The Genus *Isoëtes* L., evolution, diversification and population structure in a free-sporing heterosporous lycophyte.

by

Forrest Freund

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Integrative Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Dr. Carl Rothfels, Chair
Dr. Rasmus Nielsen
Dr. Rosemary Gillespie

Spring 2022

The Genus *Isoëtes* L., evolution, diversification and population structure in a free-sporing heterosporous lycophyte.

Abstract

The Genus *Isoëtes* L., evolution, diversification and population structure in a free-sporing heterosporous lycophyte.

by

Forrest Freund

Doctor of Philosophy in Integrative Biology

University of California, Berkeley

Dr. Carl Rothfels, Chair

Among land plants, one of the most cryptic lineages is the genus *Isoëtes* L., both in terms of public awareness and appearance. Taxonomically, *Isoëtes* are members of the vascular plant Division Lycopodiophyta, the sister lineage to the more well-known and lineage rich Euphyllophyta, which includes both the Monilophytes (ferns) and Spermatophytes (seed plants). Lycophytes are comprised of three extant lineages, the homosporous order Lycopodiales, which have the greatest extant generic diversity within the lineage, and the heterosporous Selaginellales and Isoetales, both of which are contemporarily monogeneric. Lycophytes are united by a few key features, including the microphyll, a leaf type unique to the lineage, and adaxial placement of the sporangia on the sporophylls. Like the Monilophytes, Lycophytes are free-sporing plants. As such, they have a pronounced alternation of generations, with a large, dominant diploid sporophyte producing haploid spores via meiosis, which go on to germinate into haploid gametophytes. These gametophytes produce either or both haploid eggs and sperm, which unite within the gametophyte's archegonia to produce a new diploid sporophyte embryo.

The Isoëtalean lycopods, both extinct and extant, are particularly unique among their relatives due to both a suite of unusual characters. One of the first and most obvious of them is their growth form. Unlike their fellow lycophytes, which have unipolar growth, meaning they grow and elongate via a terminal apical meristem on the shoot system, and produce adventitious roots from this stem axis, Isoëtaleans have bipolar growth like most Spermatophytes. And, like the non-monocot Spermatophytes, the Isoëtaleans undergo secondary growth, adding both cortical tissue and secondary vascular tissue via a meristematic region known as the prismatic layer. In the extinct Isoëtalians, these two traits allowed them to become arborescent, with some of the largest members, such as Lepidodendron Sternberg growing to over ten meters tall.

Today, *Isoëtes* is the only remaining member of Isoëtales Prantl. The genus appears to have

arisen in the Triassic and can be found throughout the non-polar regions of the world in seasonally to permanently hydric to aquatic habitats. Morphologically, *Isoëtes* have a highly conserved base bodyplan; almost all of them are small, perennial, semi-herbaceous geophytes whose body is comprised of a highly reduced 1-3 lobed, corm-like trunk, an apical rosette of long, simple, linear leaves, and numerous roots that emerge from their basal furrow. Species are identified through a combination of habitat, spore color, ornamentation and size, leaf morphology, corm lobe numbers, and size of the plants themselves. Because their morphology is so conserved, species identification can be difficult in areas where multiple taxa overlap, especially when the taxa in question are close relatives, as the plants are known to readily hybridize with one another, or form allopolyploids.

Modern *Isoëtes* can be broken up into five distinct sub-clades, which predominantly correspond to their geographic range. There is the Gondwanan clade, which is found in Southern Africa, South America, India and Australia. The Laurasian clade, which occur in the Mediterranean region of Europe and North Africa, North America, and India. The Italian clade, which occurs in and around the Italian peninsula. The Austro-Asian clade, found in Eastern and Southern Asia, India and Australia. And the American clade, which is found in North and South America, as well as a few circumboreal species, and the only known species in Oceania. Because *Isoëtes* is so morphologically conserved, prior to molecular phylogenetics it was assumed that spore morphology or habitat types defined the taxonomic groups, which we now know not to be the case. In fact, these traits are quite labile, particularly when polyploids are involved.

In this dissertation, I explore the evolution of one of the traits in extant *Isoëtes* through ancestral state reconstruction, as well as conduct multi-locus population genetic and phylogenetic analyses to determine if the species composition in one of the sub-clades found on the West Coast of North America.

In my first chapter, I take published morphological data on the corms of modern *Isoëtes* in conjunction with one of the most recently published genus-wide phylogenies to infer how the number of corm lobes have evolved within the genus. Contemporary Isoetes are either bi-lobate or tri-lobate, a trait that does not seem to be restricted to a single clade, but has arisen in multiple different lineages. To examine the evolution of this convergent morphology, I conducted an extensive literature review of published type descriptions in conjunction with monographic sources to code the trait of corm lobation in as many species of *Isoëtes* as possible, with special emphasis placed on finding descriptions of species included in said phylogeny. Using the program RevBayes using Reversable Jump Markov Chain Monte Cristo (RJMCMC) and the 15,000 tree posterior distribution from Larsen and Rydin 2016, we inferred which of the five evolutionary models of change in corm lobation (bi-to-tri, tri-to-bi, 1-rate reversable, 2-rate reversable, and invariant). From the ancestral state reconstruction and model analysis, the most well supported model of character evolution of corm lobation in extant *Isoëtes* an irriversable change from tri to bilobate, though the reversable models had significant non-zero probabilities as well.

In my second chapter, I generated novel molecular data for the Pacific Laruasian clade (PLC), a small, disjunct sub-clade found on the west coast of North America in Southern British Columbia, Washington, Idaho, Oregon, California and northern Baja, for the purpose of generating a clade wide phylogeny and population genetics analysis. The nearest relatives of this subclade are found in Europe and India, with the estimated divergence time between the two dated to around 34 MYA during the transition from the Eocene to Oligocene. 8-12 individuals were collected from 22 populations located across much of the range (specifically California and Oregon) of and identified to the following species based on the published keys: *Isoëtes nuttallii* and *Isoëtes orcuttii*. Full genomic DNA was extracted from a total of 136 samples from the PLC, along with two outgroup species (*I. howellii* and *I. bolanderi*) from the distantly related American clade, were sequenced using the GoFlag probe set at the University of Florida. Sequences fragments were assembled with a newly developed version of HybPiper, a next-generation sequence assembly pipeline that generate sequence assemblies and call variants from short-read DNA data assembled to the 450 reference sequences used to develop the GoFlag probes. Of these 450 reference loci, we recovered 215 single-copy genes, which were used to infer the phylogenetic relationships among the PLC using the multi-species coalescent. In addition, we used the pop-gen program ANGST to generate admixture plots, F-statistics, and Mantel test. These two types of analysis, when combined, show that the existing taxonomy of the PLC does not reflect the phylogenetic structure of the PLC.

In my third chapter, I used the results of my phylogenetics and population genetics chapter as a framework to assess the existing taxonomy of the PLC, and whether it should be reivised in light of my results. In order to do so, I measured the available morphological features of the plants commonly used for species delimitation in *Isoëtes*, as well as made qualitative assessments of their spore morphology, leaf characters and pressence of characters like phyllopodia or scales. From the quantitative data, I performed MANOVA/ANOVA and PCA analysis of their leaf and spore sizes to determine if there were morphological features within each of the clades inferred in Ch. 2 that could be used to distinguish them. Similarly, qualitative characters were compiled and compared to find morphology that supported the structure of the phylogeny. As a result, I highlight the need for taxonomic revision in the PLC, as well as identifying putative new species.

To my family and friends who helped me make it through my dissertation and helped me follow my dreams.

Contents

List of Figures

List of Tables

Acknowledgments

I wish to acknowledge the following people and sources of funding, without whom my dissertation would never have happened.

First, I would like to thank The Jepson Herbarium's Heckard Fund for providing me with funding for my initial molecular work at UC Berkeley for the Pacific Laurasian Clade of *Isoëtes*.

Next, I would like to thank the Myrtle Wolfe grant from The California Native Plants society for use in covering the costs of my collecting trips throughout California to make my collections.

I would like to thank the American Society of Plant Taxonomists Graduate Student Research Grant for providing partial funding for my GoFlag sequencing, while the remaining funds were provided by Dr. Carl Rothfels.

I would like to acknowledge the Berkeley Research Computing program at the University of California, Berkeley (supported by the UC Berkeley Chancellor, Vice Chancellor for Research, and Chief Information Officer), for the use and maintenance of the Savio computational cluster.

In addition to financial support, I would like to thank the following individuals, for both their mentorship and their emotional support throughout the process, and without whom I may not have been able to complete my work.

I would like to thank my father, Robert 'Bob' Edward Freund, who has stood by me and offered nothing but kindness and support for my dreams and my aspirations, even during the days when he and I were not on speaking terms. My brother Glenn Justin Freund, who has been a source of comfort when I'm feeling low and given me a place to feel welcomed. And, my friends Annie Blumenberg, Korey Luna, and my oath-brother Malek Ramadan, who have helped me stay grounded and gave me someone to talk to when I needed it.

Of my academic advisors and professors, I would like to acknowledge the following: Dr. Carl Rothfels, Dr. Cindy Loy, Dr. Rasmus Nielsen, Dr. Rosemary Gilespie, Dr. Jamie Deneris, Dr. Alexandru 'Mihai' Tomescue, Dr. J. Mark Porter, Dr. Carl Taylor, Dr. Elizabeth 'Liz' Zimmer, Dr. Ingrid Jordan Thaden, Sonia Nosratinia, Dr. Fay-Wei Li, Dr. Benjamin Daupan, Dr. Keir Wefferling, Dr. Mike May and Dr. Will Freyman. All scholars and teachers that have given me invaluable help in working on my dissertation, both directly and through helping me figure out what I wanted to do with my academic career.

Finally, I would like to thank the following grad-students in IB who've offered me their friendship during my time at UCB: Carrie Tribble, Jenna Ekwealor, Ixchel Gonzalez Ramirez, Javi Jauregi Lazo, Mick Song, Maryam Sedaghatpour, Briana Boaz, Betsabe Castro Escobar, Tanner Frank, Isaac Mark, David Adelhelm, and Joyce Chery. Even if I'm not the most social person around, I wanted to thank the people listed here who've taken the time to get past my anti-social tendencies and become people I'm comfortable around.

Chapter 1

Inferring the evolutionary reduction of corm lobation in *Isoëtes* using Bayesian model-averaged ancestral state reconstruction

1.1 Authors

Freund, F. D., W. A. Freyman, and C. J. Rothfels. 2018. Inferring the evolutionary reduction of corm lobation in *Isoëtes* using Bayesian model-averaged ancestral state reconstruction. American Journal of Botany 105(2): XXX.

1.2 Premise of the Study

Inferring the evolution of characters in *Isoëtes* has been problematic, as these plants are morphologically conservative and yet highly variable and homoplasious within that conserved base morphology. However, molecular phylogenies have given us a valuable tool to test hypotheses of character evolution within the genus, such as the hypothesis of ongoing morphological reductions.

Methods: We examined the reduction in lobe number on the underground trunk, or corm, by combining the most recent molecular phylogeny with morphological descriptions gathered from the literature and observations of living specimens. Ancestral character states were inferred using nonstationary evolutionary models, reversible-jump MCMC, and Bayesian model averaging.

Key Results: Our results support the hypothesis of a directional reduction in lobe number in *Isoëtes*, with the best-supported model of character evolution being one of irreversible reduction. Furthermore, the most probable ancestral corm lobe number of extant *Isoëtes* is three, and a reduction to two lobes has occurred at least six times.

Conclusions: From our results, we can infer that corm lobation, like many other traits in *Isoëtes*, shows a degree of homoplasy, and yet also shows ongoing evolutionary reduction.

1.3 Key Words:

Bayes factors; evolutionary reduction; Isoetaceae; model selection; morphological evolution; morphological simplification; RevBayes

1.4 Introduction:

Interpreting the morphological evolution of *Isoëtes* L. has troubled botanists for many years. These plants have an outwardly simple, highly conserved body plan, consisting of an apical rosette of linear sporophylls on top of a reduced corm-like trunk (Engelmann, 1882; Cox and Hickey, 1984; Hickey, 1986b,a; Taylor and Hickey, 1992; Budke et al., 2005), yet are highly variable both within (Budke et al., 2005; Liu et al., 2006) and among closely related species (Cox and Hickey, 1984; Hickey, 1986a; Taylor and Hickey, 1992; Bagella et al., 2011). Even characters once thought to be useful in delimiting natural groups within the genus, such as habitat (Engelmann, 1882) or megaspore morphology (Pfeiffer, 1922), have been found to be labile (Cox and Hickey, 1984; Taylor and Hickey, 1992; Budke et al., 2005; Bagella et al., 2011; Hickey et al., 2009). And while some characters such as the glossopodium (the portion of the ligule internal to the leaf) have shown some diagnostic potential (Sharma and Singh, 1984; Shaw and Hickey, 2005; Singh et al., 2010; Freund, 2016), actually examining and interpreting these structures requires considerable histological and computational effort, making them ill-suited for field identification. This absence of consistent, dependable characters creates a paradox: the lack of reliable traits impedes the inference of phylogenies or classifications in the genus, but without a phylogeny, examining character evolution is exceptionally difficult.

