise, Cotton *et al.* (2015) found that their data could not reject the possibility of Aristidoideae being the sister taxon to a clade comprising the remainder of PACMAD. When I re-analyzed the data of Cotton *et al.* (2015), I recovered their tree topology, but with a bootstrap value of 75% for the branch establishing Panicoideae as the sister taxon to the rest of PACMAD instead of the value of 100% in their published phylogeny. This value in their figure must be an error, as the bootstrap values for this relationship reported in the text of the results section and in the supplemental information of their paper are all between 56% and 91% for alternative outgroup sets. Given the phylogenetic ambiguity of results presented here using broad sampling with full plastomes and of the much larger taxon sampling in the three-gene phylogeny of GPWGII (2011), it appears unlikely that the exact placement of Aristidoideae will be resolved with chloroplast sequence data.

1.4.3 Ancestral PACMAD Habitat

The placement of *Alloeochaete*+*Dichaetaria* as a clade sister to the rest of the Panicoideae makes inference of ancestral habitat preferences equivocal. *Alloeochaete* is an African genus of open-savannah grasses, while *Dichaetaria* occupies shady habitats in southern India and Sri Lanka. Thus there are both open- and closed-habitat species in the early-diverging lineages of the Panicoideae. Previous studies have assumed that early panicoids all occurred in

shady environments (Bouchenak-Khelladi *et al.*, 2010; Cotton *et al.*, 2015). If Panicoideae is the sister taxon to the rest of PACMAD, then diversification of the PACMAD clade can potentially be explained at least in part by the transition to a new habitat type (Cotton *et al.*, 2015). However, because of the positions of Alloeochaete + Dichaetaria, this hypothesis receives less support. Furthermore, two of the three recovered positions of Aristidoideae would suggest open habitats as the most parsimonious ancestral condition for PACMAD. Transitions from open to shaded habitats have also occurred in other subfamilies (i.e. as must be the case for tribe Isachneae in Micrairoideae), suggesting that inferring such a transition in Panicoideae is not unreasonable. Conversely, if the fairly long branch preceding the diversification of the Aristioideae indicates time, then the condition of extant taxa in this clade may not be a reliable indicator of the state in the ancestor of this lineage. The ancestor of all PACMAD in this scenario could have been shade-adapted, with the ancestor to Aristidoideae shifting to an open habitat later in parallel with the rest of the clade. Thus, either habitat can be reasonably inferred as the ancestral one, regardless of which of the tree topologies in Figure 1.2 is chosen, depending on what assumptions are made about transition probabilities between states.

1.4.4 Divergence Date Estimates

Ages inferred in the current analysis are very roughly concordant with previously published estimates with a few key differences. However, published estimates of ages within the grasses vary widely within a broad range of plausible values, so this concordance is almost inevitable. Vincentini *et al.* (2008) reported age estimates for the ancestor of BOP and PACMAD ranging from 48 to 85 mya, while Christin *et al.* (2014) reported ages for the same divergence of 20-62 mya from plastid and 51-63 mya from nuclear sequence data across four different dating

analyses. My analysis yielded a substantially younger age of 35.59 mya for this clade, and the reasons for this discrepancy are somewhat unclear. A tempting explanation would be the increased size of our data set as compared to these studies, which were based on a small number of molecular markers. Cotton *et al.* (2015), using full plastomes with a sampling of 36 taxa across the grasses, reported an age of 32.44 mya with a 95% HPD range of 11.89-50.55 mya for the crown PACMAD clade, which fully encompasses the corresponding ages in the current study: 26.05 mya with a 95% HPD range of 22.55-30.3 mya. However, it is also probable that fossil placement and associated prior distribution parameters are strongly affecting the age estimate for this node. The phytolith described by Strömberg (2005) provides a minimum age of 35 mya for the BOP+PACMAD clade, and the multiflowered spikelet fossil described by Crepet & Feldman (1991) does the same for the clade containing all grasses except the Anomochlooideae and Pharoideae with an age of 55 mya. In the analysis of Vincentini *et al.* (2008), the age of the former clade is estimated at 51.6 mya and the latter at 66.2 mya. This is a slightly smaller time range than is recovered between these two clades in my analysis (35.59 – 55.2 mya), but both reflect the long branch leading from the common ancestor of Pharoideae and the rest of Poaceae to the divergence of Puelioideae. The values in my study for these nodes are very close to their fossil prior age estimates of 35 and 55 mya, suggesting that perhaps my prior distributions are too restrictive. On the other hand, the fact that age ranges in this part of the tree closely match branch lengths in the ML tree would seem to suggest that the fossil priors are not causing extreme stretching of the tree. Thus, while the absolute ages in my analysis may be too young because of excessively strict priors, the relative ages throughout the tree, and therefore estimates of diversification rate changes, appear not to be strongly affected.

1.4.5 Implications for Classification

Polyphyly of Arundinoideae – The placement of *Nematopoa longipes* in Chloridoideae makes some sense given its taxonomic history. The monotypic genus was separated from the chloridoid genus *Triraphis* by Hubbard (1957a), who also cited affinities of these taxa with the genus *Crinipes*. On the basis of the current phylogeny, the similarity between *Nematopoa* and *Triraphis* seems likely due to shared ancestry, while these traits are most likely of convergent origin in *Crinipes*, placed with strong support in Arundinoideae s.s. Leaf cross-sectional anatomy (Figure 1.10) shows *Nematopoa* to be C₄, supporting its placement in the largely C₄ Chloridoideae.

Phaenanthoecium koestlinii is the only member of its genus and occupies shady cliffs in northeast Africa. Its position in the Danthonioideae is supported by its hygroscopic medial awn often found in this subfamily (Humphreys *et al.*, 2010). Indeed, most authors have allied the genus with members of Danthonioideae on the basis of these characters (i.e. Clayton & Renvoize, 1986; Watson & Dallwitz, 1992; Soreng *et al.*, 2015). However, *Phaenanthoecium* was usually joined in this context by other genera, such as *Dregeochloa* and *Alloeochoete*, which possess similar characters but do not form a clade in the current tree.

As mentioned above, the position of *Dichaetaria* and *Alloeochoete* in a clade that is the sister taxon to the rest of the Panicoideae is particularly interesting given their similarity to members of more recently-derived subfamilies. The other early-diverging members of this subfamily are highly morphologically heterogeneous, and these former arundinoid genera provide a potential morphological link between the Panicoideae and the rest of PACMAD.

Arundinoideae s.s. – The genera constituting a reduced Arundinoideae s.s. still form a morphologically and ecologically heterogeneous assemblage, albeit with some commonalities

found within subgroups. Consistent with previous phylogenetic reconstructions, two separate movements into the temperate zone are recovered by our analysis, one comprising the clade (*Hakonechloa*+*Molinia*)+*Phragmites* and the other represented by the genus *Arundo*.

Morphological and ecological parallels between the reedy genera *Phragmites* and *Arundo* are striking, suggesting hybridization or remarkable parallelism.

More broadly, the current arundinoid genera can be divided into two clades, one with glumes shorter than the spikelet and the other with glumes as long as or longer than the spikelet. The former group consists of *Phragmites*, *Hakonechloa*, and *Molinia* as well as another clade of mostly African genera. "*Eragrostis*" *walteri*, formerly thought to be a unique example of reversion from C₄ to C₃ photosynthesis (Ingram *et al.*, 2011), falls in this clade and is sufficiently distinct from its sister taxon *Elytrophorus* that it should be assigned to its own genus. The other two taxa in the "Short Glumes" clade, *Styppeiochloa* and *Crinipes*, are sister taxa as suggested by their taxonomic history (the type species of *Styppeiochloa* was segregated from *Crinipes*) and supported by their similar preference for seasonally wet, rocky habitats, their 1-nerved glumes and their tendency towards spikelets with 2 florets. The "Long Glumes" clade of Arundinoideae contains, in addition to *Arundo* (5 spp.), the Australian genera *Amphipogon* (9 spp.) and *Monachather* (1 sp.) and the South African genus *Dregeochloa* (2 spp.). These last two genera occupy dry habitats and share a version of the active burial syndrome.

Unsourced Putative Arundinoids – Four genera possibly belonging in the Arundinoideae are not included in this study. The monotypic African genera *Leptagrostis* and *Piptophyllum* have insufficient collected material to conduct destructive sampling. Like *Nematopoa*, *Piptophyllum* was placed with a polyphyletic *Triraphis*-*Crinipes* group (Hubbard, 1957b). Herbarium samples

of the Indian genera *Danthonidium* (1 sp.) and *Zenkeria* (5 spp.) yielded DNA that was too degraded to be sequenced, possibly due to the circumstances under which the specimens were dried (Jankowiak *et al.*, 2005). None of these taxa possess unambiguous synapomorphies to support their placement in the current phylogeny. Linder *et al.* (1997) reported morphological and anatomical phylogenetic support for the placement of *Leptagrostis*, *Piptophyllum* and *Zenkeria* in a so-called "*Crinipes* group", but this hypothesized clade is contradicted by the current study. The spikelets of *Danthonidium* have lemmas with features similar to *Dregeochloa* and *Monachather*, including a geniculate awn, many veins, and hairs in tufts and transverse rows. However, these traits are also shared with several taxa in subfamily Danthonioideae. Soreng *et al.* (2015) treat *Danthonidium* as *incertae sedis* in this subfamily along with *Alloochaete* and *Phaenanthoecium*, which are recovered in very different clades in our analysis.

1.4.6 Conclusion

This study represents the first evolutionary analysis of spikelet burial characteristics across the PACMAD grasses as well as the largest full plastome phylogenetic study of the grass family conducted to date, including the most complete sampling of subfamily Arundinoideae. Resolving the polyphyly of this poorly-studied subfamily is shown to have substantial implications for ancestral trait estimations across the PACMAD. A straight-awned lemma with a hairy callus is found to be the most likely and parsimonious ancestral state for the clade, whose members have experienced multiple independent losses of these features as well as similar gains of a more active burial syndrome as indicated by a hygroscopic, geniculate awn. Passive burial, indicated by a loss of diaspore awns, is loosely concordant with clades having higher diversification rates, though this result requires more detailed taxon sampling to test rigorously.

Two clades, tribe Andropogoneae in subfamily Panicoideae and tribe Arundineae in subfamily Arundinoideae, stand out as particularly promising for future in-depth studies of burial syndrome evolution and its effect on the ecology and biogeography of grasses.

LITERATURE CITED

- Ansong, M. and C. Pickering. 2013. Long-distance dispersal of black spear grass (*Heteropogon contortus*) seed on socks and trouser legs by walkers in Kakadu National Park. *Ecological Management & Restoration* 14(1): 71–74.
- Bankevich, A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, et al. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19(5): 455–77.
- Barker, N. P., Linder, H. P., and E. H. Harley. 1995. Polyphyly of Arundinoideae (Poaceae): evidence from rbcL sequence data. *Systematic Botany* 20(4): 423–435.
- Barker, N. P., H. P. Linder, E.H. Harley, C. Town, P. Bag, & M. Lavin. 1998. Sequences of the grass-specific insert in the chloroplast rpoC2 gene elucidate generic relationships of the Arundinoideae (Poaceae). *Systematic Botany* 23(3): 327–350.
- Beaulieu, J. M., B. C. O’Meara, and M. J. Donoghue. 2013. Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. *Systematic Biology* 62(5): 725–37.
- Besnard, G., P.-A. Christin, P.-J. G. Malé, E. Lhuillier, C. Lauzeral, E. Coissac, and M. S. Vorontsova. 2014. From museums to genomics: old herbarium specimens shed light on a C₃ to C₄ Transition. *Journal of Experimental Botany* 65(22): 6711–21.
- Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15): 2114–20.
- Bouchenak-Khelladi, Y., G. A. Verboom, V. Savolainen, and T.R. Hodkinson. 2010. Biogeography of the grasses (Poaceae): a phylogenetic approach to reveal evolutionary history in geographical space and geological time. *Botanical Journal of the Linnean Society* 162(4): 543–57.
- Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* 17: 540–552.
- Christin, P.-A., E. Spriggs, C. P. Osborne, C. A. E. Strömberg, N. Salamin, and E. J. Edwards. 2014. Molecular dating, evolutionary rates, and the age of the grasses. *Systematic Biology* 63(2): 153–65.
- Clark, L. G., W. Zhang, and J. F. Wendel. 1995. A phylogeny of the grass family (Poaceae) Based on ndhF Sequence Data. *Systematic Botany* 20(4): 436–460.
- Clayton, W. D., and S. A. Renvoize. 1999. *Genera Graminum*. London: Royal Botanic Gardens, Kew.

- Conert, H.J. 1987. Current concepts in the systematics of the Arundinoideae. Pp. 239-250 in *Grass Systematics and Evolution*, eds. T.R. Soderstrom, K.W. Hilu, C.S. Campbell, and M.E. Barkworth. Washington, DC: Smithsonian Institution Press.
- Cota-Sánchez, J. H., K. Remarchuk, and K. Ubayasena. 2006. Ready-to-use DNA extracted with a CTAB method adapted for herbarium specimens and mucilaginous plant tissue. *Plant Molecular Biology Reporter* 24(2): 161–67.
- Cotton, J. L., W. P. Wysocki, L. G. Clark, S. A. Kelchner, J. C. Pires, P. P. Edger, D. Mayfield-Jones, and M. R. Duvall. 2015. Resolving deep relationships of PACMAD grasses: a phylogenomic approach. *BMC Plant Biology* 15(1): 1–11.
- Cousens, R., R. Law, and C. Dytham. 2008. *Dispersal in Plants: A Population Perspective*. Oxford: OUP Oxford.
- Crepet, W. L. and G. D. Feldman. 1991. The earliest remains of grasses in the fossil record. *American Journal of Botany* 78: 1010–1014.
- Davidse, G. 1987. Fruit dispersal in the Poaceae. Pp. 143-155 in *Grass Systematics and Evolution*, eds. T.R. Soderstrom, K.W. Hilu, C.S. Campbell, and M.E. Barkworth. Washington, DC: Smithsonian Institution Press.
- Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29: 1969–1973.
- Elbaum, R., L. Zaltzman, I. Burgert, P. Fratzl, L. Zaltzman, and I. Burgert. 2007. The role of wheat awns in the seed dispersal unit. *Science* 316(5826): 884–886.
- Elias, M. K. 1942. Tertiary prairie grasses and other herbs from the High Plains. *Geological Society of America Special Paper (Regular Studies)* 41: 1-176.
- Estep, M. C., M. R. McKain, D. V. Diaz, J. Zhong, J. G. Hodge, T. R. Hodkinson, D. J. Layton, S. T. Malcomber, *et al.* 2014. Allopolyploidy, diversification, and the Miocene grassland expansion. *PNAS* 111 (42): 15149–15154.
- Garnier, L. K. M., and I. Dajoz. 2001. Evolutionary significance of awn length variation in a clonal grass of fire-prone savannas. *Ecology* 82(6): 1720–33.
- Grass Phylogeny Working Group II. 2011. New grass phylogeny resolves deep evolutionary relationships and discovers C₄ origins. *The New Phytologist* 193 (2): 304–312.
- Green, D. S. 1980. The terminal velocity and dispersal of spinning samaras. *American Journal of Botany* 67(8): 1218–1224.
- Hamrick, J. L., and J. M. Lee. 1987. Effect of soil surface topography and litter cover on the germination, survival, and growth of musk thistle (*Carduus nutans*). *American Journal of*

- Botany* 74(3): 451.
- Harper, J. L., J. T. Williams, and G. R. Sagar. 1965. The behaviour of seeds in soil: I. the heterogeneity of soil surfaces and its role in determining the establishment of plants from seed. *The Journal of Ecology* 53(2): 273.
- Hayashi, M., K. L. Feilich, and D. J. Ellerby. 2009. The mechanics of explosive seed dispersal in orange jewelweed (*Impatiens capensis*). *Journal of Experimental Botany* 60(7): 2045–2053.
- Hubbard, C.E. 1957a. Notes on African grasses: XXV. *Nematopoa*, a new genus from Southern Rhodesia. *Kew Bulletin* 12(1): 51-52.
- Hubbard, C.E. 1957b. Notes on African grasses: XXVI. *Piptophyllum*, a new genus from Angola. *Kew Bulletin* 12(1): 52-53.
- Humphreys, A. M., A. Antonelli, M. D. Pirie, and H. P. Linder. 2010. Ecology and evolution of the diaspore burial syndrome. *Evolution* 65(4): 1163–1180.
- Ingram, A. L., P. Christin, and C. P. Osborne. 2011. Molecular phylogenies disprove a hypothesized C₄ reversion in *Eragrostis walteri* (Poaceae). *Annals of Botany* 321–325.
- Jankowiak, K., K. Buczkowska, and Z. Szweykowska-Kulinska. 2005. Successful extraction of DNA from 100-year-old herbarium specimens of the liverwort *Bazzania trilobata*. *Taxon* 54(2): 335–36.
- Johnson, E.E., and Z. Baruch. 2013. Awn length variation and its effect on dispersal unit burial of *Trachypogon spicatus* (Poaceae). *Revista de Biología Tropical/International Journal of Tropical Biology and Conservation* 62: 321–326.
- Katoh, S. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30: 772-780.
- Kellogg, E. A. 2015. Flowering Plants. Vol. 13. Monocots: Poaceae. In *The Families and Genera of Vascular Plants Series*, ed. K. Kubitzki. Heidelberg, Germany: Springer.
- Kiviniemi, K. 1996. A study of adhesive seed dispersal of three species under natural conditions. *Acta Botanica Neerlandica* 45(1): 73–83.
- Kulić, I. M., M. Mani, H. Mohrbach, R. Thaokar, and L. Mahadevan. 2009. Botanical ratchets. *Proceedings of the Royal Society: Biological Sciences* 276(1665): 2243–47.
- Linder, H.P., G. A. Verboom, and N.P. Barker. 1997. Phylogeny and evolution in the *Crinipes* group of grasses (Arundinoideae: Poaceae). *Kew Bulletin* 52(1): 91-110.
- Maddison, W. P., and R. G. FitzJohn. 2015. The unsolved challenge to phylogenetic correlation

- tests for categorical characters. *Systematic Biology* 64(1): 127–36.
- Maddison, W. P. and D.R. Maddison. 2015. Mesquite: a modular system for evolutionary analysis. Version 3.04 <http://mesquiteproject.org>
- Miller M. A., W. Pfeiffer, T. Schwartz. 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. In: SC10 Workshop on Gateway Computing Environments (GCE10).
- Morris, L. M., and M. R. Duvall. 2010. The chloroplast genome of *Anomochloa marantoidea* (Anomochlooideae; Poaceae) comprises a mixture of grass-like and unique features. *American Journal of Botany* 97(4): 620–27.
- Pagel, M. 1994. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society: Biological Sciences* 255(1342): 37–45.
- Peart, M. H. 1979. Experiments on the biological significance of the morphology of seed dispersal units in grasses. *Journal of Ecology* 67(3): 843–863.
- Peart, M.H. 1981. Further experiments on the biological significance of the morphology of seed-dispersal units in grasses. *Journal of Ecology* 69(2): 425-436.
- Peart, M. H. 1984. The effects of morphology, orientation and position of grass diaspores on seedling survival. *Journal of Ecology* 72(2): 437–453.
- Peart, M. H., and H. T. Clifford. 1987. The influence of diaspore morphology and soil-surface properties on the distribution of grasses. *Journal of Ecology* 75(2):569-576.
- Popescu, A.-A., K. T. Huber, and E. Paradis. 2012. Ape 3.0: new tools for distance-based phylogenetics and evolutionary analysis in R. *Bioinformatics* 28(11): 1536–37.
- R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Rabosky, D. L. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* 9(2): e89543.
- Rabosky, D. L., M. Grudler, C. Anderson, P. Title, J. J. Shi, J. W. Brown, H. Huang, and J. G. Larson. 2014. BAMMtools: An R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution* 5(7): 701–7.
- Renvoize, S.A. 1981. The sub-family Arundinoideae and its position in relation to a general classification of the Gramineae. *Kew Bulletin* 36(1): 85-102.
- Revell, L. J. 2012. phytools: An R package for phylogenetic comparative biology (and other

- things). *Methods Ecol. Evol.* 3: 217-223.
- Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M. A. Suchard, and J. P. Huelsenbeck. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61(3): 539–42.
- Särkinen, T., M. Staats, J. E. Richardson, R. S. Cowan, F. T. Bakker, and D. Caramelli. 2012. How to open the treasure chest? Optimising DNA extraction from herbarium specimens. *PLoS ONE* 7(8): 1–9.
- Schöning, C., X. Espadaler, I. Hensen, and F. Roces. 2004. Seed predation of the tussock-grass *Stipa Tenacissima* L. by ants (*Messor spp.*) in South-Eastern Spain: the adaptive value of trypanocarpus. *Journal of Arid Environments* 56(1): 43–61.
- Sheldon, J. C. 1974. The behaviour of seeds in soil: III. The influence of seed morphology and the behaviour of seedlings on the establishment of plants from surface-lying seeds. *The Journal of Ecology* 62(1): 47.
- Shimodaira, H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution* 16(8): 1114.
- Soreng, R. J., P. M. Peterson, K. Romaschenko, G. Davidse, F. O. Zuloaga, E. J. Judziewicz, T. S. Filgueiras, et al. 2015. A worldwide phylogenetic classification of the Poaceae (Gramineae): phylogenetic classification of the grasses. *Journal of Systematics and Evolution* 53(2): 117–37.
- Sorensen, A. E. 1986. Seed dispersal by adhesion. *Annual Review of Ecology and Systematics* 17: 443–63.
- Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9): 1312–1313.
- Strömberg, C. A. E. 2005. Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. *Proceedings of the National Academy of Sciences USA* 102: 11980–11984.
- Van der Pijl, L. 1982. *Principles of Dispersal in Higher Plants*. Berlin: Springer-Verlag.
- Vicentini, A., J. C. Barber, S. S. Aliscioni, L. M. Giussani, and E. A. Kellogg. 2008. The age of the grasses and clusters of origins of C₄ photosynthesis. *Global Change Biology* 14(12): 2963–2977.
- Washburn, J. D., J. C. Schnable, G. Davidse, and J. C. Pires. 2015. Phylogeny and photosynthesis of the grass tribe Paniceae. *American Journal of Botany* 102: 1493–1505.

Watson, L., and Dallwitz, M.J. 1992 onwards. The grass genera of the world: descriptions, illustrations, identification, and information retrieval; including synonyms, morphology, anatomy, physiology, phytochemistry, cytology, classification, pathogens, world and local distribution, and references. Version: 7th December 2015. <http://delta-intkey.com>.

Wolgemuth, C. W. 2009. Plant biomechanics: using shape to steal motion. *Current Biology* 19(10): R409–410.

FIGURES AND TABLES

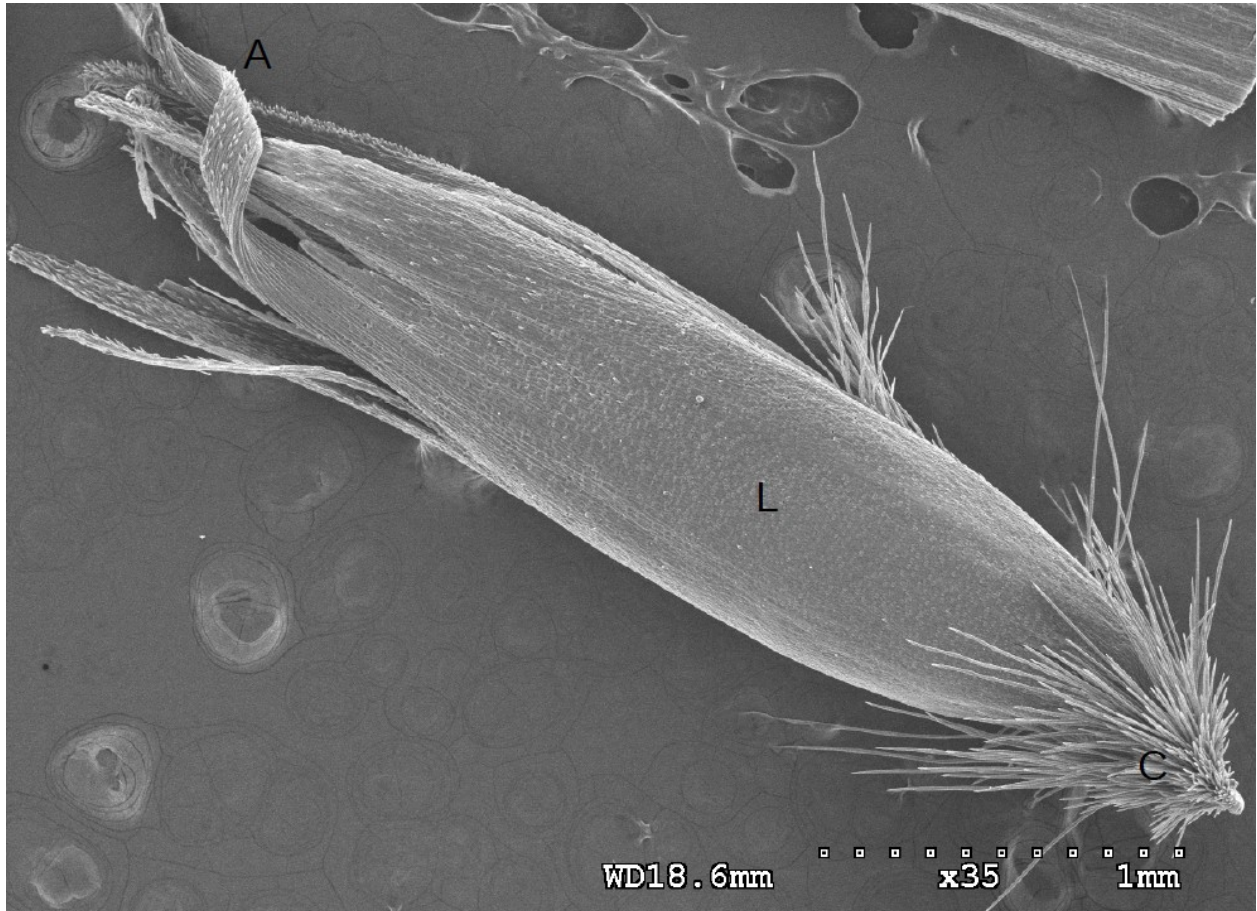


Figure 1.1. SEM image of floret/diaspore from *Alloochaete gracillima* positioned with the back of the lemma facing upwards. Letters in figure are as follows: C – callus, hairy and slightly pointed in this species; L – lemma, with tufts of hairs on either margin approximately 1/3 of the length of the lemma from the callus; A – awn, which in this species is twisted.

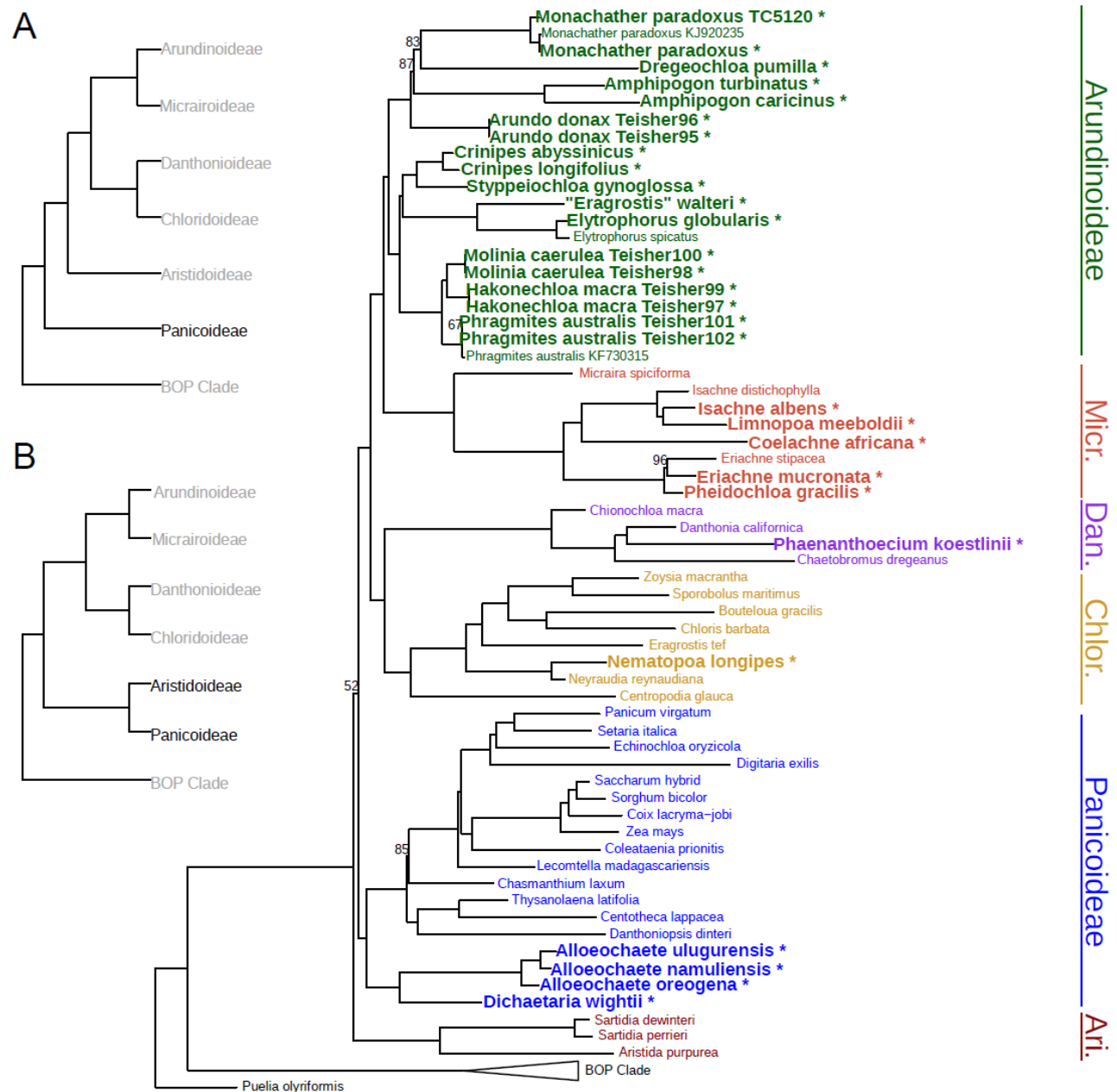


Figure 1.2. Maximum likelihood phylogeny of untrimmed alignment of 88 full plastomes, with *Anomochloa* and *Pharus* removed and members of the BOP clade collapsed for clarity. Bootstrap values below 100 are shown above nodes. Subfamilies in PACMAD are grouped by color. Samples in bold with asterisks were generated for the current study. Alternative topologies that cannot be rejected with a SH-test are shown in A and B inserts.

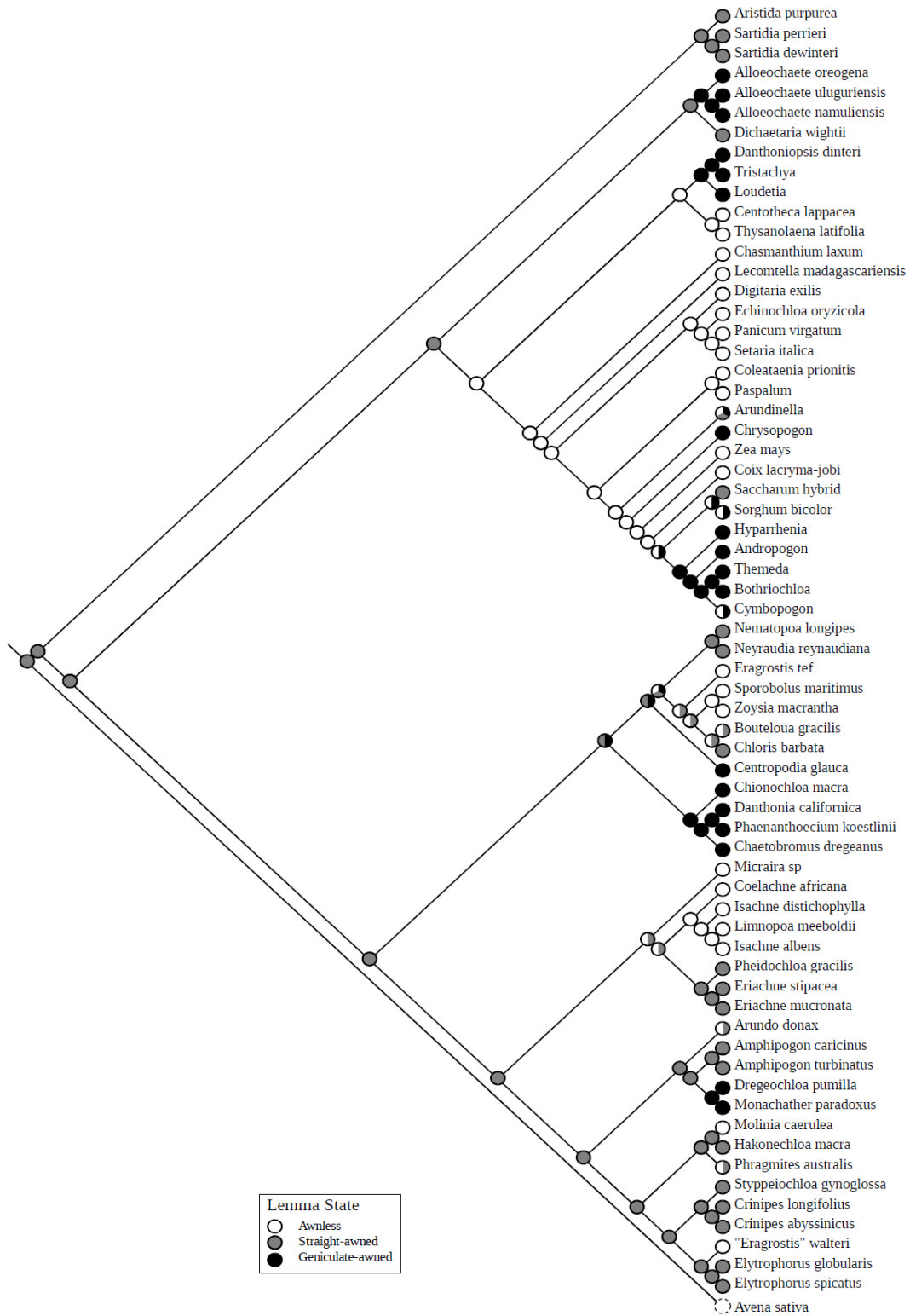


Figure 1.3. Parsimony ancestral states of awn presence and type in PACMAD. Data are coded as missing for outgroup *Avena sativa*.

