

**Untangling the evolution of CAM photosynthesis within Australian
Calandrinia (Montiaceae)**

By Lillian Pine Hancock

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of
Philosophy in the Department of Ecology and Evolutionary Biology at Brown University

Providence, Rhode Island
May 2018

© Copyright 2018 by Lillian Pine Hancock

This dissertation by Lillian Pine Hancock is accepted in its present form by the
Department of Ecology and Evolutionary Biology as satisfying the dissertation
requirement for the degree of Doctor of Philosophy.

Date_____

Erika J. Edwards, Ph.D., Advisor

Recommended to the Graduate Council

Date_____

James R. Kellner, Ph.D., Reader

Date_____

Daniel M. Weinreich, Ph.D., Reader

Date_____

Joseph A.M. Holtum, Ph.D., Reader

Approved by the Graduate Council

Date_____

Andrew G. Campbell, Dean of the Graduate School

Lillian Pine Hancock

Curriculum Vitae

Education

Ph.D. Ecology and Evolutionary Biology	2012- 2017
---	------------

Brown University (Providence, RI) Advisor: Erika Edwards

Thesis: Untangling the evolution of CAM photosynthesis within Australian *Calandrinia* (Montiaceae)

M.S. Biological Sciences	2008-2010
---------------------------------	-----------

University of Rhode Island (Kingstown, RI) Advisor: Christopher Lane

Thesis: Unraveling the role of mitochondria in red algal parasite evolution

B.S. Biology	2004- 2008
---------------------	------------

Trinity College (Hartford, CT) Faculty Honors

Thesis: Shell investigation, foraging behavior & sexual receptiveness; priorities in *P. longicarpus*

Appointments

Postdoctoral Researcher – James Cook University	Jan 18 – present
--	------------------

Teaching assistant, Brown University

- | | |
|-----------------------|-------------|
| • Desert Ecology | spring 2016 |
| • Evolution | fall 2015 |
| • Plan Eco-Physiology | spring 2014 |

Research technician, New York Botanical Garden, Bronx, NY	2010-2012
--	-----------

Teaching assistant, University of Rhode Island

- | | |
|------------------------|-----------|
| • Introductory Biology | 2009-2010 |
|------------------------|-----------|

Sustainable Farming Internship – Spannocchia Foundation, Italy	summer 2008
---	-------------

Supplementary Instructor, Trinity College

- | | |
|------------------------|-----------|
| • Introductory Biology | 2007-2008 |
|------------------------|-----------|

Publications

Hancock LP, Holtum JAM, and EJ Edwards. The evolution of a spectrum of CAM phenotypes in Australian *Calandrinia* (Montiaceae). *New Phytologist*, In Review.

Hancock LP, Obbens F, Moore AJ, Thiele K, de Vos JM, West J, Holtum JAM, and EJ Edwards. 2018. Phylogeny, evolution, and biogeographic history of *Calandrinia* (Montiaceae). *American Journal of Botany*. Accepted.

Holtum JAM, **LP Hancock**, EJ Edwards, K Winter. 2018. Crassulacean acid metabolism (CAM) in the Basellaceae (Caryophyllales). *Plant Biology*: DOI: 10.1111/plb.12698.

Goolsby, E, AJ Moore, **LP Hancock**, J De Vos, EJ Edwards. 2018. Molecular evolution of photosynthesis genes during transitions to C₄ and CAM metabolism. *American Journal of Botany*. DOI: 10.1002/ajb2.1051.

Thiele K, F Obbens, **LP Hancock**, EJ Edwards, JG West. 2018. Proposal to conserve the name *Parakeelya* against *Rumicastrum* (Montiaceae). *Taxon* 67: 214-215.

Moore AJ, J De Vos, **LP Hancock**, E Goolsby, EJ Edwards. 2018. Targeted enrichment of large gene families for phylogenetic inference: phylogeny and molecular evolution of photosynthesis genes in the portullugo (Caryophyllales). *Systematic Biology* 67: 367-383.

Holtum JAM, **LP Hancock**, EJ Edwards, K Winter. 2017. Facultative CAM photosynthesis (crassulacean acid metabolism) in four species of *Calandrinia*, ephemeral succulents of arid Australia. *Photosynthesis Research* 134: 1-21.

Obbens F, **LP Hancock**, KC Thiele, EJ Edwards. 2017. *Calandrinia holtumii* (Portulacaceae), a new and unusual species from arid Western Australia. *Nutsya* 28: 217-223.

Holtum JAM, **LP Hancock**, EJ Edwards, K Winter. 2017. Optional use of CAM photosynthesis in two C₄ species, *Portulaca cyclophylla* and *Portulaca digyna*. *Journal of Plant Physiology* 212: 91-96.

Holtum JAM, **LP Hancock**, EJ Edwards, MD Crisp, DM Crayn, R Sage, and K Winter. 2016. Australia lacks stem succulents but is it depauperate in plants with crassulacean acid metabolism (CAM)? *Current Opinions in Plant Biology* 31: 109-117.

Hancock LP and EJ Edwards. 2014. Phylogeny and the inference of evolutionary trajectories. *Journal of Experimental Botany* 65 (13): 3491-3498.

Christin PA, M Arakaki, CP Osborne, A Brautigam, RF Sage, JM Hibberd, S Kelley, S Covshoff, GK Wong, **LP Hancock**, and EJ Edwards. 2014. Shared origins of a key enzyme during the evolution of C₄ and CAM metabolism. *Journal of Experimental Botany* 65 (13): 3609-3621.

Hancock LP, L Goff, and CE Lane. 2010. Red algae lose key mitochondrial genes in response to becoming parasitic. *Genome Biology and Evolution* 2: 897-910.

Publications in preparation

Hancock LP, Ogburn M, Martin E, Holtum JAM, and EJ Edwards. CAM biochemistry precedes the evolution of full CAM morphological and anatomical traits in the Portulacineae. In preparation.

Fellowships and grants

IBES Research Training & Travel award (\$3,600)	2017
Doctoral Dissertation Improvement Grant (\$20,700)	2016
IBES Research Training & Travel award (\$3,500)	2015
American Society of Plant Taxonomists, William R. Anderson Award (\$1000)	2015
Brown University Graduate Travel Award (\$650)	2015
EAPSI NSF Fellowship (Australia) (\$7000)	2015
Society of Systematic Biology satellite conference travel award (\$500)	2015
Brown University Graduate Travel Award (\$650)	2014
New Phytologist, Systems biology and ecology of CAM plants (\$800)	2014
Oliver Cromwell Gorton Arnold Biological Pre-doctoral Fellow	2013- '14
IGERT Fellow, Reverse Ecology (2 years graduate funding)	2012- '14
EPSCoR travel award (\$1000)	2010
REU (Research Experience for Undergraduates) Anacortes, WA	2007

Selected presentations

Lillian Hancock, Abigail Moore, Jurriaan de Vos, Erika Edwards, 2017. Phylogeny, evolution, and biogeographic history of *Calandrinia* (Montiaceae). Ernst Mayr Symposium, Evolution, Portland, OR

Lillian Hancock and EJ Edwards. 2017. Phylogeny, evolution, biogeographic history of *Calandrinia* (Montiaceae). Plants in New England. Massachusetts Institute of Technology, MA.

Lillian Hancock, Abigail Moore, Joseph Holtum, and Erika Edwards. 2016. Phylogeny and evolution of Australian *Calandrinia* (Montiaceae): a CAM-evolving, succulent plant lineage from Australia, and the coolest plant radiation you've never heard of. Botany. Savannah, GA.

Lillian Hancock, Abigail Moore, Joseph Holtum, and Erika Edwards. 2016. A phylogenetic perspective reveals evolutionary lability in CAM evolution in *Calandrinia* (Montiaceae). Evolution. Austin, TX.

Lillian Hancock, Emily Hollenbeck, and August Gaung. 2015. Untangling the morphotype and latitudinal variation in *Spartina alterniflora*. Brown University, RI.

Lillian Hancock and Erika Edwards. 2015. Parakeelya, the amazing Australian radiation you have never heard about. Society for Systematic Biology. Ann Arbor, MI

Lillian Hancock, Sam Church, and Erika Edwards. 2014. A new look at an old method for characterizing CAM photosynthesis. New Phytologist, Systems biology and ecology of CAM plants. Lake Tahoe, CA.

Lillian Hancock and Erika Edwards. 2013. Exploring the phylogenetic approach to detect the CAM evolutionary trajectory. C4/ CAM symposium. Champaign, IL.

Lillian Hancock. 2012. Transitioning to Tracheophytes. Brown University, RI.

Lillian Hancock & Chris Lane. 2010. Comparative mitochondrial genomics in red algal parasite evolution. Evolution, Portland, OR.

Lillian Hancock & Chris Lane. 2008 Unraveling the role of mitochondria in red algal parasite evolution. Society for Molecular Biology and Evolution, Iowa City, IA.

Mentoring and community outreach

Plant to farm to table program (2017)

- Started a collaborative program with a local RI middle school science teacher, farm, and restaurant to teach kids about how plants grow and where food comes from.
- Installed a small pseudo-greenhouse at the school so that students can grow vegetables year round

Undergraduate students advised

Church, Sam	Brown University 2014
Lichtenburg, Regan	Brown University 2014
O'Brian, Megan	URI 2010

ACKNOWLEDGEMENTS

For this opportunity, I am deeply grateful to my advisor Erika Edwards who gave me both the freedom and the support to develop a dissertation project that inspired and challenged me as scientist and as a person. Her thoughtfulness, scientific integrity, and overall curiosity was infectious and it was because of her guidance and friendship that I was able to pursue and complete such an ambitious thesis project. Such mentorship as found in Erika is rare and I feel humbled to have spent five years working alongside her.

I am also grateful to my committee member, advisor, and friend Joe Holtum. Joe welcomed me to Australia with such generosity and spirit, and essentially provided me with the connections and tools necessary to run a successful field campaign in a foreign country. His unconditional confidence and support over the past five years has been a true gift and this dissertation would not be what it is without him. In addition, I want to thank his wife, Rosemary Dunn, who opened her house to me many times and who's love and warmth made Australia feel like home.

As with any successful PhD project, there are many others who provide invaluable support and assistance. I want to express thanks to my committee members Jim Kellner and Dan Weinreich for their guidance and confidence over the past five years. Thanks to Frank Obbens for teaching me to both *see* and identify Australian *Calandrinia*, and for spending many days driving across Australia with me. Thank you to Kevin Thiele, Leesa Baker, Wendy and Steve Powles, Hazel Jensen, and Alex Cheeseman for their hospitality, support, feedback, and companionship while traveling in Australia. Thank you to Fred Jackson for his greenhouse support, the EEB staff for their assistance and patience, and the EEB department for their teachings, support, and

friendship. I want to say a special thank you to the Edwards Lab (Felipe, Radika, Juri, Abby, Monica, David, Elissa, Morgan, Laura, and Beth) who listened to and supported my research goals over the past five years. Being a part of the Edwards lab group was a true honor and a wonderful environment in which to learn and grow as a scientist.

I would not be so curious or captivated about the natural world if it wasn't for my parents, Bill and Lindsay Hancock, who have not only spent their lives exploring, appreciating, and protecting the natural world but have instilled this spirit in all of their children. I am undeniably fortunate and forever grateful to my family and to have grown up in a home full of flowers and home-grown veggies, where education and learning were encouraged, and the where world was budding with opportunity. I would not have embarked on this PhD – or have completed it for that matter – without the support and confidence of my parents: thank you!

Finally, much gratitude and love to my husband, Drew. Although we spent most of my PhD apart, Drew was always my intellectual, emotional, and adventuring companion over the past five years. There are few people as bright, thoughtful, and complex as Drew, and I am truly blessed to have a partner who not only read and edited my thesis but provided me with some of the most insightful discussions on photosynthesis and evolution that I had over the course of my PhD.

TABLE OF CONTENTS

PREFACE AND ACKNOWLEDGMENTS.....	viii
LIST OF TABLES.....	xi
LIST OF ILLUSTRATIONS.....	xiii
CHAPTER 1: Phylogeny and the inference of evolutionary trajectories	1
Abstract.....	2
Introduction.....	3
Methods.....	6
Results.....	9
Discussion.....	11
CHAPTER 2: Phylogeny, evolution, and biogeographic history of <i>Calandrinia</i> (<i>Montiaceae</i>)	32
Abstract.....	33
Introduction.....	34
Methods.....	37
Results.....	47
Discussion.....	55
Conclusion.....	61
CHAPTER 3: The evolution of a spectrum of CAM phenotypes in Australian <i>Calandrinia</i> (<i>Montiaceae</i>)	100
Abstract.....	101
Introduction.....	102

Methods.....	106
Results.....	111
Discussion.....	115
CHAPTER 4: CAM biochemistry precedes the evolution of full CAM morphological and anatomical traits in the Portulacineae.....	145
Abstract.....	146
Introduction.....	147
Methods.....	151
Results.....	157
Discussion.....	162

LIST OF TABLES

CHAPTER 1

Table 1A/ 1B. Ordered Q matrices. Examples of Q matrices (equal rates) used to simulate an ordered evolutionary trajectory.....	22
Table 2A/2B. Not ordered Q matrices. Examples of Q matrices (equal rates) used to simulate an unordered evolutionary trajectory.....	23
Table 3. Ancestral reconstruction transition rate matrix.....	24
Supplemental Table S1. Ordered Q matrices.....	30
Supplemental Table S2. Unordered Q matrices.....	31

CHAPTER 2

Appendix S1. Voucher specimen table.....	82
Appendix S2. Bootstrap support values for Montiaceae and Australian <i>Calandrinia</i> relationships.....	83
Appendix S11. Concordance factors from BUCKy analysis.....	92

CHAPTER 3

Table 1. Voucher specimen table and associated measurements.....	138
---	-----

CHAPTER 4

Table 1. Average values for cross sectional anatomical traits, SWC, d13C and titratable acidity measurements.....	180
Table 2. Average values of anatomical traits for photosynthetic mode.....	181

LIST OF ILLUSTRATIONS

CHAPTER 1

- Figure 1.** Interpreting the evolution of intermediate phenotypes along the C4 evolutionary trajectory.....25
- Figure 2.** Schematic representation of methods used to simulate and analyze ordered and unordered evolutionary trajectories.....26
- Figure 3.** Probability of accurately inferring an ordered when the generating data is ordered.....27
- Figure 4.** Probability of accurately inferring an ordered when the generating data is unordered (type 1 error).....28
- Figure 5.** Transition rates between character states 1-4 for the fully sampled, 1000-taxon dataset.....29

CHAPTER 2

- Figure 1.** Species representatives of Australian *Calandrinia*.....76
- Figure 2.** Specimen collections of Australian *Calandrinia*.....77
- Figure 3.** Summary trees depicting the major clades in Montiaceae. Bootstrap support percentages are below branches.....78
- Figure 4.** Bayesian concordance analysis primary concordance tree.....79
- Figure 5.** Australian *Calandrinia* phylogeny reconstructed with the g2 loci dataset (297 loci) with g2 RAxML branch lengths.....80
- Figure 6.** Time-calibrated phylogeny made with BEAST of Australian *Calandrinia* and its sister lineages.....81

Appendix S3. Density plots of Australian <i>Calandrinia</i> internal and terminal branch lengths.....	84
Appendix S4. ASTRAL topology from the g7 loci dataset (116 individuals, 116 loci).....	85
Appendix S5. Maximum likelihood RAxML topology from the g7 loci dataset (116 individuals, 116 loci).	86
Appendix S6. ASTRAL topology from the i98 loci dataset (116 individuals, 74 loci).	87
Appendix S7. Maximum likelihood RAxML topology from the i98 loci dataset (116 individuals, 74 loci).....	88
Appendix S8. ASTRAL topology from the i40 loci dataset (116 individuals, 162 loci).	89
Appendix S9. Maximum likelihood RAxML topology from the i40 loci dataset (116 individuals, 162 loci).	90
Appendix S10. Maximum likelihood RAxML topology from the g2 loci dataset (116 individuals, 297 loci)	91
Appendix S12. Ancestral character state reconstructions for life history in Australian <i>Calandrinia</i>	93
Appendix S13. Ancestral character state reconstructions for fruit type in Australian <i>Calandrinia</i>	94
Appendix S14. Ancestral character state reconstructions for petal number in Australian <i>Calandrinia</i>	95

Appendix S15. Ancestral character state reconstructions for valve number in Australian <i>Calandrinia</i>	96
Appendix S16. Australian <i>Calandrinia</i> Lineage through time plot.....	97
Appendix S17. Bayesian analysis of macro-evolutionary mixtures speciation through time plot.....	98
Appendix S18. Collapsed Montiaceae time-calibrated phylogeny with biogeographic reconstructions.....	99

CHAPTER 3

Figure 1: The Australian <i>Calandrinia</i> coalescent based, RAxML branch length phylogeny.....	139
Figure 2a, b. Summary of nocturnal CO ₂ fixation (ΔH^+ $\mu\text{mol g}^{-1} \text{fw}$) when plants are watered (grey) and droughted (black).....	140
Figure 3a, b, c, d. Net CO ₂ exchange by a leaf for four Australian <i>Calandrinia</i> species.	141
Figure 4. Ancestral character state reconstructions using the 22-tip Australian <i>Calandrinia</i> species tree.....	142
Figure 5. Delta H ⁺ mapped as a continuous character across the 22-tip Australian <i>Calandrinia</i> phylogeny.....	143
Figure S1: Principle component analysis of 8 climatic variables for Australian <i>Calandrinia</i>	144

CHAPTER 4

Figure 1. 42-taxa, four-marker (<i>matK</i> , <i>rbcL</i> , <i>ndhF</i> , and <i>ITS</i>) combined analysis maximum likelihood phylogeny.....	182
---	-----

Figure 2. Anatomical cross-sections of species representing the C ₃ to CAM phenotypic spectrum.....	183
Figure 3. Macro-evolutionary correlations among anatomical traits, photosynthetic mode, venation type, and climatic variables.....	184
Figure 4. Ancestral character state reconstructions across the four-marker ML portullugo phylogeny.....	185
Figure 5. Linear discriminant analysis of 16 BioClim climatic variables for species included in the study.....	186
Figure 6. Ancestral character state reconstructions of MAP (a) and MAT (b) across the four-marker ML portullugo phylogeny.....	187
Figure S1. Ancestral character state reconstructions of carbon isotope ratio (d13C).....	188
Figure S2. Ancestral character state reconstructions of precipitation seasonality.	189

Chapter 1: Phylogeny and the inference of evolutionary trajectories

Lillian Hancock and Erika J. Edwards

Published: Hancock LP and EJ Edwards. 2014. Phylogeny and the inference of evolutionary trajectories. *Journal of Experimental Botany* 65 (13): 3491-3498.

ABSTRACT

Most important organismal adaptations are not actually single traits, but complex trait syndromes that are evolutionarily integrated into a single ‘emergent’ phenotype. Two alternative photosynthetic pathways, C4 photosynthesis and Crassulacean Acid Metabolism (CAM), are primary plant adaptations of this sort, each requiring multiple biochemical and anatomical modifications. Phylogenetic methods are a promising approach for teasing apart the order of character acquisition during the evolution of complex traits, and the phylogenetic placement of intermediate phenotypes as sister taxa to fully optimized syndromes has been taken as good evidence of an ‘ordered’ evolutionary trajectory. But how much power does the phylogenetic approach have to detect ordered evolution? In this study, we simulated ordered and unordered character evolution across a diverse set of phylogenetic trees to understand how tree size, models of evolution, and sampling efforts influence our ability to detect an evolutionary trajectory. We find that small trees (15 taxa) do not contain enough information to correctly infer either an ordered or unordered trajectory, though inference improves as tree size and sampling increases. However, even when working with a 1000-taxon tree, the possibility of inferring the incorrect evolutionary model (type I/ type II error) remains. We urge caution in interpreting the phylogenetic placement of intermediate phenotypes, especially in small lineages. Such patterns can provide a line of evidence for the existence of a particular evolutionary trajectory, but they should be coupled with other types of data to infer the stepwise evolution of a complex character trait.

INTRODUCTION

One of the most compelling and enduring problems in evolution is the origin and assembly of complex character traits or syndromes. How does natural selection gradually modify genes and phenotypes to assemble new and fully integrated suites of characters? How do we reconstruct these serial modifications or ‘steps’ along an evolutionary trajectory? Crassulacean Acid Metabolism (CAM) and C4 photosynthesis in plants are two key examples of such integrated trait syndromes. Evolving numerous times in distantly related lineages, these morphologically and physiologically distinct adaptive syndromes use an internal plant carbon-concentrating mechanism (CCM) that improves the efficiency of C3 photosynthesis under conditions of high heat, drought, and/or low atmospheric CO₂ (Edwards and Ogburn, 2012). The genetic and morphological underpinnings of C4 photosynthesis are especially well studied, and importantly, many ‘intermediate’ phenotypes have been discovered and analyzed (Kennedy and Laetsch, 1974; Powell 1978; Hattersley, 1986; Monson *et al.*, 1986; Monson and Moore, 1989; Kellogg, 1999; Besnard *et al.*, 2009; Feodorova *et al.*, 2010; Christin *et al.*, 2011; Sage, 2012). These C3-C4 intermediates have been essential in the construction of a theoretical ‘C4 evolutionary trajectory’, which maps out the transitions from C3 to C4 metabolism as a series of ordered events, with anatomical changes preceding most of the biochemical changes, and a precursor, photorespiratory pump CCM preceding the activation of PEP Carboxylase and the C4 cycle (Sage, 2004, 2012). When C3, intermediate, and C4 phenotypes are mapped across phylogenies that span the C3 to C4 spectrum, the intermediates are often placed as sister to a fully C4 group, and the intermediate state is then interpreted as the ancestral condition of their shared node (i.e. Molluginaceae,

Neurachninae, and *Flaveria*) (McKown *et al.*, 2005; Christin *et al.*, 2011, 2012). The further interpretation is that to evolve a full C4 syndrome, a lineage must first pass through this particular intermediate condition.

Similar to C4 photosynthesis, ‘intermediate’ CAM-like phenotypes have been discovered and analyzed (Sternberg *et al.*, 1983, 1984; Ting, 1985; Harris and Martin, 1991; Guralnick and Jackson, 2001; Winter *et al.*, 2008; Silveria *et al.*, 2010). These ‘intermediate’ phenotypes fall along a C3 to CAM spectrum and are often difficult to characterize as they are variably expressed and appear indistinguishable from C3 species under non-stressful conditions (Ting, 1985; Cushman, 2001; Dodd *et al.*, 2001, Lüttge, 2004). Nevertheless, hypotheses posit that these CAM-intermediates (CAM cycling and facultative CAM) may act as evolutionary steps between a typical C3 plant and a fully expressed CAM syndrome (Monson, 1989; Guralnick and Jackson, 2001; Sage, 2002; Edwards and Ogburn, 2013), though their phylogenetic distribution with respect to C3 and CAM lineages is less clear.

We agree that there are likely to be a fairly limited set of functional intermediate phenotypes that are frequently passed through during the C3 to C4 transition, and the same is likely to be true during CAM evolution as well. Our question is not whether a trajectory (or a set of trajectories) exists, but rather, how useful is the phylogenetic approach for uncovering them? The situation is more complicated than it appears at first glance (Fig. 1), because all of the information we have on intermediate states are found in currently living species- taxa that are not ancestral to any other taxa. We use the distribution of traits among the living ‘tips’ of phylogenies to infer the tempo and mode of character evolution and to reconstruct past evolutionary events- but in small trees with

few character transitions, there is precious, little information to work with. For example, in the common scenario pictured in Figure 1, it is typically inferred that the intermediate phenotype evolved once, prior to the divergence of the C3-C4 and C4 lineages, and so the extant C3-C4 intermediate represents the ancestral condition of the C4 clade. (e.g., Fig. 1A). That is certainly a plausible scenario, but no more plausible than the alternatives in 1B and 1C- all of these result in only 2 evolutionary transitions, but only figure 1A supports the hypothesized ordered trajectory of C4 evolution. In general, a phylogenetic approach would need a richer ‘tree-trait space’ to discern the relative likelihood of these three alternatives- many more branches, and many more character transitions. While the three taxon case in Figure 1 is an especially information-poor scenario, unfortunately, many empirical datasets do not really contain many more transitions, as intermediate phenotypes are generally rare in comparison to the numbers of species with fully evolved syndromes, and also require significant effort to detect (McKown *et al.*, 2005; Marshall *et al.*, 2007; Vogan *et al.*, 2007; Christin *et al.*, 2011, 2012; Ocampo *et al.*, 2013)

In this study we explored the power of the phylogenetic approach to identify ordered and unordered evolutionary trajectories by simulating character evolution under different evolutionary scenarios and across trees of various sizes and shapes. Because complete taxon sampling of a focal clade is very rare in phylogenetic studies, we also investigated how limited taxon sampling across a tree influences the evolutionary signal of these ordered and unordered data sets. We tested for both the ability to detect an ordered evolutionary trajectory when it exists, as well as the possibility of inferring an ordered trajectory when there isn’t one. Our findings suggest that the taxonomic scale of many phylogenetic studies is inadequate for inferring ordered evolution, and surprisingly,

increasing tree size and the number of evolutionary transitions has a relatively modest positive effect on the accuracy of model inference. On the positive side, limited taxon sampling does not seem to greatly reduce the overall evolutionary signal of the data sets.

METHODS

Simulating Character Evolution

To simulate ordered, discrete character data, eight unique Q matrices were built representing four character states along an evolutionary trajectory: state 1 $\rightarrow$ state 2 $\rightarrow$ state 3 $\rightarrow$ state 4. A Q matrix is an instantaneous transition rate matrix (Table 1; Table 2) where each cell within the matrix corresponds to the probability of transitioning between two states. To specify order within the Q matrix, an instantaneous transition rate of ‘zero’ was assigned to all matrix cells that represent a ‘skip’ across states (e.g., 1 $\rightarrow$ 3, 2 $\rightarrow$ 4, etc.), whereas possible transitions were assigned integers 1-3 (Table 1); integers correspond to relative, but not absolute, transition rates. Since all rows must sum to zero in a Q matrix, the diagonal cells (e.g. 1 $\rightarrow$ 1, 2 $\rightarrow$ 2) within the matrix are equal to the sum of the off-diagonal elements within the rows (i.e. the diagonal cells can be negative so that rows sum to zero). The eight matrices were built to simulate different models of character evolution (i.e. with all transitions occurring at equal rates (ER), with all rates different (ARD), with forward and reverse transitions between two characters at equal rates (SYM), and with high transition rates out of intermediate character states) (Supplemental Table S1). Integers 1-3 were placed in matrix cells to reflect these four models of evolution. Four of the eight matrices were ordered-reversible while the remaining matrices were not reversible (meaning that a transition from state 1 to state 2

was allowed, but a transition from state 2 to state 1 was not allowed) (Table 1b). All analyses were done run on the R Project for Statistical Computing (R).

We also simulated unordered trait evolution using the same four character states with the goal of understanding the probability of incorrectly inferring an ordered trajectory. Q matrices were constructed to imply unordered evolution between each of the four character states, 1, 2, 3, and 4 (Table 2; Supplemental Table S2). In other words, transitions between states 1 to 3, 1 to 4, and 2 to 4 (impossible under the ordered model) were allowed to happen. All matrix cells were assigned integers 1, 2, or 3 (Table 2a) unless the matrix was non-reversible (4 of the 8 simulating matrices) (Table 2b).

Using the *rtree* function in the APE package (Paradis *et al.*, 2004), twenty unique ultrametric, coalescent 15-taxon, 50-taxon, 100-taxon, and 1000-taxon trees were randomly generated for the ordered and the unordered character simulations (Fig. 2). We acknowledge that trees generated with the coalescent generally capture a smaller fraction of tree shape than do birth-death models (Mooers *et al.* 2007). With the *sim.char* function in the Geiger package (Harmon *et al.*, 2008), character evolution was simulated across each tree using our constructed matrices (Fig 2). For a given Q matrix and tree, character data was independently simulated 200 times, and the first ten simulations containing all four states at the tree tips (1, 2, 3, 4) were selected for downstream analyses. More data was simulated than used because all four character states were not always present across extant tips in each simulation.

Sensitivity analysis

To test for the effect of poor sampling on our ability to infer character evolution, we generated a second, ‘pruned’ dataset where 50% of the taxa (tips) and corresponding

data were randomly removed from each of the original simulated data sets across ordered and unordered 50-taxon, 100-taxon, and 1000-taxon trees (Fig. 2). Fifty percent of the taxa were not removed from the 15-taxon data set as this tree is already quite small.

Approximately 1200 ordered and unordered data simulations were generated for each of the 15-taxon, 50-taxon, 100-taxon, and 1000-taxon “groups”. This number is lower than the expected 1600 (20 trees x 8 matrices x 10 simulations of each matrix) because simulated data was removed when 1) all four character states were not realized among the 10 simulations when 50% of the data/ tips were removed, and 2) when the transition rates between states were too infinitely low (but not zero) to calculate (non-linear minimization error). When these errors ensued, simulations were removed from the data sets so the analysis could continue to run.

Inferring character transition rates of simulated data

In this next step, we used the simulated data as if we were beginning a typical phylogenetic analysis, with only a phylogeny and character data at the tips and no a priori knowledge of the underlying model of character evolution. Using the *ace* function in the package APE (Paradis *et al.*, 2004), we calculated the maximum likelihood estimates of each transition rate within the four-by-four matrix under a model that allowed each rate to vary independently (Fig. 2). Since the diagonal cells within the matrix are the sum of the off-diagonal elements within the rows, twelve transition rates corresponding to the remaining matrix cells (A-L) were saved to relevant tables (Table 3), which could then be compared to expected rates according to the generating matrix.

Data analysis

Our primary goal was to evaluate how well phylogenetic methods can detect an evolutionary trajectory using only the phylogenetic relationships and trait distributions of living taxa. We focused our analyses on only comparing the forward transition rates among character states (cells A, B, C, E, F, and I in Table 3). Matrix cells B (transition of state 1 $\rightarrow$ state 3), C (state 1 $\rightarrow$ state 4), and F (state 2 $\rightarrow$ state 4) represent the transition rates that were expected to be zero when simulated under an ordered model (Table 1). For each simulation, we evaluated how close to zero these rates were, using multiple thresholds for designating an estimated rate as essentially ‘zero’. In the most conservative threshold, B, C, and F cells of the estimated matrix all had to be zero for the inferred evolutionary model to be considered ‘ordered’. Using a more liberal threshold, all B, C, and F cells must be lower than 0.1; and in the most liberal threshold, all B, C, and F cells must have rates that were at least one magnitude lower than the average transition rate of cells A, E, and I. We then designated each simulated dataset to be an inferred ‘ordered’ or ‘unordered’ evolutionary model according to each threshold, and compared this to the actual model used to simulate the data.

RESULTS

Using our most conservative threshold (all ‘zero’ cells must have inferred transition rates to be zero), it was exceedingly difficult to infer an ordered evolutionary trajectory, with only 22% of our 15-taxon tree simulations accepting an ordered model, which increased to only 44% in a 1000-taxon tree (Fig. 3). Relaxing the threshold of rates to be larger than zero, but close to zero, greatly improved our ability to infer the generating model: using a threshold where the ‘zero’ cells must be inferred to be at least

one order of magnitude lower than the average rate of the ‘non-zero’ cells, we could classify many more of our ordered simulations as ordered, up to 60% in the 15-taxon case and 83% in the 1000-taxon case.

However, relaxing the threshold had the additional complicating effect of also drastically increasing the frequency of assigning an ‘ordered’ model to simulations that were actually generated using an unordered matrix (type I error) (Fig. 4). Our conservative threshold produced negligible type I error, with the exception of the 15-taxon tree (5.7% acceptance of an ‘ordered’ model when the underlying data were unordered). However, relaxing the threshold consistently increased our propensity to label unordered data as ordered, from 14% in our 1000-taxon trees to an alarming 22% in our 15-taxon trees.

Increasing tree size consistently reduced the probability of both type I and type II error, but this improvement was surprisingly modest. The 656% increase in tree size between the 15-taxon and the 1000-taxon groups resulted in only a 20% gain in inference accuracy. In a similar vein, reducing taxon sampling produced its predicted effect of reducing accuracy, but the magnitude of the effect was small, typically only reducing the number of correctly inferred models by several percentage points (Figs 3,4). The effect of missing data was larger in smaller datasets: in the 50-taxon dataset, simulating missing data resulted in an average 12% decrease in accuracy across the three thresholds, but this reduction was only 3% in the 1000-taxon dataset.

We also looked at the individual transition rates generated from each of the eight Q matrices across all tree size classes separately, to determine whether data simulated under particular models of evolution (inferred from the given Q matrices; Supplemental

Table S1) were more indicative of an ordered evolutionary trajectory. No general patterns were detected across all matrices, and consistent with our threshold analyses, transition rates were generally much higher in the ‘evolving’ cells, while the cells not allowed to evolve had rates close to or at zero (Fig. 5). By looking at the individual rates we were able to detect a consistent difference among the ‘zero’ cells: cell C, which represents a transition from state 1 to state 4, thereby skipping all intermediate character states, had much lower transition rates across all matrices than the other ‘zero’ cells (transitions 1 → 3, or 2 → 4). This pattern was observed consistently across all tree by matrix combinations.

DISCUSSION

Often, we use the distribution of ‘intermediate’ character states across the tips of a phylogeny to infer the stepwise evolution of integrated trait syndromes, such as C4 photosynthesis. In this study we asked, how much power does the phylogenetic approach have to detect ordered trait evolution? By simulating ordered and unordered character evolution across a diverse set of phylogenetic trees, we were able to investigate how tree size, model of evolution, and sampling efforts influence the ability of standard phylogenetic comparative tools to detect an evolutionary trajectory.

Our findings highlight some significant limitations to this approach. Under the most data-poor scenario (a 15-taxon tree) our methods of inference were, at best, 83% effective at detecting ordered evolution- but this came with the serious cost of also *incorrectly* inferring order over 20% of the time. Increasing clade size improved the situation, but not as much as one might hope. Two results are somewhat uplifting. First, increasing clade size from 15 to 50 tips provided nearly the same benefit as increasing to

100 tips, suggesting that the biggest returns relative to investments might be at the small end of the spectrum. Second, in very large clades (e.g., 1000 taxa) the effect of missing data appears to be quite negligible. This is especially good news for studies of CAM evolution, as many CAM-evolving clades are overwhelmingly speciose (e.g., orchids, bromeliads) and exhaustive taxon sampling is, at this stage, simply not feasible. While this is heartening, it comes with the caveat that our ‘missing data’ were simulated randomly with respect to character state. In both C4 and CAM syndromes, we are likely to know much more about the distribution of fully evolved pathways, as they are easily detectable with stable carbon isotopes (Bender, 1971; Bender *et al.*, 1973; Sternberg *et al.*, 1984; O’leary, 1988; Farquhar *et al.*, 1989). The sampling in empirical studies may therefore be biased toward including known C4 and CAM plants at the exclusion of other (potentially still unidentified) intermediates, and this particular pattern of sampling bias was not addressed in our study.

Admittedly, there is more than one way to infer an ‘ordered trajectory’ based on phylogenetic patterns, and our analyses are fairly simplistic. In future work it might be useful to compare various phylogenetic approaches. An alternative to our estimated rates approach might be an actual tally of inferred transitions, using stochastic mapping (Bollback, 2006; for implementation in grasses see Roalson, 2012) or parsimony reconstruction. Another might be to simply compare the fit of a suite of character evolution models and choose the best-fit model with likelihood scores. We tested this alternative approach on a subset of our data using the Geiger module `fitDiscrete` in R (Harmon *et al.*, 2008). Preliminary analyses indicate that our rate matrix approach may work better with smaller trees, and the model-fitting approach increases in accuracy more

quickly as trees become larger. At the same time (and possibly related), a model-fitting approach appears more sensitive to poor taxon sampling. While these different methods clearly need further examination, one benefit to the approach we took here is that it allows for a closer look at all the possible transitions, and thus allows for a better intuition of where the methods are working and where they are not. For instance, we found that matrix cell C (Table 3), which represents a transition from state 1 $\rightarrow$ state 4, had the lowest transition rate across all simulations (Fig. 3). This is an important detail, as it suggests that type II error is likely caused by incorrectly inferring transitions into and out of intermediate states. Thus, in the 4 character state scenario that we tested, our methods successfully identified that intermediate states were passed through from the initial first state to the final fourth state, however, these methods were not successful in identifying the exact order of the intermediate states.

We still feel that the phylogenetic placement of intermediate states is an *essential* component of inferring evolutionary trajectories; we have simply illustrated that phylogenetic inference, like anything else, is not infallible, and should not be given any sort of ‘primary’ importance when evaluating different scenarios supported by different kinds of evidence. A key example of this ‘priority’ type of treatment is the recent study by Ocampo *et al.* (2013), who found a C3-C4 intermediate lineage nested within C4 *Portulaca*. In spite of acknowledging the anatomical and biochemical differences among the C4 lineages that surround this intermediate, the authors seem to eventually believe that, because ancestral state reconstruction supported a reversion from C4 to C3-C4, this is the order of trait evolution that we should accept. Our simulation study suggests that the number of character transitions in *Portulaca* are too few to provide the information

necessary to feel confident in a phylogenetic reconstruction. And indeed, as evidence grows about this particular example of C4 evolution, it seems increasingly clear that there have been multiple parallel realizations of C4 in different *Portulaca* clades (Christin *et al.* in review).

Two recent studies (Heckman *et al.*, 2013; Williams *et al.*, 2013) use alternative approaches to tease apart the assembly and evolution of C4 photosynthesis. Heckman *et al.* (2013) inferred the relative fitness gain of each biochemical change along the C4 trajectory, using modeled photosynthetic rate as a measure of fitness, to create the ‘adaptive landscape’ of C4 evolution. They found that the intermediate C3-C4 phenotypes are indeed transitory states, as the relative fitness gain associated with the realization of these intermediate phenotypes were minor compared to the fitness gain of a fully optimized C4 syndrome. Williams *et al.* (2013) took an entirely different approach, and using empirical data from 43 studies, characterized 16 biochemical, anatomical and cellular characteristics associated with C4 photosynthesis and C3-C4 intermediates to build a transition network connecting C3 and C4 photosynthesis. Then, using a bayesian approach, they sampled the biologically relevant ‘paths’ through the network to infer the order in which phenotypic changes occur along the C4 trajectory. Findings from this study suggest that the evolutionary trajectory of C4 photosynthesis is somewhat ordered but trait acquisition along the trajectory is flexible.

Both of these studies were explicitly ‘non-phylogenetic’ - and we see the incorporation of phylogenetic information, especially in the approach of Williams *et al.*, as an exciting opportunity for improvement. Unquestionably, phylogeny has a central role in the study of both C4 and CAM evolution - it identifies new lineages for research,

can provide a timeline for transitions, and identifies ecological and organismal correlates. As phylogenetic biologists, we truly think that phylogeny can and should be integrated into all areas of biological research. It is simply important to understand the caveats and limitations inherent in inferring past evolutionary events with only a handful of data from currently living species. A phylogenetic approach provides valuable evidence for a particular evolutionary trajectory, but should be considered for what it is- an informed hypothesis, that can be supported or refuted with additional data.

ACKNOWLEDGEMENTS

The authors would like to thank Radika Bhaskar, David Chatelet, Pascal-Antoine Christin, Abigail Moore, Jurriaan de Vos, and Felipe Zapata for their programming assistance and helpful comments. This work was funded in part by a National Science Foundation grant DEB-1252901 to EJE.

LITERATURE CITED

Bender MM. 1971. Variation in $^{13}\text{C}/^{12}\text{C}$ ratios of plants in relation to the pathway of photosynthetic carbon dioxide fixation. *Phytochemistry* **10**, 1239–1344.

Bender MM, Ronham I, Vines HM, Black CC. 1973. $^{13}\text{C}/^{12}\text{C}$ ratio changes in Crassulacean acid metabolism plants. *Plant Physiology* **53**, 427–430.

Besnard G, Muasya AM, Russier F, Roalson EH, Salamin N, Christin PA. 2009. Phylogenomics of C_4 photosynthesis in sedges (Cyperaceae): multiple appearances and genetic convergence. *Molecular Biology and Evolution* **26**, 1909–1919.

Bollback JP. 2006. SIMMAP: Stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics* **7**, 88.

Christin PA, Freckleton R P, and Osborne, CP. 2010. Can phylogenetics identify C_4 origins and reversals? *Trends in Ecology & Evolution* **25**, 403–409.

Christin P, Sage TL, Edwards EJ, Ogburn RM, Khoshravesh R, Sage RF. 2011. Complex evolutionary transitions and the significance of C_3 - C_4 intermediate forms of photosynthesis in Molluginaceae. *Evolution* **65**, 643–660.

Christin PA, Wallace MJ, Clayton H, Edwards EJ, Furbank RT, Hattersley PW, Sage RF, Macfarlane TD, Lugwid M. 2012. Multiple photosynthetic transitions, polyploidy, and lateral gene transfer in the grass subtribe Neurachninae. *Journal of Experimental Botany* **63**, 6297–6308.

Christin PA, Arakaki M, Osborne CP, Brautigam A, Sage RF, Hibberd JM, Kelly S, Covshoff S, Wong G, Edwards EJ. 2014. Shared origins of a key enzyme during the evolution of C_4 and CAM metabolism. *Journal of Experimental Botany*, in press.

- Cushman, JC.** 2001. Crassulacean acid metabolism. A plastic photosynthetic adaptation to arid environments. *Plant Physiology* **127**, 1439–1448.
- Dodd AN, Borland AM, Haslam RP, Griffiths H, Maxwell K.** 2002. Crassulacean acid metabolism: plastic, fantastic. *Journal of experimental botany* **53**, 569–580.
- Edwards EJ, Ogburn RM.** 2012. Angiosperm responses to a low-CO₂ world: CAM and C₄ Photosynthesis as Parallel Evolutionary Trajectories. *International Journal of Plant Sciences* **173**, 724–727.
- Farquhar GD, Ehleringer JR, Hubick KT.** 1989. Carbon isotope discrimination and photosynthesis. *Annual Review of Plant Physiology* **40**, 503–537.
- Feodorova TA, Voznesenskaya EV, Edwards GE, Roalson EH.** 2010. Biogeographic patterns of diversification and the origins of C₄ in Cleome (Cleomaceae). *Systematic Botany* **35**, 811–826.
- Guralnick LJ, Jackson MD.** 2001. The occurrence and phylogenetics of Crassulacean acid metabolism in the Portulacaceae, *International Journal of Plant Sciences* **162**, 257–262.
- Hattersley PW, Wong SC, Perry S, Roksandic Z.** 1986. Comparative ultrastructure and gas-exchange characteristics of the C₃–C₄ intermediate *Neurachne minor*. *Plant, Cell, and Environment* **9**, 217–233.
- Harris FS, Martin CE.** 1991. Correlation between CAM-Cycling and photosynthetic gas exchange in five species of *Talinum* (Portulacaceae). *Plant Physiology* **96**, 1118–1124.
- Harmon LJ, Weir JT, Brock CD, Glor RE, and Wendell C.** 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics* **24**, 129–131.

- Heckmann D, Schulze S, Denton A, Gowik U, Westhoff P, Weber APM, Lercher MJ.** 2013. Predicting C4 photosynthesis evolution: modular, individually adaptive steps on a Mount Fuji fitness landscape. *Cell* **153**, 1579–1588.
- Kadereit G, Freitag H.** 2011. Molecular phylogeny of Camphorosmeae (Camphorosmoideae, Chenopodiaceae): implications for biogeography, evolution of C4-photosynthesis and taxonomy. *Taxon* **60**, 51–78.
- Kellogg EA.** 1999. Phylogenetic aspects of the evolution of C4 photosynthesis. In: Sage FR, Monson RK, eds. *C4 Plant Biology*. San Diego: Academic, pp. 411–44.
- Kennedy RA, Laetsch WM.** 1974. Plant species intermediate for C3, C4 photosynthesis. *Science* **184**, 1087–1089.
- Lüttge U.** 2004. Ecophysiology of Crassulacean Acid Metabolism (CAM). *Annals of Botany* **93**, 629–652.
- Marshall DM, Muhaidat R, Brown N, Liu Z, Stanley S, Griffiths H, Sage R, Hibberd J.** 2007. Cleome, a genus closely related to Arabidopsis, contains species spanning a developmental progression from C3 to C4 photosynthesis. *The Plant Journal* **51**, 886–896.
- McKown AD, Moncalvo JM, Dengler NG.** 2005. Phylogeny of Flaveria (Asteraceae) and inference of C4 photosynthesis evolution. *American journal of botany* **92**, 1911–1928.
- Monson R.** 1989. On the evolutionary pathways resulting in C4 photosynthesis and crassulacean acid metabolism (CAM), *Advances in Ecological Research* **19**, 57–110.

- Monson RK, Moore BD.** 1989. On the significance of C3–C4 intermediate photosynthesis to the evolution of C4 photosynthesis. *Plant, Cell, and Environment*. **12**, 689–699.
- Mooers, AO, Harmon LJ, Blum MGB, Wong DHJ, Heard, SB.** 2007. Some models of phylogenetic tree shape. Pp 149-170 in *Reconstructing Evolution: new mathematical and computational advances*. Ed. O Gascuel and M Steel, Oxford University Press, Oxford.
- Ocampo G, Koteyeva NK, Voznesenskaya E V, Edwards GE, Sage TL, Sage RF, Columbus JT.** 2013. Evolution of leaf anatomy and photosynthetic pathways in *Portulacaceae*. *American Journal of Botany* **100**, 1–15.
- O’leary MH.** 1988. Carbon isotopes in photosynthesis. *BioScience* **38**, 328–3.
- Paradis E, Claude J, Strimmer K.** 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290.
- Powell AM.** 1978. Systematics of *Flaveria* (*Flaveriinae* Asteraceae). *An. Mo. Bot. Gard.* **65**, 590–636.
- R Development Core Team.** 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Roalson, E. H.** 2011. C4 photosynthesis origins in the monocots: a review and reanalysis. Pp. 319–338 in *C4 Photosynthesis and Related CO2 Concentrating Mechanisms*. Raghavendra, A. S. and R. F. Sage (eds.). Dordrecht, The Netherlands, Kluwer.

- Smith JAC, Winter K.** 1996. Taxonomic distribution of crassulacean acid metabolism. In: Winter K, Smith JAC (eds) Crassulacean acid metabolism. Biochemistry, ecophysiology and evolution. Springer, 427–436.
- Sage RF.** 2001. Environmental and Evolutionary Preconditions for the Origin and Diversification of the C₄ Photosynthetic Syndrome. *Plant Biology* **3**, 202–213.
- Sage RF.** 2002. Are crassulacean acid metabolism and C-4 photosynthesis incompatible? *Functional Plant Biology* **29**, 775–785.
- Sage RF.** 2004. The evolution of C₄ photosynthesis. *New Phytologist* **161**, 341–470.
- Sage RF, Sage TL, Kocacinar F.** 2012. Photorespiration and the evolution of C₄ photosynthesis, *Annu. Rev. Plant Biol.* **63**, 19–47.
- Silvera K, Neubig K, Whitten W, Williams N, Winter K, Cushman J.** 2010. Evolution along the crassulacean acid metabolism continuum. *Functional Plant Biology* **37**: 995–1010.
- Sternberg L, Deniro MJ, Johnson HB.** 1984. Isotope ratios of cellulose from plants having different photosynthetic pathways. *Plant Physiology* **74**, 557–561.
- Ting, IP.** 1985. Crassulacean Acid Metabolism. *Ann. Rev. Plant Physiol.* **36**, 595–622.
- Vogan PJ, Frohlich MW, Sage RF.** 2007. The functional significance of C₃–C₄ intermediate traits in *Heliotropium* L. (Boraginaceae): gas exchange perspectives. *Plant, Cell, and Environment* **30**, 1337–1345.
- Williams BP, Johnston IG, Covshoff S, Hibberd JM.** 2013. Phenotypic landscape inference reveals multiple evolutionary paths to C₄ photosynthesis. *eLife* **2**, e00961.

Winter K, Garcia M, Holtum M. 2008. On the nature of facultative and constitutive CAM: environmental and developmental control of CAM expression during early growth of *Clusia*, *Kalanchoe*, and *Opuntia*. *Journal of Experimental Botany* **59**, 1829–1840.

Table 1A/ 1B. Ordered Q matrices. Examples of Q matrices (equal rates) used to simulate an ordered evolutionary trajectory. Character evolution was modeled using a continuous-time markov model, where each transition between character states is assigned an instantaneous rate. Table 1A is an ordered reversible Q matrix and Table 1B a nonreversible Q matrix.

Table A	State 1	State 2	State 3	State 4
State 1	-1	1	0	0
State 2	1	-1	1	0
State 3	0	1	-1	1
State 4	0	0	1	-1

Table B	State 1	State 2	State 3	State 4
State 1	-1	1	0	0
State 2	0	-1	1	0
State 3	0	0	-1	1
State 4	0	0	0	0

Table 2A/ 2B. Not Ordered Q Matrices. Example Q matrices (equal rates) used to simulate an unordered evolutionary trajectory. Character evolution was modeled using a continuous-time markov model, where each transition between character states is assigned an instantaneous rate. Table 1A is an unordered reversible Q matrix and Table 1B a nonreversible Q matrix.

Table A	State 1	State 2	State 3	State 4
State 1	-3	1	1	1
State 2	1	-3	1	1
State 3	1	1	-3	1
State 4	1	1	1	-3

Table B	State 1	State 2	State 3	State 4
State 1	-3	1	1	1
State 2	0	-2	1	1
State 3	0	0	-1	1
State 4	0	0	0	0

Table 3. Ancestral Reconstruction Transition Rate Matrix. Letters A through L refer to the transition rates generated for each ancestral character state reconstruction. Transitions BCF and GJK are not possible under an ordered model. All transitions are possible under an unordered model.