However, molecular phylogenetics has vastly improved our understanding of *Isoëtes* (Rydin and Wikström, 2002; Hoot et al., 2004; Schuettpelz and Hoot, 2006; Larsén and Rydin, 2016) and has provided evidence for five major clades within the genus: Gondwanan (Clade A), Laurasian (Clade B), Italian (Clade C), Austro-Asian (Clade D), and New World or American (Clade E) (Hoot et al., 2004; Larsén and Rydin, 2016)). This phylogeny has completely overturned the old morphological and ecological systems of classification and has also provided a framework to begin earnestly studying character evolution in the genus. While characters such as spore ornamentation are highly homoplastic (Cox and Hickey, 1984; Larsén and Rydin, 2016), there are other changes that may be informative, such as serial reduction of the corm (Karrfalt and Eggert, 1977b,a; Pigg, 1992; Grauvogel-Stamm and Lugardon, 2001).

The isoetalians have a long fossil history, which has spanned over 300 million yrs (Pigg, 1992; Gensel and Pigg, 2010; Taylor et al., 2009). Due to this long history, and the abundance of fossil representatives, there is a long-standing comparison between the corm of extant

Isoëtes and the stigmarian appendages of their extinct arborescent relatives. This comparison is supported by the unique suite of characters that are found only in the isoetalian lycophytes (Stewart, 1947; Karrfalt and Eggert, 1977b,a; Karrfalt, 1984a; Jennings et al., 1983; Pigg, 1992). These features include a unique and independently evolved form of secondary growth that is fundamentally different from the secondary growth of spermatophytes (Scott and Hill, 1900; Stokey, 1909; Pfeiffer, 1922; Stewart, 1947; Karrfalt and Eggert, 1977b,a; Pigg and Rothwell, 1985; Ash and Pigg, 1991; Pigg, 1992; Gensel and Pigg, 2010). Their rooting structures are also unique and have been compared to a dichotomous branching system, with their rootlets arising from modified leaves Rothwell and Erwin (1985). Some fossil isoetalians, such as the arborescent Protostigmaria Jennings (Jennings et al., 1983) or geophytic Nathorstiana Richter Richter (1910); Mägdefrau (1932); Taylor et al. (2009), have subterranean morphology comparable to that of modern *Isoëtes* (i.e., not producing elongated stigmarian appendages), with the size, elaboration, or arborescent habit of the plants separating them from the contemporary taxa Stewart (1947); Karrfalt (1984a); Pigg (1992). Protostigmaria, while having a generally comparable morphology to extant *Isoëtes*, differs by having as many as 13 lobes on its rooting axis. *Nathorstiana*, while having corms that are quite similar to those of extant *Isoëtes*, shows high levels of corm lobe variability, like some extant species of *Isoëtes*, yet little correlation of plant size with lobe number (some of the small plants have four lobes, and some larger plants are unlobed; Richter, 1910; Mägdefrau, 1932; Jennings et al., 1983; Rothwell, 1984; Rothwell and Erwin, 1985; Taylor et al., 2009). Extant *Isoëtes*, and potentially the extinct forms, also possess a unique mechanism to penetrate into the substrate by laterally displacing the soil and then pulling the corm deeper through a combination of secondary growth and development of new rootlets. While functionally similar to the contractile roots of monocots, this mechanism does not compress the soil below the plant, allowing for deeper substrate penetration (Karrfalt and Eggert, 1977b,a). Due to these shared features, it has been hypothesized that modern *Isoëtes* are the result of continuous reduction of the once arborescent isoetalian body plan, with the reduced corm lobe numbers in modern *Isoëtes* continuing this trend (Stewart, 1947; Karrfalt and Eggert, 1977b,a).

These corm lobes are formed by the basal, rootlet-producing meristems (Engelmann, 1882; Stokey, 1909; Osborn, 1922; Pfeiffer, 1922; Bhambie, 1963; Karrfalt and Eggert, 1977b; Budke et al., 2005); extant plants have either a trilobate or bilobate base morphology. This morphology appears early in ontology and has been observed in sporelings (F. D. Freund, personal observation). In trilobate species, there are three basal furrows, which run down the lateral face of the corm and join together at the distal end (Fig. 1.1 A-C). These basal furrows divide the corm into three sections where secondary growth ultimately results in the formation of the triple corm lobes 1963;(Stokey, 1909; Osborn, 1922; Bhambie, 1963, F. D. Freund, personal observation). By contrast, the bilobate species have only a single furrow, which runs in a line across the base of the corm, dividing it into two halves (Fig. 1.1 D–F). While the plants do occasionally acquire additional lobes as they age, the base morphology—the minimum number of lobes the plants have before any elaboration—is consistent (Karrfalt and Eggert, 1977b,a). Also, there are other, rarer morphologies, such as

the rhizomatous, mat-forming *I. tegetiformans* Rury (Rury, 1978) and the monolobate *I. andicola* (Amstutz) L.D. Gómez (formerly treated as *Stylites andicola* Amstutz; (Amstutz, 1957)). However, both *I. andicola* and *I. tegetiformans* begin life with a bilobate morphology before secondarily acquiring these alternate states (Rury, 1978; Tryon, 1994); their base morphology is bilobate.

Here, we explore the evolution of corm lobation across the most recent *Isoëtes* phylogeny (Larsén and Rydin, 2016). If *Isoëtes* has undergone evolutionary reduction in corm lobation number as (Karrfalt and Eggert, 1977a) hypothesized, we expected the ancestral state for extant *Isoëtes* to be trilobate, and the rate of transition from trilobate to bilobate to be higher than the rate of bilobate to trilobate. To test this hypothesis, we used reversible-jump Markov chain Monte Carlo (MCMC Green (1995)) to explore multiple nonstationary statistical models of morphological evolution (Klopfstein et al., 2015) and to assess whether there is evidence for directional or even irreversible evolution in corm lobation morphology. Finally, we used simulations to examine our power to detect irreversible evolution on datasets of this size.

1.5 Methods

Corm lobation characterization

Corm lobe numbers were collected from observations of fresh material, herbarium specimens (Appendix 1), and published descriptions (Appendix 2). We characterized lobation from fresh plants for eight species (*Isoëtes appalachiana* D.F. Brunt. & D.M. Britton [four individuals]; *I. bolanderi* Engelm. [20+ individuals]; *I. eatonii* R. Dodge [three individuals]; *I. englemannii* A. Braun [four individuals]; *I. howellii* Engelm. [200+ individuals]; *I. nuttallii* A. Braun ex Engelm. [400+ individuals]; *I. occidentalis* L.F. Hend. [15+ individuals]; and *I. orcuttii* A.A. Eaton [50+ individuals]). Each individual was cleaned of encrusting soil, then examined to determine the number of basal furrows and to assess the symmetry of the corm lobes. Plants of different ages were observed for each species when possible to get a sense of variability in lobation pattern as the plants aged, as well as to identify any unusual morphological outliers such as plants with asymmetric lobation or nonlinear basal furrows that may have occurred due to advanced age or damage to the corm. The *I. nuttallii* samples include two sporelings, one of which was still connected to the megaspore, which already showed distinctly trilobate arrangement of their rootlets. Taxa in which the plants have a single contiguous furrow and two symmetrical lobes were scored as "bilobate," while species with three linear furrows with symmetrical lobation around the base of the corm were scored as "trilobate" (Fig. 1.1).

For lobation numbers garnered from the literature, species with a definitive, single reported corm lobation value were assigned that number, while species with a range of reported corm lobe numbers were scored as "undefined" due to the possibility that the description is of a taxon with a range of corm morphologies, such as *I. tuckermanii* A. Br. (Karrfalt

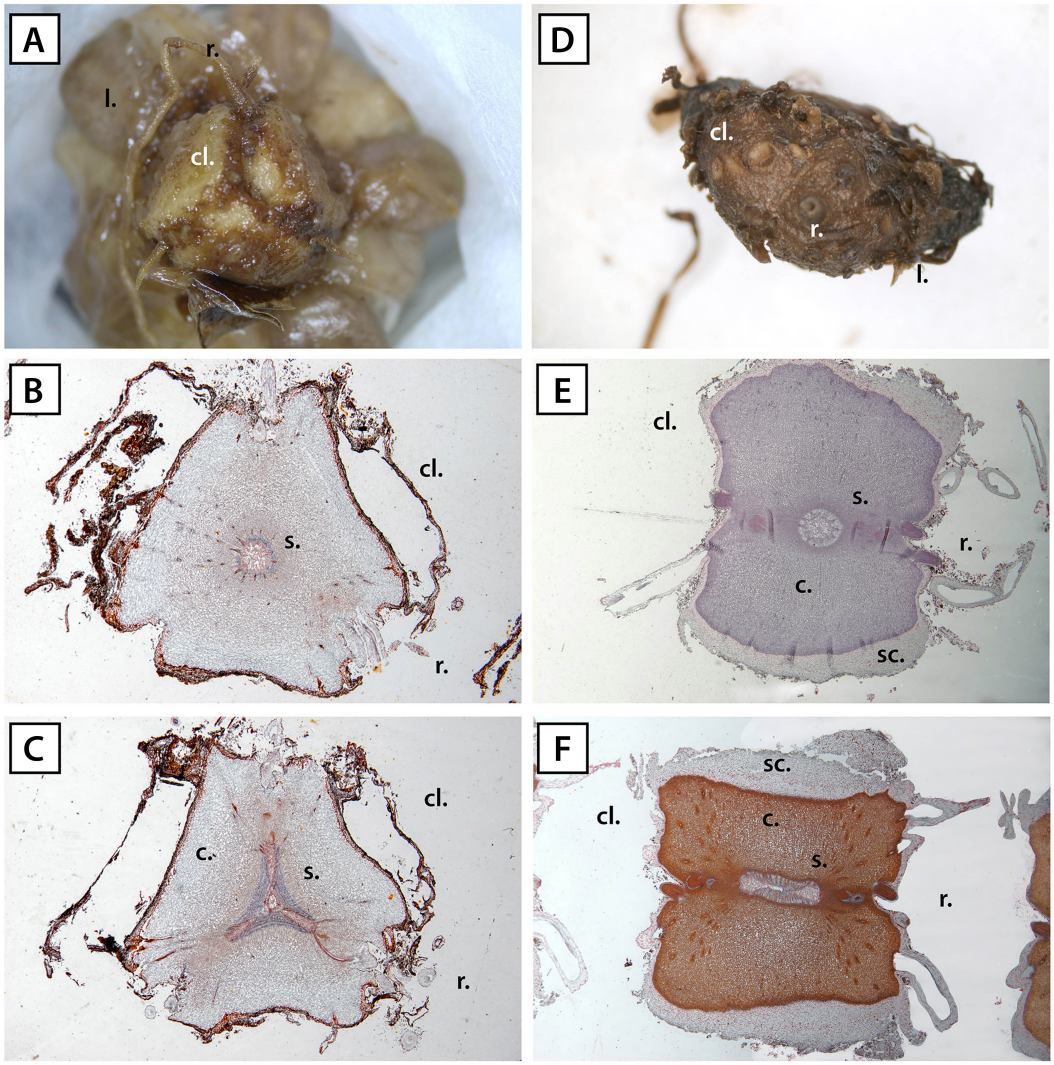

Figure 1.1: Comparative morphology of trilobate and bilobate corms. Trilobed corm: (A) *Isoëtes nuttallii* (F. Freund 266, UC) in proximal view with senesced cortical tissue removed (8.1×, 29 mm field of view); (B–C) transverse sections of *I. nuttallii* (F. Freund 169, RSA796374) below apical rosette (B) and near basal furrows (C) (32.5×, 7.1 mm field of view). Bilobed corm: (D) *I. howellii* (F. Freund 266.1, UC) in proximal view (24×, 9.5 mm field of view); (E–F) cross sections of *I. bolanderi* (F. Freund 10, RSA811643) below apical rosette (E) and near basal furrow (F) (24.5×, 9.3 mm field of view). All images were adjusted in Adobe Photoshop to improve visibility by using the image levels tool and the dodge tool to even out lighting and coloration. For herbarium acronyms, see Index Herbariorum (http://sweetgum.nybg.org/science/ih/). Labels: c. = cortex ground tissue, cl. = corm lobe, r. = rootlets, l. = leaves, s. = stele, and sc. = senesced cortical ground tissue.

and Eggert, 1977a; Croft, 1980), or due to the inclusion of multiple cryptic species within a single taxon (Table 1.1).