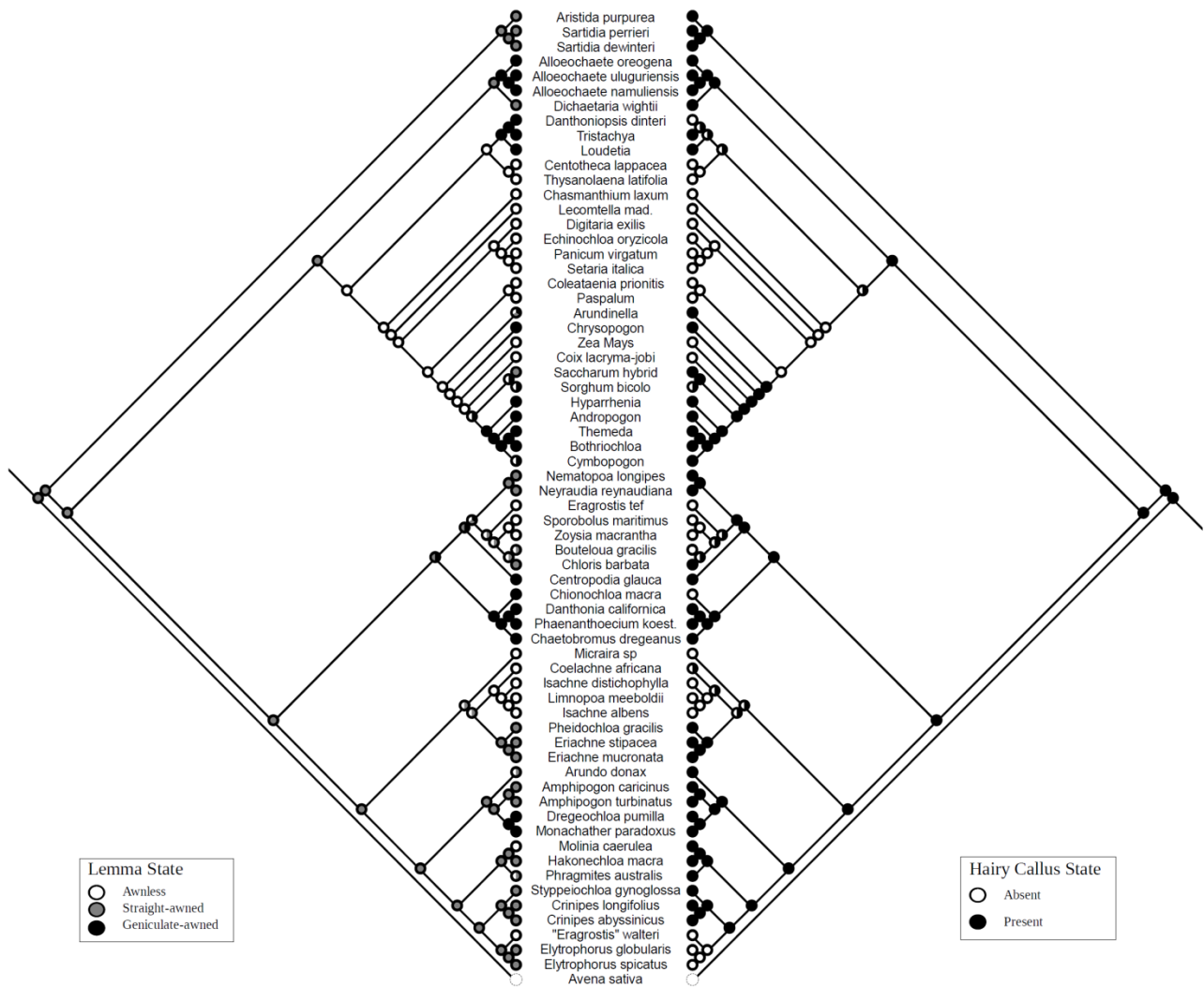


Figure 1.4. Parsimony ancestral states of awn presence/type vs. hairy callus presence in PACMAD. Missing data are coded by a dashed empty circle.

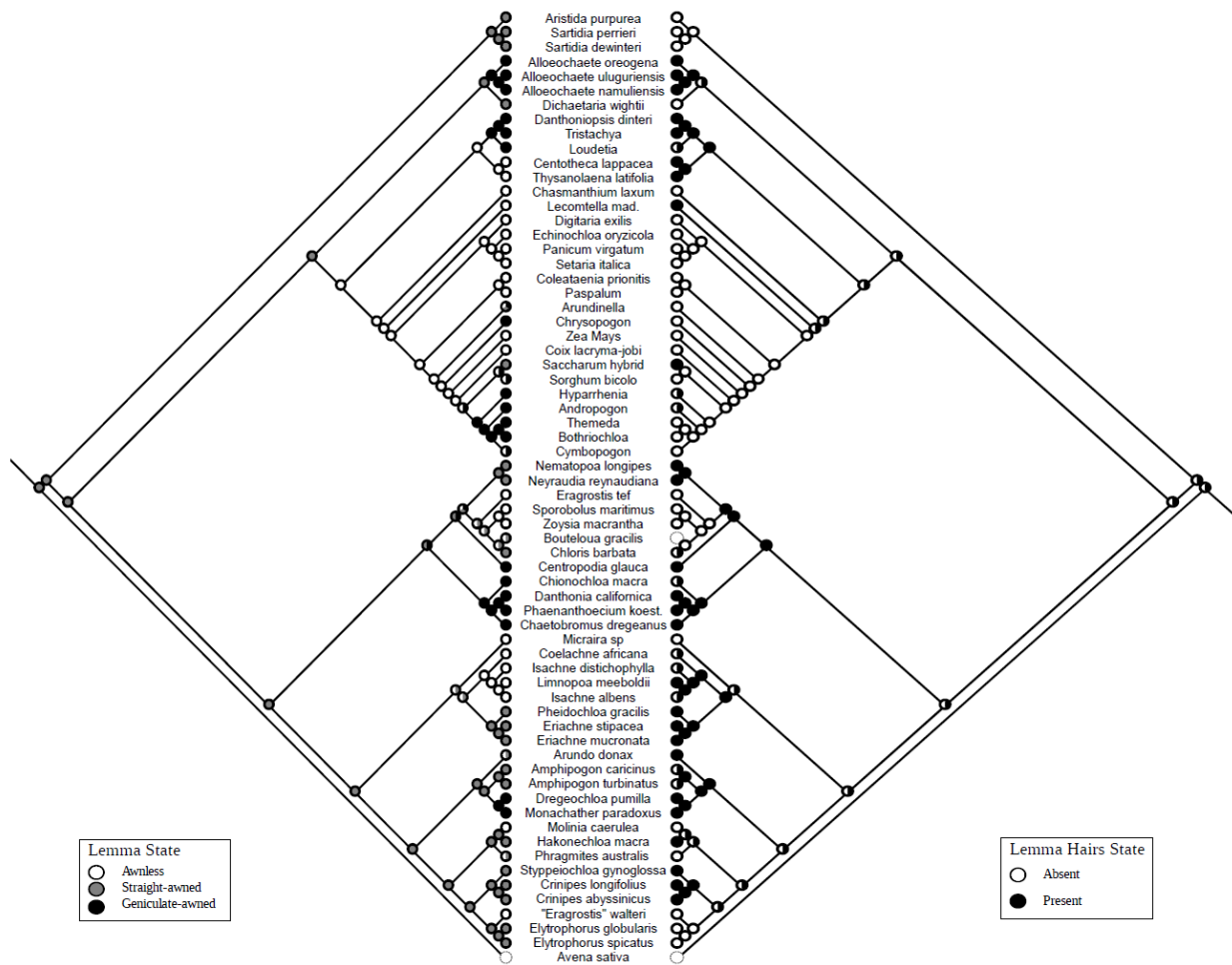


Figure 1.5. Parsimony ancestral states of awn presence/type vs. lemma body hair presence in PACMAD. Missing data are coded by a dashed empty circle.

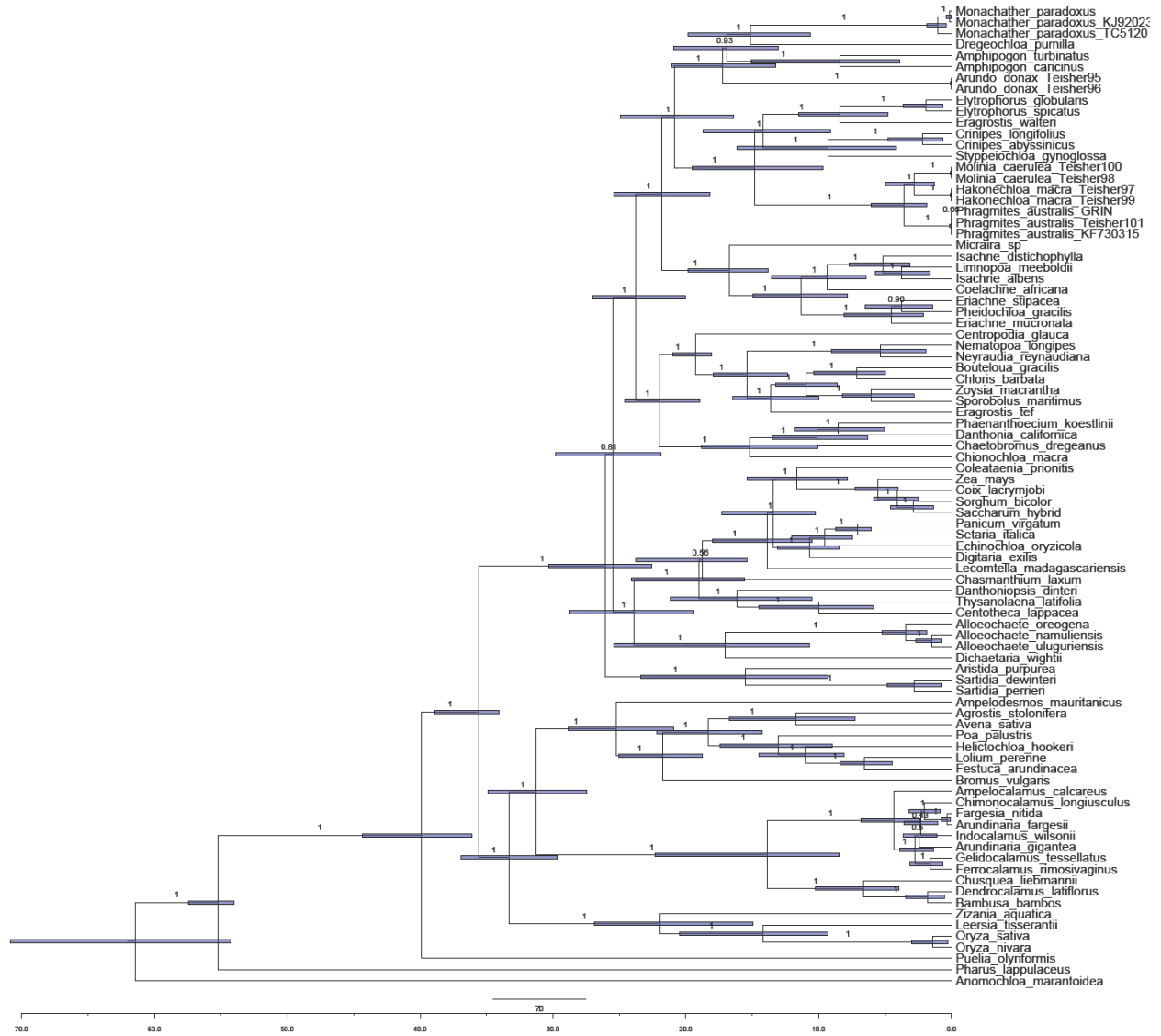


Figure 1.6. BEAST optimal ultrametric, dated phylogeny based on gapless plastome alignment and four fossil calibrations. 95% HPD ranges are depicted by bars above nodes. Numbers above branches are posterior probabilities.



Figure 1.7. The nine highest posterior probability shift sets identified by BAMM.

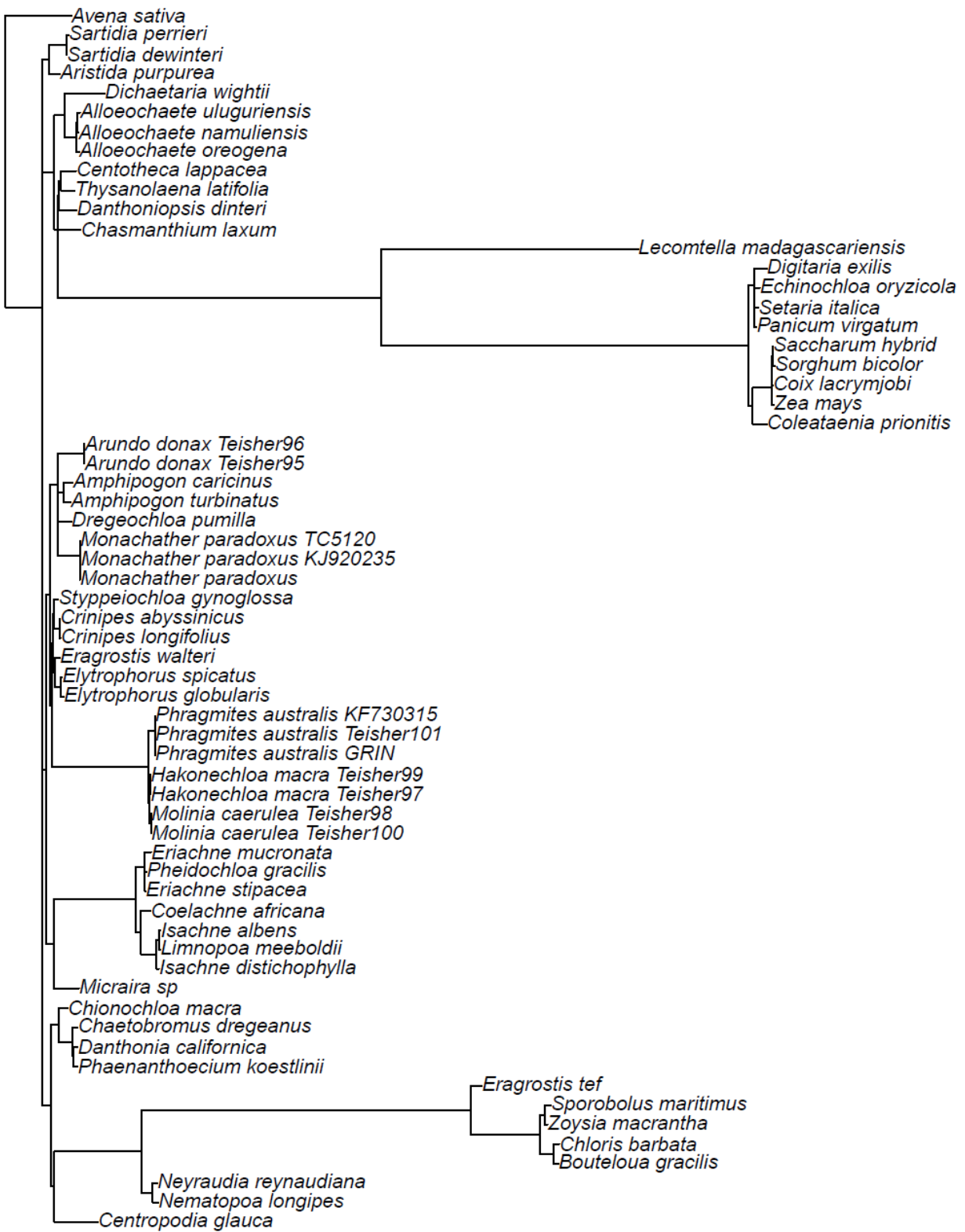


Figure 1.8. ML topology of PACMAD clade with branch lengths proportional to the frequency of diversification rate shifts across posterior distribution in BAMM analysis.

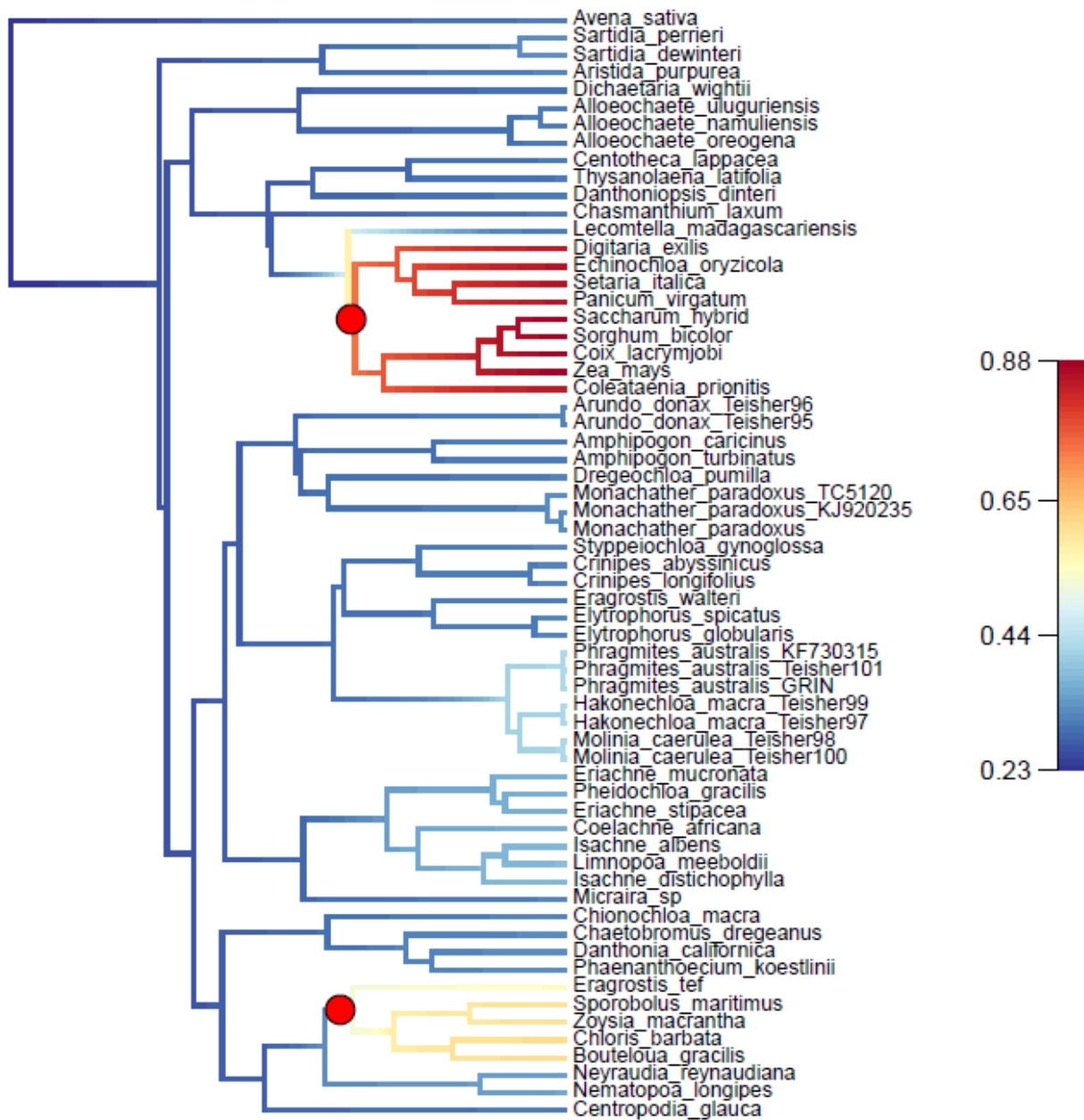


Figure 1.9. Diversification rate shift set with the highest posterior probability (16%) found in BAMM. Branches are colored according to inferred diversification rate, and two rate shifts are identified by red circles.

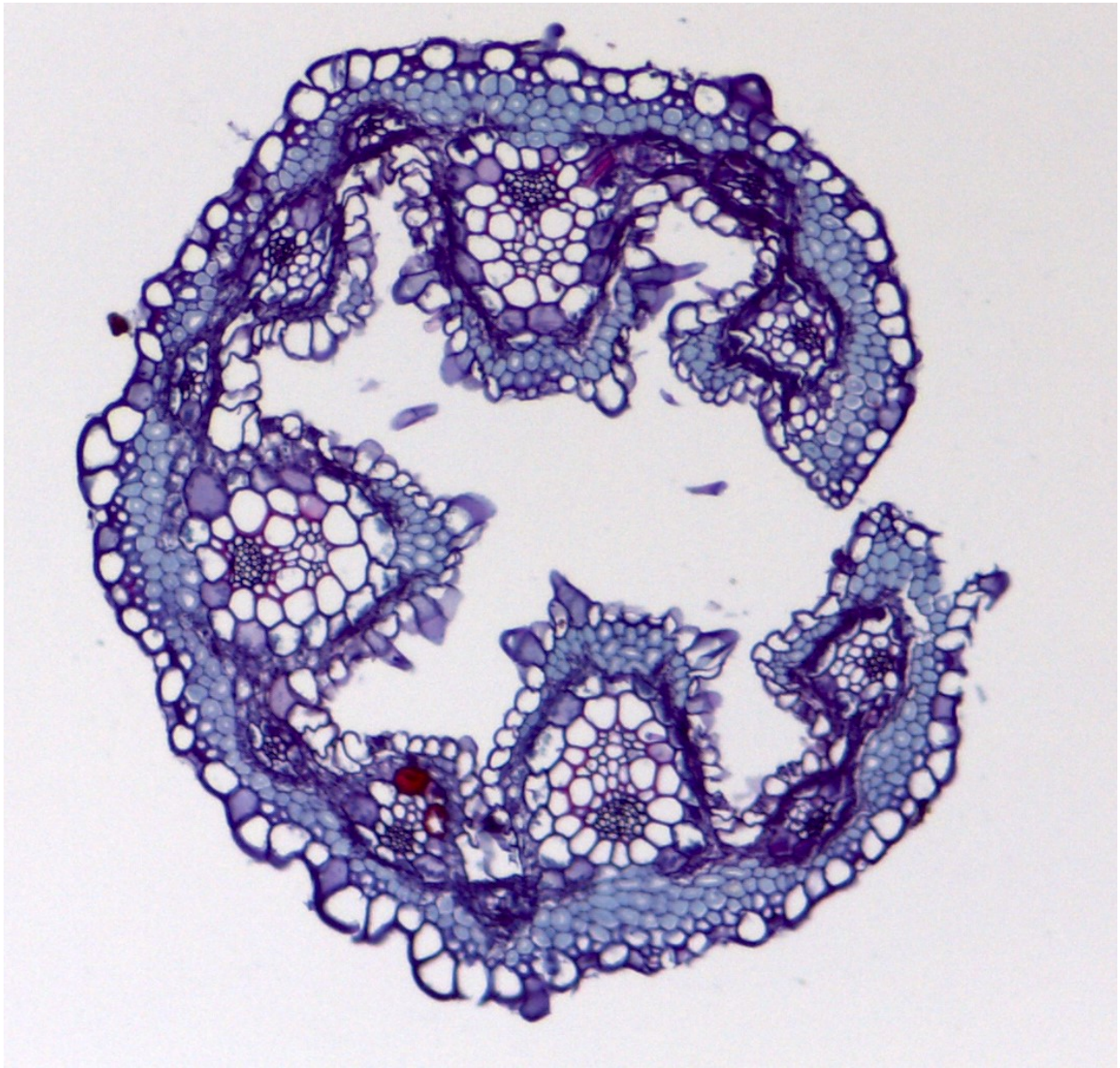


Figure 1.10. Cross-section of *Nematopoa longipes* taken from herbarium material and stained with Safranin-Fast Green.

Table 1.1. Alternative alignments and their effects on the placement of Aristidoideae.

Alignment	Length	Topology Supported
Unedited	156,968	Aristidoideae sister to PCMAD
Gblocks All Gaps	120,422	Aristidoideae+Panicoideae sister to (CMAD)
Gblocks Half Gaps	103,233	Aristidoideae+Panicoideae sister to (CMAD)
Gblocks No Gaps	79,909	Aristidoideae sister to PCMAD

Table 1.2. Ages of Select Clades from BEAST Analysis.

Clade	Age (mya)	95% HPD	
		Lower	Higher
BOP+PACMAD	35.59	34.04	38.9
Crown PACMAD	26.05	22.55	30.3
Panicoideae + CMAD	25.51	21.86	29.78
Arundinoideae + Micrairoideae	21.83	18.16	25.42
Crown Arundinoideae	20.86	16.42	24.92
Crown Micrairoideae	16.76	13.83	19.86
Crown Danthonioideae	15.19	10.03	18.84
Crown Chloridoideae	19.27	18.05	20.99
Crown Panicoideae	23.89	19.4	28.71
Panicoideae s.s.	19.01	15.59	24.07
<i>Dichaetaria+Alloeochoete</i>	17.04	10.7	25.41
Crown Aristidoideae	15.54	9.28	23.37

Table S1.1. Whole-plastome samples used in the phylogenetic analysis, with assembly statistics for plastomes generated in the current study.

Taxon	Voucher	Source	LSC Length	SSC Length	IR Length	Assembly Length	LSC Coverage	SSC Coverage	IR Coverage	GenBank Accession No.
<i>Alloochaete namuliensis</i>	Harris 192	This Study	81420	12246	22762	139,190	31	31	61	-
<i>Alloochaete oreogena</i>	Chapman 19	This Study	80953	12330	21398	136,079	36	33	84	-
<i>Alloochaete uluguriensis</i>	Abeid 3692	This Study	81520	12351	22761	139,395	171	189	335	-
<i>Amphipogon caricinus</i>	Latz 12610	This Study	81102	12640	21126	135,994	163	193	248	-
<i>Amphipogon turbinatus</i>	Kellogg 1027	This Study	80603	12759	21138	135,638	35	29	78	-
<i>Arundo donax</i>	Teisher 95	This Study	82061	12623	21244	137,172	452	648	772	-
<i>Arundo donax</i>	Teisher 96	This Study	82057	12623	21246	137,172	165	229	266	-
<i>Coelachne africana</i>	Thomas 3935	This Study	80171	12732	21112	135,127	73	60	126	-
<i>Crinipes abyssinicus</i>	De Wilde 206	This Study	82707	12636	21283	137,909	185	229	311	-
<i>Crinipes longifolius</i>	Gilbert 1009	This Study	82781	12679	21285	138,030	109	105	212	-
<i>Dichaetaria wightii</i>	Soderstrom 1739	This Study	82486	12666	20048	135,248	36	32	74	-
<i>Dregeochloa pumilla</i>	Barker	This Study	81441	12725	21290	136,746	68	38	346	-
<i>Elytrophorus globularis</i>	Ingram 692	This Study	81286	12536	21290	136,402	249	238	598	-
<i>Eragrostis walteri</i>	Ingram 656	This Study	81766	12691	21300	137,057	181	201	364	-
<i>Eriachne mucronata</i>	Latz 13498	This Study	80148	12591	21188	135,115	129	142	221	-
<i>Hakonechloa macra</i>	Teisher 97	This Study	82662	12702	21284	137,932	281	347	478	-
<i>Hakonechloa macra</i>	Teisher 99	This Study	82661	12700	21285	137,931	188	244	321	-
<i>Isachne albens</i>	Lammers 8452	This Study	81078	12663	21155	136,051	113	111	209	-
<i>Limnopoa meeboldii</i>	Cook 282	This Study	79897	12663	21082	134,724	88	89	160	-
<i>Molinia caerulea</i>	Teisher 98	This Study	82320	12739	21275	137,609	45	43	107	-
<i>Molinia caerulea</i>	Teisher 100	This Study	82333	12735	21277	137,622	71	70	158	-
<i>Monachather paradoxus</i>	Columbus 5120	This Study	82418	12653	21225	137,521	42	39	100	-
<i>Monachather paradoxus</i>	Lazarides 149	This Study	82566	12626	21230	137,652	367	398	696	-
<i>Nematopoa longipes</i>	Simon 2353	This Study	79885	12541	21016	134,458	120	118	268	-
<i>Phaenanthoecium koestlinii</i>	Wood 1305	This Study	79379	12398	20775	133,327	168	182	314	-
<i>Pheidochloa gracilis</i>	Sharp 391	This Study	80081	12604	21147	134,979	37	37	76	-
<i>Phragmites australis</i>	Teisher 101	This Study	82449	12697	21269	137,688	64	60	174	-
<i>Phragmites australis</i>	Teisher 102	This Study	82461	12691	21276	137,704	125	116	328	-
<i>Styppeichloa gynoglossa</i>	Hilliard 18774	This Study	82307	12605	21272	137,456	188	185	332	-
<i>Agrostis stolonifera</i>	-	GenBank	-	-	-	-	-	-	-	EF115543
<i>Ampelocalamus calcareus</i>	-	GenBank	-	-	-	-	-	-	-	KJ496369
<i>Ampelodesmos mauritanicus</i>	-	GenBank	-	-	-	-	-	-	-	KM974731
<i>Anomochloa marantoidea</i>	-	GenBank	-	-	-	-	-	-	-	GQ329703
<i>Aristida purpurea</i>	-	GenBank	-	-	-	-	-	-	-	KJ920224
<i>Arundinaria fargesii</i>	-	GenBank	-	-	-	-	-	-	-	JX513413
<i>Arundinaria gigantea</i>	-	GenBank	-	-	-	-	-	-	-	JX235347
<i>Avena sativa</i>	-	GenBank	-	-	-	-	-	-	-	KM974733
<i>Bambusa bambos</i>	-	GenBank	-	-	-	-	-	-	-	KJ870988

<i>Bouteloua gracilis</i>	-	Duvall Lab	-	-	-	-	-	-	-	KT168386
<i>Bromus vulgaris</i>	-	GenBank	-	-	-	-	-	-	-	KM974737
<i>Centotheca lappacea</i>	-	GenBank	-	-	-	-	-	-	-	KJ920225
<i>Centropodia glauca</i>	-	Duvall Lab	-	-	-	-	-	-	-	KT168383
<i>Chaetobromus dregeanus</i>	-	GenBank	-	-	-	-	-	-	-	KJ920226
<i>Chasmothium laxum</i>	-	Kellogg Lab	-	-	-	-	-	-	-	-
<i>Chimonocalamus longiusculus</i>	-	GenBank	-	-	-	-	-	-	-	JX513415
<i>Chionochloa macra</i>	-	GenBank	-	-	-	-	-	-	-	KJ920227
<i>Chloris barbata</i>	-	Duvall Lab	-	-	-	-	-	-	-	KT168393
<i>Chusquea liebmannii</i>	-	GenBank	-	-	-	-	-	-	-	KJ871001
<i>Coix lacrym-jobi</i>	-	GenBank	-	-	-	-	-	-	-	FJ261955
<i>Coleataenia prionitis</i>	-	GenBank	-	-	-	-	-	-	-	KJ920228
<i>Danthonia californica</i>	-	GenBank	-	-	-	-	-	-	-	KJ920229
<i>Danthoniopsis dinteri</i>	-	Washburn et al. (2015)	-	-	-	-	-	-	-	-
<i>Dendrocalamus latiflorus</i>	-	GenBank	-	-	-	-	-	-	-	FJ970916
<i>Digitaria exilis</i>	-	GenBank	-	-	-	-	-	-	-	NC024176
<i>Echinochloa oryzicola</i>	-	GenBank	-	-	-	-	-	-	-	KJ000048
<i>Elytrophorus spicatus</i>	-	GenBank	-	-	-	-	-	-	-	KJ920230
<i>Eragrostis tef</i>	-	GenBank	-	-	-	-	-	-	-	KT168385
<i>Eriachne stipacea</i>	-	GenBank	-	-	-	-	-	-	-	KJ920231
<i>Fargesia nitida</i>	-	GenBank	-	-	-	-	-	-	-	JX513416
<i>Ferocalamus rimosivaginus</i>	-	GenBank	-	-	-	-	-	-	-	HQ337794
<i>Festuca arundinacea</i>	-	GenBank	-	-	-	-	-	-	-	FJ466687.2
<i>Gelidocalamus tessellatus</i>	-	GenBank	-	-	-	-	-	-	-	JX513420
<i>Helictochloa hookeri</i>	-	GenBank	-	-	-	-	-	-	-	KM974734
<i>Indocalamus wilsonii</i>	-	GenBank	-	-	-	-	-	-	-	JX513421
<i>Isachne distichophylla</i>	-	GenBank	-	-	-	-	-	-	-	NC_025236
<i>Lecomtella madagascariensis</i>	-	GenBank	-	-	-	-	-	-	-	HF543599
<i>Leersia tisserantii</i>	-	GenBank	-	-	-	-	-	-	-	JN415112
<i>Lolium perenne</i>	-	GenBank	-	-	-	-	-	-	-	AM777385
<i>Micraira sp.</i>	-	GenBank	-	-	-	-	-	-	-	KJ920234
<i>Monachather paradoxus</i>	-	GenBank	-	-	-	-	-	-	-	KJ920235
<i>Neyraudia reynaudiana</i>	-	GenBank	-	-	-	-	-	-	-	KF356382
<i>Oryza nivara</i>	-	GenBank	-	-	-	-	-	-	-	NC005973
<i>Oryza sativa</i>	-	GenBank	-	-	-	-	-	-	-	NC001320
<i>Panicum virgatum</i>	-	GenBank	-	-	-	-	-	-	-	HQ822121
<i>Pharus lapullaceus</i>	-	GenBank	-	-	-	-	-	-	-	KC311467
<i>Phragmites australis</i>	-	GenBank	-	-	-	-	-	-	-	KF730315
<i>Poa palustris</i>	-	GenBank	-	-	-	-	-	-	-	KM974749
<i>Puella olyrifformis</i>	-	GenBank	-	-	-	-	-	-	-	KC534841
<i>Saccharum hybrid</i>	-	GenBank	-	-	-	-	-	-	-	AP006714

<i>Sartidia dewinteri</i>	-	GenBank	-	-	-	-	-	-	-	KJ819550
<i>Sartidia perrieri</i>	-	GenBank	-	-	-	-	-	-	-	KJ819549
<i>Setaria italic</i>	-	GenBank	-	-	-	-	-	-	-	NC022850
<i>Sorghum bicolor</i>	-	GenBank	-	-	-	-	-	-	-	EF115542
<i>Sporobolus maritimus</i>	-	GenBank	-	-	-	-	-	-	-	KP176438
<i>Thysanolaena latifolia</i>	-	GenBank	-	-	-	-	-	-	-	KJ920236
<i>Zea mays</i>	-	GenBank	-	-	-	-	-	-	-	X86563
<i>Zizania aquatica</i>	-	GenBank	-	-	-	-	-	-	-	KJ870999
<i>Zoysia macrantha</i>	-	Duvall Lab	-	-	-	-	-	-	-	KT168390

Table S1.2. Sampling fractions used in the BAMM analysis, using Kellogg (2015) as a guide to total species numbers in each tribe of PACMAD.