	State 1	State 2	State 3	State 4
State 1	-	A	B	C
State 2	D	-	E	F
State 3	G	H	-	I
State 4	J	K	L	-

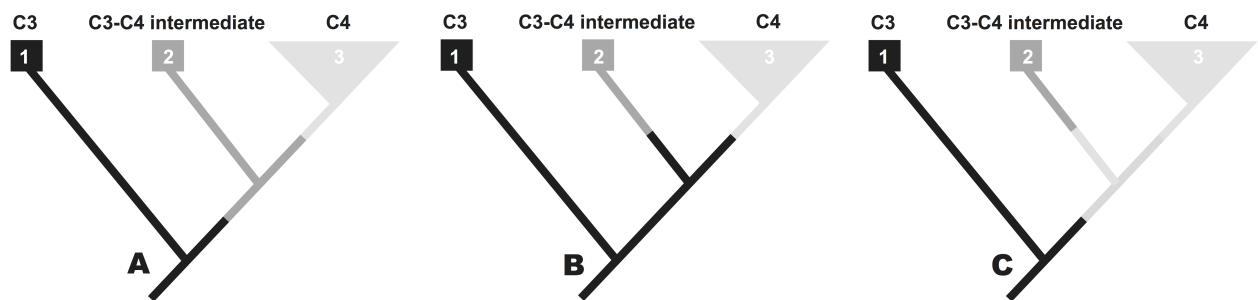


Figure 1: Interpreting the evolution of intermediate phenotypes along the C4 evolutionary trajectory. 1A) The intermediate phenotype (C3-C4 intermediate) evolved once, prior to the divergence of lineages 2 and 3, and so the extant intermediate represents the ancestral condition of lineage 3, the C4 clade. 1B) The C3-C4 intermediate and the C4 phenotype each evolved once after the divergence of lineage 2 and 3, and so the extant intermediate does not represent the ancestral condition of lineage 3 (C4 clade) rather the C3 phenotype represents the ancestral condition of both lineage 2 and 3. 1C) The C4 phenotype evolved prior to the divergence of lineage 2 and 3 and the C3-C4 intermediate evolved after the divergence of lineage 2. In this scenario, the C4 phenotype represents the ancestral condition of the C3-C4 intermediate.

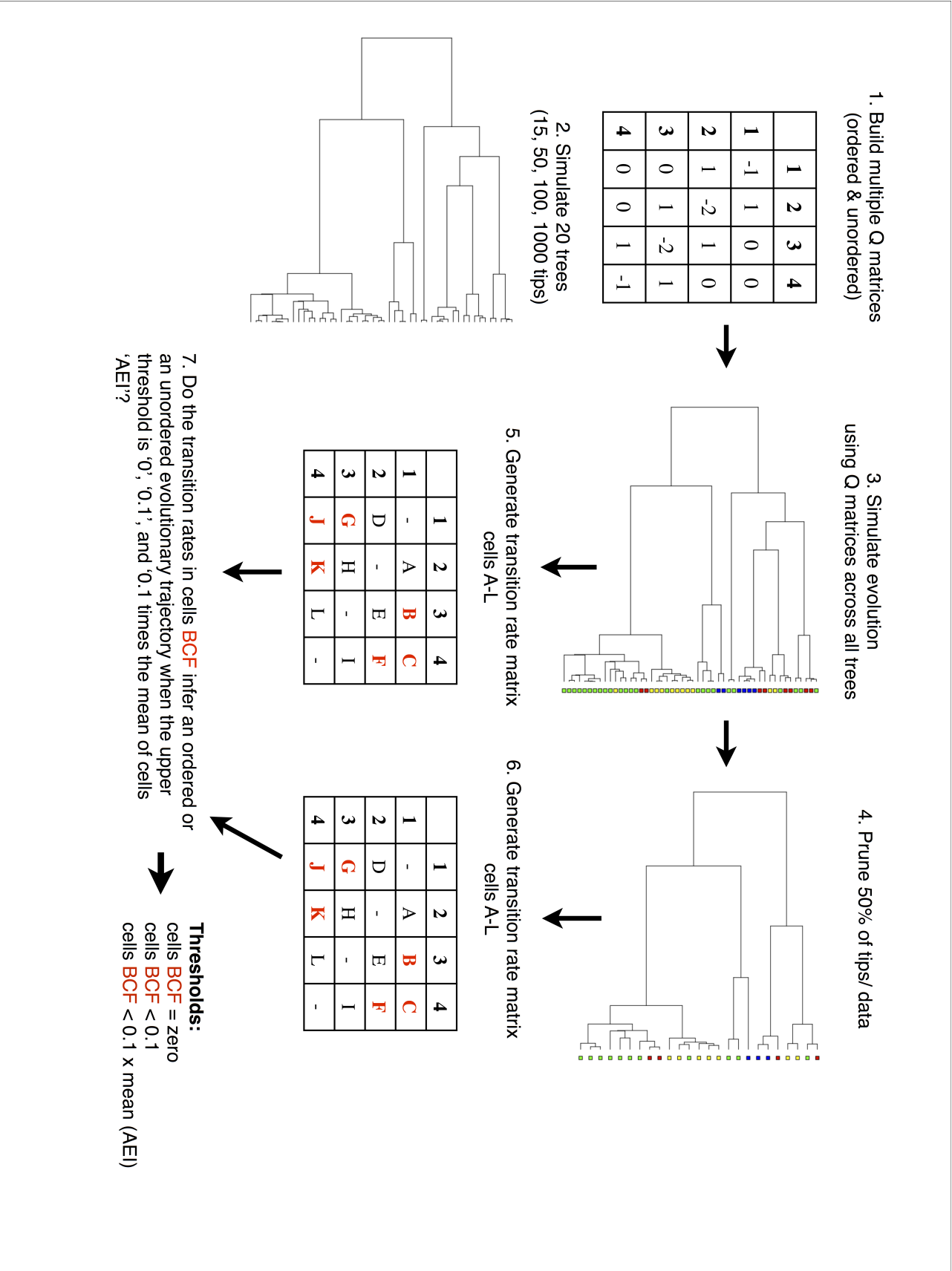


Figure 2: Schematic representation of methods used to simulate and analyze ordered and unordered character data.

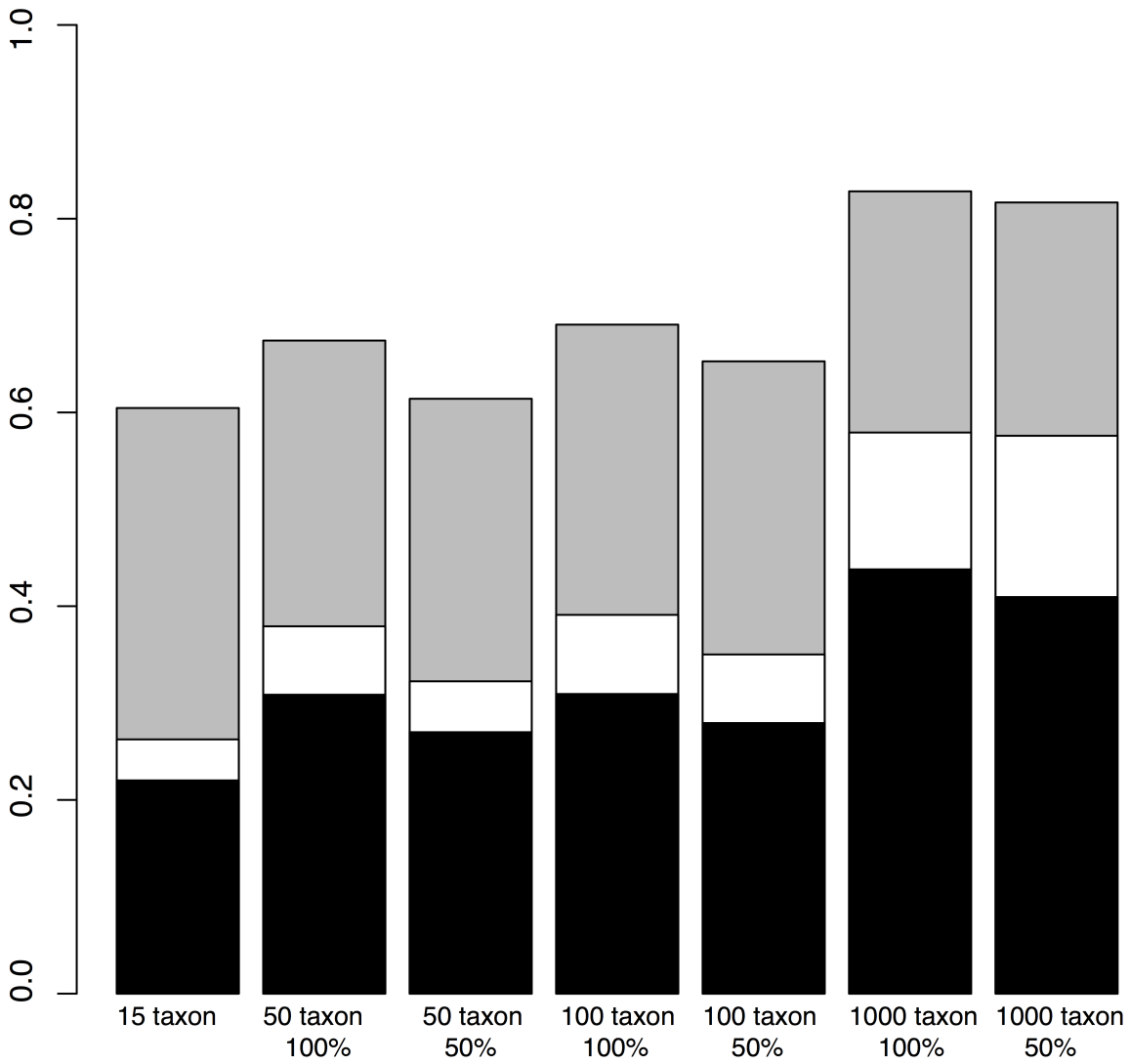


Figure 3: Bars represent the probability of accurately inferring an ordered trajectory when the generating data is ordered for each taxon group at the three predefined thresholds: 1) cells B, C, and F must be 0 (black); 2) cells B, C, and F must be less than 0.1 (white); cells B, C, and F cells have rates at least one magnitude less than the average transition rate of cells A, E, and I (grey).

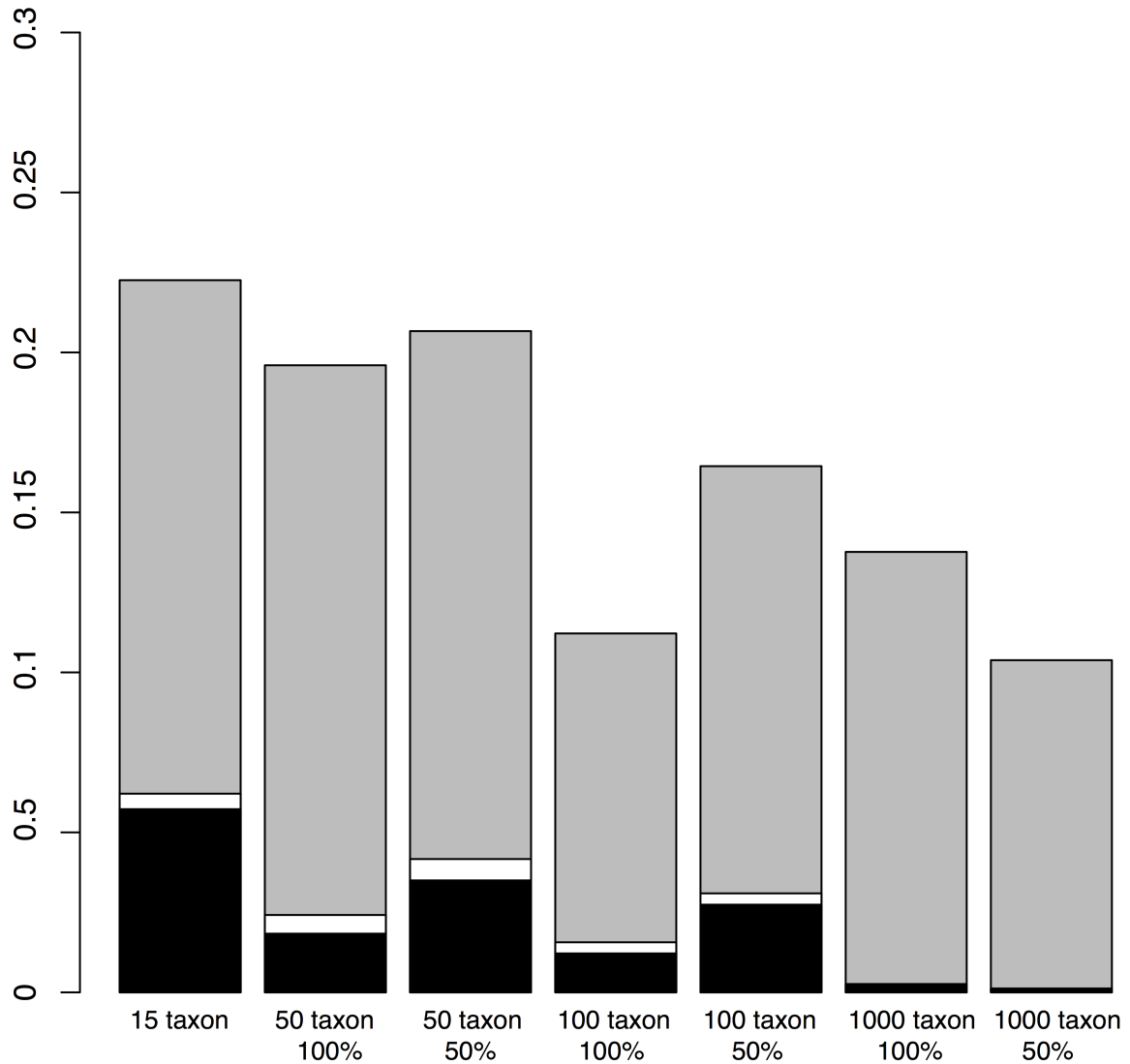


Figure 4: Bars represent the probability of inferring an ordered evolutionary trajectory when the generating data is unordered (type I error) for each taxon group at the three predefined thresholds: 1) cells B, C, and F must be 0 (black); 2) cells B, C, and F must be less than 0.1 (white); cells B, C, and F cells have rates at least one magnitude less than the average transition rate of cells A, E, and I (grey).

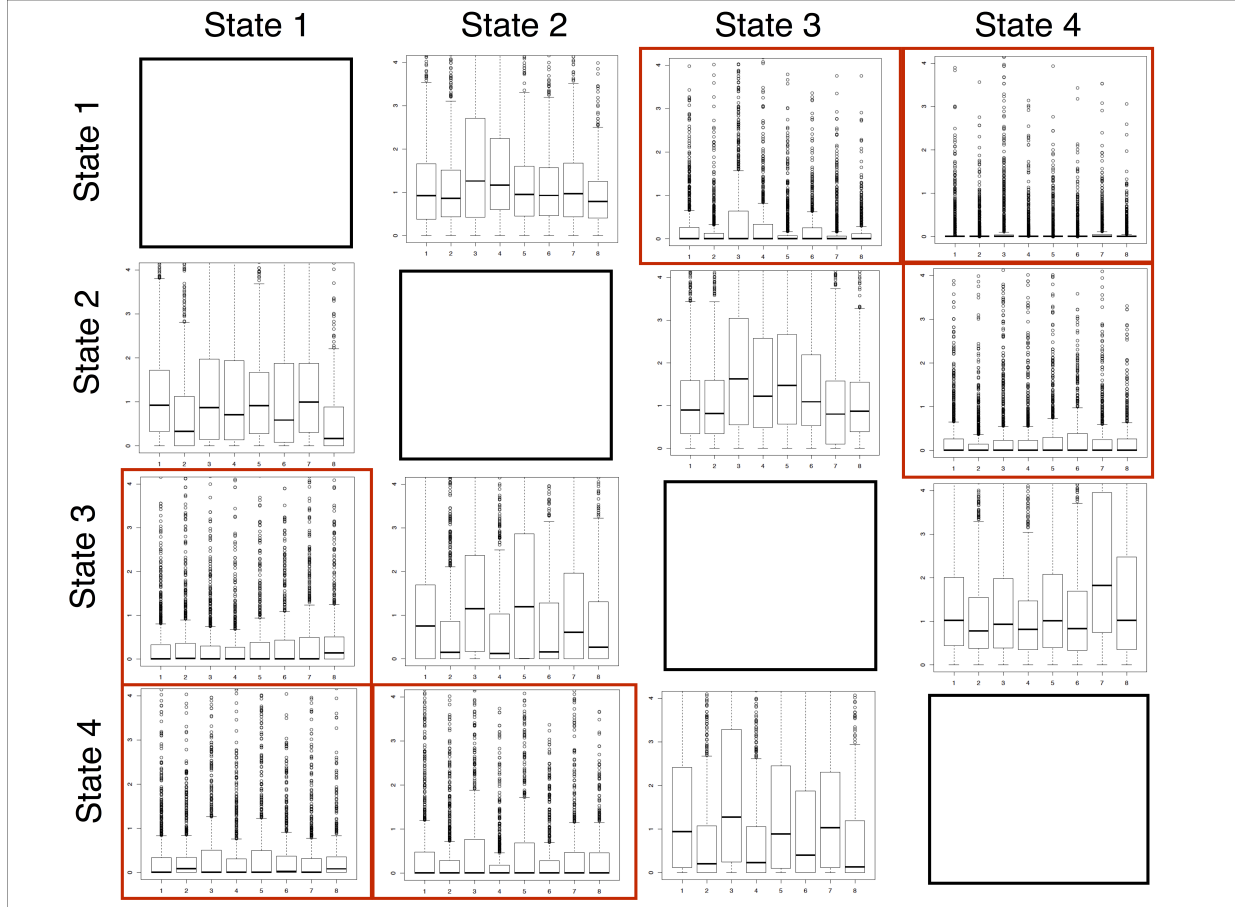


Figure 5: Four-by-four matrix represents the transition rates between character states 1-4 for the fully sampled, 1000-taxon ordered dataset. Boxplots within each cell represent transition rates for each of the eight matrices. Forward and reverse “zero cells” (not allowed to evolve under an ordered model) are outlined.

Supplemental Table S1. Ordered Q matrices.

Matrix 1	State 1	State 2	State 3	State 4
State 1	-1	1	0	0
State 2	1	-2	1	0
State 3	0	1	-2	1
State 4	0	0	1	-1

Matrix 2	State 1	State 2	State 3	State 4
State 1	-1	1	0	0
State 2	0	-1	1	0
State 3	0	0	-1	1
State 4	0	0	0	0

Matrix 3	State 1	State 2	State 3	State 4
State 1	-2	2	0	0
State 2	1	-4	3	0
State 3	0	2	-3	1
State 4	0	0	3	-3

Matrix 4	State 1	State 2	State 3	State 4
State 1	-2	2	0	0
State 2	0	-3	3	0
State 3	0	0	-1	1
State 4	0	0	0	0

Matrix 5	State 1	State 2	State 3	State 4
State 1	-1	1	0	0
State 2	1	-3	2	0
State 3	0	2	-3	1
State 4	0	0	1	-1

Matrix 6	State 1	State 2	State 3	State 4
State 1	-1	1	0	0
State 2	0	-2	2	0
State 3	0	0	-1	1
State 4	0	0	0	0

Matrix 7	State 1	State 2	State 3	State 4
State 1	-1	1	0	0
State 2	1	-2	1	0
State 3	0	1	-4	3
State 4	0	0	1	-1

Matrix 8	State 1	State 2	State 3	State 4
State 1	-1	1	0	0
State 2	0	-1	1	0
State 3	0	0	-3	3
State 4	0	0	0	0

Supplemental Table S2. Unordered Q matrices

Matrix 1	State 1	State 2	State 3	State 4
State 1	-3	1	1	1
State 2	1	-3	1	1
State 3	1	1	-3	1
State 4	1	1	1	-3

Matrix 2	State 1	State 2	State 3	State 4
State 1	-3	1	1	1
State 2	0	-2	1	1
State 3	0	0	-1	1
State 4	0	0	0	0

Matrix 3	State 1	State 2	State 3	State 4
State 1	-5	2	1	2
State 2	1	-5	3	1
State 3	2	2	-7	3
State 4	1	1	3	-4

Matrix 4	State 1	State 2	State 3	State 4
State 1	-5	2	1	2
State 2	0	-4	3	1
State 3	0	0	-3	3
State 4	0	0	0	0

Matrix 5	State 1	State 2	State 3	State 4
State 1	-6	1	2	3
State 2	1	-5	2	2
State 3	2	2	-5	1
State 4	3	2	1	-6

Matrix 6	State 1	State 2	State 3	State 4
State 1	-6	1	2	3
State 2	0	-4	2	2
State 3	0	0	-1	1
State 4	0	0	0	0

Matrix 7	State 1	State 2	State 3	State 4
State 1	-5	1	1	3
State 2	1	-5	1	3
State 3	1	1	-5	3
State 4	1	1	1	-3

Matrix 8	State 1	State 2	State 3	State 4
State 1	-5	1	1	3
State 2	0	-4	1	3
State 3	0	0	-3	3
State 4	0	0	0	0

Chapter 2: Phylogeny, evolution, and biogeographic history of
***Calandrinia* (Montiaceae)**

Lillian P. Hancock, Frank Obbens, Abigail J. Moore, Kevin Thiele, Jurriaan M. de Vos,
Judy West, Joseph A. M. Holtum, and Erika J. Edwards

Accepted: Hancock LP, Obbens F, Moore AJ, Thiele K, de Vos JM, West J, Holtum
JAM, and EJ Edwards. 2018. Phylogeny, evolution, and biogeographic history of
Calandrinia (Montiaceae). *American Journal of Botany*.

Abstract

PREMISE OF THE STUDY: *Calandrinia* are small, succulent herbs that have broad variation in habitat, morphology, life history, and photosynthetic metabolism. The lineage is placed within the Montiaceae, which in turn is sister to the rest of the Portulacineae (Caryophyllales). *Calandrinia* occupies two distinct biogeographic regions, one in the Americas (~14 species), and one in Australia (~74 species). Past analyses of the Montiaceae present conflicting hypotheses for the phylogenetic placement and monophyly of *Calandrinia*, and to date, there has been no molecular phylogenetic analysis of the Australian species.

METHODS: Using a targeted gene enrichment approach, we sequenced 297 loci from multiple gene families across the Montiaceae, including all named and 16 putative new species of Australian *Calandrinia*, and the enigmatic monotypic genus *Rumicastrum*.

KEY RESULTS: All datasets and analyses reject the monophyly of *Calandrinia*, with Australian and New World *Calandrinia* each comprising distinct and well supported clades, and *Rumicastrum* nested within Australian *Calandrinia*. We provide the first well-supported phylogeny for Australian *Calandrinia*, which includes all named species of Australian *Calandrinia* and several phrase-named taxa.

CONCLUSIONS: This study brings much needed clarity to relationships within Montiaceae, and confirms that New World and Australian *Calandrinia* do not form a clade. Australian *Calandrinia* is a longtime resident of the continent, having diverged from its sister lineage ~ 30 Ma, concurrent with separation of Australia from Antarctica. Most diversification occurred during the middle Miocene, with lowered speciation and/or

higher extinction rates coincident with the establishment of severe aridity by the late Miocene.

Keywords: *Australia; aridity; Australian Calandrinia; crassulacean acid metabolism (CAM); Montiaceae; Portulacineae; Rumicistrum; targeted bait enrichment*

INTRODUCTION

Over the last ~150 years there has been considerable debate regarding the monophyly, phylogenetic placement, and naming of *Calandrinia* Kunth, an enigmatic group of small, succulent herbs that display broad variation in habitat (Tahir and Ashton, 1989), vegetative morphology (Hershkovitz, 1993; Obbens, 2006; 2011; Ogburn and Edwards, 2015), and photosynthetic metabolism (Winter and Holtum, 2011; 2014; Holtum et al., 2016). *Calandrinia* is a member of the Montiaceae, which in turn is sister to the rest of the Portulacineae (Caryophyllales) (Nyffeler and Eggli, 2010; Ogburn and Edwards, 2015), a clade containing many arid-adapted, succulent plants (Arakaki et al., 2011; Ogburn and Edwards, 2013). *Calandrinia* as currently circumscribed is tri-continental, with *c.* 14 species in North and South America and the bulk of species (*c.* 74) endemic to Australia. In the New World, species occur in arid and alpine climates along the western cordillera from British Columbia to Chile (Arroyo et al., 1990; Elvebakk et al., 2015; Ogburn and Edwards, 2015), while in Australia, *Calandrinia* is widely distributed across the continent, typically inhabiting arid to semiarid environments, but extending into temperate regions in Tasmania and Victoria, and the monsoonal tropics of northern Australia (Obbens, 2006; Tahir and Carolin, 2011; West and Chinnock, 2013).

The monophyly of *Calandrinia* has been in question since Carolin (1987). Following a morphological cladistic analysis, Carolin (1987) concluded that Australian *Calandrinia* were not closely related to New World *Calandrinia*, and that the monotypic genus *Rumicastrum* Ulbr, originally described in the family Chenopodiaceae, was closely related to the Australian clade. He recommended that the Australian species be transferred to *Rumicastrum*, but failed to make new combinations. Hershkovitz (Hershkovitz, 1991; 1993; Hershkovitz and Zimmer, 1997; Hershkovitz, 1998) supported Carolin's separation of the Australian and New World *Calandrinia* but, believing that *Rumicastrum* was indeed in the Chenopodiaceae, erected the genus *Parakeelya* Hershk., a name taken from an aboriginal vernacular for the Australian species of *Calandrinia*. Although he provided new combinations for all species known at that time into *Parakeelya*, Hershkovitz's genus was not adopted in Australia (e.g., Obbens, 2006; 2011; 2017; West and Chinook 2013), as the non-monophyly of *Calandrinia* was not confidently resolved and the correct placement of *Rumicastrum* (which has nomenclatural priority) was regarded as uncertain (Obbens, 2006).

More recently, molecular phylogenetic analyses of the Montiaceae (Ogburn and Edwards, 2015) showed contrasting hypotheses for the placement and monophyly of *Calandrinia*. In a three-gene (*matK*, *ndhF*, and *phyC*) analysis, *Calandrinia* was supported as monophyletic, albeit with relatively low support (ML bootstrap = 65%, Bayesian posterior probability = 0.90). In a five-gene (*matK*, *ndhF*, *phyC*, ITS and *ycf3*) analysis, however, *Calandrinia* was paraphyletic, with Australian *Calandrinia* sister to a clade comprising Hectorelleae + Montieae. However, taxon sampling for Australian

Calandrinia was extremely poor in both analyses, with only one and five species represented in the three-gene and five-gene analyses, respectively.

The present study had multiple goals. First, we aimed to establish whether *Calandrinia* s.l. is monophyletic, and resolve the placement of Australian and New World species within the Montiaceae. Second, we aimed to completely sample all named species of Australian *Calandrinia* as representing a potentially important arid plant radiation that has received scarce attention. Australia is the driest vegetated continent, yet it harbors very few native succulent plant lineages, and only 0.6% of its vascular plants are known to use crassulacean acid metabolism (CAM) photosynthesis (Holtum et al., 2017). This is perplexing, as both succulence and CAM are common adaptations to aridity and have evolved numerous times across a diversity of plant lineages that inhabit similar deserts to those found in Australia (Winter and Smith, 1996; Ogburn and Edwards, 2010; Edwards and Ogburn, 2012). *Calandrinia* provides a key opportunity to explore this conundrum, as it is one of the most speciose succulent lineages in Australia and has evolved CAM photosynthesis (Winter and Holtum 2011; 2014). Third, we aimed to explore the biogeographic history and radiation of *Calandrinia* across Australia, assess how and when the lineage likely arrived there, and evaluate how aridification of the continent in the mid-late Miocene may have influenced patterns of lineage diversification. To address these issues, we used a targeted gene enrichment approach (Moore et al., 2018) to sequence hundreds of loci from multiple gene families of interest across *Calandrinia* and the Montiaceae. We provide the first well-supported phylogeny for Australian *Calandrinia*, which includes all named species of Australian *Calandrinia*, several phrase-named taxa, and the enigmatic monotypic genus *Rumicastrum*.

MATERIALS AND METHODS

Taxon Sampling—

We collected silica-dried leaf, stem, or flower material from across the Montiaceae (Appendix S1: see the Supplementary Data with this article), including multiple species from all recognized genera except *Hectorella* Hook.f., *Lyallia* Hook.f., and *Schreiteria* Carolin, three monotypic, geographically remote genera. Seven of the ~14 described New World *Calandrinia* species (nine individuals total) were included, including multiple accessions of the type species, *Calandrinia ciliata* (Ruiz & Pav.) DC. Material from a specimen of *Rumicastrum chamaecladum* (Diels) Ulbr. was obtained from the National Herbarium of Victoria (MEL). In addition to the Montiaceae, we included outgroup taxa from additional Portulacineae and Molluginaceae, using material and sequence data originally collected for Moore et al. (2018).

As a primary objective of the study was to resolve species relationships within Australian *Calandrinia* (Fig 1), we collected extensively in this region (Fig 2). We included 64 of the ~74 species currently recognized in Australia (see Australian Virtual Herbarium: http://avh.ala.org.au/occurrences/search?taxa=Calandrinia#tab_mapView): all 48 described species and 16 of 26 undescribed but recognized phrase-named taxa. We sequenced multiple individuals from across the range of species with large geographic distributions, for a total of 77 Australian *Calandrinia* individuals. Of these, 45 were newly collected in the field over the course of six field campaigns between August 2014 and August 2015, while 32 were sequenced from plant material obtained from herbarium specimens at the Western Australian Herbarium (PERTH) and the National Herbarium of

Victoria (MEL). In addition, we received field-collected plant material of several taxa from Bob Chinnock (State Herbarium of South Australia) and Attila Kapitany (Melbourne, Australia).

Molecular Methods—

Genomic DNA was extracted from dried material using a two-step DNA extraction protocol. First, we used the FastDNA Kit (MP Biomedicals, Santa Ana, CA) to extract DNA from 20–40 mg of plant tissue. We followed the manufacturer’s protocol, allowing samples to sit on the bench top for ~2 hours following homogenization with both CLS-VF and PPS buffers. After the initial extraction, samples were eluted twice in 75µl of distilled water, cleaned using a QIAquick PCR Cleanup Kit (Qiagen, INC., Valencia, CA), and eluted twice in 50µl of EB buffer.

To prepare DNA for library preparation and subsequent bait hybridization, ~500 ng of sample DNA was sonicated using a Covaris S220 (Covaris, Inc., Woburn, MA) at the Brown University Genomics Core Facility. The following parameters were implemented to generate a mean fragment length of 400 bp: peak power 140.0, duty factor 10.0, and cycles/burst 200 for 50 seconds. Library preparation was done using the NEBNext Ultra DNA Library Prep Kit or NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA) following the manufacturer’s protocol. Following library preparation, samples were pooled in groups of 8–9 and combined for hybridization. For bait hybridization, a custom MYbaits kit was used from MYcroarray. Both probe design and bait hybridization were carried out according to Moore et al., 2018. Quantification, combination, and sequencing were performed at the Brown

University Genomics Core Facility and the Oklahoma Medical Research Foundation Genomics Facility on Illumina HiSeq 2000 or 2500, to obtain 100-bp paired-end reads.

Sequence assembly—

Reads were initially filtered and assigned to individuals using their inline barcodes. Paired reads in which neither barcode matched exactly were discarded. The last five bases and the barcodes were then removed from the remaining reads. Finally, trimmed reads with more than one low-quality base were also discarded. Our sequence processing pipeline (Moore et al. 2018) consisted of three steps: step 1) extracted all relevant reads for each gene family and then assembled the reads into contigs, step 2) assigned contigs into putative paralogs, and step 3) identified gene duplications within gene families and extracted individual loci. These loci were then used to infer phylogeny with both concatenation and species tree approaches.

Briefly, step 1 classified reads into gene families by BLASTing (Altschul et al., 1990) them against a local library of sequences used to design the baits. The “captured” reads were then pooled among individuals belonging to the same family or subgroups within the Montiaceae. SPAdes version 3.1.0 (Bankevich et al., 2012) was used to assemble pooled reads into contigs and the contigs were then added to the bait sequence database. The trimmed reads were then BLASTed against this combined database (baits reads + preliminary SPAdes assembly) and separate assemblies were made for each individual for each gene family. This produced a fasta file for each gene family containing all contigs, labeled by individual. Finally, these files were BLASTed against

the bait sequence database to delineate exons. Introns were removed and only exons were used in subsequent analyses.

Step 2 classified the assembled contigs from step 1 according to paralog within gene families and combined contigs that represent different regions of the same paralog into a single gene sequence. Classification was executed by placing the contigs in a ‘backbone’ phylogeny of each gene family, built from the original bait sequences and additional sequences generated from prior analyses (Moore et. al 2018). The backbone tree was iteratively refined using the add fragments algorithm in MAFFT version 7.017 (Kato and Standley, 2013; using the localpair option because the fragments to be aligned covered only a portion of the backbone alignment) and the short read classification (-f v) algorithm in RAxML version 8.0.22 (Berger and Stamatakis, 2011; Stamatakis, 2014).

These steps produced gene-family trees containing the original sequences plus clusters of contigs (i.e. putative paralogs). For each of these clusters, the contigs from each individual were combined to form a consensus sequence. These consensus sequences were further analyzed and if they met certain criteria (i.e. > 75 bases long and 75% of the mean length of all consensus sequences from that cluster) they were added to the backbone alignments for each gene family.

Step 3 used a preliminary species tree to 1) identify the location of gene duplications within gene families, 2) extract paralogs arising from these gene duplications as separate loci, and 3) build individual gene trees and concatenated matrices for subsequent analyses. The initial species tree included all individuals in the analysis plus sequences for *Arabidopsis thaliana*, *Glycine max*, *Oryza sativa*, *Populus trichocarpa*, *Solanum tuberosum*, *Vitis vinifera* (derived from genomes for these model organisms)

and was constructed from three plastid loci (*matK*, *ndhF*, and *rbcL*) and the nuclear internal transcribed spacer (ITS) region. Before this species tree was made, gene trees of these three loci plus ITS were inspected by eye for individuals that were out of place and these sequences (likely pseudogenes) were removed from the tree. NOTUNG version 2.8.1.6 (Chen et al., 2000; Stolzer et al., 2012) was used to find gene duplications in the gene family trees based on the preliminary species tree, which delimited sets of orthologous sequences, i.e., phylogenetic loci. After a series of quality-control steps (see Moore et al. 2018), alignments were made using the local pair algorithm in MAFFT for the selected loci and sets of individuals. To reduce missing data, all sites with >90% missing data were removed from the alignments. Separate gene trees were then inferred for each locus using RAxML version 8.0.22 (GTRCAT, 1000 rapid bootstrap replicates); these gene trees were then used for coalescent-based species tree analyses.

After the initial run of step 3 of the pipeline (as described above), the plastid/ITS species tree was replaced with the reconstructed ASTRAL species tree (see below) and step 3 was rerun in its entirety. In addition to the updated species tree, we reran step 3 with 116 individuals divided into eight groupings (Appendix S1): Caryophyllales outgroups, Molluginaceae, Portulacineae except for Montiaceae, *Phemeranthus*, *Calandrinia* s.s., Australian *Calandrinia*, Montieae (*Lewisia* + *Montia* + *Claytonia* + *Lewisiopsis*) and the CCM clade (*Cistanthe* + *Calyptridium* + *Montiopsis*).

Species tree reconstruction—

We inferred phylogenetic relationships using both concatenation and coalescent-based approaches. Concatenation analyses were performed in RAxML version 8.0.22 (Stamatakis, 2006) using a GTRCAT substitution model, with 1000 rapid bootstrap

replicates. Coalescent-based species trees were estimated with ASTRAL II 4.10.2 (Mirarab et al., 2014), a package that analyzes individual gene trees (made with RAxML in step 3 of the pipeline). We used site-only multi-locus bootstrapping to generate bootstrap support values. Both ASTRAL and maximum likelihood concatenated RAxML species trees were reconstructed using four unique datasets, each with different amounts of missing data. The first two datasets varied in the number of loci present in a given percentage of individuals: all loci present in 85% or more of the individuals (minimum of 98 of 116 individuals, 74 loci; named i98) and all loci present in 35% or more of the individuals (minimum of 40 of 116 individuals, 165 loci; named i40). The other two datasets varied in loci present per major grouping (8 groups, see above): all loci present in two or more groups (297 loci; named g2) and all loci present in seven or more groups (116 loci; named g7). In all datasets, we removed sites (nucleotides) present in < 10% of individuals. Finally, to more fully understand the effects of missing data on phylogenetic relationships and branch lengths we removed all sites present in <50% of the individuals and reran concatenation and coalescent-based phylogenetic analyses using the reduced g2, g7, i40, and i98 datasets. To compare the internal and the terminal branch lengths of the g2, g7, i40, and i98 Australian *Calandrinia* concatenated phylogenies, when sites present in <10% of the individuals were removed and when sites present in <50% of the individuals were removed, we used a two-tailed, paired t-test as implemented in R.

Bayesian concordance analysis was performed using BUCKy version 1.4.4 (Larget et al., 2010) based on the posterior distribution of gene trees from analyses in MrBayes 3.2 (Ronquist et al., 2012). BUCKy estimates the genomic support as a concordance factor (CF) for each relationship found across analyses of all individual loci

(Ane et al., 2006; Baum, 2007), thereby detecting groups of genes supporting the same topology while accounting for uncertainty in gene tree estimates. We focused BUCKy analyses on the relationships among 12 major Montiaceae lineages: *Phemeranthus*, *Cistanthe* s.s. (sensu Ogburn and Edwards 2015), *Calyptridium*, *Montiopsis*, Montieae, *Calandrinia* s.s. and the six Australian *Calandrinia* sub-clades (see below). BUCKy does not allow for missing data, and most of our loci were not sampled from across all 116 terminals. However, as we were primarily interested in how these major lineages were related, we could include all loci that were sampled from at least one taxon in each of the lineages. To construct this dataset, we renamed taxa to their major lineages, and then pruned all but one random exemplar for each lineage from each sample of the posterior distribution of MrBayes trees (as in Moore et al., 2018). BUCKy analyses used four independent runs under default settings, with the option enabled to compute the posterior probability that pairs of loci support the same topology.

Ancestral character state reconstruction—

Ancestral character state reconstructions within Australian *Calandrinia* were based on the topology inferred using ASTRAL and the g2 dataset (297 loci), with RAxML branch lengths. This choice was guided by the result that coalescent-based species tree estimations were topologically more congruent across all datasets (g2, g7, i40, i98) than trees inferred using concatenated data, which showed considerable topological variation depending on the loci used (Appendix S2). Moreover, the concordance analyses also supported this topology (Fig 4). Four morphological traits were characterized for each Australian *Calandrinia* species and mapped onto the tree:

stigmatic lobe number, petal number, capsule type (dehiscent vs. indehiscent), and life-history (annual vs. perennial). We chose to focus our analyses on these four morphological traits as they were originally used by von Poellnitz (1934), Tahir & Ashton (1989), and Tahir & Carolin (1990) to differentiate Australian *Calandrinia* clades. Ancestral character state probabilities at all internal nodes were inferred based on maximum likelihood, using the ape package version 4.1 of the statistical software R (R Core team, 2017; Paradis et al., 2004). For each analysis, we modelled these characters as discrete traits, allowed transition rates between states to vary equally, differently, or symmetrically. The model with an AIC >2 above the alternative was chosen as the preferred model.

Dating, diversification rates, and biogeographic analyses—

Divergence times were estimated using a Bayesian approach as implemented in BEAST v2.4.5. (Bouckaert et al., 2014) using the g7 concatenated matrix (116 individuals, 116 loci, 158,436 bps) as it was the shortest. We fixed the tree topology using the g2 ASTRAL topology and estimated branch lengths under a log-normal relaxed clock (Drummond et al., 2006). Since there are no reliable and suitably old fossils of Portulacineae or Molluginaceae, we relied on a secondary calibration to date the divergence of Portulacineae; the stem age of Portulacineae was set to 53 MYA with a standard deviation of +/- 2.1 million years (Arakaki et al., 2011) and parameterized under a normal prior, because it was a secondary calibration point. Of studies that estimate divergence times across Portulacineae lineages (Ocampo and Columbus, 2010; Arakaki et al., 2011; Hernandez-Hernandez et al., 2014), we chose to use divergence times (and

standard deviations) from Arakaki et al. (2011) as their dating analyses are informed by a larger number of fossil calibration points. We used the GTR model of substitution with 4 gamma rate categories for the substitution rate. A birth-death speciation process was specified with default priors. The analysis was run twice for 10,000,000 generations, sampled every 1000 generations. Convergence and adequacy of burn-in were inspected visually in Tracer v1.5 (Rambaut and Drummond, 2009), and a maximum clade credibility tree was computed in TreeAnnotator.

We examined both the geographic history of the Montiaceae and Australian *Calandrinia* using a likelihood-based dispersal-extinction-cladogenesis (DEC) ancestral range reconstruction in Lagrange 20130526 (Ree and Smith, 2008). For the Montiaceae analysis we coded species as North American, South American, and/or Australian. For the Australian *Calandrinia* analysis we allowed each tip to occupy one or more of the following five geographic regions: south western Australia (SW), western Australia excluding the SW (WA), northern Australia (N), central Australia (C) and east/southeastern Australia (SE). We chose these five regions as they 1) represent floristically distinct and/or geographically separate zones in Australia (Crisp et al. 2004a), and 2) allowed us to reconstruct informative ancestral nodes. We separated the Eremean region of Australia into a western and central zone in efforts to further understand dispersal across the Nullabor Plain, a large limestone karst that is a geographic barrier to dispersal (Crisp and Cook 2007). We allowed for migration between all adjacent regions using an assigned rate of 1.0. For non-adjacent regions (i.e. N-SW, SE-SW and WA-SE), where migration between regions required passing through

another region, we used a “stepping-stone” model (Fine et al., 2014), assigning a rate of 0.5 to these transitions.

To identify potential shifts in diversification rate within the Australian *Calandrinia* clade, we ran MEDUSA (Alfaro et al., 2009) using a birth-death model, as implemented in the R package GEIGER 2.0.6 (Harmon et al., 2008). We removed all tips except for a species representative for each Australian *Calandrinia* taxa from 500 BEAST posterior trees and ran MEDUSA on the resulting 67-tip species trees. Since taxonomic sampling of Australian *Calandrinia* is essentially complete, with the exception of several undescribed taxa, we did not build a species richness table to account for missing taxa.

We also evaluated diversification in Australian *Calandrinia* using Bayesian analysis of macroevolutionary mixtures (BAMM) (Rabosky, 2014). We set the priors for the analyses using the setBAMMpriors function in BAMMtools (Rabosky et al., 2014). These priors included: expectedNumberofShifts = 1.0; lambdaInitPrior = 1.71; lambdaShiftPrior = 0.038; muInitPrior = 1.78. We ran BAMM on the maximum clade credibility Australian *Calandrinia* topology (with 67 tips) and allowed possible rate-change events to occur on all branches (minCladeSizeForShift = 1.0). BAMM was run for 10 million generations sampling every 1000. BAMMtools was used to assess convergence, effective sample size, and to identify shifts within the 95% credible shift configuration. In addition, we constructed a lineage through time (LTT) plot based on 500 posterior distribution trees from the BEAST analysis to visualize broad shifts in diversification rate through time.

RESULTS

Sequencing results and assembly—

After running step 1 and 2 of the pipeline, a total of 1018 loci were available for analysis. The number of loci per gene family (82 gene families in total) ranged between 1 and 39 (mean = 11.4; SD = 8.74). The number of loci recovered per individual ranged between 12 and 364 with the average being 151 loci (Appendix S1). Average locus length was 666.5 (SE +/- 16.89) nucleotide sites, ranging between 152 and 3468 base pairs.

Phylogenetic reconstruction—

The concatenated matrices of the four datasets, where sites present in <10% of the individuals were removed, had the following properties: dataset i98 = 164,559 nucleotides (bp), 91,811 parsimony-informative sites (is), 61.95% missing data (md); dataset i40 = 317,622 bp, 173,194 is, 67.99% md; dataset g7 = 158,436 bp, 105,972 is, 49.88% md; dataset g2 = 244,257 bp, 152,108 is, 58.05% md. To investigate the effect of missing data on phylogenetic analyses, particularly on the length of terminal branches, we then removed all sites present in <50% of the individuals. The concatenated matrices of the four resulting datasets had the following properties: dataset i98 = 65,050 bp, 43,388 is, 25.87% md; dataset i40 = 91,466 bp, 62,413 is, 28.96% md; dataset g7 = 81,068 bp, 54,897 is, 27.91% md; dataset g2 = 91,466 bp, 62,413 is, 28.96% md. Although these datasets had less missing data, they also had fewer phylogenetically informative sites, resulting in inconsistent and poorly resolved phylogenetic relationships along the Montiaceae backbone (Appendix S2). No difference in topology or bootstrap

support was found across the Australian *Calandrinia* backbone when reconstructing the phylogenies with the reduced matrices (Appendix S2). Furthermore, we found no significant difference in Australian *Calandrinia* branch lengths (both internal and terminal branch lengths) when reconstructing the phylogenies with the reduced matrices versus the original matrices (<10%) except for in the i40 ($t=2.40$, $df=66$, $p=0.02$) and i98 ($t=3.23$, $df=66$, $p=0.002$) datasets, where the terminal branches were significantly shorter with less missing data (Appendix S3). Thus, we chose to focus the remainder of our results and discussion on the phylogenetic trees reconstructed with more missing data (i.e. where sites present in <10% of the individuals were removed) and informative sites.

Paraphyly of Calandrinia— There is strong and consistent support for two distinct monophyletic *Calandrinia* clades: New World *Calandrinia* (hereafter referred to as *Calandrinia* s.s.) and Australian *Calandrinia*. These clades are strongly supported and distinct across all datasets and methods of analysis (Fig 3 and Appendices S4-S10). The type species for *Calandrinia*, *C. ciliata* (Ruiz & Pav.) DC, falls within the New World *Calandrinia* clade. In addition to resolving the monophyly of Australian *Calandrinia*, we confidently place *Rumicastrum chamaecladum* (Diels) Ulbr. within Australian *Calandrinia*. This strongly supported result requires new combinations for all Australian *Calandrinia* species, either into *Rumicastrum* or, if a proposal to conserve the name *Parakeelya* is successful, into *Parakeelya* (Thiele et al. in prep).

Relations among major Montiaceae lineages— With the exception of *Calandrinia*, all recognized genera within Montiaceae are well-supported (100% bootstrap support) in both concatenated and ASTRAL analyses (Fig 3 and Appendices S4-S10). It should be noted, however, that section *Philippiamra* Kuntz within *Cistanthe*

was not sampled in this study, and in previous studies it has been shown to be distinct from other *Cistanthe* with strong support (Ogburn and Edwards, 2013; 2015).

Phemeranthus is resolved as sister to the rest of Montiaceae (100% bootstrap support), and *Lewisiopsis*, *Lewisia*, *Claytonia*, and *Montia* form a well-supported clade (Montieae), with *Lewisiopsis* consistently sister to *Lewisia* + *Montia* + *Claytonia* across analyses and datasets (100% support in all cases). *Cistanthe s.s.*, *Calyptridium*, and *Montiopsis* also form a well-supported clade (hereafter called ‘CCM’). However, the inferred relationships among these major groups vary notably across analyses and datasets.

Overall, ASTRAL produced consistent, well-supported topologies across datasets, while topologies from concatenation analyses differed (Fig 3 and Appendices S2, S4, S6, S8). In ASTRAL analyses, the CCM clade is recovered as sister (100%) to Montieae + *Calandrinia s.s.*, and Australian *Calandrinia*. The only conflict observed in the backbone tree using ASTRAL are relationships within the CCM clade: *Calyptridium* is resolved as sister (100%) to *Cistanthe s.s.* + *Montiopsis* using the g7 and i98 datasets, but *Cistanthe s.s.* is sister (100%) to *Montiopsis* + *Calyptridium* with the other two datasets (g2 and i40). All datasets show low support for *Cistanthe s.s.* + *Montiopsis* (g7, 70%; i98, 62%) and *Montiopsis* + *Calyptridium* (g2, 56%; i40, 60%). Across the concatenated trees estimated with RAxML there is also conflict among these relationships. Specifically, *Calyptridium* is sister to *Cistanthe s.s.* + *Montiopsis* in the i98 and g7 trees, while *Cistanthe s.s.* is sister to *Montiopsis* + *Calyptridium* with the remaining two datasets. The bootstrap support for these relationships is > 99% across concatenated trees in all analyses except for g7, where the *Cistanthe s.s.* + *Montiopsis* node received only 86% bootstrap support (Fig 3). Results from the Bayesian concordance analysis illuminated

these differences: while the primary concordance tree recovered *Cistanthe s.s.* as sister to *Montiopsis* + *Calyptridium*, the concordance factor for this relationship (CF = 0.221) was only marginally higher than the one for the competing relationship *Cistanthe s.s.* + *Montiopsis* (CF = 0.211), indicating genome-wide conflict among loci at this node (Fig 4 and Appendix S11).