Table 1.1: Corm lobation numbers for *Isoëtes*. Clade designations follow Larsén and Rydin (2016). For herbarium acronyms, see Index Herbariorum (http://sweetgum.nybg.org/science/ih/).

Species	Observed by F. Freund	Clade	Lobes	Locality	Sources
Isoëtes abyssinica Choiv. – synonym of *I. welwitschii* (sensu Verdcourt and Beentje, 2002)		B-3	3	Africa, Ethiopia	Pfeiffer (1922); Cook (2004); Crouch (2012)
Isoëtes acadiensis Kott		?	2	USA (MA, ME, NH, NJ, NY, VT), CAN (NB, NF, NS)	Taylor (1993)
Isoëtes adspersa A. Braun		?	3	Algeria	Pfeiffer (1922)
Isoëtes aequinoctialis A. Braun		B-3	3	Angola, South Africa, Mali, Ghana, Tanzania, Zambia, Zimbabwe, south to RSA: NAM, NC	Pfeiffer (1922); Cook (2004); Crouch (2012)
Isoëtes alpina Kirk		D-2	3	New Zealand	Pfeiffer (1922)
Isoëtes alstonii C.F. Reed & Verdc.		?	2	Egypt; Mozambique; Namibia; Sudan; Tanzania, United Republic of; Zambia; Zimbabwe	Crouch (2012)
Isoëtes amazonica A. Braun ex Kuhn		A-3	3	Brazil (north [Pará])	Pfeiffer (1922)
Isoëtes anatolica Prada & Rolleri		C	3	Turkey	Prada and Rolleri (2005)
Isoëtes andicola (Amstutz) L.D. Gómez (formerly *Stylites andicola* Amstutz)	Live plants	E-2	2	Peru (Lima, Pasco, Junin, Cuzco, Puno), Bolivia	Karrfalt (1984b); Tryon (1994)
Isoëtes andina Spruce ex Hook.– synonym of I. triquetra		E-2	2	South America, Peru	Pfeiffer (1922)
Isoëtes appalachiana D.F. Brunt. & D.M. Britton	Live plants	E-2	2	USA (FL, GA, NC, PA, SC, VA)	Brunton and Britton (1997)
Isoëtes araucaniana Macluf & Hickey		?	2	Malleco (Chile)	Macluf and Hickey (2007)
Isoëtes asiatica Makino		E-2	2	Kamchatka, Sakhalin, the Kuriles, and Japan	Yi and Kato (2001)
Isoëtes australis R.O. Williams		A-2, D-2	2	Bruce Rock, West Australia	Williams (1944)
Isoëtes azorica Durieu ex Milde		?	2	Islands of Azores	Pfeiffer (1922)
Isoëtes bolanderi Engelm.	RSA (811637, 881638, 811639, 811643)	E-2	2	USA (AZ, CA, CO, FL, ID, MT, NM, NV, OR, UT, WA, WY), CAN (AB)	Pfeiffer (1922); Taylor (1993)
Isoëtes boliviensis U. Weber		?	2	Peru (Cajamarca, San Martin, Ancash, Lima, Pasco, Junin, Ayacucho, Cuzco, Puno), Bolivia (La Paz)	Tryon (1994)
Isoëtes boryana Durieu		?	3	France (Landes)	Pfeiffer (1922)

Isoëtes bradei Herter		A-3	3	Brazil, Sudeste (São Paulo)	Brazilian Flora 2020 (in construction). Rio de Janeiro Botanical Garden. Available at http://floradobrasil.jbrj.gov.br/, accessed Aug 15, 2016; Hickey, 1990
Isoëtes braunii Unger		?	2	USA (NH, VT, MA), CAN	Pfeiffer (1922)
Isoëtes brevicula E.R.L. Johnson		D-2	3	Western Australia	FloraBase, the Western Australian Flora, Western Australian Herbarium. https://florabase.dpaw.wa.gov.au/, accessed Aug 15, 2016
Isoëtes brochoni Motelay		?	2	France	Pfeiffer (1922)
Isoëtes butleri Engelm.		?	2	USA (AL, AR, GA, IL, KS, KY, MO, OK, TN, TX)	Pfeiffer (1922)
Isoëtes capensis A.V. Duthie		A-1	3	South Africa (Western Cape)	Cook (2004); Crouch (2012)
Isoëtes caroli E.R.L. Johnson		D-2	3	Western Australia	FloraBase, the Western Australian Flora, Western Australian Herbarium. https://florabase.dpaw.wa.gov.au/, accessed Aug 15, 2016
Isoëtes caroliniana (A.A. Eaton) Luebke		?	2	USA (NC, TN , VA, WV)	Taylor (1993)
Isoëtes coreana Y.H. Chung & H.K. Choi		D-3	3	Korea	Jeong and Choe (1986)
Isoëtes coromandeliana L.f.		A-4	3	India	Pfeiffer (1922)
Isoëtes cubana Engelm. ex Baker		A-3	3	Cuba (Pinao del Rio), Belize, Mexico (Yucatan)	Pfeiffer (1922)
Isoëtes dispora Hickey		?	2	Laguna Tembladera, Lambayeque, Peru	Tryon (1994)
Isoëtes dixitei Shende		B-3	3	Panchgani, India	Pant and Srivastava (1962)
Isoëtes drummondii A. Braun		D-2	3	Australia (South Australia, Western Australia, Victoria, New South Wales)	Osborn (1922); Pfeiffer (1922)
Isoëtes duriei Bory		B-1	3	Algeria, Corsica, France, Italy, Turkey	Pfeiffer (1922)
Isoëtes eatonii R. Dodge	Live plants.	E-2	2	USA (CT, MA, NH, NJ, NY, PA, VT), CAN (ON)	Pfeiffer (1922)

Isoëtes echinospora Durieu		E-2	2	CAN (AB, BC, MB, NB, NT, NS, ON, PE, QC, SK, YT), USA (AK, CA, CO, ID, ME, MA, MI, MN, MT, NH, NJ, OH, OR, PA, VT, WA, WI), British Isles, Circum boreal	Pfeiffer (1922)
Isoëtes elatior A. Braun		?	3	Tasmania	Pfeiffer (1922)
Isoëtes eludens J.P.Roux, Hopper & Rhian J.Sm.		?	3	South Africa (Kamiesberg Mountains, Namaqualand)	Roux et al. (2009); Crouch (2012)
Isoëtes engelmannii A. Braun (includes I. valida)	Live plants.	E-2	2	CAN (ON), USA (Al, AR, CT, DE, FL, GA, IL, IN, KY, MD, MA, MI, MO, NH, NJ, NY, NC, OH, PA, RI, SC, TN, VT, VA, WV)	Pfeiffer (1922); Taylor (1993)
Isoëtes flaccida A. Braun		E-2	2	USA (FL, GA)	Pfeiffer (1922); Taylor (1993)
Isoëtes flettii (A.A. Eaton) N. Pfeiff.		?	2	USA (WA - Spanaway Lake)	Pfeiffer (1922)
Isoëtes foveolata A.A. Eaton ex R. Dodge		?	2	USA (CT, MA, NH), CAN (ONT)	Pfeiffer (1922)
Isoëtes gardneriana Kunze ex Mett.		?	3	Brazil (Midwest [Goias]), Paraguay	Pfeiffer (1922)
Isoëtes georgiana Luebke		E-2	2	USA (GA)	Taylor (1993)
Isoëtes giessii Launert		?	3	Namibia (Erongo Mountains and in seasonally wed depressions in acacia scrub)	Crouch (2012)
Isoëtes gunnii A. Braun		D-2	3	Tasmania (Lake Fenton on Mt. Field)	Pfeiffer (1922)
Isoëtes habbemensis Alston		D-3	(2)–3–(4)	New Guinea	Croft (1980)
Isoëtes hallasanensis H.K. Choi, Ch. Kim & J. Jung		D-3	3	Jeju Island, Korea	Choi et al. (2008)
Isoëtes hawaiiensis W.C. Taylor & W.H. Wagner		E-2	2	Hawaii	Taylor (1993)
Isoëtes heldreichii Wettst.		?	3	Greece (plains of Thessaly, base of Pindus Mts.)	Pfeiffer (1922)
Isoëtes herzogii U. Weber		E-2	2	Bolivia	Weber (1922)
Isoëtes hewitsonii Hickey		?	2	Celendin, Cajamarca, Peru	Tryon (1994)
Isoëtes histrix Bory & Durieu		E-1, E-2	3	Algeria, Italy, France, islands of the Mediterranean	Pfeiffer (1922)
Isoëtes howellii Engelm.	RSA (796366, 796367, 796368, 796369. 796370, 796371)	E-2	2	USA (MT, ID, WA, OR, CA)	Pfeiffer (1922); Taylor (1993)
Isoëtes humilior A. Braun		D-1	2	Tasmania (So. River Esk)	Pfeiffer (1922); Taylor (1993)

Isoëtes inflata E.R.L. Johnson		?	2	Australia (West Australia)	FloraBase, the Western Australian Flora, Western Australian Herbarium. https://florabase.dpaw.wa.gov.au/, accessed Aug 15, 2016
Isoëtes japonica A. Braun		D-3	3	Japan (Yokohama)	Pfeiffer (1922)
Isoëtes jejuensis H.K. Choi, Ch. Kim & J. Jung		D-3	3	Jeju Island, Korea	Choi et al. (2008)
Isoëtes kirkii A. Braun		D-1	3	New Zealand	Pfeiffer (1922)
Isoëtes labri-draconis N.R. Crouch		?	3	South Africa (Drakensberg Range in KwaZulu-Natal)	Crouch (2012)
Isoëtes lacustris L.		E-2	2	CAN (MB, NB, NL, NT, NS, ON, OQ, SK), USA (ME, MA, MI, MN, NH, NY, VT, WI)	Pfeiffer (1922); Taylor (1993)
Isoëtes laosiensis C. Kim & H.K. Choi		A-4	3	Laos	Kim et al. (2010)
Isoëtes lechleri Mett.		?	2	Argentina, Equador, Colombia, Peru	Pfeiffer (1922); Tryon (1994)
Isoëtes lithophila N. Pfeiff.		E-2	2	USA (TX)	Pfeiffer (1922); Tryon (1994)
Isoëtes longissima Bory & Dur. – As *I. velata* forma *longissima* (sensu Pfeiffer (1922))		B-3	3	Algeria	Pfeiffer (1922)
Isoëtes macrospora Durieu		?	2	Newfoundland to USA (MN)	Pfeiffer (1922)
Isoëtes malinverniana Ces. & De Not.		C	3	Italy	Pfeiffer (1922)
Isoëtes maritima Underw.		E-2	2	Alaska, British Columbia, Washington	Underwood (1888)
Isoëtes martii A. Braun ex Kuhn		E-2	2	Brazil (Southeast [Minas Gerais, Rio de Janeiro], South [Rio Grande do Sul])	Brazilian Flora 2020 (in construction). Rio de Janeiro Botanical Garden. Available at http://floradobrasil.jbrj.gov.br/, accessed Aug 15, 2016
Isoëtes maxima Hickey, Macluf & Link-Pérez		?	2	Brazil (South [Rio Grande do Sul])	Hickey et al. (2009)
Isoëtes melanopoda Gay & Durieu		E-2	2	USA (AL, AR, GA, IA, ID, IL, IN, KS, KY, LA, MN, MO, MS, MT, NC, NE, NJ, OK, SC, SD, TN, TX, UT, VA)	Pfeiffer (1922); Taylor (1993)
Isoëtes melanospora Engelm.		E-2	2	USA (GA - Stone Mtn.)	Pfeiffer (1922); Taylor (1993)
Isoëtes mexicana Underw.		E-2	2	Mexico (Chihuahua, Hidalgo, Mexico, Morelos, Michoacan)	Pfeiffer (1922)