Tree Sample	Tribe	Fraction of Species	Tree Sample	Tribe	Fraction of Species
<i>Aristida_purpurea</i>	Aristidoideae	0.0082	<i>Centropodia_glauca</i>	Centropodieae	0.2
<i>Sartidia_perrieri</i>	Aristidoideae	0.0082	<i>Neyraudia_reynaudiana</i>	Triraphideae	0.1429
<i>Sartidia_dewinteri</i>	Aristidoideae	0.0082	<i>Nematopoa_longipes</i>	Triraphideae	0.1429
<i>Pheidochloa_gracilis</i>	Eriachneae	0.06	<i>Eragrostis_tef</i>	Eragrostideae	0.002
<i>Eriachne_mucronata</i>	Eriachneae	0.06	<i>Chloris_barbata</i>	Cynodonteae	0.0022
<i>Eriachne_stipacea</i>	Eriachneae	0.06	<i>Bouteloua_gracilis</i>	Cynodonteae	0.0022
<i>Coelachne_africana</i>	Isachneae	0.0339	<i>Sporobolus_maritimus</i>	Zoysieae	0.0085
<i>Limnopoa_meeboldii</i>	Isachneae	0.0339	<i>Zoysia_macrantha</i>	Zoysieae	0.0085
<i>Isachne_albens</i>	Isachneae	0.0339	<i>Chaetobromus_dregeanus</i>	Danthonioideae	0.0142
<i>Isachne_distichophylla</i>	Isachneae	0.0339	<i>Phaenanthoecium_koestlinii</i>	Danthonioideae	0.0142
<i>Micraira_sp</i>	Micraireae	0.0667	<i>Danthonia_californica</i>	Danthonioideae	0.0142
<i>Phragmites_australis_KF730315</i>	Arundinoideae	0.3784	<i>Chionochloa_macra</i>	Danthonioideae	0.0142
<i>Phragmites_australis_GRIN</i>	Arundinoideae	0.3784	<i>Dichaetaria_wightii</i>	Dichaetarieae	0.6667
<i>Phragmites_australis_Teisher101</i>	Arundinoideae	0.3784	<i>Alloeochaete_oreogena</i>	Dichaetarieae	0.6667
<i>Hakonechloa_macra_Teisher97</i>	Arundinoideae	0.3784	<i>Alloeochaete_namuliensis</i>	Dichaetarieae	0.6667
<i>Hakonechloa_macra_Teisher99</i>	Arundinoideae	0.3784	<i>Alloeochaete_uluguriensis</i>	Dichaetarieae	0.6667
<i>Molinia_caerulea_Teisher98</i>	Arundinoideae	0.3784	<i>Danthoniopsis_dinteri</i>	Tristachyideae	0.0112
<i>Molinia_caerulea_Teisher100</i>	Arundinoideae	0.3784	<i>Centotheca_lappacea</i>	Centothecaeae	0.25
<i>Elytrophorus_spicatus</i>	Arundinoideae	0.3784	<i>Thysanolaena_latifolia</i>	Centothecaeae	0.25
<i>Elytrophorus_globularis</i>	Arundinoideae	0.3784	<i>Chasmanthium_laxum</i>	Chasmanthieae	0.04
<i>Eragrostis_walteri</i>	Arundinoideae	0.3784	<i>Lecomtella_madagascariensis</i>	Lecomtelleae	1
<i>Styppeiochloa_gynoglossa</i>	Arundinoideae	0.3784	<i>Coleataenia_prionitis</i>	Paspaleae	0.0015
<i>Crinipes_longifolius</i>	Arundinoideae	0.3784	<i>Zea_mays</i>	Andropogoneae	0.0031
<i>Crinipes_abyssinicus</i>	Arundinoideae	0.3784	<i>Coix_lacrymjobi</i>	Andropogoneae	0.0031
<i>Arundo_donax_Teisher95</i>	Arundinoideae	0.3784	<i>Sorghum_bicolor</i>	Andropogoneae	0.0031
<i>Arundo_donax_Teisher96</i>	Arundinoideae	0.3784	<i>Saccharum_hybrid</i>	Andropogoneae	0.0031
<i>Amphipogon_caricinus</i>	Arundinoideae	0.3784	<i>Digitaria_exilis</i>	Paniceae	0.0032
<i>Amphipogon_turbinatus</i>	Arundinoideae	0.3784	<i>Echinochloa_oryzicola</i>	Paniceae	0.0032
<i>Dregeochloa_pumilla</i>	Arundinoideae	0.3784	<i>Setaria_italica</i>	Paniceae	0.0032
<i>Monachather_paradoxus</i>	Arundinoideae	0.3784	<i>Panicum_virgatum</i>	Paniceae	0.0032
<i>Monachather_paradoxus_KJ920235</i>	Arundinoideae	0.3784	<i>Avena_sativa</i>	Pooideae	0.0021
<i>Monachather_paradoxus_TC5120</i>	Arundinoideae	0.3784			

Table S1.3. Log-likelihoods for evolution of lemma awns in PACMAD grasses under three models of trait evolution using maximum likelihood with R function *rayDISC* in package **corHMM** and under stochastic character mapping with package **phytools**.

	Equal Rates	Symmetric Rates	All Rates Different
Maximum Likelihood	-43.87	-41.11	-39.91
Stochastic Mapping	-47.74	-44.98	-43.79

CHAPTER 2

Phylogenetic Analysis of Polyploidy in Temperate Arundinoideae

2.1 INTRODUCTION

Polyploidy has been and continues to be a major source of genetic variation in plants. All flowering plants share ancient whole genome duplications or WGDs (Cui *et al.*, 2006; Jiao *et al.*, 2011; Arrigo & Barker, 2012) and many major crops are relatively recently-formed polyploids (Renny-Byfield & Wendel, 2014). Even the model species *Arabidopsis thaliana*, with a relatively small genome for a flowering plant, shows evidence of genome doubling (Blanc & Wolfe, 2004). Flowering plants are all ancient polyploids, and it is impossible to understand their evolution without addressing the history of genome duplications.

The grass family, Poaceae, is particularly noteworthy for its genome duplications, with 80% of species estimated as being the result of relatively recent polyploidy (Hunziker & Stebbins, 1986). Three WGD events have been identified in the lineages leading to the grasses: one early in the history of monocots called *tau* (Jiao *et al.*, 2014), another near the origin of Poales called *sigma* (Tang *et al.*, 2010), and a third duplication just prior to the origin of Poaceae called *rho* (Salse *et al.*, 2008; McKain *et al.*, 2016). The bamboos (subfamily Bambusoideae) experienced multiple allopolyploid events and reticulate evolution via intergeneric hybridization (Triplett *et al.*, 2014). Similarly, the tribe Andropogoneae in subfamily Panicoideae has experienced at least 34 WGD events, with a minimum of 32% of species resulting from allopolyploidy, although the actual number could be much higher (Estep *et al.*, 2014). Linder & Barker (2014) identified numerous nested polyploid events in subfamily Danthonioideae, with at least 23% of species having multiple ploidy levels.

Despite ubiquitous polyploidy in nature and the increased attention the phenomenon has gained since the advent of high-throughput molecular sequencing technologies, there is surprisingly little consensus as to why genome duplication is so common in plant evolution

(Soltis *et al.*, 2010; Madlung, 2013). Polyploid taxa have been hypothesized to have advantages over their diploid progenitors under extreme or variable conditions, for example in the arctic (Brochman *et al.*, 2004) and during the K-T extinction (Fawcett *et al.*, 2009; Lohaus & Van de Peer, 2016) though these advantages have been difficult to test because of confounding factors (Fawcett & Van de Peer, 2010). Polyploidy has been found to be more common among invasive plants and less common among rare plants (Pandit *et al.*, 2011), and there is some evidence that polyploid grasses are more successful at long-distance dispersal, potentially due to increased establishment ability in polyploids as compared to diploids (Linder & Barker, 2014). However, Martin and Husband (2009) examined three species from each of 144 North American plant genera and found that phylogenetic history strongly influences geographic and ecological ranges of species without a significant difference between ploidy levels.

Recent polyploid lineages have been associated with decreased diversification rates as compared to diploid ones (Mayrose *et al.*, 2011), while ancient WGDs are often associated with increased diversification rates following a lag period (Schranz *et al.*, 2012; Tank *et al.*, 2015). At least part of the cause of this apparent paradox is due to limitations in the methods used to identify polyploidy and model its evolution (Madlung, 2013; Kellogg, 2016). In particular, chromosome counts are frequently used to identify different ploidy levels, but this approach potentially suffers from chromosomal rearrangements (i.e. in maize; Wei *et al.*, 2007) and from an inability to distinguish between auto- and allopolyploids (Catalan *et al.*, 2012). Sequence-based approaches like K_s ratios and synteny plots do not depend on retained chromosomal structure, but the former method may have difficulty distinguishing signals from multiple historical events, and suitable quality genomes required by the latter are still sparse and concentrated in diploid species (Kellogg, 2016).

The grass subfamily Arundinoideae represents a unique opportunity in which the potential for genome duplication and possible hybridization to facilitate evolutionary success can be examined. This small subfamily contains a heterogeneous group of mostly tropical taxa, but two lineages have spread to occupy temperate habitats. The first is a clade consisting of the genera *Phragmites*, *Molinia*, and *Hakonechloa*. *Phragmites* is cosmopolitan (Haslam, 2010), whereas *Hakonechloa* is restricted to Japan, and its closest relative *Molinia* extends across Europe, western Asia, and northern Africa (Watson & Dallwitz, 1992; Taylor *et al.*, 2001); both *Hakonechloa* and *Molinia* are widely cultivated as ornamentals (Greenlee *et al.*, 1992). *Molinia* is strongly competitive and potentially invasive in wetlands, heathlands, and grasslands of Europe (Todd *et al.*, 2000; Hájková *et al.*, 2009). It is also a high polyploid, with up to 12x ploidy levels reported in the genus (Dančák *et al.*, 2012). *Hakonechloa* is reported as having a chromosome number of 50 (Tateoka, 1955) or 48 (Rice *et al.*, 2014), making it a tetraploid or octaploid depending on which base chromosome number is used for the Arundinoideae (Hardion *et al.*, 2015).

Phragmites australis is genetically and morphologically variable across its range (Hansen *et al.*, 2007). In North America, populations of *P. australis* reproduce either predominately sexually or vegetatively depending on whether they are new colonizers or well-established in an area, with sexual reproduction and seed production playing a major role in dispersal and establishment of new populations (Albert *et al.*, 2015). High ploidy levels have evolved independently multiple times in *P. australis* and the genus as a whole (Lambertini *et al.*, 2006) due potentially in part to long-distance dispersal and weak reproductive barriers between clones and species (Lambertini *et al.*, 2012). Ploidy levels from 3x to 12x as well as numerous aneuploids have been reported in this species (Clevering & Lissner, 1999). The tetraploid form is

considered predominant and presumably ancestral in the genus (Lambertini *et al.*, 2006), with higher ploidy levels the result of autopolyploidy or intrageneric allopolyploidy. However, the origins of the subgenomes in the tetraploid are unknown.

The second temperate clade in Arundinoideae is made up of the five species of *Arundo*, one of which, *Arundo donax*, is nearly as widespread as *Phragmites australis*, though slightly more restricted to warm climates. The other four species are distributed across Eurasia with a concentration in the Mediterranean region (Hardion *et al.*, 2012). In contrast with the frequently outcrossing *Phragmites*, accessions of *A. donax* are sterile and spread exclusively through vegetative propagules (Mariani *et al.*, 2010). The invasive form of this species in North America possesses a single multilocus genotype based on Sequence Related Amplification Polymorphisms and transposable element-based molecular markers (Ahmad *et al.*, 2008). Additionally, all samples of *A. donax* in the Mediterranean represent a single invasive haplotype from Asia based on hypervariable plastid DNA sites (Hardion *et al.*, 2014). In Australia, sterile stands of *A. donax* possess up to three distinct genotypes, suggesting multiple invasions via vegetative propagules (Haddadchi *et al.*, 2013). The reason for this sterility appears to be the odd ploidy level found in the species, which has 110 chromosomes whereas a close and fertile relative, *A. plinii*, has 72 (Bucci *et al.*, 2013).

Phragmites and *Arundo* share several features in common besides their geographic distribution. Both are large reeds, growing up to 6m tall in *Phragmites* and 10m tall in *Arundo*, with strongly lignified hollow culms, broad leaves, and plumose inflorescences. The species of these genera spread vegetatively through rhizomes and tend to occupy wetland habitats, although their climatic distributions are quite broad. Phylogenetic analyses of chloroplast genomes show that at least the maternal genomes of these genera are not sister to each other (Barker *et al.*, 1995;

Clark *et al.*, 1995; GPWG, 2001; GPWGII, 2011; Chapter 1, this dissertation). However, both *Arundo* and *Phragmites* are complex polyploids, with chromosome numbers up to 110 and 144, representing ploidy levels of 9-10x and 12x, respectively (Bucci *et al.*, 2013; Clevering & Lissner, 1999), and the nuclear phylogenetic relationships in Arundinoideae have not been investigated.

The physiology and ecology of *P. australis* and *A. donax* have been explored extensively (see Lambert *et al.*, 2010 for a review of ecology), due in large part to the species' potential for ecological invasiveness (i.e. Saltonstall, 2002; Ahmad *et al.*, 2008) and biofuel production (Laurent *et al.*, 2015). Transcriptomic data has also been generated for both species (He *et al.*, 2012; Barrero *et al.*, 2015) in an attempt to elucidate genes responsible for invasiveness. However, the origin of the sub-genomes of these species is unknown, leaving open questions regarding the possible role of WGD and hybridization in facilitating the success of these large reeds.

In this study, I conduct a phylogenetic analysis of newly-generated transcriptomic sequences of *Arundo*, *Hakonechloa*, *Molinia*, and *Phragmites* with published and unpublished coding DNA sequences (CDS) from transcriptomes and sequenced genomes for the grass subfamilies Anomochlooideae, Panicoideae, Aristidoideae, Oryzoideae, and Bambusoideae. Using these data, I trace the history of the Arundinoideae subgenomes and explore whether the morphological and ecological similarities between *Arundo* and *Phragmites* can possibly be explained by uniquely shared genomic elements. This analysis also represents the largest sampling of transcriptomes from PACMAD subfamilies to date, so implications for genome evolution in this clade are also explored. Lastly, I discuss limitations and advantages of the

phylogenetic approach as compared to other common methods for investigating polyploidy in the context of PACMAD grasses.

2.2 MATERIALS AND METHODS

2.2.1 Taxon Sampling

Plants for this study were either wild specimens or were grown at the Tyson Research Center or the Jeanette Goldfarb Plant Growth Facility at Washington University in St. Louis (Table 2.1). Two samples each of *Arundo donax*, *Hakonechloa macra*, *Molinia caerulea*, and *Phragmites australis* were sampled at the vegetative apex, including at least one mature leaf along with the meristem. Samples were immediately immersed in liquid nitrogen and transported to a -80°C freezer for storage.

2.2.2 RNA Extraction and Sequencing

Total RNA was extracted for all samples using a protocol developed by Simon Malcomber for the Kellogg Lab. Tissue was ground in liquid nitrogen using a mortar and pestle, and ca. 500 µL of Invitrogen TRIzol (Thermo Fisher Scientific) was added to each sample while still cold. Samples were allowed to thaw, were ground further in TRIzol, and incubated at room temperature for 10 mins. RNase-free chloroform was added to each sample in a 1:2 ratio (chloroform:TRIzol), and the combined samples were vortexed, incubated at room temperature for an additional 10 mins, and centrifuged at 12,000xg for 15 mins at 4°C. The aqueous layers from each tube were transferred to a new tube and combined with an equal volume of nuclease-free water. Ice-cold RNase-free isopropanol was added to this mixture in a 1:1 ratio and mixed by inversion of the tube before incubating samples for 10 mins at room temperature and again

vortexing at 12,000xg for 15 mins at 4°C. The supernatant was decanted, and the pellet was washed with cold, freshly-made 80% ethanol. The tubes were centrifuged a third time at 12,000xg for 5 minutes at 4°C to secure the pellet. The supernatant was again decanted, and samples were allowed to air dry for approximately 10 mins before being suspended in 50µL water. Qiagen DNase 1 was used to remove DNA from the samples according to the manufacturer's protocol.

Eight cDNA libraries – two from each species – were prepared using the NEBNext Poly(A) mRNA Magnetic Isolation Module and Ultra Directional RNA Library Prep Kit for Illumina following the manufacturer's protocols (New England Biolabs, Inc.). Libraries were size selected to a total size of approximately 500-700 base pairs (bp) and purified using AMPure XP Beads (Beckman Coulter, Inc.). Transcriptomic sequences were generated using an Illumina 2x150 paired-end HiSeq run at the University of Illinois at Urbana-Champaign Roy J. Carver Biotechnology Center.

2.2.3 Transcriptome Assembly

Illumina reads for the two samples of each taxon were pooled and cleaned using Trimmomatic version 0.32 for TruSeq3-PE adapters with a sliding window of 10 bp, a phred score of 20, and a minimum read length of 40 bp (Bolger *et al.*, 2014). The resulting trimmed reads were assembled with Trinity v. 2.0.6 using the direction library setting and normalizing reads with a max read coverage of 15. The abundances of the assembled reads were measured using the RSEM method in Trinity, and sequences with less than 1% abundance were removed based on FPKM (Fragments Per Kilobase Million) values. The further reduced sequences were translated using RefTrans (<https://github.com/mrmckain/RefTrans>). This program conducts a

tblastx analysis of all assembly contigs against the primary transcripts from PACMAD genomes, including *Zea mays*, *Sorghum bicolor*, *Panicum virgatum*, *Setaria viridis*, and *Dichanthelium oligosanthes*, using an e-value cutoff of 1e-10. Hits were filtered to include only those contigs with bidirectional coverage of at least 85% to a known gene, and the best hits were used by Genewise v.2.20 as models for translation (see McKain *et al.*, 2016). The outputs of these assemblies were summarized using the TrinityStats perl script that is provided with version 2.0.6 of Trinity.

2.2.4 K_s Plots

K_s frequency plots were constructed using the FASTKs pipeline (<https://github.com/mrmckain/FASTKs>) following the methodology of McKain *et al.* (2016). Amino acid sequences from each taxon were blasted against themselves using an e-value cutoff of 1e-40 to identify putative pairs, discarding any pairs in which the sequences were identical, had fewer than 300 bp overlap, or less than 40% identity. The retained pairs were aligned using MUSCLE (Edgar, 2004) and translated back to DNA sequences using the program PAL2NAL v. 14 (Suyama *et al.*, 2006). These cDNA sequences were used to calculate numbers of changes per site for synonymous (K_s) and nonsynonymous (K_a) sites and their ratios (K_a/K_s) using the codeml program in PAML v. 4.8 (Yang, 2007) using the paired sequence settings (yn00, Yang & Nielsen, 2000) and the F3x4 model (Goldman & Yang, 1994) as outlined in McKain *et al.* (2012). The program mclust v. 5.0.2 (Fraley *et al.*, 2012) was then used in R to estimate normal mixture models for K_s values. A peak in the distribution of K_s values is considered evidence for a WGD, since such an event creates thousands of paralogues at the same time (Lynch & Connery, 2000).

2.2.5 Gene Clustering, Alignment, and Phylogenetic Analysis

Translated sequences from the four arundinoid taxa were also combined with those from nine other grasses for phylogenetic analyses (Table 2.1). Orthogroups for this combined sequence set were identified using OrthoFinder v. 0.4 with the default settings (Emms & Kelly, 2015), retaining only those groups with at least one sequence from each of the 13 species. An orthogroup contains all genes descending from a single ancestral gene and may include all paralogous genes resulting from duplications after that ancestor (Wapinski *et al.*, 2007), so the fact that an orthogroup contains sequences from all thirteen species in our analysis does not imply that it contains only thirteen sequences. For example, if a species in the analysis is a large autopolyploid, there could be many copies of a gene falling in the same orthogroup. For the purposes of simplicity, "orthogroup" and "gene" will be used interchangeably in this paper, so a "gene tree" represents a phylogenetic analysis of all of the sequences of an orthogroup.

Peptide sequences in filtered orthogroups were aligned using MAFFT v.7.029b (Kato, 2013) with the default settings. Nucleic acid sequences were mapped to the peptide alignments via codons using PAL2NAL v. 14 (Suyama *et al.*, 2006). The resulting DNA alignments were used to construct gene trees with RAxML v.8.0.22 under a GTRGAMMA model of base pair substitution, treating *Streptochoeta* as an outgroup and bootstrapping each tree 500 times.

Gene trees were analyzed for signal of whole genome duplications using the program PUG (Phylogenetic Placement of Polyploidy Using Genomes, <https://github.com/mrmckain/PUG>, McKain *et al.*, 2016). This program compares gene trees to a user-specified species tree to identify nodes at which gene pairs from the same species coalesce. The first step is identification of a pair of sequences from the same species within an orthogroup.

The position of these sequences in the gene tree is checked, and sequences that are in clades with only sequences from the same species are ignored. By excluding sequences that are duplicated within a single terminal taxon, PUG will not identify a polyploidization event within that taxon so unshared polyploidization events are ignored here. If a gene copy from another species separates the focal sequence pair, the topology of the most exclusive clade in the gene tree that contains the focal pair is compared to the topology of the species tree. Sequence pairs in subtrees that violate the species tree topology are ignored, while pairs in subtrees that are concordant with the species tree are recorded by PUG. Repeating this analysis for all gene pairs in all orthogroups yields a list of gene trees and the nodes at which gene pairs in those trees coalesce.

An example may prove useful to illustrate this process. Figure 2.1 shows a hypothetical gene tree with species labelled by letter and gene copy by number. Four species, lettered A-D, have between 1 and 3 gene copies in this orthogroup. Assume PUG starts with gene copy A1 and compares it to A2. This pair is ignored because there are no gene copies from other species contained within the clade above the node at which A1 and A2 share a common ancestor (labelled with x). Now suppose A1 is compared with A3. These gene copies share a common ancestor at the node labelled by y. The total set of gene copies descending from this node includes those from species B, C, and D, so the gene tree topology is compared with the species tree, shown in the insert in Figure 2.1. PUG does not require all relationships between gene copies to mirror the gene tree, but rather that the species appearing in the subtree form a monophyletic group in the species tree. In our example, the subtree starting at y contains species A, B, C, and D, and does not contain species O. This agrees with our species tree, where species O is an outgroup to the other four. Thus, PUG would count our hypothetical gene tree as having a gene coalescence at node y. A similar logic applies to gene copies from species C, in which

C2+C3 would be ignored by PUG, but C1+C2 and C1+C3 would count as a coalescence point at node z, since all samples descending from this node come from species that form a monophyletic group in the species tree.

Once PUG has finished analyzing all gene pairs in all orthogroups, the information from the list can be summarized in several different ways. All gene pairs filtered by the program can be counted, or we can restrict each gene tree so that it can count only once for a particular node. In our example above this would mean that either A1+A3 or A2+A3 would contribute towards the count of coalescence events at node y, but not both. The gene pairs can also be filtered by bootstrap values, only allowing a gene pair to contribute to the count of a coalescent event at a particular node if the bootstrap support for that node in the corresponding gene tree is greater than a user-specified value. In our analyses, 50% and 80% bootstrap cutoffs were used to filter gene pairs from PUG, with only the counts from the 80% filter shown below. Going back to the example in Figure 1, the coalescence event between A1+A3 or A2+A3 (node y) would be counted under the 50% cutoff, but not under the 80% cutoff since the branch leading to that node has a bootstrap value of 75%. The coalescence of C1+C2 or C1+C3 at node z would be accepted under both cutoff levels since the branch leading to that node has a bootstrap value of 97%.

Because the relationships between subgenomes in the polyploid Arundinoideae are unknown, PUG was run multiple times using alternative topologies of the members of this clade: one following the chloroplast topology and five others treating two taxa as sister taxa with the other two unresolved (Figure 2.2). PUG was also run without restricting the number of times an orthogroup can be counted in support of a coalescence event at a given node to measure the relative proportion of gene copies from each taxon that coalesce to each node. This procedure helps to identify whether the coalescence points identified by PUG are supported evenly across

the members of the corresponding clade. For example, suppose that node z in Figure 2.1 is associated with 200 unique gene trees identified by PUG, including the gene tree in that figure. We could ask how many gene pairs from species B or D also contribute to this number, but since each gene tree counts only once for a given node, there is no guarantee that the numbers of gene trees counting gene pairs from each species are proportional to the total numbers of gene pairs from those species that coalesce to the node in question. In other words, if the hypothetical gene tree in Figure 2.1 also had gene pairs from species B and D that coalesce to node z, only one pair from one of the species (C, B, or D) would be counted. While node z may have 200 unique gene trees with approximately equal numbers of pairs from species B, C, and D, it may have 800 total gene pairs, of which 600 come from species C. This higher number may indicate a much higher copy number in general in species C, or it could suggest that species C is an allopolyploid resulting from a cross between species B and species D. In this particular example, we would have to examine gene trees to try to determine whether the 200 gene pairs coalescing to node z are trustworthy, for example by examining bootstrap values and the species composition of the relevant subtrees.

Figures depicting the results of PUG analyses were created using the R script `PUG_Figure_Maker` packaged with PUG (<https://github.com/mrmckain/PUG>), in which the species tree is plotted with branches beneath nodes colored according to the number of gene trees coalescing to that node. A value of 10% of the maximum value for any branch on the tree is used as a cutoff, with branches corresponding to nodes having fewer than this number of coalescing gene trees being left black. This cutoff is arbitrary and is used for visualization, so it is important also to compare the actual counts of coalescing gene pairs. It is also vital to keep in mind that the events identified by PUG are coalescence events between gene copies, not necessarily whole

genome duplications. Hybridization between non-sister species and phylogenetic uncertainty between members of a clade can both push gene pair coalescence deeper into the phylogeny. These alternative explanations will be explored for the current study in the Discussion section below.

All assemblies, alignments, and analyses were performed on the Apollo computing cluster at the Donald Danforth Plant Science Center. Raw data from transcriptome sequencing will be deposited on the Sequenced Read Archive (SRA) of the National Center for Biotechnology Information (NCBI), and the full transcriptome assemblies and all analyses will be stored on Dryad (datadryad.org).

2.3 RESULTS

2.3.1 Transcriptome Assembly and Orthogroup Analysis

Assembly summaries of the translated transcripts for the four arundinoid species are given in Table 2.2. Total assembly length and total transcript number were significantly lower in *Hakonechloa macra*, with roughly half as many base pairs and translated sequences as *Arundo donax*. As a result of this lower sequence coverage, the mean and median total contig lengths and the N50 value from TrinityStats (the average length of contigs making up 50% of the total assembly length) are longer in *Hakonechloa*. This phenomenon is likely due to the fact that high expression sequences are assembled in full at relatively low sequencing depths, but that low expression sequences only become partially assembled even at relatively high sequencing depths. Thus, average length of the assembly would increase up to a certain sequencing depth and then start to decrease as these rare fragments are incorporated.

Orthofinder identified 3,381 orthogroups containing at least one sequence from all 13 species included in this study. Figure 2.3 shows the numbers of contigs for each of the four arundinoid taxa that were placed in orthogroups containing 1-13 species. These distributions were similar for *Arundo*, *Molinia*, and *Phragmites*, with the majority of contigs being in orthogroups with either very few or most of the species in the analysis. Most of the genes expressed in an organism are housekeeping genes that are common across species. The large number of contigs that are unique to a single species is likely due in part to assembly errors, which create contigs that cannot be aligned across species, but since only contigs with sequences from all species in the analysis are used, these errors do not pose a problem for this study. Consistent with the lower total transcript number and assembly size, *Hakonechloa* contig counts are lower across all species number categories except for the full set of 13. The lack of a substantial number of contigs unique to *Hakonechloa* (the low species number peak seen in the other taxa) can similarly be explained by the lower sequence coverage, as more sequences would tend to add lower copy transcripts as well as errors, both of which would inflate the number of contigs that are unique to one or two species.

2.3.2 K_s Plots and PUG Analyses

K_s plots for the four arundinoid taxa are given in Figure 2.4. All samples show a signal of polyploidy, with peaks inferred by normal mixture models placed at values of K_s between ~ 0 and 0.7. The latter value is consistent with previous estimates of ρ (McKain *et al.*, 2016), but the signal for this event is weakened by more recent polyploid events in the sampled taxa. The other three peaks inferred by the models are clustered around low values of K_s and represent support for at least one WGD event.

PUG analyses identified coalescence points between substantial numbers of gene pairs corresponding to several nodes on the species tree that are consistent across alternative topologies in Arundinoideae. Many pairs coalesce at the base of the species tree, presumably reflecting the signal of the grass duplication *rho* combined with rooting issues due to the use of *Steptochaeta* as an outgroup (M. R. McKain, Donald Danforth Plant Science Center, pers. communication). Thus, gene pairs supporting this event were ignored in subsequent analyses. The PUG results based on the chloroplast species topology are shown in Figure 2.5. The branch on the species tree associated with the largest number of coalescing gene pairs in unique gene trees is the branch leading to *Zea+Sorghum*. However, this result can best be explained by a known genome duplication shared by *Zea* and *Tripsacum* combined with short branches in this part of the phylogeny. Swiganova *et al.* (2004) analyzed 11 orthologous genes in maize, sorghum and rice to identify the progenitor genomes of tetraploid maize. Their study confirmed that maize is of tetraploid origin and showed that the two maize progenitor genomes diverged from one another around the same time that they diverged from sorghum. However, Estep *et al.* (2014), using four low-copy nuclear loci in 100 species in tribe Andropogoneae, found that the maize tetraploidy occurred after the divergence from sorghum but before the origins of the genera *Zea* and *Tripsacum*. Thus, when phylogenetic analysis of maize and sorghum relatives is sparse, the timing of the maize allopolyploidy event is estimated to be older than it is under denser taxon sampling. The current study includes only maize and sorghum in the Andropogoneae and would thus be expected to recover results similar to Swiganova *et al.* (2004) in which the timing of the *Zea* duplication is difficult to disentangle from the divergence between *Zea* and *Sorghum*.

One coalescence event at the base of the PACMAD clade is associated with 410 unique gene trees and 2,151 gene pairs, although these pairs are not equally distributed across all species

in the clade. Sequences from *Panicum* make up 525 of the 2,151 pairs identified by PUG as coalescing to this node, while *Zea*, *Sorghum*, and *Dichanthelium* constitute only 61, 90, and 81 pairs, respectively. The remaining taxa possess between 122 and 345 gene pairs that coalesce to nodes in their respective gene trees corresponding to the base of PACMAD.

The node connecting *Phragmites*, *Hakonechloa* and *Molinia* is associated with coalescence of 1,631 total gene pairs in 381 unique gene trees, with gene pairs coming roughly equally from all three taxa. The node below this point in the species tree, representing the ancestor of all members of Arundinoideae in the current sampling, corresponds to the site of coalescence of 1,256 gene pairs in 231 unique gene trees. *Arundo* has the fewest gene pairs coalescing to this point out of the four arundinoid taxa. Of the total 1,256 gene pairs, only 165 are from *Arundo*, with *Phragmites*, *Molinia*, and *Hakonechloa* contributing 291, 416, and 284 pairs, respectively. Additionally, in 105 of the 165 *Arundo* pairs one of the gene copies is in a subclade by itself.

In alternative topologies of Arundinoideae, unique coalescing gene copies in unique gene trees are found in the following pairs: *Arundo*+*Hakonechloa* – 2 trees (Figure 2.6A); *Arundo*+*Molinia* – 2 trees (Figure 2.6B); *Arundo*+*Phragmites* – 0 trees (Figure 2.6C); *Hakonechloa*+*Molinia* – 91 trees (Figure 2.7A); *Hakonechloa*+*Phragmites* – 51 trees (Figure 2.7B); *Molinia*+*Phragmites* – 104 trees (Figure 2.7C). It is noteworthy that in the alternative tree in which *Hakonechloa* and *Molinia* are sister and the relationship of this clade with *Arundo* and *Phragmites* is left unresolved (Figure 2.7A), 91 unique gene trees have at least one gene pair within one of these taxa coalescing to the common ancestor of both taxa, while only 39 unique trees display this pattern in the chloroplast tree. Examination of the gene trees reveals that the 52 extra gene trees supporting a coalescence event in the former case lack any sequences from

Phragmites in the relevant subtree and would thus violate the chloroplast species tree. By allowing either *Phragmites* or *Arundo* to serve as the outgroup, the alternative tree includes gene pairs that could either coalesce to the common ancestor of *Hakonechloa* and *Molinia* or to the base of *Phragmites*+*Hakonechloa*+*Molinia*. The same phenomenon occurs in the alternative tree treating *Molinia* and *Phragmites* as sister (Figure 2.7C); of the 104 unique gene trees identifying a coalescence point at the node connecting this pair between gene pairs in one species, 60 are missing *Hakonechloa* sequences from the relevant subtree.

Two coalescence events in the Panicoideae are associated with a lower number of unique gene trees: one including all sampled members of Panicoideae and another including only *Panicum* and *Setaria*. The event at the base of Panicoideae is supported by 161 unique gene trees, while the one shared by *Panicum* and *Setaria* is supported by 105 such trees. It is difficult to say whether or not these values represent significant support for coalescence events. The highest value in the chloroplast topology that is not highlighted in the PUG plot is 49 unique gene trees corresponding to Panicoideae+Arundinoideae. The lowest value in this tree that is identified as an event in the PUG plot is 231 unique gene trees corresponding to the Arundinoideae. Thus, the two Panicoideae event values are intermediate between the highest "nonsignificant" and the lowest "significant" values. However, it is noteworthy that when examining all orthologous pairs in all gene trees, sequences from *Panicum* make up 71% and 80% of the pairs supporting the Panicoideae and *Panicum*+*Setaria* events, respectively. This result is consistent with what would be expected if *Panicum* were the result of a hybridization involving a distant relative. Gene copies from this parent (presumably the father, since it is not reflected in the chloroplast tree) would fall outside the Paniceae or possibly outside the Panicoideae, while copies from the other parent (mother) would follow the chloroplast topology.

2.4 DISCUSSION

This study represents the first phylogenetic analysis of the polyploid genomes in Giant and Common Reed, *Arundo donax* and *Phragmites australis*. These two species are ecological heavyweights, dominating wetland habitats across the globe, and are of considerable economic interest due to use of their culms to make shelter and musical reeds, potential use as biofuels, and invasive tendencies (Lewandowski *et al.*, 2003; Haslam, 2010; Lambert *et al.*, 2010). The results of this study are discussed below in the context of the geography, ecology, and evolution of *Arundo*, *Phragmites*, and their relatives. Since this is also the first study to examine transcriptomes from three PACMAD subfamilies in a phylogenetic context, the broader implications for genome evolution in this highly speciose, successful, and economically important clade are also examined. In the following sections, I often refer to gene pairs "supporting an event" at a particular position in the species tree. This is a shorthand used for convenience and can be interpreted as gene pairs coalescing to a node/branch in their respective gene trees that corresponds to the node/branch in question in the species tree. Hybridization between taxa from different clades in a phylogenetic analysis like the one performed by PUG can cause coalescence of gene pairs from the descendant hybrid offspring to occur deep in the tree, generating patterns resembling those seen in WGDs, so it is important to keep in mind that gene pair coalescence does not necessarily imply a WGD. Also, coalescence points in this analysis can only be identified with respect to the taxa that are sampled, so while a phylogenetic approach of sequence data provides more information regarding historical placement of putative genome duplications than non-phylogenetic approaches, this advantage is proportional to the taxon sampling in the analysis.

2.4.1 Evolution of Reedy Arundinoideae

The results of the K_s plots and PUG analyses suggest at least one coalescence event in the Arundinoideae. The putative event identified by the K_s plots is identical between *Arundo* and the other arundinoid taxa, even though there is fairly strong support for a WGD shared between *Phragmites*, *Molinia*, and *Hakonechloa* that is not shared by *Arundo*. The K_s plots thus confirm the polyploid nature of the arundinoid species, but are unable to provide details regarding the number of independent recent events or the timing of specific events, highlighting one of the limitations of this approach – namely, the inability to distinguish between signal from events occurring close to one another in time (Doyle & Egan, 2010; Kellogg, 2016). Analysis of the full set of gene pairs supporting the event at the base of Arundinoideae show that all four taxa contribute substantial numbers of orthologues, lending credibility to the claim that the event represents a shared ancestral WGD rather than hybridization between *Arundo* and one of the other arundinoids in this study. The extremely low numbers of gene trees supporting coalescence events between *Arundo* and any of the other arundinoid taxa supports this conclusion. Another possibility would be that the common ancestor of *Phragmites*, *Molinia* and *Hakonechloa* was a hybrid between distant relatives that subsequently underwent a WGD. Paralogous genes from this latter duplication would coalesce to the base of *Phragmites-Molinia-Hakonechloa*, while copies from the two parents would coalesce deeper in the tree. Broader sampling in the Arundinoideae and closely related subfamilies like Micrairoideae, Chloridoideae, and Danthonioideae is needed to determine whether such a hybridization event occurred and which lineages participated in the cross.