In contrast to the ASTRAL analyses, topologies based on concatenation and RAxML analysis differed across datasets (Fig 3 and Appendices S2, S5, S7, S9, S10). Three nodes that exhibit considerable conflict are highlighted here. First, *Calandrinia s.s.* is recovered as sister to the CCM clade (>96% bootstrap support) across all datasets except g7 (Appendix S5), with *Calandrinia s.s.* sister to Montieae, albeit with low support (BS=82%). Secondly, the placement of the CCM clade + *Calandrinia s.s.* varied across matrices: in the i40 and i98 trees, the CCM clade + *Calandrinia s.s.* is resolved as sister to Montieae + Australian *Calandrinia* with 100% bootstrap support, while in the g2 tree, CCM + *Calandrinia s.s.* is sister to Australian *Calandrinia* (BS=66%). Third, the Montieae clade is placed as sister to Australian *Calandrinia* in the i40 and i98 trees (BS=96%), while species tree reconstructions with the g2 dataset place Montieae as sister to CCM + *Calandrinia s.s.* and Australian *Calandrinia*. The concatenation analysis based on dataset g7 resolved a very similar topology as recovered with the coalescent based reconstructions, placing Montieae + *Calandrinia s.s.* as sister to Australian *Calandrinia* (Fig 3).

Bayesian Concordance Analysis resulted in a primary concordance tree that was congruent with the topology recovered by ASTRAL and g7 Concatenation (Fig 4). Specifically, it supported *Calandrinia s.s.* as sister to Montieae, with these jointly being

sister to Australian *Calandrinia*. Although CF values are typically quite low, indicating that significant portions of the genome support relationships that deviate from the dominant signal (Appendix S11), there are no alternative resolutions that are equally well supported as those in the primary concordance tree. The only exception to this pattern was within the CCM clade, highlighted above.

Australian Calandrinia— Species tree reconstructions using concatenation, coalescence, and Bayesian inference recover six strongly-supported clades within Australian *Calandrinia* (Fig 5). Three of these correspond to the previously described sections *Tuberosae*, *Pseudodanthoidiae* and *Basales*, while three (clades 3, 4, and 5) have not been previously recognized taxonomically. There is both strong and uniform support for the monophyly and placement of these six clades across all datasets (Appendices S4-S10). Furthermore, the topologies within *Tuberosae* and clades 3, 4 and 5 are congruent and well supported across all datasets and analyses. Within the *Pseudodanthoidiae* and *Basales* clades, however, there is considerably more uncertainty in species relationships (Fig 5); this is true regardless of dataset and/or method of analysis. Below we provide a brief description for each of the six clades.

Clade 1 aligns well with section *Tuberosae* as circumscribed by von Poellnitz (1934), Tahir & Ashton (1989) and Tahir & Carolin (1990). These species are perennials with a tuberous rootstock, five-petaled, three-carpellate flowers with numerous stamens, and dehiscent capsules. Most species occur in semi-arid to arid regions of Western Australia.

Clade 2 includes most species placed in section *Pseudodanthoideae* by von Poellnitz (1934), Tahir and Ashton (1989) and Tahir and Carolin (1990). These are semi-

erect to erect (rarely prostrate to decumbent) annuals (rarely perennials), with five-petaled, three-carpellate flowers that are usually medium to large, mostly with numerous stamens, and have dehiscent, many-seeded capsules. Most occur in temperate to semi-arid parts of southern and central Australia. Though the clade is well supported, many species relationships within it are poorly resolved.

Clade 3 includes *C. disperma*, *C. brevipedata*, *C. corrigioloides*, *C. liniflora*, and *Rumic astrum chamaecladum*, as well as three currently undescribed taxa (*C. sp.* Truncate capsules, *C. sp.* The Pink Hills, *C. sp.* Bungalbin). The named species, with the exception of *R. chamaecladum*, were placed by von Poellnitz 1934 and Tahir and Carolin 1989 in section *Pseudodianthoideae*. Most are small, prostrate to decumbent annuals often with many, relatively small, three-carpellate flowers with five or fewer petals and stamens. Most have several-seeded, dehiscent capsules, except *C. sp.* The Pink Hills, *R. chamaecladum*, and *C. disperma*, which have one- or two-seeded, indehiscent capsules. All are endemic to southwestern Australia except for *C. disperma*, which grows through western and central Australia.

Clade 4 comprises *C. pickeringii* and *C. pumila*. Both are prostrate to decumbent, five-petaled, pink-flowered, 3-carpellate annuals that range widely across temperate arid Australia into eastern Queensland and New South Wales. No obvious morphological synapomorphies unite this clade.

Clade 5 is a morphologically diverse group, also with no apparent synapomorphy that differentiates it from the other five clades. Most species are prostrate to decumbent, diminutive annuals with small to very small, three-carpellate, five-petaled flowers. All have several- to many-seeded dehiscent capsules except *C. sp.* Piawaning, which has

indehiscent capsules usually containing two seeds. Most species are endemic in southern Western Australia except *C. granulifera*, which is widespread across southern Australia including Tasmania.

Clade 6 aligns well with section *Basales*, as circumscribed by von Poellnitz (1934) and Tahir and Ashton (1989). All are four-carpellate except *C. pentavalvis* and *C. strophiolata*, which are five- and six-carpellate respectively. All species are decumbent to erect annuals, mostly with medium to large flowers with six or more pink, purple or white petals, and usually numerous stamens. Capsules are generally many-seeded (*C. oblonga* has few, large seeds). Capsules in *C. porifera* open by a terminal pore, while all other species are valvate. Many species in this clade occur in northern, subtropical (monsoonal) Australia, with others occurring in the arid interior.

Ancestral character state reconstructions—

Character state reconstructions indicate that ancestral Australian *Calandrinia* was likely an annual plant with three-stigmatic lobes, five petals, and dehiscent capsules (Fig 5). When ML analyses were conducted for life-history and fruit type reconstructions, the simplest model with equal transitions rates was favored, though AIC scores between models were typically very similar, and our actual reconstructions inferred shifts in only one direction. Life history state reconstructions under an equal rates model (ER AIC = 33.57; ARD AIC= 35.03) recovered three shifts from annual to perennial; one shift at the base of section *Tuberosae*, and two in *Pseudodanthoidieae* (Appendix S12). Evolution of indehiscent fruits, under the favored ER model (ER AIC = 31.71; ARD AIC 33.73), likely occurred four separate times with three of these switches arising in clade 3 and one

in clade 4 (Appendix S13). There is considerable intraspecific variation in petal number when a species has six or more petals; for this reason, we chose to bin petal number as five and six plus. Under the favored ER model (ER=37.56; ARD=39.51), the derived six plus petal state evolved twice, once in clade 4 and once at the base of *Basales*; there are also several reversals back to five petals from 6 plus petals (Appendix S14). Based on AIC values we chose the more complex, symmetrical transition rates model to reconstruct valve number. We found one shift to a four-valve state from the ancestral 3-valved state at the base of section *Basales*, and two shifts to four plus valves within this clade (ER AIC=45.32; SYM AIC=43.21; ARD AIC=52) (Appendix S15).

Divergence time estimation, rate shifts, and biogeographic analyses—

The estimated crown age of Australian *Calandrinia* is late Oligocene (27.03 +/- 2.95 my), with most diversification events occurring during the Miocene (Fig 6). MEDUSA and BAMM analyses indicate no clade-specific shifts in diversification rate within Australian *Calandrinia*, although BAMM shows a lineage-wide general slowdown or decrease in speciation rate with time (Appendix S17). The inferred slowdown is not surprising given the near absence of any topological divergence events from the Pliocene onward (Fig 6). The LTT plot (Appendix S16) shows an abrupt and dramatic decline in lineage accumulation at 10 Ma.

Biogeographic reconstruction of the Montiaceae using the time-calibrated phylogeny (Fig. 6, Supp. Fig. S17) indicate that Australian *Calandrinia* most likely entered Australia via South America/Antarctica close to the estimated Australia/Antarctica split at ~33 Ma (Crisp et al. 2004). LaGrange recovered a South

America/Australia separation (100% probability) at the node separating Australian *Calandrinia* from Montieae + *Calandrinia s.s.*, indicating an ancestral South American distribution. Australian *Calandrinia* likely originated in Western Australia (78% probability) and moved east and south into central and southwestern Australia, respectively, multiple times. There were only two dispersals into northern Australia, and both in section *Basales*; one roughly ~18 Ma, and one within the last few million years (*C. arenicola*).

DISCUSSION

Phylogenetic relationships within the Montiaceae—

The most recent phylogenetic evaluation of the Montiaceae (Ogburn and Edwards, 2015), using three-marker (*matK*, *ndhF*, and *phyC*) and five-marker (including ITS and *ycf3*) combined analyses, is congruent with many of the relationships presented here: *Phemeranthus* is sister to the rest of Montiaceae, with *Cistanthe*, *Calyptridium*, and *Montiopsis* forming a well-supported clade (the CCM clade) and *Claytonia*, *Montia*, *Lewisia* and *Lewisiopsis* (the Montieae) also well-supported as monophyletic.

Although the paraphyly of *Calandrinia s.l.* has long been suspected (Carolin, 1987; Hershkovitz, 1998) only here do we finally confirm that *Calandrinia s.s.* and Australian *Calandrinia* are distinct lineages and do not form a clade. Concatenation and coalescent-based methods support alternative scenarios for relationships within the major lineages of Montiaceae, however concordance analyses, despite only being run on a smaller subset of our loci, show that the top 10 splits (i.e. bifurcations) converge onto the same underlying phylogenetic relationships, those recovered in the g7 RAxML and all

ASTRAL analyses (Fig 3a, 3b, 3e and Appendices S2, S6, S8). Discordance in the placement of *Calandrinia s.s.* and the sister lineage to Australian *Calandrinia* do not appear to result from strongly supported conflict at the nodes, as seen in the relationships among *Cistanthe*, *Calyptridium*, and *Montiopsis*, but rather from a low signal-to-noise ratio in the gene trees. It is likely that individual gene trees are resolved with low support because of missing data across many individuals. Similarly, we suspect that the concatenation analyses were only congruent with the ASTRAL backbone topology when analyzing the g7 matrix because the loci represented in this dataset (see Fig 3c compared to Fig 3f) have the least amount of missing data across the Montiaceae. Concatenation analyses, in general, appeared more sensitive to missing data than ASTRAL analyses. Furthermore, although genomic support for many clades is quite low (typically CF of 0.2–0.3), the entire primary concordance tree (Fig 4) assembled from the top ten best-supported splits is perfectly congruent with the ASTRAL topology.

Phylogenetic relationships within Australian Calandrinia—

Phylogenetic analyses using concatenation, coalescence, and Bayesian inference recover six strongly-supported clades within Australian *Calandrinia* (Fig 5), with *Rumicastrum chamaecladum* nested within clade 3 with strong support. Although the backbone topology of the lineage is consistent across datasets and all three methods of inference, there remains considerable topological uncertainty within certain clades, specifically within *Basales* and *Pseudodianthoideae*. The recovery of low and inconsistent support for many of these relationships may indicate widespread gene flow and introgression during past speciation events. Many of these poorly supported species

relationships represent taxa that overlap geographically and share similar, or the same, habitat and niche space.

In addition to hybridization and introgression, low support values and topological uncertainty at the tips may also reflect several complexes of poorly-defined species within these two clades, centered on the (mostly) geographically widespread *C.*

quadrivalvis, *C. eremaea*, *C. baccata*, *C. polyandra*, and *C. remota*. For example, the *C. eremaea* complex, which is not monophyletic in our analyses, occurs across the southern half of Australia, from Mediterranean climate areas of southwestern Australia and Victoria, to more arid regions of western and central Australia. *C. eremaea* s.s. is from central southern Queensland and is a small, decumbent to erect annual herb with three stigmas, five petals, eight stamens and reniform to sub-reniform, colliculate to finely papillate seeds. Most variants deviate notably from the original species description, including individuals with a perennial rather than annual growth form and others with smooth to dull, non-colliculate seeds.

Finally, low support for some relationships may be a product of poor sequence capture with the bait hybridization approach. For example, it is likely that placement of *C. oblonga* is compromised by the fact that we recovered the fewest loci for this taxon (12 loci).

Dating and the biogeographic history of Australian Calandrinia—

While we are fully aware of the shortcomings (see Schenk 2016) of utilizing secondary node calibrations to estimate divergence times, there are no suitable fossils available for our analyses. We feel that an attempt to date this Australian radiation with

secondary calibration is preferable to none at all and recommend interpreting our inferred ages with a reasonable level of skepticism, as we do.

The evolution and radiation of Australian *Calandrinia* has received little attention, as species relationships within the lineage have been largely unresolved. Tahir and Carolin (1990) suggested that Australian *Calandrinia* were Gondwanan in origin (i.e. colonizing Australia before it separated from Antarctica) rather than arriving via long distance dispersal from South America. They argued that dispersal (e.g., by birds, ocean currents) would have left a “trail” of related species on islands between South America and Australia, including New Zealand, which provide suitable climate and habitat.

Using modern molecular and phylogenetic analyses (Fig 6 and Appendix S18), we provide some support for Tahir and Carolin’s biogeographic hypothesis for the radiation of *Calandrinia*. In our BEAST analysis, the node separating the Australian *Calandrinia* clade from its sister lineage (the Montieae clade plus *Calandrinia s.s.*) is estimated to be $\sim 30 \pm 2.34$ Ma, which aligns well with the final separation of Australia from Antarctica at ~ 33 Mya (Scher and Martin, 2006). The Montieae + *Calandrinia s.s.* clade is primarily distributed in North and South America, with *Calandrinia s.s.* extending into southern Chile, Argentina, and the Falkland Islands (*C. feltonii*) and two *Montia* species (*M. fontana* and *M. australasica*) into Australia. Not included in our phylogenetic analyses are the monotypic *Hectorella* and *Lyallia* lineages. *Hectorella* is recovered as sister to the Montieae in Ogburn and Edwards (2015) and *Lyallia* is placed as sister to *Hectorella*, albeit with low support, by Arakaki *et al.* (2011). *Lyallia* is endemic to the sub-Antarctic Kerguelen Islands, while *Hectorella* is endemic to the South Island of New Zealand. If these were sister to the Australian *Calandrinia* clade, they

could comprise the “trail” of taxa, anticipated by Tahir and Carolin (1990), indicating east to west long-distance dispersal. However, given the dates in our analysis, and their uncertainty, we find it difficult to strongly argue for either hypothesis. The dispersal and radiation of plants between South America and Australia via Antarctica is well known and documented by a rich record of both macro and micro fossils (Eklund et al., 2004; Sanmartín and Ronquist, 2004; Wagstaff and Hennion, 2007; Crisp and Cook, 2013), and this seems to also be a reasonable scenario for the Montiaceae. The phylogenetic placement of *Lyallia* and *Hectorella* is critical to these discussions and including these taxa in subsequent work is a high priority.

Regardless of whether Australian *Calandrinia* formed due to vicariance or long distance dispersal, the lineage is a longtime resident of Australia. Our analyses indicate that the stem lineage of Australian *Calandrinia* originated in mesic, temperate climates of Western Australia during the late Oligocene. The crown lineage then diversified over the next 15 million years, during the early to mid-Miocene, when the climate in Australia was generally wet, warm, and stable (Martin, 2006; Byrne et al., 2011). Australian *Calandrinia* radiated across the continent, moving mostly eastwards from Western Australia into central Australia and southwards to southwestern Australia. The lineage – specifically section *Basales* – reached northern Australia ~18 Ma, during which time the Australian plate was colliding with the Asian plate (Hall 2012). Once in northern Australia, section *Basales* radiated across the tropics (Fig 6), the only Australian *Calandrinia* clade to do so.

The dramatic paleoclimatic history of Australia provides a framework in which to consider Australian *Calandrinia* diversification. Although Australia became gradually

drier as it drifted northward from Antarctica, due both to the establishment of the Antarctic Circumpolar Current (Scher and Martin, 2006; Lyle et al., 2007) and the precipitous decline in global CO₂ levels (DeConto and Pollard, 2003; Edwards and Ogburn, 2012), aridity did not become widely established in Australia until the mid to late Miocene (~13-6 Mya) (Crisp et al., 2004; Martin, 2006; Byrne et al., 2008). During this time Australia experienced sea-level subsidence, the end of regular flows in paleo-drainage systems, the drying-up of seasonal lakes, and general climatic instability (Hill et al., 1999; Martin, 2006; Byrne et al., 2008). The earlier part of this time period (in general, ~ 20 to 10 Ma) corresponds with diversification of Australian *Calandrinia* in both western and central Australia (Fig 6). Although MEDUSA and BAMM analyses recover no clade-specific shifts in diversification rate during this time, LTT plots demonstrate a substantial decline in diversification rate from ~10 Ma to present day (Appendix S16). Even though the beginning of the Pliocene saw slightly wetter and warmer temperatures, the mid-Pliocene into the Pleistocene marks the beginning of severe aridity in Australia (Crisp and Cook, 2007). Stony deserts began to form across western and central Australia, and the tropics of northern Australia became drier (Fujioka et al., 2005; Byrne et al., 2008; Fujioka et al., 2009). Major oscillations between glacial and interglacial climates that characterized the Pleistocene enhanced the expansion of aridity in Australia (Byrne et al., 2008; Fujioka et al., 2009). BAMM analyses also detected an overall decrease in speciation with time (Appendix S17), which is not surprising given the virtual absence of divergence events in our phylogeny from 6 million years onwards (Fig 6).

A decline in diversification rates has also been reported for other Australian plant lineages that had originally diversified and radiated during the mid-Miocene, e.g. *Tetralochea* (Crayn and Rossetto, 2006), Goodeniaceae (Jabaily et al., 2014), Proteaceae (Sauquet et al., 2009; Mast et al., 2012) and Myrtaceae (Thornhill et al., 2015). Many of these lineages likely responded to extreme desertification of the continent during glacial maxima by retreating to refugia. This seems to be a period of either extinction or species maintenance for many lineages (Crisp et al., 2004; Byrne et al. 2008). While these “Gondwanan” lineages were likely constrained to refugia, especially during glacial maxima, arid-adapted migrant lineages – groups that had newly arrived during the mid-late Miocene and Pliocene epochs, e.g. Chenopodiaceae (Shepherd et al., 2004; Kadereit et al., 2010), Poaceae (Gillespie et al., 2009), Triodiinae (Toon et al., 2015), *Lepidium* (Mummenhoff et al., 2001; Crisp, Cook and Steane, 2004a), *Ptilotus* (Hammer et al., 2015), and certain Asteraceae, including Gnaphalieae (Bergh and Linder, 2009) and *Olearia* and *Celmisia* (Wagstaff et al., 2010), radiated across the continent. These newer arrivals, many of which are short-lived, fast-growing, and often C₄ annuals, may have been better adapted to expanding arid conditions. Expansion of these immigrant lineages across Australia may have fundamentally changed community structure, ecosystem properties, and perhaps the competitive ability of established lineages.

CONCLUSION

The apparent lack of diversification in Australian *Calandrinia* from the late Miocene onwards is paradoxical: why did this succulent, facultative CAM lineage (Winter and Holtum, 2011; 2014; Holtum et al., 2016), which seems well-adapted to arid

landscapes, not diversify as arid environments became more widespread? Did the extreme climate during the Pliocene-Pleistocene simply exceed the climatic range in which the species could thrive? Or was Australian *Calandrinia* outcompeted by more recent arrivals, such as C₄ grasses and chenopods? Many biogeographic studies of arid ecosystems are hampered by a lack of fossil evidence, and the main sources of biogeographic hypotheses for the assembly of the arid Australian flora have been contemporary species distributions. Lacking physical evidence, such as fossils, the continued generation of robust, well-supported, taxon-rich, time-calibrated phylogenies (e.g., Owen et al., 2017; Toon et al., 2017) are imperative to our understanding of the Australian arid zone, as comparisons between multiple, thoroughly sampled model lineages can reveal general patterns in the timing of species radiations and extinctions.

DATA ACCESSIBILITY

Voucher information is available in Appendix S1. Raw reads are deposited in the NCBI Short Read Archive (PRJNA417446). All trees and final alignments are available in GitHub [https://github.com/lillypine/Calandrinia_phylogeny].

ACKNOWLEDGEMENTS

The authors thank the Edwards lab members, both past and present, for comments on the manuscript, Elissa Martin for figure and illustration assistance, and two anonymous reviewers for suggestions on an earlier draft of the manuscript. Special thanks to the Western Australia Herbarium, The Melbourne Botanic Gardens, Northern

Tropical Herbarium, Leesa Baker, and Alex Cheeseman for field-collection assistance.
Thanks to Bob Chinnock and Attila Kapitany for collecting and sending plant material.
Heartfelt gratitude to Rosemary Dunn, Wendy Powles, Steve Powles, and Hazel Jensen.
This research was funded by the National Science Foundation [DEB-1252901 to E.J.E.],
the Australian Research Council (DP16010098), L.P.H. was supported in part by NSF
IGERT grant DGE-11 0966060, a NSF Doctorate Dissertation Improvement Grant
(GR5260110), and a NSF EAPSI Australia award (1514963).

LITERATURE CITED

- AVH. 2017. Australia's Virtual Herbarium, Council of Heads of Australasian Herbaria, <http://avh.chah.org.au>, accessed October 10, 2015.
- ALFARO, M.E., F. SANTINI, C. BROCK, H. ALAMILLO, A. DORNBURG, D.L. RABOSKY, G. CARNEVALE, AND L.J. HARMON. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences of the United States of America* 106: 13410–13414.
- ALTSCHUL, S.F., W. GISH, W. MILLER, E.W. MYERS, AND D.J. LIPMAN. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410.
- ANE, C., B. LARGET, D.A. BAUM, S.D. SMITH, AND A. ROKAS. 2006. Bayesian estimation of concordance among gene trees. *Molecular Biology and Evolution* 24: 412–426.
- ARAKAKI, M., P.A. CHRISTIN, R. NYFFELER, A. LENDEL, U. EGGLI, R.M. OGBURN, E. SPRIGGS, ET AL. 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. *Proceedings of the National Academy of Sciences* 108: 8379–8384.
- ARROYO, M.K., E. MEDINA, AND H. ZIEGLER. 1990. Distribution and $\delta^{13}\text{C}$ values of Portulacaceae species of the high Andes in northern Chile. *Botanica Acta* 103: 291–295.
- 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: 455–

477.

- BAUM, D.A. 2007. Concordance trees, concordance factors, and the exploration of reticulate Genealogy. *Taxon* 56: 417–426.
- BERGER, S.A., AND A. STAMATAKIS. 2011. Aligning short reads to reference alignments and trees. *Bioinformatics* 27: 2068–2075.
- BERGH, N.G., AND H. PETER LINDER. 2009. Cape diversification and repeated out-of-southern-Africa dispersal in paper daisies (Asteraceae–Gnaphalieae). *Molecular Phylogenetics and Evolution* 51: 5–18.
- BOUCKAERT, R., J. HELED, D. KÜHNERT, T. VAUGHAN, C.-H. WU, D. XIE, M.A. SUCHARD, ET AL. 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* 10: e1003537–6.
- BYRNE, M., D.A. STEANE, L. JOSEPH, D.K. YEATES, G.J. JORDAN, D. CRAYN, K. APLIN, ET AL. 2011. Decline of a biome: evolution, contraction, fragmentation, extinction and invasion of the Australian mesic zone biota. *Journal of Biogeography* 38: 1635–1656.
- BYRNE, M., D.K. YEATES, L. JOSEPH, M. KEARNEY, J. BOWLER, M.A.J. WILLIAMS, S. COOPER, C. DONNELLAN, J.S. KEOGH, R. LEYS, J. MELVILLE, D.J. MURPHY, N. PORCH, K.H. WYRWOLL 2008. Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Molecular Ecology* 17: 4398–4417.
- CAROLIN, R. 1987. A review of the family Portulacaceae. *Australian Journal of Botany*

35: 383–412.

CHEN, K., D. DURAND, AND M. FARACH-COLTON. 2000. NOTUNG: a program for dating gene duplications and optimizing gene family trees. *Journal of Computational Biology* 7: 429–447.

CRAYN, D.M., M. ROSSETTO, AND D.J. MAYNARD. 2006. Molecular phylogeny and dating reveals an Oligo-Miocene radiation of dry-adapted shrubs (former Tremandraceae) from rainforest tree progenitors (Elaeocarpaceae) in Australia. *American Journal of Botany* 93: 1328–1342.

CRISP, M., L. COOK, AND D. STEANE. 2004. Radiation of the Australian flora: what can comparisons of molecular phylogenies across multiple taxa tell us about the evolution of diversity in present-day communities? *Philosophical Transactions of the Royal Society B* 359: 1551–1571.

CRISP, M.D., AND L.G. COOK. 2007. A congruent molecular signature of vicariance across multiple plant lineages. *Molecular Phylogenetics and Evolution* 43: 1106–1117.

DECONTO, R.M., AND D. POLLARD. 2003. Rapid Cenozoic glaciation of Antarctica induced by declining atmospheric CO₂. *Nature* 421: 245–249.

DRUMMOND, A.J., S.Y.W. HO, M.J. PHILLIPS, AND A. RAMBAUT. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4: e88–12.

EDWARDS, E.J., AND R.M. OGBURN. 2012. Angiosperm responses to a low-CO₂ world:

- CAM and C₄ photosynthesis as parallel evolutionary trajectories. *International Journal of Plant Sciences* 173: 724–733.
- EKLUND H., AND J.E. CANTRILL. 2004. Late Cretaceous plant mesofossils from Table Nunatak, Antarctica. *Cretaceous Research* 25: 211–28.
- ELVEBAKK, A., A. FLORES, AND J.M. WATSON. 2015. Revisions in the South American *Calandrinia caespitosa* complex (Montiaceae). *Phytotaxa* 203: 001–023.
- FINE, P.V.A., F. ZAPATA, AND D.C. DALY. 2014. Investigating processes of neotropical rain forest tree diversification by examining the evolution and historical biogeography of the Protieae (Burseraceae). *Evolution* 68: 1988–2004.
- FUJIOKA, T., J. CHAPPELL, L.K. FIFIELD, AND E.J. RHODES. 2009. Australian desert dune fields initiated with Pliocene-Pleistocene global climatic shift. *Geology* 37: 51–54.
- FUJIOKA, T., J. CHAPPELL, M. HONDA, I. YATSEVICH, K. FIFIELD, AND D. FABEL. 2005. Global cooling initiated stony deserts in central Australia 2–4 Ma, dated by cosmogenic ²¹Ne-¹⁰Be. *Geology* 33: 993–996.
- GILLESPIE, L.J., R.J. SORENG, AND S.W.L. JACOBS. 2009. Phylogenetic relationships of Australian *Poa* (Poaceae: Poinae), including molecular evidence for two new genera, *Saxipoa* and *Sylvipoa*. *Australian Systematic Botany* 22: 413–25.
- HALL, R. 2012. Sundaland and Wallacea: geology, plate tectonics and palaeogeography. In 'Biotic Evolution and Environmental Change in Southeast Asia. (Eds D Gower, K Johnson, J Richardson, L R,ber and S Williams). Cambridge University Press:

Cambridge

- HAMMER, T., R. DAVIS, AND K. THIELE. 2015. A molecular framework phylogeny for *Ptilotus* (Amaranthaceae): evidence for the rapid diversification of an arid Australian genus. *Taxon* 64: 272–285.
- HARMON, L.J., J.T. WEIR, C.D. BROCK, R.E. GLOR, AND W. CHALLENGER. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics* 24: 129–131.
- HERNÁNDEZ-HERNÁNDEZ, T., J.W. BROWN, B.O. SCHLUMPBERGER, L.E. EGUIARTE, AND S. MAGALLÓN. 2014. Beyond aridification: multiple explanations for the elevated diversification of cacti in the New World Succulent Biome. *New Phytologist* 202: 1382–1397
- HERSHKOVITZ, M.A. 1998. *Parakeelya*: a new genus segregated from *Calandrinia* (Portulacaceae). *Phytologia* 84: 98–106.
- HERSHKOVITZ, M.A. 1991. Phylogenetic assessment and revised circumscription of *Cistanthe* Spach (Portulacaceae). *Annals of the Missouri Botanical Garden* 78: 1009–1021.
- HERSHKOVITZ, M.A. 1993. Revised circumscriptions and subgeneric taxonomies of *Calandrinia* and *Montiopsis* (Portulacaceae) with notes on phylogeny of the portulacaceous alliance. *Annals of the Missouri Botanical Garden* 80: 333–365.
- HERSHKOVITZ, M.A., AND E.A. ZIMMER. 1997. On the evolutionary origins of the cacti. *Taxon* 46: 217–232.

- HILL R.S., E.M TRUSTWELL, S. MCLOUGHLIN AND M.E. DETTMAN. 1999. The evolution of the Australian flora: fossil evidence. In: Orchard A.E. editor. *Flora of Australia*. Melbourne, Australia: CSIRO Publishing: 251–320.
- HOLTUM, J.A., L.P. HANCOCK, E.J. EDWARDS, M.D. CRISP, D.M. CRAYN, R. SAGE, AND K. WINTER. 2016. Australia lacks stem succulents but is it depauperate in plants with crassulacean acid metabolism (CAM)? *Current Opinion in Plant Biology* 31: 109–117.
- HOLTUM, J.A.M., L. HANCOCK, E.J. EDWARDS, AND K. WINTER. 2017. Facultative CAM photosynthesis (crassulacean acid metabolism) in four species of *Calandrinia*, ephemeral succulents of arid Australia. *Photosynthesis Research* 134: 17–25.
- JABAILY, R.S., K.A. SHEPHERD, A.G. GARDNER, M.H.G. GUSTAFSSON, D.G. HOWARTH, AND T.J. MOTLEY. 2014. Historical biogeography of the predominantly Australian plant family Goodeniaceae. *Journal of Biogeography* 41: 2057–2067.
- KADEREIT, G., E.V. MAVRODIEV, E.H. ZACHARIAS, AND A.P. SUKHORUKOV. 2010. Molecular phylogeny of Atripliceae (Chenopodioideae, Chenopodiaceae): implications for systematics, biogeography, flower and fruit evolution, and the origin of C₄ photosynthesis. *American Journal of Botany* 97: 1664–1687.
- KATOH, K., AND D.M. STANDLEY. 2013. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780.
- LARGET, B.R., S.K. KOTHA, C.N. DEWEY, AND C. ANÉ. 2010. BUCKy: gene tree/species

- tree reconciliation with Bayesian concordance analysis. *Bioinformatics* 26: 2910–2911.
- LYLE, M., S. GIBBS, T.C. MOORE, AND D.K. REA. 2007. Late Oligocene initiation of the Antarctic Circumpolar Current: evidence from the South Pacific. *Geology* 35: 691–704.
- MARTIN, H.A. 2006. Cenozoic climatic change and the development of the arid vegetation in Australia. *Journal of Arid Environments* 66: 533–563.
- MAST, A.R., E.F. MILTON, E.H. JONES, R.M. BARKER, W.R. BARKER, AND P.H. WESTON. 2012. Time-calibrated phylogeny of the woody Australian genus *Hakea* (Proteaceae) supports multiple origins of insect-pollination among bird-pollinated ancestors. *American Journal of Botany* 99: 472–487.
- MIRARAB, S., R. REAZ, M.S. BAYZID, T. ZIMMERMANN, M.S. SWENSON, AND T. WARNOW. 2014. ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics* 30: 541–548.
- MOORE, A. J., J. M. DE VOS, L. P. HANCOCK, E. W. GOOLSBY, AND E. J. EDWARDS. 2018. Targeted enrichment of large gene families for phylogenetic inference: phylogeny and molecular evolution of photosynthesis genes in the portullugo clade (Caryophyllales). *Systematic Biology*: doi: 10.1093/sysbio/syx078.
- MUMMENHOFF, K., H. BRÜGGEMANN, AND J.L. BOWMAN. 2001. Chloroplast DNA phylogeny and biogeography of *lepidium* (Brassicaceae). *American Journal of Botany* 88: 2051–2063.

- NYFFELER, R., AND U. EGGLE. 2010. Disintegrating Portulacaceae: a new familial classification of the suborder Portulacineae (Caryophyllales) based on molecular and morphological data. *Taxon* 59: 227-240.
- OBBENS, F. 2006. A review of the tuberous *Calandrinia* species (section Tuberosae), including three new species for Western Australia. *Nutsyia* 16: 95–115.
- OBBENS, F.J. 2011. Five new species of *Calandrinia* (Portulacaceae) from Western Australia with additional information on morphological observations. *Nutsyia* 21: 1–23.
- OCAMPO, G., AND J.T. COLUMBUS. 2010. Molecular phylogenetics of suborder Cactineae (Caryophyllales), including insights into photosynthetic diversification and historical biogeography. *American Journal of Botany* 97: 1827–1847.
- OGBURN, M.R., AND E.J. EDWARDS. 2015. Life history lability underlies rapid climate niche evolution in the angiosperm clade Montiaceae. *Molecular Phylogenetics and Evolution* 92: 181–192.
- OGBURN, R.M., AND E.J. EDWARDS. 2013. Repeated origin of three-dimensional leaf venation releases constraints on the evolution of succulence in plants. *Current biology* 23: 722–726.
- OGBURN, R.M., AND E.J. EDWARDS. 2010. The ecological water-use strategies of succulent plants. *Advances in Botanical Research* 55: 179-255.
- OWEN, C.L., D.C. MARSHALL, K.B.R. HILL, AND C. SIMON. 2017. How the aridification

- of Australia structured the biogeography and influenced the diversification of a large lineage of Australian Cicadas. *Systematic Biology* 66: 569–589.
- PARADIS, E., J. CLAUDE, AND K. STRIMMER. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289–290.
- RABOSKY, D.L. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* 9: e89543–15.
- RABOSKY, D.L., M. GRUNDLER, 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: 701–707.
- RAMBAUT, A., AND A. J. DRUMMOND. 2007. Tracer v1.4, Available from <https://beast.bio.ed.ac.uk/Tracer>
- REE, R.H., AND S.A. SMITH. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* 57: 4–14.
- RONQUIST, F., M. TESLENKO, P. VAN DER MARK, D. L. AYRES, A. DARLING, S. HÖHNA, B. LARGET, ET AL. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.
- SANMARTÍN, I., AND F. RONQUIST. 2004. Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Systematic Biology* 53: 216–243.

- SAUQUET, H., P.H. WESTON, C.L. ANDERSON, N.P. BARKER, D.J. CANTRILL, A.R. MAST, AND V. SAVOLAINEN. 2009. Contrasted patterns of hyperdiversification in Mediterranean hotspots. *Proceedings of the National Academy of Sciences* 106: 221–225.
- SCHER, H.D., AND E.E. MARTIN. 2006. Timing and climatic consequences of the opening of Drake Passage. *Science* 312: 428–430.
- SHEPHERD, K.A., M. WAYCOTT, AND A. CALLADINE. 2004. Radiation of the Australian Salicornioideae (Chenopodiaceae)—based on evidence from nuclear and chloroplast DNA sequences. *American Journal of Botany* 91: 1387–1397.
- STAMATAKIS, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.
- STAMATAKIS, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.
- STOLZER, M., H. LAI, M. XU, D. SATHAYE, B. VERNOT, AND D. DURAND. 2012. Inferring duplications, losses, transfers and incomplete lineage sorting with nonbinary species trees. *Bioinformatics* 28: 409–415.
- TAHIR, S.S., AND P.S. ASHTON. 1989. A cladistic analysis of *Calandrinia* (Portulacaceae). *Pakistan Journal of Botany* 21: 158–169.
- TAHIR, S.S., AND R.C. CAROLIN. 2011. A new species of *Calandrinia* (Portulacaceae) from Northern Territory, Australia. *Proceedings of the Linnean Society of New South*

Wales 133: 11–14.

TAHIR, S.S., AND R.C. CAROLIN. 1990. Phylogeny and origin of Australian (Portulacaceae). *Pakistan Journal of Botany* 22: 1–4.

THORNHILL, A.H., S.Y.W. HO, C. KÜLHEIM, AND M.D. CRISP. 2015. Interpreting the modern distribution of Myrtaceae using a dated molecular phylogeny. *Molecular Phylogenetics and Evolution* 93: 29–43.

TOON, A., M.D. M.D. CRISP, H. GAMAGE, J. MANT, D.C. MORRIS, S. SCHMIDT, AND L.G. COOK. 2015. Key innovation or adaptive change? A test of leaf traits using Triodiinae in Australia. *Scientific Reports* 5: 12398.

VOON POELNITZ, K. 1934. Die *Calandrinia* Arten Australiens. Feddes. Rep., 35: 161–173.

WAGSTAFF, S.J., AND F. HENNION. 2007. Evolution and biogeography of *Lyallia* and *Hectorella* (Portulacaceae), geographically isolated sisters from the Southern Hemisphere. *Antarctic Science* 19: 417–426.

WAGSTAFF, S.J., I. BREITWIESER, AND M. ITO. 2010. Evolution and biogeography of *Pleurophyllum* (Astereae, Asteraceae), a small genus of megaherbs endemic to the subantarctic islands. *American Journal of Botany* 98: 62–75.

WEST, J.G., AND R.J. CHINNOCK. 2013. *Calandrinia mirabilis* (Portulacaceae), a spectacular new species from Western Australia with notes on its ecology, seed germination and horticultural potential. *Journal of the Adelaide Botanic Garden* 26: 97–102.

WINTER, K., AND J.A.C. SMITH. 1996. Crassulacean Acid Metabolism. K. Winter, and J. A. C. Smith [eds.],. Ecological Studies, Berlin, Heidelberg.

WINTER, K., AND J.A.M. HOLTUM. 2011. Induction and reversal of crassulacean acid metabolism in *Calandrinia polyandra*: effects of soil moisture and nutrients. *Functional Plant Biology* 38: 576–582.

WINTER, K., AND J.A.M. HOLTUM. 2014. Facultative crassulacean acid metabolism (CAM) plants: powerful tools for unravelling the functional elements of CAM photosynthesis. *Journal of Experimental Botany* 65: 3425–3441.



Figure 1. Species representatives of Australian *Calandrinia*. (A) *C. sp.* Pink Hills
 (B) *C. tumida* (C) *C. papillata* (D) *C. holtumii* (E) *C. granulifera* (F) *C. pumila*
 (G) *C. polyandra* (H) *C. schistorhiza*

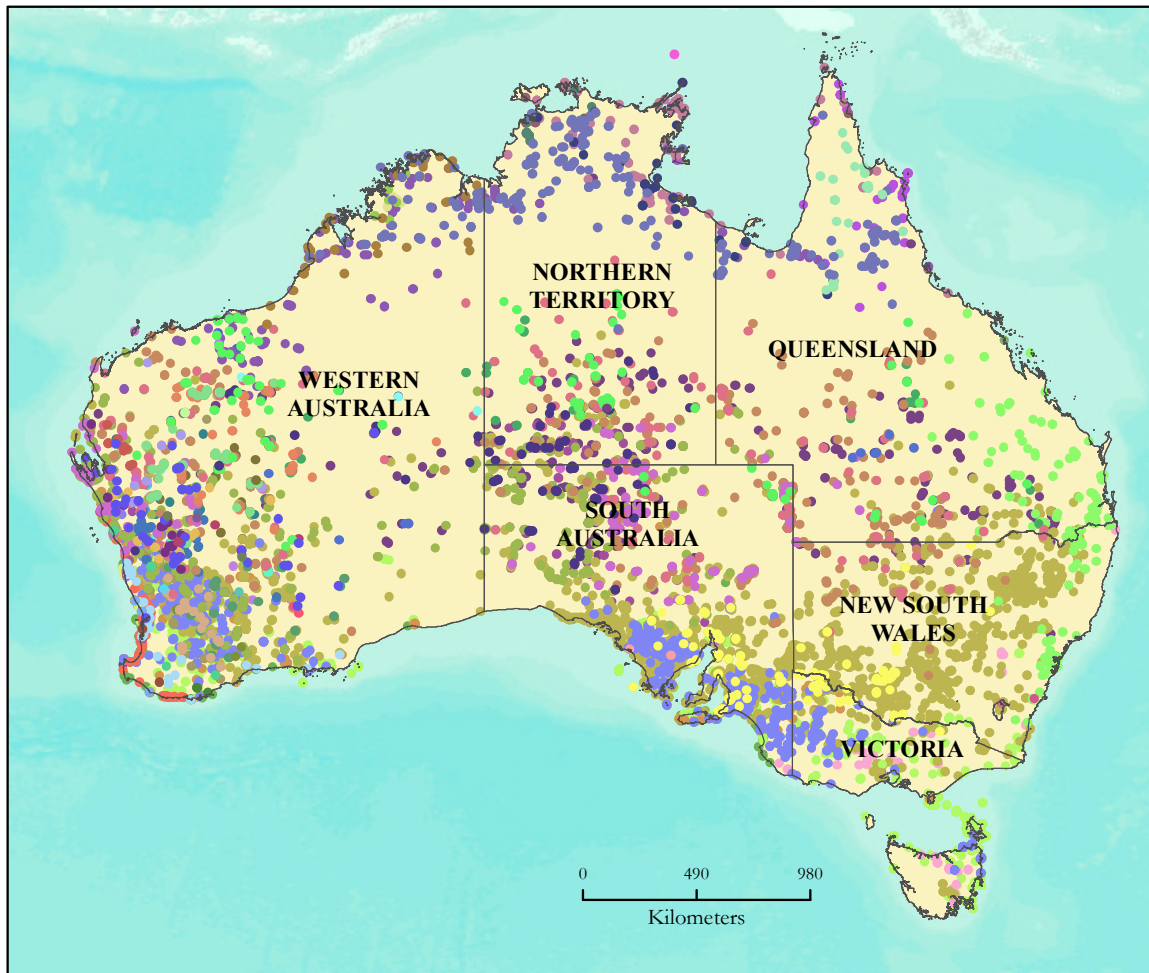


Figure 2. Specimen collections of Australian *Calandrinia*. Different colored circles represent the 74 described and undescribed Australian *Calandrinia* species.

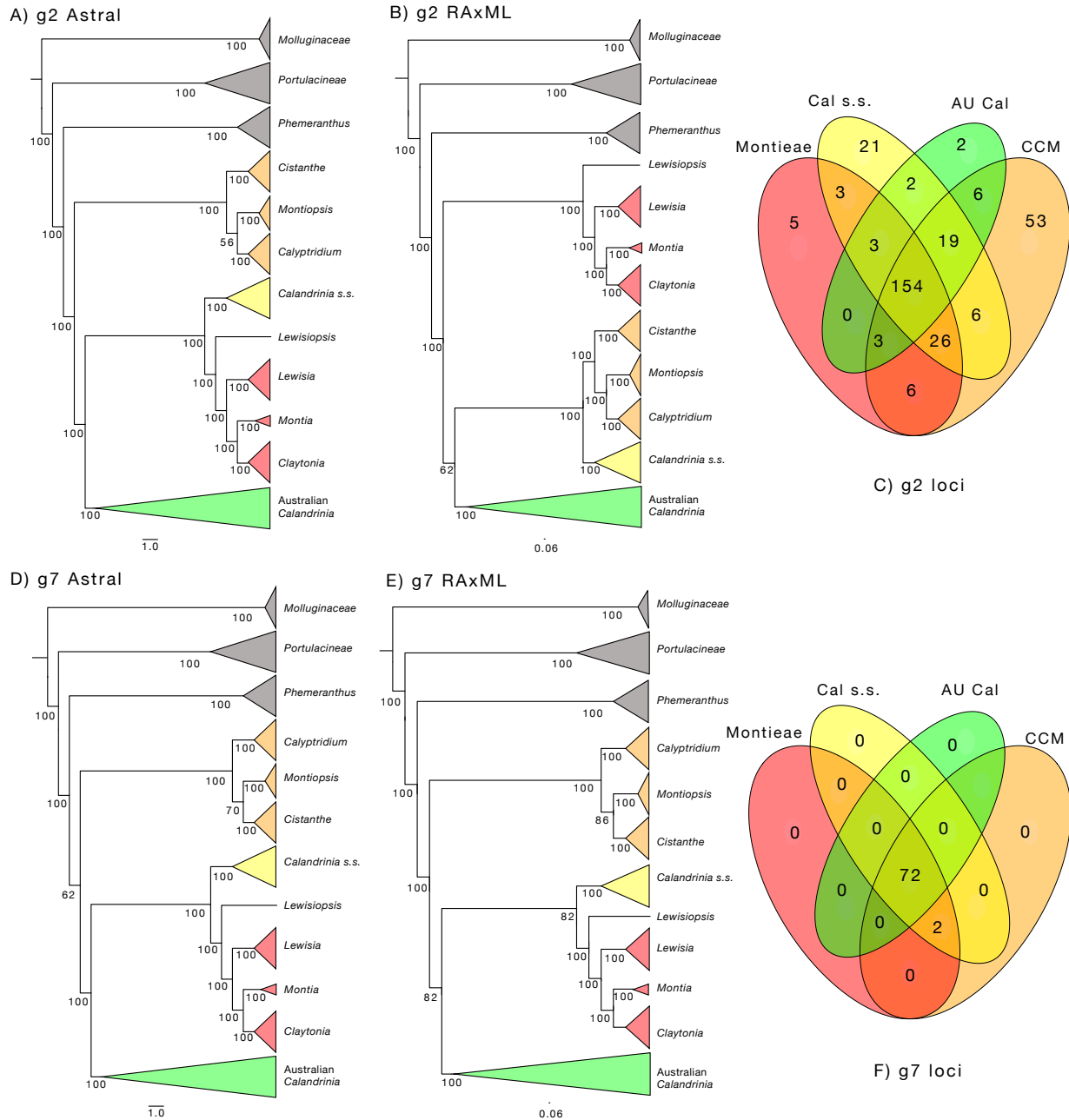


Figure 3. Summary trees depicting the major clades in Montiaceae. Bootstrap support percentages are below branches. (a) g2 dataset representing 297 loci, ASTRAL consensus topology. (b) g2 dataset representing 297 loci, maximum likelihood consensus topology. (c) Venn diagram showing the distribution of unique and shared loci in the g2 dataset across the groups that show topological conflict: CCM group (*Cistanthe* + *Calyptidium* + *Montiopsis*), *Calandrinia* s.s., Australian *Calandrinia*, and Montieae (*Lewisia*, + *Claytonia* + *Montia* + *Lewisiopsis*). (d) g7 dataset representing 115 loci, ASTRAL consensus topology. (e) g7 dataset representing 115 loci, maximum likelihood consensus topology. (f) Venn diagram showing the distribution of loci across groups (CCM, *Calandrinia* s.s., Montieae, Australian *Calandrinia*) that show topological conflict using the g7 dataset.

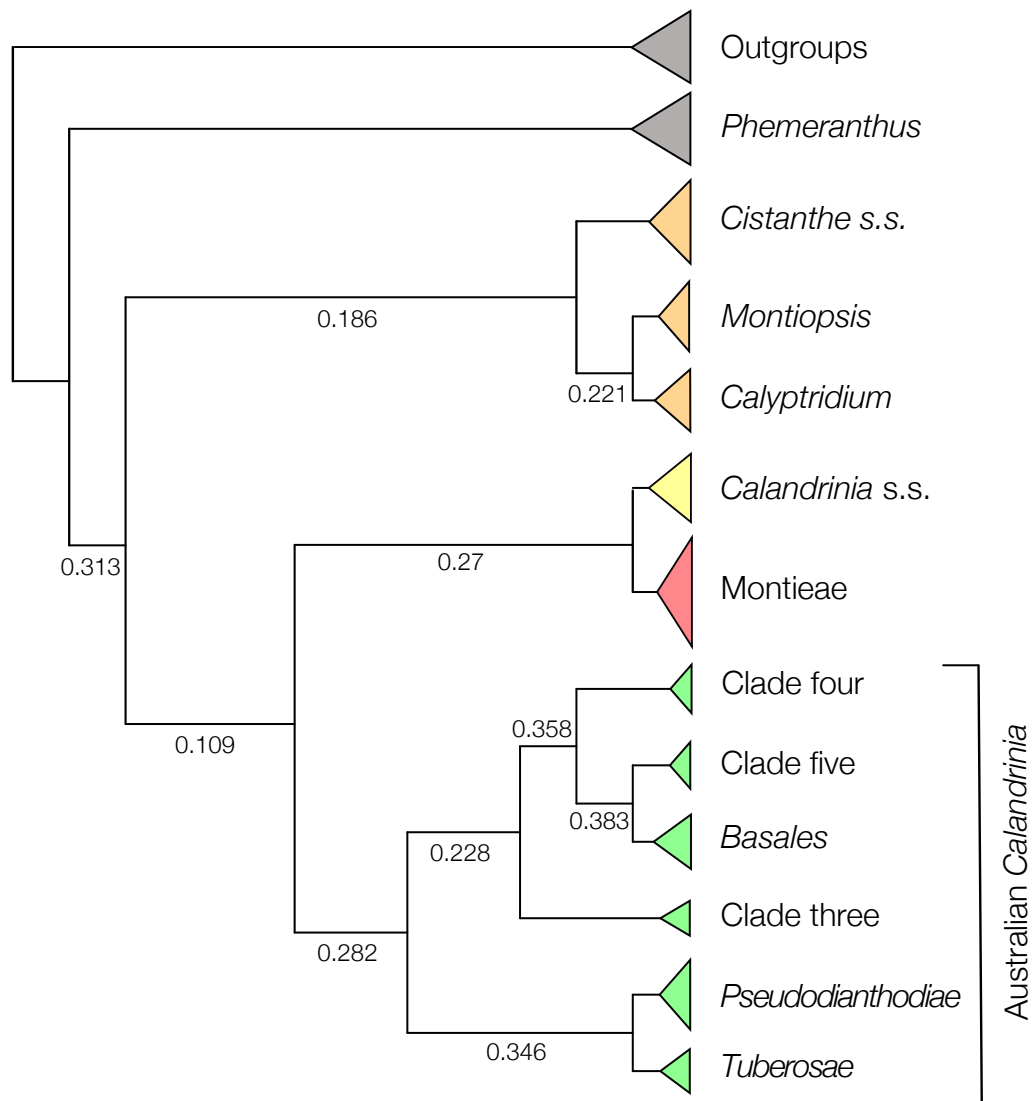


Figure 4. Bayesian Concordance Analysis primary concordance tree for Montiaceae backbone relationships including the six Australian *Calandrinia* clades. The concordance factor for a given split is below the node.