Isoëtes mongerensis E.R.L. Johnson		?	3	Western Australia	FloraBase, the Western Australian Flora, Western Australian Herbarium. https://florabase.dpaw.wa.gov.au/, accessed Aug 15, 2016
Isoëtes muelleri A. Braun		D-3	3	Eastern Australia (Rockhampton)	Pfeiffer (1922)
Isoëtes muricata Durieu		E-2	2	North America	Pfeiffer (1922), as I. braunii
Isoëtes nigritiana A. Braun		?	3	Nigeria (along the Niger River, Nupe)	Pfeiffer (1922)
Isoëtes novo-granadensis H.P. Fuchs		E-2	2	Columbia, Peru	Tryon (1994)
Isoëtes nuttallii A. Braun ex Engelm.	RSA (796374, 796375, 796376)	B-2	3	USA (CA, OR, WA), CAN (Vancouver)	Pfeiffer (1922); Taylor (1993)
Isoëtes occidentalis L.F. Hend.	RSA (811640, 811641, 811642)	E-2	2	USA (CA, WY, CO, ID)	Pfeiffer (1922); Taylor (1993)
Isoëtes olympica A. Braun		B-3	3		Pfeiffer (1922)
Isoëtes orcuttii A.A. Eaton	Live plants	B-2	3	USA (CA)	Pfeiffer (1922); Taylor (1993)
Isoëtes ovata N. Pfeiff.		?	3	French Guiana, Guyana	Pfeiffer (1922)
Isoëtes panamensis Maxon & C.V. Morton	Museo Nacional de Costa Rica (sheet no. not recorded)	A-3	2	Brazil (northeast [Maranhão, Bahia], midwest [Mato Grosso])	Brazilian Flora 2020 (in construction). Rio de Janeiro Botanical Garden. Available at http://floradobrasil.jbrj.gov.br/, accessed Aug 15, 2016
Isoëtes parvula Hickey		?	2	Laguna Yaurihuiri, Ayacucho, Perú	Tryon (1994)
Isoëtes pedersenii H.P. Fuchs ex E.I. Meza & Macluf		?	2	Argentina (Corrientes - Mburucuyá National Park)	Macluf et al. (2010)
Isoëtes philippinensis Merr. & R.H. Perry		D-3	3	Philippine islands	Merrill and Perry (1940)
Isoëtes piperi A.A. Eaton		?	2	USA (WA)	Pfeiffer (1922)
Isoëtes pringlei Underw.		?	2	Mexico (Guadalajara, state of Jalisco)	Pfeiffer (1922)
Isoëtes prototypus D.M. Britton		?	2	CAN (NB, NS), USA (MA)	Taylor (1993)
Isoëtes pseudojaponica M. Takamiya, Mits. Watan. & K. Ono		D-3	3	Japan	Takamiya et al. (1998)
Isoëtes quiririensis J.B.S. Pereira & Labiak		?	2	Brazil (Serra do Quoriri)	Pereira and Labiak (2013)
Isoëtes riparia Engelm. ex A. Braun		?	2	Can (southern region), USA (New England south to DE and PA)	Pfeiffer (1922)
Isoëtes saccharata Engelm.		?	2	USA (DE, DC, MD, VA)	Pfeiffer (1922)

Isoëtes sampathkumarinii L.N. Rao		D-2	2	India	Rao (1944)
Isoëtes saracochensis Hickey		?	2	Laguna Saracocha, Puno, Peru	Tryon (1994)
Isoëtes savatieri Franch.		E-2	2	Argentina (Peurto Beuno)	Hickey et al. (2003)
Isoëtes schweinfurthii Baker		A-4	3	South Africa (Namibia, Botswana, Limpopo), Sudan, Madagascar, Zambia	Pfeiffer (1922); Cook (2004); Crouch (2012)
Isoëtes setacea Lam.		E-1	3	France (Montpellier, Hérault, Pyréneés-Orientales); Morocco; Portugal; Spain (Baleares)	Pfeiffer (1922)
Isoëtes stellenbossiensis A.V. Duthie		A-1	3	South Africa (western part of Western Cape province)	Cook (2004); Crouch (2012)
Isoëtes stephansenii A.V. Duthie		A-1	2	South Africa (flats around Stellenbosch in the Western Cape)	Crouch (2012)
Isoëtes stevensii J.R. Croft		D-3	(2)–3–(4)	New Guinea	Croft (1980)
I. storkii T.C. Palmer	GH00021453, CM0102, NY00144272 (virtual herbarium sheets)	E-2	2	Costa Rica	JStor Global Plants virtual herbarium. http://plants.jstor.org/search?filter=name&so=ps_group_by_genus_species+asc&Query=isoetes+storkii, accessed Dec 28, 2016
Isoëtes subinermis Cesca – synonym of I. histrix		B-1	3		Cesca and Peruzzi (2001)
Isoëtes taiwanensis De Vol	Live plants	D-3	3	Taiwan	Chiang (1976)
Isoëtes tegetiformans Rury		?	2	USA (GA)	Rury (1978); Taylor (1993)
Isoëtes toximontana Musselman & J.P. Roux		A-1	3	South Africa (Giftberg endemic)	Cook (2004); Crouch (2012)
Isoëtes transvaalensis Jermy & Schelpe		?	3	Lesotho, South Africa (Limpopo, Mpumalanga, Free State, KwaZulu-Natal)	Cook (2004); Crouch (2012)
Isoëtes triquetra A. Braun		E-2	2	Peru	Pfeiffer (1922)
Isoëtes truncata (A.A. Eaton) Clute		?	2	CAN (Vancouver Island), USA (Alaska)	Pfeiffer (1922)
Isoëtes tuerckheimii Brause		E-2	2	Dominican Republic (Santo Domingo)	Pfeiffer (1922)
Isoëtes valida (Engelm.) Clute		E-2	2	USA (NC, TN, VA, WV)	Taylor (1993)
Isoëtes velata A. Braun		B-3	3	Italy (Algeria, Corsica, Sicily)	Pfeiffer (1922)
Isoëtes welwitschii A. Braun ex Kuhn		?	3	Angola, South Africa (Mpumalanga, KwaZulu-Natal), Madagascar	Pfeiffer (1922); Cook (2004); Crouch (2012)

Isoëtes wormaldii R. Sim		?	3	South Africa (Eastern Cape - small area between Grahamstown and East London)	Pfeiffer (1922); Cook (2004); Crouch (2012)
Isoëtes yunguiensis Q.F. Wang & W.C. Taylor		D-3	3	China	Qing-Feng et al. (2002)

Additionally, three accessions were coded as "unknown" for lobation due to identification uncertainties: (1) One accession of *I. savatieri* Franchet was collected in Uruguay, outside of the currently accepted range of the species Hickey et al. (2003), and we were unable to examine the specimen to determine its identity or lobation. (2) *Isoëtes histrix* Bory & Durieu is a known species complex (Bagella et al., 2011, 2015), and specimens fall in two different areas of our phylogeny. One is sister to *I. setacea* Lam., a position that is consistent with morphology and with existing hypotheses of their relationship Hoot et al. (2006); Bagella et al. (2011); Troia and Greuter (2014); we treated this accession as correctly identified. The second accession falls phylogenetically distant from the first, and presumably is misidentified; we were unable to ascertain its true identity or morphology, so we coded it as unknown. (3) *Isoëtes australis* R.O. Williams also shows up in two places in the phylogeny: Clade A and Clade D. The Clade A plant was collected and identified by Dr. Carl Taylor, an authority on *Isoëtes*, so we treated this accession as correctly identified and coded the Clade D plant as unknown. All character states and literature sources can be found in (Table 1.1).

Phylogenetic modeling

Ancestral character state reconstructions were performed on a posterior sample of 15,000 trees from Larsén and Rydin (2016) that we rooted on the bipartition between Clade A and the remainder of the genus (following Larsén and Rydin (2016)).

We employed reversible-jump MCMC (Green, 1995) in RevBayes (Höhna et al., 2016) to explore the space of all five possible continuous-time Markov models of phenotypic character evolution and to infer ancestral states. The reversible-jump MCMC sampled from the five models in proportion to their posterior probability. This approach enabled model-fit comparisons through Bayes factors (Kass and Raftery, 1995) and provided the opportunity to account for model uncertainty by making model-averaged ancestral state and parameter estimates (Madigan and Raftery, 1994; Kass and Raftery, 1995; Huelsenbeck et al., 2004; Freyman and Höhna, 2018). The five models of corm lobation evolution considered were as follows: a model with the rate of lobation gain and loss set to be equal (the 1-rate model); a model where the rates of lobation gain and loss are independent and nonzero (the 2-rate model); two irreversible models where the rate of either lobation gain or loss was fixed to zero; and lastly a model where both rates were fixed to zero. To test for directional evolution, we used nonstationary models of character evolution with root state frequencies that differed from the stationary frequencies of the process (Klopfstein et al., 2015).

Each of the five models was assigned an equal prior probability using a uniform set-partitioning prior. The root state frequencies were estimated using a flat Dirichlet prior.

The rates of corm lobation gain and loss were drawn from an exponential distribution with a mean of one expected character state transition over the tree ($\lambda = \tau/1$, where τ is the length of the tree).

The MCMC was run for 22,000 iterations, where each iteration consisted of 48 MCMC proposals. The 48 proposals were scheduled randomly from six different Metropolis-Hastings moves that updated the sampled tree, root frequencies, and corm lobation gain and loss rate parameters. The first 2000 iterations were discarded as burn-in, and samples were logged every 10 iterations. Convergence of the MCMC was confirmed by ensuring that the effective sample size of all parameters was over 600. The results were summarized and plotted using the RevGadgets R package (https://github.com/revbayes/RevGadgets). The scripts that specify our model, run the analysis, and summarize results are available in the code repository at https://github.com/wf8/isoetes.

Simulations

To test how many observed characters are necessary to reliably infer irreversible evolution on a phylogeny the size of ours, we simulated 10 datasets with each of 1, 5, 10, 50, or 100 characters per tip (for a total of 50 simulations; note that our empirical dataset has a single character per tip). Each dataset was simulated under an irreversible model with the mean rate of corm lobation loss set to the value estimated by the irreversible model using the observed corm lobation data (2.39 changes per unit branch length). We performed the simulations using RevBayes over the maximum a posteriori phylogeny from the same tree distribution used to infer the ancestral states. For each of the 50 simulated datasets, an MCMC analysis was run for 11,000 iterations, with the first 1000 iterations dropped as burn-in. The model used was identical to that used for the observed corm lobation dataset, except that for the simulated datasets we fixed the maximum a posteriori phylogeny instead of integrating over the posterior distribution of trees.

Figure editing

Figure plates of prepared slides and preserved materials were imported into Adobe Photoshop and adjusted using the levels and dodge tools to improve image definition and to reduce uneven color levels resulting from variations in background exposure. No features that were part of the original images were removed.

1.6 Results

Model-fit comparisons

The maximum a posteriori model of corm lobation evolution was the tri- to bi- irreversible model (which did not allow transitions from the bilobate to the trilobate state) with a

posterior probability of 0.38 (Table 1.2). This tri- to bi- irreversible model was weakly sup- ported over the 1-rate and 2-rate reversible models (Bayes factor = 1.26 and 1.21, respectively;(Kass and Raftery, 1995)); however, all three models were strongly supported over the bi- to tri- irreversible model. Since the Bayes factor support for the best-supported model over the next two was negligible, we focus mostly on the model-averaged parameter estimates and ancestral states.

Table 1.2: Comparisons of models of corm lobation evolution.

	Bayes factors				
Model	Model posterior probability	Bi- to tri-irreversible	Tri- to bi-irreversible	1-rate	2-rate
Bi- to tri-irreversible	0	–	<1	<1	<1
Tri- to bi-irreversible	0.38	>1000	–	1.26	1.21
1-rate	0.3	>1000	<1	–	<1
2-rate	0.32	>1000	<1	1.04	–

Model-averaged parameter estimates and ancestral states

The model-averaged estimated rate of transition from trilobate to bilobate forms was significantly nonzero (mean = 2.17 changes per unit branch length, 95% HPD interval: 0.015–5.69), whereas the rate of bilobate to trilobate transitions was not significantly nonzero (mean = 0.82, 95% HPD interval: 0.0–3.35; Fig. 1.2). The model-averaged maximum a posteriori ancestral state of *Isoëtes* was trilobate with a posterior probability of 1.0 (Fig. 1.3). The ancestral state of the New World clade ("Clade E-2"; Fig. 1.3) was bilobate with a posterior probability of 0.99. The bilobate morphology arose independently in six places over the phylogeny (Fig. 1.3). No reversals from bilobate to trilobate were inferred. All species with unknown base corm lobation characters (*I. stevensii* J.R. Croft, *I. habbemensis* Alston, and *I. hallasanensis* H.K. Choi, Ch. Kim & J. Jung) were derived from a trilobate most recent common ancestor, with a posterior probability near 1.0.