The other event identified by the K_s plot in *Hakonechloa*, *Molinia*, and *Phragmites* is a separate coalescence shared by those taxa and not by *Arundo*. The alternative topologies analyzed with PUG show that this event is shared by all three taxa and thus most likely also represents a WGD event. While some gene trees support an event between *Hakonechloa* and *Molinia* and another between *Phragmites* and *Molinia*, the fact that the majority of these trees are missing the relevant outgroup in the corresponding subtree suggests that this moderate signal may be an artifact of transcriptome sampling. However, gene flow between these three taxa has not been tested, and hybridization between different ploidy levels is common in *Phragmites* and *Molinia* (Clevering & Lissner, 1999; Dančák *et al.*, 2012), so the possibility that members of these genera are also hybridizing cannot be ruled out. Another possible explanation is incomplete lineage sorting of gene copies, which is especially likely given the relatively young age of this clade (ca. 3.5 mya according to BEAST dating in Chapter 1).

The K_s plots hint at a possible WGD event in *Arundo* that has yet to be identified or placed by phylogenetic analysis. This putative event is not the result of hybridization between *Arundo* and any other arundinoid taxa in the current study as evidenced by the lack of significant gene tree support in any of the PUG analyses of alternative topologies. The data fail to support the hypothesis of Bucci *et al.* (2013) on odd ploidy in *Arundo donax*. Specifically, the authors hypothesize a cross between *P. australis* with 96 chromosomes and a tetraploid resulting from genome doubling in *A. plinii* with 72 chromosomes to yield an *Arundo*-like hybrid offspring with 120 chromosomes that are then reduced through aneuploidy to 110. This somewhat complicated scenario would be expected to produce at least some gene pairs shared between *A. donax* and *P. australis* that are not shared by *Hakonechloa* or *Molinia*, but we find no evidence of such pairs. Since *A. plinii* was not sampled in the current analysis, the alternative hypothesis of Bucci *et al.*

– a cross between diploid *A. plinii* and its tetraploid offspring combined with a gain of two chromosomes – cannot be tested here.

The absence of shared unique gene pairs between *Arundo* and *Phragmites* also means that shared genomes alone cannot be the cause of their convergent morphology. It is still possible that the similar features between these two taxa are due to the same genes shared via a WGD event at the base of Arundinoideae, but that those genes are not expressed in *Hakonechloa* and *Molinia*, or at least are not expressed in the same way.

The role that polyploidy has played in facilitating the geographic spread and ecological success of *Arundo* and *Phragmites* cannot be determined with the limited sampling in this study. Within the *Hakonechloa-Molinia-Phragmites* clade, *Phragmites* and *Molinia* possess much broader geographic and ploidy ranges than *Hakonechloa*, but causality in this relationship is unclear. If a WGD at the base of this clade facilitated the spread of *Molinia caerulea* and *Phragmites australis*, it remains to be explained why other members of these genera and *Hakonechloa macra* have maintained much more limited ranges. A similar problem exists for interpreting the putative WGD identified at the base of all Arundinoideae in our study. The majority of arundinoid species have limited geographic ranges (Kellogg, 2015), so if the genome duplication really did occur in the lineage leading to the subfamily, other factors are needed to explain why two clades have been so ecologically successful while the others are comparably restricted.

Arundo and *Phragmites* also present a case in which we must examine our definition of evolutionary success. These genera are relatively species-poor, supporting the story that polyploidy is an evolutionary dead-end. However, the contemporary success of *A. donax* and *P. australis* is undeniable and impressive. Equally striking is the fact that the spread of these species

has occurred via different population dynamics, with *A. donax* being sterile and *P. australis* preferentially reproducing sexually during invasions to new habitats. Despite their many morphological and ecological similarities, we find no evidence for hybridization playing a role in the comparable success of these large reeds.

2.4.2 PACMAD Genome Evolution

The support in our study for a PACMAD WGD event is a surprising result that needs broader sampling to evaluate fully. The PACMAD grasses constitute an enormously successful clade that lacks a clear distinguishing synapomorphy. Taxa in this clade have diversified to occupy a wide range of habitats, becoming dominant in C₄ grasslands and tropical and subtropical savannahs. All origins of the C₄ photosynthetic pathway in grasses occur in the PACMAD clade, which has been explained at least in part by an ancestral anatomical preadaptation (Christin *et al.*, 2013). The potential existence of a genome duplication preceding the origins of this clade is an exciting opportunity to explore the potential for polyploidy to drive or facilitate major long-term radiations.

However, this event has not been found by multiple previous studies using various approaches (i.e. Wei *et al.*, 2007; McKain *et al.*, 2016). None of these studies used a phylogenetic analysis of transcriptomic data from more than two PACMAD subfamilies, but this putative WGD event has not been seen in detailed genomic studies of *Sorghum*, *Setaria*, and *Zea*. It is possible that these methods, including K_s plots and synteny analyses, have difficulty distinguishing between multiple ancient WGDs (Kellogg, 2016). An explicitly phylogenetic analysis is in some ways a fundamentally different approach to identifying genome doubling events, so the possibility that the event identified in this study is real should not be ignored. That

being said, half of the subfamilies in PACMAD have not been included in the phylogeny, and adding clades could substantially affect the placement of the putative event.

The sparse sampling of outgroup lineages in our analysis may also be contributing to apparent gene coalescence at the base of PACMAD. Since all grasses share a WGD event, *rho*, a fully-sampled gene tree in this family would be expected to be mirrored, with all lineages represented at least once in each of two family-wide clades. However, in our analysis, since *Streptochoaeta* is used as the outgroup, this mirroring is distorted, and this distortion is the cause of the coalescence of gene pairs at the base of BOP+PACMAD that was ignored in our analysis (see Materials and Methods section). This same rooting problem could cause gene pair coalescence at the base of PACMAD if sequences from the two BOP lineages, *Oryza* and *Dendrocalamus*, cluster together at the base of gene trees, or if one of the pairs of gene copies in these lineages resulting from *rho* is missing from a gene tree. Both of these circumstances would cause the two *rho* clades of PACMAD to appear to share a common ancestor excluding the other grasses in our sample, thus generating a substantial number of gene pairs coalescing at the base of this clade. Transcriptomes from more BOP lineages and better outgroup sampling would greatly reduce or eliminate this problem.

A final issue in interpreting the PACMAD event is the ongoing difficulty in placing subfamily Aristidoideae in the phylogeny. Phylogenies based on chloroplast sequence data have recovered this subfamily in three different positions near the base of PACMAD (GPWGII, 2011; Cotton *et al.*, 2015; Chapter 1, this dissertation), and this uncertainty is reflected among the gene trees in our current analysis. The consistency of this problem across data sets supports a rapid radiation at the base of PACMAD. When using phylogenetic polyploidy analyses like PUG it is important to check individual gene trees because phylogenetic uncertainty can create artificial

coalescence events. In this case, *Aristida* does change phylogenetic position between different gene trees, but this inconsistency is insufficient to explain the signal for a possible PACMAD WGD event. Additional sampling across other PACMAD subfamilies could also help identify the placement of Aristidoideae with greater confidence, especially if nuclear genes could be combined with those from the plastome for representatives from all PACMAD subfamilies.

2.4.3 Approaches to Identifying WGDs

The results presented here highlight the need for multiple data sets and approaches to studying ancient polyploidy. Chromosome counts vary widely for *Arundo* and *Phragmites* (Connor & Dawson, 1993; Rice *et al.*, 2014), complicating ancestral reconstructions of chromosome number for the Arundinoideae. The frequency of recent eu- and aneuploidy in these taxa masks the older WGD identified by phylogenetic analysis of gene orthogroups. K_s plots of the four species generated for this study can only identify a single putative WGD that appears to be shared by all Arundinoideae, despite strong evidence from PUG for an event shared by *Phragmites*, *Molinia*, and *Hakonechloa* that is not shared by *Arundo*. On the other hand, coalescence events identified by phylogenetic analyses of gene trees are not guaranteed evidence of WGD (Doyle & Egan, 2010). Several evolutionary phenomena can lead to inferred coalescence between gene copies at a particular node in the phylogeny, and there is as yet no explicit statistical framework for evaluating support for a given event in PUG. Coupling phylogenetic analyses of putative polyploid events with synteny analyses of completed genomes can help identify genome duplications with greater confidence (i.e. McKain *et al.*, 2016). A logical next step for the possible PACMAD event found in this study would be to identify which gene trees coalesce to this point and then evaluate those genes for synteny in one of the available

PACMAD genomes like *Setaria* or *Sorghum*. If these genes are found in syntenic blocks, it would suggest that previous analyses of synteny in this genome could not distinguish between the PACMAD event and other WGDs in lineages leading to this species. Alternatively, if the genes in question are found scattered across the genome, other causes of their inferred shared coalescence will need to be explored.

2.4.4 Conclusion

This study presents a preliminary exploration into WGD events in the history of subfamily Arundinoideae and the PACMAD clade of grasses using a phylogenetic approach with transcriptomic data sets. A possible WGD is shared by all PACMAD taxa in the current analysis, though this result is most likely due to problems caused by rooting the tree with *Streptochaeta* as well as uncertainty in the phylogenetic position of *Aristida*. Members of Arundinoideae share two separate events, although there is little evidence at present that these events are causally related to the success of the large-statured invasive reeds *Arundo donax* and *Phragmites australis*. The morphological and ecological convergence between these species is not attributable to possession of uniquely shared genes that are not shared by other members of Arundinoideae. The addition of other arundinoid genera to the analysis would help in determining the nature of the putative event at the base of this subfamily.

LITERATURE CITED

- Ahmad, R., P.-S. Liow, D. F. Spencer, and M. Jasieniuk. 2008. Molecular evidence for a single genetic clone of invasive *Arundo donax* in the United States. *Aquatic Botany* 88: 113–120.
- Albert, A., J. Brisson, F. Belzile, J. Turgeon, and C. Lavoie. 2015. Strategies for a successful plant invasion: the reproduction of *Phragmites australis* in north-eastern North America. *Journal of Ecology* 103: 1529–1537.
- Arrigo, N., and M. S. Barker. 2012. Rarely successful polyploids and their legacy in plant genomes. *Current Opinion in Plant Biology* 15: 140–146.
- Barker, N. P., Linder, H. P., and E. H. Harley. 1995. Polyphyly of Arundinoideae (Poaceae): evidence from rbcL sequence data. *Systematic Botany* 20(4): 423–435.
- Barrero, R. A., F. D. Guerrero, P. Moolhuijzen, J. A. Goolsby, J. Tidwell, S. E. Bellgard, and M. I. Bellgard. 2015. Shoot transcriptome of the giant reed, *Arundo donax*. *Data in Brief* 3: 1–6.
- Blanc, G., and K. H. Wolfe. 2004. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *The Plant Cell* 16: 1667–1678.
- Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15): 2114–2120.
- Brochmann, C., A. K. Brysting, I. G. Alsos, L. Borgen, H. H. Grundt, A.-C. Scheen, and R. Elven. 2004. Polyploidy in arctic plants. *Biological Journal of the Linnean Society* 82: 521–536.
- Bucci, A., E. Cassani, M. Landoni, E. Cantaluppi, and R. Pilu. 2013. Analysis of chromosome number and speculations on the origin of *Arundo donax* L. (Giant Reed). *Cytology and Genetics* 47: 237–241.
- Catalán, P., J. Müller, R. Hasterok, G. Jenkins, L. A. J. Mur, T. Langdon, A. Betekhtin, et al. 2012. Evolution and taxonomic split of the model grass *Brachypodium distachyon*. *Annals of Botany* 109: 385–405.
- Christin, P.-A., C. P. Osborne, D. S. Chatelet, J. T. Columbus, G. Besnard, T. R. Hodkinson, L. M. Garrison, et al. 2013. Anatomical enablers and the evolution of C₄ photosynthesis in grasses. *Proceedings of the National Academy of Sciences of the United States of America* 110 (4): 1381–86.
- Clark, L. G., W. Zhang, and J. F. Wendel. 1995. A phylogeny of the grass family (Poaceae) Based on ndhF Sequence Data. *Systematic Botany* 20(4): 436–460.

- Clevering, O. A., and J. Lissner. 1999. Taxonomy, chromosome numbers, clonal diversity and population dynamics of *Phragmites australis*. *Aquatic Botany* 64: 185–208.
- Connor, H. E., and M. I. Dawson. 1993. Evolution of reproduction in *Lamprothyrus* (Arundineae: Gramineae). *Annals of the Missouri Botanical Garden* 80: 512–517.
- Cotton, J. L., W. P. Wysocki, L. G. Clark, S. A. Kelchner, J. C. Pires, P. P. Edger, D. Mayfield-Jones, and M. R. Duvall. 2015. Resolving deep relationships of PACMAD grasses: a phylogenomic approach. *BMC Plant Biology* 15(1): 1–11.
- Cui, L., P. K. Wall, J. H. Leebens-Mack, B. G. Lindsay, D. E. Soltis, J. J. Doyle, P. S. Soltis, et al. 2006. Widespread genome duplications throughout the history of flowering plants. *Genome Research* 16: 738–749.
- Dančák, M., M. Duchoslav, and B. Trávníček. 2012. Taxonomy and cytogeography of the *Molinia caerulea* complex in central Europe. *Preslia* 84: 351–374.
- Doyle, J. J., and A. N. Egan. 2010. Dating the origins of polyploidy events. *New Phytologist* 186: 73–85.
- Edgar, R. R. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.
- Emms, D. M. and S. Kelly. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology* 16:157.
- Estep, M. C., M. R. McKain, D. Vela Diaz, J. Zhong, J. G. Hodge, T. R. Hodkinson, D. J. Layton, et al. 2014. Allopolyploidy, diversification, and the Miocene grassland expansion. *Proceedings of the National Academy of Sciences of the United States of America* 111: 15149–15154.
- Fawcett, J. A., S. Maere, Y. Van de Peer, and M. C. E. Van Montagu. 2009. Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. *Proceedings of the National Academy of Sciences of the United States of America* 106: 5737–5742.
- Fawcett, J. A., and Y. Van de Peer. 2010. Angiosperm polyploids and their road to evolutionary success. *Trends in Evolutionary Biology* 2: 16–21.
- Fraley, C., A. E. Raftery, T. B. Murphy, and L. Scrucca. 2012. Mclust version 4 for R: normal mixture modeling for model-based clustering, classification, and density estimation.
- Goldman, N. and Z. Yang. 1994. A codon-based model of nucleotide substitutions for protein-coding DNA sequences. *Molecular Biology and Evolution* 11: 725–736.

- Grass Phylogeny Working Group (GPWG). 2001. Phylogeny and subfamilial classification of the grasses (Poaceae). *Annals of the Missouri Botanical Garden* 88: 373–457.
- Grass Phylogeny Working Group II (GPWGII). 2011. New grass phylogeny resolves deep evolutionary relationships and discovers C₄ origins. *The New Phytologist* 193 (2): 304–312.
- Greenlee, J., D. Fell, and W. Oehme. 1992. *The Encyclopedia of Ornamental Grasses: How to Grow and Use Over 250 Beautiful and Versatile Plants*. Emmaus, PA: Rodale Books.
- Haddadchi, A., C. L. Gross, and M. Fatemi. 2013. The expansion of sterile *Arundo donax* (Poaceae) in southeastern Australia is accompanied by genotypic variation. *Aquatic Botany* 104: 153–161.
- Hájková, P., M. Hájek, and K. Kintrová. 2009. How can we effectively restore species richness and natural composition of a *Molinia*-invaded fen? *Journal of Applied Ecology* 46: 417–425.
- Hansen, D. L., C. Lambertini, A. Jampeetong, and H. Brix. 2007. Clone-specific differences in *Phragmites australis*: effects of ploidy level and geographic origin. *Aquatic Botany* 86: 269–279.
- Hardion, L., R. Verlaque, A. Baumel, M. Juin, and B. Vila. 2012. Revised systematics of Mediterranean *Arundo* (Poaceae) based on AFLP fingerprints and morphology. *Taxon* 61: 1217–1226.
- Hardion, L., R. Verlaque, K. Saltonstall, A. Leriche, and B. Vila. 2014. Origin of the invasive *Arundo donax* (Poaceae): a trans-Asian expedition in herbaria. *Annals of Botany* 114: 455–462.
- Haslam, S. M. 2010. *A Book of Reed (Phragmites australis (Cav.) Trin. Ex Steudel, formerly Phragmites communis Trin.)*. Ceredigion, UK: Forrest Text.
- He, R., M.-J. Kim, W. Nelson, T. S. Balbuena, R. Kim, R. Kramer, J. A. Crow, *et al.* 2012. Next-generation sequencing-based transcriptomic and proteomic analysis of the common reed, *Phragmites australis* (Poaceae), reveals genes involved in invasiveness and rhizome specificity. *American Journal of Botany* 99: 232–247.
- Hunziker, J. H. and G. L. Stebbins. 1986. Chromosomal evolution in the Gramineae. Pp. 179–187 in *Grass Systematics and Evolution*, eds. T. R. Soderstrom, K. W. Hilu, C. S. Campbell, and M. E. Barkworth. Washington: Smithsonian Institution Press.
- Jiao, Y., J. Li, H. Tang, and A. H. Paterson. 2014. Integrated syntenic and phylogenomic analyses reveal an ancient genome duplication in monocots. *The Plant Cell* 26: 2792–2802.
- Jiao, Y., N. J. Wickett, S. Ayyampalayam, A. S. Chanderbali, L. Landherr, P. E. Ralph, L. P.

- Tomsho, *et al.* 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473: 97–100.
- Katoh, S. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780.
- Kellogg, E. A. 2016. Has the connection between polyploidy and diversification actually been tested? *Current Opinion in Plant Biology* 30: 25–32.
- Lambert, A. M., T. L. Dudley, and K. Saltonstall. 2010. Ecology and impacts of the large-statured invasive grasses *Arundo donax* and *Phragmites australis* in North America. *Invasive Plant Science and Management* 3: 489–494.
- Lambertini, C., M. H. G. Gustafsson, J. Frydenberg, J. Lissner, M. Speranza, and H. Brix. 2006. A phylogeographic study of the cosmopolitan genus *Phragmites* (Poaceae) based on AFLPs. *Plant Systematics & Evolution* 258: 161–182.
- Lambertini, C., I. A. Mendelssohn, M. H. G. Gustafsson, B. Olesen, T. Riis, B. K. Sorrell, and H. Brix. 2012. Tracing the origin of Gulf Coast *Phragmites* (Poaceae): a story of long-distance dispersal and hybridization. *American Journal of Botany* 99: 538–551.
- Laurent, A., E. Pelzer, C. Loyce, and D. Makowski. 2015. Ranking yields of energy crops: a meta-analysis using direct and indirect comparisons. *Renewable and Sustainable Energy Reviews* 46: 41–50.
- Lewandowski I., J. M. O. Scurlock, E. Lindvall, and M. Christou. 2003 The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe. *Biomass Bioenergy* 25: 335–361.
- Linder, H. P., and N. P. Barker. 2014. Does polyploidy facilitate long-distance dispersal? *Annals of Botany* 113: 1175–1183.
- Lohaus, R., and Y. Van de Peer. 2016. Of dups and dinos: evolution at the K/Pg boundary. *Current Opinion in Plant Biology* 30: 62–69.
- Lynch, M. and J. S. Conery. 2000. The evolutionary fate and consequences of duplicate genes. *Science* 290(5494): 1151–1155.
- Madlung, A. 2013. Polyploidy and its effect on evolutionary success: old questions revisited with new tools. *Heredity* 110: 99–104.
- Mariani, C., R. Cabrini, A. Danin, P. Piffanelli, A. Fricano, S. Gomasasca, M. Dicandilo, *et al.* 2010. Origin, diffusion and reproduction of the giant reed (*Arundo donax* L.): a promising weedy energy crop. *Annals of Applied Biology* 157: 191–202.
- Martin, S. L., and B. C. Husband. 2009. Influence of phylogeny and ploidy on species ranges of

- North American angiosperms. *Journal of Ecology* 97: 913–922.
- Mayrose, I., S. H. Zhan, C. J. Rothfels, K. Magnuson-Ford, M. S. Barker, L. H. Rieseberg, and S. P. Otto. 2011. Recently formed polyploid plants diversify at lower rates. *Science* 333: 1257–1257.
- McKain, M. R., N. Wickett, Y. Zhang, S. Ayyampalayam, W. R. McCombie, M. W. Chase, J. C. Pires, *et al.* 2012. Phylogenomic analysis of transcriptome data elucidates co-occurrence of a paleopolyploid event and the origin of bimodal karyotypes in Agavoideae (Asparagaceae). *American Journal of Botany* 99: 397–406.
- McKain, M. R., H. Tang, J. R. McNeal, S. Ayyampalayam, J. I. Davis, C. W. dePamphilis, T. J. Givnish, *et al.* 2016. A phylogenomic assessment of ancient polyploidy and genome evolution across the Poales. *Genome Biology and Evolution* 8(4): 1150–1164.
- Pandit, M. K., M. J. O. Pockock, and W. E. Kunin. 2011. Ploidy influences rarity and invasiveness in plants. *Journal of Ecology* 99: 1108–1115.
- Renny-Byfield, S., and J. F. Wendel. 2014. Doubling down on genomes: polyploidy and crop plants. *American Journal of Botany* 101: 1711–1725.
- Rice, A., L. Glick, S. Abadi, M. Einhorn, N. M. Kopelman, A. Salman-Minkov, J. Mayzel, *et al.* 2014. The Chromosome Counts Database (CCDB) – a community resource of plant chromosome numbers. *New Phytologist* doi: 10.1111/nph.13191.
- Salse, J., S. Bolot, M. Throude, V. Jouffe, B. Piegu, U. M. Quraishi, T. Calcagno, *et al.* 2008. Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution. *The Plant Cell* 20: 11–24.
- Saltonstall, K. 2002. Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. *Proceedings of the National Academy of Sciences of the United States of America* 99: 2445–2449.
- Schranz, E. M., S. Mohammadin, and P. P. Edger. 2012. Ancient whole genome duplications, novelty and diversification: the WGD Radiation Lag-Time Model. *Current Opinion in Plant Biology* 15(2): 147–153.
- Soltis, D. E., R. J. A. Buggs, J. J. Doyle, and P. S. Soltis. 2010. What we still don't know about polyploidy. *Taxon* 59: 1387–1403.
- Suyama, M., D. Torrents, and P. Bork. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research* 34: W609–W612.
- Swigoňová, Z., J. Lai, J. Ma, W. Ramakrishna, V. Llaca, J. L. Bennetzen, and J. Messing. 2004. Close split of sorghum and maize genome progenitors. *Genome Research* 14: 1916–1923.

- Tang, H., J. E. Bowers, X. Wang, and A. H. Paterson. 2010. Angiosperm genome comparisons reveal early polyploidy in the monocot lineage. *Proceedings of the National Academy of Sciences of the United States of America* 107: 472–477.
- Tank, D. C., J. M. Eastman, M. W. Pennell, P. S. Soltis, D. E. Soltis, C. E. Hinchliff, J. W. Brown, *et al.* 2015. Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytologist* 207: 454–467.
- Tateoka, T. 1955. Karyotaxonomy in Poaceae III. Further studies of somatic chromosomes. *Cytologia* 20(4): 296–306.
- Taylor, K., A. P. Rowland, and H. E. Jones. 2001. *Molinia caerulea* (L.) Moench. *Journal of Ecology* 89: 126–144.
- Todd, P. A., J. D. P. Phillips, P. D. Putwain, and R. H. Marrs. 2000. Control of *Molinia caerulea* on moorland. *Grass and Forage Science* 55: 181–191.
- Triplett, J. K., L. G. Clark, A. E. Fisher, and J. Wen. 2014. Independent allopolyploidization events preceded speciation in the temperate and tropical woody bamboos. *New Phytologist* 204: 66–73.
- Wapinski, I., A. Pfeffer, N. Friedman, and A. Regev. 2007. Automatic genome-wide reconstruction of phylogenetic gene trees. *Bioinformatics* 23: i549–i558.
- Watson, L. and M. J. Dallwitz. 1992. *The Grass Genera of the World*. Wallingford: CAB International.
- Wei, F., E. Coe, W. Nelson, A. K. Bharti, F. Engler, E. Butler, H. Kim, *et al.* 2007. Physical and genetic structure of the maize genome reflects its complex evolutionary history. *PLoS Genetics* 3(7): 1254–1263.
- Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* 24: 1586–1591.
- Yang, Z., and R. Nielsen. 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Molecular Biology and Evolution* 17: 32–43.

FIGURES AND TABLES

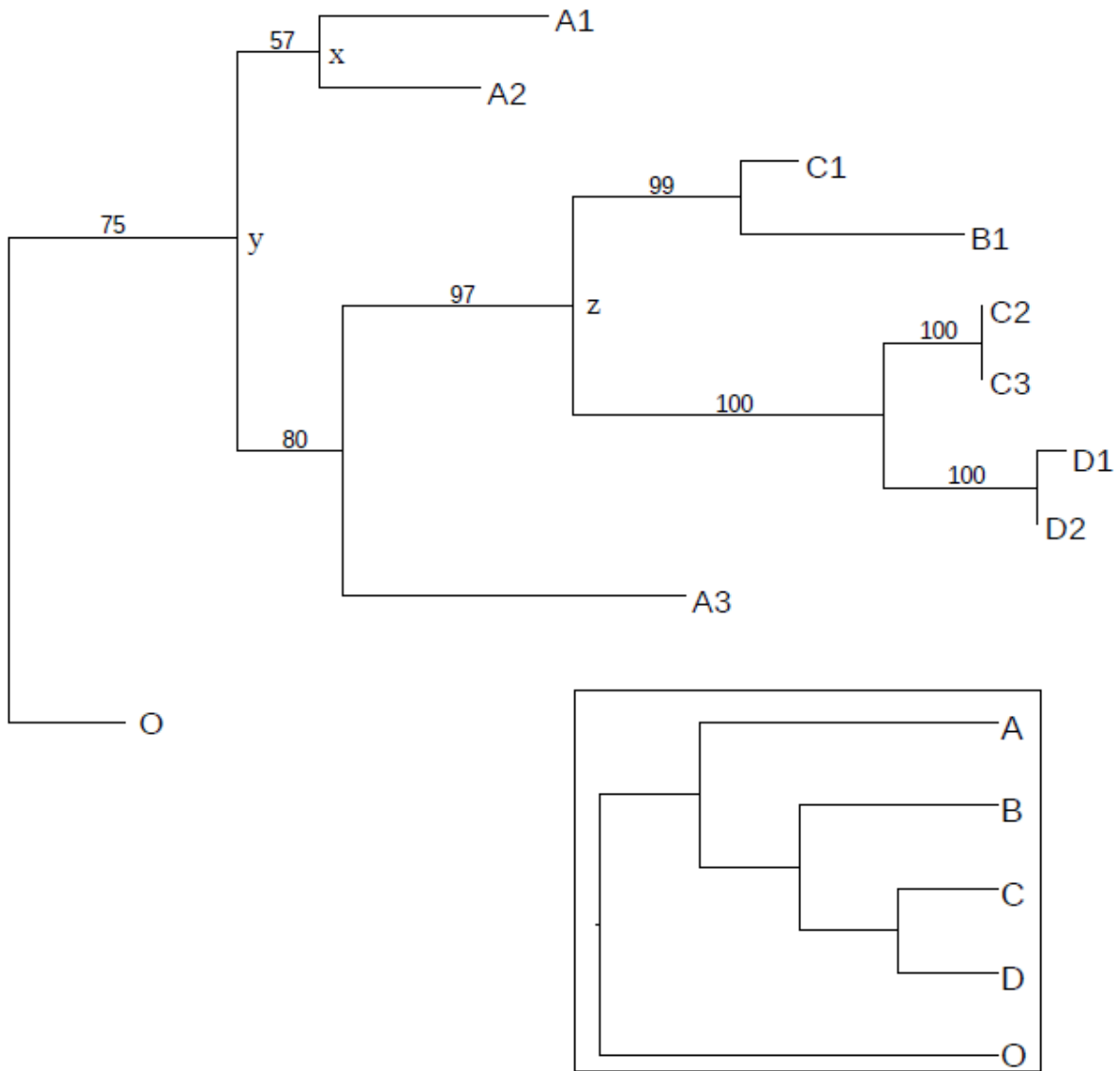


Figure 2.1. Hypothetical gene tree and species tree (inside box) for five species labelled A-D with gene copies labelled 1-3. The outgroup taxon is represented by O.

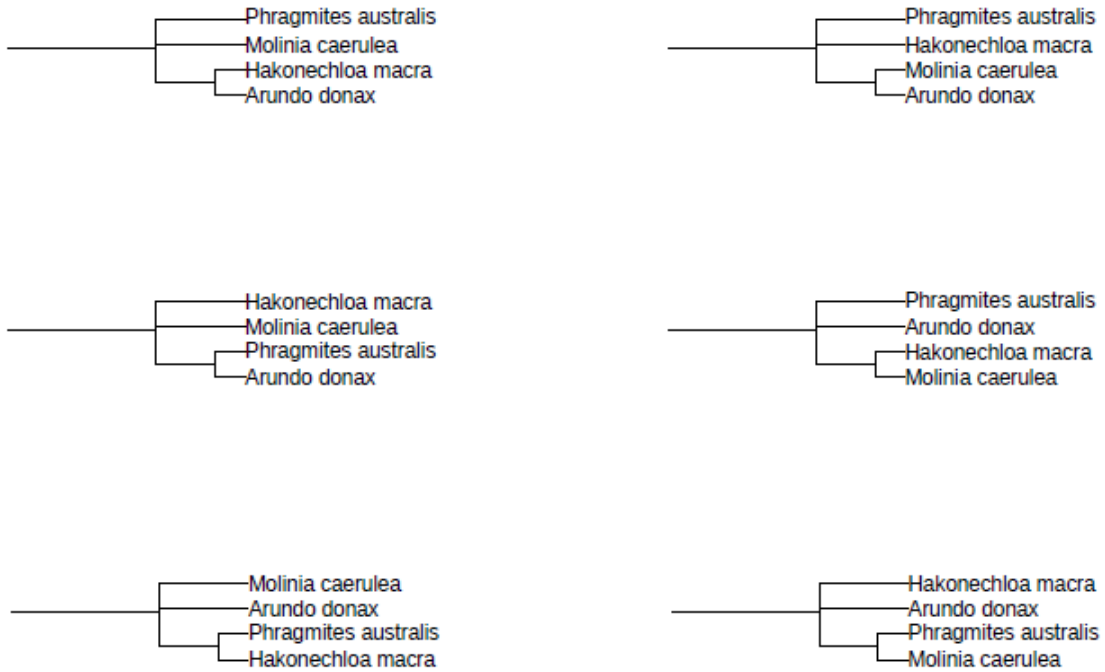
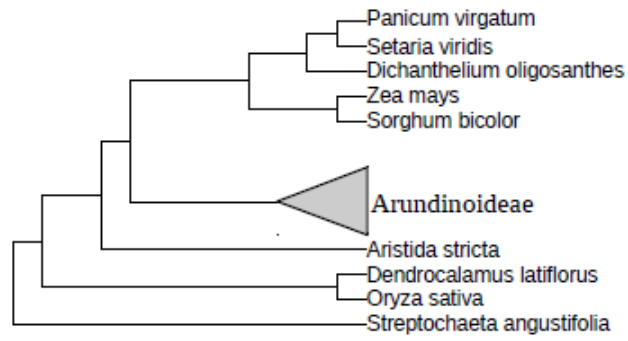


Figure 2.2. Alternative topologies used for samples of Arundinoideae in PUG analyses.

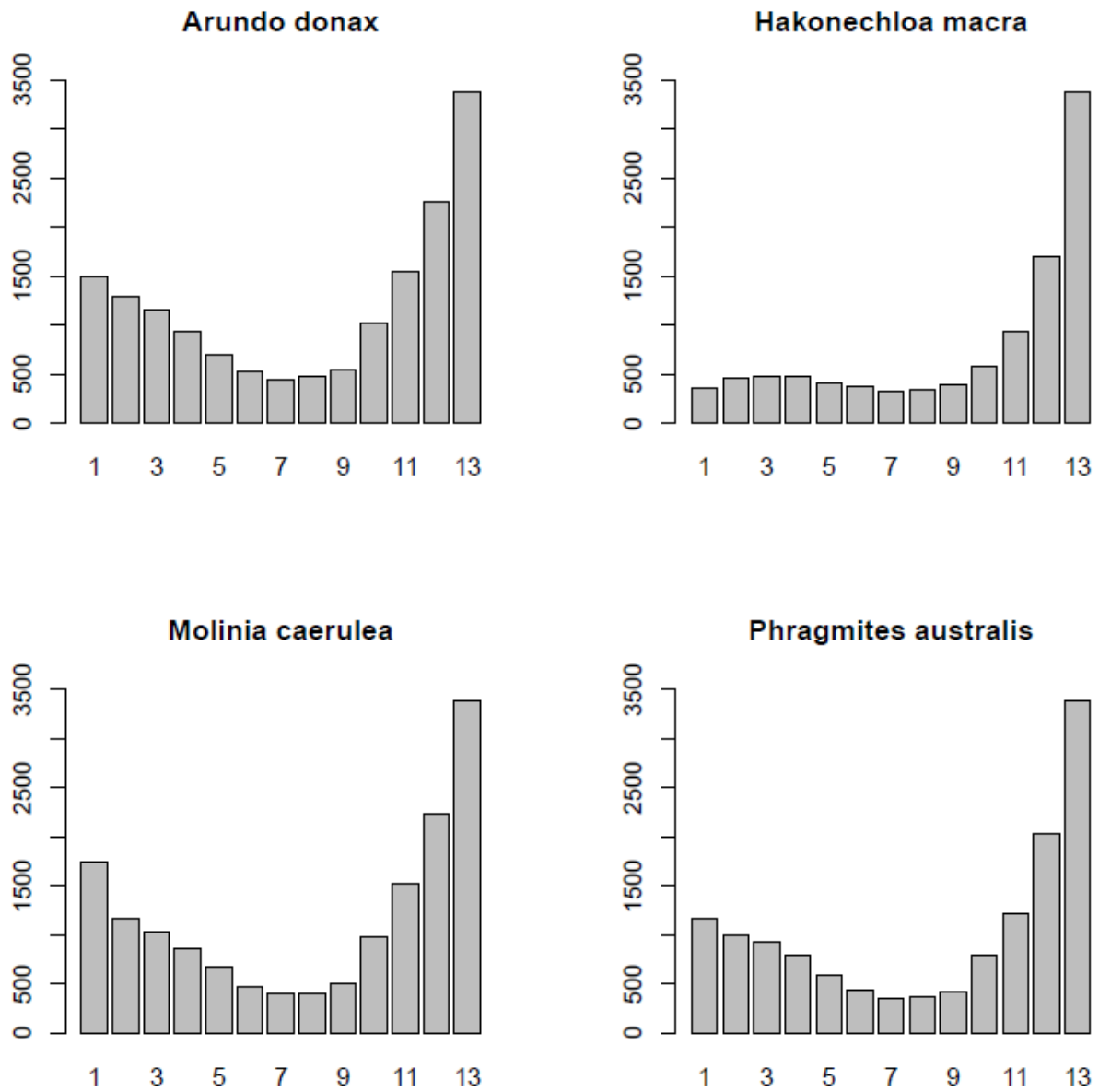


Figure 2.3. Numbers of genes from each arundinoid species in orthogroups containing genes from between one and all thirteen species in the analysis.