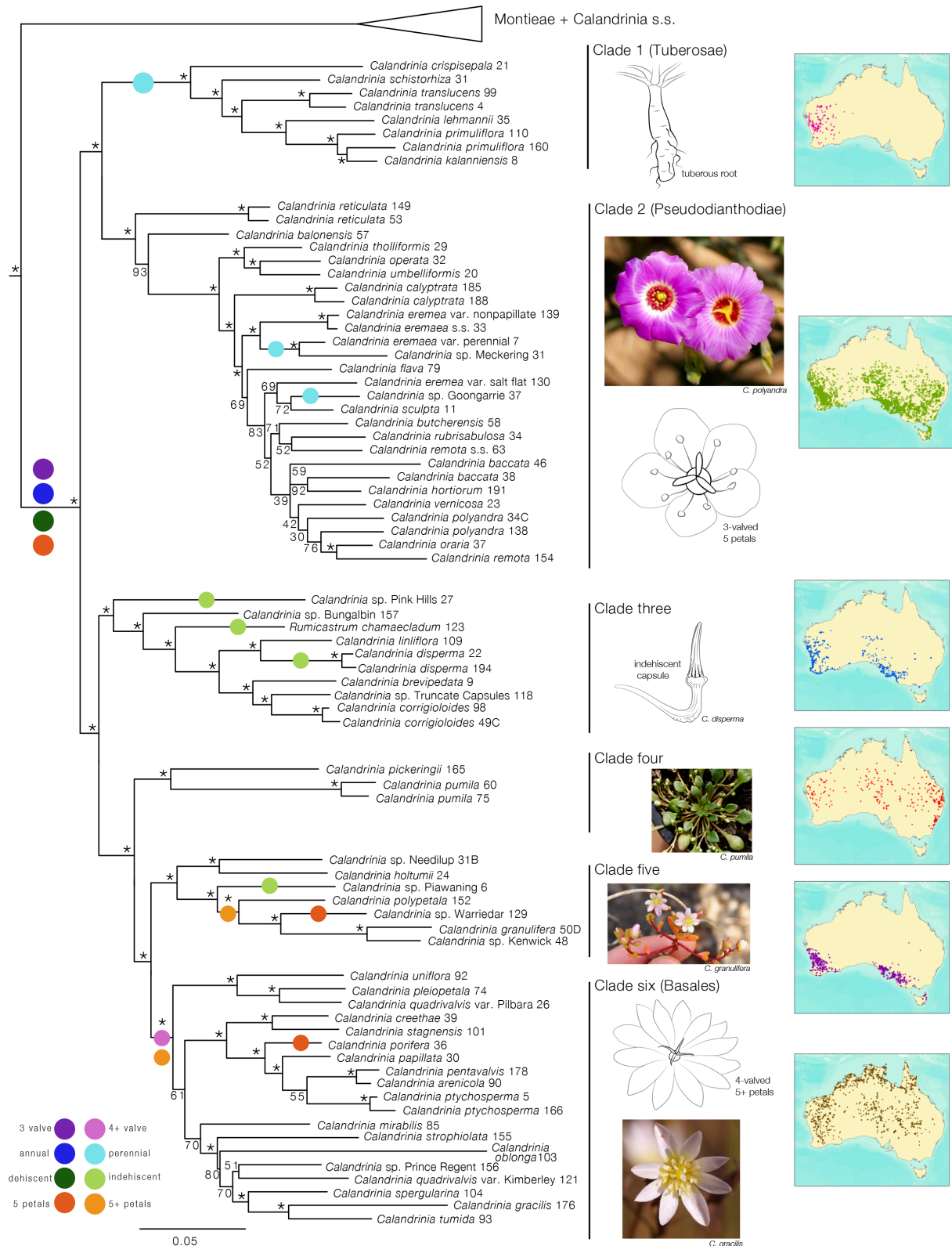


Figure 5. Australian *Calandrinia* phylogeny reconstructed with the g2 loci dataset (297 loci) with g2 RAXML branch lengths. Bootstrap support values with greater than 95% support indicated with a star. Ancestral character states and shifts in character traits are mapped onto the phylogeny, with purple circles representing valve number, blue circles life history habit, green circles capsule type, and orange circles petal number. Maps correspond to each of the six clades, with colored dots representing species distributions for the given clade. For more ancestral character state reconstruction detail see Appendices S10 – S13.

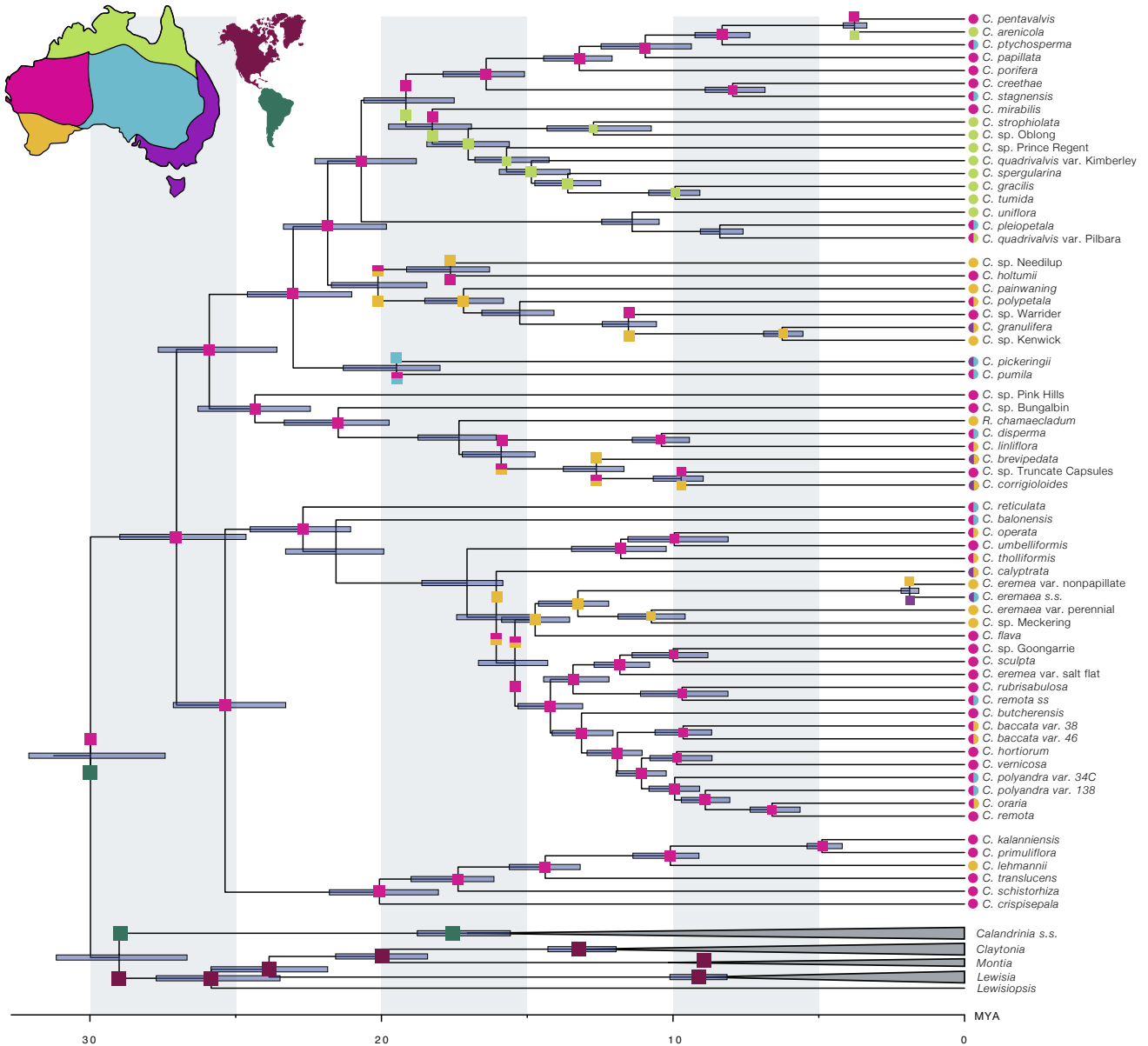


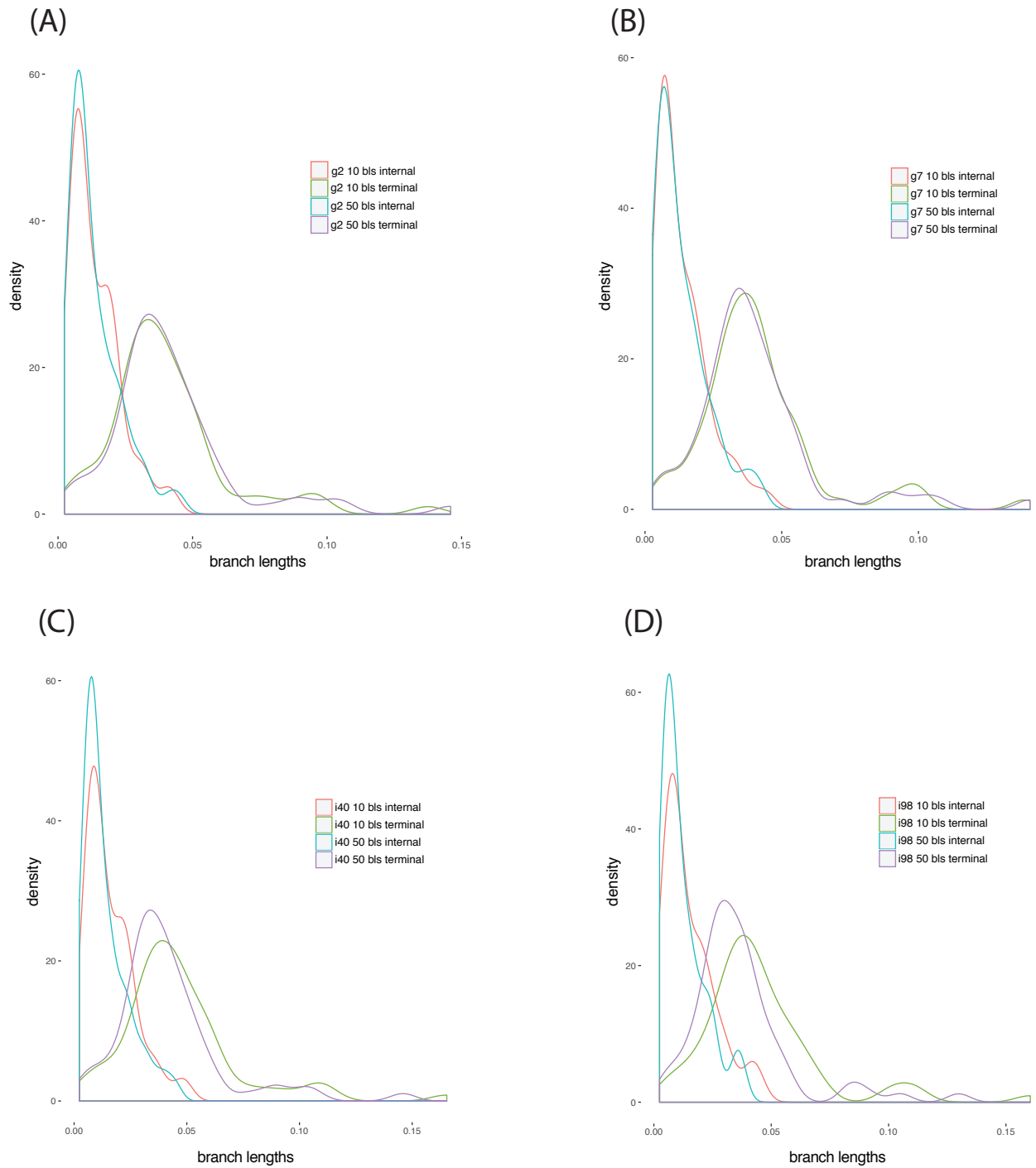
Figure 6. A) Map of Australia colored by geographic region plus North and South America. B) Time-calibrated phylogeny made with BEAST of Australian *Calandrinia* and its sister lineages. Colored boxes reconstructed at the nodes correspond to the geographic regions as denoted in figure 6.A. Only the ancestral states with greater than 50% probability are reconstructed on the tree. If there was equal probability of a region at a node, both states were reconstructed.

Appendix S1. Voucher Specimen Table. Individuals included in the present study with corresponding voucher number, DNA extraction identifier, and number of loci recovered, and SRA numbers.

Family	Genus	Species	Author	Taxon	Voucher number	DNA extraction #	Taxonomic group	Total loci	SRA number
Montiaceae	<i>Calandrinia</i>	<i>arenicola</i>	Sydes	<i>Calandrinia arenicola</i> Sydes	JAH 137 (JCU)	iq949	AU Calandrinia	165	SAMN07987712
Montiaceae	<i>Calandrinia</i>	<i>baccata</i>	Obbens	<i>Calandrinia baccata</i> Obbens	LPH 46 (BRU)	iqh17	AU Calandrinia	121	SAMN07987714
Montiaceae	<i>Calandrinia</i>	<i>baccata</i>	Obbens	<i>Calandrinia baccata</i> Obbens	PERTH06017991	iqh38	AU Calandrinia	15	SAMN07987713
Montiaceae	<i>Calandrinia</i>	<i>balsamestris</i>	Lindl.	<i>Calandrinia balsamestris</i> Lindl.	PERTH08297730	iqh57	AU Calandrinia	181	SAMN07987712
Montiaceae	<i>Calandrinia</i>	<i>brevipedata</i>	F.Muell.	<i>Calandrinia brevipedata</i> F.Muell.	LPH 50 A (BRU)	iq949	AU Calandrinia	163	SAMN07987716
Montiaceae	<i>Calandrinia</i>	<i>buckneriana</i>	Obbens	<i>Calandrinia buckneriana</i> Obbens	PERTH08336474	iqh58	AU Calandrinia	153	SAMN07987717
Montiaceae	<i>Calandrinia</i>	<i>calyptrata</i>	Hook.f.	<i>Calandrinia calyptrata</i> Hook.f.	LPH 52 (BRU)	iqh185	AU Calandrinia	167	SAMN07987718
Montiaceae	<i>Calandrinia</i>	<i>calyptrata</i>	Hook.f.	<i>Calandrinia calyptrata</i> Hook.f.	JAH 46 (JCU)	iqh188	AU Calandrinia	103	SAMN07159085
Montiaceae	<i>Calandrinia</i>	<i>corrigioloides</i>	Benth.	<i>Calandrinia corrigioloides</i> Benth. (island)	PERTH08136142	iqh98	AU Calandrinia	91	SAMN07987724
Montiaceae	<i>Calandrinia</i>	<i>corrigioloides</i>	Benth.	<i>Calandrinia corrigioloides</i> Benth. (coastal)	LPH 49 C (BRU)	iqh18	AU Calandrinia	178	SAMN07987725
Montiaceae	<i>Calandrinia</i>	<i>crebbae</i>	Morrison	<i>Calandrinia crebbae</i> Morrison	LPH 170 (BRU)	iqh161	AU Calandrinia	182	SAMN07839381
Montiaceae	<i>Calandrinia</i>	<i>crispipetala</i>	Obbens	<i>Calandrinia crispipetala</i> Obbens	PERTH08136415	iqh21	AU Calandrinia	130	SAMN07987727
Montiaceae	<i>Calandrinia</i>	<i>disperma</i>	J.M.Black	<i>Calandrinia disperma</i> J.M.Black	PERTH08018172	iqh22	AU Calandrinia	162	SAMN07987729
Montiaceae	<i>Calandrinia</i>	<i>disperma</i>	J.M.Black	<i>Calandrinia disperma</i> J.M.Black	PERTH07357443	iqh194	AU Calandrinia	169	SAMN07987728
Montiaceae	<i>Calandrinia</i>	<i>eremaea</i>	Ewart	<i>Calandrinia eremaea</i> Ewart (perennial)	LPH 41 (BRU)	iqh67	AU Calandrinia	171	SAMN07987730
Montiaceae	<i>Calandrinia</i>	<i>eremaea</i>	Ewart	<i>Calandrinia eremaea</i> Ewart (non-papillate seeds)	PERTH08597871	iqh139	AU Calandrinia	162	SAMN07987732
Montiaceae	<i>Calandrinia</i>	<i>eremaea</i>	Ewart	<i>Calandrinia eremaea</i> Ewart (salt-flats)	PERTH08470178	iqh130	AU Calandrinia	129	SAMN07987733
Montiaceae	<i>Calandrinia</i>	<i>eremaea</i>	Ewart	<i>Calandrinia eremaea</i> Ewart s.s.	PERTH08336369	iqh23	AU Calandrinia	169	SAMN07987731
Montiaceae	<i>Calandrinia</i>	<i>flava</i>	Obbens	<i>Calandrinia flava</i> Obbens	JAH 6 (JCU)	iqh79	AU Calandrinia	160	SAMN07159086
Montiaceae	<i>Calandrinia</i>	<i>gracilis</i>	Benth.	<i>Calandrinia gracilis</i> Benth.	LPH 124 (BRU)	iqh176	AU Calandrinia	200	SAMN07987734
Montiaceae	<i>Calandrinia</i>	<i>graciliflora</i>	Benth.	<i>Calandrinia graciliflora</i> Benth.	LPH 50 B (BRU)	iqh14	AU Calandrinia	206	SAMN07987735
Montiaceae	<i>Calandrinia</i>	<i>hololoma</i>	Obbens	<i>Calandrinia hololoma</i>	PERTH0873928	iqh24	AU Calandrinia	156	SAMN07987765
Montiaceae	<i>Calandrinia</i>	<i>hortorum</i>	Obbens	<i>Calandrinia hortorum</i> Obbens	LPH 190 (BRU)	iqh191	AU Calandrinia	155	SAMN07987736
Montiaceae	<i>Calandrinia</i>	<i>kalaueriana</i>	Obbens	<i>Calandrinia kalaueriana</i> Obbens	LPH 36 (BRU)	iqh8	AU Calandrinia	133	SAMN07987737
Montiaceae	<i>Calandrinia</i>	<i>lehmannii</i>	Endl.	<i>Calandrinia lehmannii</i> Endl.	PERTH06707955	iqh35	AU Calandrinia	91	SAMN07159088
Montiaceae	<i>Calandrinia</i>	<i>inflata</i>	Fend.	<i>Calandrinia inflata</i> Fend.	LPH 49 A (BRU)	iqh109	AU Calandrinia	142	SAMN07987738
Montiaceae	<i>Calandrinia</i>	<i>invalabilis</i>	Chinnock & J.G.West	<i>Calandrinia invalabilis</i> Chinnock & J.G.West	LPH 24 C (BRU)	iqh45	AU Calandrinia	186	SAMN07159089
Montiaceae	<i>Calandrinia</i>	<i>oblonga</i>	Sydes & Carolin	<i>Calandrinia oblonga</i> Sydes & Carolin	LPH 120 (BRU)	iqh103	AU Calandrinia	12	SAMN07987771
Montiaceae	<i>Calandrinia</i>	<i>operta</i>	Obbens	<i>Calandrinia operta</i> Obbens	PERTH08336185	iqh32	AU Calandrinia	138	SAMN07987739
Montiaceae	<i>Calandrinia</i>	<i>ovata</i>	Obbens	<i>Calandrinia ovata</i> Obbens	PERTH08019249	iqh177	AU Calandrinia	150	SAMN07987740
Montiaceae	<i>Calandrinia</i>	<i>papillata</i>	Sydes	<i>Calandrinia papillata</i> Sydes	PERTH08477302	iqh30	AU Calandrinia	98	SAMN07987742
Montiaceae	<i>Calandrinia</i>	<i>penicillata</i>	Obbens	<i>Calandrinia penicillata</i> Obbens	LPH 157 (BRU)	iqh178	AU Calandrinia	151	SAMN07987743
Montiaceae	<i>Calandrinia</i>	<i>pickersleri</i>	A.Gray	<i>Calandrinia pickersleri</i> A.Gray	JAH 40 (JCU)	iqh16	AU Calandrinia	227	SAMN07987742
Montiaceae	<i>Calandrinia</i>	<i>placopetala</i>	F.Muell.	<i>Calandrinia placopetala</i> F.Muell.	JAH 3 (JCU)	iqh74	AU Calandrinia	157	SAMN07987744
Montiaceae	<i>Calandrinia</i>	<i>polypandra</i>	(Hook.) Benth.	<i>Calandrinia polypandra</i> (Hook.) Benth.	LPH 175 (BRU)	iqh138	AU Calandrinia	135	SAMN07987745
Montiaceae	<i>Calandrinia</i>	<i>polypandra</i>	(Hook.) Benth.	<i>Calandrinia polypandra</i> (Hook.) Benth.	iqh16	iqh16	AU Calandrinia	150	SAMN07987741
Montiaceae	<i>Calandrinia</i>	<i>polyneala</i>	(Hook.) Benth.	<i>Calandrinia polyneala</i> (Hook.) Benth.	PERTH06708013	iqh152	AU Calandrinia	140	SAMN07987747
Montiaceae	<i>Calandrinia</i>	<i>porifera</i>	Sydes	<i>Calandrinia porifera</i> Sydes	PERTH08772599	iqh36	AU Calandrinia	265	SAMN07987748
Montiaceae	<i>Calandrinia</i>	<i>primulaeflora</i>	Diels	<i>Calandrinia primulaeflora</i> Diels	PERTH08114668	iqh10	AU Calandrinia	135	SAMN07987749
Montiaceae	<i>Calandrinia</i>	<i>primulaeflora</i>	Diels	<i>Calandrinia primulaeflora</i> Diels	PERTH08174229	iqh169	AU Calandrinia	175	SAMN07987750
Montiaceae	<i>Calandrinia</i>	<i>psycosperma</i>	F.Muell.	<i>Calandrinia psycosperma</i> F.Muell.	LPH 169 (BRU)	iqh166	AU Calandrinia	242	SAMN07987751
Montiaceae	<i>Calandrinia</i>	<i>psycosperma</i>	F.Muell.	<i>Calandrinia psycosperma</i> F.Muell.	LPH 139 (BRU)	iqh173	AU Calandrinia	173	SAMN07987752
Montiaceae	<i>Calandrinia</i>	<i>pumila</i>	F.Muell.	<i>Calandrinia pumila</i> F.Muell.	KRT 4824 (PERTH)	iqh60	AU Calandrinia	152	SAMN07159090
Montiaceae	<i>Calandrinia</i>	<i>pumila</i>	F.Muell.	<i>Calandrinia pumila</i> F.Muell.	JAH 1 (JCU)	iqh75	AU Calandrinia	167	SAMN07987753
Montiaceae	<i>Calandrinia</i>	<i>quadrivalvis</i>	F.Muell.	<i>Calandrinia quadrivalvis</i> F.Muell. s.s.	LPH 124 (BRU)	iqh26	AU Calandrinia	130	SAMN07987755
Montiaceae	<i>Calandrinia</i>	<i>quadrivalvis</i>	F.Muell.	<i>Calandrinia quadrivalvis</i> F.Muell. (Kimberley)	PERTH08072132	iqh121	AU Calandrinia	145	SAMN07987754
Montiaceae	<i>Calandrinia</i>	<i>remota</i>	J.M.Black	<i>Calandrinia remota</i> J.M.Black s.s.	LPH 126 (BRU)	iqh154	AU Calandrinia	138	SAMN07987756
Montiaceae	<i>Calandrinia</i>	<i>remota</i>	J.M.Black	<i>Calandrinia remota</i> J.M.Black s.s.	PERTH06708315	iqh63	AU Calandrinia	159	SAMN07987757
Montiaceae	<i>Calandrinia</i>	<i>reticulata</i>	Sydes	<i>Calandrinia reticulata</i> Sydes	JAH 12 (JCU)	iqh53	AU Calandrinia	189	SAMN07987759
Montiaceae	<i>Calandrinia</i>	<i>reticulata</i>	Sydes	<i>Calandrinia reticulata</i> Sydes	LPH 126 (BRU)	iqh48	AU Calandrinia	180	SAMN07987758
Montiaceae	<i>Calandrinia</i>	<i>rubrastralis</i>	Obbens	<i>Calandrinia rubrastralis</i> Obbens	PERTH088336547	iqh34	AU Calandrinia	139	SAMN07987760
Montiaceae	<i>Calandrinia</i>	<i>schottii</i>	Morrison	<i>Calandrinia schottii</i> Morrison	LPH 148 (BRU)	iqh117	AU Calandrinia	130	SAMN07839384
Montiaceae	<i>Calandrinia</i>	<i>scutellaria</i>	Obbens & J.G.West	<i>Calandrinia scutellaria</i> Obbens & J.G.West	LPH 200 (BRU)	iqh161	AU Calandrinia	160	SAMN07987761
Montiaceae	<i>Calandrinia</i>	<i>sp.</i>	Bungabin (G.J.Keighery & N.Gibson 1656)	<i>Calandrinia sp.</i> Bungabin (G.J.Keighery & N.Gibson 1656)	PERTH0874455	iqh17	AU Calandrinia	152	SAMN07987766
Montiaceae	<i>Calandrinia</i>	<i>sp.</i>	Goongaree (F. Obbens, F. Hort & J. Hort FO 1813)	<i>Calandrinia sp.</i> Goongaree (F. Obbens, F. Hort & J. Hort FO 1813)	LPH 37 (BRU)	iqh15	AU Calandrinia	77	SAMN07987767
Montiaceae	<i>Calandrinia</i>	<i>sp.</i>	Kewick (G.J.Keighery 10905)	<i>Calandrinia sp.</i> Kewick (G.J.Keighery 10905)	PERTH074903	iqh16	AU Calandrinia	161	SAMN07987768
Montiaceae	<i>Calandrinia</i>	<i>sp.</i>	Meckering (F. Obbens 4202)	<i>Calandrinia sp.</i> Meckering (F. Obbens 4202)	LPH 31 A (BRU)	iqh1	AU Calandrinia	40	SAMN07987769
Montiaceae	<i>Calandrinia</i>	<i>sp.</i>	Needilup (K.R.Newbey 4892)	<i>Calandrinia sp.</i> Needilup (K.R.Newbey 4892)	LPH 31 B (BRU)	iqh3	AU Calandrinia	162	SAMN07987770
Montiaceae	<i>Calandrinia</i>	<i>sp.</i>	Needilup (K.R.Newbey 4892)	<i>Calandrinia sp.</i> Needilup (K.R.Newbey 4892)	LPH 178 (BRU)	iqh217	AU Calandrinia	132	SAMN07987771
Montiaceae	<i>Calandrinia</i>	<i>sp.</i>	The Pink Hills (F.Obbens FO 1906)	<i>Calandrinia sp.</i> The Pink Hills (F.Obbens FO 1906)	LPH 178 (BRU)	iqh151	AU Calandrinia	161	SAMN07987772
Montiaceae	<i>Calandrinia</i>	<i>sp.</i>	Pawaring (A.C.Boagbiche 12257)	<i>Calandrinia sp.</i> Pawaring (A.C.Boagbiche 12257)	LPH 28 (BRU)	iqh6	AU Calandrinia	98	SAMN07987741
Montiaceae	<i>Calandrinia</i>	<i>sp.</i>	Prince Regent (M.D. Barrett 1590)	<i>Calandrinia sp.</i> Prince Regent (M.D. Barrett 1590)	PERTH074903	iqh18	AU Calandrinia	161	SAMN07987741
Montiaceae	<i>Calandrinia</i>	<i>sp.</i>	Truncate Cupules (A.Markey & S.Dillon 1474)	<i>Calandrinia sp.</i> Truncate Cupules (A.Markey & S.Dillon 1474)	LPH 194 (BRU)	iqh118	AU Calandrinia	176	SAMN07987774
Montiaceae	<i>Calandrinia</i>	<i>sp.</i>	Wardlaw (F.Obbens 0409)	<i>Calandrinia sp.</i> Wardlaw (F.Obbens 0409)	PERTH08195	iqh129	AU Calandrinia	159	SAMN07987775
Montiaceae	<i>Calandrinia</i>	<i>spregeliana</i>	F.Muell.	<i>Calandrinia spregeliana</i> F.Muell.	LPH 156 (BRU)	iqh104	AU Calandrinia	164	SAMN07987777
Montiaceae	<i>Calandrinia</i>	<i>sugensii</i>	(F.Muell.) Ewart, B.Rees & B.Wood	<i>Calandrinia sugensii</i> (F.Muell.) Ewart, B.Rees & B.Wood	LPH 90 (BRU)	iqh101	AU Calandrinia	158	SAMN07987777
Montiaceae	<i>Calandrinia</i>	<i>tholiformis</i>	Obbens	<i>Calandrinia tholiformis</i> Obbens	PERTH07184174	iqh165	AU Calandrinia	79	SAMN07987778
Montiaceae	<i>Calandrinia</i>	<i>translucens</i>	Obbens	<i>Calandrinia translucens</i> Obbens	LPH 34 B (BRU)	iqh4	AU Calandrinia	146	SAMN07987780
Montiaceae	<i>Calandrinia</i>	<i>translucens</i>	Obbens	<i>Calandrinia translucens</i> Obbens	PERTH08473434	iqh99	AU Calandrinia	161	SAMN07987781
Montiaceae	<i>Calandrinia</i>	<i>namda</i>	Sydes	<i>Calandrinia namda</i> Sydes	JAH 138 (JCU)	iqh93	AU Calandrinia	149	SAMN07987782
Montiaceae	<i>Calandrinia</i>	<i>umbelliformis</i>	Obbens	<i>Calandrinia umbelliformis</i> Obbens	PERTH08373660	iqh20	AU Calandrinia	169	SAMN07987783
Montiaceae	<i>Calandrinia</i>	<i>uniflora</i>	F.Muell.	<i>Calandrinia uniflora</i> F.Muell.	JAH 7 (JCU)	iqh92	AU Calandrinia	154	SAMN07987784
Montiaceae	<i>Calandrinia</i>	<i>vernicosa</i>	Obbens	<i>Calandrinia vernicosa</i> Obbens	PERTH08336393	iqh23	AU Calandrinia	144	SAMN07987785
Montiaceae	<i>Rumicrastrum</i>	<i>chamaecladum</i>	(Diels) Ulbr.	<i>Rumicrastrum chamaecladum</i> (Diels) Ulbr.	PERTH08323148	iqh123	AU Calandrinia	211	SAMN07987794
Montiaceae	<i>Calandrinia</i>	<i>affinis</i>	Gillies ex Arn.	<i>Calandrinia affinis</i> Gillies ex Arn.	MO 253 (BRU)	iqh145	Calandrinia s.s.	138	SAMN07987771
Montiaceae	<i>Calandrinia</i>	<i>clausa</i>	DC.	<i>Calandrinia clausa</i> DC.	MO 273 (BRU)	iqh2014	Calandrinia s.s.	224	SAMN07987719
Montiaceae	<i>Calandrinia</i>	<i>clausa</i>	DC.	<i>Calandrinia clausa</i> DC.	Vanderveit72074 (HUSA1)	iqh215	Calandrinia s.s.	210	SAMN07987720
Montiaceae	<i>Calandrinia</i>	<i>clausa</i>	DC.	<i>Calandrinia clausa</i> DC.	F701 (BRU)	iqh216	Calandrinia s.s.	216	SAMN07987721
Montiaceae	<i>Calandrinia</i>	<i>colchaguensis</i>	Barnoud	<i>Calandrinia colchaguensis</i> Barnoud	MO 273 (BRU)	iqh146	Calandrinia s.s.	142	SAMN07987722
Montiaceae	<i>Calandrinia</i>	<i>compressa</i>	Schrad. ex DC.	<i>Calandrinia compressa</i> Schrad. ex DC.	MO 276 (BRU)	iqh144	Calandrinia s.s.	173	SAMN07987723
Montiaceae	<i>Calandrinia</i>	<i>monandra</i>	DC.	<i>Calandrinia monandra</i> DC.	MO 275 (BRU)	iqh143	Calandrinia s.s.	146	SAMN07987724
Montiaceae	<i>Calandrinia</i>	<i>skottbergii</i>	Gand.	<i>Calandrinia skottbergii</i> Gand.	PCU 88604 (CONC)	iqh222	Calandrinia s.s.	166	SAMN07987763
Montiaceae	<i>Calandrinia</i>	<i>sp.</i>	(L.) Thulin	<i>Calandrinia sp.</i> (L.) Thulin	LPH 29 (BRU)	iqh220	Calandrinia s.s.	152	SAMN07987764
Molluginaceae	<i>Triglochin</i>	<i>arcta</i>	L.	<i>Triglochin arcta</i> L.	LPH 117 (BRU)	iqh148	Molluginaceae	109	SAMN07159111
Molluginaceae	<i>verruculata</i>		L.	<i>verruculata</i>	Moore & Dewenter 1631 (BRU)	AM56	Molluginaceae	121	SAMN07159106
Montiaceae	<i>Calyptidium</i>	<i>monospermum</i>	Greene	<i>Calyptidium monospermum</i> Greene	LPH 221 (BRU)	iqh108	Montiaceae	274	SAMN07987786
Montiaceae	<i>Calyptidium</i>	<i>parryi</i>	A.Gray	<i>Calyptidium parryi</i> A.Gray	LPH 209 (BRU)	iqh170	Montiaceae	364	SAMN07159091
Montiaceae	<i>Calyptidium</i>	<i>umbellatum</i>	Greene	<i>Calyptidium umbellatum</i> Greene	LPH 17 (BRU)	iqh193	Montiaceae	271	SAMN07159092
Montiaceae	<i>Cuscuta</i>	<i>gondiflora</i>	(Lindl.) Schindl.	<i>Cuscuta gondiflora</i> (Lindl.) Schindl.	MO194 (BRU)	iqh199	Montiaceae	165	SAMN07159094
Montiaceae	<i>Cuscuta</i>	<i>parviflora</i>	(DC.) Carolin ex Herbk.	<i>Cuscuta parviflora</i> (DC.) Carolin ex Herbk.	EE 2014 (11 (BRU))	iqh223	Montiaceae	244	SAMN07987787
Montiaceae	<i>Cuscuta</i>	<i>sp.</i>	(Barnoud) D.I.Ford	<i>Cuscuta sp.</i> (Barnoud) D.I.Ford	MA 2250 (BRU)	iqh219	Montiaceae	237	SAMN07987788
Montiaceae	<i>Montipha</i>	<i>umbellata</i>	(Rau & Pav.) D.I.Ford	<i>Montipha umbellata</i> (Rau & Pav.) D.I.Ford	MO 253 (BRU)	iqh225	Montiaceae	224	SAMN07987789
Montiaceae	<i>Clytostoma</i>	<i>lanceolatum</i>	Pall. ex Pursh	<i>Clytostoma lanceolatum</i> Pall. ex Pursh	ipb 4 (BRU)	iqh106	Montiaceae	200	SAMN07987789
Montiaceae	<i>Clytostoma</i>	<i>negundo</i>	Kuntze	<i>Clytostoma negundo</i> Kuntze	LPH 230 (BRU)	iqh195	Montiaceae	118	SAMN07159095
Montiaceae	<i>Clytostoma</i>	<i>vaughaniana</i>	Sukd.	<i>Clytostoma vaughaniana</i> Sukd.	de Vos 304 (BRU)	AV26	Montiaceae	157	SAMN07159096
Montiaceae	<i>Levistia</i>	<i>brachycalyx</i>	Engelm. ex A.Gray	<i>Levistia brachycalyx</i> Engelm. ex A.Gray	de Vos 260 (BRU)	AV16	Montiaceae	130	SAMN07839350
Montiaceae	<i>Levistia</i>	<i>condensata</i>	(Rydb.) S.Gray	<i>Levistia condensata</i> (Rydb.) S.Gray	de Vos 247 (BRU)	AV17	Montiaceae	141	SAMN07159099
Montiaceae	<i>Levistia</i>	<							

Appendix S2. Bootstrap support values for backbone Montiaceae and Australian Calandrinia relationships using both concatenation and coalescent-based (ASTRAL) approaches with the g2, g7, i40, and i98 datasets. Bootstrap values are listed for the g2, g7, i40, and i98 datasets where all sites present in <10% individuals have been removed and all sites present in <50% of the individuals have been removed. Appendix S4. ASTRAL topology from the g7 loci dataset (116 individuals, 116 loci). Bootstrap support values are below branches.

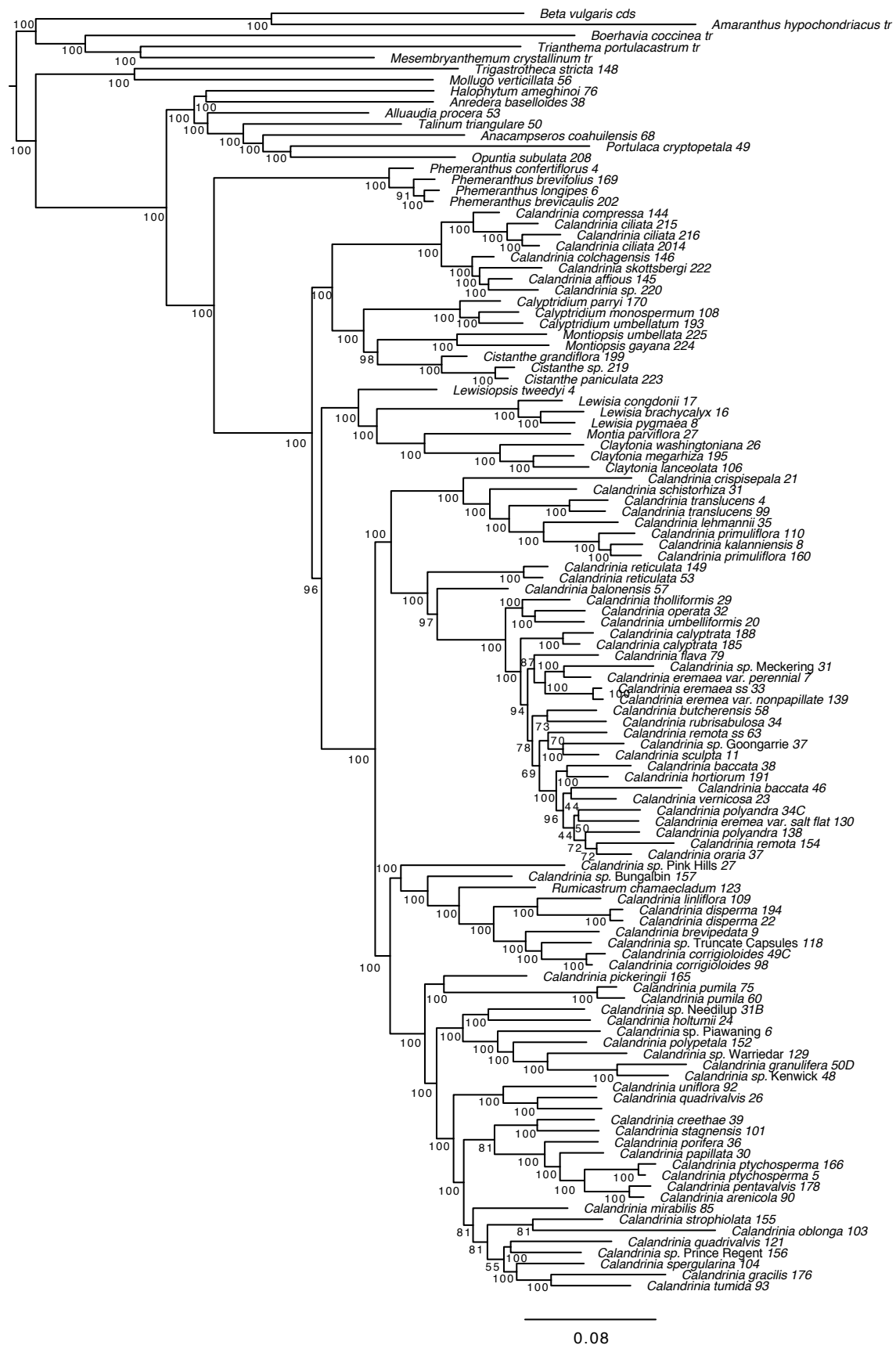
node	g2 concatenated		g7 concatenated		g2 ASTRAL		g7 ASTRAL		i40 concatenated		i98 concatenated		i40 ASTRAL		i98 ASTRAL	
	(10% missing sites removed)	(50% missing sites removed)	(10% missing sites removed)	(50% missing sites removed)	(10% missing sites removed)	(50% missing sites removed)	(10% missing sites removed)	(50% missing sites removed)	(10% missing sites removed)	(50% missing sites removed)	(10% missing sites removed)	(50% missing sites removed)	(10% missing sites removed)	(50% missing sites removed)	(10% missing sites removed)	(50% missing sites removed)
Montiaceae + (Halophytaceae, Bachelidae, Dipterocarpaceae, Tiliaceae, Anacardiaceae, Portulacaceae, Cucurbitaceae)																
<i>Phoradendron</i> + (<i>Calandrinia</i> s.s. + CCM + Montiaceae + Australian <i>Calandrinia</i>)																
<i>Montiopsis</i> + <i>Cistanche</i>	NA	NA	100	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Calytridium</i> + <i>Calytridium</i>	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Calytridium</i> + <i>Calytridium</i> + <i>Montiopsis</i>	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Calandrinia</i> s.s. + CCM	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Calandrinia</i> s.s. + CCM	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Montiaceae + (CCM + <i>Calandrinia</i> s.s. + Australian <i>Calandrinia</i>)	NA	NA	100	100	100	100	100	100	100	100	100	100	100	100	100	100
CCM + <i>Calandrinia</i> s.s. + Montiaceae + Australian <i>Calandrinia</i>	NA	NA	100	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Calytridium</i> + <i>Calytridium</i>	NA	NA	100	100	100	100	100	100	100	100	100	100	100	100	100	100
CCM + <i>Calandrinia</i> s.s. + Montiaceae + Australian <i>Calandrinia</i>	NA	NA	100	100	100	100	100	100	100	100	100	100	100	100	100	100
(CCM + <i>Calandrinia</i> s.s.) + (Montiaceae + Australian <i>Calandrinia</i>)	NA	NA	100	100	100	100	100	100	100	100	100	100	100	100	100	100
<i>Calandrinia</i> s.s. + Montiaceae	NA	NA	100	100	100	100	100	100	100	100	100	100	100	100	100	100
(<i>Calandrinia</i> s.s. + Montiaceae) + Australian <i>Calandrinia</i>	NA	NA	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Montiaceae + Australian <i>Calandrinia</i>	NA	NA	100	100	100	100	100	100	100	100	100	100	100	100	100	100
CLD 3 + (CLD 4 + CLD 5 + BAS + CLD 3)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
CLD 4 + (CLD 5 + BAS)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
CLD 5 + BAS	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100



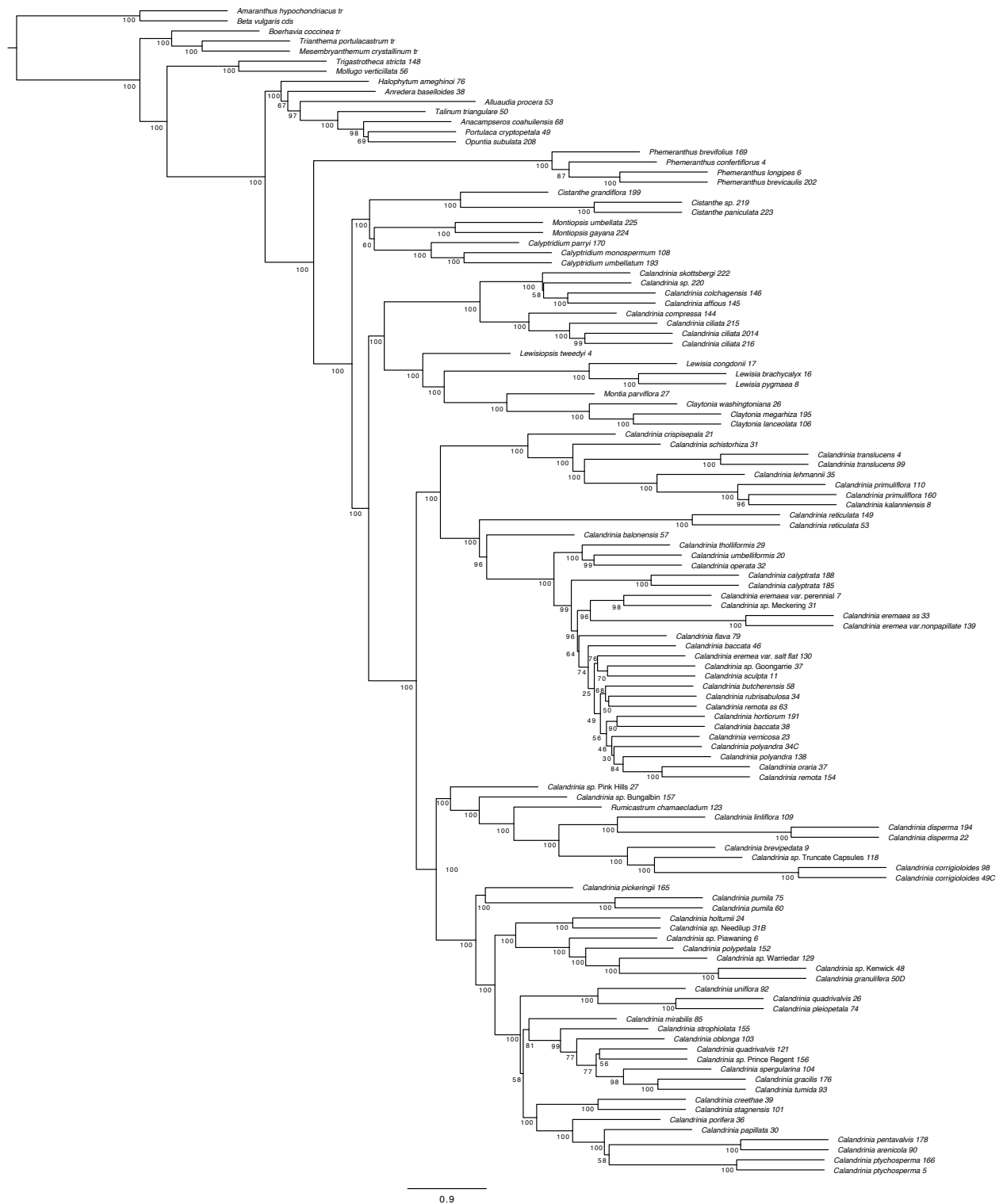
Appendix S3. Density plots of Australian *Calandrinia* internal and terminal branch lengths where all sites present in <10% individuals (denoted as “10”) have been removed and all sites present in <50% of the individuals (denoted as “50”) have been removed for (A) g2 loci matrix, (B) g7 loci matrix, (C) i40 loci matrix, and (D) i98 loci matrix.