Simulations

For simulated datasets with a single character, the true irreversible model was at best weakly supported over the other models; the mean posterior probability of the true irreversible model was 0.37 (range: 0.22–0.44). As the number of characters increased, the posterior probability of the true model increased (Fig. 1.4). With 100 characters the support for the true model was strong; the mean posterior probability was 0.85 (range: 0.80–0.92).

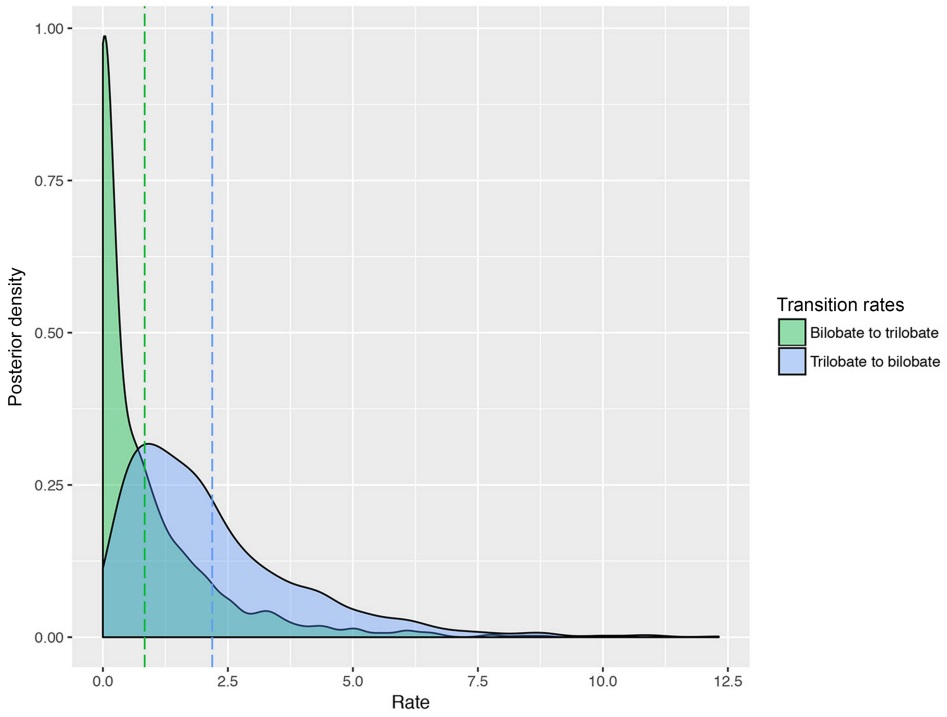

Figure 1.2: Model-averaged posterior densities of transition rates be- tween corm lobation states. Mean values are represented by dashed lines. The rate of transition from trilobate to bilobate forms was significantly nonzero (mean = 2.17, 95% HPD interval: 0.015–5.69), whereas the rate of bilobate to trilobate transitions was not significantly nonzero (mean = 0.82, 95% HPD interval: 0.0–3.35). Transition rates are reported in changes per unit branch length.

1.7 Discussion

Our analyses support the hypothesis of directional reduction in lobe number over time in *Isoëtes*, with the best-supported model being one of irreversible evolutionary reduction. Additionally, even when incorporating model and phylogenetic uncertainty and allowing for reversals, the model-averaged estimate of the transition rate from trilobate to bilobate was much higher than the estimate of the transition rate from bilobate to trilobate (the latter rate was not significantly nonzero). Furthermore, we found strong support for the ancestral state for all extant *Isoëtes* being trilobate, indicating that there have been multiple convergent reductions of the corms to bilobate. These bilobate forms are nested deeply within clades of trilobate plants. These results support the hypothesis that crown *Isoëtes* has continued a reduction in corm morphology from the larger arborescent lycophytes. Additionally, while the bilobate form has emerged multiple times throughout the phylogeny, it is the dominant morphology in only one major clade: the "American clade" (Clade E-2, sensu Larsén and Rydin, 2016; Fig. 1.3). In this clade, nearly all species have a bilobate morphology.

The simulations demonstrate that the relatively weak support for the irreversible reduc-

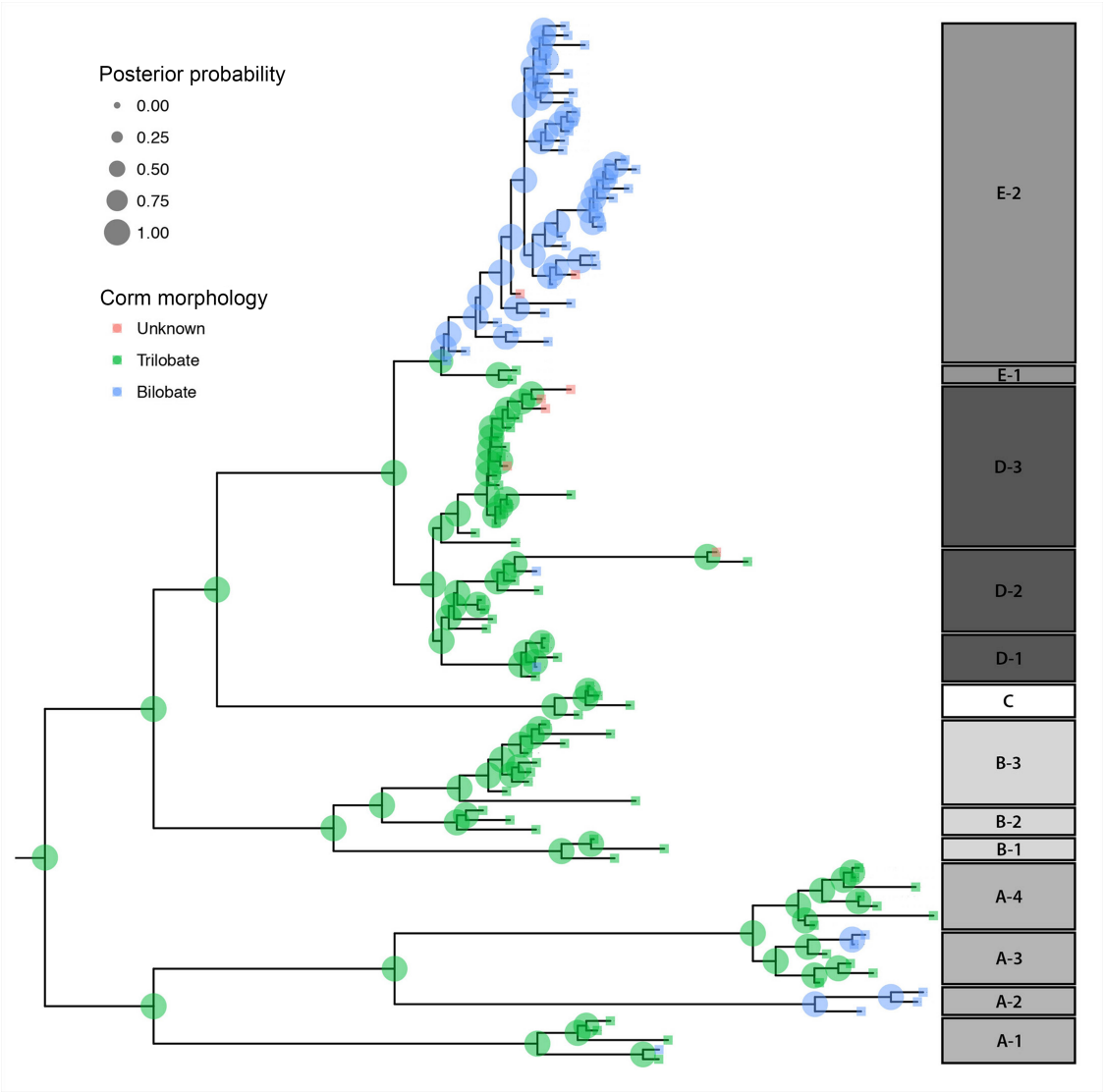

Figure 1.3: Bayesian model-averaged ancestral states of *Isoëtes* corm lobation inferred over the posterior tree sample from Larsén and Rydin (2016). Ancestral states are summarized on the maximum a posteriori phylogenetic tree. The size of the circles at each node represents the posterior prob- ability of the most probable ancestral state, and the color represents the state: green = trilobate corms, blue = bilobate corms, and red = unknown. Boxes on the right reflect subclade designations to broad subclade: A = Gondwanan, B = Laurasian, C = Italian, D = Austro-Asian, and E = New World or American. For specific placement of taxa into subclades, see Appendix 2.

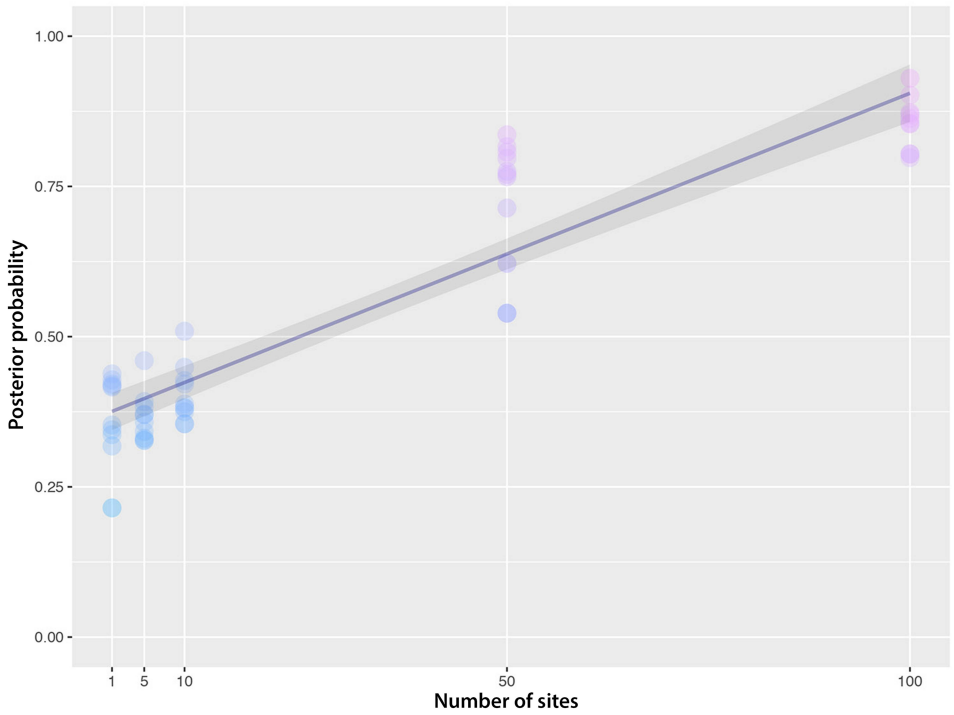

Figure 1.4: Statistical power to detect irreversible evolution as a function of the number of characters available. Each point plotted represents a different simulation replicate. The y-axis shows the posterior probability of the true irreversible model of character evolution. The x-axis shows the number of simulated characters (sites or columns in the data matrix). Ten replicates were simulated for 1, 5, 10, 50, and 100 character datasets, resulting in a total of 50 simulated datasets.

tion model over the reversible models is likely due to the inherent limitation in statistical power of a single observed morphological character over a phylogeny of this size. Repeating this analysis with a larger, more densely sampled phylogeny, or one that incorporates fossil data, might find stronger support for the irreversible model of corm morphology evolution. Nevertheless, it is in cases like this, where no single model is decisively supported over others, that reversible-jump MCMC and Bayesian model-averaging demonstrate their utility for testing phylogenetic hypotheses of character evolution (Huelsenbeck et al., 2004; Freyman and Höhna, 2018).

While the bilobate morphology dominates only Clade E-2, it does occur in several other areas of the phylogeny (Fig. 1.2 and Appendix 2). In areas where Clade E-2 co-occurs with others, such as South America and the west coast of the United States, corm lobation is useful in determining to which clade a plant belongs. However, outside the range of Clade E-2, other bilobate taxa occur as single species nested within larger trilobate clades. As such, using corm lobation outside the Americas to assign plants with unknown phylogenetic placement is not advisable, since they may represent other independent evolutions of the

character state.