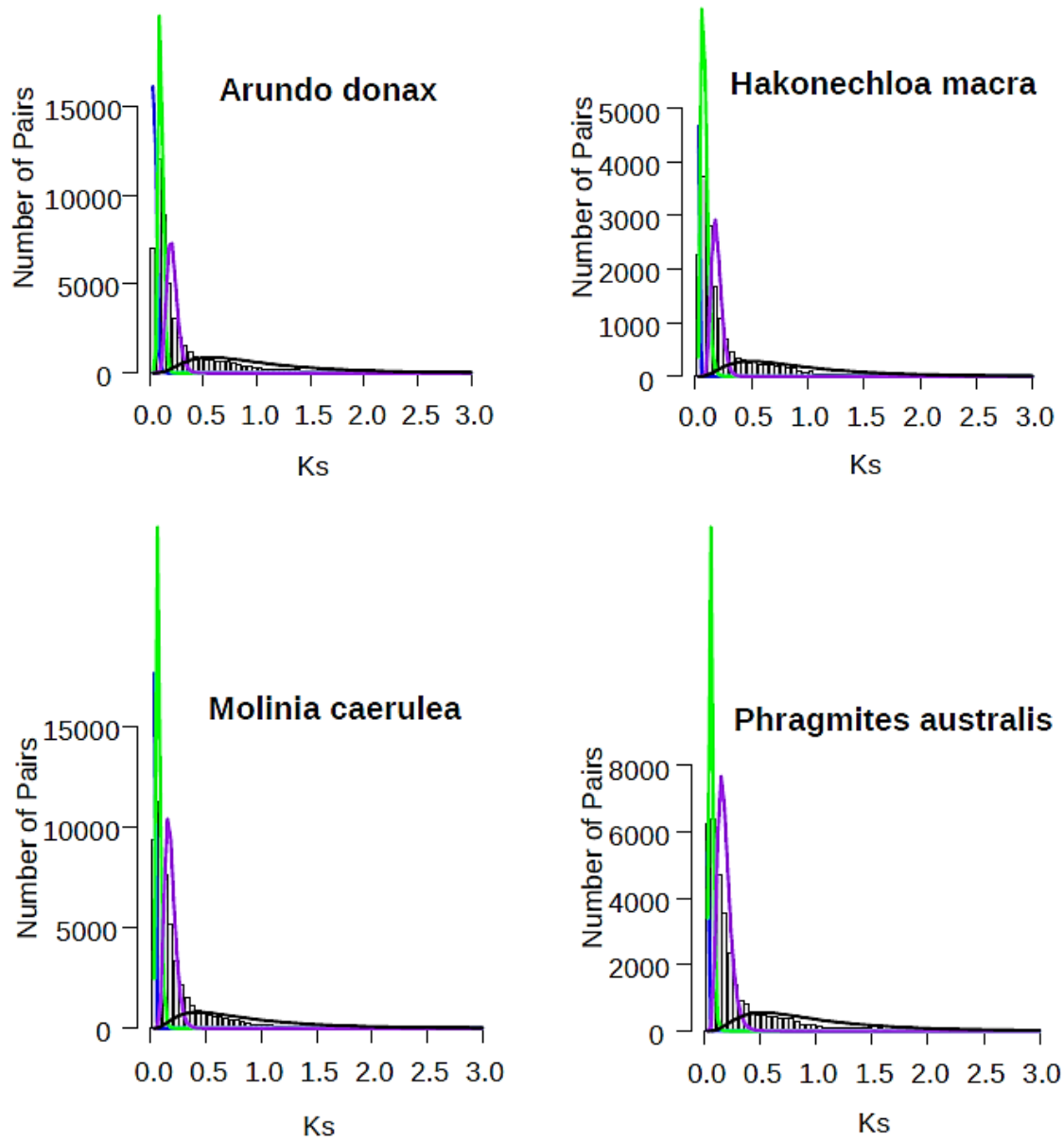


Figure 2.4. K_s plots for four arundinoid species using transcriptome sequence data.

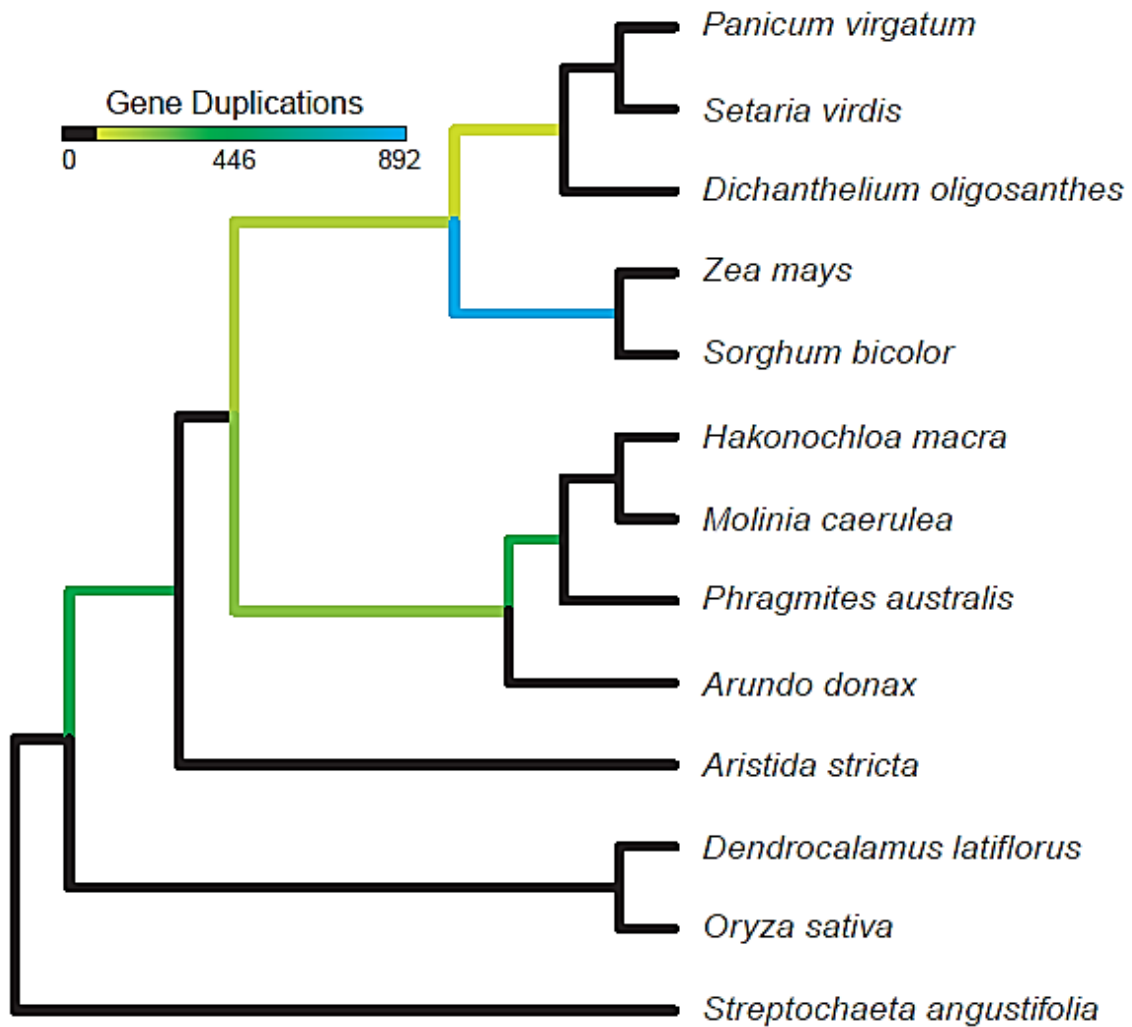


Figure 2.5. Results from PUG using relationships inferred by chloroplast phylogenies for the species tree topology. Branches in the species tree are colored according to how many unique gene trees possess gene pairs coalescing to the corresponding branch in that gene tree. Branches associated with fewer than 10% of the unique gene trees that are associated with the highest value in the tree are colored black

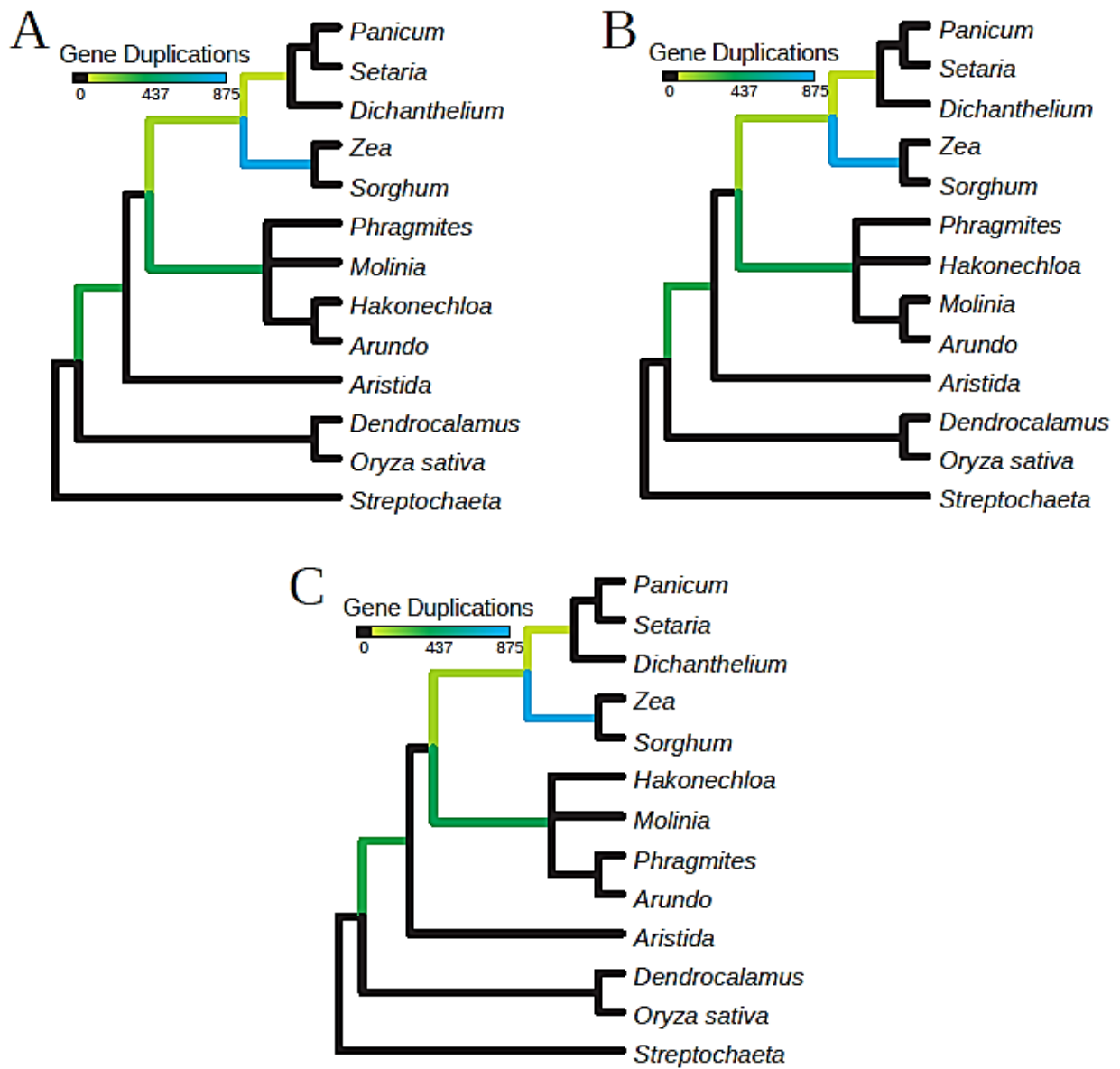


Figure 2.6. PUG results using alternative species tree topologies in which *Arundo* is treated as a sister taxon to one of the other three Arundinoideae in the study. Coloring of branches is the same as in Figure 5.

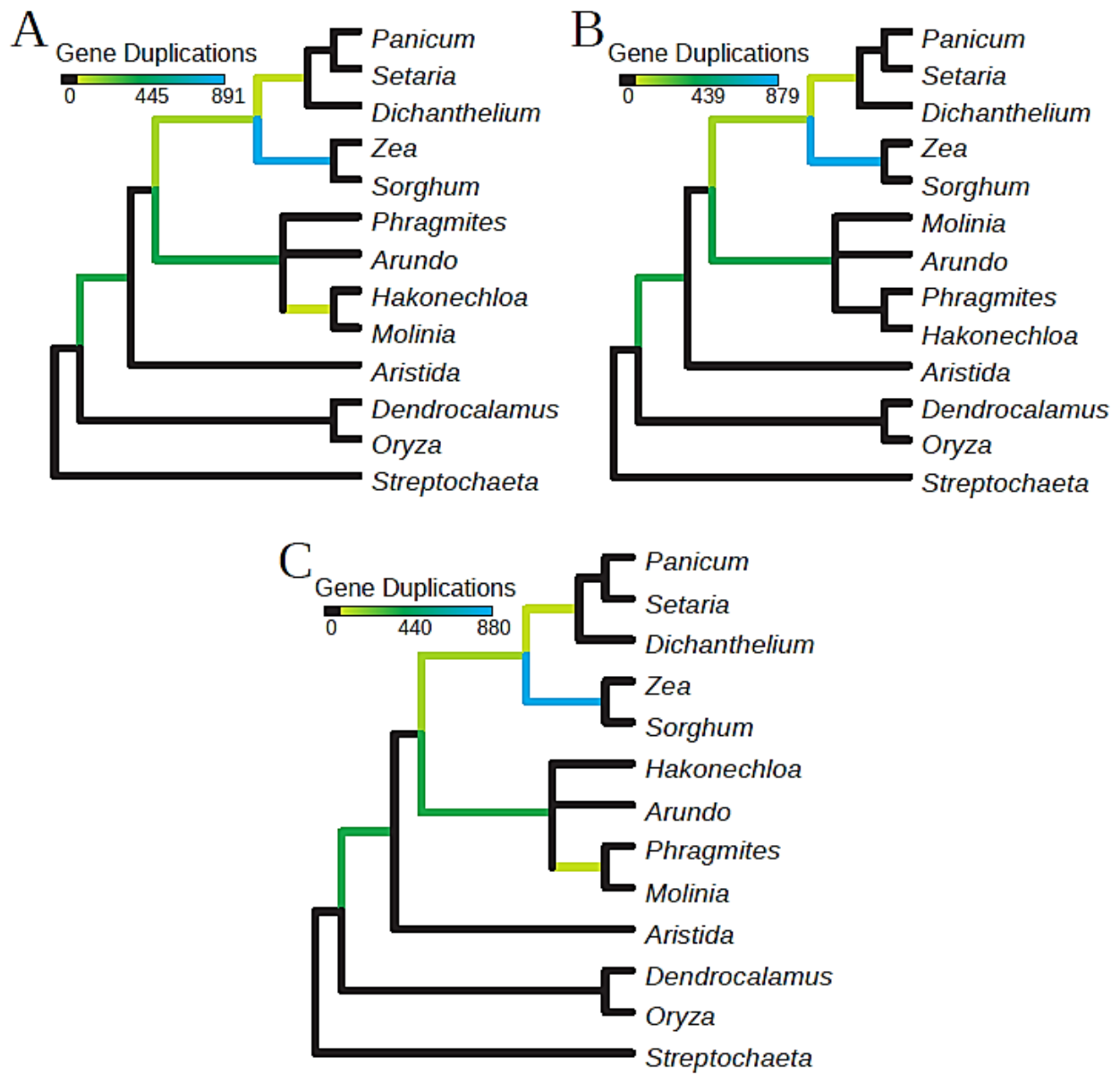


Figure 2.7. PUG results using alternative species tree topologies in which two members of *Phragmites*, *Molinia*, and *Hakonechloa* are treated as sister while the other is left in an unresolved position in Arundinoideae. Coloring is the same as in Figure 5.

Table 2.1. Sources of coding DNA sequences for taxa used in K_s and PUG analyses.

Taxon	Voucher	Source
<i>Arundo donax</i>	Teisher 95	This Study
<i>Arundo donax</i>	Teisher 96	This Study
<i>Hakonechloa macra</i>	Teisher 97	This Study
<i>Hakonechloa macra</i>	Teisher 99	This Study
<i>Molinia caerulea</i>	Teisher 98	This Study
<i>Molinia caerulea</i>	Teisher 100	This Study
<i>Phragmites australis</i>	Teisher 101	This Study
<i>Phragmites australis</i>	Teisher 102	This Study
<i>Aristida stricta</i>		McKain <i>et al.</i> , 2016 Data from SSRA, Assembly from Mckain <i>et al.</i> , 2016
<i>Dendrocalamus latiflorus</i>		
<i>Dichantherium</i>		
<i>oligosanthes</i>		Steuder <i>et al.</i> , 2016
<i>Oryza sativa</i>		Phytosome 10
<i>Panicum virgatum</i>		Phytosome 10
<i>Setaria viridis</i>		Phytosome 10
<i>Sorghum bicolor</i>		Phytosome 10
<i>Streptochaeta</i>		
<i>angustifolia</i>		McKain <i>et al.</i> , 2016
<i>Zea mays</i>		Phytosome 10

Table 2.2. Summary of translation statistics for four arundinoid transcriptomes.

Taxon	Total Assembled Bases	Total Trinity Transcripts	N50	Median Contig Length	Average Contig Length
<i>Arundo donax</i>	41,193,831	53,246	1,212	462	774
<i>Hakonechloa macra</i>	25,960,605	26,812	1,710	522	968
<i>Molinia caerulea</i>	35,585,490	50,897	1,083	405	699
<i>Phragmites australis</i>	31,575,813	41,192	1,248	432	767

CHAPTER 3

Evolution of C₄ Photosynthesis in the Micrairoideae (Poaceae)

3.1 INTRODUCTION

The over 60 independent origins of C₄ photosynthesis collectively constitute one of the most striking examples of parallel evolution in plants (Sage *et al.*, 2011). In particular, the grass family (Poaceae) alone contains at least 22 separate transitions from C₃ to C₄ photosynthesis (GPWGII, 2011), with over 4,500 C₄ species including such major crops as maize, sorghum, and sugar cane (Brown, 1999). This pathway involves increasing the concentration of carbon dioxide around the carbon-fixing enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) and restricting the expression of the enzyme to specialized cells around the vasculature, thereby maximizing the enzyme's efficiency (Sage, 2004; Kellogg, 2013). The efficiency of the C₄ pathway has generated considerable interest in understanding how it has evolved, not only because C₄ species are such dominant members of so many ecosystems, but also due to the potential to increase crop yield via genetically engineering C₃ crops, such as rice and soybean, to use C₄ (i.e. Sage & Zhu, 2011; Denton *et al.*, 2013; Slewinski, 2013; Wang *et al.*, 2014).

Exploring both the commonalities and unique qualities of the various C₄ origins is vital to understanding the role this complex trait has played in evolutionary history and how it can be exploited to benefit humanity. Several mechanisms have been proposed to explain the apparent ease with which the transition to C₄ is made, including conserved regulatory elements in key genes (Brown *et al.*, 2011), lateral gene transfer from C₄ to C₃ taxa (Christin *et al.*, 2012), and anatomical preadaptations (Christin *et al.*, 2013). Other studies have highlighted the differences between various C₄ lineages, demonstrating that extensive physiological and ecological variation can exist between plants using this pathway (Sinha & Kellogg, 1996; Liu & Osborne, 2015).

The advantages of C₄ photosynthesis over the ancestral C₃ pathway under certain environmental conditions have been well-documented. Water use efficiency is generally higher

in C₄ plants (Kokacinar, 2015), giving them a competitive advantage in arid environments and under more intense light (Taylor *et al.*, 2014). Edwards *et al.* (2010) showed that C₄ photosynthesis represents an adaptation to open habitats and a potential preadaptation to arid ones. Edwards & Still (2008) similarly found that C₄ grasses in Hawaii have an adaptive advantage in dry habitats, but that their tendency to occupy warmer areas could simply be the result of evolutionary history in that the C₃ relatives of these taxa are also warm-adapted. Spriggs *et al.* (2014) found that speciation rates are higher on average in C₄ taxa than in their sister C₃ clades, but similarly emphasize the role that historical contingency plays in shaping subsequent evolutionary events. Increases in diversification rates are frequently associated with C₄ origins following a lag period, suggesting that the pathway interacts with other factors to influence speciation and/or extinction rates.

Subfamily Micrairoideae contains a particularly poorly-studied origin of C₄ photosynthesis that seems to defy generalizations. The subfamily contains about 188 species divided into three tribes: Micraireae (15 species in one genus), Isachneae (119 species in 6 genera), and Eriachneae (50 species in 2 genera) (Kellogg, 2015). Eriachneae is C₄ and is sister to Isachneae (Sanchez-Ken *et al.*, 2007; GPWGII, 2011), which is C₃, so at least in this system C₄ photosynthesis has not led to a net increase in species diversification. This is not to say that change in photosynthetic pathway has not facilitated diversification at all, however. There is no logical reason an adaptive radiation must lead to a large number of species. If a novel trait opens a new range of habitats, one might expect the lineage possessing the trait to fill those habitats relatively quickly. Should the number of habitat types be small, the expected number of species would also be small, assuming close relatives compete with each other and cannot infinitely partition resources through specialization. Thus, the acquisition of C₄ photosynthesis in

Eriachneae may have allowed members of this tribe to occupy new habitats, leading to an adaptive radiation even if the total number of species is not very high.

This explanation is also complicated in the Micrairoideae, however, in that the direction of habitat evolution appears to be reversed compared to the more general story in grasses (Spriggs *et al.*, 2014). The C₃ Tribe Micraireae is sister to the rest of the subfamily and is characterized by short, densely mat-forming species with spiral phyllotaxis (Philipson, 1935), which is one of only two known occurrence of this trait in the entire grass family, the other being the South American genus *Arundoclaytonia* (Davidse & Ellis, 1987). This moss-like growth habit likely serves to reduce water loss (Glime, 2015), and, coupled with the ability to resurrect after dehydration (Gaff & Latz, 1978), helps the species of *Micraira* survive in open habitats with infrequent water. These habitats are also typically dominated by C₄ species, including members of tribe Eriachneae. Thus, both C₃ and C₄ Micrairoideae occur in open arid habitats, and this appears to be the ancestral condition in the subfamily. Alternatively, Eriachneae and Micraireae may have both adapted to dry open habitats independently, with Isachneae retaining ancestral habitat preference for shadier and wetter environments.

Some basic questions regarding photosynthetic pathway in the Micrairoideae need to be answered before more specific evolutionary hypotheses can be tested. First, it is common practice to assume that all members of a genus share the same photosynthetic pathway in the absence of evidence to the contrary (i.e. Osborne *et al.*, 2014). The C₄ pathway is reported as being limited to members of Eriachneae, but only a few species in this tribe and in Isachneae have been examined for this trait (Smith & Brown, 1973; Brown, 1977; Ehleringer *et al.*, 1987). Additionally, members of Isachneae possess chlorenchyma that radiates out from the vascular bundle (Watson & Dallwitz, 1992 onwards), a trait typically affiliated with C₄ photosynthesis

(Lundgren *et al.*, 2014). Thus, clarifying the phylogenetic boundaries of C₄ is a necessary first step toward further characterization of its effects on evolutionary dynamics in the subfamily.

Carbon isotope ratios have been shown to be reliable and convenient indicators for discrimination between C₃ and C₄ plants (O'Leary, 1983; Farquhar *et al.*, 1989) and for measuring water use efficiency in general (Maguas & Griffiths, 2003; Caemmerer *et al.*, 2014). The source of this ability stems from the relative preference of the primary carbon-fixing enzymes in C₃ versus C₄ pathways. Rubisco, which is responsible for initial carbon dioxide capture in the mesophyll of C₃ plants, preferentially binds molecules with ¹²C rather than ¹³C (O'Leary, 1988). In the case of the C₄ pathway, phosphoenolpyruvate carboxylase (PEPC) does not discriminate as strongly against C¹³-containing molecules, so both isotopes become fixed and transported to the bundle sheath cells. Within these cells, the selectivity of Rubisco is eventually overcome as the ratio of ¹²C:¹³C decreases. Thus, when carbon from both kinds of plants is compared to the atmospheric carbon isotope ratio using known standards (a value known as δ¹³C and measured in parts per thousand = per mil), material coming from a C₄ plant will have a less negative value as compared to that from a C₃ plant. Specifically, two non-overlapping ranges of δ¹³C are found: -20 to -9 per mil with an average of -14 per mil for C₄ and -35 to -21 per mil with an average of -28 per mil for C₃ plants (O'Leary, 1988). This feature is particularly useful for identifying C₄ species from dried material, for which anatomical details of the leaves can be difficult to discern. Carbon isotope ratios can be measured from herbarium specimens using very little material (Dawson *et al.*, 2002), making this technique ideal for the current study.

A second issue confronting an exploration of C₄ evolution in Micrairoideae stems from the lack of quantitative analyses of geographic distribution patterns in the subfamily. Are the habitats of Eriachneae and Micraireae quantitatively more similar than those of Isachneae? Does

the broad geographic distribution of Isachneae correspond to a similarly broad range of habitat preferences as compared to its more narrowly-distributed sister Eriachneae? An inordinate number of variables contribute to "habitat", and it can be difficult to determine a priori which of these variables will be important in species distributions. However, geographic analysis of variables related to precipitation and temperature can be a valuable and expedient method of quantifying patterns in species distributions (Barbet-Massin & Jetz, 2014; Duan *et al.*, 2014). The WorldClim database maintains a 30 square arcsecond raster of 19 BioClim variables that can be extracted using geographical coordinates (Hijmans *et al.*, 2005), which can be downloaded from the Global Biodiversity Inventory Facility (GBIF.org) for many species in the Micrairoideae.

Perhaps the greatest obstacle to understanding C₄ evolution in Eriachneae is the lack of a substantive phylogeny for the tribe. The largest sampling to date of Micrairoideae for molecular phylogenetics was conducted by GPWGII (2011) and included six species of *Eriachne*, four species of Isachneae, and two species of *Micraira* in a phylogeny of the entire grass family based on three chloroplast genes. Their phylogeny and the one in Chapter 1 of this dissertation confirm the monophyly of Micrairoideae and sister relationship with Arundinoideae identified by Sánchez-Ken *et al.* (2007). A broader taxon sampling within *Eriachne* is needed to test hypotheses of an adaptive radiation associated with the acquisition of the C₄ pathway.

In this study, I explore the evolution of habitat occupation in Micrairoideae with emphasis on the C₄ tribe Eriachneae. First, I clarify the patterns of bioclimatic preference and photosynthetic pathway among species in the three tribes of this subfamily using BioClim data from the WorldClim database as well as carbon isotopes from dried leaf samples. Then I test the hypothesis that C₄ photosynthesis has acted as a key innovation in the Micrairoideae using a

phylogeny of 62 whole-plastome samples representing 30 species, including 53 new plastomes of 24 species of *Eriachne*. This phylogeny is the first to contain a significant sampling of the Eriachneae and thus constitutes a substantial step forward toward understanding this unique C₄ lineage.

3.2 MATERIALS AND METHODS

3.2.1 Collection of Material

Two collecting trips were undertaken in Northern Territory and Western Australia to gather material of the genus *Eriachne* (Figure 3.1). These locations were chosen because they maximize the number of species available in a minimal geographic range. Leaf material was dried in the field in silica gel or salt, with no detectable difference in DNA quality (Carrió & Rossello, 2014). In total, 76 specimens from at least 23 species of *Eriachne* were collected, of which 48 specimens representing all species were included in the phylogenetic analysis.

Additional samples were received from T. Columbus at Rancho Santa Ana Botanic Garden and B.K. Simon, formerly of the Queensland Department of Environment and Resource Management. Plastomes for *Pheidochloa* and several outgroups were used from Chapter 2 of this dissertation. We also included two plastomes from *E. mucronata* and *E. stipacea* available on GenBank (Table 3.1), totaling 63 samples for phylogenetic analysis.

3.2.2 DNA Isolation and Sequencing

Total DNA was extracted from field-dried material using either a QIAGEN EasyDNA Plant Mini Kit or a modified CTAB protocol (Cota-Sánchez *et al.*, 2006), with no consistent differences in DNA quality detected between the two approaches. DNA was mechanically

sheared using a Covaris S220 sonicator under the following conditions: peak power 175, duty factor 5.0, 200 cycles for 30 seconds. Fragments of size 400-500bp were isolated and purified using AMPure XP Beads (Beckman Coulter, Inc.), and a NEBNext Ultra DNA Library Prep Kit for Illumina (New England BioLabs, Inc.) was used to prepare libraries according to the manufacturer's instructions. The resulting libraries were sequenced using an Illumina 2x250 paired-end HiSeq run at the University of Illinois at Urbana-Champaign Roy J. Carver Biotechnology Center.

3.2.3 Plastome Assembly

Plastome assemblies were performed either on the Apollo Cluster at the Donald Danforth Plant Science Center or on Google Cloud. Raw sequence reads were cleaned with Trimmomatic version 0.32 for TruSeq3-PE adapters, using a sliding window of 10bp with a cutoff phred score of 20 and keeping fragments of minimum length 40 (Bolger *et al.*, 2014). Trimmed fragments were assembled with SPAdes version 3.1.0 using k values of 55, 87 and 121 (Bankevich *et al.*, 2012); the resulting contigs were extended on either end using afin (bitbucket.org/benine/afin) with a stop extension value of 0.1, an initial trim of 100 bp from contigs, a maximum extension of 100 bp per loop, and 50 search loops. Contigs generated by afin were connected by hand in Sequencher version 5.3 (Gene Codes Corporation) by manually searching trimmed reads to connect any remaining fragments. Gaps for which no reads could be found in the final alignment were filled with N's. Boundaries between the quadripartite regions (Large Single Copy – LSC, Inverted Repeat B – IRB, Short Single Copy – SSC, and Inverted Repeat A – IRA) were identified, and the IR region was duplicated to serve as both IRA and IRB since any differences between these regions cannot be reliably phased. Plastomes were checked for accuracy by

searching a 20 bp sliding window against the trimmed reads for that sample and plotting the resulting coverage. Low coverage areas were compared with other completed plastomes on DOGMA (Wyman *et al.*, 2004) and with the distribution of overlapping trimmed reads to correct any errors in the assembly. Finished plastomes start at the beginning of the Large Single Copy (LSC) and end with IRA. Sequences were annotated in Verdant (McKain *et al.*, submitted; verdant.iplantcollaborative.org) and submitted to GenBank.

3.2.4 Alignment and Phylogenetic Analysis

Plastomes were aligned by quadripartite region with the IRA removed. MAFFT version 7.029b was run on each region using default parameters (Kato, 2013). The three alignments were then combined to form a single alignment. A maximum likelihood tree was calculated using RAxML version 8.0.22 with 500 bootstrap replicates (Stamatakis, 2014), and MrBayes version 3.2.6 (Ronquist & Huelsenbeck, 2003) with 5 million generations was used for Bayesian analysis. Trees were visualized and edited using FigTree version 1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>) and with the *plot.phylo* function in R package **ape** (Paradis & Strimmer, 2004).

3.2.5 BioClimatic Data

Latitude and longitude coordinates were downloaded from the Global Biodiversity Inventory Facility (GBIF.org) for all available species in subfamily Micrairoideae. Quality control was done using Microsoft Excel and the R package **maptools** (Bivand & Lewin-Koh, 2015) to remove entries lacking coordinate data, duplicate entries and those with doubtful coordinates. Additionally, any species represented by fewer than five samples was removed from

the analysis, with the exception of four specimens of *Limnopoa* that were georeferenced approximately based on label information. Nineteen BioClim variables with 30 second resolution were downloaded from the WorldClim database (Hijmans *et al.*, 2005), and values were extracted for cleaned coordinates using the R package **raster** (Hijmans, 2015). Principal components analysis using the *prcomp* function in the **stats** package was performed to reduce the dimensionality of the data. Mean and median values for species were also calculated from the rotated BioClim data set and compared to help account for outlier individual records. Bioclimatic disparity for Eriachneae and Isachneae was compared using a principal coordinates analysis of average species values with the R function *betadisper* in package **vegan** (Oksanen *et al.*, 2016). This function takes a distance matrix, in this case the pairwise distances between species averages from the full set of BioClim variables calculated from GBIF localities above, and constructs a principal coordinate space to maximize the variation in distances that is captured by the fewest possible number of axes. Samples in the analysis are assigned to groups by the user, and the average distances of all samples within each group to that group's centroid (calculated as the spatial median of the group samples) are calculated. These average distances are compared with an ANOVA to test for significant differences in disparity while controlling for sample size.

3.2.6 Carbon Isotope Discrimination

Carbon isotope ratios were measured from field-dried or herbarium material of 26 species of *Eriachne*, two species of *Micraira*, one species each of *Pheidochloa*, *Coelachne* and *Limnopoa*, and 16 species of *Isachne* (Table 3.2). 400 micrograms of each sample were run in two batches of 200 micrograms each to assess measurement replicability, and wherever possible multiple samples per species were included to test the consistency of $\delta^{13}\text{C}$ values between

closely related individuals. All isotopic measurements were conducted in the laboratory of D. Fike in the Department of Earth and Planetary Sciences at Washington University in St. Louis using acetanilide, cellulose, graphite, and sucrose as carbon standards.

3.2.7 Trait Evolution

BioClim data were extracted for each of the samples of *Eriachne* and *Pheidochloa* sampled in the current phylogeny using the same methods as outlined above. Individual variables from these data were mapped onto the plastome phylogeny under a Brownian motion model using the "fastML" method in the *contMap* function in R package **phytools** (Revell, 2012). The outgroup *Arundo donax* was treated as missing data because of the extremely wide climatic range of this species, and species or tribe averages were used for the Isachneae and Micraireae. Bioclimate variables were also fitted to a reduced plastome phylogeny consisting only of Eriachneae using several different models in the *fitContinuous* function in R package **geiger** (Harmon *et al.*, 2008). Akaike Information Criteria (Akaike, 1973), or AIC, generated from this function were compared to evaluate the statistical support for competing evolutionary hypotheses, including Brownian motion (Felsenstein, 1973), early burst (Harmon *et al.*, 2010), delta (Pagel, 1999), and white noise models. These models were chosen to distinguish between a constant rate of evolution (Brownian motion), an increasing or decreasing rate (exponential in the Early Burst model, linear in the delta model), and absence of phylogenetic signal (the white noise model). An adaptive radiation of *Eriachne*, in which species radiate into new habitat types quickly, would imply high evolutionary rates early in the history of the genus and low rates towards the tips of the tree, which would be supported by the early burst model or the delta model with negative values for the change in evolutionary rate of bioclimatic preferences.