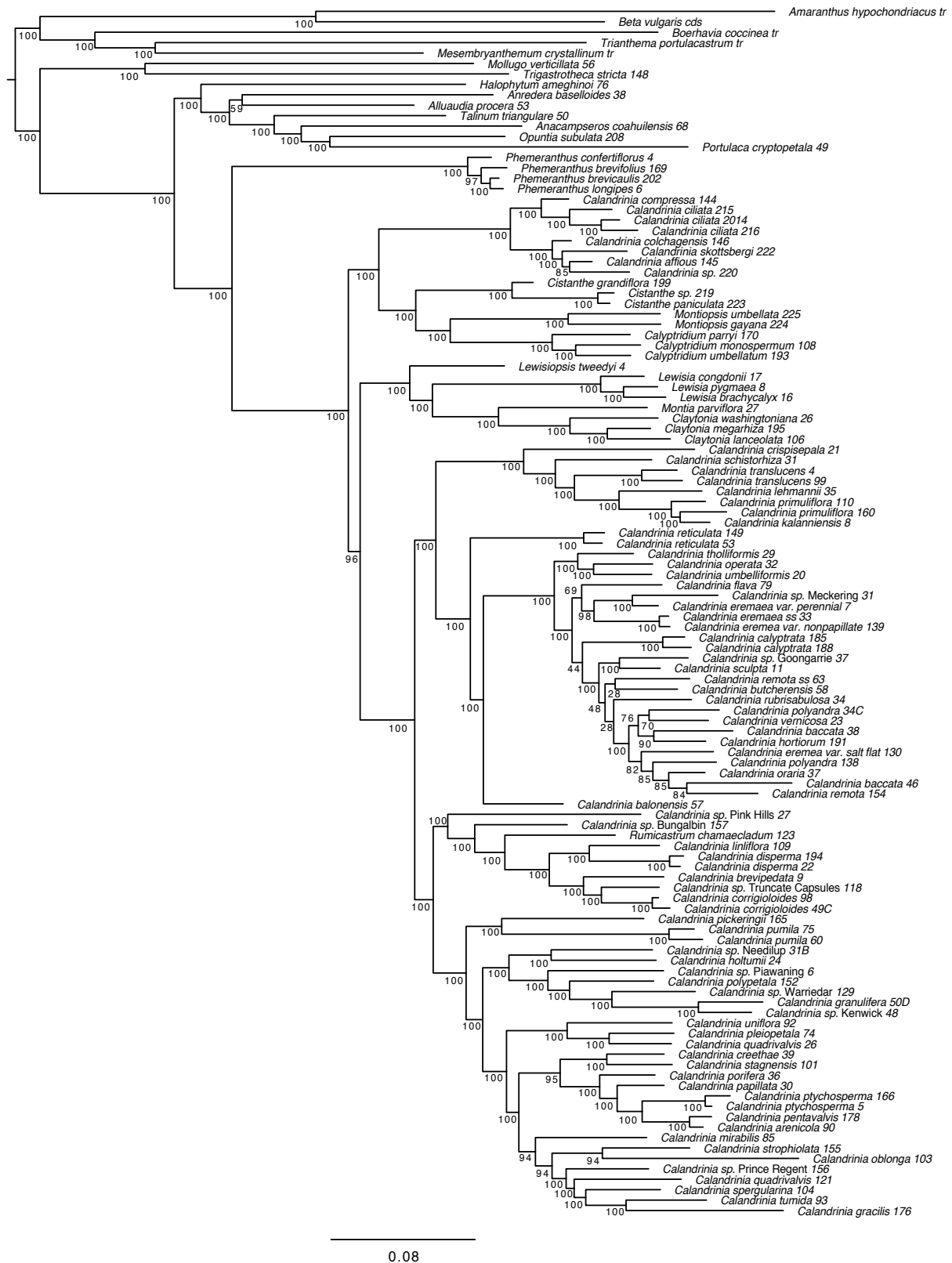
Supplemental appendix S5. RAxML topology from the g7 loci dataset (116 individuals, 116 loci)



Supplemental appendix S7. RAxML topology from the i98 loci dataset (116 individuals, 74 loci)



Supplemental appendix S8. ASTRAL topology from the i40 loci dataset (116 individuals, 162 loci)



Supplemental appendix S9. RAxML topology from the i40 loci dataset (116 individuals, 162 loci)



Supplemental appendix S10. RAxML topology from the g2 loci dataset (116 individuals, 297 loci)

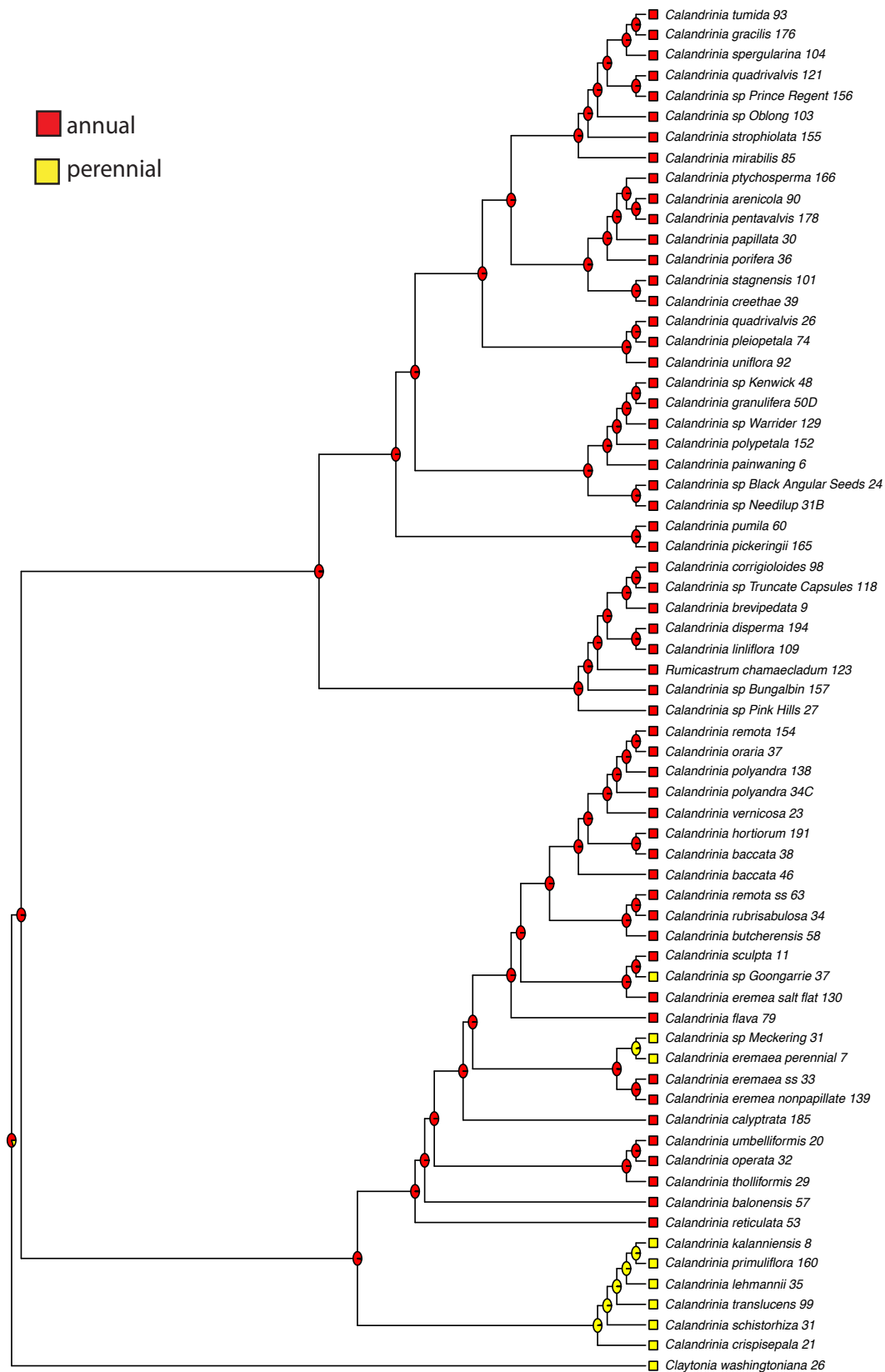
Supplemental appendix S11. Concordance factors from BUCKy analysis for all putative clades with at least 5% genome-wide support

Analysis ⁽¹⁾	Genome-wide CF (95% CI) ⁽²⁾	Clade ⁽³⁾	In primary concordance tree?
All groups	0.380 (0.257,0.510)	CLD 5 + BAS	yes
All groups	0.354 (0.245,0.470)	CLD 4 + CLD 5 + BAS	yes
All groups	0.343 (0.230,0.464)	TUB + PSEUDO	yes
All groups	0.310 (0.209,0.421)	PHE + OUT	yes
All groups	0.279 (0.184,0.384)	CLD 4 + TUB + CLD 5 + BAS + CLD 3 + PSEUDO	yes
All groups	0.267 (0.167,0.380)	MONTIAE + CAL s.s.	yes
All groups	0.226 (0.137,0.329)	CLD 4 + CLD 5 + BAS + CLD 3	yes
All groups	0.219 (0.121,0.334)	MONT + CALYP	yes
All groups	0.184 (0.103,0.280)	MONT + CALYP + CIST	yes
All groups	0.108 (0.047,0.188)	MONT + CALYP + CIST + PHEM + OUT	yes
All groups	0.209 (0.113,0.320)	MONT + CIST	no
All groups	0.175 (0.078,0.290)	CLD 4 + BAS	no
All groups	0.123 (0.043,0.225)	CLD 4 + CLD 5	no
All groups	0.085 (0.023,0.170)	CALYP + CIST	no
All groups	0.081 (0.023,0.161)	MONT + CAL s.s.	no
All groups	0.070 (0.013,0.153)	CLD 3 + PSEUDO	no
All groups	0.064 (0.007,0.150)	TUB + CLD 3 + PSEUDO	no
All groups	0.064 (0.010,0.145)	BAS + CLD 3	no
All groups	0.060 (0.009,0.139)	CLD 5 + CLD 3	no
All groups	0.054 (0.005,0.132)	TUB + CLD 3	no
All groups	0.053 (0.005,0.130)	CLD 4 + CLD 3	no
All groups	0.050 (0.012,0.110)	MONT + PHEM + CALYP + MONTIAE + CAL s.s. + OUT	no
All groups	0.050 (0.005,0.123)	CLD 4 + PSEUDO	no

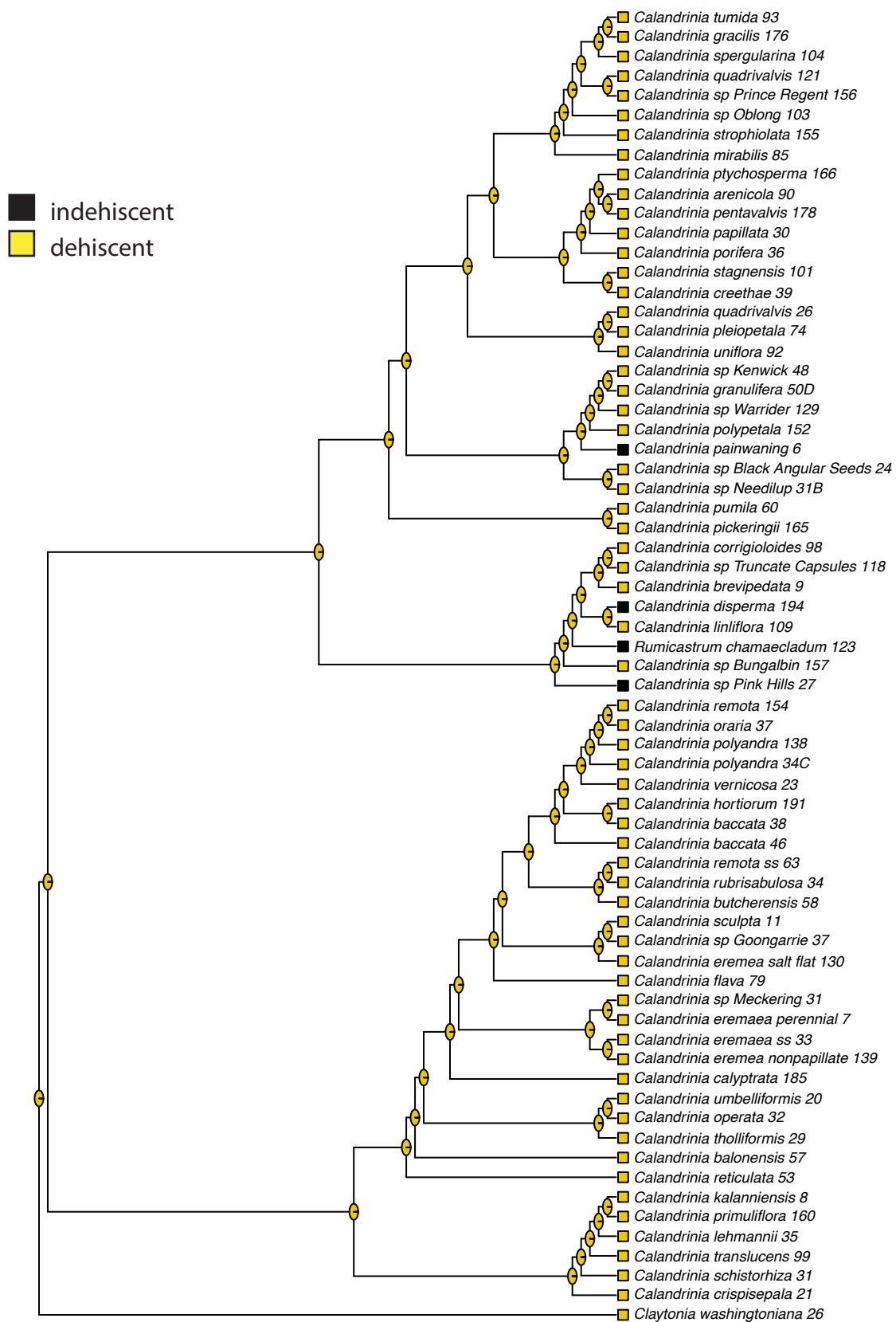
(1): All groups – analysis with all groups, which includes the 6 clades of Australian Calandrinia, Montiae, Calyptidium, Montiopsis, Cistanthe, Phemeranthus, Calandrinia s.s. and Outgroups.

(2): Genome wide concordance factor with 95% credibility interval

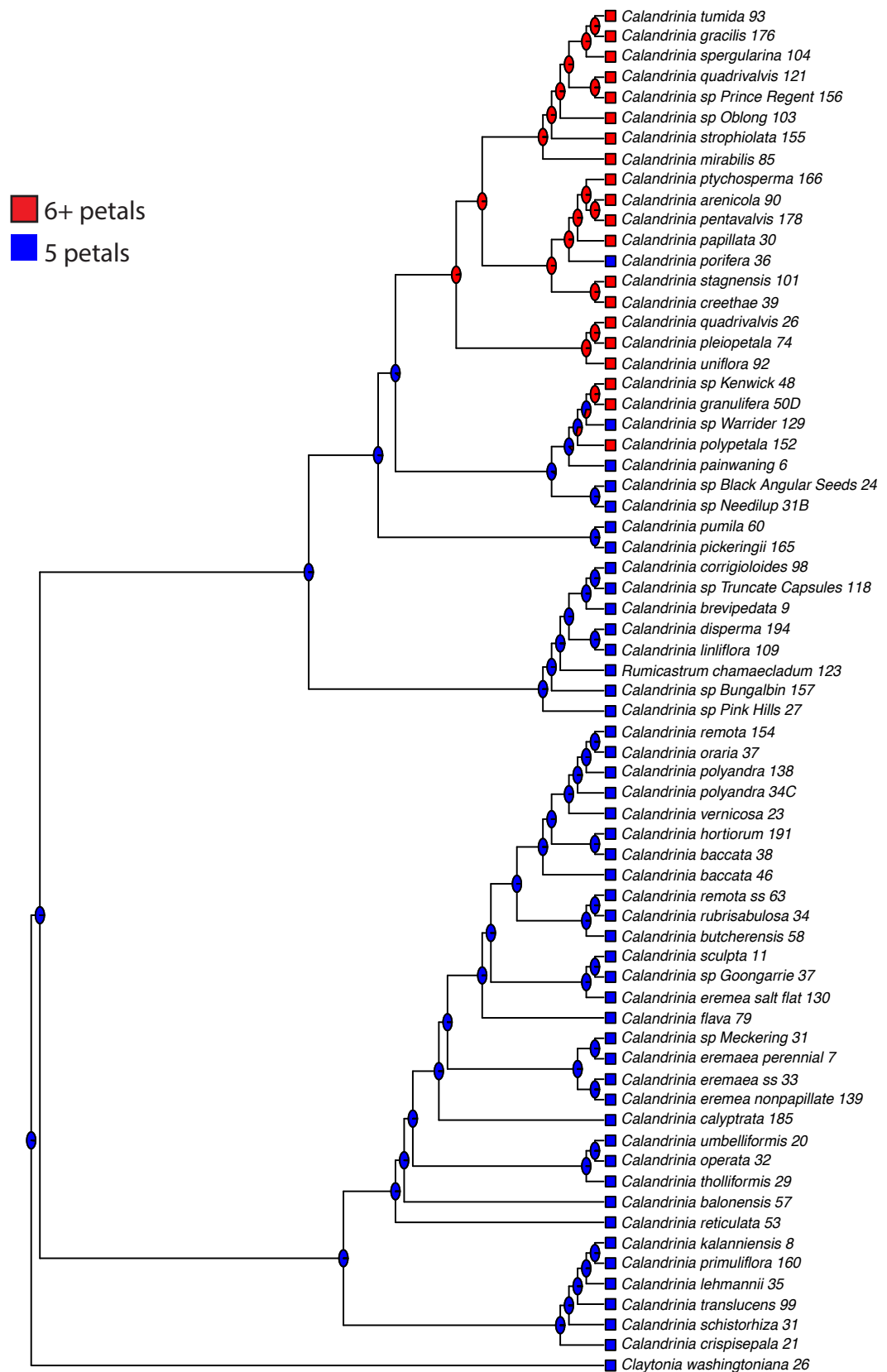
(3): Clades are abbreviated by the first three to four letters of their group except for Montiae, which is written out



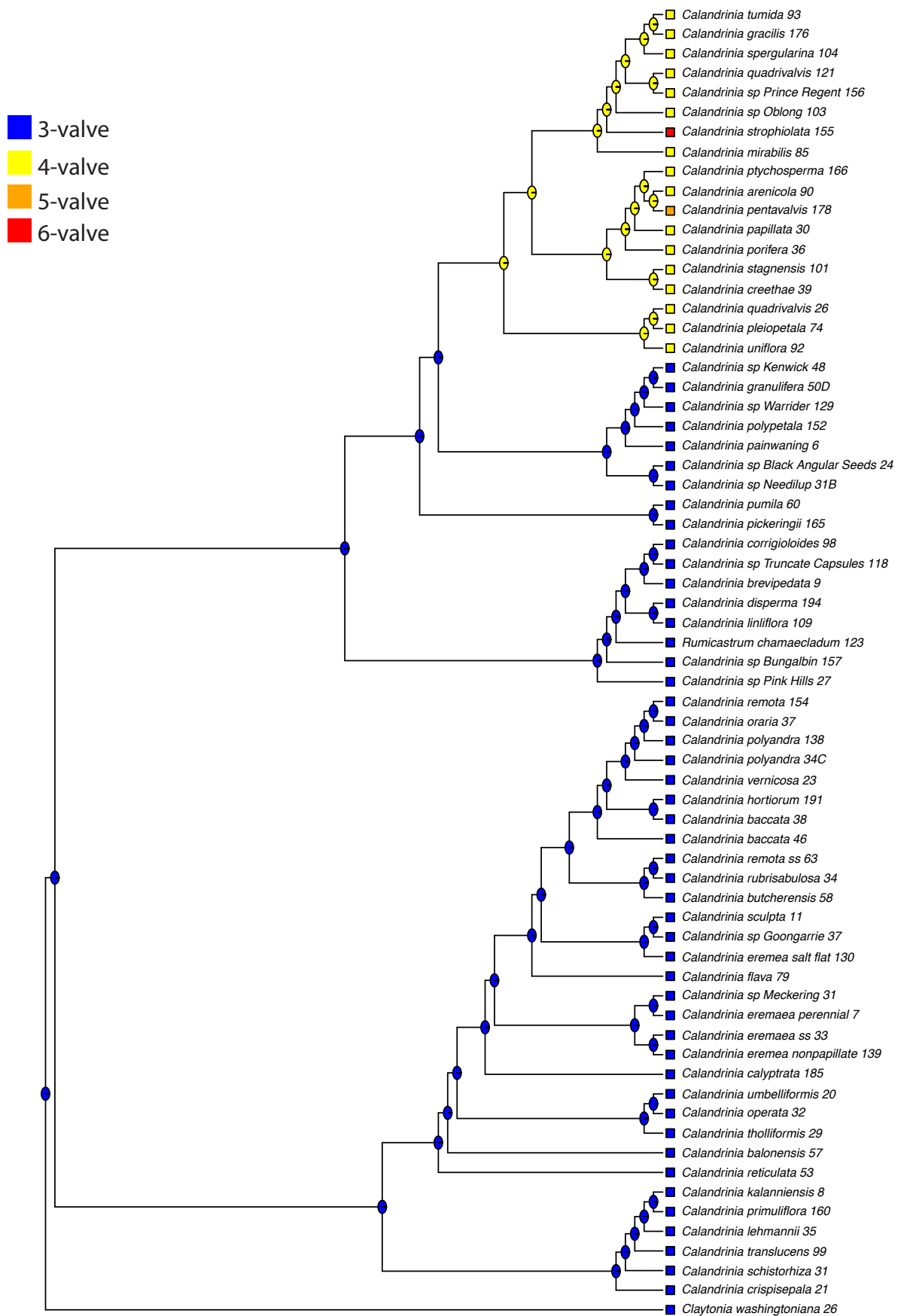
Supplemental appendix S12. Ancestral character state reconstruction of life history



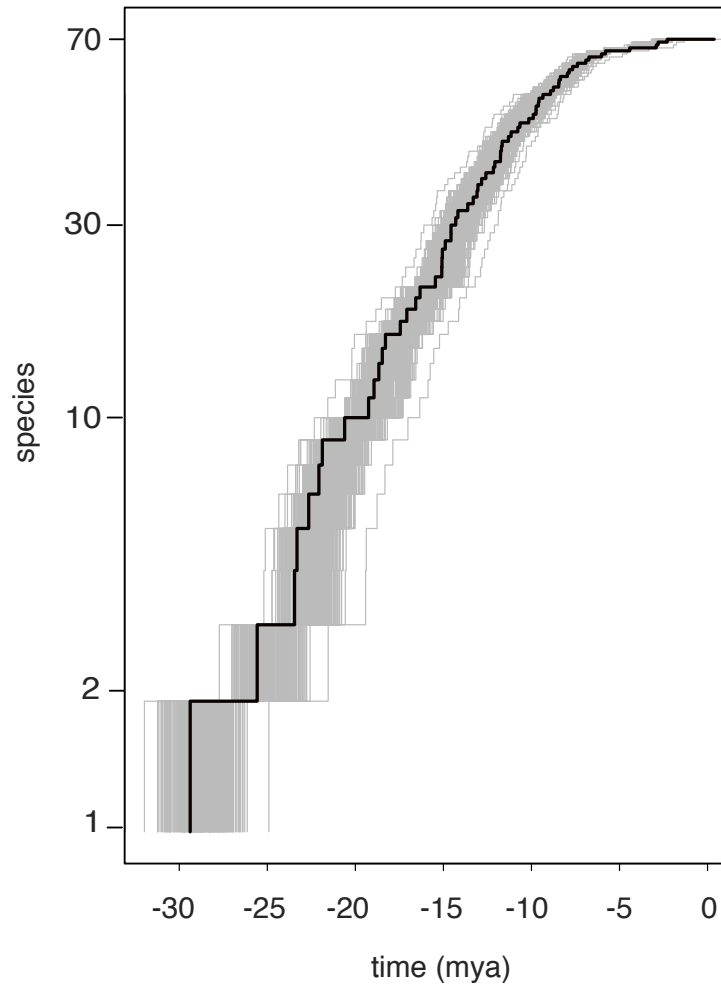
Supplemental appendix S13. Ancestral character state reconstruction of fruit type



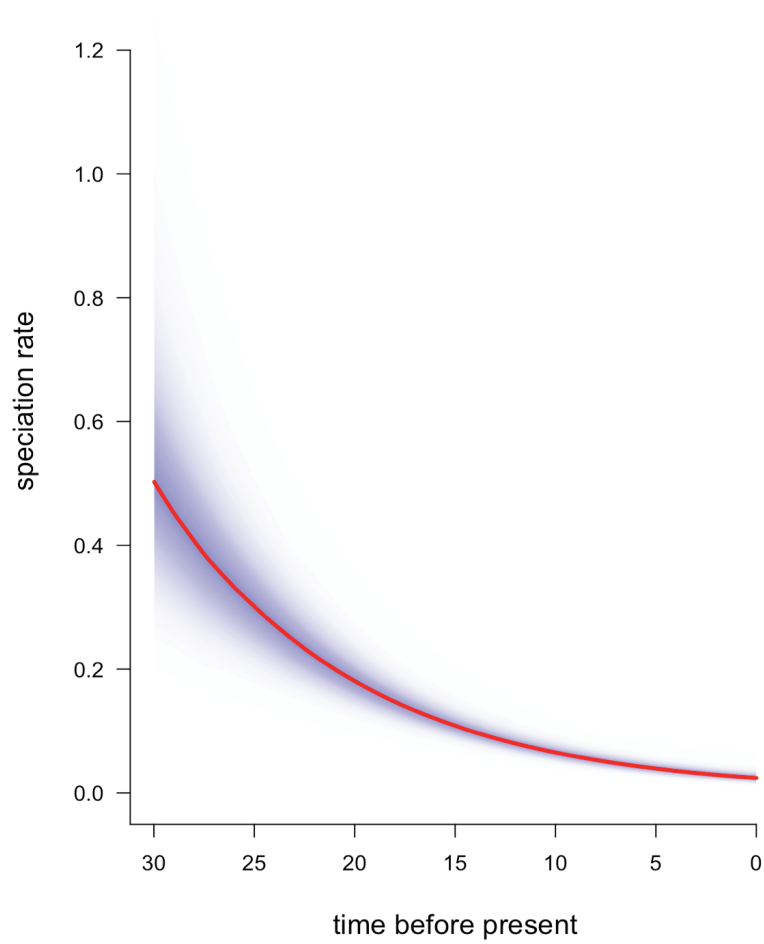
Supplemental appendix S14. Ancestral character state reconstruction of petal number



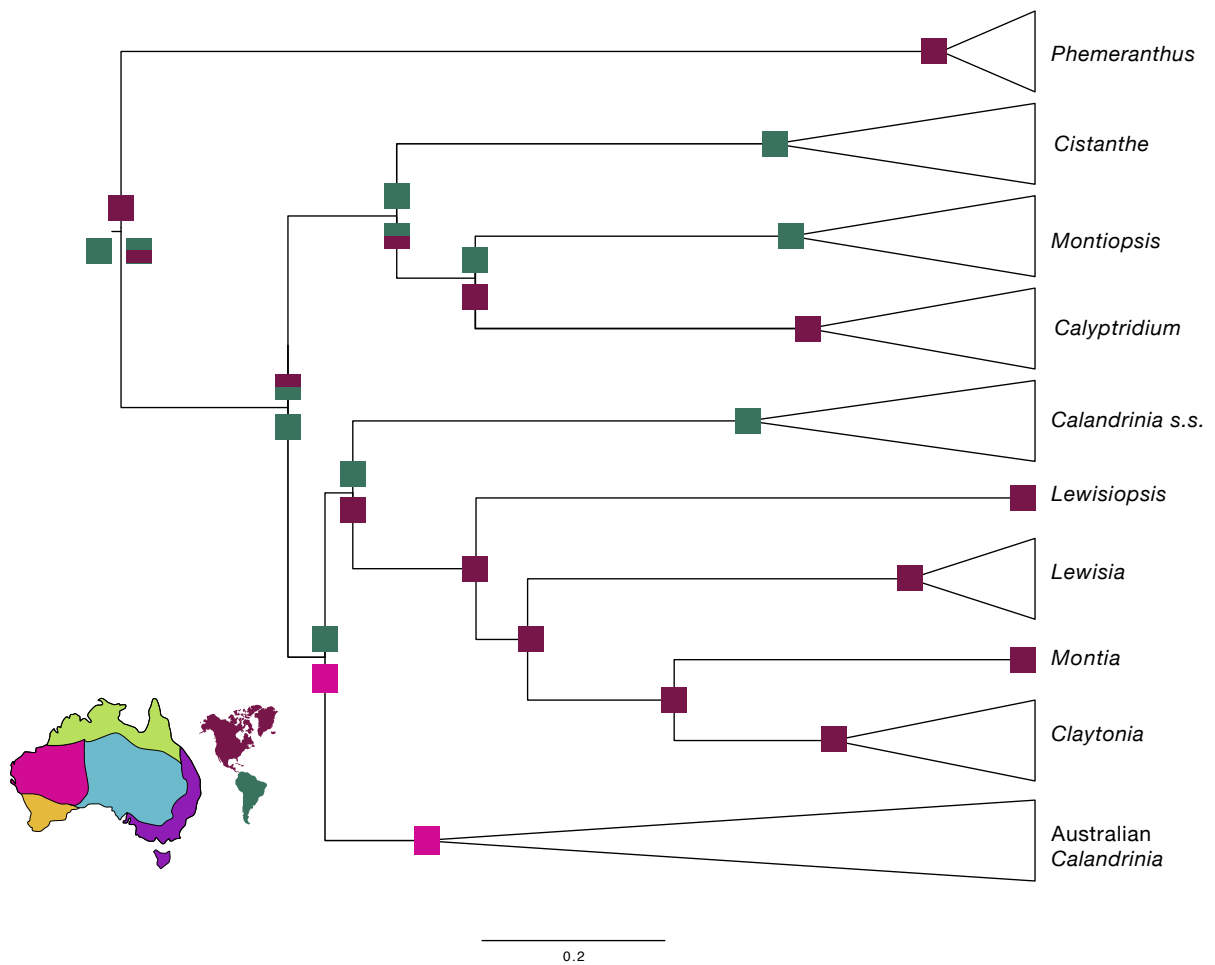
Supplemental appendix S15. Ancestral character state reconstruction of valve number



Supplemental appendix S16. Australian *Calandrinia* lineage through time plot. The black line indicates the log (number of species) over time for Australian *Calandrinia* based on the maximum clade credibility BEAST tree. The gray lines show results for 500 posterior distribution BEAST trees.



Supplemental appendix S17. Bayesian analysis of macro-evolutionary mixtures speciation through time



Supplementa appendix S18. Collapsed Montiaceae time-calibrated phylogeny with biogeographic reconstructions, i.e. North America (red), South America (green), and Australia (pink) at each node. Reconstructed nodes have a greater than 50% relative probability.

**Chapter 3: The evolution of a spectrum of CAM phenotypes in Australian
Calandrinia (Montiaceae)**

Lillian P Hancock, Joseph AM Holtum, and Erika J Edwards

In Review: Hancock LP, Holtum JAM, and EJ Edwards. The evolution of a spectrum of CAM phenotypes in Australian *Calandrinia* (Montiaceae). New Phytologist.

ABSTRACT

- Our understanding of CAM evolution is based largely on the characterization of photosynthetic mode using the carbon isotope ratio ($\delta^{13}\text{C}$) and the mapping of this data onto large, often poorly resolved phylogenies. Although useful, $\delta^{13}\text{C}$ cannot detect low-level and facultative CAM phenotypes. It is hypothesized that these “C₃-CAM” phenotypes represent early steps in the evolution of full CAM.
- To investigate the evolution of C₃-CAM phenotypes, we cultivated 22 species of Australian *Calandrinia* and conducted a drought experiment. We used physiological measurements and $\delta^{13}\text{C}$ to characterize photosynthetic mode and characterized the climatic envelopes of all species in their native ranges.
- Australian *Calandrinia* is a lineage of plants capable of employing C₃, low-level and facultative CAM, though no full CAM behavior was detected. Results highlight the limitations of $\delta^{13}\text{C}$ surveys in detecting C₃-CAM phenotypes. Ancestral character state reconstructions demonstrate considerable evolutionary lability among C₃-CAM phenotypes and indicate several evolutionary losses of CAM capability.
- Our analyses suggest a model of CAM evolution that includes substantial evolutionary lability among C₃-CAM and C₃ phenotypes, with no clear ecological distinction between low-level and facultative CAM species. We suggest that the spectrum of C₃-CAM phenotypes may not yet be optimally circumscribed into discrete character states.

KEY WORDS: Australian *Calandrinia*, C₃-CAM phenotypes, stable carbon isotope ratio, CAM evolution, CAM photosynthesis, phylogenetics

INTRODUCTION

Understanding how complex adaptations evolve is an enduring problem in evolutionary biology. In plants, Crassulacean Acid Metabolism (CAM) photosynthesis is a complex trait that improves water-use efficiency and has evolved multiple times, across a diversity of lineages, from the more common C₃ photosynthetic pathway (Winter & Smith, 1996; Silvera *et al.*, 2010; Edwards & Ogburn, 2012). CAM differs from the C₃ pathway in the primary step of CO₂ assimilation: in the C₃ pathway CO₂ fixation is catalyzed during daylight hours by ribulose 1,5-bisphosphate carboxylase-oxygenase (Rubisco), however in the CAM pathway CO₂ is first fixed at night using PEP carboxylase (PEPC) (Osmond, 1978). The carbon, which is stored overnight as malic acid, is then decarboxylated and the CO₂ re-fixed and converted to carbohydrate via the C₃ pathway during the following light period (Holtum & Osmond, 1981; Ting, 1985; Holtum *et al.*, 2005). Depending upon the light intensity, the catabolism of malic acid is usually complete by midday. Characteristics of tissues with CAM therefore include nocturnal acidification of photosynthetic tissues and nocturnal CO₂ fixation.

It is nonetheless difficult to describe a definitive CAM phenotype because, in well-watered plants that express CAM, CAM is rarely the sole contributor to carbon gain. It almost always co-exists with a component of C₃ photosynthesis (or, rarely, C₄ photosynthesis) that is expressed following the depletion of the malic acid pool in the light (Koch & Kennedy, 1982; Winter *et al.*, 2008; Holtum *et al.*, 2017a). Depending upon species, the proportion of carbon obtained during the daylight with the C₃ pathway

may contribute anywhere between close to zero percent and 100 percent of the 24 hour carbon uptake (Winter & Holtum, 2002; 2014). Plants with CAM therefore include a spectrum of “CAM+C₃” and “C₃+CAM” phenotypes.

Categorization of the CAM phenotype is further complicated because the expression of CAM is also phenotypically plastic, and often varies in response to environmental stress (Griffiths, 1989; Cushman, 2001; Winter *et al.*, 2015; Holtum *et al.*, 2017b). In plants in which CAM expression is regulated developmentally (i.e. constitutive CAM), existing nocturnal acidification may be upregulated in response to stress. Alternatively, in otherwise C₃ plants CAM may be induced when stressors such as drought or salinity are imposed. If the stress-related increase in acidification is reversed when the stress is removed, the upregulated or induced CAM-component is said to be facultative (Winter *et al.*, 2008).

Although the expression of C₃-CAM is clearly a continuous trait (Zotz, 2002; Winter *et al.*, 2015), it may also be true that there are distinct categories of CAM behavior. For example, there might be significant genetic, ecological or functional repercussions of engaging in constitutive vs. purely facultative CAM, quite independent of the level of CAM expression. For example, the constitutive CAM category includes plants where nocturnal CO₂ fixation with the CAM pathway contributes to the majority of carbon gain (e.g., cacti and agaves, what we refer to as full CAM), but at the same time, includes plants that primarily fix carbon with the C₃ pathway, but *always* engage in a small component of CAM regardless of environmental stressors (what we refer to as low-level CAM). Constitutive CAM can be up- or-down regulated but is always present. Purely facultative CAM refers to the induction of CAM in response to environmental

conditions, such as drought or salinity. When the environmental stressor is removed (e.g., with re-watering) the plant reverts to C₃ photosynthesis (Winter & Holtum, 2014).

Similar to the variation between low-level and full constitutive CAM species, the amount of CAM that a facultative CAM plant performs when stressed varies considerably across species (Holtum & Winter, 1999; Cushman & Borland, 2002; Silvera *et al.*, 2005; Lüttge, 2006; Winter & Holtum, 2006; Winter *et al.*, 2008; Winter & Holtum, 2014). In this study, we analyze the evolution of CAM behavior as both a qualitative (constitutive vs. facultative) and a quantitative (amount of malic acid generated) character.

The discovery and characterization of potential C₃-CAM phenotypes, particularly in cases when the expression of CAM is facultative, generally requires measurements of day-night changes in titratable acidity and 24 hour gas-exchange; a relatively laborious process during which plant responses are examined under watered, water-limited, and re-watered conditions. The gas-exchange measurements capture patterns of CO₂ exchange and stomatal conductance, while nocturnal acidification quantifies β -carboxylation and the overnight storage of malic acid. The measurement of stable carbon isotope ratios, while useful for screening for more pronounced CAM (Winter & Holtum, 2002), is not suitable for confidently detecting and differentiating C₃-CAM intermediates (Sternberg *et al.*, 1984; Winter *et al.*, 2015). Determination of $\delta^{13}\text{C}$ may quantify the proportion of plant carbon contributed by Rubisco and PEPC during the life of the tissue, but when the long-term contribution of CAM to plant carbon gain is small, such as it is in constitutive low-level CAM plants and in facultative CAM plants, $\delta^{13}\text{C}$ values may not be intermediate between those expected for C₃ and CAM plants, and can be

indistinguishable from C₃-type values (Winter & Holtum, 2002; Silvera *et al.*, 2005; Winter *et al.*, 2015).

It has been hypothesized that the C₃-CAM phenotypes in which C₃ photosynthesis contributes the majority of carbon to the plant may represent early steps in the evolution of full CAM photosynthesis (Guralnick & Jackson, 2001; Silvera *et al.*, 2010; Edwards & Ogburn, 2012; Hancock & Edwards, 2014). This hypothesis is based upon three assumptions: 1) the CAM phenotype is too complex to appear *de novo*, so it must be assembled incrementally, 2) C₃-CAM intermediates are found in close phylogenetic association with constitutive CAM species (Guralnick *et al.*, 2007; Silvera *et al.*, 2010; Crayn *et al.*, 2015; Goolsby *et al.*, 2018), and 3) in the evolution of C₄ photosynthesis - a related photosynthetic pathway - C₃-C₄ intermediate phenotypes represent transitional states along the C₄ evolutionary trajectory (Christin *et al.*, 2011; Sage *et al.*, 2012). The characterization of C₃-C₄ intermediates, combined with densely sampled and well-resolved phylogenies, has provided significant insight into the evolutionary assembly of the C₄ syndrome in the Molluginaceae (Christin *et al.*, 2011), Neurachninae (Christin *et al.*, 2012), Portulacaceae (Ocampo & Columbus, 2012), Chenopodiaceae (Kadereit *et al.*, 2003; 2010), Poaceae (Christin *et al.*, 2013), and *Flaveria* (McKown *et al.*, 2005). In contrast to C₄, the evolutionary dynamics of CAM are not well understood.

In efforts to investigate the expression and evolution of these C₃-CAM phenotypes, we examined the capacity for nocturnal acidification in 22 species of Australian *Calandrinia* (Montiaceae) during the course of a drought experiment. We selected this lineage because 1) a comprehensive and well-resolved phylogeny is available for the group (Hancock *et al.*, in revision), 2) the lineage contains only ca. 70

species and complete physiological testing is ultimately feasible, and 3) previous physiological and isotopic research had already demonstrated variation in the expression of C₃ and CAM photosynthesis across the group (Winter *et al.*, 2008; Winter & Holtum, 2011; Holtum *et al.*, 2017b). Furthermore, Australian *Calandrinia* is widely distributed across Australia – from the central arid regions of Northern Territory and Western Australia to the cool, temperate regions of Victoria and Tasmania, to the monsoon-dominated tropics of coastal Australia– making it an ideal lineage in which to investigate the relationship between the expression of CAM and the environment in which these plants live.

MATERIALS AND METHODS

Study species, growth conditions and experimental parameters: Species included in the experiment represent the phylogenetic diversity of the Australian *Calandrinia* lineage (Fig. 1). Plant material and seeds were obtained from wild plants, plants in cultivation (garden), from herbarium specimens, and herbarium seed collections (Table 1). Seeds were germinated in 24-cell seedling trays, with Osmocote seed raising mix during the winter of 2015 (July-August). Roughly three weeks after germination, seedlings were transplanted to 1L plastic pots, with Osmocote potting mixture. All plants were grown under semi-natural conditions in a closed-sided shade house at James Cook University (Queensland, Australia). Daily temperature within the shade house ranged from 36.2°C to 23°C and light availability was close to 50% ambient. All plants were watered daily until the experiment commenced.

The drying down, re-watering experiment was staggered over the months of September and October 2015. Staggering the experiment was necessary 1) as species germinated and matured at different rates, 2) to ensure that all plants were in the same location within the shade house over the duration of the experiment, and 3) to facilitate rapid sampling of time-sensitive biochemistry (i.e. malic acid degradation with sunrise). Because of variation in germination success, we sampled a range of 3 to 10 individuals per species. When we had greater than seven individuals per species, we left 3 of the total plants per species watered (control) over the duration of the experiment. For the experimental plants, watering was withheld on day 2 and recommenced on day ~ 12. Leaves were harvested from control and experimental plants at dusk on day 1 and dawn on day 2 (well-watered), at dusk on day 11 and dawn on day 12 (drought), and dusk on day 20 and dawn on day 21 (re-watered).

Titrateable acidity: Harvested leaves were immediately measured for fresh weight and were then frozen in liquid nitrogen. Organic acids were extracted by boiling leaf tissue in ~ 80ml of 50% ethanol in distilled water. Once half of the total volume boiled off, an additional ~ 40mL of distilled water was added to each beaker. Beakers were removed from the hotplates when the final volume reached 40mL. The H^+ pools were determined by measuring the volume of 5 mM NaOH required to neutralize each sample to pH 6.5. To calculate nocturnal increase in tissue acidity (ΔH^+) we subtracted the titrateable acidity (H^+) at the end of the light period from the titrateable acidity at the end of the dark period (H^+). ΔH^+ was then averaged for each species at each time point (i.e. watered, droughted, and re-watered). A student's t-test was performed to determine if ΔH^+ values were significantly different and greater than zero, as well as to check for

significant differences when plants were watered, droughted, and re-watered. A significant ΔH^+ value indicated CAM. A species was characterized as facultative CAM if there ΔH^+ values were not significant when watered but significant when droughted. Low-level CAM plants showed significant ΔH^+ regardless of experimental conditions.

Net CO₂ exchange: Gas exchange was monitored for four Australian *Calandrinia* species: *C. balonensis*, *C. mirabilis*, *C. pickeringii*, and *C. ptychosperma*. Plants were cultivated in 1L plastic pots. During the experiment, performed September – November 2015 (concurrent with the titration experiment), watering was withheld on day 2 and recommenced on ~ day 12. Gas exchange on one leaf was measured continuously throughout the watered, drought, re-watered experiment (~22 days) using a Li-6400 portable gas analyzer with values recorded every 10 minutes. The value recorded every 10 minutes was the average of 600 measurements, +/- standard deviation, made over the course of that time period. Mature leaf tissue was placed into the 6400-08 standard clear bottom chamber. The leaf was exposed to ~12h, ~ 500 μ mol photons m⁻² s⁻¹, 28°C/ 22°C light/ dark cycles.

Stable isotope analysis: $\delta^{13}\text{C}$ values of dried, finely-ground leaves collected from whole plants were measured using a Finnigan Delta plusXL mass spectrometer. For the majority of species, $\delta^{13}\text{C}$ values were measured for two to five individuals and the average (+/- SD) value was calculated (Table 1). To differentiate phenotypes based on $\delta^{13}\text{C}$ values we used figures and values from Winter and Holtum (2002) and Winter *et al.* (2015). Full CAM species or CAM plants have $\delta^{13}\text{C}$ values that are c. -20‰ and less negative while C₃ species have $\delta^{13}\text{C}$ values that are c. -22‰ and more negative. Species that are C₃-CAM have $\delta^{13}\text{C}$ values that can fall anywhere within the $\delta^{13}\text{C}$ C₃ ‘cluster’ but

often (based on Figure 7 from Winter & Holtum, 2002 and Figure 1 from Winter *et al.*, 2015) present $\delta^{13}\text{C}$ values between c. -24‰ and c. -20‰. Based on these data and the observation that -22‰ is a natural break in full CAM and C_3 $\delta^{13}\text{C}$ values displayed in Figure 7 (Winter & Holtum, 2002), we used -22‰ as our C_3 -CAM cutoff.

Ancestral character state reconstruction: For analyses of character evolution, we used the Australian *Calandrinia* phylogeny from Hancock *et al.*, (in revision). To reconstruct ancestral character states at all internal nodes for our qualitative character states, inferred from both the carbon isotope ratio and physiological measurements, we used ML reconstruction in the APE package as implemented in the R Project for Statistical Computing (Paradis *et al.*, 2004). For both character state reconstructions, we compared an evolutionary model with equal transition rates between states (ER) with models that allowed the transition rates to vary (ARD and SYM). We chose the model with the lowest AIC value when reconstructing ancestral character states.

Since the C_3 -CAM phenotypic spectrum can also be interpreted as a continuous trait, with ΔH^+ as the measure of CAM activity, we plotted ΔH^+ when plants were watered and droughted across the tips of the phylogeny. We used Phytools (Revell, 2012) as implemented in R to reconstruct ancestral character states under each condition.

Locality and climate data: 13, 590 georeferenced records for Australian *Calandrinia* were downloaded from the Australian Virtual Herbarium (<http://avh.chah.org.au/>) on 10 February 2016. We followed a number of steps to clean these records (i.e., records with identical coordinates for a given taxon), records with latitude and longitude coordinates of 0, 0, and records that matched the lat./ long. of herbaria and botanical gardens (Edwards *et al.*, 2015);

https://github.com/ejedwards/reanalysis_zanne2014). After cleaning we were left with 11, 087 unique records. The remaining localities were imported into ArcGIS 4.0 and overlain on a map of Australia, and any records that conflicted with country land boundaries (e.g. were in the ocean) were removed.

Climate layers (ASCII format) for mean annual/ semi-annual temperature, mean annual/ semi-annual precipitation (MAP), precipitation variability, evapotranspiration, and relative humidity were downloaded from the Australian Government Bureau of Meteorology (<http://Calandrinia.bom.gov.au/>). Layers were converted into Rasters using ArcMAP 4.0.1 and clipped to match the Australian continent, which included a 35km buffer in efforts to capture surrounding islands. Finally, an attributes table was constructed containing all georeferenced localities, with corresponding species name and the climate data.

Species means for each climate variable were calculated for further comparative analysis. Principal components analysis (Fig. S1) showed that temperature and precipitation explained most of the variation in the data, and since all Australian *Calandrinia* (even perennials) are seasonal in their growth, we focused our analyses on mean growing season temperature (MGST) and mean growing season precipitation (MGSP). Mean growing season temperature and precipitation were calculated for each species based on species descriptions, distribution maps from the literature (Obbens, 2006; 2011; Tahir & Carolin, 2011; Obbens, 2012; West & Chinnock, 2013), and the Australian Virtual herbarium. To visualize any relationship between photosynthesis (i.e. C₃ or C₃-CAM) and climate, we plotted mean values ($\pm$ SD) in an X-Y coordinate space with MGST on the X-axis and MGSP on the Y-axis. A MANOVA and Tukey Test were

run to determine if the phenotypes (C₃, low-level, and facultative CAM) occupy significantly distinct temperature and precipitation climatic space.

To test for these trait correlations within a phylogenetic context, we used the *brunch* function in the R package Caper (Orme, 2013). The brunch function performs phylogenetically independent contrast (PIC) correlations between discrete and continuous traits by assessing contrasts of continuous variables on nodes of the tree where sister taxa have contrasting values for the discrete trait. To estimate the evolutionary correlation between continuous characters ΔH^+ and MGST and MGSP using PIC, we used the *pic* function in the R package Phytools (Revell, 2012).

RESULTS

Titrateable acidity measurements

Watered: Measurements of nocturnal acid accumulation (ΔH^+) were used as an indicator of the presence or absence of CAM activity in leaves (Table 1; Figure 2a, 2b). The ΔH^+ values ranged from 0 to 13.7 $\mu\text{mol H}^+$ per unit fresh leaf mass (g^{-1} fwt) when plants were well watered. Ten of the 22 *Calandrinia* species exhibited significant ($P < 0.05$) differences between evening and morning titrateable acidities (Table 1; Fig. 2). The remaining 12 species did not have significant ΔH^+ values.

Drought: After ~ 12 days of drought there was an overall increase (Fig. 2a) in titrateable acidity across the majority of species (17 of 22), and nocturnal acid accumulation ranged between 0 and 82.2 $\mu\text{mol H}^+$ g^{-1} fwt (Table 1; Fig. 2). Of the ten species that showed significant ΔH^+ when watered, six of them showed a significant increase in titrateable acidities (g^{-1} fwt) when droughted: *C. crispisepala*, *C. eremaea*, *C.*

flava, *C. kalannensis*, *C. pleiopetala*, and *C. ptychosperma*. The other four species, *C. balonensis*, *C. pickeringii*, *C. pumila*, and *C. reticulata*, showed no significant difference in ΔH^+ when water-limited compared to well-watered.

Nine of the 12 species that had non-significant ΔH^+ when watered, had significant ΔH^+ when water-limited (i.e. exhibited facultative CAM). These species include *C. calyptrata*, *C. creethae*, *C. mirabilis*, *C. monosperma*, *C. papillata*, *C. uniflora*, *C. quadrivalvis*, *C. spergularina*, and *C. volubilis*. ΔH^+ ranged between 7.6 and 82.2 $\mu\text{mol g}^{-1}$ fwt. Nocturnal acid accumulation was not detected in three species of *Calandrinia* (*C. arenicola*, *C. granulifera*, and *C. tumida*) either when the plants were watered or when water-limited (Fig. 2a, b).

Re-watered: After re-watering, nocturnal acid accumulation generally declined (Table 1, Fig. 2b). Of the ten species with significant ΔH^+ values when watered, all went back to low-level CAM except for *C. balonensis*, *C. crispisepala*, and *C. pumila*. These three species no longer significantly accumulated malic acid overnight after re-watering. In the case of *C. balonensis* and *C. pumila* it is likely that these droughted, re-watered individuals were no longer healthy, as the leaves had lost their turgidity and appeared to be senescing.

Of the nine species that showed a significant increase in titratable acidity (facultative CAM) with drought stress but no CAM activity when watered, seven of them showed no appreciable ($P>0.05$) nocturnal acid accumulation when re-watered (Fig. 2b). The other two facultative CAM species, *C. creethae*, and *C. volubilis*, continued accumulating H^+ at night when re-watered (Table 1); it is possible that our sampling

protocol – titratable acidity measurements made on day 20 – simply did not capture the reversal of CAM in these two species.

Gas exchange: Continuous whole-leaf gas exchange measurements on four species of Australian *Calandrinia* over the course experiment demonstrated net CO₂ gas-exchange patterns indicative of C₃, CAM-cycling (a.k.a. the re-fixation of respiratory CO₂), and facultative CAM photosynthesis (Fig. 3a-d). All four species engaged in daytime CO₂ assimilation (C₃ photosynthesis) when watered. With the induction of drought, there was a precipitous decline in daytime net CO₂ fixation and a general decrease in the amount of CO₂ lost through dark respiration, a product of both less daytime photosynthesis and the re-fixation of respiratory CO₂. After ~5 days of water-limitation all of the species showed net nocturnal CO₂ uptake, although the amount varied considerably by species (Fig. 3a-d), and no species switched to solely nocturnal CO₂ uptake. With re-watering, all four species returned to primarily daytime CO₂ uptake (C₃ photosynthesis).