When assessing corm lobation numbers, it is important to determine whether the corm lobe numbers are the base lobe numbers or additional lobes that have developed as the plant ages (Stokey, 1909; Karrfalt and Eggert, 1977b, F. D. Freund, personal observation). While (Karrfalt and Eggert, 1977a) reported a propensity for gaining additional lobes in their study of *I. tuckermanii* (68% bilobate, 30% trilobate, 2% tetralobate), other observers, working on other taxa, have not found this degree of variability (Engelmann, 1882, F. D. Freund, personal observation of >200 specimens of *I. howellii* and >400 specimens of *I. nuttallii*). In fact, we have observed a total of <10 *I. howellii* and *I. nuttallii* specimens that were not either bilobate or trilobate, respectively (F. D. Freund, personal observation). In *I. nuttallii*, the trilobate morphology was readily observable even in very young plants, including two sporelings, one of which was still attached to the megaspore (F. D. Freund, personal observation). These results suggest that the degree of variability of *I. tuckermanii* may be species-specific, and they add further support to the value of sampling multiple individuals when examining and designating corm lobation numbers within a species. If at all possible, it is additionally worthwhile to look at very young plants, especially sporelings, to determine the base lobe number.

1.8 Conclusions

Our results support the hypothesis that crown *Isoëtes* has continued an evolutionary reduction in corm morphology from the larger arborescent lycophytes, with the best-supported model being one of irreversible evolutionary reduction. However, results from our simulation study showed that a dataset of this size only has weak statistical power to support irreversible models of character evolution, emphasizing the need for broader sampling of extant taxa. When we accounted for the uncertainty in character evolution models by making model-averaged estimates, we found strong support for the hypothesis of directional evolutionary reduction in corm number, with the rate of lobe loss estimated to be much higher than the rate of lobe gain. Furthermore, we found strong support that the ancestral state for all extant *Isoëtes* was trilobate, indicating that there have been multiple convergent reductions of the corms to a bilobate state.

1.9 Acknowledgements

We thank E. Larsén and C. Rydin for providing the phylogenies used in this analysis; the UC Berkeley Herbaria and Department of Integrative Biology, the Rancho Santa Anna Botanic Garden and Herbarium, and the National Herbarium of Costa Rica for use of their collections to confirm corm lobation numbers; the reviewers for their valuable feedback; and the National Forest Service, Bureau of Land Management, California Department of Fish and Wildlife, and Sonoma Land Trust for allowing us to collect specimens on their properties.

Chapter 2

Evolution and demography of the Pacific Laurasian Clade of *Isoëtes*

2.1 Introduction

Patchy distributions can have a profound effect on evolutionary dymnamics at both the intra- and infra-specific levels (Aguilar et al., 2008; Belletti et al., 2008; Finger et al., 2013; Matesanz et al., 2017; Wood et al., 2018). A common effect of patchy distributions is the development of population structuring, where the process of natural selection and/or genetic drift can fix different genotypes at an accelerated rate due to the smaller population size and localized conditions at each site (McCraney et al., 2010; Pither et al., 2003). Given enough time, such structuring and isolation can lead to the accumulation of enough differences between populations that these sub-populations become new species (De Queiroz, 2007; Rannala and Yang, 2017). Isolation and structuringcan happen even in small sub-populations that are isolated from larger central ones (Aars and Ims, 2000; Roland et al., 2000; Bollmer et al., 2005; Whiteman et al., 2007; Hedrick, 2019), in naturally scattered and fragmented populations (Gaudeul et al., 2000; Belletti et al., 2008; Kuss et al., 2008), anthropogenically degraded ones (Segelbacher et al., 2003; Van Rossum and Triest, 2006; Scobie and Wilcock, 2009; Gordon et al., 2012; Sloop et al., 2012; Bovee et al., 2018), or a combination of any of the three.

Reproductive strategy can also play a part in population structuring and phylogenetic divergence. Species with patchy distributions in specialized habitats may appear to have small local populations that are primed for speciation, but be functionally part of a larger, mostly continuous population due to good dispersibility of either individuals or reproductive propagules (Ibrahim et al., 1996; Pither et al., 2003; MacDonald et al., 2018). Conversely, species with apparently large, widespread distributions may have highly structured populations due to poor dispersibility, unexpected barriers to geneflow, limitations on the movement of individuals across the range, or local adaptations due to habitat heterogeneity selecting against certain genotypes at a local level, even with gene flow between populations (Rasmussen and

Brødsgaard, 1992; Gehring and Delph, 1999). Effective population size is also important to the study of patchy populations, as it is possible to have few individuals that have high genetic diversity / low inbreeding among them, or have many individuals with depauperate genetic diversity due to inbreeding, genetic drift, or recent colonization (Segelbacher et al., 2003; Van Rossum and Triest, 2006). All of these factors can mean that predicting the genetic structure among organisms with patchy distributions without understanding their dispersibility, lifespan, or ecology can be very difficult (Gaudeul et al., 2000; Kuss et al., 2008; Zhu et al., 2009; Hou and Lou, 2011).

For sedentary organisms such as plants, having fragmented populations creates many challenges when it comes to maintaining genetic diversity and interconnectivity within their species. Unlike most terrestrial and many marine animals, plants can only disperse via propagules (spores, pollen, seeds, or vegetative structures like gemma or plantlets). Depending on the nature and efficiency of these dispersal vectors, this movement can be either short or long range, and can even differ if there are multiple types of propagules within a given taxon (e.g., mega- and microspores for free-sporing heterosporous plants, or pollen and seeds for spermatophytes; Rasmussen and Brødsgaard, 1992; Van Rossum and Triest, 2006; Grivet et al., 2009; Jordano et al., 2010).

Measuring and predicting the effects of fragmentation on the evolution of plant species is also complicated by the ability of many plant species to asexually reproduce either vegetatively, or apomictically. In the former, the plant is able to create identical clones of itself by dividing its body in some way, producing small vegetative propagules such as gemmae or plantlets that it releases into the environment and which then grow into new individuals, or by producing large clusters of clonal individuals through rhizomes or stolons. In apomictic reproduction, the process of meiosis is arrested in the sporocytes, and instead produces embryos that are genetically identical to the parent plant. Some plants are also capable of self-fertilization, where the plant will sexually reproduce with itself, which will result in higher-than-normal levels of inbreeding. These reproductive strategies can greatly decrease standing genetic diversity, particularly if asexual reproduction or selfing is the primary form of reproduction within a population. Conversely, some plants have adaptations that discourage selfing, such as genetic self-incompatibility, dioecy (separate male and female plants), dichogamy (different timing of male and female part maturation / receptivity within a flower or on an individual) or herkogamy (spatial separation of stamens and carpels within a hermaphroditic flower), which can increase or guarantee outcrossing, depending on how strong the effect is (Bawa, 1980; Thomson and Barrett, 1981; Takebayashi et al., 2006; Kitamoto et al., 2006; de Vos et al., 2012; Kalisz et al., 2012). While these properties are generally thought to increase outbreeding and therefore help preserve genetic diversity, their effects may be minimized if the plant's ability to move propagules outside a local area is too low (Rasmussen and Brødsgaard, 1992; Van Rossum and Triest, 2006), or in the case of biotically pollinated species like many flowering plants and some gymnosperms, if their pollinators do not move between populations at a rate sufficient to offset localized inbreeding Schmitt (1980).

Taxa that have evolved to either promote outcrossing or selfing can face genetic conse-

quences for these adaptations: asexual reproduction and relaxed selfing barriers can mean a greater chance of successfully reproducing when there are fewer individuals in a population, but at the costs of reducing the genetic diversity potentially needed to adapt to natural selection in the long term, while stronger outcrossing can potentially lead to greater adaptability and reduced inbreeding, but can also result in population collapse if there aren't enough individuals to successfully outcross with, or in the case of pollinator-dependant reproducers, ones whose pollinators do not readily move between sub-populations (Kitamoto et al., 2006; Van Rossum and Triest, 2006). While naturally fragmented populations must cope with the above mentioned dangers of excess selfing or small popualtions in obligate outcrossers, they are a much larger threat to taxa that have had their range anthropogenically fragmented, as they are unlikely to have had the time to adapt to their newly reduced range. The genus *Isoëtes* L., which is the focus of this study, is one such group of organisms: naturally fragmented populations that have had their original range both drastically reduced and further fragmented by anthropogenic destruction of the environment.

Isoëtes are small, perennial, semi-woody geophytes that occupy mesic to aquatic habitats on every continent other than Antarctica, as well as occurring on several oceanic islands (Pfeiffer, 1922; Larsén and Rydin, 2016; Troia et al., 2016, 2019). These plants are inconspicuous, and are often mistaken for grasses, rushes, or sedges (pers obs of herbarium sheets and on iNaturalist.com). However, unlike the groups they are often mistaken for, these are not seed plants, but free-sporing, heterosporous lycophytes. Taxonomically, lycophytes are the sister lineage to the euphyllophytes, which consists of the spermatophytes (extant seed plants) and monilophytes (eusporangiate and leptosporangiate ferns, horsetails, and whisk-ferns). As heterosporous plants, *Isoëtes* differ from most other free-sporing lineages by producing two highly dimorphic spores: a small male microspore (a spore that will produce a gametophyte that will only ever produce sperm), and a large female megaspore (a spore that will produce a gametophyte that will only ever produce egg cells). *Isoëtes* reproduce by releasing both their spores into the environment where their free swimming, flagellated sperm, produced by the microgametophyte, must swim to the archegonia of the megagametophyte. As a consequence, both of their spores must germinate in an area with enough water for the plants to complete their sexual reproduction. In aquatic species of *Isoëtes*, this isn't as strong of a limitation, while terrestrial/amphibious species must account for periods where standing water is unavailable. In addition both megaspores and microspores must mature in the same area at the same time to produce a new sporophyte. This differs from the desynchronized dispersal in seed plants, where the microgametophytes are dispersed in the form of pollen and deliver sperm via a pollen tube to fertilize the eggs of a megagametophyte retained on the parent plant, which produces an embryonic sporophyte that is dispersed later in the form of a seed.

The question of how these spores move between populations, and its effect on their evolutionary history has only been studied sparingly. In the case of riparian and riverine species, upstream to downstream movement of spores is considered to be one possible route of gene-flow, as the water more easily moves the spores with the flow of the system (Chen et al., 2010). In aquatic environments, such as lakes, ponds, and vernal pools, intra-lacustrine

movement of spores occurs when the sporophylls are shed and the sporangia rupture, while inter-lacustrine movements have been speculated to occur by rare flooding events that connect the waterbodies, or by zoochory as the plants have been observed being eaten of ducks and geese, potentially moving spores to other water bodies through their guano, or by spores clinging to their feet and feathers (Freund, pers obs; Musselman et al., 1997; Pfeiffer, 1922; Schafran, 2019). In terrestrial species, the mechanism of spore movement is less well understood, though it has been observed that sporangia remain in the soil during the dry season and are expelled when the rain returns via expansion of hydrophyllic mucilage produced at the abscission zone of the previous year's leaf crown (Osborn, 1922), though how new populations are established or the amount of gene-flow between populations is still unknown.

Regardless of how the spores move, *Isoëtes* populations are naturally fragmented in the landscapes. These plants typically occur at low to high densities with a patchy, scattered distribution in a matrix of other vegetation or inhospitable xeric conditions, though in some cases they can form large, nearly monogeneric patches (Fig. 2.1). For some species, these conditions are highly specialized and restricted, such as *I. tegetiformans* Rury, *I. inflata* E.R.L. Johnson, *I. lithophila* N.E. Pfeiffer, or *I. eludens* J.P.Roux, Hopper & Rhian J.Sm., which all occur in ephemeral pools on granitic outcrops, while other species, such as *I. nuttallii* Englm., can be found throughout large, mesic meadows in addition to small, patchy seeps and thin mossy banks along creeks and springs (Fig. 2.1 A–N). This fragmented distribution has become more pronounced with the destruction of their habitats by anthropogenic development, changes to local hydrology, or indirectly through reduced precipitation as a consequence of climate change. These alterations can lead to disruptions in gene-flow between populations, potentially having a drastic effect on the evolutionary trajectory of any given species, or leaving it vulnerable to loss of standing genetic diversity.