Positive values of the rate parameters for these models would indicate that evolution of bioclimatic niche has increased through time, as might be expected under a model of recent climate change. Phylogenetic signal in the bioclimate variables was also estimated and tested using Pagel's lambda (Pagel, 1999) and Blomberg's K (Blomberg *et al.*, 2003) with the *phylosig* function in R package **phytools**.

3.3 RESULTS

3.3.1 Plastome Assembly, Alignment and Phylogenetic Analysis

53 plastomes representing 24 species of *Eriachne* were successfully assembled from Illumina trimmed reads (Table 3.1). The length of the final assemblies was highly consistent across samples, ranging from a low of 134,445 bp to a high of 135,081 bp and with a mean of 134,740 bp. The average single copy coverage within samples ranged from 6X to 283X with a mean across all samples of 52X.

The alignment of all plastomes is 123,690 bp long including the IR region only once. Maximum likelihood and Bayesian analyses of the unedited alignment produced trees with nearly identical topology and with strong support across most of the backbone and for most species relationships (Figure 3.2). Species in the genus fall out into five morphologically cohesive clades, with *E. compacta* resolved as sister to the rest of the genus. Plant habit, spikelet size and the presence of awns are consistent within these clades, although other traits like life history (annual or perennial) are shared by distantly-related species. The genus *Pheidochloa* is recovered within a paraphyletic *Eriachne*, specifically in a clade with *E. pallescens* and *E. trisetata*.

3.3.2 Habitat Breadth in Micrairoideae

The cleaned matrix of localities from GBIF contained 22,292 specimens in 101 species (Figure 3.3). Principal components analysis yielded two axes that contained a combined 82% of the variance in temperature and precipitation variables among these specimens. Variable loadings for these axes are given in Table 3.3. Values for these first two axes are plotted for all specimens in Figure 3.4A, with specimens colored according to tribe: Eriachneae (red), Isachneae (blue) and Micraireae (green). Micraireae occupies a space closer to Eriachneae, with the exception of *M. subulifolia*, which has a distribution extending along Australia's northeast coast. These results hold true when average values of each principal component for each species are used instead of those for all individuals (Figure 3.4B). When species values are calculated by averaging over the original unrotated data, over 75% of the variation in pairwise distances between those species can be captured in two principal coordinate axes, and only the first four such axes have eigenvalues greater than 100 (Figure 3.5A). Plotting species values in these axes shows that members of Isachneae occupy a broader overall climate space than Eriachneae, a result confirmed by an ANOVA of the average distances from species to the tribe centroid in the rotated space ($F=33.677$, $P=7.238e-08$, $df=1$, Figure 3.4).

3.3.3 Carbon Isotope Ratios

Values for $\delta^{13}\text{C}$ for all samples are given in Table 3.2 and plotted in Figure 3.6. All measured taxa in the Micraireae and Isachneae have values in the typical C_3 range, with a range of -32 to -23 per mil and an average value of -29. Similarly, specimens of Eriachneae possess $\delta^{13}\text{C}$ values in the normal range for C_4 species, ranging from -17 to -11 per mil and with an average of -13.5. Standard deviations between the two replicate measurements for each sample

range from close to 0 to 0.89 with an average of 0.12, with most samples below the recommended standard deviation of 0.20. Considerable variation was found between samples from the same species. In particular, values for two samples of *Isachne mauritiana* differed by more than 7 per mil, though this is an extreme case. The average within-species variance across species with multiple samples is 4.99 for Isachneae and 0.66 for Eriachneae.

3.3.4 Habitat Evolution in Eriachneae

Figure 3.7 shows the range of mean annual temperatures and mean annual precipitation levels occupied by the samples of Eriachneae in the phylogenetic tree versus samples with locality information in GBIF. The phylogeny mostly samples specimens in the hotter and drier range of total Eriachneae values. The values for annual mean temperature and annual mean precipitation are plotted according to a Brownian Model of evolution on the phylogeny of Micrairoideae in Figures 3.8 and 3.9, respectively. These figures indicate that most of the change in these two climate traits occurred at the base of either the Eriachneae or Isachneae or both and that change within Eriachneae was clustered toward the tips, with internal branches within this clade showing relatively low rates of change. This pattern within the tribe is especially apparent when the same values are plotted on a reduced phylogeny, as shown in Figures 3.10 and 3.11.

The results of fitting different models of trait evolution using maximum likelihood are given in Table 3.4. The white noise model, indicating a lack of phylogenetic signal, was the best fit according to Akaike's Information Criterion in 15 out of the 19 BioClim variables, highlighting the large amount of variability in these values within clades and between members of the same species. The other 4 BioClim variables – Isothermality, Mean Temperature of Wettest Quarter, Annual Precipitation, and Precipitation of Wettest Quarter – all have the delta

model as their preferred model, although in the latter two cases it is not significantly better than the Brownian motion model. In all four cases, the value of delta is >1 , indicating that evolution towards the tips of the tree has been faster than towards the root. The early burst model was not favored for any of the climate variables.

As shown in Figure 3.12, the infrageneric classification erected by Lazarides (1995) is at odds with the chloroplast phylogeny. All groups are found to be nonmonophyletic with the exception of Group D, for which only *E. glauca* is included in the phylogeny. A few species are also recovered as nonmonophyletic in the current phylogeny, but some caution is needed in interpreting these results (See the Notes on Classification section in the Discussion).

3.4 DISCUSSION

Unlike most C_4 lineages, Eriachneae has fewer species than its C_3 sister Isachneae and appears to occupy a habitat more similar to the ancestral condition for the subfamily. While the origin of most C_4 lineages correlated with a shift into more open and drier habitats and an increased net diversification rate (Edwards *et al.*, 2010; Spriggs *et al.*, 2014), evolution of the C_4 photosynthetic pathway in tribe Eriachneae of subfamily Micrairoideae appears to oppose these broad generalizations. The results of this study support this story using a survey of carbon isotope ratios across the Micrairoideae, a molecular phylogeny of whole plastome sequences of Eriachneae, and evolutionary analyses of bioclimatic data extracted from the WorldClim database. In the following sections I elaborate on possible interpretations of these results, including limitations of the data and implications for the taxonomy of Eriachneae.

3.4.1 C_4 Photosynthesis and Habitat Breadth in Micrairoideae

The carbon isotope analysis confirms the origin of C₄ photosynthesis in the ancestor of the Eriachneae and rejects the possibility that the radiate chlorenchyma in Isachneae contribute to an intermediate C₃-C₄ pathway. Additionally, $\delta^{13}\text{C}$ values for *Micraira* fall within the range for Isachneae, despite the fact that members of this genus occupy habitats with more similar bioclimatic profiles to the C₄ Eriachneae (Figure 3.4). *Micraira's* moss-like habit, resurrection abilities, and ability to grow in very shallow soils on rocks likely help its species live under climate conditions that are otherwise unfavorable to C₃ grasses (Philipson, 1935; Gaff & Latz, 1978). The most parsimonious interpretation is that the ancestor of the Micrairoideae occupied relatively hot and dry climates rather than that both Micraireae and Eriachneae moved into such climates independently. However, a broader phylogenetic sampling in Isachneae and Micraireae would be needed to attempt a formal ancestral state estimation of climatic niche, especially given the wide range of climates occupied by Isachneae.

The wider bioclimatic niche breadth found in the Isachneae as compared to the Eriachneae (Figures 3.5) also contradicts the pattern found by Christin & Osborne (2014) in Hawaiian Paniceae. In their study, C₃ lineages remained inside a relatively narrow climate space as compared to their C₄ sister lineages, which diversified into a wide range of habits from deserts to tropical rainforests. In contrast, C₄ Eriachneae occupies habitats more similar to the inferred ancestral ones based on its proximity to the Micraireae in BioClim space, while C₃ Isachneae appears to have diversified following a shift in preferred climate. The radiate chlorenchyma shared by both tribes may have been an "anatomical enabler" of the C₄ pathway (Christin *et al.*, 2013), but its contribution to evolutionary success in Isachneae is unclear. However, significantly more work on the Isachneae is required to test this scenario, as there is no taxonomic revision or phylogeny for the genus *Isachne* across its full range.

3.4.2 Evolution of Habitat Preference in Eriachneae

Evolutionary model fitting of the bioclimate variables does not support the hypothesis of an adaptive radiation following acquisition of C₄ photosynthesis in Eriachneae. Most of the variables explained well by a model lacking any phylogenetic structure, and the few variables that show such structure are better explained by an evolutionary model with increasing rates through time. If bioclimatic niches were evolving according to an adaptive radiation model, we would expect either the Early-burst or delta model to be favored with decreasing rates through time, signifying an early filling of novel niches made available by the acquisition of C₄ photosynthesis followed by comparatively small modifications once these niches were occupied.

One possible explanation for the unimpressive diversification associated with the C₄ pathway in Eriachneae may be lack of sufficient time. As noted by Spriggs *et al.* (2014), increases in diversification rate are often separated from C₄ origins by a significant lag period. Christin *et al.* (2008) estimated the split between Eriachneae and Isachneae as occurring approximately 11 mya, making it one of the younger C₄ clades in the PACMAD grasses (Vincentini *et al.*, 2008). Perhaps the necessary conditions that would lead to increased diversification rate in this clade simply have not had time to arise. Such interpretations need to be treated with caution, however, as too much flexibility in a model makes its rejection difficult or impossible. Also, diversification dates for grasses vary considerably due to a lack of reliable fossils for calibration (Christin *et al.*, 2014), so hypotheses requiring accurate absolute ages are perhaps not well-suited to the family.

Another possibility is that not all C₄ types are equally prone to diversification. Eriachneae possess a unique form of C₄ that couples use of NADP-ME for decarboxylation in the bundle

sheath with presence of two well-defined bundle sheaths in which chloroplast density in the outer sheath is extremely high (Sinha & Kellogg, 1996). This type has not been studied extensively, so whether it involves strong fitness trade-offs that prevent it from outcompeting other grasses is unknown. Additionally, C_4 taxa are compared to their closest C_3 relatives, which in this case possess C_4 -like radiate chlorenchyma (Watson & Dallwitz, 1992 onwards). No difference in water use efficiency between *Isachneae* and *Micraireae* was detectable from carbon isotopes, but given the high levels of intraspecific and even intrasample variation and the small sample size in *Micraireae*, differences that are difficult to measure but might be biologically meaningful might not be visible with the current data set.

Lastly, several caveats need to be addressed regarding evolutionary analysis of bioclimatic data. First, as shown in Figure 3.7, only a portion of the climatic ranges occupied by *Eriachneae* is represented by the current phylogenetic sampling. Theoretically, this sampling error could prevent the recognition of major evolution along early-diverging branches in the tree. This is unlikely to be the case in the current study, as the most extreme climatic outliers are represented by widespread species included in our phylogeny, such as *E. trisetata* and *E. burkittii*. If representatives of these taxa from significantly different habitats were added to the tree, they would only increase the amount of variation among the tips within clades. On the other hand, the fact that a large portion of the variance in climatic variables is contained within species means that taxa represented by a single specimen in the current tree may have values that are not representative of most of that species' members.

Another major issue involved in analysis data from the BioClim database is limited resolution of the raster. Species may occupy microhabitats with substantially different climatic conditions as compared to the average value for the surrounding square kilometer, for example a

short-lived annual plant growing in vernal pools. With very large samples, this discrepancy can potentially be accounted for by using mean or median species values, but in a phylogeny of moderate size such control is not possible. Large measurement error of tip values can cause artificial reduction of phylogenetic signal (Silvestro *et al.*, 2015).

These complications may account for the frequency with which the white noise model is chosen as the most likely among the bioclimate variables. Measurement error could also explain why the delta model is supported in some variables showing significant phylogenetic signal and why the value of delta is positive. More variation at the tips of the tree would tend to lead to an increase in the inferred evolutionary rate among shallow branches, which could resemble an overall accelerating rate across the tree.

3.4.3 Notes on Classification

In general, the plastome phylogeny reflects morphological themes in the genus reasonably well. The taxonomic groups outlined by Lazarides (1995) display varying degrees of compatibility with molecular data, with his Group A, made up of awnless perennials occurring in arid habitats, having the broadest representation across the phylogeny. The possibly paraphyletic nature of this group was recognized by Lazarides, who thought that the simple morphology and wide climatic ranges of its members could signify an ancestral condition from which the rest of the genus may have emerged. His Group B, consisting of mesophytic long-awned perennials, is represented in the current phylogeny by several closely-related members, including the *schultziiana-stipacea-triodioides* complex. *E. trisetata* and *E. pallescens* are placed in Group B by Lazarides, but fall out in a clade with *Pheidochloa gracilis*, described below. *Eriachne basedowii* is recovered in a somewhat isolated position sister to *E. melicacea-avenacea-agrostidea*, though

this clade does not possess obvious morphological synapomorphies. Group C forms two sister clades in our phylogeny, with three taxa being closely related to members of other groups.

Eriachne armitii was identified by Lazarides as being similar to *E. stipacea*, and this resemblance is supported by the chloroplast tree. It is unclear why the two species were placed in different groups in his classification. *Eriachne nodosa* is compared to *E. major* and *E. obtusa* in terms of its spikelet morphology by Lazarides, and despite differing from these two species in being an annual, it is placed as their sister in the tree. Group D is difficult to evaluate because *E. glauca* is the only representative in our current phylogeny.

The position of *Pheidochloa gracilis* suggests that this genus should be synonymized into *Eriachne*. *Pheidochloa* possesses two species that are distinguished from *Eriachne* by their cylindrical spikelets and caryopses and their markedly unequal glumes (Van Eck-Borsboom, 1980). However, it is like *E. trisetia* and *E. pallescens* in that all three species are slender-culmed, tussock-forming plants with delicate, long-awned spikelets. *Pheidochloa* is unique in this clade in being an annual and in lacking awns on the palea, but these characters vary throughout *Eriachne*.

Eriachne ciliata and *E. semiciliata* form a clade in our tree, with the latter derived from within the former. Lazarides (1995) distinguished the two species based on the relative size of the floret to the glumes, the shape of the glume apex, the apical extent of the lemma margin indumentum, and the orientation of prickles on the culms and foliage, but these characters were inconsistent in our sampling. *Teisher 91* was identified as *E. semiciliata* due to the striking retrorse orientation of the culm and foliage prickles and the markedly more acuminate glumes, but this glume shape was shared by *Teisher 74*, which has antrorsely oriented prickles. The other two samples of *E. ciliata* included in this study match the description given by Lazarides fairly

well except that the indumentum along the lemma margins frequently ends well below the apex, a character that is supposedly restricted to *E. semiciliata*. Our data thus suggest that the two species should be synonymized under the name *E. ciliata*.

A few specimens are difficult to identify or have morphological characters that are inconsistent with their molecular placement. *Teisher 58* was collected in a remote location on the Burrup Peninsula in Western Australia. Unfortunately, the spikelets on this specimen were not well-preserved, so identification is difficult. Its morphology supports placement near the *melicacea-avenacea-agrostidea-basedowii* clade as it shares with most of these species a densely tufted habit with narrow leaves and glumes of medium length. The placement of *Teisher 84* as sister to the *schultziiana-triodioides-stipacea-armitii* complex is a bit odd given the morphology of this specimen, which strongly resembles *E. melicacea* and shares no obvious features with the aforementioned clade. However, its unique and isolated phylogenetic position outside of this complex suggests that the result is not simply a case of sample mix-up. Virtually nothing is known about hybridization in *Eriachne*, so the possibility that the chloroplast phylogeny may be incompatible with certain elements of morphology and physiology is a reasonable one. Two other slightly strange phylogenetic placements are *Thompson GAL360* and *Columbus 5125*. The former is identified as *E. stipacea*, and the latter is identified as *E. aristidea*, but according to the plastid phylogeny they are sister and more closely related to *schultziiana-triodioides-stipacea-armitii*. The vouchers for these specimens were unavailable and thus their identities could not be confirmed. However, *E. aristidea* is a striking and easily identified taxon, and the other two specimens of the species in this study form a closely-related clade, so the possibility that the chloroplast tree is at odds with the species tree again cannot be ignored. *Teisher 82* resembles *E. glauca* on the basis of growth form, lemma awn shape, and spikelet size and indumentum, but

according to the plastome phylogeny it groups very strongly with specimens of *E. obtusa*. The vouchers for *Teisher 56* and *66* did not preserve well, so their identity was left ambiguous, but their vegetative morphology is consistent with that of *E. obtusa*.

3.4.4 Conclusion

The current study confirms an unusual pattern of C₄ evolution in the Micrairoideae. I confirm the greater bioclimatic range of C₃ Isachneae compared to its C₄ sister-taxon Eriachneae and the greater similarity in climatic niche between Eriachneae and the outgroup C₃ tribe Micraireae. I confirm the photosynthetic pathway of all three tribes using carbon isotopic evidence. Thus, the origin of C₄ photosynthesis does not appear to have driven the Eriachneae into a unique climatic zone as compared to sister C₃ taxa, though the range of habitats occupied by Eriachneae is larger and drier than that found in Micraireae. Isachneae may have experienced a parallel radiation into colder and wetter habitats, overshadowing the expansion of Eriachneae driven by C₄ photosynthesis. However, there is little evidence that Eriachneae experienced an adaptive radiation, though these results may at least in part be due to the uncertainties inherent in BioClim data.

LITERATURE CITED

- Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Pp. 267-281 in *2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971*, eds. B. N. Petrov and F. Csáki. Budapest: Akadémiai Kiadó.
- Bankevich, A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, et al. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19(5): 455–477.
- Barbet-Massin, M., and W. Jetz. 2014. A 40-year, continent-wide, multispecies assessment of relevant climate predictors for species distribution modelling. *Diversity & Distributions* 20(11): 1285–1295.
- Bivand, R. and N. Lewin-Koh. 2015. Maptools: tools for reading and handling spatial objects. R package version 0.8-36. <http://CRAN.R-project.org/package=maptools>.
- Blomberg, S. P., T. Garland, and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57:717–745.
- Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15): 2114–2120.
- Brown, N. J., C. A. Newell, S. Stanley, J. E. Chen, A. J. Perrin, K. Kajala, and J. M. Hibberd. 2011. Independent and parallel recruitment of preexisting mechanisms underlying C₄ photosynthesis. *Science* 331(6023): 1436–1439.
- Brown, R. H. 1999. Agronomic implications of C₄ photosynthesis. Pp. 473-508 in *C₄ Plant Biology*, eds. R. F. Sage and R. K. Monson. San Diego, CA: Academic Press.
- Brown, W. V. 1977. The Kranz syndrome and its subtypes in grass systematics. *Memoirs of the Torrey Botanical Club* 23(3): 1-97.
- Caemmerer, S. von, O. Ghannoum, J. J. L. Pengelly, and A. B. Cousins. 2014. Carbon isotope discrimination as a tool to explore C₄ photosynthesis. *Journal of Experimental Botany* 65(13): 1–12.
- Carrió, E., and J. A. Rosselló. 2014. Salt drying: a low-cost, simple and efficient method for storing plants in the field and preserving biological repositories for DNA diversity research. *Molecular Ecology Resources* 14(2): 344–351.
- Christin, P.-A., G. Besnard, E. Samaritani, M. R. Duvall, T. R. Hodkinson, V. Savolainen, and N. Salamin. 2008. Oligocene CO₂ decline promoted C₄ photosynthesis in grasses. *Current Biology* 18(1): 37–43.
- Christin, P.-A., E. J. Edwards, G. Besnard, S. F. Boxall, R. Gregory, E. A. Kellogg, J. Hartwell,

- and C. P. Osborne. 2012. Adaptive evolution of C₄ photosynthesis through recurrent lateral gene transfer. *Current Biology* 22(5): 445–49.
- Christin, P.-A., and C. P. Osborne. 2014. The evolutionary ecology of C₄ plants. *New Phytologist* 204(4): 765–81.
- Christin, P.-A., C. P. Osborne, D. S. Chatelet, J. T. Columbus, G. Besnard, T. R. Hodkinson, L. M. Garrison, *et al.* 2013. Anatomical enablers and the evolution of C₄ photosynthesis in grasses. *Proceedings of the National Academy of Sciences of the United States of America* 110 (4): 1381–86.
- Christin, P.-A., E. Spriggs, C. P. Osborne, C. A. E. Strömberg, N. Salamin, and E. J. Edwards. 2014. Molecular dating, evolutionary rates, and the age of the grasses. *Systematic Biology* 63(2): 153–65.
- Cota-Sánchez, J. H., K. Remarchuk, and K. Ubayasena. 2006. Ready-to-use DNA extracted with a CTAB method adapted for herbarium specimens and mucilaginous plant tissue. *Plant Molecular Biology Reporter* 24(2): 161–67.
- Davidse, G., and R. P. Ellis. 1987. *Arundoclaytonia*, a new genus of the Steyermarkochloaeae (Poaceae: Arundinoideae) from Brazil. *Annals of the Missouri Botanical Garden* 74(3): 479–490.
- Dawson, T. E., S. Mambelli, A. H. Plamboeck, P. H. Templer, and K. P. Tu. 2002. Stable isotopes in plant ecology. *Annual Review of Ecology & Systematics* 33: 507–559.
- Denton, A. K., R. Simon, and A. P. M. Weber. 2013. C₄ photosynthesis: from evolutionary analyses to strategies for synthetic reconstruction of the trait. *Current Opinion in Plant Biology, Physiology and Metabolism* 16 (3): 315–21.
- Duan, R.-Y., X.-Q. Kong, M.-Y. Huang, W.-Y. Fan, and Z.-G. Wang. 2014. The predictive performance and stability of six species distribution models. *PLoS ONE* 9(11): 1–8.
- Edwards, E. J., and C. J. Still. 2008. Climate, phylogeny and the ecological distribution of C₄ grasses. *Ecology Letters* 11(3): 266–76.
- Edwards, E. J., S. A. Smith, and M. J. Donoghue. 2010. Phylogenetic analyses reveal the shady history of C₄ grasses. *Proceedings of the National Academy of Sciences of the United States of America* 107(6): 2532–2537.
- Ehleringer, J. R., Z. F. Lin, C. B. Field, G. C. Sun, and C. Y. Kuo. 1987. Leaf carbon isotope ratios of plants from a subtropical monsoon forest. *Oecologia* 71(1): 109–114.
- Farquhar, G. D., J. R. Ehleringer, and K. T. Hubick. 1989. Carbon isotope discrimination and photosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology* 40: 503–537.

- Felsenstein, J. 1973. Maximum likelihood estimation of evolutionary trees from continuous characters. *American Journal of Human Genetics* 25: 471–492.
- Gaff, D. and P. Latz. 1978. The occurrence of resurrection plants in the Australian flora. *Australian Journal of Botany* 26(4): 485–492.
- Glime, J. M. 2015. Water Relations: Plant Strategies. Chapt. 7-3 in Glime, J. M. Bryophyte Ecology. Volume 1. Physiological Ecology. Ebook sponsored by Michigan Technological University and the International Association of Bryologists. Last updated 19 April 2015 and available at <www.bryoecol.mtu.edu>.
- Grass Phylogeny Working Group II. 2011. New grass phylogeny resolves deep evolutionary relationships and discovers C₄ origins. *New Phytologist* 193(2): 304–312.
- Griffith, D. M., T. M. Anderson, C. P. Osborne, C. A. E. Strömberg, E. J. Forrestel, and C. J. Still. 2015. Biogeographically distinct controls on C₃ and C₄ grass distributions: merging community and physiological ecology. *Global Ecology & Biogeography* 24(3): 304–313.
- Harmon, L. J., J. T. Weir, C. D. Brock, R. E. Glor, and W. Challenger. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics* 24:129–131.
- Harmon, L. J., J. B. Losos, T. J. Davies, R. G. Gillespie, J. L. Gittleman, W. B. Jennings, K. H. Kozak, *et al.* 2010. Early bursts of body size and shape evolution are rare in comparative data. *Evolution* 64: 2385–2396.
- Hijmans, R. J. 2015. Raster: geographic data analysis and modeling. R package version 2.5-2. <http://CRAN.R-project.org/package=raster>.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.
- Katoh, S. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780.
- Kellogg, E. A. 2013. C₄ photosynthesis. *Current Biology* 23: R594–R599.
- Kocacinar, F. 2015. Photosynthetic, hydraulic and biomass properties in closely related C₃ and C₄ species. *Physiologia Plantarum* 153(3): 454–466.
- Lazarides, M. 1995. The genus *Eriachne* (Eriachneae, Poaceae). *Australian Journal of Botany* 8: 355–452.
- Liu, H., and C. P. Osborne. 2015. Water relations traits of C₄ grasses depend on phylogenetic lineage, photosynthetic pathway, and habitat water availability. *Journal of Experimental*

- Botany* 66(3): 761–73.
- Lundgren, M.R., C. P. Osborne, P.-A. Christin. 2014. Deconstructing Kranz anatomy to understand C₄ evolution. *Journal of Experimental Botany* 65: 1–13.
- Máguas, C. and H. Griffiths. 2003. Applications of stable isotopes in plant ecology. *Progress in Botany* 64:472–505.
- Osborne, C. P., A. Salomaa, T. A. Kluyver, V. Visser, E. A. Kellogg, O. Morrone, M. S. Vorontsova, *et al.* 2014. A global database of C₄ photosynthesis in grasses. *New Phytologist* 204(3): 441–46.
- Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, *et al.* 2016. Vegan: community ecology package. R package version 2.3-4.
<http://CRAN.R-project.org/package=vegan>
- O'Leary, M. H. 1988. Carbon isotopes in photosynthesis. *Bioscience* 38(5): 328–336.
- Pagel, M. 1999. Inferring the historical patterns of biological evolution. *Nature* 401: 877–884.
- Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289–290.
- Philipson, W.R. (1935). A grass with spiral phyllotaxis: *Micraira subulifolia*. *Bulletin of Miscellaneous Information (Royal Botanic Gardens, Kew)* 1935(5): 324–326.
- Revell, L. J. 2012. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3: 217–223.
- Ronquist, F. and J. P. Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
- Sage, R. F. 2004. The evolution of C₄ photosynthesis. *New Phytologist* 161(2): 341–370.
- Sage, R. F., P.-A. Christin, and E. J. Edwards. 2011. The C₄ plant lineages of planet earth. *Journal of Experimental Botany* 62(11): 3155–69.
- Sage, R. F., and X.-G. Zhu. 2011. Exploiting the engine of C₄ photosynthesis. *Journal of Experimental Botany* 62(11): 2989–3000.
- Sánchez-Ken, J. G., L. G. Clark, E. A. Kellogg, and E. E. Kay. 2007. Reinstatement and emendation of subfamily Micrairoideae (Poaceae). *Systematic Botany* 32(1): 71–80.
- Silvestro, D., A. Kostikova, G. Litsios, P. B. Pearman, and N. Salamin. 2015. Measurement errors should always be incorporated in phylogenetic comparative analysis. *Methods in Ecology and Evolution* 6: 340–346.

- Sinha, N. R., and E. A. Kellogg. 1996. Parallelism and diversity in multiple origins of C₄ photosynthesis in the grass family. *American Journal of Botany* 83(11): 1458–1470.
- Slewinski, T. L. 2013. Using evolution as a guide to engineer Kranz-type C₄ photosynthesis. *Frontiers in Plant Science* 4: 1–13.
- Smith, B. N., and W. V. Brown. 1973. The Kranz syndrome in the Gramineae as indicated by carbon isotopic ratios. *American Journal of Botany* 60(6): 505–513.
- Spriggs, E. L., P.-A. Christin, and E. J. Edwards. 2014. C₄ photosynthesis promoted species diversification during the Miocene grassland expansion. *PLoS ONE* 9(5): 1–10.
- Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9): 1312–1313.
- Taylor, S. H., B. S. Ripley, T. Martin, L.-A. De-Wet, F. I. Woodward, and C. P. Osborne. 2014. Physiological advantages of C₄ grasses in the field: a comparative experiment demonstrating the importance of drought. *Global Change Biology* 20(6): 1992–2003.
- van Eck-Borsboom, M. H. J. 1980. A Revision of Eriachne R. Br.(Gramineae). *Blumea* 26 (1): 127–138.
- Vicentini, Alberto, Janet C. Barber, Sandra S. Aliscioni, Liliana M. Giussani, and Elizabeth A. Kellogg. 2008. The Age of the Grasses and Clusters of Origins of C₄ Photosynthesis. *Global Change Biology* 14 (12): 2963–77.
- Wang, L., R. A. Mertz, L. K. Dedow, D. W. Bryant, S. Weissmann, A. Studer, Y. Shao, *et al.* 2014. Comparative analyses of C₄ and C₃ photosynthesis in developing leaves of maize and rice. *Nature Biotechnology* 32(11): 1158–65.
- Watson, L., and Dallwitz, M.J. 1992 onwards. The grass genera of the world: descriptions, illustrations, identification, and information retrieval; including synonyms, morphology, anatomy, physiology, phytochemistry, cytology, classification, pathogens, world and local distribution, and references. Version: 7th December 2015. <http://delta-intkey.com>.
- Wyman, S. K., R. K. Jansen, and J. L. Boore. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* 20(17): 3252–3255.

FIGURES AND TABLES

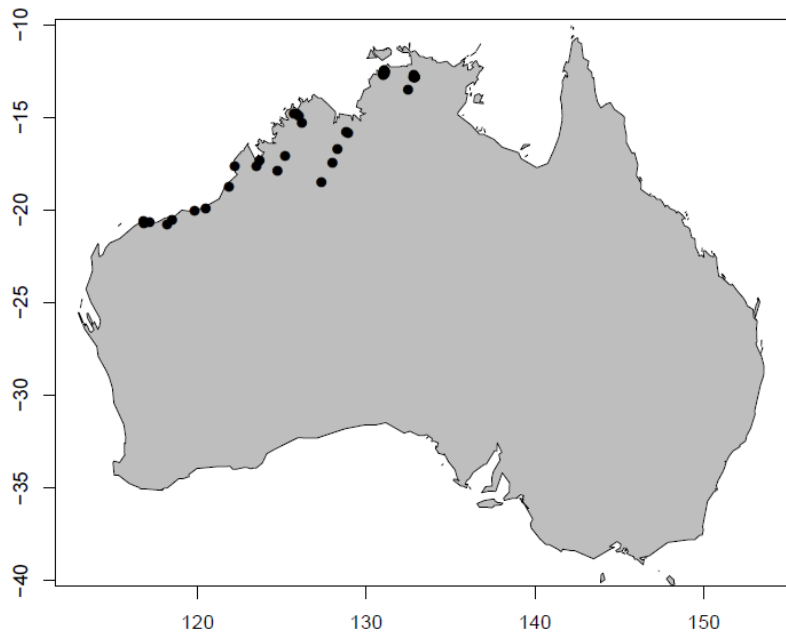


Figure 3.1. Map of collecting sites during two field expeditions in Northern Territory and Western Australia, Australia.

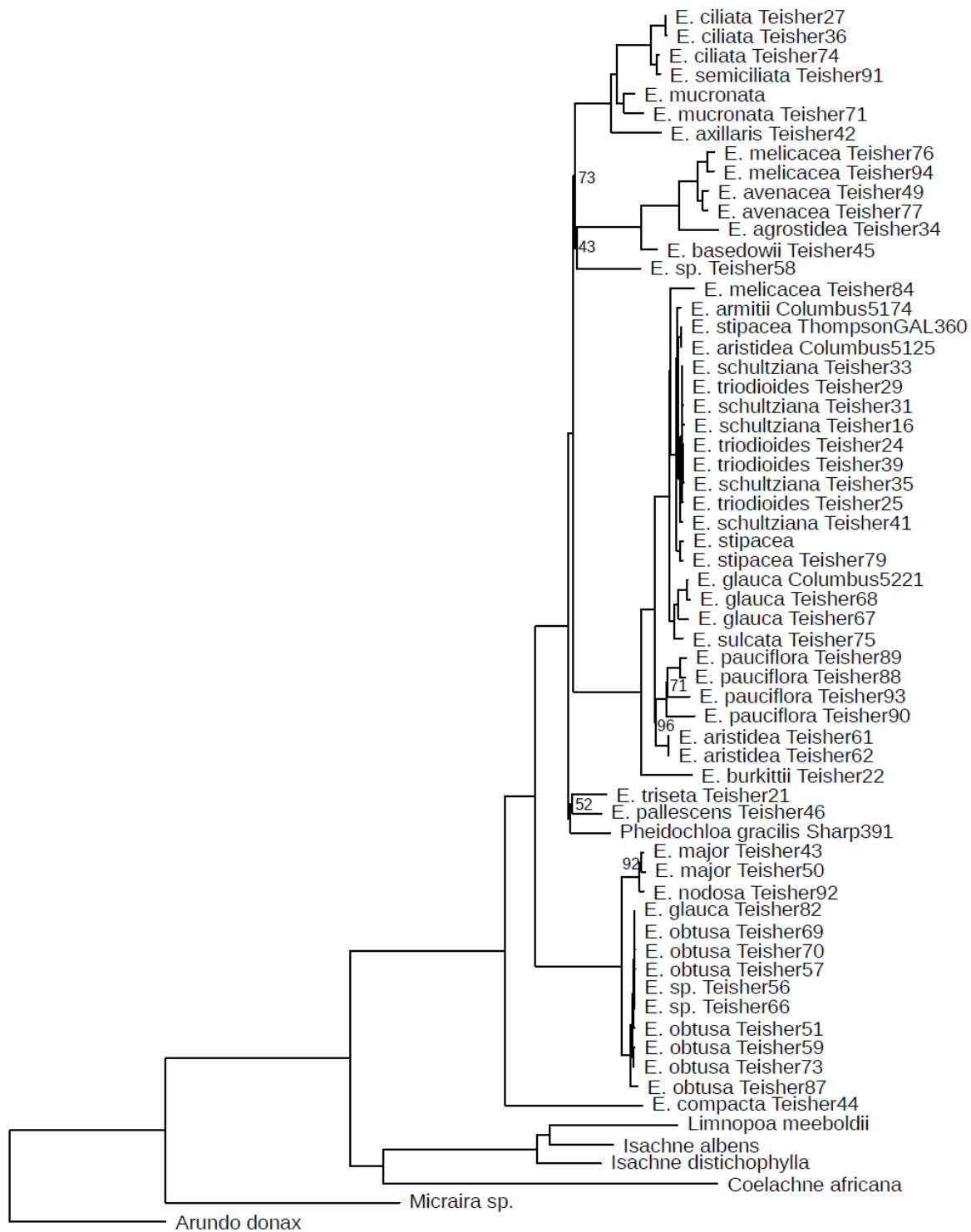


Figure 3.2. Maximum-likelihood phylogeny of Eriachneae based on whole-chloroplast genome sequences with *Arundo donax* (Arundinoideae) as an outgroup. Numbers next to nodes represent bootstrap values using 500 replicates. Nodes without numbers have 100% bootstrap support, except in the *E. schultziiana* and *E. obtusa* clusters, which have lower values that were removed for clarity.