Photosynthetic types and ancestral character state reconstruction: Species were assigned to discrete photosynthetic categories based on both physiological measurements (titratable acidity and gas-exchange, when available) and stable carbon isotope content ($\delta^{13}\text{C}$). On the basis of physiological measurements, species were characterized as either C₃, constitutive low-level CAM, or facultative CAM (Fig. 4). Species phenotyped as C₃ did not have significant ΔH^+ values when watered or water-limited and included three species, *C. arenicola*, *C. granulifera*, and *C. tumida*. 10 species were scored as constitutive low-level CAM: *C. balonensis*, *C. crispisepala*, *C. eremaea*, *C. flava*, *C. kalannensis*, *C. pickeringii*, *C. pleiopetala*, *C. ptychosperma*, *C. pumila*, and *C.*

reticulata. Species characterized as facultative CAM showed an induction of CAM activity only when water-limited (i.e. significant ΔH^+ values) and, in most cases, significant reduction of CAM activity (i.e. non-significant ΔH^+ values) when re-watering resumed. Based on these criteria, nine of the 22 species were scored as facultative CAM: *C. calyptrata*, *C. creethae*, *C. mirabilis*, *C. sp.* Pink Hills, *C. papillata*, *C. uniflora*, *C. quadrivalvis*, *C. spergularina*, and *C. volubilis*. We also used $\delta^{13}\text{C}$ values to score species as either C_3 or $\text{C}_3\text{-CAM}$; $\delta^{13}\text{C}$ cannot differentiate between constitutive low-level CAM and facultative CAM. *C. flava* (-22‰) was the only species with a minimum $\delta^{13}\text{C}$ value equal to or higher than -22‰. The remaining 21 species were categorized as C_3 .

Not surprisingly, ancestral character state reconstructions of photosynthetic mode differed significantly whether using $\delta^{13}\text{C}$ or physiological measurements (Fig. 4). Based on $\delta^{13}\text{C}$ values, the ancestral character state of Australian *Calandrinia* (under the slightly favored ARD transition model) is C_3 and there has been one transition to $\text{C}_3\text{-CAM}$ (*C. flava*). When using physiological methods, the ancestral state under an equal rates model ($\text{AIC} > 2$) is reconstructed as low-level constitutive CAM (scaled likelihood=0.71), with multiple shifts from constitutive low-level CAM to facultative CAM. There are also shifts back from facultative CAM to low-level constitutive CAM, as seen in *C. pleiopetala* and *C. ptychosperma* (Fig. 4). Analyses indicate multiple evolutionary reversals from $\text{C}_3\text{-CAM}$ back to C_3 photosynthesis (i.e. *C. tumida*, *C. arenicola*, and *C. granulifera*). The ancestral character state reconstructions of CAM activity as a continuous trait (ΔH^+) tells a similar evolutionary story (Fig. 5): the degree of CAM expression, as measured by the magnitude of ΔH^+ , is evolutionarily labile, with the

highest levels of H^+ mostly clustering in one subclade (lineages Tuberosae + Pseudodianthodiae, Hancock *et al.*, in revision).

CAM types and environment: Australian *Calandrinia* species included in the study span the entire temperature/ precipitation climatic envelope of the lineage, which occupies considerable breadth in climatic space, from localities that experience over 1500 mm of rain to places with long-term averages of only 20mm of rain per year (Fig. 6a, b). Only three species (*C. tumida*, *C. spergularina*, and *C. arenicola*) live in the wet regions of the northern tropics, while most species grow in the dry, hot regions of Australia (lower right quadrant of Fig. 6a, b). Results from the MANOVA and Tukey Test show that precipitation (p-value <0.001), but not temperature (p-value = 0.7), significantly varies between photosynthetic types, with both low-level CAM (p-value = 0.005) and facultative CAM (p-value = 0.023) occupying significantly drier regions than C_3 species. There was no difference in precipitation or temperature between low-level and facultative CAM phenotypes (p-value = 0.63 and 0.41, respectively); however, facultative CAM species occupy a broader climate range than low-level CAM species; low-level CAM is nested within the facultative CAM range (Fig. S1). When we looked at C_3 and C_3 -CAM phenotypes within a phylogenetic framework there was no significant difference in the precipitation and temperature space occupied by C_3 species compared to C_3 -CAM species. Phylogenetically independent contrasts between ΔH^+ and MSGT and ΔH^+ and MSGP were not significant, although there was a stronger relationship between ΔH^+ and MSGP (p-value = 0.11) than ΔH^+ and MSGT (p-value = 0.75).

DISCUSSION

C3 to CAM phenotypic variation in Australian Calandrinia

Australian *Calandrinia* is a lineage of plants capable of employing C₃ and various amounts of CAM photosynthesis. A full CAM phenotype was not detected in any of the species included in the drought experiment; all species use the C₃ pathway as their primary mode of carbon fixation. Although additional physiological investigation of a few candidate species is warranted (i.e., species with more positive $\delta^{13}\text{C}$ values, not presented here) we do not believe, based on current evidence, that full CAM has evolved in Australian *Calandrinia*. This hypothesis is based on both the present findings and previous work in which only low-level and facultative CAM were detected in four additional Australian *Calandrinia* species (Winter & Holtum, 2011; 2014; Holtum *et al.*, 2017b); thus, of the ~70 Australian *Calandrinia* species, three are currently phenotyped as C₃, twenty-three as C₃-CAM, and zero as full CAM. The demonstration of pervasive low-level and facultative CAM in Australian *Calandrinia* adds to a growing body of research uncovering the prevalence of these phenotypes across a diversity of plant lineages, including Agavoideae (Heyduk *et al.*, in press), *Clusia* (Borland *et al.*, 1992; Lüttge, 1999; Holtum *et al.*, 2004), Crassulaceae (Schuber and Kluge, 1981; Castillo, 1996), *Euphorbia* (Herrera, 2013), other Montiaceae (Harris & Martin, 1991; Guralnick & Jackson, 2001), Piperaceae (Ting and Sipes, 1985), Portulacaceae (Koch & Kennedy, 1982; Holtum *et al.*, 2017a), Talinaceae (Herrera, 2009), and Cactaceae (Edwards & Diaz, 2006; Edwards & Donoghue, 2006).

Based on nocturnal acidification measurements, three Australian *Calandrinia* species are phenotyped as C₃, ten as low-level CAM, and nine as facultative CAM, though six low-level CAM also upregulated CAM expression in response to drought. The

coupling of 24-hour, ~ 22-day gas-exchange data (Fig. 3a-d) with titratable acidity measurements (Fig. 2a, b) provides an additional layer of information about the source of CO₂ being fixed. Plants may fix and store both atmospheric and respiratory CO₂ using CAM biochemistry (Ting, 1985). Gas exchange profiles for constitutive low-level CAM species *C. balonensis*, *C. pickeringii* and *C. ptychosperma* indicate that these species show no net CO₂ uptake during the night when well-watered but have significant nocturnal leaf acidification (Fig. 3a, c, d). For this pattern to be observed, plants must be fixing carbon at a rate low enough to simply balance respiratory CO₂ release (Szarek *et al.*, 1973; Flexas *et al.*, 2007). The use of respiratory CO₂ as a substrate for nocturnal CO₂ acid synthesis is commonly observed in C₃-CAM plants and is thought to be a mechanism for carbon conservation, especially during periods of extended drought when little atmospheric CO₂ is being exchanged (Winter *et al.*, 1986; Martin *et al.*, 1988; Harris & Martin, 1991; Guralnick & Gladsky, 2017).

The necessity of experiments for identifying C₃-CAM phenotypes

Characterizing species into discrete photosynthetic phenotypes– and the ancestral state reconstructions they inform – is, unsurprisingly, quite sensitive to the method used to detect CAM. We inferred some CAM activity in only one of the 22 species included in the greenhouse experiment using $\delta^{13}\text{C}$, while titratable acidity measurements revealed CAM activity in 19 species (Fig. 4). Such disparity in signal illustrates the shortcomings of the carbon isotope ratio method to detect C₃-CAM phenotypes and reinforces the necessity of physiological measurements to truly understand the phylogenetic distribution and prevalence of C₃-CAM plants. With that said, there is value to carbon isotope

screens, as they uncover strong CAM species, and thus guide us to fruitful phylogenetic neighborhoods in which to investigate CAM evolution.

Reconstructing the ancestral character state with physiological measurements (i.e. titratable acidity and 24-hour gas exchange) tells a different story about the expression and evolution of CAM in Australian *Calandrinia* (Fig. 4). We show that the ancestral state of Australian *Calandrinia* is likely to be low-level, constitutive CAM, with multiple transitions to a purely facultative state (and several reversions back to low-level CAM). Remarkably, we have documented several clear losses of all CAM activity, with three putative C₃-only species (*C. tumida*, *C. arenicola* and *C. granulifera*) nested within CAM clades. Reversion to C₃ physiology from CAM, particularly full CAM, is likely to be rare. Although it has been reported at least once in the Agavoideae (Heyduk *et al.*, 2016; although see a re-analysis by Heyduck *et al.*, in press) and Bromeliaceae (Crayn *et al.*, 2004) and multiple times in the Orchidaceae (Silvera *et al.*, 2009) and *Clusia* (Vaasen *et al.*, 2002; Gehrig *et al.*, 2003; Silvera *et al.*, 2010), phylogenetic and photosynthetic phenotypic sampling (all of which was done with $\delta^{13}\text{C}$ values) within these studies were insufficient to confidently confirm the reversion of CAM.

Here, we have discovered C₃ species nested well within C₃-CAM lineages, using a well-resolved phylogeny and experimental physiological data, as opposed to isotopic measurements. The multiple reversals to C₃ and the considerable lability detected between C₃, low-level, and are facultative CAM suggests that transitions between these character states are not uncommon, and are evolutionarily accessible. Continuous character state reconstructions of ΔH^+ support this deduction and indicate that species from the Tuberosae clade (*C. crispisepala* and *C. kalanniensis*) and Pseudodanthodiae

clade (*C. reticulata*, *C. balonensis*, *C. flava*, *C. calyptrata*, *C. volubilis* and *C. eremaea*) generally have higher ΔH^+ values regardless of environmental conditions than the rest of Australian *Calandrinia* studied. Both Pseudodanthodiae and Tuberosae (an entirely perennial clade) live in particularly arid regions of western and central Australia (Obbens, 2006; 2011; Obbens *et al.*, 2017) (refer to bottom right corner of Fig. 6a, b) where more CAM activity, whether it is low-level or facultative, would likely be advantageous.

Since we did not detect full CAM in Australian *Calandrinia*, we cannot corroborate or refute hypotheses that these C₃-CAM phenotypes represent likely states along a C₃ to CAM evolutionary trajectory. We do, however, suggest a model of C₃-to-CAM evolution which allows for lability and reversibility among C₃-CAM phenotypes and C₃ photosynthesis. Continued phylogenetic and physiological investigation into both Australian *Calandrinia* and the rest of the Montiaceae and Portulacineae may help to further resolve the evolutionary significance of C₃-CAM variation.

Climatic space of C₃-CAM phenotypes

Non-phylogenetic climate space analyses of the 22 species included in the study indicate that both low-level constitutive CAM species and facultative CAM species occupy significantly drier regions than C₃ species, and facultative CAM species occupy a broader climate envelope than low-level CAM species. In spite of these general patterns, independent contrasts showed no relationship between photosynthetic type, when coded as either a qualitative or quantitative trait, and climate. These findings are not entirely surprising given 1) the lack of statistical power that results from having only three contrasts (three C₃ species) in the data, and 2) our experiment included only a sub-sample of Australian *Calandrinia* (i.e. only 22 of ~ 70 species), so inferred sister species

relationships in the analysis are not necessarily true sister species. Finally, the lack of correlation between photosynthesis and climate clearly reflects the occupation of C₃ species in two distinct climate zones: either hot and wet, or cold and dry. Two of the C₃ species, *C. arenicola* and *C. tumida*, inhabit the hot, wet tropics of northern Queensland, while the third C₃ species, *C. granulifera*, inhabits cool, drier climates of southern Australia, including northern Tasmania (Fig. 6a, b). All three species germinate at the end of the wet season, when temperatures are cooler and water is readily available. C₃ plants are well adapted to moist and cool climates and so it is not surprising, from an ecological perspective, that these three species do not engage in CAM photosynthesis.

Facultative CAM plants tend to inhabit arid environments where precipitation is variable or distinctly seasonal, though in some cases (i.e. *Clusia*), facultative CAM plants are epiphytic or hemi-epiphytic in wet forests (Zotz & Winter, 1993; 1994; Winter & Holtum, 2014). Under such environmental conditions, plants use C₃ photosynthesis when water is available, allowing for higher rates of CO₂ fixation and vegetative growth than when CAM. When water becomes limiting, plants engage a CAM cycle, which provides prolonged net carbon gain at a lower water cost, presumably allowing for higher reproductive output. This increase in fecundity with CAM is best demonstrated in the facultative CAM annual *Mesembryanthemum crystallinum* (Winter & Zeigler, 1992). In one experiment, *M. crystallinum* exposed to air without CO₂ during the night – and therefore unable to engage fully in CAM – exhibited a 90% decline in seed production compared to those plants that were provided CO₂ at night. Furthermore, it has been shown that mutants without facultative CAM had lower fecundity than wild-type individuals (Cushman *et al.*, 2008). Beyond prolonged carbon gain and increases in

fecundity, switching to CAM may help plants survive periods of stress (i.e. drought) until they can start using C_3 photosynthesis again. The adaptive advantage of constitutive low-level CAM, however, is not fully understood. One hypothesis is that by engaging in low-level CAM, plants can reduce daytime transpiration rates and thereby increase water use efficiency and overall plant fitness (Harris & Martin, 1991; Herrera, 2009). Another hypothesis is that by always engaging in low-level CAM, these plants can quickly upregulate CAM expression in the face of stress (e.g., water deficit).

Despite the possible advantages of low-level CAM, there are also likely to be trade-offs associated with maintaining this phenotype, as it has presumably been lost multiple times in Australian *Calandrinia* (Fig. 4). One trade-off to consider is that CAM is energetically more expensive than C_3 , requiring an additional 2.5 or 3.5 more ATP per CO_2 than required by the C_3 pathway (assuming zero photorespiration in C_3) (Nobel, 1991; Winer & Smith, 1996). Another possible detriment may involve anatomical traits that enhance the degree or efficiency of CAM but limit the efficiency of C_3 photosynthesis. Reduction in intercellular airspace is known to be correlated to CAM function, as this physical trait limits CO_2 flux and thereby increases the likelihood of CO_2 recapture from both respiration and malate decarboxylation (Maxwell *et al.*, 1997; Nelson *et al.*, 2005; Nelson & Sage, 2007; Males 2018). However, the reduction in mesophyll conductance associated with reduced intercellular air space (IAS) limits the uptake of atmospheric CO_2 , the concentration of CO_2 around the active sites of Rubisco, and thus the photosynthetic efficiency of the C_3 pathway (Evans and von Caemmerer, 1996; Maxwell *et al.*, 1997).

It might just be that maintaining or engaging in low-level CAM, regardless of environmental conditions, may not be of any benefit for species growing in regions that do not experience intermittent drought but instead have short life-cycles that occur during a period of gradual drying. For example, annuals *C. uniflora* and *C. spergularina* germinate following the monsoonal summer rains that characterize the tropics of northern Australia. These species typically complete their life-cycle in 6 to 8 weeks, during which time the soil slowly dries out. Only towards the end of their lives do these species really contend with water limitation. It is conceivable, then, that these plants (which use facultative CAM) only engage CAM at the end of life thereby extending carbon gain for reproductive output.

Finally, there appear to be no known full CAM species that are also annuals. Although there are perennial Australian *Calandrinia* species (i.e. section *Tuberosae* and species in *Pseudodanthodiae*), these perennials are seasonal in their growth. We hypothesize that full CAM may not have evolved in this lineage because all of the species are essentially annual, and full CAM does not appear to be selected for in species with this life history strategy.

CAM as a discrete versus a continuous trait

Defining and differentiating CAM phenotypes is difficult, in part because CAM is never really the sole contributor to carbon gain in a plant; the pathway always co-exists with C_3 photosynthesis, or in the case of *Portulaca*, C_4 photosynthesis (Koch & Kennedy, 1982; Christin *et al.*, 2014; Holtum *et al.*, 2017a). Additionally, the expression and contribution of CAM often varies within a species in response to environmental stress, and this response is not always uniform across a species range (Gurevitch *et al.*, 1986;

Guralnick & Gladsky, 2017). Thus, the classification of species into discrete phenotypes is not straight forward. Differentiating between low-level and facultative CAM can be particularly problematic, as small amounts of acid accumulation can be difficult to detect. For example, *C. volubilis* and *C. quadrivalvis* were classified as facultative CAM, even though their ΔH^+ values ($3.2 \mu\text{mol g}^{-1} \text{fw}$ and $3.0 \mu\text{mol g}^{-1} \text{fw}$, respectively) were similar to those recovered for low-level CAM plants *C. eremaea* and *C. kalanenensis* ($3.8 \mu\text{mol g}^{-1}$ and $3.3 \mu\text{mol g}^{-1}$, respectively) when droughted. What differentiates these species is a significance value of 0.05 between PM and AM titratable acidity measurements (ΔH^+), made on only a handful of individuals in varied physiological states. Given these limitations, and considering that CAM expression can vary across a species range (Zotz, 2002; Winter *et al.*, 2015; Guralnick & Gladsky, 2017), we are not altogether certain that our discrete classification is warranted. We suspect that the most useful approach in discerning truly qualitatively different states of C_3 -CAM photosynthesis will be detailed studies of comparative gene expression in closely related C_3 , low-level constitutive CAM, and facultative CAM plants (e.g., Heyduk *et al.*, in press).

C₃-CAM biogeography in Australia

The presence and type of CAM activity and its relationship to climatic space in Australian *Calandrinia* remains puzzling. Australia is the most arid vegetated continent and yet it harbors no native, large stem-succulents, and less than 1% of Australia's plant species are known to use CAM photosynthesis (Holtum *et al.*, 2016). Most of these CAM species are epiphytes, while only twelve are terrestrial (Holtum *et al.*, 2016). Of these terrestrial CAM plants, only one is known to be a full constitutive CAM plant

(*Cynanchum viminale* L.), while the rest are C₃-CAM (Holtum *et al.*, 2016). This study not only triples the number of known C₃-CAM plants native to Australia, but highlights how C₃-CAM phenotypes, not full CAM, dominate the Australian CAM flora.

Why there are so few full CAM plants and considerably more C₃-CAM plants in Australia is still unclear. Ellenberg (1991) suggested that precipitation in arid Australia is too variable to support the large succulent life form, which is closely associated with CAM photosynthesis. By comparing the climate space of regions that harbor large succulents around the world (i.e. the Americas and Africa), Ellenberg identified a rainfall predictability envelope that he hypothesized was required for succulent life forms such as the cacti and *Agave*. Only a small region in South Australia and Victoria fits within this climatic envelope. However, non-native succulent CAM plants have become invasive weeds (i.e. *Opuntia stricta*) following introduction to Australia, and flourish outside of Ellenberg's rainfall predictability envelope (Mann, 1970; Holtum *et al.*, 2016).

There is an alternative argument: perhaps Australia's climate is perfectly suited to full CAM, but the lineages with the propensity to evolve full CAM are simply not present. This does not necessarily appear to be the case either, given that there are lineages in Australia where full CAM and succulence have evolved, but only in species inhabiting different continents [e.g., *Euphorbia* (Horn *et al.*, 2014), *Peperomia* (Holthe *et al.*, 1992), *Anacampseros* (Guralnick *et al.*, 2007), Crassulaceae (Ranson & Thomas, 1960; Caemmerer & Griffiths, 2009)]. In spite of the long history of research into CAM biology (Ranson & Thomas, 1960), we still know very little about the evolution of a full CAM syndrome- how often it has evolved, whether it is reversible, and if there are precise anatomical requirements that facilitate its assembly. We predict that, in contrast to

the considerable evolutionary lability among C₃ and C₃-CAM phenotypes we have demonstrated here, the transition to a full CAM phenotype may have happened more rarely, and requires significant anatomical evolution (e.g., Edwards & Donoghue, 2006).

ACKNOWLEDGEMENTS

The authors thank the Edwards lab members for comments on the manuscript and Elissa Martin for figure and illustration assistance. Special thanks to Rosemary Dunn for hosting LPH while in Australia. This research was funded by the National Science Foundation [DEB-1252901] and the Australian Research Council [DP160100098]. L.P.H. was supported in part by NSF IGERT grant DGE-11 0966060 and by an NSF EAPSI Australia (2015) award.

REFERENCES

- Borland AM, Griffiths H, Maxwell C, Broadmeadow MSJ, Griffiths NM, Barnes JD. 1992.** On the ecophysiology of the Clusiaceae in Trinidad: expression of CAM in *Clusia minor* L. during the transition from wet to dry season and characterization of three endemic species. *New Phytologist* **122**: 349–357.
- Castillo FJ. 1996.** Antioxidative protection in the inducible CAM plant *Sedum album* L following the imposition of severe water stress and recovery. *Oecologia* **107**: 469–477.
- Christin P-A, Osborne CP, Chatelet DS, Columbus JT, Besnard G, Hodkinson TR, Garrison LM, Vorontsova MS, Edwards EJ. 2013.** Anatomical enablers and the evolution of C₄ photosynthesis in grasses. *Proceedings of the National Academy of Sciences* **110**: 1381–1386.
- Christin P-A, Sage TL, Edwards EJ, Ogburn RM, Khoshravesh R, Sage RF. 2011.** Complex evolutionary transitions and the significance of C₃-C₄ intermediate forms of photosynthesis in Molluginaceae. *Evolution* **65**: 643–660.
- Christin P-A, Wallace MJ, Clayton H, Edwards EJ, Furbank RT, Hattersley PW, Sage RF, Macfarlane TD, Ludwig M. 2012.** Multiple photosynthetic transitions, polyploidy, and lateral gene transfer in the grass subtribe Neurachninae. *Journal of Experimental Botany* **63**: 6297–6308.
- Crayn DM, Winter K, Smith JAC. 2004.** Multiple origins of crassulacean acid metabolism and the epiphytic habit in the Neotropical family Bromeliaceae. *Proceedings of the National Academy of Sciences* **101**: 3703–3708.

- Crayn DM, Winter K, Schulte K, Smith JAC. 2015.** Photosynthetic pathways in Bromeliaceae: phylogenetic and ecological significance of CAM and C₃ based on carbon isotope ratios for 1893 species. *Botanical Journal of the Linnean Society* **178**: 169–221.
- Cushman JC. 2001.** Crassulacean acid metabolism. A plastic photosynthetic adaptation to Arid environments. *Plant Physiology* **127**: 1439–1448.
- Cushman JC, Borland AM. 2002.** Induction of cassulacean acid metabolism by water limitation. *Plant, Cell & Environment* **25**: 295–310.
- Cushman JC, Tillett RL, Wood JA, Branco JM, Schlauch KA. 2008.** Large-scale mRNA expression profiling in the common ice plant, *Mesembryanthemum crystallinum*, performing C₃ photosynthesis and crassulacean acid metabolism (CAM). *Journal of Experimental Botany* **59**: 1875–1894.
- Edwards EJ, Diaz M. 2006.** Ecological physiology of *Pereskia guamacho*, a cactus with leaves. *Plant, Cell & Environment* **29**: 247–256.
- Edwards EJ, Donoghue MJ. 2006.** *Pereskia* and the origin of the cactus life-form. *The American Naturalist* **167**: 777–793.
- Edwards EJ, Ogburn RM. 2012.** Angiosperm responses to a low-CO₂ world: CAM and C₄ photosynthesis as parallel evolutionary trajectories. *International Journal of Plant Sciences* **173**: 724–733.
- Edwards EJ, de Vos JM, Donoghue MJ. 2015.** Doubtful pathways to cold tolerance in

plants. *Nature* **521**: E5–6.

Ellenberg H. 1981. Reasons for stem succulents being present or absent in the arid regions of the world. *Flora* **171**:114-169.

Evans JR, von Caemmerer S. 1996. Carbon dioxide diffusion inside leaves. *Plant Physiology* **110**: 339–346.

Flexas J, Diaz-Espejo A, Berry J, Cifre J, Galmes J, Kaldenhoff R, Medrano H, Ribas-Carbo M. 2007. Analysis of leakage in IRGA's leaf chambers of open gas exchange systems: quantification and its effects in photosynthesis parameterization. *Journal of Experimental Botany* **58**: 1533–1543.

Gehrig HH, Aranda J, Cushman MA, Plant AV. 2003. Cladogram of Panamanian *Clusia* based on nuclear DNA: implications for the origins of crassulacean acid metabolism. *Plant Biology* **5**: 59–90.

Griffiths H. 1989. Crassulacean acid metabolism: a re-appraisal of physiological plasticity in form and function. Advances in Botanical Research. *Advances in Botanical Research* **15**: 43–92.

Guralnick LJ, Gladsky K. 2017. Crassulacean acid metabolism as a continuous trait: variability in the contribution of Crassulacean acid metabolism (CAM) in populations of *Portulacaria afra*. *Heliyon*: e00293.

Guralnick LJ, Jackson MD. 2001. The occurrence and phylogenetics of crassulacean acid metabolism in the Portulacaceae. *International Journal of Plant Sciences* **162**: 257–

Guralnick LJ, Cline A, Smith M, Sage RF. 2007. Evolutionary physiology: the extent of C₄ and CAM photosynthesis in the genera *Anacampseros* and *Grahamia* of the Portulacaceae. *Journal of Experimental Botany* **59**: 1735–1742.

Hancock LP, Edwards EJ. 2014. Phylogeny and the inference of evolutionary trajectories. *Journal of Experimental Botany* **65**: 3491–3498.

Hancock LP, Obbens F, Moore A, Thiele K, de Vos J, West J, Holtum JAM, Edwards E. 2018. Phylogeny, evolution, and biogeographic history of *Calandrinia* (Montiaceae). *American Journal of Botany*, in revision.

Harris FS, Martin CE. 1991. Correlation between CAM-cycling and photosynthetic gas exchange in five species of *Talinum* (Portulacaceae). *Plant Physiology* **96**: 1118-1124.

Herrera A. 2009. Crassulacean acid metabolism and fitness under water deficit stress: if not for carbon gain, what is facultative CAM good for? *Annals of Botany* **103**: 645–653.

Herrera A. 2013. Crassulacean acid metabolism-cycling in *Euphorbia milii*. *AoB plants*: 1-9.

Heyduk K, McKain MR, Lalani F, Leebens-Mack J. 2016. Evolution of a CAM anatomy predates the origins of Crassulacean acid metabolism in the Agavoideae (Asparagaceae). *Molecular Phylogenetics and Evolution* **105**: 102–113.

Heyduk K, Ray NJ, Ayyampalayam S, and Leebens-Mack J. 2018. Shifting gene

expression profiles associated with weak and strong Crassulacean acid metabolism.

American Journal of Botany, in press.

Holthe PA, Patel A, Ting IP. 1992. The occurrence of CAM in *Peperomia. Selbyana* **13**: 77–87.

Holtum J, Osmond CB. 1981. The gluconeogenic metabolism of pyruvate during deacidification in plants with crassulacean acid metabolism. *Australian Journal of Plant Physiology* **8**: 31–44.

Holtum JAM, Hancock LP, Edwards EJ, Crisp MD, Crayn DM, Sage R, Winter K. 2016. Australia lacks stem succulents but is it depauperate in plants with crassulacean acid metabolism (CAM)? *Current Opinion in Plant Biology* **31**: 109–117.

Holtum JAM, Winter K. 1999. Degrees of crassulacean acid metabolism in tropical epiphytic and lithophytic ferns. *Functional Plant Biology* **26**: 749–757.

Holtum JAM, Hancock L, Edwards EJ, Winter K. 2017a. Optional use of CAM photosynthesis in two C₄ species, *Portulaca cyclophylla* and *Portulaca digyna*. *Plant Physiology* **214**: 91–96.

Holtum JAM, Hancock L, Edwards EJ, Winter K. 2017b. Facultative CAM photosynthesis (crassulacean acid metabolism) in four species of *Calandrinia*, ephemeral succulents of arid Australia. *Photosynthesis Research* **134**: 1–21.

Holtum JAM, Smith JAC, Neuhaus HE. 2005. Intracellular transport and pathways of carbon flow in plants with crassulacean acid metabolism. *Functional Plant Biology* **32**:

429–449.

Horn JW, Xi Z, Riina R, Peirson JA, Yang Y, Dorsey BL, Berry PE, Davis CC, Wurdack KJ. 2014. Evolutionary bursts in *Euphorbia* (Euphorbiaceae) are linked with photosynthetic pathway. *Evolution* **68**: 3485–3504.

Kadereit G, Borsch T, Weising K, Freitag H. 2003. Phylogeny of Amaranthaceae and Chenopodiaceae and the evolution of C₄ photosynthesis. *International Journal of Plant* **164**: 959-986.

Kadereit G, Mavrodiev EV, Zacharias EH, Sukhorukov AP. 2010. Molecular phylogeny of Atripliceae (Chenopodioideae, Chenopodiaceae): implications for systematics, biogeography, flower and fruit evolution, and the origin of C₄ photosynthesis. *American Journal of Botany* **97**: 1664–1687.

Koch KE, Kennedy RA. 1982. Crassulacean Acid Metabolism in the succulent C₄ Dicot, *Portulaca oleracea* under natural environmental conditions. *Plant Physiology* **69**: 757–761.

Lüttg U. 1999. One morphotype, three physiotypes: sympatric species of *Clusia* with obligate C₃ photosynthesis, obligate CAM and C₃-CAM intermediate behaviour. *Plant Biology* **1**: 138–148.

Lüttg U. 2006. Photosynthetic flexibility and ecophysiological plasticity: questions and lessons from *Clusia*, the only CAM tree, in the neotropics. *The New phytologist* **171**: 7–

25.

Martin CE, Higley M, Wang WZ. 1988. Ecophysiological significance of CO₂-recycling via Crassulacean acid metabolism in *Talinum calycinum* Engelm. (Portulacaceae). *Plant Physiology* **86**: 562-588.

Mann J. 1990. Cacti naturalised in Australia and their control. Brisbane: S.G. Reid, Government Printer.

Maxwell K, von Caemmerer S, Evans JR. 1997. Is a low internal conductance to CO₂ diffusion a consequence of succulence in plants with crassulacean acid metabolism? *Australian Journal of Plant Physiology* **24**: 777.

McKown AD, Moncalvo J-M, Dengler NG. 2005. Phylogeny of *Flaveria* (Asteraceae) and inference of C₄ photosynthesis evolution. *American Journal of Botany* **92**: 1911–1928.

Nelson EA, Sage TL, Sage RF. 2005. Functional leaf anatomy of plants with crassulacean acid metabolism. *Functional Plant Biology* **32**: 409-419.

Nelson EA, Sage RF. 2007. Functional constraints of CAM leaf anatomy: tight cell packing is associated with increased CAM function across a gradient of CAM expression. *Journal of Experimental Botany* **59**: 1841–1850.

Nobel PS. 1991. Achievable productivities of certain CAM plants: basis for high values compared with C₃ and C₄ plants. *New Phytologist* **119**: 183–205.

Obbens F. 2006. A review of the tuberous *Calandrinia* species (section Tuberosae),

including three new species for Western Australia. *Nutsyia* **16**: 95–115.

Obbens FJ. 2011. Five new species of *Calandrinia* (Portulacaceae) from Western Australia with additional information on morphological observations. *Nutsyia* **21**: 1–23.

Obbens FJ. 2012. Three new species of *Calandrinia* (Portulacaceae) from the Eremaean and South West Botanical Provinces of Western Australia. *Nutsyia* **22**: 1–2.

Obbens FJ, Hancock LP, Edwards E, Thiele KR. 2017. *Calandrinia holtumii* (Portulacaceae), a new and unusual species from arid Western Australia. *Nutsyia* **28**: 217–223

Ocampo G, Columbus JT. 2012. Molecular phylogenetics, historical biogeography, and chromosome number evolution of *Portulaca* (Portulacaceae). *Molecular Phylogenetics and Evolution* **63**: 97–112.

Orme D. 2013. *The Caper Package: Comparative Analysis of Phylogenetics and Evolution in R*.

Osmond CB. 1978. Crassulacean acid metabolism: a curiosity in context. *Annual Review of Plant Physiology* **29**: 397–414.

Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.

Ranson SL, Thomas M. 1960. Crassulacean acid metabolism. *Annual Review of Plant Physiology and Plant Molecular Biology* **11**: 81–110.

Revell LJ. 2012. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**: 217-223.

Sage RF, Sage TL, Kocacinar F. 2012. Photorespiration and the evolution of C₄ photosynthesis. *Annual Review of Plant Biology* **63**: 19–47.

Schuber M, Kluge M. 1981. In situ studies on crassulacean acid metabolism in *Sedum acre* L. and *Sedum mite* Gil. *Oecologia* **50**: 82-87.

Silvera K, Neubig KM, Whitten WM, Williams NH, Winter K, Cushman JC. 2010. Evolution along the crassulacean acid metabolism continuum. *Functional Plant Biology* **37**: 995–1010.

Silvera K, Santiago LS, Winter K. 2005. Distribution of crassulacean acid metabolism in orchids of Panama: evidence of selection for weak and strong modes. *Functional Plant Biology* **32**: 397-407.

Silvera K, Santiago LS, Cushman JC, Winter K. 2010. The incidence of crassulacean acid metabolism in the Orchidaceae derived from carbon isotope ratios: a checklist of the flora of Panama and Costa Rica. *Botanical Journal of the Linnean Society* **163**: 194–222.

Sternberg LO, DeNiro MJ, Johnson HB. 1984. Isotope ratios of cellulose from plants having different photosynthetic pathways. *Plant Physiology* **74**: 557–561.

Szarek SR, Johnson HB, Ting IP. 1973. Drought adaptation in *Opuntia basilaris*. *Plant Physiology* **52**: 539–541.

- Tahir SS, Carolin RC. 2011.** A new species of *Calandrinia* (Portulacaceae) from Northern Territory, Australia. *Proceedings of the Linnean Society of New South Wales*: 11–14.
- Ting IP. 1985.** Crassulacean acid metabolism. *Annual Review of Plant Physiology* **36**: 595–622.
- Ting IP, Sipes D. 1985.** Metabolic modifications of crassulacean acid metabolism - CAM-idling and CAM cycling. In: Burris RH, Ludden PW, Burris JE, eds. *Nitrogen fixation and CO₂ metabolism*. Elsevier, New York: 371–378.
- Vaasen A, Begerow D, Luttge U, Biology RHP. 2002.** The genus *Clusia* L.: molecular evidence for independent evolution of photosynthetic flexibility. *Plant Biology* **4**: 86–93.
- von Caemmerer S, Griffiths H. 2009.** Stomatal responses to CO₂ during a diel Crassulacean acid metabolism cycle in *Kalanchoe daigremontiana* and *Kalanchoe pinnata*. *Plant, Cell & Environment* **32**: 567–576.
- West JG, Chinnock RJ. 2013.** *Calandrinia mirabilis* (Portulacaceae), a spectacular new species from Western Australia with notes on its ecology, seed germination and horticultural potential. *Journal of the Adelaide Botanic Garden* **26**: 97–102.
- Winter K, Holtum JAM. 2002.** How closely do the $\delta^{13}\text{C}$ values of Crassulacean acid metabolism plants reflect the proportion of CO₂ fixed during day and night? *Plant Physiology* **129**: 1843–1851.
- Winter K, Holtum JAM. 2006.** Environment or development? Lifetime net

CO₂ exchange and control of the expression of crassulacean acid metabolism in *Mesembryanthemum crystallinum*. *Plant Physiology* **143**: 98–107.

Winter K, Holtum JAM. 2011. Induction and reversal of Crassulacean acid metabolism in *Calandrinia polyandra*: effects of soil moisture and nutrients. *Functional Plant Biology* **38**: 576–582.

Winter K, Holtum JAM. 2014. Facultative crassulacean acid metabolism (CAM) plants: powerful tools for unravelling the functional elements of CAM photosynthesis. *Journal of Experimental Botany* **65**: 3425–3441.

Winter K, Smith JAC. 1996. *Crassulacean Acid Metabolism* (K Winter and JAC Smith, Eds.). Berlin, Heidelberg: Ecological Studies.

Winter K, Garcia M, Holtum JAM. 2008. On the nature of facultative and constitutive CAM: environmental and developmental control of CAM expression during early growth of *Clusia*, *Kalanchoë*, and *Opuntia*. *Journal of Experimental Botany* **59**: 1829–1840.

Winter K, Holtum JAM, Smith JAC. 2015. Crassulacean acid metabolism: a continuous or discrete trait? *New Phytologist* **208**: 73–78.

Winter K, Schröppel-Meier G, Caldwell MM. 1986. Respiratory CO₂ as carbon source for nocturnal acid synthesis at high temperatures in three species exhibiting crassulacean acid metabolism. *Plant Physiology* **81**: 390–394.

Winter K and Ziegler H. 1992. Induction of crassulacean acid metabolism in *Mesembryanthemum crystallinum* increases reproductive success under conditions of

drought and salinity stress. *Oecologia* **92**: 475–479.

Zotz G and Winter K. 1993. Short-term regulation of crassulacean acid metabolism activity in a tropical hemiepiphyte, *Clusia uvitana*. *Plant Physiology* **102**: 835-841.

Zotz G and Winter K. 1994. One year study on carbon, water, and nutrient relationships in a tropic C₃-CAM hemi-epiphyte *Clusia uvitana* Pittier. *New Phytologist* **127**: 45-60.

Zotz G. 2002. Categories and CAM – blurring divisions, increasing understanding? *New Phytologist* **156**: 4–6.

Table 1. Voucher specimen table and species mean values for $\delta^{13}\text{C}$ and titratable acidity of extracts of leaves from over the course of the watered, drought, and re-watered experiment

Species	Voucher ID	Collector	Origin	Latitude (S)	Longitude (E)	Min. $\delta^{13}\text{C}$	Mean $\delta^{13}\text{C}$	sd $\delta^{13}\text{C}$	Watered		Drought		Re-Watered		Phenotype
									Mean ΔH^+	sd ΔH^+	Mean ΔH^+	sd ΔH^+	Mean ΔH^+	sd ΔH^+	
<i>C. arenicola</i>	JAH 137	Leesa	~5km NW of Laura, QLD	-15.501	144.484	-27.12	-27.93	0.74	0.00	0.00	-0.83	3.01	0.00	0.00	C3
<i>C. balonensis</i>	LPH 85	LPH	~7 km south of Alice Springs	-23.813	133.779	-27.00	-27.70	0.82	10.23	2.63	12.01	1.90	0.64	0.52	low-level
<i>C. calypttrata</i>	JAH191	LPH	Long Forest Reserve, Victoria	-37.644	144.497	-26.56	-27.90	1.41	0.26	0.26	8.59	2.61	1.76	1.22	fac. CAM
<i>C. creethae</i>	JAH 16	AK	Willuna, WA	-26.591	120.152	-26.24	-28.35	1.62	-0.58	2.44	8.10	2.08	10.18	0.93	fac. CAM
<i>C. crispispala</i>	LPH 196	LPH	South of Yalgoo, WA	-28.999	117.125	-23.82	-24.76	1.33	5.93	1.61	45.82	2.07	1.07	2.33	low-level
<i>C. eremaea</i>	JAH 5	AK	You Yang Ranges, VIC	-37.953	144.423	-23.51	-25.65	2.30	3.75	1.10	10.83	2.54	3.70	1.42	low-level
<i>C. flava</i>	JAH 6	AK	Mt. Magnet, WA	-28.065	117.848	-21.68	-23.09	2.00	13.10	2.00	27.60	2.18	24.28	6.98	low-level
<i>C. granulifera</i>	JAH 7	AK	from AK, east of Northam WA	-31.655	116.671	-29.04	-29.46	0.49	1.67	1.19	-0.05	1.80	4.91	2.03	C3
			41 km N along Yalgoo Road from junction w/ Geraldton- Mt. Magnet Rd												
<i>C. kalanennsis</i>	LPH 198	LPH		-28.013	116.721	-27.00	-27.60	0.85	3.31	0.70	42.62	4.20	4.88	2.09	low-level
<i>C. mirabilis</i>	JAH 21	AK	from AK, RJC10295			-25.11	-25.73	0.70	0.39	1.55	11.33	3.45	-4.98	0.58	fac. CAM
			~5.3km N of Cobra												
<i>C. sp. The Pink Hills</i>	LPH 178	LPH	Daily Creek Homestead	-25.258	115.890	-25.01	-25.83	1.17	0.17	0.22	13.19	0.19	6.08	1.80	fac. CAM
			3.2 km S of Meka												
<i>C. papillata</i>	LPH 191	LPH	Homestead	-27.423	116.827	-24.90	-24.90	NA	0.24	0.13	22.64	1.84	3.40	1.25	fac. CAM
			Near peak of Mt Bald,												
<i>C. pickeringii</i>	JAH 40	JAH	northern NSW.	-28.852	152.045	-28.24	-28.24	NA	12.04	1.81	9.32	1.69	11.01	1.23	low-level
			QH 568091, QH699641												
<i>Calandrinia pleiopen</i>	JAH 3	JAH	or QH726961			-25.60	-27.20	1.85	5.21	2.01	22.00	3.37	6.87	3.53	low-level
			From seeds in BNE (AQ0736128 - 202.7 km by road W of St George on road to Cunnamulla) from AK, Sth Wongan Hills WA												
<i>C. ptychosperma</i>	JAH 36	JAH		-28.004	146.538	-25.95	-27.45	0.97	1.20	0.70	7.50	1.00	5.60	0.50	low-level
<i>C. pumila</i>	JAH 10	JAH		-31.034	116.930	-27.79	-28.55	1.06	5.02	2.19	8.76	2.67	-2.34	1.13	low-level
			Durak River crossing about 100 m west side of xing - side of road on southern side - in sand deposition.												
<i>C. quadrivalvis</i>	JAH 66	JAH	from AK, Meekatharra, WA	-15.942	127.219	-28.01	-28.43	0.66	2.97	2.45	12.18	3.06	3.89	1.95	fac. CAM
<i>C. reticulata</i>	JAH 12	AK		-26.025	118.698	-23.19	-26.47	2.48	13.74	1.50	6.32	2.20	5.70	1.10	low-level
			Along Savannah Way, east of Burktown, NT												
<i>C. spurgularina</i>	JAH 204	LPH 130		-10.043	136.399	-27.31	-28.71	1.31	0.00	0.00	7.55	2.52	0.99	0.99	fac. CAM
			~5km NW of Laura, QLD												
<i>C. tumida</i>	JAH 138	Leesa		-15.501	144.484	-31.16	-31.16	NA	0.00	0.00	0.00	0.00	0.00	0.00	C3
			Stuart Hwy roadside - ca 1.5 m from road - 43 km past Larramie												
<i>C. uniflora</i>	JAH 56	JAH		-15.207	133.105	-27.19	-28.83	1.35	-5.78	1.30	7.55	2.70	2.91	1.33	fac. CAM
			Bob Chinnock, Adelaide, SA												
<i>C. volubilis</i>	JAH 194	BC		-34.938	138.841	-23.52	-25.70	3.00	3.16	1.90	82.23	11.61	23.79	2.71	fac. CAM

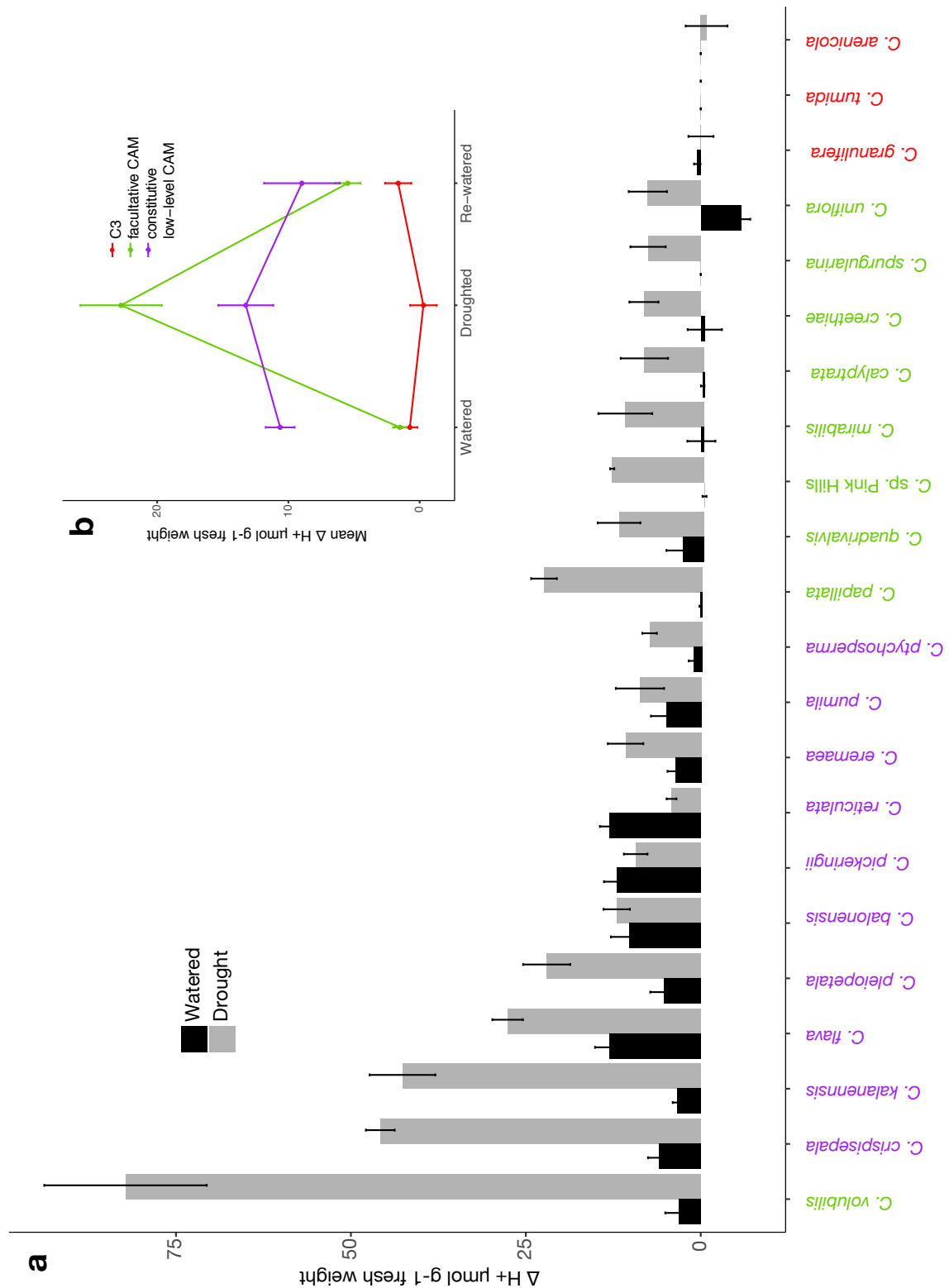


Figure 2. **a**) Barplot showing mean ($\pm$ s.e.) nocturnal CO₂ fixation (ΔH^+ $\mu\text{mol g}^{-1}$ fresh weight) when plants are watered (grey) versus droughted (black). Species names colored according to phenotypes in figure 2b. **b**) Mean ($\pm$ s.e.) nocturnal CO₂ fixation (ΔH^+ $\mu\text{mol g}^{-1}$ fresh weight) when species are watered, droughted, and re-watered. Species phenotyped as C₃ (red), constitutive low-level CAM (purple), and facultative CAM (green) based on ΔH^+ values over the course of the experiment.

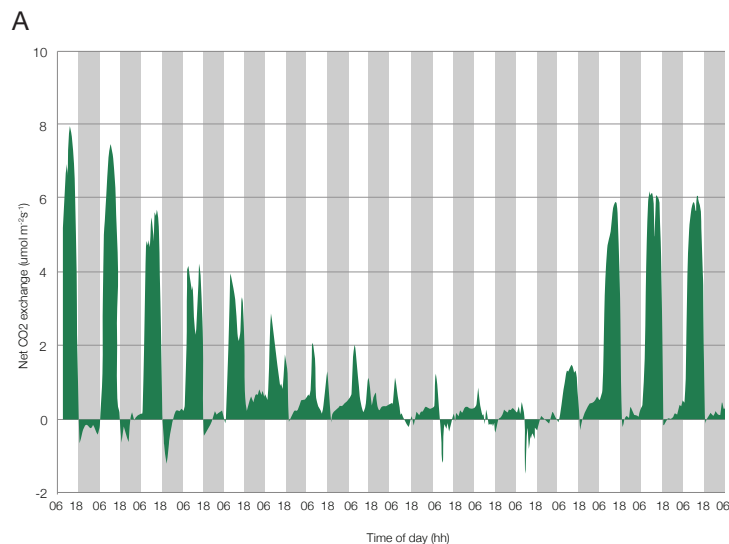


Figure 3a. Sixteen days of net CO₂ exchange by a leaf of a ~ 2-month-old *C. balonensis* plant exposed to drought. Measurements were performed at ambient CO₂ in the 6400-08 standard clear to bottom chamber. The leaf was exposed to ~12h, ~ 500 μmol photons m⁻² s⁻¹, 28C/, 22C light/dark cycles. Watering was withheld from day 2 onwards. Shaded areas indicate the dark periods. Re-watering commenced on day 13.

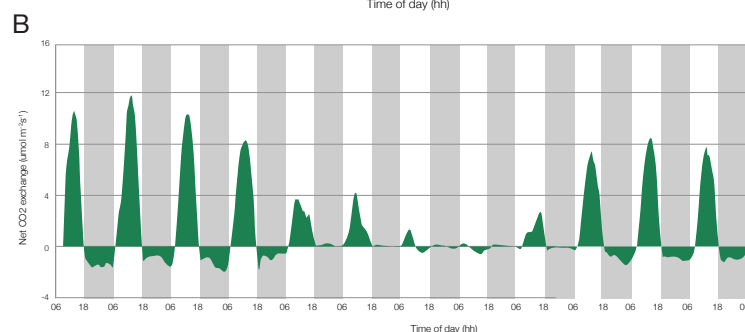


Figure 3b. Twelve days of net CO₂ exchange by a leaf of a ~ 2-month-old *C. mirabilis* plant exposed to drought. Measurements were performed at ambient CO₂ in the 6400-08 standard clear to bottom chamber. The leaf was exposed to ~12h, ~ 500 μmol photons m⁻² s⁻¹, 28C/22C light/dark cycles. Watering was withheld from day 2 onwards. Shaded areas indicate the dark periods. Re-watering commenced on day 9.