Four of the described species occurring in the west coast region of North America (*Isoëtes bolanderi* Engelm., *I. echinospora* Durieu, *I. howellii* Engelm. and *I. occidentalis* L.F. Hend) belong to the American or New World clade, which occurs throughout North and South America, but also has members with a circumboreal distribution, and includes the only native *Isoëtes* on the Hawaiian archipelago (Baldwin et al., 2012; Hoot et al., 2006; Larsén and Rydin, 2016; Rydin and Wikström, 2002; Pfeiffer, 1922; Troia et al., 2016). The other three species (*I. minima* A.A. Eaton, *I. nuttallii* and *I. orcuttii* A.A. Eaton) are part of a highly disjunct sub-clade nested within the Mediterranean or Laurasian clade, the majority of which occurs in the Mediterranean region of Europe and North Africa as well as India (Hoot et al., 2006; Rydin and Wikström, 2002; Larsén and Rydin, 2016; Pfeiffer, 1922). These clades are highly phylogenetically divergent, having last shared a common ancestor ca. 103 MYA (Rydin and Wikström (2002); Hoot et al. (2006); Larsén and Rydin (2016); Troia et al. (2019)). Members of this disjunct group, which I refer to as the Pacific Laurasian Clade (PLC) will be the focus of this study.

Of these three named species, the most widespread is *I. nuttallii*, which occurs from southern British Columbia to Northern Baja California west of the Cascades, Sierra Nevada, and Peninsular Range, with a small number of sites east of the Cascade Range (Fig. 2.2). *I. minima*, the taxon with the fewest known occurrences, can be found east of the Cascades

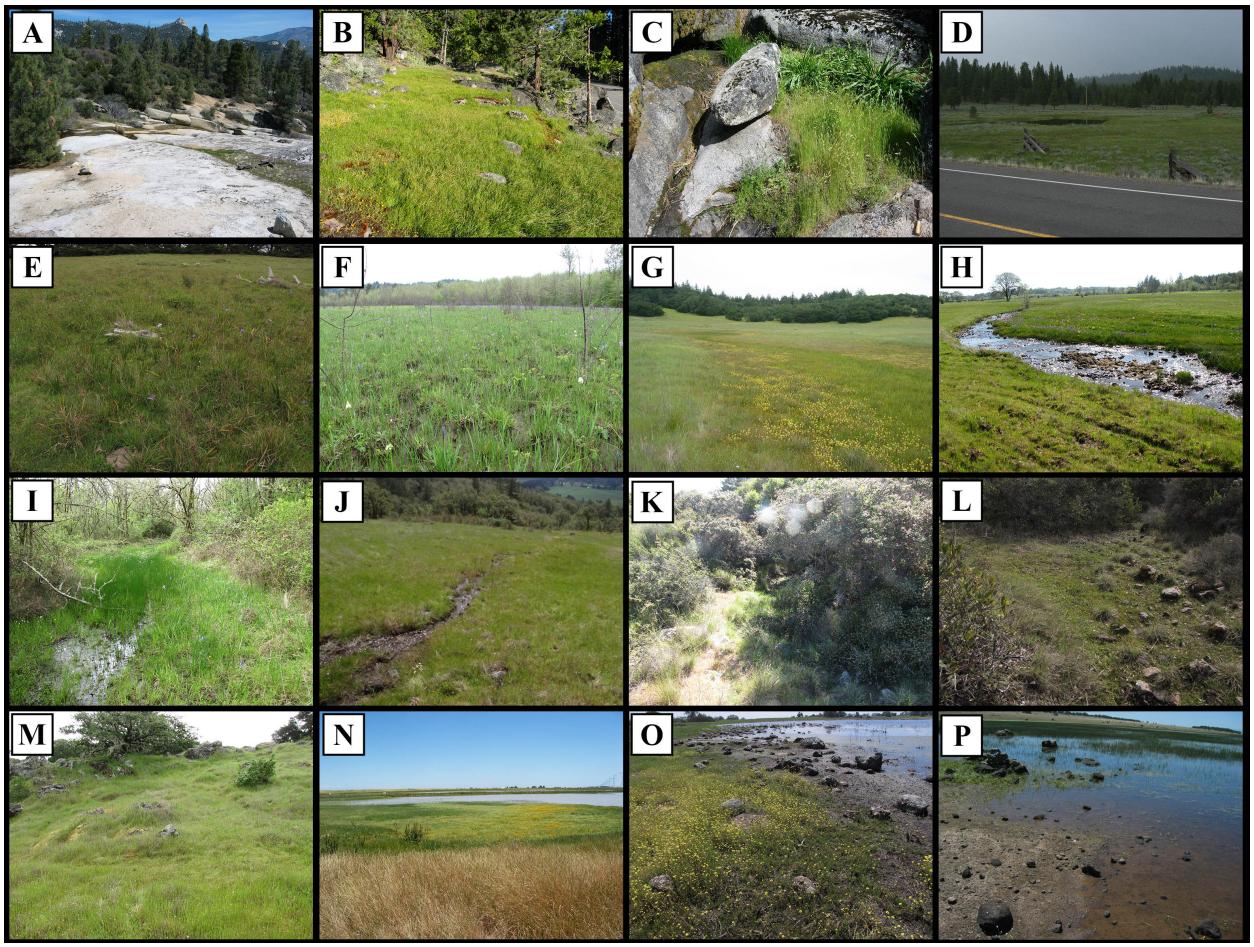

Figure 2.1: Habitats where Pacific Laurasian Clade plants can be found growing. A-C: Mossy banks; D–G: Wet meadows; H-J: Ripparian; J–M: Wet hillsides; N–P: Vernal Pools

along the edge of the Colorado Plateau and Great Basin Desert in Washington, Oregon, California, and Idaho (Fig. 2.2. Both *I. nuttallii* and *I. minima* occupy predominantly terrestrial habitats, though in some areas, *I. nuttallii* can be found growing submergent in creeks and wetlands (Fig. 2.1, Table 2.1). The third described member of the PLC is *I. orcuttii*, a California Floristic Province endemic found throughout the Central Valley, Sierra Nevada Foothills, Los Angeles Basin, and into northern Baja California (Fig. 2.2). Unlike the previous two members of the clade, *I. orcuttii* is described as an amphibious vernal pool endemic (Fig. 2.2). All PLC plants are distinguishable from their American Clade cousins by their trilobate corms (Freund et al., 2018), while the individual species can be identified through a combination of spore, leaf, and habitat characters (Table 2.1). Despite their large ranges, individual populations of these plants tend to be highly discontinuous and fragmented, with some occupying areas less than ca. 15 square meters (Freund, pers obs). Coupled with human development, much of their historical distribution has been extirpated,

severely reduced, or fragmented into discontinuous patches (Bauder and McMillan, 1998; Gordon et al., 2012).

Table 2.1: Morphological characters of the three described species of Isoëtes in the Pacific Laurasian Clade.

Character	*Isoëtes nuttallii*	*Isoëtes orcuttii*	*Isoëtes minima*
Habit	Terrestrial to aquatic	amphibious	Terrestrial to amphibious
Habitat	Damp prairie, wet meadows, seeps, mossy banks of creeks flowing over granite, submergent aquatics in shallow creeks and waterways.	vernal pools	Damp prairie, wet meadows, seeps and vernal pools.
Latitudinal range	32.721500° : 49.950000°	31.541700° : 42.453409°	41.986479° : 49.300000°
Longitudinal range	-123.729700° : -116.561900°	-122.961900° : -116.453333°	-120.176320° : -116.249435°
Leaf number	13-60	6-14	6-12
Leaf length	8-17 cm	2-6.5 cm (rarely 10cm)	2-4cm
Leaf diameter	>1 mm	<1mm	.67-.74mm
Leaf description	rigid, ± brittle, tapered to tip	long, fine, errect, soft, flexible, tapered to tip	long, round, slender
Leaf scales	present	present	?
Sporangium length	4-7mm	2-5mm	4mm
Sporangium coloration	unpigmented	unpigmented	unspeckled
Megaspore ornamentation	± tubercled, ± shiny	small, indistinct papillae or smooth	short, slender, blunt, distinct spinules
Megaspore size	400-528 µm	216-360 µm	290-350 µm
Megaspore color	white to light grey	grey at maturity, brown/black when wet	grey to white
Microspore ornamentation	papillose	spinulose to smooth	papillose or sparsingly spinulose
Microspore length	25-20µm	21-27µm	26-31µm
Microspore color	brown	brown	white
Velum coverage	complete	complete	66-75%
Corm length	?	?	2-3mm
Corm width	?	?	3-4mm
Stomata	numerous	present	present
Bast bundles	3	0-2	4
Ligule	small, triangular	triangular	triangular, slightly elongated
sources	Pfeiffer 1922, Frye and Jackson 1913, Jepson Online, Flora of North America, personal observations	Pfeiffer 1922, Jepson Online, Flora of North America, personal observations	Pfeiffer 1922, Eaton 1898, Frye and Jackson 1913, WNHP 2005

In order to better understand how this fragmented distribution has affected the evolutionary history of these species, I will be taking a two pronged approach to this study. The first aspect is the phylogenetic relationships among and within the described species, and whether the taxa as described reflect an accurate picture of the systematic diversity of the PLC. In order to do so, I will be utilizing multi-gene phylogenetics in a coalescent framework (Kingman, 1982; Nordborg, 2019; Hudson et al., 1990) to infer both the intra- and infra-species relationships within the PLC. While more time and computationally intensive, coalescent based phylogenetic methods avoid several pitfalls of the faster single gene methods. Specifically, coalescence based phylogenies seek to mitigate or outright negate the effects of gene-tree/species-tree discordance, which arise through incomplete lineage sorting, introgression, and/or horizontal gene transfer (Rannala and Yang, 2017, 2003). By developing this phylogeny, I will be able to test not only identify which, if any, of the described taxa reflect true species, but also illuminate the relationships between them, the patterns and ages of the diversification within the clade, and the potential effects the plant's scattered distribution has had on their infra-species evolution. The second aspect will be to examine the

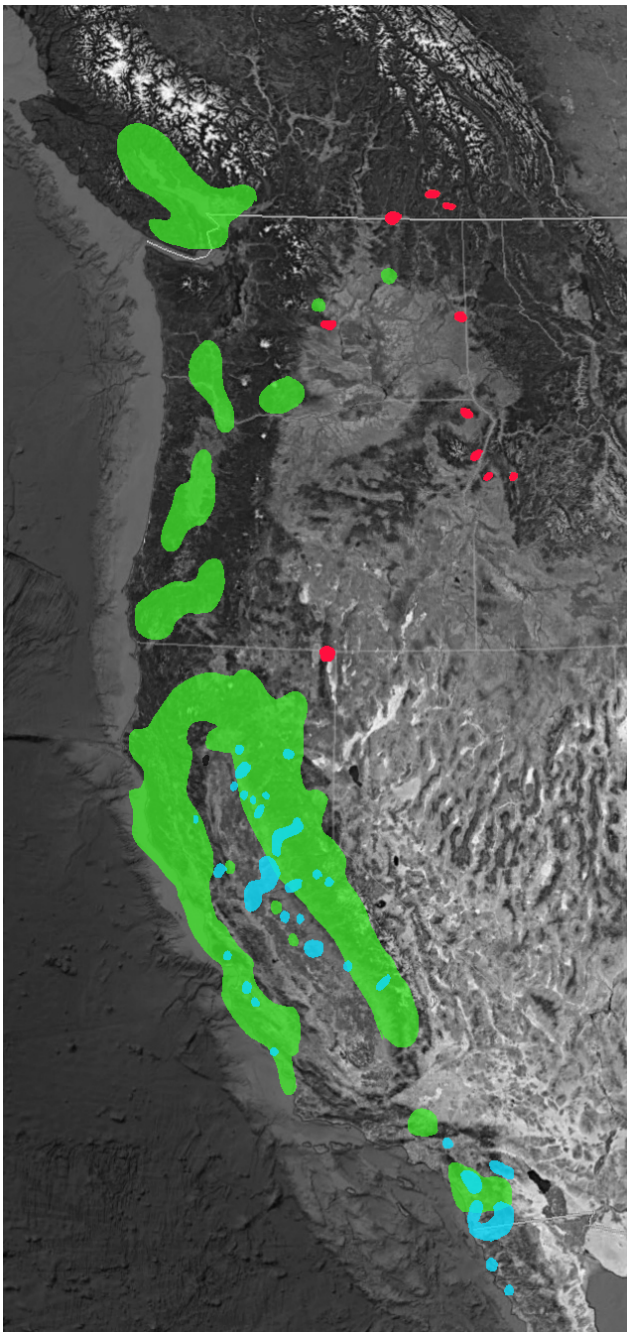

Figure 2.2: Known distribution of *Isoëtes* species from the Pacific Mediterranean Clade based on historical and contemporary collections. Range data gathered from Consortium of California Herbaria, Oregon State University and University of Washington. Green - *Isoëtes nuttallii*, Blue - *Isoëtes orcuttii*, Red - *Isoëtes minima*.

within population structuring and genetic diversity of the PLC through admixture plotting

(Korneliussen et al., 2014; Peter, 2016). In doing so, I intend to determine the effects of the plant's scattered distribution on local genetic diversity, the degree of intermixing between the different populations. Combined, the phylogeny and the population genetic analyses can inform us about the natural history of the PLC, their capacity for adaptation, and the range of its ability to disperse and colonize new areas.