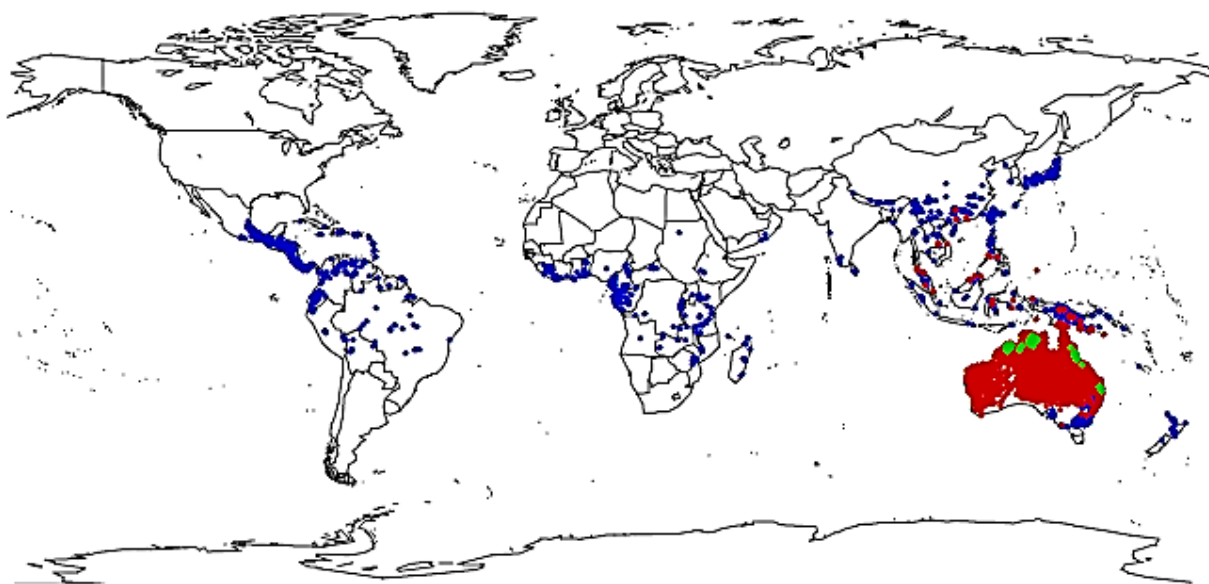
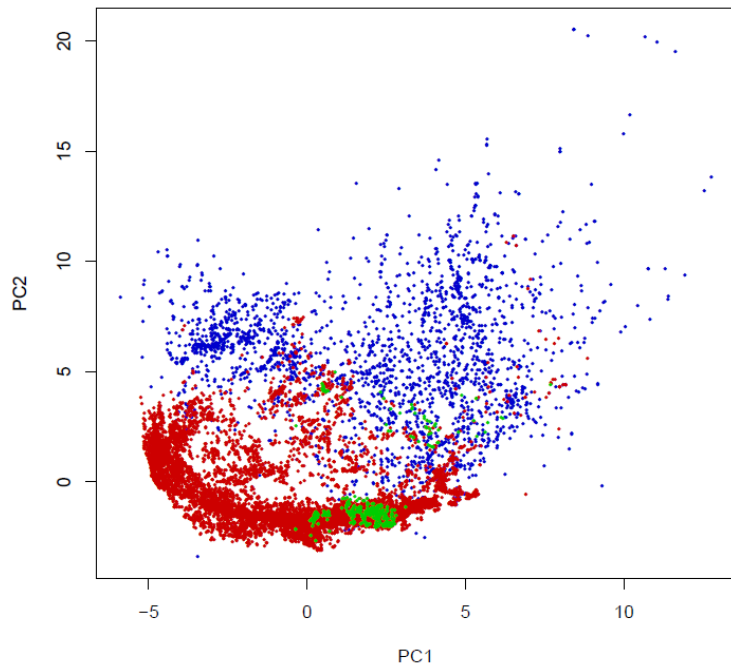


Figure 3.3. Map of 22,292 localities of 101 species of Isachneae (blue), Eriachneae (red), and Micraireae (green) downloaded from GBIF.

A.



B.

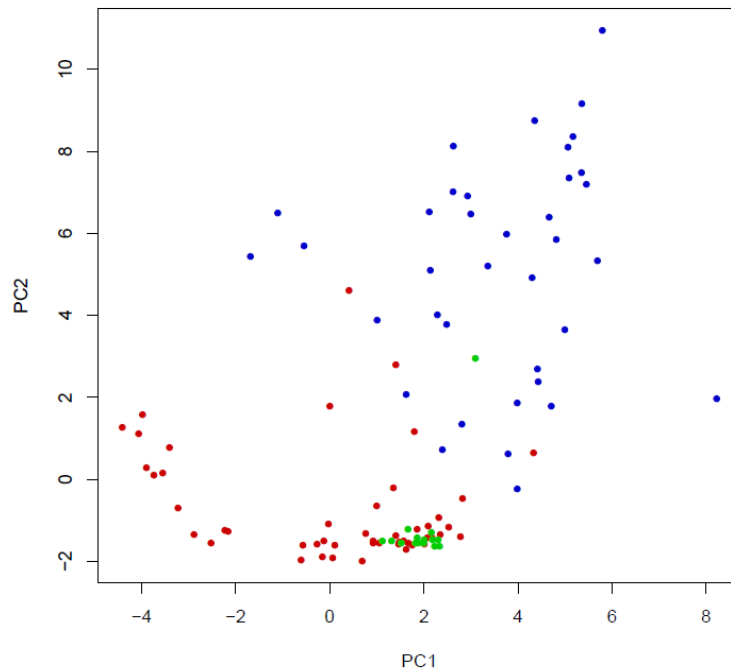


Figure 3.4. A) First and second principal components of 19 climatic variables extracted for 22,292 GBIF localities in Isachneae (blue), Eriachneae (red), and Micraireae (green). B) The same two axes with the average value plotted for each species in A.

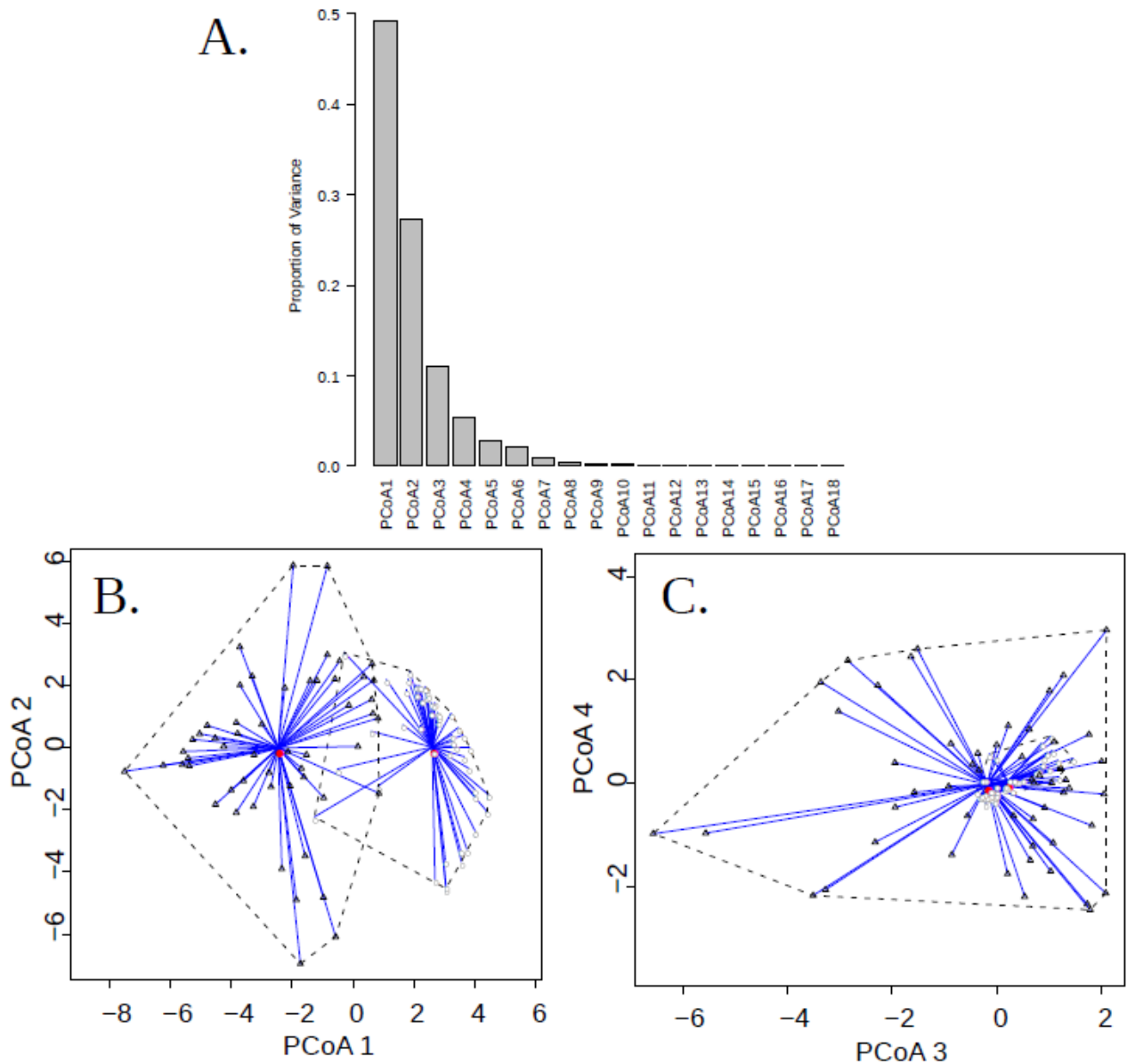


Figure 3.5. Results of disparity analysis using R package *vegan*. A) Scree plot of the proportion of variance between species contained within each principal coordinate axis. B) Principal coordinate axes 1 and 2 plotted for species of *Isachneae* (open triangles) and *Eriachneae* (open circles) with distances to tribe spatial median shown by blue lines. C) Same as B for principal coordinate axes 3 and 4.

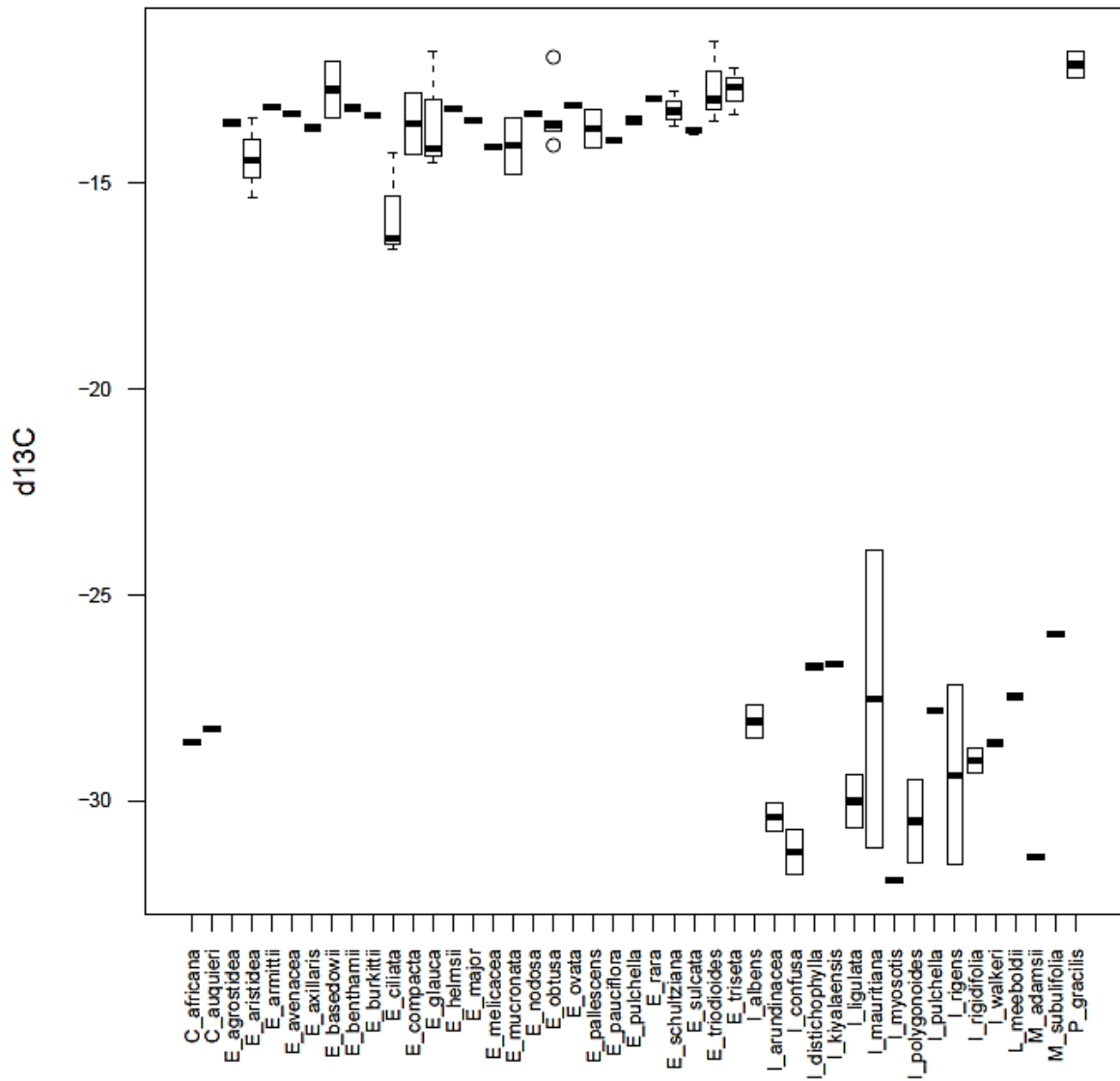


Figure 3.6. Carbon isotope ratios for 26 species of *Eriachne*, two species of *Micraira*, one species each of *Pheidochloa*, *Coelachne* and *Limnopoa*, and 16 species of *Isachne*.

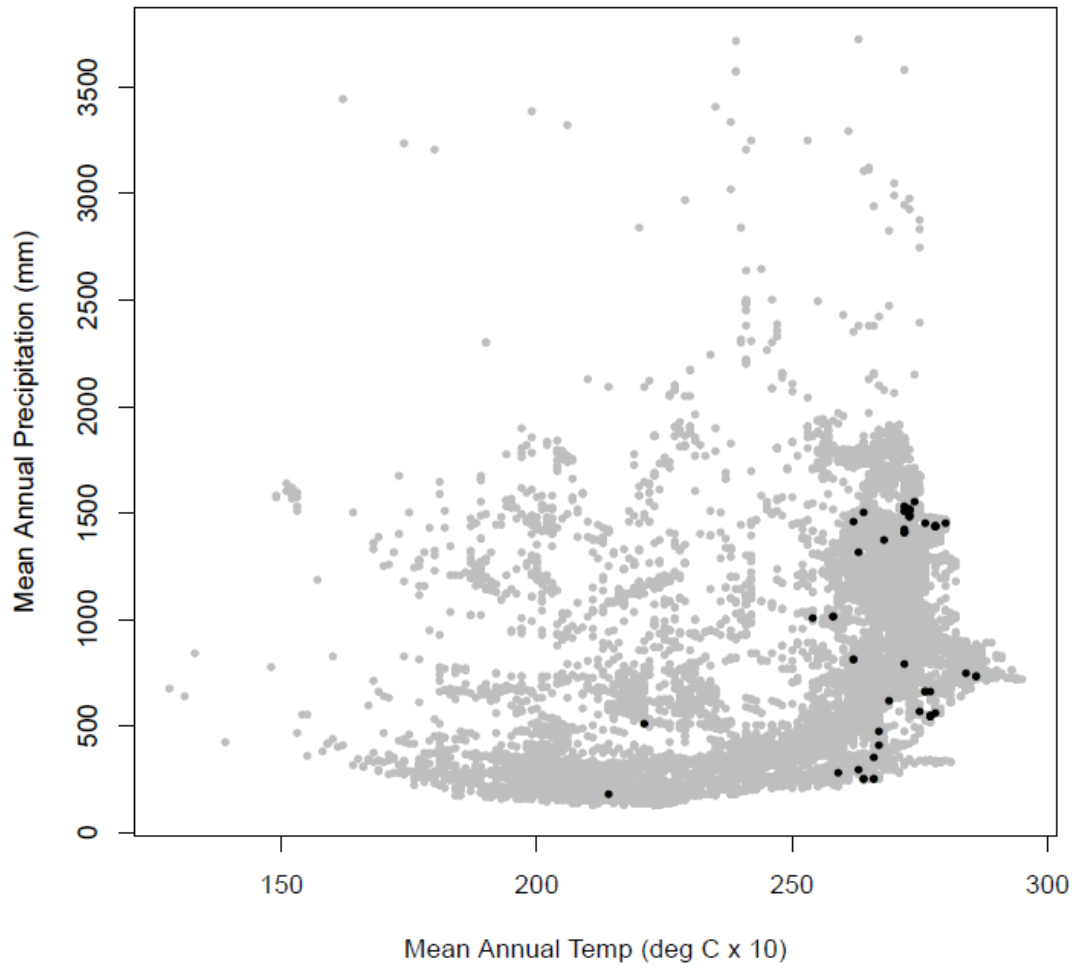


Figure 3.7. Values for mean annual temperature and mean annual precipitation plotted for Eriachneae samples in the phylogeny (black) and from GBIF (grey).

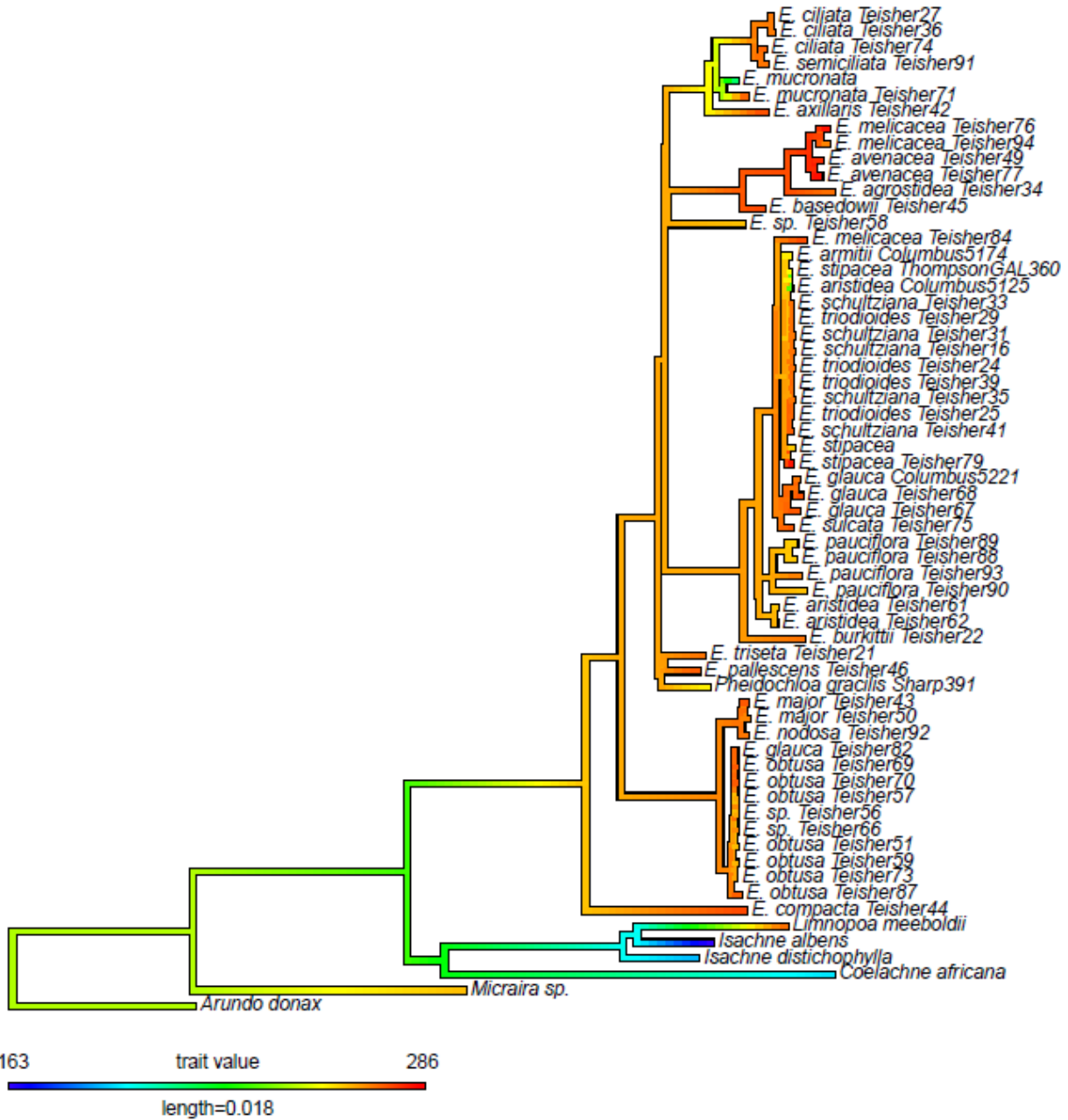


Figure 3.8. Mean annual temperature plotted across the full Maximum-likelihood phylogeny under a Brownian motion model of evolution using the function *contMap* in R package *phytools*.

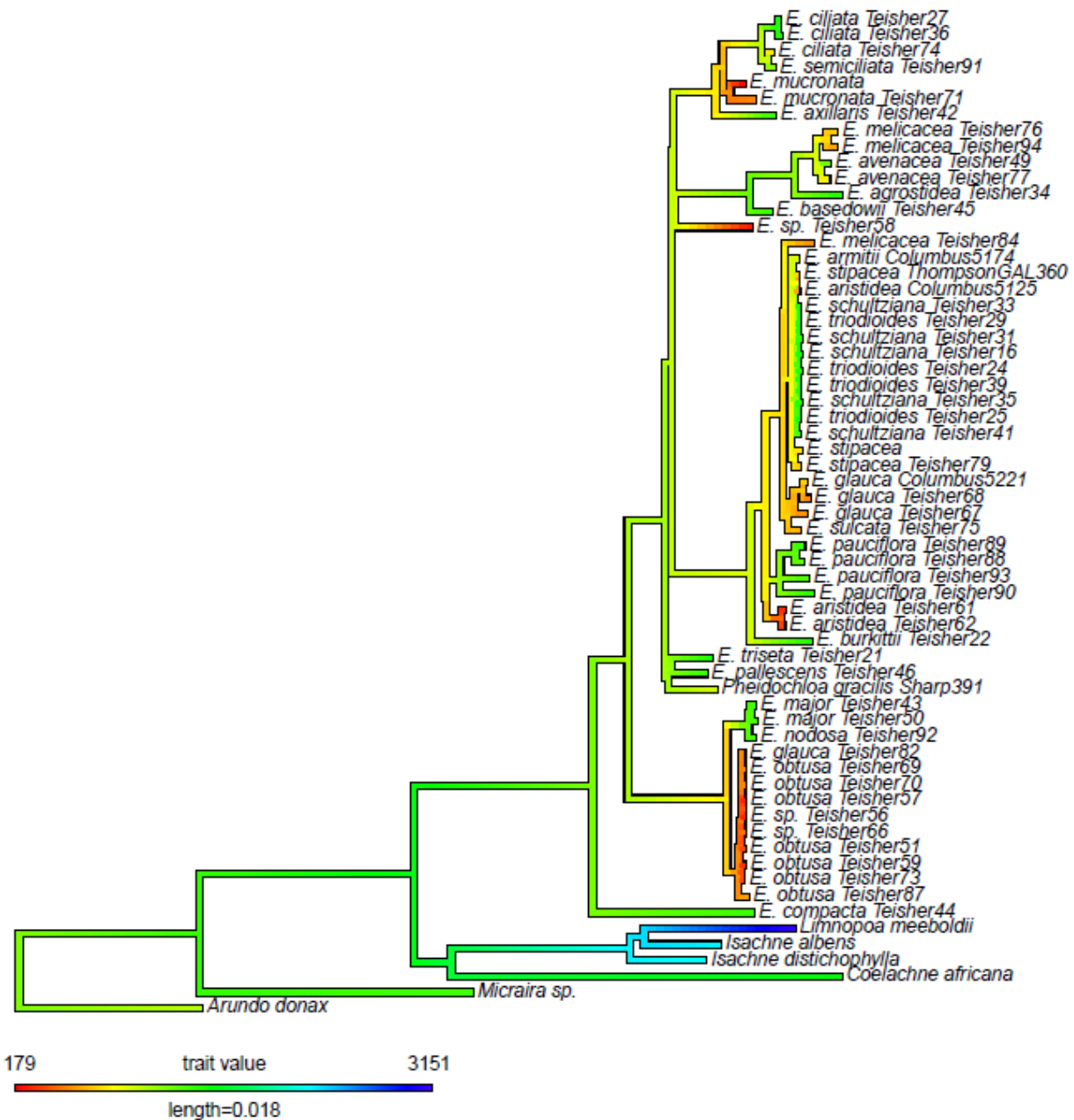


Figure 3.9. Mean annual precipitation plotted across the full maximum-likelihood phylogeny under a Brownian motion model of evolution using the function *contMap* in R package *phytools*.

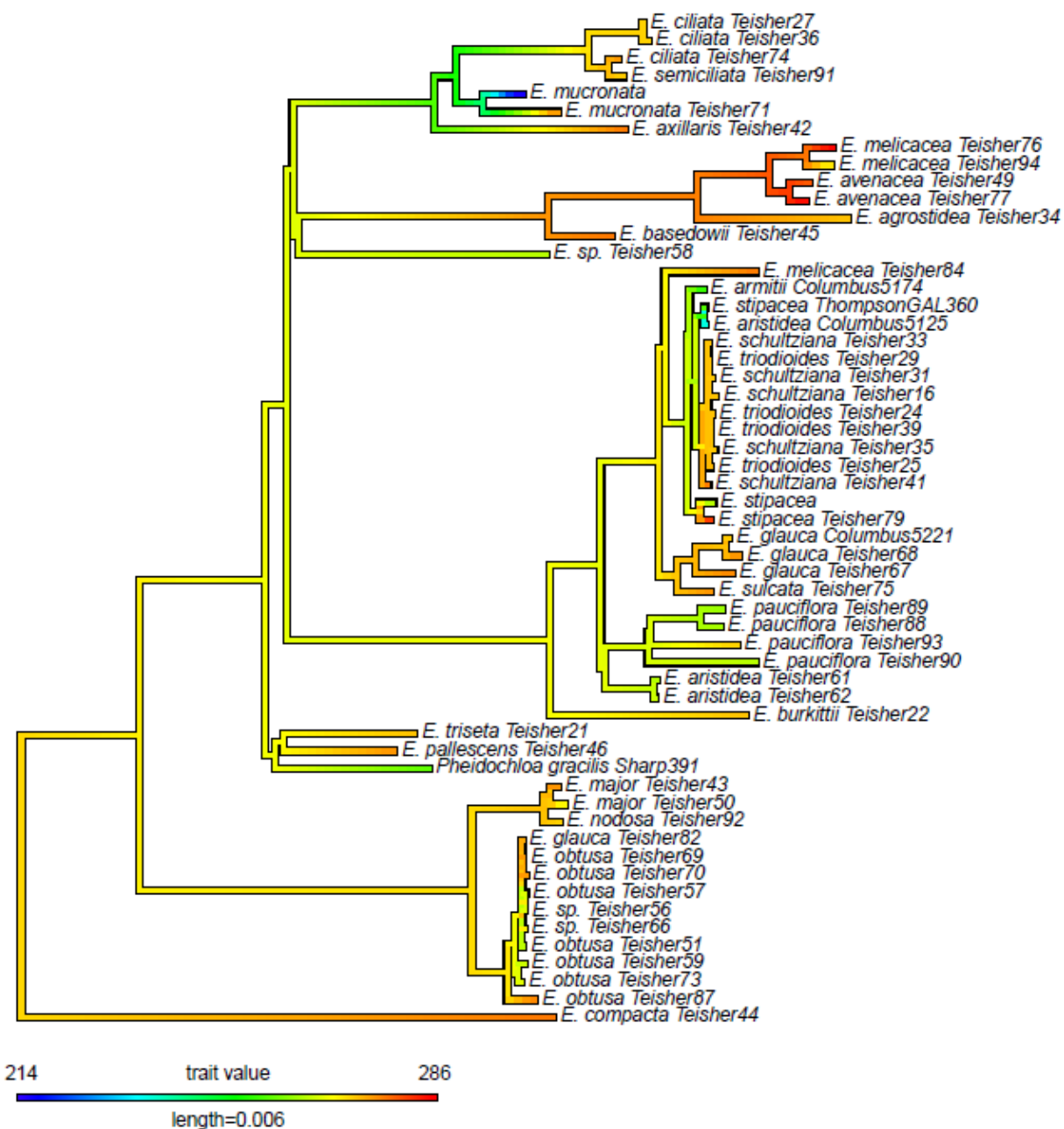


Figure 3.10. Mean annual temperature plotted across a reduced maximum-likelihood phylogeny of Eriachneae under a Brownian motion model of evolution using the function *contMap* in R package phytools.

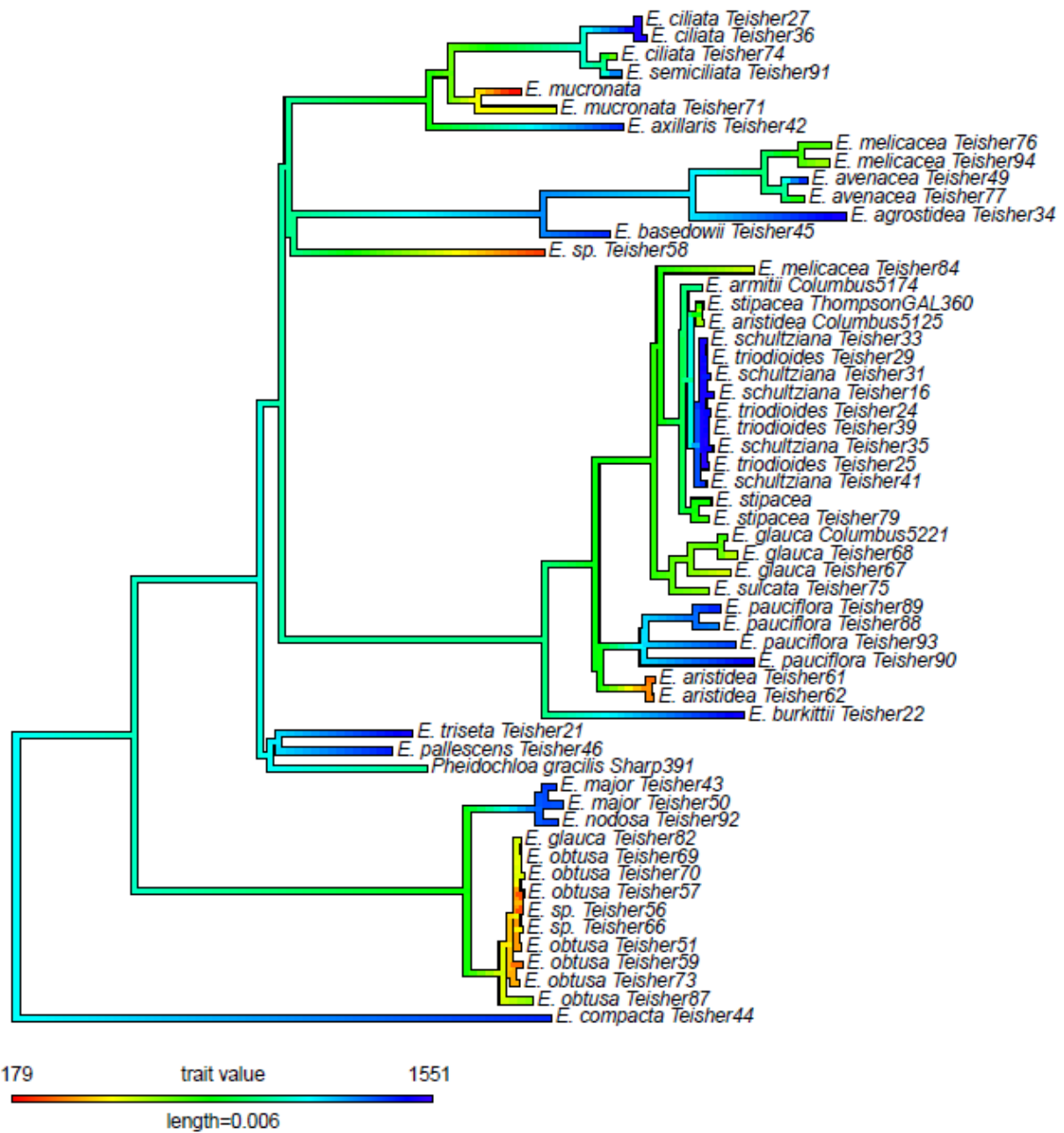


Figure 3.11. Mean annual precipitation plotted across a reduced maximum-likelihood phylogeny of Eriachneae under a Brownian motion model of evolution using the function *contMap* in R package phytools.

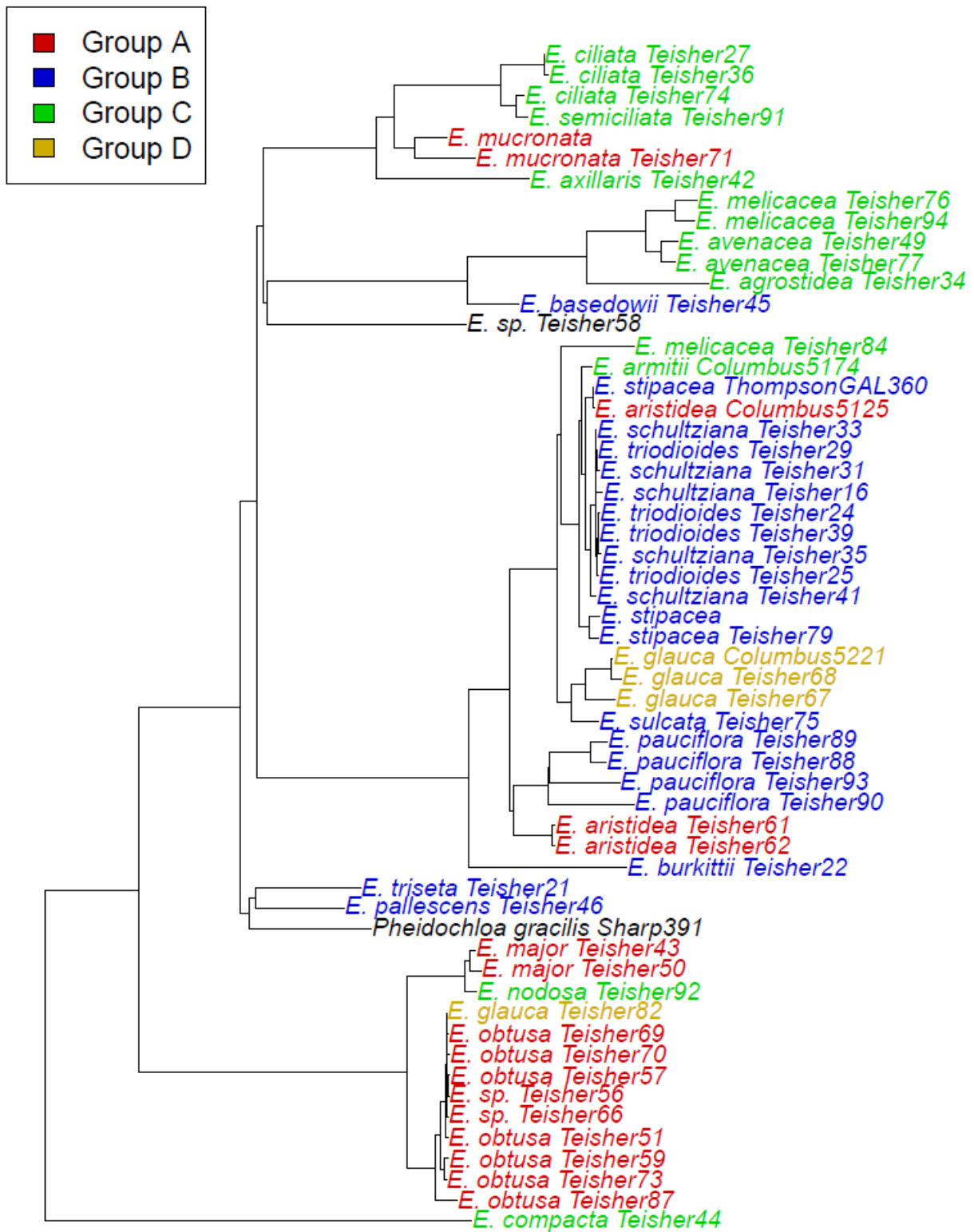


Figure 3.12. Maximum-likelihood topology of Eriachneae with taxon labels colored according to classification of Lazarides (1995).