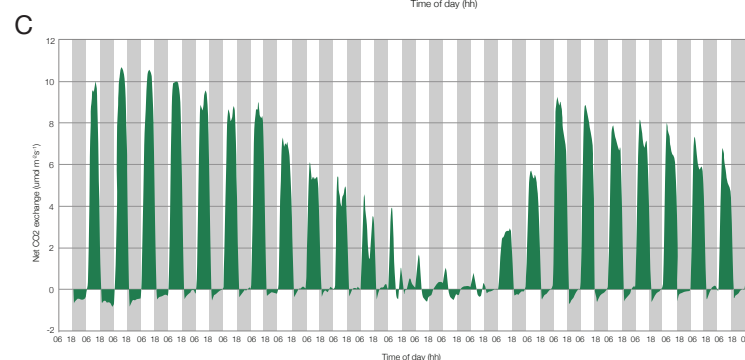


Figure 3c. Twenty-four days of net CO₂ exchange by a leaf of a ~ 2-month-old *C. pickeringii* plant exposed to drought. Measurements were performed at ambient CO₂ in the 6400-08 standard clear to bottom chamber. The leaf was exposed to ~12h, ~ 500 μmol photons m⁻² s⁻¹, 28C/12C, light/dark cycles. Watering was withheld from day 2 onwards. Shaded areas indicate the dark periods. Re-watering commenced on day 17.

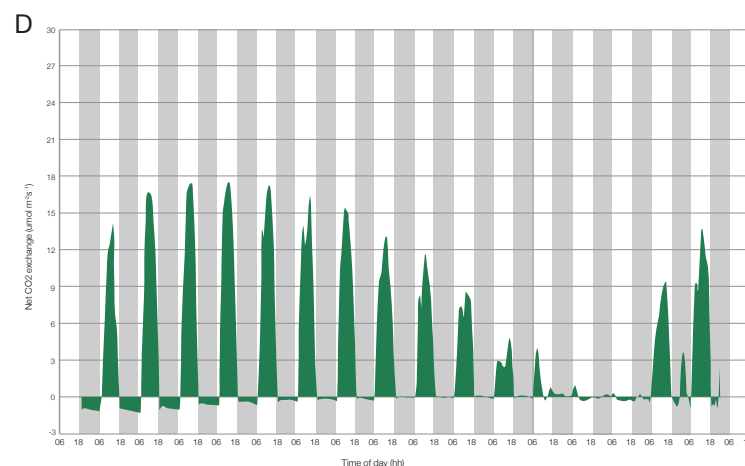


Figure 3d. Seventeen days of net CO₂ exchange by a leaf of a ~ 2-month-old *C. ptychosperma* plant exposed to drought. Measurements were performed at ambient CO₂ in the 6400-08 standard clear to bottom chamber. The leaf was exposed to ~12h, ~ 500 μmol photons m⁻² s⁻¹, 28C/12C, light/dark cycles. Watering was withheld from day 3 onwards. Shaded areas indicate the dark periods. Re-watering commenced on day 15.

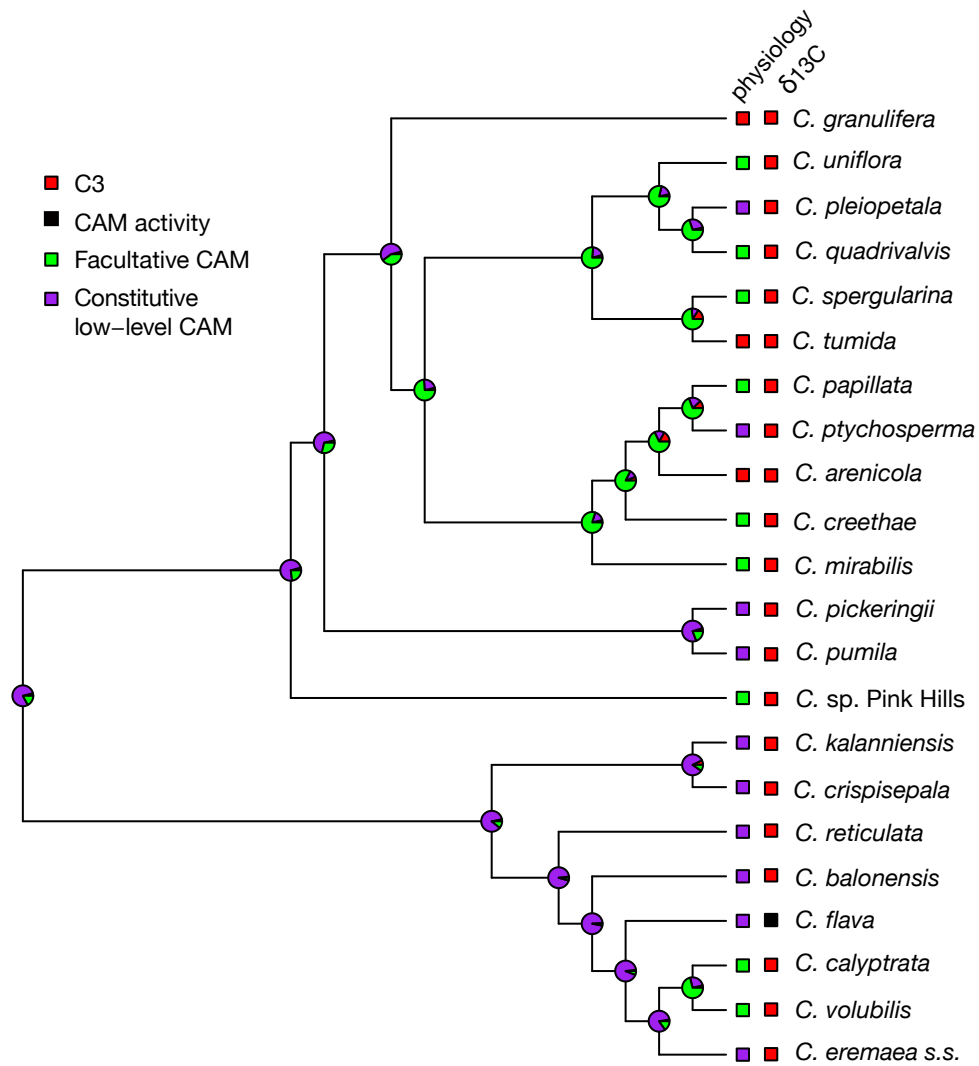


Figure 4. Australian *Calandrinia* species tree investigated in the study. The colored boxes (red=C3, black = CAM activity, purple = constitutive low-level CAM, green = facultative CAM) mapped onto the tips indicate the discrete photosynthetic phenotype for each species derived from physiological methods (first column) compared to $\delta^{13}\text{C}$ values (second column). Ancestral character states were reconstructed in APE under an equal rates model using the photosynthetic phenotype derived from $\delta^{13}\text{C}$ values. Ancestral character states were reconstructed in APE under an equal rates model using the photosynthetic phenotype derived from physiological measurements.

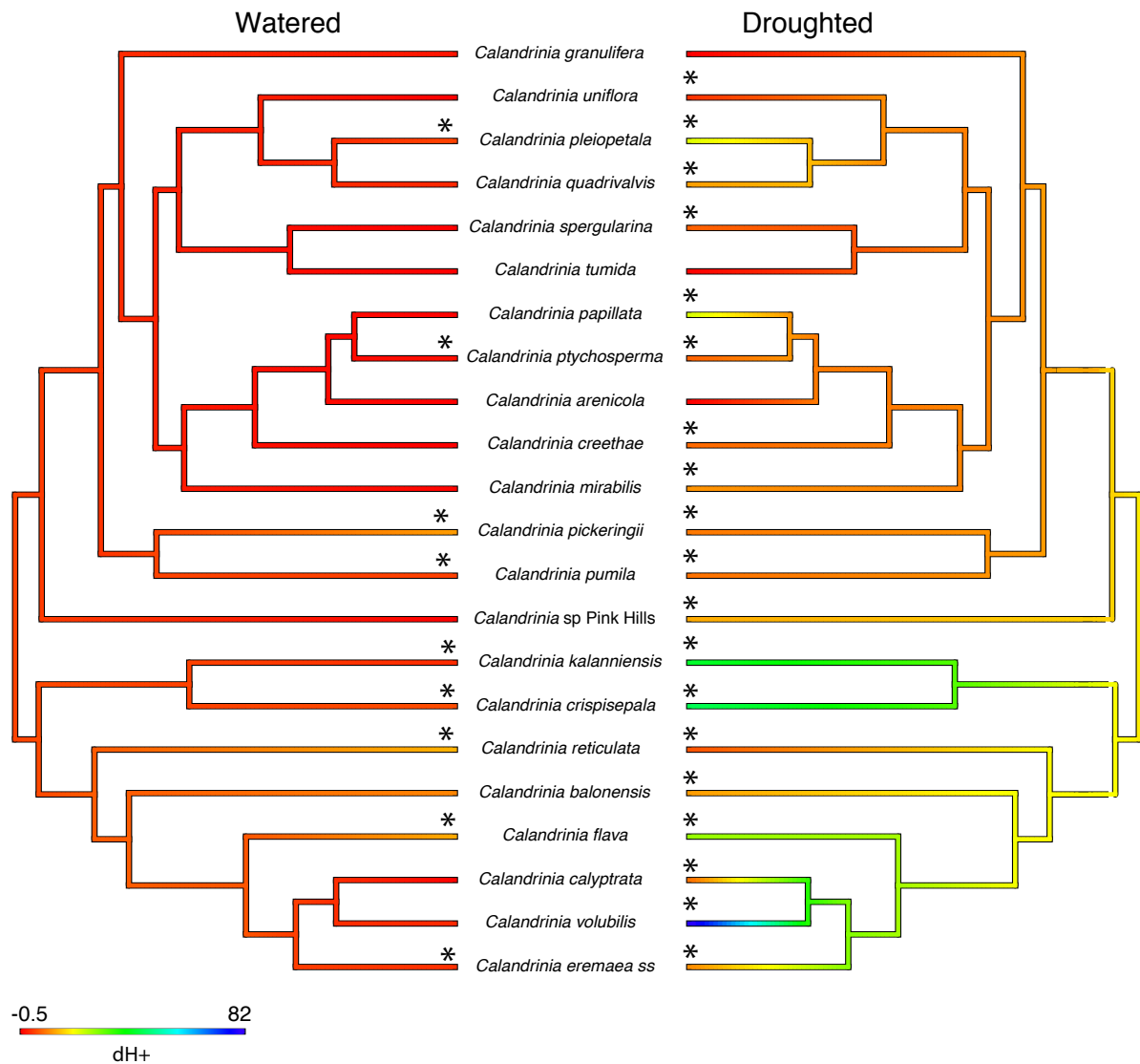
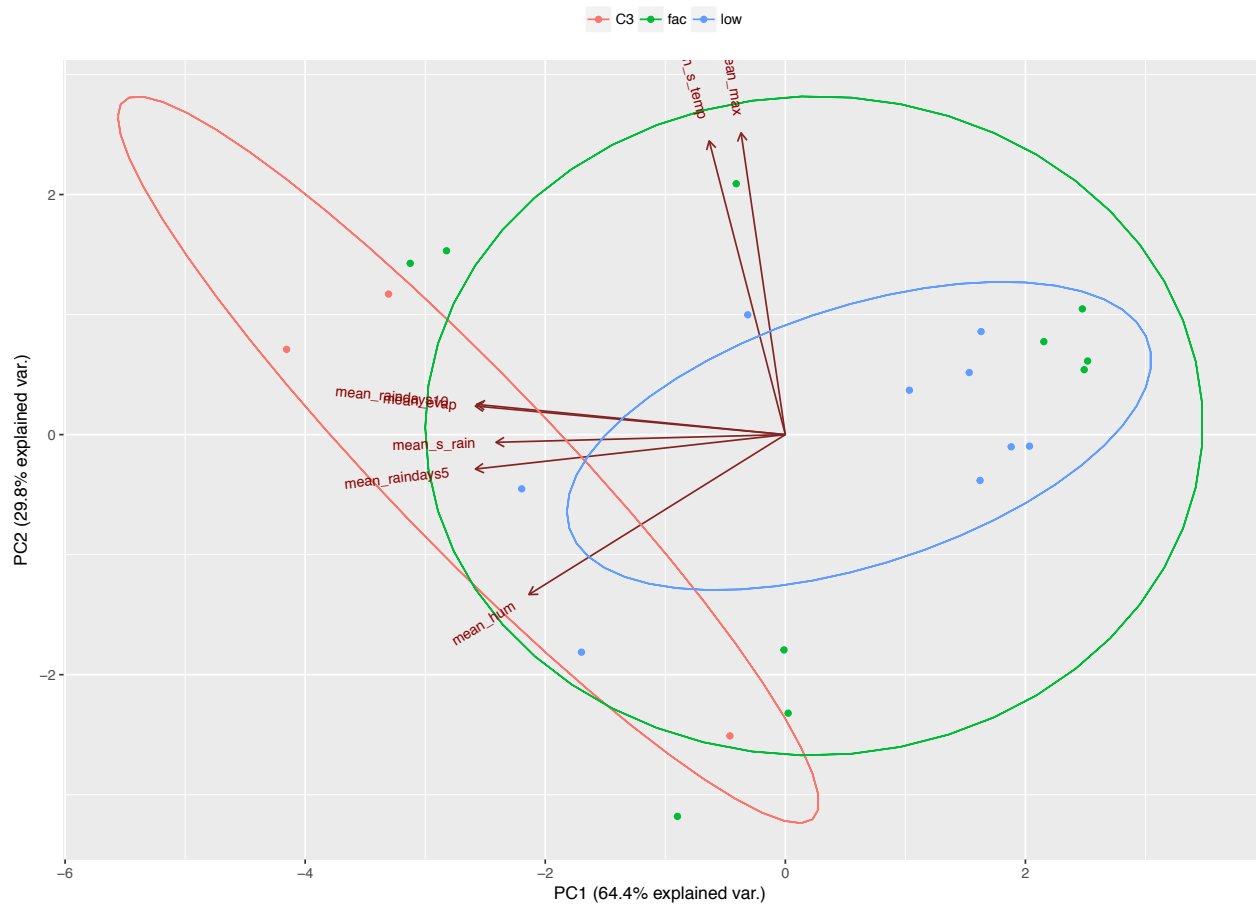


Figure 5. Delta H+ mapped as a continuous character across the 22-tip Australian *Calandrinia* phylogeny when plants were watered (left) versus droughted (right). Stars above branches indicate significant $\Delta H+$ values.



Supplemental Figure 1. Principle component analysis of 8 climatic variables for species in Australian *Calandrinia*. Species are coded by whether they are C3 (red), low-level CAM (blue), or facultative CAM (green).

Chapter 4: CAM biochemistry precedes the evolution of full CAM morphological and anatomical traits in the Portulacineae

Lillian P Hancock, Elissa Martin, Matthew R Ogburn, and Erika J Edwards

In Preparation: Hancock LP, Ogburn M, Martin E, Holtum JAM, and EJ Edwards. CAM biochemistry precedes the evolution of full CAM morphological and anatomical traits in the Portulacineae.

ABSTRACT

Understanding how complex adaptations evolve is an enduring and compelling question in evolutionary biology. In plants, Crassulacean Acid Metabolism (CAM) is a complex trait that improves water use and photosynthetic efficiency under drought and temperature stress. Though it is an important ecological adaptation, having evolved hundreds of times from the more common C_3 photosynthetic pathway, we still know very little about its evolutionary assembly, whether it is reversible or not, and if C_3 -CAM phenotypes represent early steps in the evolution of full CAM photosynthesis. The occurrence and degree of CAM expression is often correlated with transitions in tissue morphology (e.g. increased leaf thickness and succulence) and anatomy (e.g. increased cell size and more tightly packed cells). Hypotheses posit that such morphological and anatomical changes are important for efficient CAM function. To further understand the evolution of these traits, and their evolutionary relationships to C_3 , C_3 -CAM and CAM, we characterized leaf thickness, cell size, percent intercellular airspace, succulence, and venation type across the Portulacineae, a known hot spot of C_3 -to-CAM evolution. Macro-evolutionary trait correlations show that morphological and anatomical traits are not only tightly correlated with each other, but with photosynthetic mode as well, and that C_3 -CAM plants occupy a unique phenotypic trait space. Ancestral character state reconstructions show that C_3 -CAM evolved at the base of the Portulacineae, prior to the evolution of anatomical and morphological modifications correlated with the evolution of full CAM. These analyses uncover one shift to full CAM and multiple reversals back to

C₃, which are likely linked to a changing environment, particularly a change in temperature.

INTRODUCTION

Most green plants use C₃ photosynthesis to fix CO₂ during the day via the carboxylation of ribulose-1 5-bisphosphate by the enzyme Rubisco. When CO₂ concentrations are low and temperatures high, Rubisco will bind to O₂ instead of CO₂, resulting in an energetically wasteful process of photorespiration. To mitigate photorespiration plants have evolved mechanisms to concentrate CO₂ around the active site of Rubisco. Crassulacean Acid Metabolism is one of two carbon concentrating mechanisms (CCMs – the other is C₄ photosynthesis) to have evolved from C₃ ancestors. Although CAM is found in 6-7% of vascular plants and has evolved many times across a diversity of plant lineages (Winter and Smith, 1996), we still know very little about the evolutionary assembly of this ecologically important syndrome.

CAM differs from the C₃ pathway in the primary step of CO₂ assimilation: in CAM, CO₂ fixation is not catalyzed during the light by Rubisco, but rather it is incorporated at night using PEP carboxylase (PEPC). The resulting 4-carbon acid is stored overnight as malic acid in the plant cell vacuole. When daylight ensues, the stomata close and the malic acid is shuttled out of the vacuole where it is subsequently decarboxylated. The released CO₂ is then re-fixed via the C₃ pathway with Rubisco (Osmond, 1978) and the remaining 3-carbon skeleton conserved via gluconeogenesis (Holtum *et al.*, 2005). Depending upon the light intensity, the catabolism of malic acid is usually complete by midday. At this point, typically referred to as the fourth phase of

CAM, plants will open their stomata and engage in a burst of Rubisco-mediated atmospheric CO₂ uptake (a.k.a. C₃ photosynthesis). Characteristics of tissues with CAM therefore include nocturnal acidification (from malic acid) and nocturnal CO₂ fixation (Osmond, 1978; Ting, 1985; Winter, 1985), though all CAM plants still have a fully functional C₃ metabolism.

There are also striking anatomical differences between C₃ and full CAM plants, and species that use CAM often exhibit anatomical traits which likely reflect functional constraints in photosynthetic efficiency (Maxwell *et al.*, 1997; Nelson *et al.*, 2005; Nelson and Sage, 2007; Heyduk *et al.*, 2016). A general feature of CAM anatomy is leaf succulence, which is often characterized by thicker leaves with larger, often undifferentiated mesophyll cells. Increased cell size most often is associated with increased vacuole size because larger vacuoles are required for both water and C₄ acid (i.e. malic acid) storage. An increase in cell size within a given tissue generally results in more tightly packed cells and thus a lower percentage of intercellular air space (IAS) and a reduction in leaf mesophyll surface exposed to IAS ($L_{\text{mes}}/\text{area}$) (Smith and Heuer 1981). This decrease in IAS and $L_{\text{mes}}/\text{area}$ means a more tortuous path of CO₂ diffusion into chloroplasts, resulting in reduced internal CO₂ conductance (lower g_i) (Maxwell *et al.*, 1997). Previous work by Nelson *et al.* (2007) showed that reduced internal CO₂ conductance benefits CAM function as it reduces CO₂ efflux during the third phase of CAM: the phase characterized by the decarboxylation of malate and release of CO₂ around the active site of Rubisco. In CAM plants, the evolution of thicker leaves with more tightly packed, larger cells decreases the efficiency of C₃ photosynthesis, for the same reason that these traits restrict CO₂ diffusion. In the case of C₃, this results in

limiting the amount of CO₂ available to Rubisco and therefore increasing the likelihood of photorespiration. Thus there appears to be a significant anatomical tradeoff between C₃ and CAM function that rests on IAS and mesophyll conductance.

In addition to C₃ and constitutive full CAM (fully expressed CAM), there are other well-known phenotypes that perform small amounts of CAM in an otherwise C₃ background (i.e. C₃-CAM intermediates known as low-level and facultative CAM) (Martin and Zee, 1983; Guralnick and Jackson, 2001; Holtum *et al.*, 2004; Silvera *et al.*, 2005; Winter and Holtum, 2014). Previous research (Nelson and Sage, 2007) has shown that anatomical variation in leaf anatomy reflects the degree of physiological variation in CAM activity: i.e. plants that perform more CAM have more CAM-like anatomy. Moreover, it is hypothesized that these C₃-CAM intermediate phenotypes represent transitional states along the CAM evolutionary trajectory and that transitions from C₃ photosynthesis to CAM include the evolution of anatomical modifications that allow for optimal CAM function (Silvera *et al.*, 2010; Edwards and Ogburn 2012; Heyduk *et al.*, 2016).

The most common tool used to screen for the presence of CAM photosynthesis is the ¹³C/¹²C carbon isotope ratio of dried leaf material. This ratio distinguishes between C₃ and constitutive CAM photosynthesis because Rubisco discriminates against the heavier isotope (¹³C), while PEPC does not; therefore, C₃ plants are depleted in ¹³C relative to ¹²C when compared to constitutive CAM plants (Farquhar and Ehleringer, 1989). Plants that use both C₃ and CAM have Δ¹³C values that typically fall somewhere in between the C₃ and CAM clusters, but can also present pure C₃-like Δ¹³C values (Winter and Holtum, 2002), depending on how often that particular plant has utilized the CAM pathway, and

how much atmospheric CO₂ was fixed by PEPC. Because of this ambiguity in the isotopic method, measurements of nocturnal CO₂ uptake and acidification (measured as titratable acidity) are preferred for C₃-CAM phenotypic detection. Given the amount of time and work it takes to make these physiological measurements (as they need to be made when plants are both watered and droughted), they are rarely employed.

Previous research by Hancock et al. (in review) explored the expression and evolution of CAM within Australian *Calandrinia* (Montiaceae), a lineage of ~ 70 species endemic to Australia. Characterization of photosynthesis using physiological tools (i.e. titratable acidity measurements of nocturnal CO₂ uptake) across 22 species showed that Australian *Calandrinia* is a lineage of plants capable of employing C₃ and various amounts of CAM photosynthesis, however the full constitutive CAM phenotype was not detected in any species investigated. In efforts to include the entire C₃ to CAM spectrum in our analyses of the evolutionary assembly of the CAM phenotype, we expanded species sampling to include lineages with both fully C₃ and fully CAM species (Martin and Wallace, 2000; Arakaki et al., 2011). Photosynthetic and anatomical exploration of CAM evolution within the Portulacineae was also motivated by the fact we now have a well-resolved phylogeny for the lineage (Moore et al., 2017; Goolsby et al., 2018), thus providing us an excellent framework in which to investigate trait evolution.

In efforts to further understand the evolutionary relationships among anatomical traits implicated in the evolution of C₃-CAM and full CAM we quantified leaf succulence, % IAS, cell size, leaf thickness, and venation type (2D or 3D venation arrangement) across 42 species representing the major Portulacineae and Montiaceae lineages plus outgroups (i.e. representatives in the Molluginaceae and Caryophyllales).

Since we had physiological measurements of CAM expression within Australian *Calandrinia*, we sampled more extensively within this lineage. Species included in the analysis span the C₃ to CAM phenotypic spectrum, including a few species that use C₃-C₄ and C₄ photosynthesis. The ancestral character states for all anatomical traits, including photosynthetic mode, were reconstructed across the phylogeny in efforts to assess 1) whether anatomical traits are prerequisites for the evolution of CAM as reported in the Agavodieae (Heyduk *et al.*, 2016), 2) whether C₃-CAM intermediate photosynthesis is a “stepping-stone” along the C₃-to-CAM evolutionary trajectory, and 3) if the CAM phenotype is reversible.

Furthermore, we wanted to understand what climatic factors contribute to the evolution of CAM, C₃-CAM and C₃ in the Portulacineae. Since nocturnal CO₂ uptake dramatically improves water use efficiency (WUE) CAM plants tend to live in hot, arid environments. Hypotheses predict that C₃-CAM intermediates inhabit environments where precipitation is variable or distinctly seasonal (Holtum *et al.*, 2016). These C₃-CAM plants would use different modes of photosynthesis depending on conditions: when water is plentiful plants primarily use C₃ photosynthesis, which allows for more rapid growth, but when water becomes limiting these plants upregulate (switch to) CAM photosynthesis, which provides prolonged carbon gain at a lower water cost. While this theory has sound ecological implications, there has been no thorough analysis of the climatic niche occupied by C₃-CAM intermediate species and how this niche compares with those of C₃ and CAM.

MATERIALS AND METHODS

Phylogenetic Analysis

To address the major objectives of the study, we focused our analyses on outgroup Molluginaceae and Caryophyllales species and on species representatives from all Portulacineae families (except monotypic Halophytaceae) and major Montiaceae lineages. Additional Australian *Calandrinia* taxa were included in phylogenetic analyses as we wanted to combine previous physiological data from Hancock et al. (in review) with anatomical measurements. The combined data set consisted of sequences for 42 taxa from with 5,501 sites. Gene sequences for three chloroplast loci (*matK*, *ndhF*, and *rbcL*) and the nuclear internal transcribed spacer (ITS) region were extracted from sequence data generated through a targeted gene capture approach, as outlined in Moore et al. (2018) and subsequently Goolsby et al. (in press) and Hancock et al. (in review). Gene sequences for individuals not included in these three studies, were downloaded from NCBI. Each gene was aligned separately using default settings in Aliview v.1.19 (Larsson, 2014) and concatenated with Phyutility (Smith and Dunn, 2008). Phylogenetic trees were inferred with maximum likelihood using RAxML (Stamatakis, 2014) for 1000 bootstrap replicates under a GTR + model of sequence of evolution and separate partitions for each gene (Fig. 1).

Leaf Anatomy

Leaf tissue was harvested from plants growing in the field, the Brown University greenhouse (Providence, RI), and the James Cook University greenhouse (Townsville, Australia). Tissue was initially fixed in Formaldehyde-acetic acid alcohol (FAA), Karnovsky's solution (paraformaldehyde, glutaraldehyde, and sodium cacodylate buffer), or ~ 70% ethanol depending on where and when the leaf tissue was collected (e.g., most

tissue collected in the field was placed in ~90% ethanol as FAA and Karnovsky's were not available. Tissue placed in ethanol while in the field was transferred to FAA or Karnovsky's once back in the lab). Australian *Calandrinia* leaf samples were sent to the University of Georgia Veterinary Histology Laboratory, where the leaf tissue was embedded in agarose, sliced with a microtome, mounted, and stained with Tulouidine blue. The remaining Portulacineae, Molluginaceae, and outgroup leaf samples were processed for anatomical investigation according to Ogburn and Edwards 2013. Slides were imaged with a light microscope (Nikon Eclipse E600) and mounted digital camera (Nikon DXM1200C).

All traits were measured independently for each cross section (e.g., Fig 2). We used both Photoshop and ImageJ V# (National Institutes of Health, USA) to make anatomical measurements for each section. If there were multiple sections per species, traits were averaged for each species. To measure intercellular airspace (IAS), all sections were transferred into Adobe Photoshop 2017.1.0 and the IAS filled with green pigment ($r=36$, $b=234$, $g=25$). If there was differentiation between the mesophyll and palisade cells, the IAS around the two cell types was colored differently: the palisade IAS remained green ($r=36$, $b=234$, $g=25$) while the mesophyll IAS was transformed to a magenta ($r=198$, $g=112$, $b=119$). The mesophyll, palisade and total IAS were calculated by dividing the respective IAS values by the total area. Cell size was measured for each cross-section. The area of the ten largest cells in each cross section was measured and an average taken. If the cell types were differentiated, we measured the top 10 largest cells in each layer and took the overall average of those 20 cells, to get the average cell size for the cross section. Leaf thickness was measured (a general correlate of leaf succulence) by

taking the average of at least three abaxial to adaxial measurements made towards the middle, thickest part of the leaf. Finally, we characterized venation type (i.e. 2D, 3D type 1, and 3D type 2) for each species (Ogburn and Edwards, 2013).

Stable isotope analysis

The $\Delta^{13}\text{C}$ values of dried, finely-ground leaves collected from whole plants were measured using a Finnigan Delta plusXL mass spectrometer housed at both Brown University (Providence, RI) and Boston University (Boston, MA). For those species in which we did not have $\delta^{13}\text{C}$ measurements, we collected measurements through a literature search.

Titrateable acidity

The titrateable acidity measurements for Australian *Calandrinia* were generated by Hancock et al. 2018. Briefly, these ΔH^+ values result from a drought experiment implemented at the James Cook University greenhouse (Townsville, AU) during the spring of 2015. Leaves were harvested from control and experimental plants at dusk on day 1 and dawn on day 2 (well-watered) and at dusk on day 11 and dawn on day 12 (drought). Harvested leaves were immediately measured for fresh weight and were then frozen in liquid nitrogen. Organic acids were extracted by boiling leaf tissue in ~ 80ml of 50% ethanol in distilled water. Once half of the total volume boiled off, an additional ~40mL of distilled water was added to each beaker. Beakers were removed from the hotplates once the final volume reached 40mL. The H^+ pools were determined by measuring the volume of 5 mM NaOH required to neutralize each sample to pH 6.5. To calculate nocturnal increase in tissue acidity (ΔH^+) we subtracted the titrateable acidity (H^+) at the end of the light period from the titrateable acidity at the end of the dark period

(H⁺). ΔH^+ was then averaged for each species when plants were watered and droughted. A student's t-test was performed to determine if PM and AM ΔH^+ values were significantly different and greater than zero.

Locality and climate data

Georeferenced records for the 42 species included in the study were downloaded from the Australian Virtual Herbarium (<http://avh.chah.org.au/>) and from the Global Biodiversity Information Facility (www.gbif.org). We followed a number of steps to clean these records (i.e., records with identical coordinates for a given taxon), records with latitude and longitude coordinates of 0, 0, and records that matched the lat./ long. of herbaria and botanical gardens (Edwards *et al.*, 2015); https://github.com/ejedwards/reanalysis_zanne2014). For those remaining records, 19 bioclimatic variables were downloaded from the WorldClim database (<http://www.worldclim.org/bioclim>) using the raster package in R (Hijmans and van Etten, 2012) with a medium-scale spatial resolution of 2.5 min of a degree. Species means were calculated for each variable for further comparative analyses.

Since bioclim variables bio3, bio5, bio7 co-vary, we dropped them from our climate dataset. We conducted a linear discriminate analysis (LDA) using the *lda* function in the R package *MASS* (Venables and Ripley 2002) to find the linear combinations of the remaining 16 bioclim variables that best separated the photosynthetic phenotypes C₃, C₃-CAM, CAM. Based on the LDA and the fact that certain climatic variables have been shown to have ecological consequences for photosynthetic mode (Edwards and Ogburn, 2012), we chose to focus further climate analyses on mean annual temperature (MAT), mean annual precipitation (MAP), and precipitation seasonality.

Photosynthesis designations and ancestral character state reconstructions

All phylogenetic analyses were performed with the 42-taxon maximum likelihood phylogeny. In efforts to identify associations between morphological and anatomical variation and photosynthetic pathway, we classified all species into photosynthetic phenotypes. We based phenotype categorization on published literature (Winter, 1979; Martin and Wallace, 2000; Guralnick and Jackson, 2001; Guralnick et al., 2008; Ocampo and Columbus, 2010; Winter and Holtum, 2012; Holtum et al., 2017a; Guralnick 2017, Goolsby et al 2018, Hancock et al. 2018) and on our own unpublished physiological data (i.e., drought experiments, isotope surveys). When lacking either of these sources, in the case of a few species, we used plant morphology (succulence, or lack thereof), known habitat, and phylogenetic proximity to other photosynthetic modes. We classified included species into C₃, C₃-CAM, and CAM. Although there are known phenotypic differences within the C₃-CAM category (Winter and Holtum, 2014; Winter *et al.*, 2015), we chose to bin these phenotypes since further designation requires detailed physiological work for all species (Hancock et al. in review).

To reconstruct ancestral character states at all internal nodes for photosynthesis type, we used ML reconstruction in the *APE* package as implemented in the R Project for Statistical Computing (Paradis *et al.*, 2004). We constrained the model to equal transition rates between states and then allowed the rates to vary symmetrically (SYM), differently (ARD), and wherein transitions occur in a stepwise fashion (order). We chose the model with the lowest AIC value when reconstructing ancestral character states. We used the same approach as previously outlined to reconstruct venation type (i.e. 2D, 3D type 1, and 3D type 2) across the phylogeny. Maximum likelihood ancestral state reconstructions

for continuous traits (% IAS, cell size, leaf thickness, $\Delta^{13}\text{C}$, MAT, MAP and precipitation seasonality) were performed with the *contMap* function in the R package *Phytools* (Revell, 2012).

Trait correlations

To assess pairwise macro-evolutionary correlations among continuous traits and between continuous traits and climatic variables we used the phylogenetically independent contrasts (*pic*) function in the *APE* package (Paradis *et al.*, 2004). Given the large number of pairwise correlations being estimated, we implemented the method of Benjamini and Hochberg (Benjamini and Hochberg, 1995) to control for false discovery rate, i.e. the expected proportion of false discoveries among rejected hypotheses. To assess trait differences between C_3 and CAM (which includes all C_3 -CAM species) photosynthesis and venation type (2D or 3D), we performed a phylogenetic ANOVA using the *phylANOVA* function in the R package *Phytools* (Revell, 2012). To adjust for p-values, we specified 2000 simulations (random permutation of data at the tips) and performed posthoc tests using the BH method. To assess correlation between discrete binary traits presence/ absence of a CCM vs. 2D or 3D venation, we used the *fitPagel* function in the R package *Phytools* (Revell, 2012) and the *corDisc* function in the R package *corHMM*.

RESULTS

Phylogenetic relationships

The 42-taxa, four-marker combined analysis of *matK*, *rbcL*, *ndhF*, and *ITS* produced a generally well-supported tree for the portullugo (Portulacineae +

Molluginaceae). The backbone topology is mostly concordant with previous phylogenetic analyses of the portullugo (Moore *et al.*, 2017), albeit lower bootstrap support for many relationships. The main difference between the Moore et al. (2018) portullugo phylogeny and the four-marker ML tree (Fig 1) is that Molluginaceae is not monophyletic; this discordance likely results from the fact that Molluginaceae is distantly related to *Limeum* as well as limited sequence data. As shown in Ogburn and Edwards 2015, we recovered *Cistanthe* as paraphyletic, with *Cistanthe s.s.* sister to *Montiopsis* and *Cistanthe salsoloides* (of the *Phillippiamra Cisthanthe* lineage) as sister to *Calyptridium umbellatum*. The major relationships among Australia *Calandrinia* and the rest of the Montiaceae are consistent with Hancock et al. in review.

Comparative evolutionary analyses

The only “full” CAM plant, *Quiabentia verticillata* (Martin and Wallace 2000) had the thickest leaves (4.49mm), largest cells (0.0561 mm²) and lowest IAS (0.06%) of any other species included in the study (Table 3; Fig 4b-d). The leaf anatomy for species that fall within the C₃-CAM photosynthetic category have leaves that are less thick (1.34 ± 0.67), with smaller cells (0.0151mm² ± 0.0010), and more IAS (3.47% ± 2.75) than the full CAM plant. The values for these traits are significantly different in C₃ species, which have as far smaller cells (0.0036mm² ± 0.0035), thinner leaves, (0.7 ± 0.34) and a greater percentage of IAS (9.7% ± 5.4).

Macro-evolutionary comparative analyses between continuous traits (i.e., % IAS, cell size, leaf thickness and $\Delta^{13}\text{C}$), using phylogenetically independent contrasts, indicate that many traits are tightly correlated (Fig 3). Plants with thicker leaves have significantly larger cells (p-value=0.02) and a lower percentage of IAS (p-value <0.001); interestingly,

no significant relationship was found between percent IAS and cell size (p-value=0.43). Species with higher, more positive $\Delta^{13}\text{C}$ values (i.e. more CAM-like) had significantly thicker leaves (p-value <0.001) and larger cells (p-value=0.01). Australian *Calandrinia* with higher ΔH^+ values tend to have larger cells (p-value=0.049), though the correlation between these variables was no longer significant after adjusting for p-value (p-value=0.195).

Since we only had one full CAM plant in our study, we grouped CAM with C_3 -CAM for further comparisons between photosynthesis type and continuous morphological and anatomical traits. Phylogenetic ANOVAs showed a significant relationship between C_3 -CAM and leaf thickness, cell size, and % IAS; plants with the capacity to do CAM have significantly thicker leaves (p-value=0.074), larger cells (p-value=0.047), and a lower percentage of IAS (p-value=0.01) (Fig 3, Fig 4b-d). $\Delta^{13}\text{C}$ values were not significantly correlated with photosynthetic mode (p-value=0.23) (Fig 3; Fig S1). Although $\Delta^{13}\text{C}$ values guide us to phylogenetic lineages in which to investigate and further explore CAM, they often fail to uncover C_3 -CAM intermediates, which often present C_3 like $\Delta^{13}\text{C}$ values. Although succulent water storage is a prominent feature among plants that use CAM, we found no significant link between SWC and photosynthetic mode or any other trait, for that matter.

All three venation types (2D, 3D type 1, and 3D type 2) were found within the Portulacineae (Table 1). For comparative purposes we binned 3D type 1 and 3D type 2 into the same category: 3D venation. 3D venation was strongly correlated with both cell size (p-value = 0.002) and leaf thickness (p-value = 0.003). Thus, species with thicker leaves and larger cells were far more likely to have 3D venation. A previous hypothesis

suggests that extreme succulence could only evolve by developing 3D venation to reset internal leaf hydraulic distances, maintaining moderate vein density in succulent tissues (Ogburn and Edwards, 2013).

Discrete ancestral character state reconstructions (photosynthesis and venation)

The ancestral state of the Portulacineae was reconstructed as C₃-CAM under the preferred ER model of evolution with multiple reversions back to C₃: once in Australian *Calandrinia*, once at the base of Montieae (includes *Claytonia*, *Lewisia*, and *Montia*) and twice in the *Montiopsis*, *Calyptridium* and *Cistanthe* clade. A shift to full CAM in the core Cactaceae and C₄ in the Portulacaceae was detected (Fig 4a). The base of the Portulacineae was reconstructed as 2D under the slightly preferred ordered (meristic) model of evolution (Fig 4a). There have been multiple shifts to 3D type 1 venation: three times in Australian *Calandrinia* and once in *Phemeranthus*, *Talinaceae*, *Portulacaceae*, and *Cactaceae*. There has been one shift to 3D type 2 venation at the base of Anacampserotaceae. Only one shift back to 2D venation was detected, in *Calandrinia ptychosperma*. Added phylogenetic and character resolution within the Portulacineae would likely provide further insight into venation and photosynthesis evolution, likely altering our phylogenetic reconstructions and inferred losses and gains of photosynthetic mode and venation type. Phylogenetic comparisons using both corHMM (Beaulieu et al., 2013) and phytools (Revell, 2012) between photosynthesis mode and venation type showed no significant relationship, as many of our C₃-CAM species exhibit typical 2D venation (Fig 3).

Climate data

An LDA analysis using 16 climatic variables with plants phenotyped as C₃, C₃-CAM, and CAM showed that LD1 explained 82.9% of the variation and LD2 17.1% of the variation (Fig 5). The variation in LD1 was mostly explained by MAT, max temperature of warmest month, and mean temperature of warmest quarter. The variation in LD2 was explained by MAT and max temperature of the warmest month. Although precipitation (MAP) and precipitation seasonality did not discriminate phenotypes, these two climatic variables have been shown and or hypothesized to have ecological consequences for photosynthetic mode (Herrera, 2009; Holtum et al., 2016), therefore we included these two variables, along with MAT, max temperature of the warmest month, and mean temperature of warmest quarter in further analyses.

Phylogenetically independent contrasts showed that MAT, max temperature of warmest month, and mean temperature of warmest quarter were strongly correlated (p-value <0.001 for all three variables). Because of this we chose to focus further PICs on MAT (as it showed the strongest discrimination in the LDA analysis) MAP, and precipitation seasonality. We found that MAT, after adjusting for p-value, was significantly correlated with leaf thickness (p-value=0.02) and cell size (p-value=0.017). MAP was not significantly correlated with any of the traits included when p-value was adjusted for, though it continued to show a strong relationship with leaf thickness, cell size, and $\delta^{13}\text{C}$. Precipitation seasonality showed a strong relationship with an increase in cell size, leaf thickness, and MAT. After adjusting for p-value, however, only leaf thickness (p-value=0.017) and MAT (p-value=0.012) had a significant relationship with precipitation seasonality. Phylogenetic ANOVAs between climatic variables and

photosynthesis mode and venation type revealed a strong correlation between MAT and photosynthesis mode (p-value = 0.013) and venation type (p-value <0.001).

To visualize the climate space of C₃, C₃-CAM, and CAM species, we mapped mean species values for MAT, MAP and precipitation seasonality onto the tips of the 42-tip ML tree (Fig 6a, 6b; Fig S2). The ancestral reconstruction of these climatic variables allowed us to further investigate the relationship between climate and photosynthetic mode. For example, the shift from C₃-CAM to C₃ at the base of Montieae and *Montiopsis* coincides with a shift to colder and somewhat wetter temperatures (Fig 6a, 6b), while the shift to full CAM in *Quiabentia verticillata* corresponds to a shift towards hotter temperatures and less precipitation (Fig 6a, 6b). Species with CAM capacity live in significantly hotter climates than C₃ species. Although there is no strong relationship between MAP and photosynthesis mode, species with the ability to do at least low-level CAM live in the driest regions (e.g., *Ceraria fruticulosa*, *Cistanthe salsoloides*, *Cistanthe picta*, and *Calandrinia schistorhiza*).

DISCUSSION

Untangling the evolution of CAM requires piecing together the genetic, biochemical, anatomical, physiological, and ecological aspects of this complex trait and reconstructing the evolutionary order of its assembly. In this study, we focus on a few of these properties, specifically the anatomy, morphology, and ecology associated with CAM evolution. Previous research on photosynthesis expression and evolution within the Portulacineae has focused primarily on molecular evolution and phenotyping species (Martin et al., 1988; Arroyo et al., 1990; Martin and Wallace, 2000; Rundel et al., 2003;

Guralnick et al., 2007; Winter and Holtum, 2014; Holtum et al., 2016; Guralnick and Gladsky, 2017; Goolsby et al., 2018). With this foundational research in place, and a well-resolved phylogeny for the lineage (Moore et al., 2018), we are poised to investigate and tease apart how CAM evolved within the Portulacineae; answering specific questions about functional constraints associated with photosynthesis evolution, the evolutionary significance C₃-CAM intermediate phenotypes, whether CAM is reversible or not, and if anatomical changes are a prerequisite for photosynthesis evolution.

Anatomical evolution

The correlation between anatomical traits and photosynthetic pathway indicate that specific modifications such as an increase in cell size and leaf thickness and a reduction in IAS are associated, and perhaps requisite, for CAM function (both full CAM and C₃-CAM). These trait modifications have been observed across a diversity of CAM lineages (e.g., Asclepiadaceae, Aizoaceae, Agavaceae, Clusiaceae, Portulacineae), further indicating that CAM anatomy is constrained by functional requirements that promote CAM efficiency (Maxwell et al., 1997; Nelson et al., 2005; Nelson and Sage, 2007; Heyduk et al., 2016). In contrast, species that use only C₃ photosynthesis have significantly smaller cells, a greater percentage of IAS, and thinner leaves (Fig 4b-4d). Given this dichotomy in leaf anatomy and what we know about the biochemistry and physiology of these two pathways, certain trait modifications clearly promote efficiency in either C₃ or CAM.

Species that do both C₃ and CAM photosynthesis (i.e. C₃-CAM) have trait values for cell size, leaf thickness, and percent IAS that generally fall in between C₃ and full CAM values (Table 2). Since most of these species primarily use the C₃ pathway for

carbon fixation, generally upregulating and/ or switching to CAM with the onset of stress, maintaining anatomical features that promote C₃ photosynthesis but aid CAM seems optimal. Furthermore, plants that do more CAM have more CAM-like anatomy: somewhat larger cells and thicker leaves, as demonstrated by ΔH^+ values within Australian *Calandrinia*. Given that an increase in cell size generally means an increase in vacuole size (a.k.a. the ability of a plant to store malic acid overnight) it is not surprising that these physiological measurements of nocturnal acid accumulation (ΔH^+) would have an anatomical correlation.

Beyond the Australian *Calandrinia* example, we know from previous research that many C₃-CAM species within the Anacampersotaceae and Didiereaceae engage in more CAM photosynthesis than many C₃-CAM species within the Basellaceae and Montieaceae (e.g., *Phemeranthus* and *Calandrinia*) (Guralnick et al., 2007; Winter and Holtum, 2014; Holtum et al. 2017; 2016; Guralnick and Gladsky, 2017). Both anatomical measurements (representatives in Fig 2b, 2c) and ancestral character state reconstructions (Fig 4b-d) indicate that these species that do more CAM have more CAM-like anatomy: larger cells, less intercellular air space, and thicker leaves. These findings coincide with early work by Nelson et al. (2007), in which the authors showed that species from across a diversity of lineages showed similarities in trait evolution associated with the degree of CAM expression. Further physiological investigation into these species that span the C₃ to CAM phenotypic space and represent the known C₃-CAM intermediate phenotypes (e.g., low-level CAM and facultative CAM) would greatly improve our understanding of how subtle changes in anatomy potentially facilitate shifts in photosynthesis.

Cell size and intercellular airspace evolution

CAM species typically have photosynthetic tissue with large, highly vacuolated cells (Winter and Smith 1996). Large vacuoles are an integral component of the CAM pathway as they provide both capacitance and nocturnal storage of C4 acids (i.e. malic acid). The larger the vacuole, the more CO₂ that can be taken up by PEPc and stored overnight, thereby directly increasing carbon gain in CAM species. An increase in vacuole size typically means an increase in cell size (Kluge and Ting, 1978). Although an increase in cell size can result in more tightly packed cells and an overall reduction in % IAS, we did not find a significant relationship between cell size and % IAS. What we did uncover were two complimentary significant relationships: between cell size and CAM, and between % IAS and CAM. Since there is no physical reason for IAS to correlate tightly with cell size, due to the wide range of potential cell geometries, the finding that both these traits correlate with CAM suggests that they are both selected for. Similar studies have found the same patterns: Nelson and Sage (2007) and Heyduk et al. (2016) find that CAM plants have larger cells and a lower percentage of IAS, while IAS does not correlate with cell size. They also suggest that that low IAS is not simply a bi-product of larger cells being more tightly packed, but is selected for amongst these plants to promote CAM efficiency and carbon economy (Nelson and Sage, 2007). Clearly, shifts in in these key anatomical traits facilitated the evolution of CAM within the Portulacineae. What remains unknown is at what point did shifts in these traits facilitate transitions from C₃-CAM to full CAM?

Climate space and photosynthetic mode

Over broad geographical scales the distribution of CAM is highly influenced by aridity (Ehleringer and Monson, 1993). This makes sense, not only does CAM increase

water use efficiency (a.k.a. the ratio of carbon fixed to water lost through transpiration), but in most cases CAM plants are succulents with increased water storage tissue. Plants that use C₃-CAM photosynthesis are hypothesized to inhabit arid/ semi-arid environments where precipitation is variable or distinctly seasonal (Herrera, 2009). Under such environmental conditions, plants use C₃ photosynthesis when water is available, which allows for higher rates of CO₂ fixation and vegetative growth than CAM. When water becomes limiting plants switch to CAM, which provides prolonged net carbon gain at a lower water cost, enabling a higher reproductive output under these conditions (Winter and Ziegler, 1992).

Both LDA and phylogenetic ANOVAs did not uncover a relationship between climatic variables MAP and precipitation seasonality with the occurrence of CAM capacity (Fig 3; Fig 5). However, an increase in leaf thickness and cell size was correlated with MAP and precipitation seasonality (Fig 3). As leaf thickness is a general proxy for succulence and an increase in cell size reflects not only CAM capacity but an increase in water storage, these trait modifications (a.k.a. succulence) likely enable species to inhabit drier, more seasonable climates (Ogburn and Edwards, 2010). Although these analyses indicate that C₃-CAM does not confer any ecological advantage to prolonged or intermittent aridity, such climate measurements using the WorldClim database are broad-scale and do not capture the actual micro-climates in which these plants live (Ogburn and Edwards, 2015). For example, many of the C₃-CAM Australian *Calandrinia* species included in this study germinate following either winter or summer rainfall or grow along the sandy edges of ephemeral water sources. These species typically complete their life-cycle in 6 to 8 weeks, during which time the soil slowly dries

out or the water subsides. Only towards the end of their lives do these species contend with water limitation (Hancock et al. 2018). Such subtleties in climate and habitat are not detected using broad-scale climate data. Furthermore, we suspect that the inclusion of Montieae, which grow in particularly cold regions, may be masking our ability to detect precipitation as an important discriminating variable in photosynthetic climate space.