Table 3.1. Whole-chloroplast genome samples used in the phylogenetic analysis with plastome assembly statistics for samples generated in this study.

Taxon	Voucher	Source	LSC Length	SSC Length	IR Length	Assembly Length	LSC Coverage	SSC Coverage	IR Coverage
<i>E. schultzi</i>	Teisher 16	This Study	79861	12537	21162	134722	6	6	14
<i>E. triset</i>	Teisher 21	This Study	79893	12521	21150	134714	59	59	119
<i>E. burkittii</i>	Teisher 22	This Study	79558	12576	21161	134456	12	14	24
<i>E. triodioides</i>	Teisher 24	This Study	79888	12593	21142	134765	10	10	19
<i>E. triodioides</i>	Teisher 25	This Study	79901	12566	21162	134791	18	18	39
<i>E. ciliata</i>	Teisher 27	This Study	79911	12603	21136	134786	20	17	47
<i>E. triodioides</i>	Teisher 29	This Study	79892	12567	21163	134785	22	17	59
<i>E. schultzi</i>	Teisher 31	This Study	79898	12567	21163	134791	17	15	42
<i>E. schultzi</i>	Teisher 33	This Study	79887	12593	21164	134808	23	23	42
<i>E. agrostidea</i>	Teisher 34	This Study	80030	12601	21144	134919	11	10	24
<i>E. schultzi</i>	Teisher 35	This Study	79894	12593	21164	134815	24	24	43
<i>E. ciliata</i>	Teisher 36	This Study	79903	12603	21138	134782	12	12	25
<i>E. triodioides</i>	Teisher 39	This Study	79872	12593	21163	134791	68	75	136
<i>E. schultzi</i>	Teisher 41	This Study	79876	12567	21167	134777	58	52	137
<i>E. axillaris</i>	Teisher 42	This Study	80147	12624	21151	135073	28	27	53
<i>E. major</i>	Teisher 43	This Study	79772	12678	21149	134748	23	24	45
<i>E. compacta</i>	Teisher 44	This Study	80139	12585	21079	134882	40	37	80
<i>E. basedowii</i>	Teisher 45	This Study	80018	12619	21148	134933	65	70	139
<i>E. pallescens</i>	Teisher 46	This Study	80032	12643	21137	134949	27	29	54
<i>E. avenacea</i>	Teisher 49	This Study	79926	12641	21145	134857	96	102	188
<i>E. major</i>	Teisher 50	This Study	79758	12678	21150	134736	17	12	44
<i>E. obtusa</i>	Teisher 51	This Study	79818	12594	21148	134708	59	58	113
<i>E. sp.</i>	Teisher 56	This Study	79732	12594	21150	134626	68	63	152
<i>E. obtusa</i>	Teisher 57	This Study	79736	12594	21147	134624	59	56	123
<i>E. sp.</i>	Teisher 58	This Study	79898	12634	21169	134870	142	137	329
<i>E. obtusa</i>	Teisher 59	This Study	79744	12595	21150	134639	34	28	78
<i>E. aristidea</i>	Teisher 61	This Study	79656	12573	21146	134521	27	24	63
<i>E. aristidea</i>	Teisher 62	This Study	79654	12600	21146	134546	108	106	232
<i>E. sp.</i>	Teisher 66	This Study	79728	12594	21150	134622	48	45	92
<i>E. glauca</i>	Teisher 67	This Study	79893	12622	21164	134843	59	55	136
<i>E. glauca</i>	Teisher 68	This Study	79879	12618	21164	134825	53	47	129
<i>E. obtusa</i>	Teisher 69	This Study	79735	12594	21150	134629	55	51	107
<i>E. obtusa</i>	Teisher 70	This Study	79734	12595	21150	134629	181	184	363

<i>E. mucronata</i>	Teisher 71	This Study	80148	12621	21156	135081	105	109	209
<i>E. obtusa</i>	Teisher 73	This Study	79743	12591	21153	134640	44	40	101
<i>E. ciliata</i>	Teisher 74	This Study	79884	12599	21127	134737	48	45	107
<i>E. sulcata</i>	Teisher 75	This Study	79797	12590	21147	134681	63	55	148
<i>E. melicacea</i>	Teisher 76	This Study	79964	12665	21141	134911	38	34	98
<i>E. avenacea</i>	Teisher 77	This Study	79863	12558	21153	134727	71	73	150
<i>E. stipacea</i>	Teisher 79	This Study	79847	12595	21164	134770	35	35	65
<i>E. glauca</i>	Teisher 82	This Study	79736	12594	21150	134630	73	68	179
<i>E. melicacea</i>	Teisher 84	This Study	79758	12643	21164	134729	283	270	690
<i>E. obtusa</i>	Teisher 87	This Study	79736	12592	21150	134628	50	49	93
<i>E. pauciflora</i>	Teisher 88	This Study	79639	12548	21166	134519	33	31	78
<i>E. pauciflora</i>	Teisher 89	This Study	79730	12547	21165	134607	26	25	61
<i>E. pauciflora</i>	Teisher 90	This Study	79582	12620	21133	134468	73	76	144
<i>E. ciliata</i>	Teisher 91	This Study	79813	12633	21138	134722	54	46	116
<i>E. nodosa</i>	Teisher 92	This Study	79734	12646	21148	134676	11	8	33
<i>E. pauciflora</i>	Teisher 93	This Study	79675	12664	21053	134445	17	16	41
<i>E. melicacea</i>	Teisher 94	This Study	79989	12585	21156	134886	75	76	165
<i>E. aristidea</i>	Columbus 5125	This Study	79882	12598	21164	134808	35	33	78
<i>E. armitii</i>	Columbus 5174	This Study	79847	12599	21163	134772	15	14	35
<i>E. glauca</i>	Columbus 5221	This Study	79885	12617	21164	134830	36	34	90
<i>E. stipacea</i>	ThompsonGAL360	This Study	79890	12594	21167	134818	64	64	114
<i>E. stipaceae</i>	-	Genbank	-	-	-	-	-	-	-
<i>E. mucronata</i>	-	Genbank	-	-	-	-	-	-	-
<i>Pheidochloa gracilis</i>	Sharp 391	Chapter 2	-	-	-	-	-	-	-
<i>Arundo donax</i>	-	Chapter 2	-	-	-	-	-	-	-
<i>Micraira sp.</i>	-	Chapter 2	-	-	-	-	-	-	-
<i>Coelachne africana</i>	-	Chapter 2	-	-	-	-	-	-	-
<i>Isachne albens</i>	-	Chapter 2	-	-	-	-	-	-	-
<i>Isachne distichophylla</i>	-	Chapter 2	-	-	-	-	-	-	-
<i>Limnopoa meeboldii</i>	-	Chapter 2	-	-	-	-	-	-	-

Table 3.2. Carbon isotope values for Isachneae, Eriachneae, and Micraireae based on leaf material.

Species	Sample	Tribe	d13C	st.dev.
Coelachne africana	Thomas 3935	Isachneae	-28.59	0.05
C. auquieri	Bouxin 774	Isachneae	-28.26	0.22
Eriachne agrostidea	Michel 2830	Eriachneae	-13.55	0.13
E. agrostidea	Teisher 34	Eriachneae	-13.51	0.04
E. aristidea	Henry 560	Eriachneae	-13.42	0.03
E. aristidea	Teisher 61	Eriachneae	-15.33	0.05
E. aristidea	Teisher 62	Eriachneae	-14.43	0.19
E. armitii	Teisher 15	Eriachneae	-13.16	0.08
E. avenacea	Teisher 49	Eriachneae	-13.32	0.16
E. axillaris	Teisher 42	Eriachneae	-13.65	0.02
E. basedowii	Perry 2646	Eriachneae	-12.03	0.01
E. basedowii	Teisher 45	Eriachneae	-13.42	0.05
E. benthamii	Purdie 1453	Eriachneae	-13.17	0.07
E. burkittii	Teisher 22	Eriachneae	-13.35	0.04
E. ciliata	Teisher 26	Eriachneae	-16.61	0.89
E. ciliata	Teisher 36	Eriachneae	-16.34	0.11
E. ciliata	Teisher 78	Eriachneae	-14.25	0
E. compacta	Risler 1582	Eriachneae	-12.8	0.18
E. compacta	Teisher 44	Eriachneae	-14.28	0.2
E. glauca	Simon 3688	Eriachneae	-11.81	NA
E. glauca	Teisher 67	Eriachneae	-14.15	0.21
E. glauca	Teisher 88	Eriachneae	-14.5	0.15
E. helmsii	Latz 13499	Eriachneae	-13.19	0.17
E. major	Teisher 43	Eriachneae	-13.46	0.12
E. melicacea	Teisher 84	Eriachneae	-14.06	0.07
E. melicacea	Teisher 94	Eriachneae	-14.19	0.15
E. mucronata	Latz 13498	Eriachneae	-13.4	0.02
E. mucronata	Teisher 71	Eriachneae	-14.77	0.06
E. nodosa	Teisher 92	Eriachneae	-13.3	0.13
E. obtusa	Lazarides 8525	Eriachneae	-11.94	0.1
E. obtusa	Teisher 50	Eriachneae	-14.08	0
E. obtusa	Teisher 64	Eriachneae	-13.54	0.07
E. obtusa	Teisher 73	Eriachneae	-13.73	0.43
E. obtusa	Teisher 86	Eriachneae	-13.57	0.38
E. ovata	Badman 2766	Eriachneae	-13.1	0.03
E. pallescens	Clarkson 4727	Eriachneae	-13.21	0.18
E. pallescens	Teisher 46	Eriachneae	-14.12	0.15
E. pauciflora	Teisher 93	Eriachneae	-13.95	0.34
E. pulchella	Teisher 69	Eriachneae	-13.59	0.05

<i>E. pulchella</i>	Teisher 70	Eriachneae	-13.37	0.11
<i>E. rara</i>	Hubbard 7679	Eriachneae	-12.95	0.37
<i>E. schultzi</i>	Teisher 16	Eriachneae	-13.25	0.08
<i>E. schultzi</i>	Teisher 33	Eriachneae	-12.75	0.09
<i>E. schultzi</i>	Teisher 41	Eriachneae	-13.62	0
<i>E. sulcata</i>	Aplin 1355	Eriachneae	-13.83	0.12
<i>E. sulcata</i>	Teisher 75	Eriachneae	-13.69	0.33
<i>E. sulcata</i>	Teisher 83	Eriachneae	-13.7	0.09
<i>E. triodioides</i>	Blake 8190	Eriachneae	-11.57	0.08
<i>E. triodioides</i>	Teisher 24	Eriachneae	-12.96	0.08
<i>E. triodioides</i>	Teisher 38	Eriachneae	-13.48	0.04
<i>E. trisetata</i>	Latz 3734	Eriachneae	-12.21	0.08
<i>E. trisetata</i>	Teisher 19	Eriachneae	-12.66	0.04
<i>E. trisetata</i>	Teisher 30	Eriachneae	-13.34	0.17
<i>Isachne albens</i>	KEKE 947	Isachneae	-28.49	0.09
<i>I. albens</i>	Li Heng 9721	Isachneae	-27.67	0.01
<i>I. arundinacea</i>	Galdames 1804	Isachneae	-30.74	0.05
<i>I. arundinacea</i>	Mora 1659	Isachneae	-30.06	0.17
<i>I. confusa</i>	Latz 10874	Isachneae	-30.69	0
<i>I. confusa</i>	Michell 4071	Isachneae	-31.78	0.21
<i>I. distichophylla</i>	Degener 33379	Isachneae	-26.74	0.03
<i>I. kiyalaensis</i>	Baldwin 10434	Isachneae	-26.68	0.02
<i>I. ligulata</i>	MacDougal 3633	Isachneae	-30.65	0.1
<i>I. ligulata</i>	Madsen 7134	Isachneae	-29.38	0.27
<i>I. mauritiana</i>	Croat 29246	Isachneae	-23.9	0
<i>I. mauritiana</i>	Gautier 3613	Isachneae	-31.15	0.13
<i>I. myosotis</i>	Gjellerup 33	Isachneae	-31.94	0
<i>I. polygonoides</i>	Ritter 2477	Isachneae	-29.5	0.01
<i>I. polygonoides</i>	Taylor 9784	Isachneae	-31.5	0.39
<i>I. pulchella</i>	Stone 5050	Isachneae	-27.8	0.06
<i>I. rigens</i>	Grignon 84239	Isachneae	-31.56	0.26
<i>I. rigens</i>	Laegaard 70579	Isachneae	-27.2	0.01
<i>I. rigidifolia</i>	Zanoni 20044	Isachneae	-28.72	0.28
<i>I. rigidifolia</i>	Zanoni 22285	Isachneae	-29.32	0.04
<i>I. walkeri</i>	Lazarides 7205	Isachneae	-28.6	0.02
<i>Limnopoia meeboldii</i>	Cook 282	Isachneae	-27.47	0.21
<i>Micraira adamsii</i>	Harwood 978	Micraireae	-31.38	0.04
<i>M. subulifolia</i>	Blake 18722	Micraireae	-25.95	0.12
<i>Pheidochloa gracilis</i>	Dunlop 5553	Eriachneae	-11.78	0.07
<i>P. gracilis</i>	Sharp 391	Eriachneae	-12.45	0.08

Table 3.3. Variable loadings for first two principal components from analysis of nineteen climate variables relating to temperature and precipitation.

BioClim Code	Variable	PC1	PC2
BIO1	Annual Mean Temperature	0.13	-0.34
BIO2	Mean Diurnal Range	-0.27	-0.11
BIO3	Isothermality (BIO2/BIO7)*100	0.26	0.00
BIO4	Temperature Seasonality	-0.32	0.05
BIO5	Max Temperature of Warmest Month	-0.15	-0.31
BIO6	Min Temperature of Coldest Month	0.29	-0.19
BIO7	Temperature Annual Range	-0.33	-0.03
	Mean Temperature of Wettest		
BIO8	Quarter	-0.04	-0.30
BIO9	Mean Temperature of Driest Quarter	0.20	-0.26
	Mean Temperature of Warmest		
BIO10	Quarter	-0.07	-0.34
	Mean Temperature of Coldest		
BIO11	Quarter	0.24	-0.26
BIO12	Annual Precipitation	0.30	0.14
BIO13	Precipitation of Wettest Month	0.32	-0.01
BIO14	Precipitation of Driest Month	0.10	0.31
BIO15	Precipitation Seasonality	0.15	-0.30
BIO16	Precipitation of Wettest Quarter	0.32	-0.01
BIO17	Precipitation of Driest Quarter	0.10	0.31
BIO18	Precipitation of Warmest Quarter	0.27	0.12
BIO19	Precipitation of Coldest Quarter	0.11	0.27

Table 3.4. Parameters from fitting alternative models of evolution to BioClim variables using the Maximum Likelihood tree of Eriachneae.

Brownian Motion																			
	bio_1	bio_2	bio_3	bio_4	bio_5	bio_6	bio_7	bio_8	bio_9	bio_10	bio_11	bio_12	bio_13	bio_14	bio_15	bio_16	bio_17	bio_18	bio_19
sigsq	1200655	1674206	22253.29	3094435208	1554203	4125557	7327505	355353.3	3855234	502532.6	2741854	293563870	33248658	75092.48	3722731	202083083	1308691	43191605	3827176
z0	271.489	120.3755	58.20379	2514.5802	365.4851	158.3676	207.1175	288.6832	242.0682	299.6375	234.3495	1152.5747	274.7948	1.091957	106.9968	766.65574	6.407109	352.3837	12.56273
lnL	-258.9137	-268.389	-145.252	-482.76675	-266.269	-294.0922	-310.464	-224.215	-292.1608	-234.091	-282.448	-415.6417	-353.566	-179.914	-291.164	-404.9993	-261.369	-361.023	-291.953
k	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
aic	521.8274	540.7783	294.5038	969.533501	536.5388	592.1844	624.9271	452.4293	588.3216	472.1825	568.8961	835.28337	711.1328	363.829	586.328	813.99864	526.7385	726.0458	587.9052
Delta																			
	bio_1	bio_2	bio_3	bio_4	bio_5	bio_6	bio_7	bio_8	bio_9	bio_10	bio_11	bio_12	bio_13	bio_14	bio_15	bio_16	bio_17	bio_18	bio_19
delta	4.931861	3.603649	3.12799	3.31737367	0.925162	5.481756	3.03851	3.630453	1.59225	3.304288	4.6123	2.3787455	2.580652	10.77313	3.110773	2.5921929	10.71016	1.888941	2.276623
sigsq	810391.6	1123742	14724.76	2048969348	1627479	2831830	4930442	234434.9	3056406	333415.4	1827214	203565518	22796391	75032.41	2503992	138125622	1339171	32079562	2701061
z0	268.8922	118.6448	58.26299	2512.00993	365.643	159.6664	204.7923	286.8611	241.415	297.3443	232.939	1130.6996	271.4552	1.223283	105.7472	750.72864	6.172327	351.9621	13.44399
lnL	-250.539	-264.718	-142.193	-479.38277	-266.264	-283.6027	-308.014	-219.979	-291.9043	-230.791	-275.0859	-414.2275	-351.945	-147.446	-288.5992	-403.2908	-230.016	-360.425	-290.912
k	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
aic	507.0781	535.4356	290.3862	964.76554	538.5275	573.2054	622.0272	445.9589	589.8086	467.5814	556.1718	834.45493	709.8897	300.8921	583.1985	812.58151	466.0322	726.8501	587.8249
Early Burst																			
	bio_1	bio_2	bio_3	bio_4	bio_5	bio_6	bio_7	bio_8	bio_9	bio_10	bio_11	bio_12	bio_13	bio_14	bio_15	bio_16	bio_17	bio_18	bio_19
a	-1.53E-06	-2.95E-06	-2.99E-06	-1.00E-06	-217.572	-4.14E-06	-2.87E-06	-6.34E-06	-57.52297	-1.10E-06	-2.03E-06	-5.77E-06	-2.47E-06	-1.25E-06	-2.92E-06	-1.79E-06	-1.37E-06	-3.55E-06	-3.00E-06
sigsq	1200581	1674469	22253.22	3094433624	7830750	4125253	7326519	355369.3	5972037	502350.6	2742230	293571969	33248699	75102.58	3723390	202082956	1308421	43189908	3826880
z0	271.489	120.3755	58.20379	2514.5802	366.4998	158.3676	207.1175	288.6832	242.2643	299.6375	234.3495	1152.5747	274.7948	1.091957	106.9968	766.65574	6.407109	352.3837	12.56273
lnL	-258.9137	-268.389	-145.252	-482.76674	-265.022	-294.0922	-310.464	-224.215	-292.0802	-234.091	-282.448	-415.6417	-353.566	-179.914	-291.164	-404.9993	-261.369	-361.023	-291.953
k	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
aic	523.8273	542.7782	296.5038	971.533485	536.0435	594.1844	626.927	454.4293	590.1604	474.1825	570.896	837.28336	713.1327	365.8289	588.328	815.99861	528.7383	728.0457	589.9051

Table 3.4 continued.

White Noise																			
	bio_1	bio_2	bio_3	bio_4	bio_5	bio_6	bio_7	bio_8	bio_9	bio_10	bio_11	bio_12	bio_13	bio_14	bio_15	bio_16	bio_17	bio_18	bio_19
sigsq	144.9978	216.5916	9.749462	1093194.11	224.3773	694.8945	1251.878	147.334	470.9923	110.7375	517.1283	233234.65	12009.18	3.208372	267.0348	100746.43	63.25823	10568.27	259.0212
z0	269.8596	120.0702	56.59649	2846.22807	361.7193	149.9825	211.7368	291.4386	240.9123	300.2281	227.6842	986.01754	248.3684	1.192982	108.0175	660.54386	8.070175	317.8947	16.52632
lnL	-222.716	-234.153	-145.78	-477.16101	-235.159	-267.3767	-284.153	-223.171	-256.2925	-215.034	-258.9558	-433.1338	-348.592	-114.104	-240.1198	-409.2098	-199.075	-344.949	-239.251
k	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
aic	449.432	472.3058	295.5601	958.32201	474.3188	538.7533	572.3058	450.343	516.585	434.0672	521.9116	870.26761	701.1843	232.2075	484.2396	822.41963	402.1508	693.8988	482.5029
Lambda																			
	bio_1	bio_2	bio_3	bio_4	bio_5	bio_6	bio_7	bio_8	bio_9	bio_10	bio_11	bio_12	bio_13	bio_14	bio_15	bio_16	bio_17	bio_18	bio_19
lambda	2.36E-211	0.070421	0.99857	0.97635921	0.132098	2.68E-120	0.081681	0.955439	4.88E-155	0.82594	4.59E-217	0.9966757	0.963289	4.98E-38	1.87E-208	0.9815752	2.54E-15	0.856821	0.118075
sigsq	18578.04	27168.22	18514.47	1010106636	27817.1	90379.5	164047.2	78932.6	61157.6	36378.85	67287.25	208723744	7678855	369.635	35554.74	75918347	7281.038	4128711	35441.8
z0	269.5193	121.2849	58.21248	2513.14173	363.6848	149.0709	213.8486	288.8929	241.1275	299.7399	226.8686	1149.978	272.2664	1.171266	107.58	762.11703	7.874071	345.6248	17.4783
lnL	-224.0993	-234.325	-144.201	-474.3889	-234.011	-269.1874	-285.412	-212.674	-258.0563	-211.773	-260.7786	-413.5941	-340.608	-112.459	-242.5985	-397.9258	-197.403	-343.251	-241.162
k	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
aic	454.1987	474.6496	294.4012	954.777806	474.0225	544.3748	576.8237	431.3481	522.1127	429.5461	527.5571	833.18811	687.2165	230.9172	491.197	801.85165	400.8064	692.5029	488.3245

CONCLUSION OF THE DISSERTATION

As a result of their economic value, grasses have become a model system for a variety of disciplines, including physiology, ecology, and genetics. The grass family, Poaceae, is also an ideal system in which evolutionary hypotheses can be tested due to the extensive genetic resources available from crop relatives and the broad diversity of its species. However, the large size of the family, with over 11,000 species, has ensured that systematic problems persist despite centuries of study. In this dissertation, I resolved relationships within two of the least well-studied subfamilies, the Arundinoideae and Micrairoideae, and used this newly-established phylogenetic framework to explore evolutionary issues within these subfamilies as well as across the PACMAD clade of grasses. The results of these efforts are described below along with their implications for grass evolution and suggestions for future research.

In Chapter 1, I conducted a phylogenetic analysis of chloroplast whole-genome sequences from 88 samples representing all twelve grass subfamilies. Sampling for this phylogeny focused on the Arundinoideae and included six genera from this subfamily that have never been included in a molecular phylogenetic analysis. I also scored three lemma characters relating to the grass "diaspore burial syndrome" (Humphreys *et al.*, 2010) for each of the samples in the phylogeny using my own observations on species in Arundinoideae and data from the literature (Watson & Dallwitz, 1992 forward) for other PACMAD taxa. I found that in the modern circumscription of Arundinoideae the subfamily is still polyphyletic, with *Nematopoa* belonging in the Chloridoideae, *Phaenanthoecium* grouping with the Danthonioideae, and *Alloeochoete* and *Dichaetaria* forming a clade that is sister to the rest of the Panicoideae. This small clade affects ancestral state estimation of the presence and type of lemma awn across the PACMAD grasses, which identifies presence of a straight awn as the most likely condition in the ancestor of this clade. Contrary to the findings of Humphreys *et al.* (2010) in the

Danthonioideae, a geniculate awn is not associated with higher diversification rates at the broader scale examined in my study. An analysis of diversification rate shifts using the program BAMM (Rabosky, 2014) on a dated, ultrametric tree created in BEAST (Drummond *et al.*, 2012) found support for two increases in diversification rate that most likely occurred near the base of core Panicoideae and crown Chloridoideae. These clades contain two of the many independent losses of lemma awns in the PACMAD clade, suggesting that passive burial is at the very least not an obstacle to lineage accumulation in these grasses. Two clades stand out as being potentially fruitful for more detailed analyses of burial traits: tribe Andropogoneae in subfamily Panicoideae and the modified tribe Arundineae, consisting of the genera *Arundo*, *Amphipogon*, *Monachather*, and *Dregeochloa*.

The reduced Arundinoideae *sensu stricto* identified in this study contains a slightly more manageable but still morphologically and ecologically heterogeneous collection of taxa. The subfamily can be divided into two tribes on the basis of relative glume length. Members of tribe Arundineae, expanded here from Soreng *et al.* (2015) to include the South African genus *Dregeochloa*, possess glumes that are typically as long as or longer than their spikelets. Tribe Molinieae, which is here the same as in Soreng *et al.* but without *Dichaetaria*, *Dregeochloa* or *Nematopoa*, has glumes that are typically shorter than the spikelets. Three putative members of this tribe still need to be sampled for molecular sequence data: the Angolan monotypic genus *Piptophyllum*, the Indian genus *Zenkeria*, and the Ethiopian monotypic genus *Leptagrostis*, which was collected only once in 1854 (Hubbard, 1939). Excluding any surprise transfers as in the case of *Eragrostis walteri* (Ingram *et al.*, 2011), the systematic relationships in subfamily Arundinoideae at long last appear to be under control.

In Chapter 2, I performed a phylogenetic analysis on newly-generated transcriptome sequence data from four species in Arundinoideae combined with existing coding sequence data of nine species from the subfamilies Panicoideae, Aristidoideae, Oryzoideae, Bambusoideae, and Anomochlooideae. The goal of this study was to identify the source(s) of the polyploid genomes found in the two clades of temperate Arundinoideae: *Phragmites*+*Molinia*+*Hakonechloa* and *Arundo*. The program PUG (McKain *et al.*, 2016) was used to identify the coalescence points of gene pairs on gene trees that correspond to nodes on a given species tree. The tree topology from chloroplast data in Chapter 1 was used as a guideline for this analysis, but alternative topologies for the members of Arundinoideae were also explored due to the possibility of hybridization between these taxa. I found strong support for a whole genome duplication in the ancestor of *Phragmites*, *Molinia*, and *Hakonechloa* as well as some evidence for such an event in the ancestor of the Arundinoideae. However, this latter event could also be the result of hybridization between one or more of the species of this subfamily in my study with a species that has not been sampled. Similarly, other nodes in the species tree associated with large numbers of coalescing gene pairs are likely the result of the way that the tree is rooted or short branches in the gene trees, which can both cause coalescence to be pushed down the tree. The event shared by *Phragmites*, *Molinia*, and *Hakonechloa* is likely to be real because sampling from that clade is relatively complete, but additional species from across the PACMAD clade are needed to identify the sources of the subgenomes in *Arundo*. However, *Arundo* and *Phragmites* do not share any gene pairs that are more closely related to one another than they are to *Molinia* or *Hakonechloa*, suggesting that a unique shared parental genome cannot be the reason for the convergence in morphology and ecology between these two large invasive reeds. Still, the inference of a possible whole genome duplication in the ancestor of a clade that transitioned

from a tropical ancestral habitat to a global distribution is of great interest for the role that polyploidy may play in dispersal (i.e. Linder & Barker, 2014) and invasiveness (Pandit *et al.*, 2011).

In Chapter 3, I explored the evolution of C₄ photosynthesis in the Micrairoideae using molecular phylogenetics, carbon isotopes, and climate variables relating to precipitation and temperature. Tribe Eriachneae represents one of the at least 22 independent origins of the C₄ pathway in the grass family and constitutes a unique combination of anatomical and biochemical subtypes (Sinha & Kellogg, 1996). Subfamily Micrairoideae is especially interesting for investigating the role of C₄ in shaping evolutionary history because the C₃ tribe Isachneae, which is sister to the Eriachneae, occupies a broader geographic range and has more species than Eriachneae. This is in opposition to the general trend for C₄ clades across the grasses, which are typically associated with a broader range of habitats (Edwards & Still, 2008; Christin & Osborne, 2014) and increased diversification rates (Spriggs *et al.*, 2014) as compared with the closest C₃ relatives.

To understand this atypical pattern in Micrairoideae, I measured carbon isotope ratios for sixteen species of Isachneae, 27 species of Eriachneae, and two species of Micraireae. This survey confirmed that members of the Eriachneae are C₄ and those of Isachneae are entirely C₃. Values for Micraireae were the same as for Isachneae, despite the fact that a principal components analysis of climate variables showed that Micraireae occupies more similar climatic conditions to Eriachneae. Thus, the C₄ pathway in Eriachneae and the moss-like growth habit in Micraireae appear to represent alternative adaptations to similar climatic conditions. These conditions are estimated to be ancestral for the subfamily, suggesting that rather than C₄

photosynthesis allowing the Eriachneae to move into drier and more open habitats, the Isachneae escaped these habitats into wetter and shadier ones, diversifying as a result.

The fact that Eriachneae does not possess as many species as Isachneae does not mean that C₄ has not been a driver or enabler of diversification, however. It is possible that this pathway did indeed allow the occupation of new habitats, but that the number of such habitats was limited. In that case, evolution of habitat preferences would be expected to occur quickly and early in the history of the C₄ clade, slowing down over time as newly available habitats are filled. I tested this hypothesis by constructing a whole-plastome phylogeny of 25 species of Eriachneae along with outgroups from the Isachneae, Micraireae, and Arundinoideae. I extracted nineteen climate variables relating to temperature and precipitation for each of these samples and tested the abilities of various evolutionary models to explain species climate preferences across the phylogeny. Species with multiple samples in the tree were associated with large ranges in climate variables, so that a large amount of variation in climate preference was found within clades. Unsurprisingly, the models that best fit this pattern were those in which climate preferences evolve according to a white noise model without any phylogenetic signal or in which the evolutionary rate of these preferences has increased through time. In either case, the hypothesis of an "early burst" model of evolution (Harmon *et al.*, 2010) is rejected for climate preferences in C₄ Eriachneae.

There are many caveats to using georeferenced herbarium specimens combined with climate variables extracted from a raster to model species preferences (Hijmans *et al.*, 2005; Newbold, 2010), and my phylogeny contains only half of the species in the Eriachneae, but this study confirms that the pattern of C₄ evolution in the Micrairoideae departs significantly from the more general pattern seen in the PACMAD grasses as a whole. Evolutionary biology is often

characterized more by exceptions than rules, and studying these exceptions can provide us with a more nuanced and complete picture of evolutionary phenomena than we could achieve by focusing only on models that conform to our expectations. In this regard, the Micrairoideae represent a unique and interesting case of C₄ evolution worthy of further study.

This dissertation has both made use of and contributed to the large body of genetic and systematic resources that make the grass family one of the best systems in which to address fundamental evolutionary questions. By resolving long-standing systematic issues in the Arundinoideae and Micrairoideae, I was able to provide insight into trait evolution across the important PACMAD clade, identify a likely whole-genome duplication corresponding to a shift to colder habitats, and to clarify a pattern of C₄ evolution that is exceptional among grasses. High-throughput sequencing of chloroplast genomes and transcriptomes was crucial in this process, as was access to the invaluable herbarium collections at the Missouri Botanical Garden and the Royal Botanic Gardens, Kew. This study also highlights the role that modern systematics has to play even in well-studied groups like the Poaceae and outlines particularly promising avenues for future study within the Arundinoideae and Micrairoideae.

LITERATURE CITED

- Christin, P.-A., and C. P. Osborne. 2014. The evolutionary ecology of C₄ plants. *New Phytologist* 204(4): 765–81.
- Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29: 1969–1973.
- Edwards, E. J., and C. J. Still. 2008. Climate, phylogeny and the ecological distribution of C₄ grasses. *Ecology Letters* 11(3): 266–76.
- Harmon, L. J., J. B. Losos, T. J. Davies, R. G. Gillespie, J. L. Gittleman, W. B. Jennings, K. H. Kozak, *et al.* 2010. Early bursts of body size and shape evolution are rare in comparative data. *Evolution* 64: 2385–2396.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.
- Hubbard, C. E. 1939. Notes on African Grasses: XXII. Tropical African Grasses. *Bulletin of Miscellaneous Information (Royal Botanic Gardens, Kew)* 1939(10): 643–655.
- Humphreys, A. M., A. Antonelli, M. D. Pirie, and H. P. Linder. 2010. Ecology and evolution of the diaspore burial syndrome. *Evolution* 65(4): 1163–1180.
- Ingram, A. L., P. Christin, and C. P. Osborne. 2011. Molecular phylogenies disprove a hypothesized C₄ reversion in *Eragrostis walteri* (Poaceae). *Annals of Botany* 321–325.
- Linder, H. P., and N. P. Barker. 2014. Does polyploidy facilitate long-distance dispersal? *Annals of Botany* 113: 1175–1183.
- McKain, M. R., H. Tang, J. R. McNeal, S. Ayyampalayam, J. I. Davis, C. W. dePamphilis, T. J. Givnish, *et al.* 2016. A phylogenomic assessment of ancient polyploidy and genome evolution across the Poales. *Genome Biology and Evolution* 8(4): 1150–1164.
- Newbold, T. 2010. Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. *Progress in Physical Geography* 34: 3–22.
- Pandit, M. K., M. J. O. Pocock, and W. E. Kunin. 2011. Ploidy influences rarity and invasiveness in plants. *Journal of Ecology* 99: 1108–1115.
- Rabosky, D. L. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* 9(2): e89543.
- Sinha, N. R., and E. A. Kellogg. 1996. Parallelism and diversity in multiple origins of C₄

- photosynthesis in the grass family. *American Journal of Botany* 83(11): 1458–1470.
- Soreng, R. J., P. M. Peterson, K. Romaschenko, G. Davidse, F. O. Zuloaga, E. J. Judziewicz, T. S. Filgueiras, *et al.* 2015. A worldwide phylogenetic classification of the Poaceae (Gramineae): phylogenetic classification of the grasses. *Journal of Systematics and Evolution* 53(2): 117–37.
- Spriggs, E. L., P.-A. Christin, and E. J. Edwards. 2014. C₄ photosynthesis promoted species diversification during the Miocene grassland expansion. *PLoS ONE* 9(5): 1–10.
- Watson, L., and Dallwitz, M.J. 1992 onwards. The grass genera of the world: descriptions, illustrations, identification, and information retrieval; including synonyms, morphology, anatomy, physiology, phytochemistry, cytology, classification, pathogens, world and local distribution, and references. Version: 7th December 2015. <http://delta-intkey.com>.