We found a significant relationship between MAT and CAM capacity, with C₃-CAM species occupying significantly hotter regions than C₃ plants. The relationship between temperature and photosynthesis is well known: the rates at which carbon is fixed increase with temperature until an optimal value is reached (Kluge and Ting, 1978; Osmond, 1978). In C₃ plants, once this optimal temperature is exceeded the CO₂ to O₂ ratio will decrease and Rubisco will bind to O₂ (Rubisco is both a carboxylase and oxygenase), resulting in photorespiration. In contrast, C₄ and CAM plants have the biochemical and physiological tools to continue assimilating carbon at higher temperatures, avoiding photorespiration, thus enabling them to colonize hotter localities. The ability of a plant to engage in CAM when temperatures are high, whether it be at a low-level or facultatively, must provide an added physiological value. This value may be in the form of increased internal CO₂ concentrations around the active site of Rubisco, as these C₃-CAM plants are storing C₄ acids overnight for subsequent daytime decarboxylation. Having this small but potentially significant rise in internal CO₂ concentration may enable plants to close their stomata for longer periods of time, potentially reducing both photorespiration and water loss through transpiration. Since the evolution of full CAM allows species to inhabit hot climates, it is not surprising that the

evolution of C₃-CAM, even if the expression of CAM is not as pronounced as full CAM, confers a selective advantage to species living in hotter climates.

CAM evolution and reversibility in the Portulacineae

Elaboration of the evolutionary dynamics of CAM has been hindered because there is no well-developed model lineage with complete phenotypic and phylogenetic resolution. Although considerable effort has been expended to detect and characterize CAM in lineages such the Orchidaceae (Silvera *et al.*, 2005; 2009), Bromeliaceae (Ceusters *et al.*, 2010; Crayn *et al.*, 2015) and Euphorbiaceae (Horn *et al.*, 2014), these lineages are extremely species rich, making detailed phenotypic and phylogenetic work a daunting task. Recently, Heyduk *et al.* (2016) investigated the evolution of CAM within the Agavoideae (~640 species), a lineage where CAM has evolved multiple times from C₃ ancestors. After characterizing anatomical and morphological traits across 17 species, the authors suggest that CAM anatomy likely developed prior to the CAM pathway, as there is considerable anatomical similarity between C₃ and CAM species. Although the anatomy-first hypothesis may very well be true for the Agavoideae, our analyses do not suggest that anatomical evolution predates the origin of the CAM pathway, particularly full CAM, within the Portulacineae.

Based on ancestral character state reconstructions, C₃-CAM photosynthesis evolved at the base of the Portulacineae (Fig 4a). Although C₃-CAM refers to species that primarily fix carbon with the C₃ pathway, the CAM pathway is present whether it be in the form of CAM-cycling, constitutively expressed at a low level, or induced with the onset of stress (Brilhaus *et al.*, 2016). Recently Heyduk *et al.* (2018) used a transcriptomic approach to investigate the gene expression profiles associated with C₃-

CAM and full CAM photosynthesis within the Agavoideae. The two C₃-CAM species in the study had gene expression patterns that fit the full CAM profile: the upregulation of key genes associated with carbon fixation (e.g., PEPC and PPCK), decarboxylation, and carbohydrate metabolism. What differentiated full CAM from C₃-CAM was a greater transcript abundance of these key genes and shifts in regulatory gene function associated with full CAM (Heyduk et al., 2018). These findings, along with previous molecular work assessing the genetic and biochemical underpinnings of both CAM and C₃-CAM (Ming et al., 2015; Brilhaus et al., 2016), further illustrate the functionality of the CAM pathway within these C₃-CAM species.

Given this recent work on C₃-CAM gene expression and our ancestral character state reconstructions (Fig 4a), we hypothesize that the CAM pathway evolved prior to many of the anatomical and morphological changes associated with the realization of the full CAM phenotype (Fig 2d; Fig 4a). As more CAM photosynthesis was selected for, as seen in the full CAM plant *Quiabentia verticilla*, the anatomical and morphological changes that enable this level of CAM expression evolved in concert; although trait modifications such as larger cells with bigger vacuoles may have facilitated and enabled this shift to full CAM, the CAM pathway, was already present, either completely or nearly completely. Adding more full CAM plants to this study, particularly within early branching Cactaceae lineages, would greatly improve resolution in the anatomical and morphological trait space between C₃-CAM and full CAM photosynthesis. Such resolution would enable us to tease apart how and why the Cactaceae were able, unlike most of the Portulacineae, to evolve full CAM. We suspect getting to this full CAM phenotype, which includes extreme succulence and a perennial life history strategy, is a

bit of a bottleneck and that once a plant gets there, reversibility to C_3 is nearly impossible. Although there have been reports of full CAM reversion in (Silvera et al., 2010), phylogenetic and physiological sampling within these studies was insufficient.

The C_3 -CAM phenotype, unlike the full CAM phenotype, is clearly reversible within the Portulacineae. The three detected shifts to C_3 from C_3 -CAM occurred within the Montiaceae, the family sister to the rest of the Portulacineae. The loss of CAM within these C_3 species significantly corresponds to anatomical and morphological shifts that facilitate C_3 photosynthesis as opposed to CAM photosynthesis (Fig 4b-4d). It is unclear whether the switch to C_3 facilitated ecological transitions to more mesic habitats (Fig 6) or if the loss of C_3 -CAM in these species was in response to climatic and habitat changes in which maintaining CAM was no longer advantageous. It should be noted that for some of these C_3 species we do not have detailed physiological measurements and so we are basing photosynthetic phenotyping on $\delta^{13}C$ values, degree of succulence, and habitat (including elevation). It is feasible that these species do some CAM under the right conditions or have a “vestigial” CAM pathway, but at this time we feel confident that they are C_3 species.

Hypotheses on the evolution of CAM

The traditional evolutionary concept of C_3 -to-CAM evolution parallels that of C_3 -to- C_4 evolution: a model which starts with C_3 photosynthesis and progresses in a step-wise manner towards the derived photosynthetic state (i.e. CAM or C_4) (Sage, 2004). In the case of C_3 -to-CAM evolution, these transitional steps are populated by C_3 -CAM phenotypes. An alternative hypothesis is that C_3 -CAM phenotypes and full CAM represent different endpoints along distinct evolutionary trajectories, with each trajectory

reflecting interactions between environmental selection and the genetic makeup of the ancestral taxa (Winter & Holtum 2014, Winter et al 2015). These data show that not only are C₃-CAM phenotypes anatomically intermediates between C₃ and full CAM, but that they are also intermediate states along the full CAM trajectory: C₃-CAM evolves prior to full CAM. However, given the number of C₃-CAM reversals we propose a model of CAM evolution that is not linear (i.e. C₃ to C₃-CAM to CAM) but allows for lability and reversibility among C₃-CAM phenotypes and C₃ photosynthesis. The lability of this model has one caveat; reversion from full CAM has never been confidently discovered (i.e. using physiological tools within a robust phylogenetic analysis) and we suspect that once a species evolves full CAM, with the associated anatomical changes such as extreme succulence and very little IAS, reversibility back to C₃ or C₃-CAM photosynthesis may no longer be evolutionarily accessible. Hence, full CAM may represent a tipping point at which evolved traits are "locked in" and cannot reverse to C₃-CAM or C₃ photosynthesis.

As a final note, we want to leave open the possibility that the C₃-CAM phenotypes, especially facultative CAM, could be end states in separate evolutionary trajectories. Further analysis with higher phenotypic resolution of these C₃-CAM phenotypes is required to investigate this possibility.

CITATIONS

- ARAKAKI, M., P.A. CHRISTIN, R. NYFFELER, A. LENDEL, U. EGGI, R.M. OGBURN, E. SPRIGGS, ET AL. 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. *Proceedings of the National Academy of Sciences* 108: 8379–8384.
- ARROYO, M.K., E. MEDINA, AND H. ZIEGLER. 1990. Distribution and $\delta^{13}\text{C}$ Values of Portulacaceae Species of the High Andes in Northern Chile. *Botanica Acta* 103: 291–295.
- 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: 725–737.
- BENJAMINI, Y., AND Y. HOCHBERG. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological* 57: 289–300.
- BRILHAUS, D., A. BRÄUTIGAM, T. METTLER-ALTMANN, K. WINTER, AND A.P.M. WEBER. 2016. Reversible Burst of Transcriptional Changes during Induction of Crassulacean Acid Metabolism in *Talinum triangulare*. *Plant Physiology* 170: 102–122.
- CEUSTERS, J., A.M. BORLAND, N. CEUSTERS, V. VERDOODT, C. GODTS, AND M.P. DE PROFT. 2010. Seasonal influences on carbohydrate metabolism in the CAM bromeliad *Aechmea* “Maya”: consequences for carbohydrate partitioning and growth. *Annals of Botany* 105: 301–309.

- CRAYN, D.M., K. WINTER, K. SCHULTE, AND J.A.C. SMITH. 2015. Photosynthetic pathways in Bromeliaceae: phylogenetic and ecological significance of CAM and C₃ based on carbon isotope ratios for 1893 species. *Botanical Journal of the Linnean Society* 178: 169–221.
- EDWARDS, E.J., AND R.M. OGBURN. 2012. Angiosperm Responses to a Low-CO₂ World: CAM and C₄ Photosynthesis as Parallel Evolutionary Trajectories. *International Journal of Plant Sciences* 173: 724–733.
- EDWARDS, E.J., J.M. DE VOS, AND M.J. DONOGHUE. 2015. Doubtful pathways to cold tolerance in plants. *Nature* 521: E5–6.
- EHLERINGER, J.R., AND R.K. MONSON. 1993. Evolutionary and ecological aspects of photosynthetic pathway variation. *Annual Review of Ecology and ...* 24: 411–439.
- FARQUHAR, G.D., AND J.R. EHLERINGER. 1989. Carbon isotope discrimination and photosynthesis. 503–537.
- GOOLSBY, E., A.J. MOORE, L.P. HANCOCK, J. DE VOS, AND E.J. EDWARDS. 2018. Molecular evolution of photosynthesis genes during transitions to C₄ and CAM metabolism. *American Journal of Botany*, in press.
- GURALNICK, L.J., A. CLINE, M. SMITH, AND R.F. SAGE. 2007. Evolutionary physiology: the extent of C₄ and CAM photosynthesis in the genera *Anacampseros* and *Grahamia* of the Portulacaceae. *Journal of Experimental Botany* 59: 1735–1742.
- GURALNICK, L.J., AND K. GLADSKY. 2017. Crassulacean acid metabolism as a continuous

trait: variability in the contribution of Crassulacean acid metabolism (CAM) in populations of *Portulacaria afra*. *Heliyone*00293.

GURALNICK, L.J., AND M.D. JACKSON. 2001. The Occurrence and Phylogenetics of Crassulacean Acid Metabolism in the Portulacaceae. *International Journal of Plant Sciences* 162: 257–262.

HERRERA, A. 2009. Crassulacean acid metabolism and fitness under water deficit stress: if not for carbon gain, what is facultative CAM good for? *Annals of Botany* 103: 645–653.

HEYDUK, K., M.R. MCKAIN, F. LALANI, AND J. LEEBENS-MACK. 2016. Evolution of a CAM anatomy predates the origins of Crassulacean acid metabolism in the Agavoideae (Asparagaceae). *Molecular Phylogenetics and Evolution* 105: 102–113.

HEYDUK, K., N.J. RAY, S. AYYAMPALAYAM, AND J. LEEBENS-MACK. 2018. Shifting gene expression profiles associated with weak and strong Crassulacean acid metabolism . *American Journal of Botany*, in press.

HIJMANS, R., AND J. VAN ETTEN. 2012. Geographic analysis and modeling with raster data.

HOLTUM, J., J. ARANDA, A. VIRGO, H.H. GEHRIG, AND K. WINTER. 2004. δ ¹³C values and crassulacean acid metabolism in *Clusia* species from Panama. *Trees* 18: 658–668.

HOLTUM, J.A., L.P. HANCOCK, E.J. EDWARDS, M.D. CRISP, D.M. CRAYN, R. SAGE, AND

- K. WINTER. 2016. Australia lacks stem succulents but is it depauperate in plants with crassulacean acid metabolism (CAM)? *Current opinion in plant biology* 31: 109–117.
- HOLTUM, J.A.M., J.A.C. SMITH, AND H.E. NEUHAUS. 2005. Intracellular transport and pathways of carbon flow in plants with crassulacean acid metabolism. *Functional Plant Biology* 32: 429–449.
- HOLTUM, J.A.M., L. HANCOCK, E.J. EDWARDS, AND K. WINTER. 2016. Facultative CAM photosynthesis (crassulacean acid metabolism) in four species of Calandrinia, ephemeral succulents of arid Australia . *Photosynthesis research* 1–21.
- HORN, J.W., Z. XI, R. RIINA, J.A. PEIRSON, Y. YANG, B.L. DORSEY, P.E. BERRY, ET AL. 2014. Evolutionary bursts in Euphorbia (Euphorbiaceae) are linked with photosynthetic pathway. *Evolution* 68: 3485–3504.
- KLUGE, M., AND I.P. TING. 1978. Crassulacean acid metabolism. Analysis of an ecological adaptation.
- LARSSON, A. 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30: 3276–3278.
- MARTIN, C.E., AND A.K. ZEE. 1983. C3 photosynthesis and crassulacean acid metabolism in a Kansas rock outcrop succulent, *Talinum calycinum* Engelm.(Portulacaceae). *Plant Physiology*.
- MARTIN, C.E., AND R.S. WALLACE. 2000. Photosynthetic pathway variation in leafy

- members of two subfamilies of the Cactaceae. *International Journal of Plant Sciences* 161: 639–650.
- MARTIN, C.E., M. HIGLEY, AND W.Z. WANG. 1988. Ecophysiological significance of CO₂-recycling via Crassulacean acid metabolism in *Talinum calycinum* Engelm.(Portulacaceae). *Plant Physiology*.
- MAXWELL, K., S.V. CAEMMERER, AND J.R. EVANS. 1997. Is a Low Internal Conductance to CO₂ Diffusion a Consequence of Succulence in Plants with Crassulacean Acid Metabolism? *Australian Journal of Plant Physiology* 24: 777.
- MING, R., R. VANBUREN, C.M. WAI, H. TANG, M.C. SCHATZ, J.E. BOWERS, E. LYONS, ET AL. 2015. The pineapple genome and the evolution of CAM photosynthesis. *Nature Publishing Group* 47: 1435–1442.
- MOORE, A.J., J. DE VOS, L.P. HANCOCK, E. GOOLSBY, AND E.J. EDWARDS. 2017. Targeted Enrichment of Large Gene Families for Phylogenetic Inference: Phylogeny and Molecular Evolution of Photosynthesis Genes in the Portullugo (Caryophyllales). *Systematic Biology*.
- NELSON, E.A., AND R.F. SAGE. 2007. Functional constraints of CAM leaf anatomy: tight cell packing is associated with increased CAM function across a gradient of CAM expression. *Journal of Experimental Botany* 59: 1841–1850.
- NELSON, E.A., T.L. SAGE, AND R.F. SAGE. 2005. Functional leaf anatomy of plants with crassulacean acid metabolism. *Functional Plant Biology* 32: 409.

- OGBURN, M.R., AND E.J. EDWARDS. 2015. Life history lability underlies rapid climate niche evolution in the angiosperm clade Montiaceae. *Molecular Phylogenetics and Evolution* 92: 181–192.
- OGBURN, R.M., AND E.J. EDWARDS. 2013. Repeated origin of three-dimensional leaf venation releases constraints on the evolution of succulence in plants. *Current biology : CB* 23: 722–726.
- OGBURN, R.M., AND E.J. EDWARDS. 2010. The ecological water-use strategies of succulent plants. *Advances in botanical research*.
- OSMOND, C.B. 1978. Crassulacean acid metabolism: a curiosity in context. *Annual Review of Plant Physiology*.
- PARADIS, E., J. CLAUDE, AND K. STRIMMER. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics (Oxford, England)*.
- REVELL, L.J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*.
- RUNDEL, P.W., A.C. GIBSON, G.S. MIDGLEY, S. WAND, B. PALMA, C. KLEIER, AND J. LAMBRINOS. 2003. Ecological and ecophysiological patterns in a pre-altiplano shrubland of the Andean Cordillera in northern Chile. *Plant Ecology* 169: 179–193.
- SAGE, R.F. 2004. The evolution of C4 photosynthesis. *New Phytologist*.
- SILVERA, K., K.M. NEUBIG, W.M. WHITTEN, N.H. WILLIAMS, K. WINTER, AND J.C. CUSHMAN. 2010. Evolution along the crassulacean acid metabolism continuum.

- Functional Plant Biology* 37: 995–1010.
- SILVERA, K., L.S. SANTIAGO, AND K. WINTER. 2005. Distribution of crassulacean acid metabolism in orchids of Panama: evidence of selection for weak and strong modes. *Functional Plant Biology* 32: 397.
- SILVERA, K., L.S. SANTIAGO, J.C. CUSHMAN, AND K. WINTER. 2009. Crassulacean acid metabolism and epiphytism linked to adaptive radiations in the Orchidaceae. *Plant Physiology* 149: 1838–1847.
- SMITH, S.A., AND C.W. DUNN. 2008. Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics (Oxford, England)* 24: 715–716.
- STAMATAKIS, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)* 30: 1312–1313.
- TING, I.P. 1985. Crassulacean acid metabolism. *Annual Review of Plant Physiology* 36: 595–622.
- WILLIAMS, B.P., I.G. JOHNSTON, S. COVSHOFF, AND J.M. HIBBERD. 2013. Phenotypic landscape inference reveals multiple evolutionary paths to C4 photosynthesis. *eLife* 2: e00961.
- WINTER, K. 1985. Photosynthetic mechanisms and the environment. Elsevier Science Publishers B. V.
- WINTER, K., AND H. ZIEGLER. 1992. Induction of crassulacean acid metabolism in *Mesembryanthemum crystallinum* increases reproductive success under conditions of

drought and salinity stress. *Oecologia* 92: 475–479.

WINTER, K., AND J.A.C. SMITH. 1996. Crassulacean Acid Metabolism. K. Winter, and J. A. C. Smith [eds.],. Ecological Studies, Berlin, Heidelberg.

WINTER, K., AND J.A.M. HOLTUM. 2014. Facultative crassulacean acid metabolism (CAM) plants: powerful tools for unravelling the functional elements of CAM photosynthesis. *Journal of Experimental Botany* 65: 3425–3441.

WINTER, K., AND J.A.M. HOLTUM. 2002. How closely do the $\delta(13)\text{C}$ values of Crassulacean Acid metabolism plants reflect the proportion of CO_2 fixed during day and night? *Plant Physiology* 129: 1843–1851.

WINTER, K., J.A.M. HOLTUM, AND J.A.C. SMITH. 2015. Crassulacean acid metabolism: a continuous or discrete trait? *New Phytologist* 208: 73–78.

Table 1. Average values for cross sectional anatomical traits, SWC, d13C and titratable acidity measurements

Family	Species	% IAS	cell size (mm)	d13C	leaf thickness (mm)	venation type	dH+ drought	dH+ drought (SD)	phenotype
Didiereaceae	Alluaudia procera	0.71	0.0063	-21.1	0.58	2D			C3-CAM
Anacampserotaceae	Anacampseros lanceolata	0.66	0.0234	-21.76	2.07	3D 2			C3-CAM
Anacampserotaceae	Anacampseros rufescens	1.38	0.0188	-21.21	1.87	3D 2			C3-CAM
Basellaceae	Anredera baselloides	3.38	0.0077	-28	0.59	2D			C3-CAM
Montiaceae	Calandrinia axilliflora	10.03	0.0061	NA	0.73	2D			C3-CAM
Montiaceae	Calandrinia brevipedata	3.75	0.0185	-28.81	1.34	2D			C3-CAM
Montiaceae	Calandrinia colchagensis	4.44	0.0036	NA	0.93	2D			C3-CAM
Montiaceae	Calandrinia flava	6.54	0.0458	-21.68	2.25	3D 1	27.6	2.18	C3-CAM
Montiaceae	Calandrinia kalanniensis	9.9	0.0263	-27.6	1.47	3D 1	42.62	4.2	C3-CAM
Montiaceae	Calandrinia liniflora	3.81	0.0082	-28.53	0.82	2D			C3-CAM
Montiaceae	Calandrinia pentavalvis	4.41	0.0255	-26.5	1.83	3D 1	22.1	2	C3-CAM
Montiaceae	Calandrinia pickeringii	2.26	0.0132	-28.2	0.85	2D	9.32	1.69	C3-CAM
Montiaceae	Calandrinia pleiopetala	5.22	0.0124	-27.2	1.72	3D 1	22	3.37	C3-CAM
Montiaceae	Calandrinia ptychosperma	1.51	0.0076	-26	0.95	2D	7.5	1	C3-CAM
Montiaceae	Calandrinia pumila	8.5	0.0107	-28.55	0.97	2D	8.76	2.67	C3-CAM
Montiaceae	Calandrinia quadrivalvis	4.01	0.018	-28.43	2.08	3D 1	12.18	3.06	C3-CAM
Montiaceae	Calandrinia schistorhiza	1.34	0.0261	-24.1	2.33	2D			C3-CAM
Montiaceae	Calandrinia spergularina	5.87	0.017	-28.71	1.68	3D 1	7.55	2.52	C3-CAM
Montiaceae	Calandrinia tumida	11.45	0.0119	-31.2	1.35	3D 1	0	0	C3
Montiaceae	Calandrinia volubilis	4.41	0.0379	-23.52	1.25	2D	82.23	11.61	C3-CAM
Montiaceae	Calyptidium umbellatum	15.57	0.0046	-25.98	1.14	2D			C3
Didiereaceae	Ceraria fruticulosa	1.74	0.0099	-13	0.85	2D			C3-CAM
Montiaceae	Cistanthe picta	4.77	0.0048	-24.84	1.36	2D			C3-CAM
Montiaceae	Cistanthe salsoloides	0.78	0.0181	-22.9	3.25	2D			C3-CAM
Montiaceae	Claytonia lanceolata	17.94	0.0003	-30.03	0.53	2D			C3
Montiaceae	Claytonia megarhiza	12.03	0.0063	-32	0.89	2D			C3
Molluginaceae	Hypertellis spergulacea	2.99	0.0062	NA	1.02	3D 1			C3-C4
Montiaceae	Lewisia nevadensis	10.35	0.0012	-28.13	0.5	2D			C3
Limeum	Limeum africanum	2.64	0.0045	NA	0.89	2D			C3
Molluginaceae	Mollugo verticillata	3.18	0.0047	NA	0.35	2D			C3-C4
Montiaceae	Montia linearis	12.02	0.0022	-30.1	0.57	2D			C3
Montiaceae	Montia parviflora	9.3	0.0035	-30.5	0.39	2D			C3
Montiaceae	Montiopsis andiocla	1.94	0.001	-27.52	0.4	2D			C3
Montiaceae	Montiopsis gayana	3.77	0.0008	NA	0.42	2D			C3
Pereskia	Pereskia aculeata	0.38	0.0034	-27.56	0.51	2D			C3-CAM
Pereskia	Pereskia guamacho	0.86	0.013	NA	1.19	2D			C3-CAM
Montiaceae	Phemeranthus napiformis	2.58	0.0149	-26.8	2.22	3D 1			C3-CAM
Montiaceae	Phemeranthus teretifolius	1.29	0.0117	-25.2	1.72	3D 1			C3-CAM

Table 2. Average values (and standard deviations) of anatomical traits for C3, C3-CAM and CAM photosynthesis.

phenotype	% IAS	cell size (mm ²)	leaf thickness (mm)
C3	9.70 +/- 5.4	0.0036 +/- 0.0035	0.71 +/- 0.34
C3-CAM	3.47 +/- 2.7	0.0151 +/- 0.01	1.43 +/- 0.67
CAM	0.06	0.0561	4.49

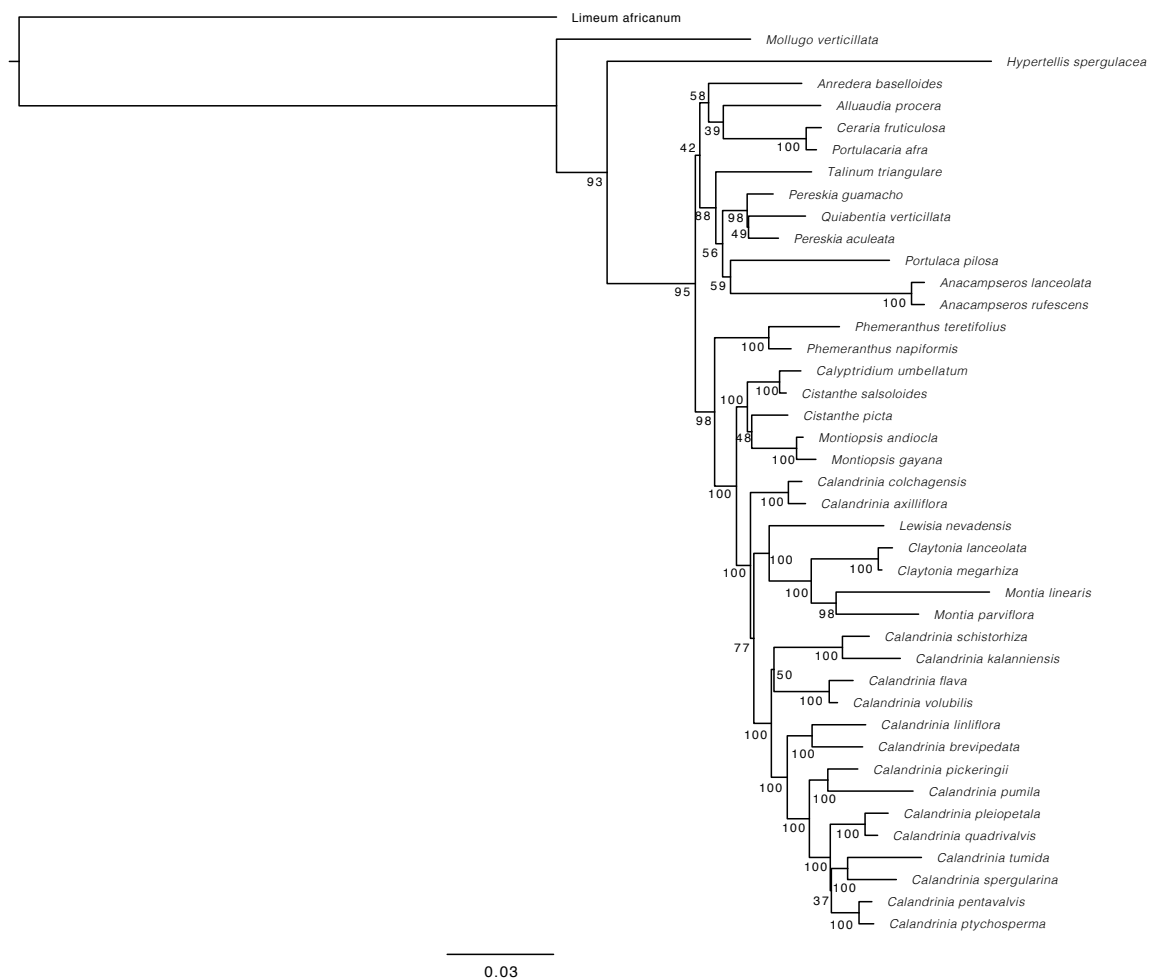


Figure 1. 42-taxa, four-marker (matK, rbcL, ndhF, and ITS) combined analysis maximum likelihood phylogeny.

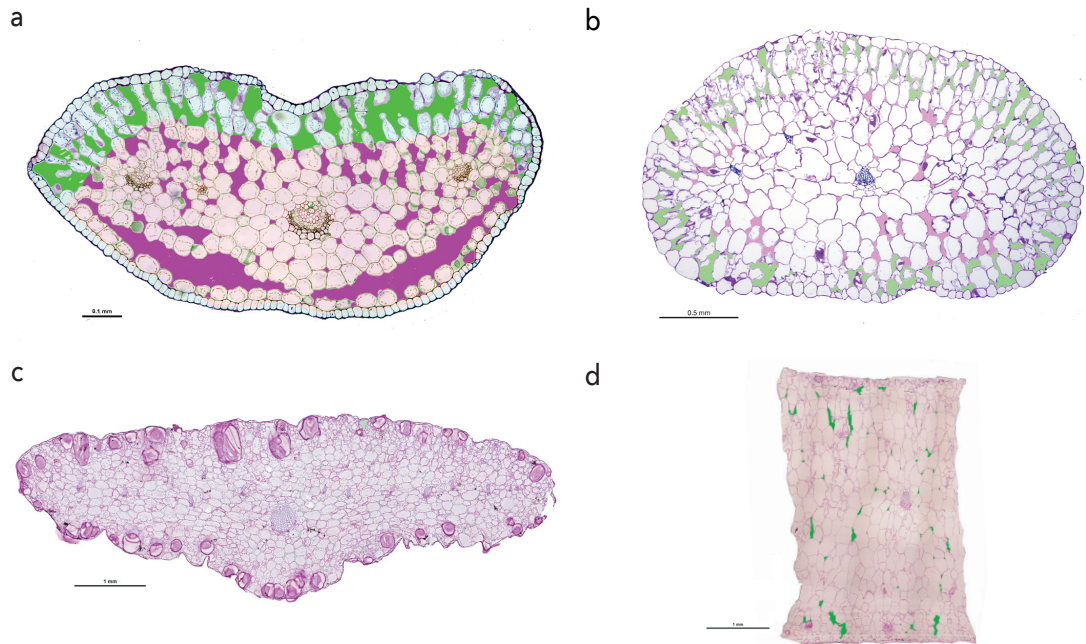


Figure 2. Anatomical cross-sections of species representing the C3 to CAM phenotypic spectrum. (a) *Montia linearis* = C3 (b) *Calandrinia spergularina* = C3-CAM (c) *Anacampseros lanceolata* = C3-CAM (d) *Quiabentia verticillata* = CAM. Green coloring denotes intercellular airspace (IAS). If the mesophyll cells are differentiated, then the spongy mesophyll is colored purple while the palisade mesophyll remains green.

	Leaf Thickness									
Leaf Thickness	na	delta H+ drought								
delta H+ drought		na	IAS	Cell Size						
IAS	*		na	Cell Size	delta 13C					
Cell Size	*	*		na	delta 13C					
delta 13C	*			*	na	MAT				
MAT	*			*		na	MAP			
MAP	*			*	*		na	Precip. seasonality		
Precip. seasonality	*			*		*		na	Venation (2D or 3D)	
Venation (2D or 3D)	*			*		*			na	Photosynthesis (C3 or CAM)
Photosynthesis (C3 or CAM)	*		*	*		*				na

Figure 3. Macro-evolutionary correlations among anatomical traits, photosynthetic mode, venation type, and climatic variables (MAP, MAT and precipitation seasonality). A black star indicates a p-value of <0.05. A red star indicates a p-value <0.05 at the more stringent multiple comparisons cutoff (false discover rate; Benjamini and Hochberg, 1998)

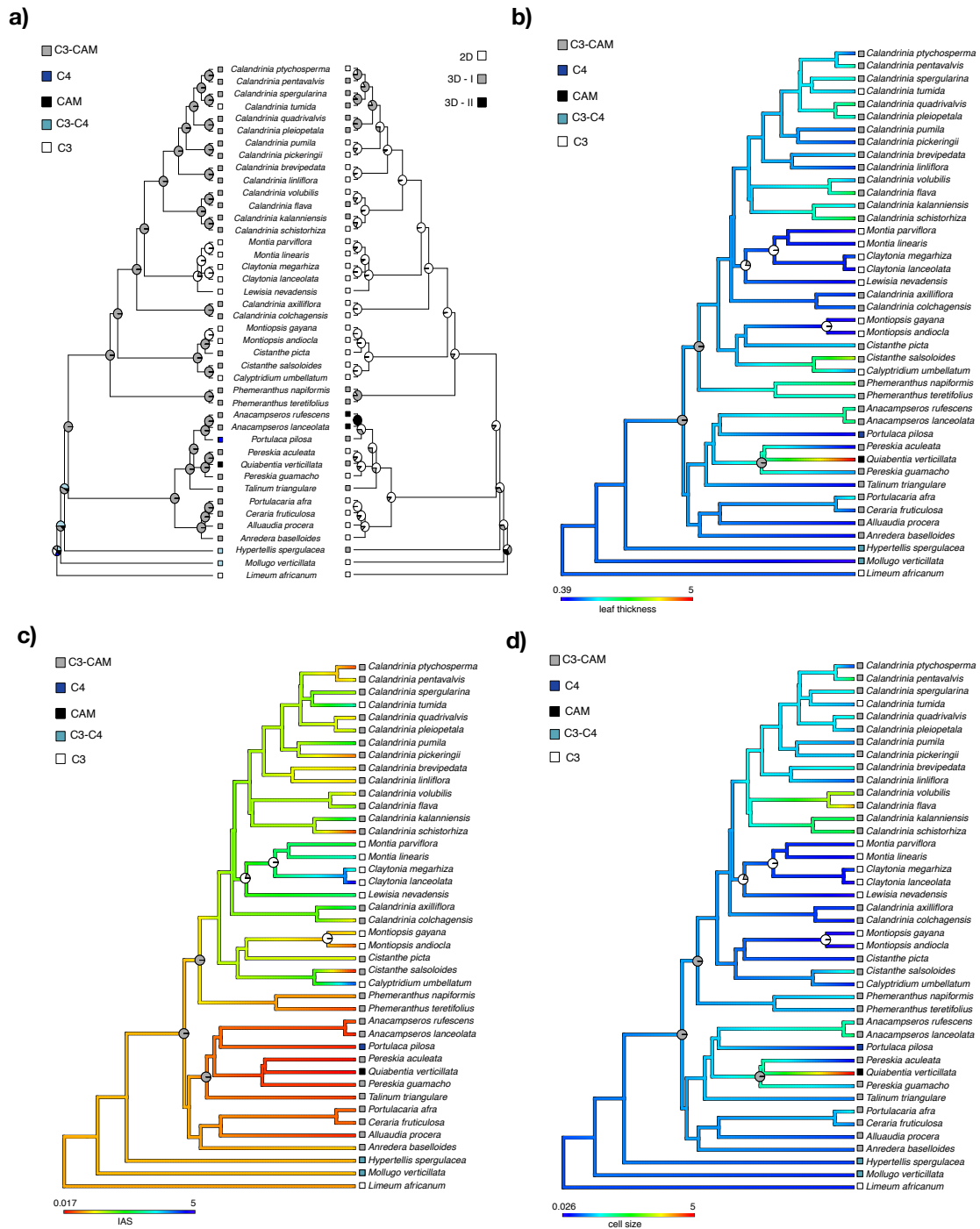


Figure 4. Ancestral character state reconstructions across the four-marker ML portulugo phylogeny. a) Photosynthetic mode and venation type reconstruction, b) leaf thickness reconstruction, c) cell size reconstruction d) intercellular airspace reconstruction. The photosynthetic mode is mapped onto the tips of figures 4b - 4d and shifts in photosynthetic mode are mapped onto corresponding nodes.

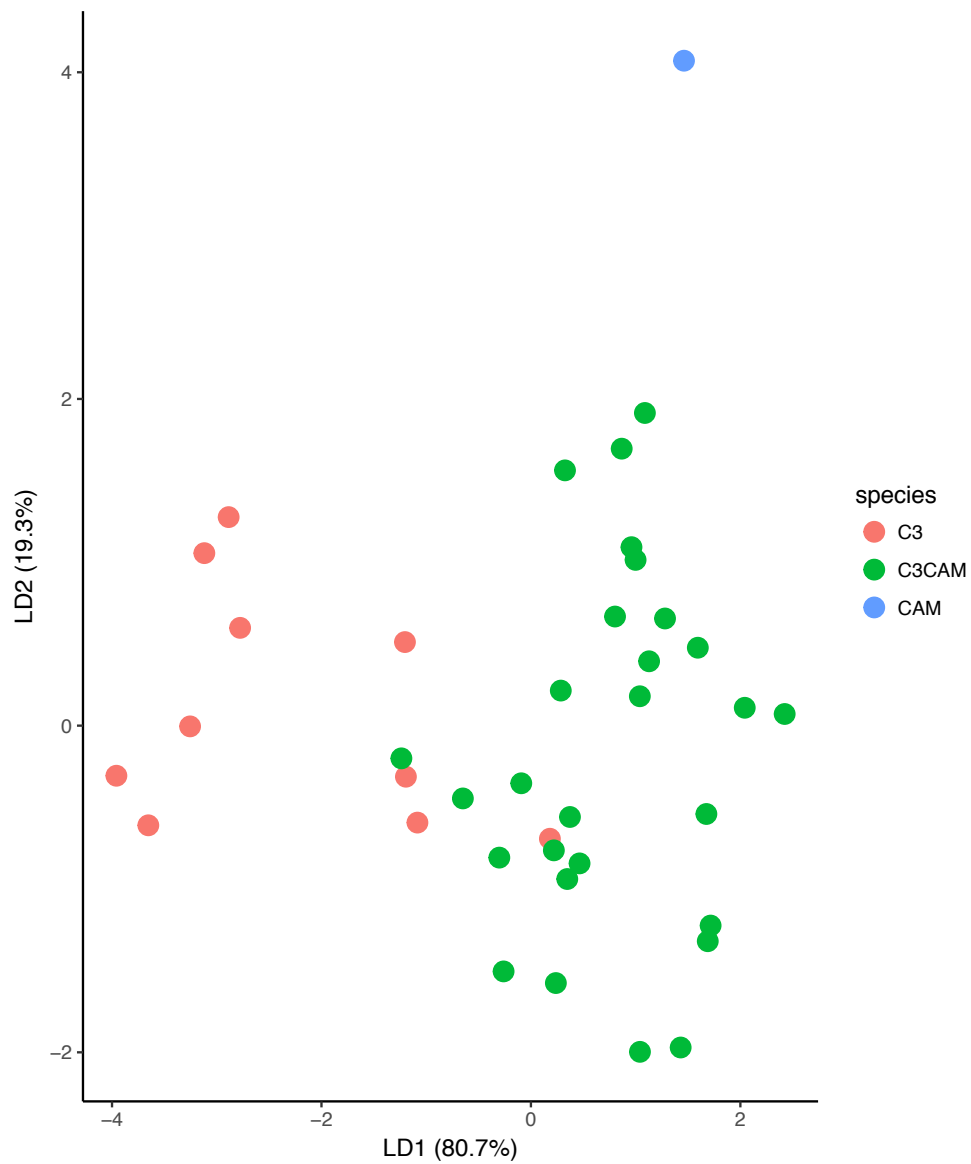


Figure 5. Linear discriminant analysis of 16 BioClim climatic variables for species included in the study. Points are coded by photosynthetic mode C3 (red), C3-CAM (green) or CAM (blue).

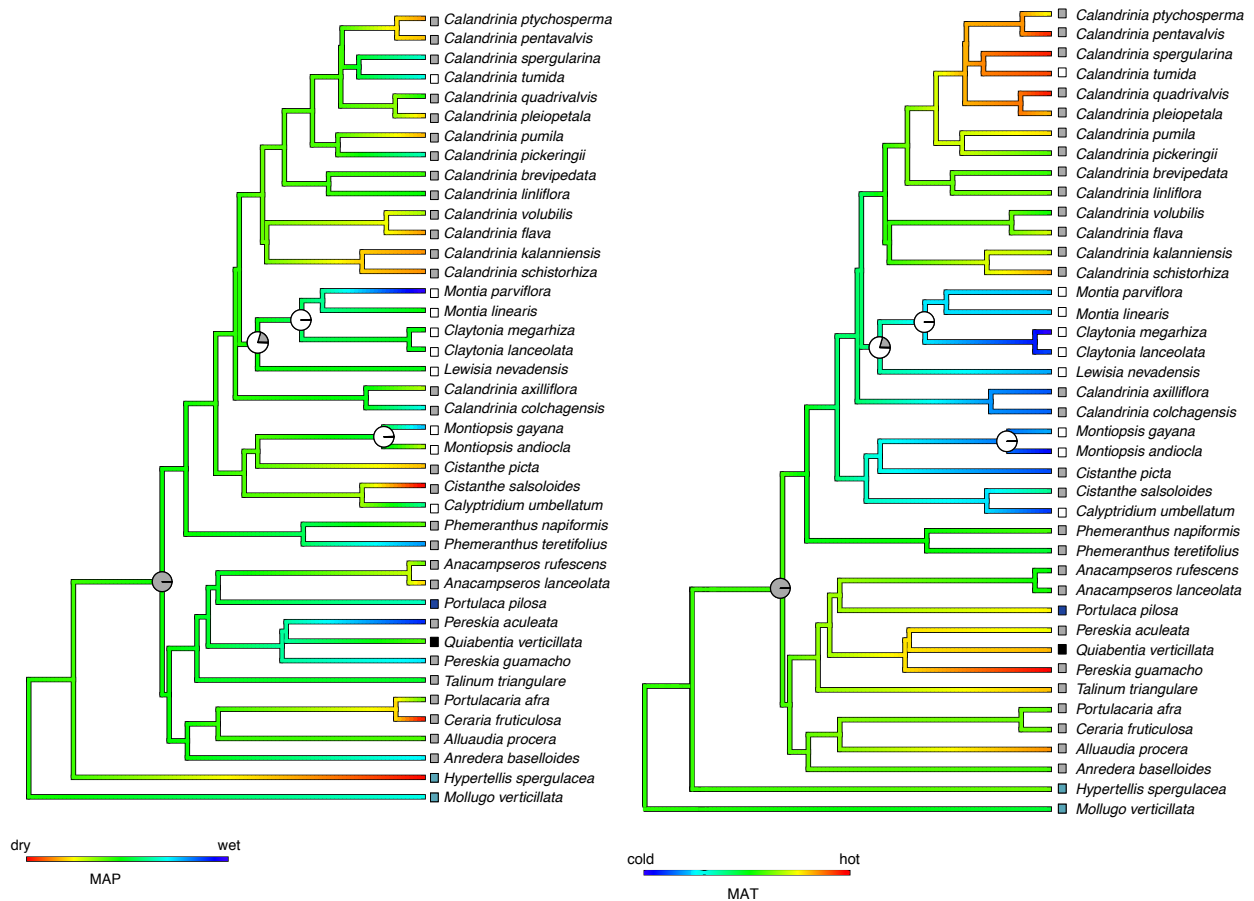


Figure 6. Ancestral character state reconstructions of MAP (a) and MAT (b) across the four-marker ML portulugo phylogeny. Photosynthetic phenotypes are mapped onto the tips. Shifts in photosynthetic mode are mapped onto corresponding nodes in the tree (see Fig 4a for more detail).

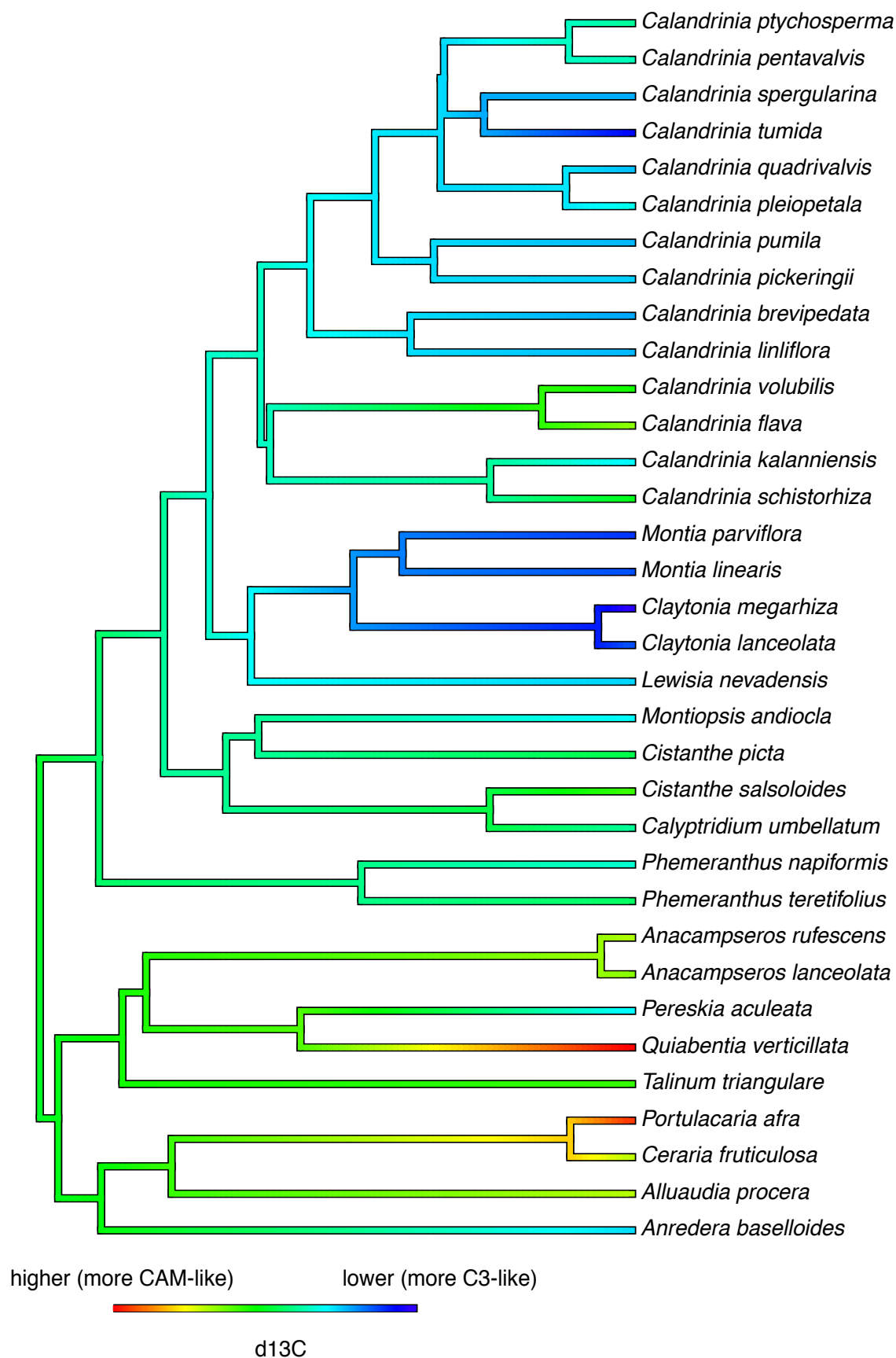


Figure S2. Ancestral character state reconstructions for d13C

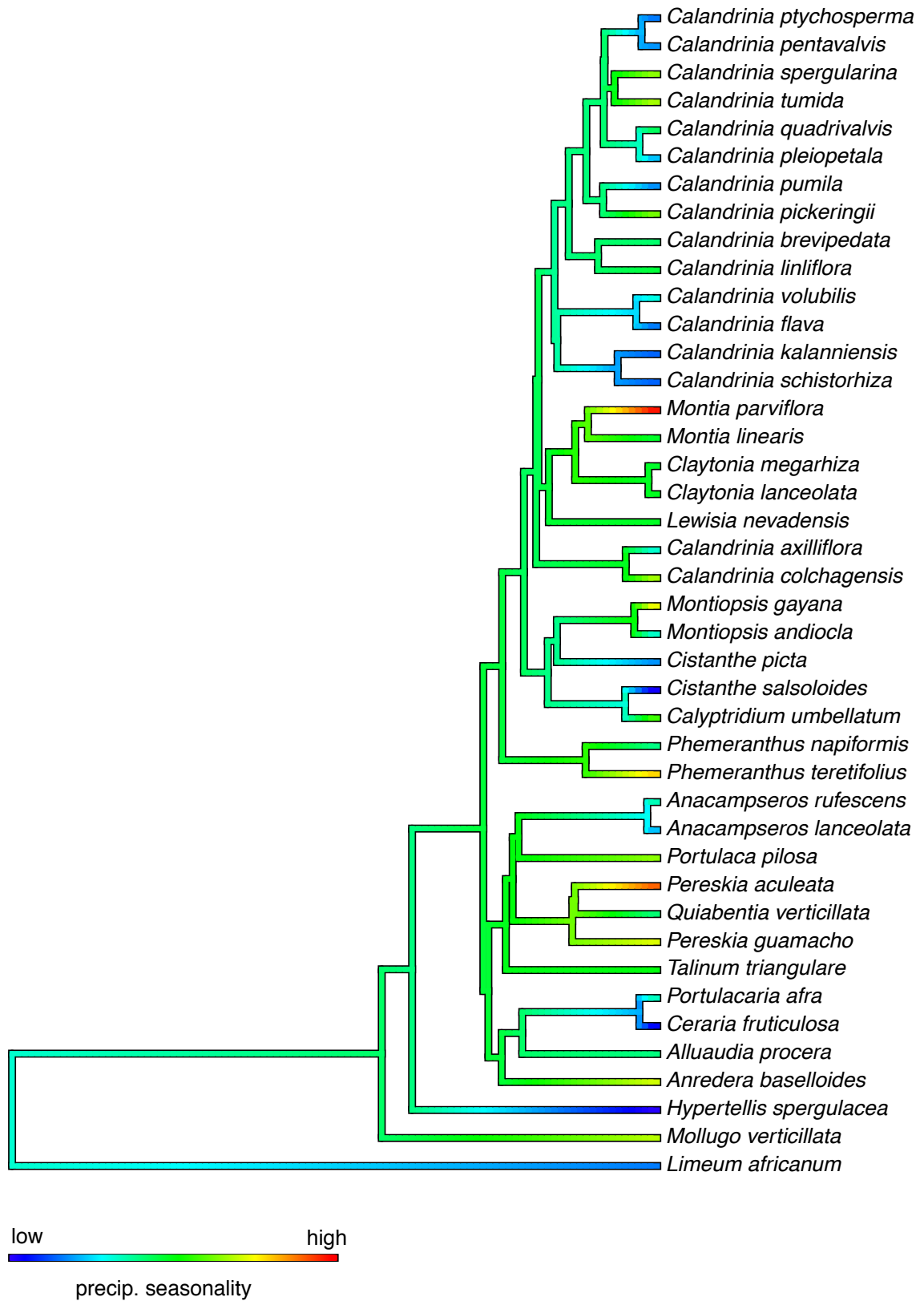


Figure S2. Ancestral character state reconstructions for precipitation seasonality