To address the phylogenetic relationships within and the population structure of the PLC, I sampled 15 localities of *I. nuttallii* and four localities of *I. orcuttii* from across most of their range, including putative type localities for both species. In addition, I received materials for a single population of *I. minima* from NE Oregon, and was able to add a small number of samples from this population to my analysis. From these populations, I hope to assess both the phylogenetic relationships between the different localities, the standing genetic diversity within them across their range, identify any admixture between the populations.

2.2 Methods

Field Collections:

Collections sites were chosen from across the geographic range of *I. nuttallii* and *I. orcuttii* using records from the Consortium of California Herbaria (https://ucjeps.berkeley.edu/consortium/), Consortium of Pacific North West Herbaria https://www.pnwherbaria.org/, Tropicos https://www.tropicos.org/home, the National Museum of Natural History https://collections.nmnh.si.edu/search/botany/, and iNaturalist https://www.inaturalist.org/. Herbaria records referenced by the Pfeiffer (1922) monograph and Tropicos.org were cross-checked with type descriptions to identify, or make a best estimate for, the type localities for both taxa.

At each field site where *Isoëtes* were found, a quick, informal survey of the size of the population was done by finding the rough extent of the area occupied by the plants. Using a weeding spike, every twentieth plant was carefully extracted from the ground, until ca. 20–30 plants were collected, or 5% of the population, whichever was lower. Sites were georeferenced using a hand-held GPS, and photographed with either a Canon PowerShot S45 or Canon PowerShot SX20 IS to record the habitat and the plants *in situ*. Specimens were gathered in zip-top bags and stored in a cooler until they could be processed for morphological observation and DNA extraction.

Specimens were entered into the Fern* Labs DNA Database (Pryer et al., 2021). Vouchered herbarium specimens were submitted to either the RSABG herbarium (prior to 2015) or UC/Jepson herbaria (post 2015). Any samples not used as part of the voucher were placed into coin envelopes and stored in silica for archival preservation.

DNA extraction and sequencing:

Genomic DNA was extracted from 1–10g of either fresh or silica preserved leaf materials using the standard CTAB extraction protocol (Appendix 2). Samples were then sent to the University of Florida for sequencing was done at RAPiD Genomics (Gainsville, Florida, USA) using the protocol from Breinholt et al. (2021). After bead-based cleanup at the facility and normalization of the DNA to 250ng and mechanical shearing to 300 bp in length, Illumina libraries were constructed by repairing the 3'-ends of the blunt-end fragments with adanine residue as described in Bentley et al. (2008). Following the repair of the 3'-ends, barcodes were ligated to the the DNA fragments, and amplified for 9-11 PCR cycles, and pooled in batches of 16 barcoded libraries equimolarly to reach a total DNA concentration of 500ng for hybridization. Samples were enriched using a modified version of Gnirke et al. (2009) using the GoFlag probes. The probes themselves are a proprietary probe-set developed by RAPiD that target 408 single-copy exonic regions using the transcriptomes produced by the 1KP project (Leebens-Mack et al., 2019). Probes were developed from exonic regions that were identified as single- or low-copy within the 1KP database, were at least 120bp in length, and had at least 65% average pairwise identity across six of the 23 genomes used in their development, which had to include at least three of the represented taxonomic groups (bryophytes, lycophytes, gymnosperms, deeply divergent angiosperms, monocots, early diverging eudicots, Caryophyllales, asterids, and rosids). Post enrichment, the samples were re-amplified for 6-12 cycles, and sequenced using an Illumina Hiseq 3000 (Illumina, San Diego, California, USA) with paired-end 100-bp reads. Adapters were removed from the Illumina sequence reads using Trim Galore! version 0.4.4 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), with only pairs of reads with both the forward and reverse read that were at least 30 bp long were retained. Once adapters were removed, the sequences were ready for assembly.

Sequence assembly:

Of the samples sent to RAPiD, 140 samples from the in-group produced usable sequence (Appendix 2). Raw sequencing reads that had been processed to remove the ligating adapters were assembled and phased using a modified version of the HybPiper pipeline. First, an initial recovery of targeted sequences was performed using HybPiper version 1.3.1 (Johnson et al., 2016) with orthologs from the GoFlag dataset. Because gene recovery was limited for many samples, we generated an *Isoëtes*-specific reference file from the individual FF290_10, which had the highest number of reads mapping to targeted loci (664,171). We generated reference sequences by using Augustus (Stanke et al., 2008) to identify coding regions in the "supercontigs" (coding regions plus flanking non-coding regions) extracted by intronerate.py in HybPiper (https://github.com/mossmatters/HybPiper. HybPiper was then re-run for all samples using this new *Isoëtes*-specific reference file.

Genes that returned two or more assembled contigs that each span >75% of the length of the target gene for the GoFlag genes were identified using the HybPiper "paralog warn-

ings" and scripts paralog_investigator.py and paralog_retreiver.py (https://github.com/mossmatters/HybPiper; this analysis revealed six genes with paralog warnings from 10 or more samples, and these genes were removed from further analysis. There were also substantial paralog warnings for two samples from the Hog Lake collection site (FF254_3 and FF254_2); further examination of these anomalous samples in AliView (Larsson, 2014) revealed that half of the recovered loci from 254_3 had features like the American Clade as opposed to the PLC, and may represent either contamination or possible inter-clade allopolyploids, while 254_2 merely had an excess of paralogs. As such, these two individuals were excluded from the downstream analysis.

For variant calling the "all-to-one" protocol was utilized (Slimp et al., 2021) using the FF290_10 individual as a reference sample. For each individual reads were mapped to the reference sample and duplicate reads were removed using MarkDuplicates in GATK v 4.0 (McKenna et al., 2010). GenotypeGVCFs was used on each individual and then HaplotypeCaller jointly for all individuals. A hard filter was run on the initial variants (QD $<$ 5.0 || FS $>$ 60.0 || SOR $>$ 3 || MQ $<$ 55.0 || MQRankSum $<$ -12.5 || ReadPosRankSum $<$ -8.0) to generate variant call files (VCFs) for all samples.

As both phylogenetic methods used in the downstream analyses either can use or require phased haplotype sequences, phased sequences were generated for gene using read-backed phasing in WhatsHap (Schrinner et al., 2020). Because WhatsHap is expecting full-length genome scaffolds, it may generate falsely phased variants in target capture data; for example, if the targeted coding region contains a large intron preventing read-backed phasing. To overcome this, the "phase blocks" annotation generated by WhatsHap was used to delete sequences outside the largest phaseable block for each gene, using haplonerate.py (available at: github.com/mossmatters/phyloscripts) (Kates et al., 2018). This generated two haplotype sequences for every individual at every gene.

Sequence alignment:

Sequences generated by HybPiper were loaded into Aliview (Larsson, 2014). Sequences were first aligned using the muscle alignment algorithm (Edgar, 2004), followed by manual realignment of mismatched regions. In the case of regions that were deemed to represented assembly or sequencing error, usually due to being present in only one or two sequences out of the entire alignment, were removed them using the exclude function. For regions that couldn't be aligned to the main sequence but appeared to not be sequencing error due to showing up in more than two sequences, these unalignable portions of the sequence were shifted to not be included in the main sequence. For sequences on the ends of the alignment, they were simply sifted out of the main frame of the sequence. For sequences that could not be moved outside the main alignment without disrupting the rest of the alignment, gaps were inserted into the main alignment of equal length to the unalignable portion, which was then shifted into the newly created gap. Using a custom R-script (Appendix 3), the .nexus files were converted into .fasta format, as the latter is necessary for downstream analysis, yet doesn't allow for excluded regions.

Sequences will be deposited in GenBank at the time of final publication.

Population genetics:

Raw sequences were demultiplexed and seperated by individual, then aligned to a quality filtered *Isoëtes* reference genome from the 1KP project using the software bwa with the defalut parameters (Li, 2013; Leebens-Mack et al., 2019). The resulting reference mapping file was then used in ANGSD (Korneliussen et al., 2014) and NGSAdmix (Skotte et al., 2013). First, ANGSD was used to call genotype liklihoods with the following arguments: -doHWE 1 -domajorminor 1 -doMaf 3 -GL 1 -doGlf 2 -SNP_pval 1e-6. This proced a genotypes liklyhood .beagle file, which was then imported into NGSAdmix. In NGSAdmix, we ran 10 replicate admixture analyses for each value K = 1 – 8. The likelihood results of these replicate runs were then input into Clumpak (Kopelman et al., 2015) to find the K value that had the highest likelihood score from the replicated. Finally, the .beagle file from the initial ANGSD analysis was rerun in NGSAdmix, with the optimal K value of 4 to infer the genetic composition of each individual in the admixture analysis.

SVDQ Analysis:

Preliminary phylogenetic analyses were run in SVDQuartets (Chifman and Kubatko, 2014, 2015), a multi-speceis coalescent method implimented in PAUP*. Prior to running SVDQ on the dataset, the individual sequence alignments had to be formatted into a paired set of "sequences" and "control" nexus files, both of which were generated using a custom R-script (Appendix 4). These control files contained all of the sequences to be included in the analysis, while the control file contained the taxa partitioning for the analysis, the partitions for the loci included in the sequences file, and the parameters of the model to be run in PAUP*. All analyses were run on UC Berkeley's SAVIO high-performance computational cluster. With SAVIO's 72hr job run-time limit, analyses were parameterized in such a way that they could be completed within these time-limits. Six different analyses were run in SVDQ, two for each partitioning scheme (described below). The first run type was the "All Quartets Examined" run, which was conducted with the following parameters: nst=6, rmatrix=estimate, basefreq=empirical, rates=gamma, ncat=4, shape=estimate, pinvar=estimate, taxpartition=species, evalQuartets=all, and all 215 genes. The second run type was the "bootstrap" run, which was done in sets of three independent replicate runs. Bootstrap runs were conducted with the same parameters as the All Quartets Examined run, with the exception of their evalQuartets paramter, which had to be lowered in order to be completed under the 72-hr computational time limit of SAVIO, and allowed to run for 1000 bootstrap replicates.

The first partitioning clustered all individuals from a given collection site into a single "species" (hereafter referred to as the "Sites Grouping"), and conducted the bootstrap runs with an evalQuartets = 570000, which equated to around 0.3058% of the possible quartets. Monophyletic clades of multiple collection sites from each analysis were identified and given

designations. If all samples from an identified species were accounted for in a single clade, the name of that species was used as the clade's name. If the prior criteria was not met, either because multiple, non-monophyletic clades were identified as a given taxon, or the clade contained specimens identified as different taxa based on their morphology, the clade was instead named based on the most common geographic area where the specimens occured: i.e. a clade that contained plants field identified as *I. orcuttii* and *I. nuttallii* from the different sites that formed a single, larger clade that mostly occur in Southern California were designated as the "Southern California" clade.

The second partitioning treated each individual plant as a "species" in the analysis (hereafter referred to as the "Individual Grouping"), with the bootstrap runs conducted on 570000 (0.2325% of the possible quartets). Monophyletic clades were named using the same criteria as the Sites Grouping. Topologies between the two were then compared to determine where the congruence between the two topologies were.

Once both the Sites and Individual Grouping analyses were done, a third set of analyses were done with each major clade identified as a putative "species" (hereafter referred to as the "Major Clade Grouping" or just "Major Clades"). The bootstrap analysis was done with a evalQuartets=100000 (equaling 0.1706% of the possible trees).

After all three analyses had run to completion, the bootstrap replicates for a given analysis type were combined, and a custom R-script was run to calculate the bootstrap support values for the "All Quartets Examined" topology (Appendix 5). Plants of these two sites (FF290 & CR5205) showed morphological characters consistent with hybridity or polyploidy in the form of malformed or irregularly sized spores, (see Figs. 3.23 and 3.24), and these sites consistently had low support for their placement within the phylogeny; to examine the effect of their exclusion on the inferred topologies, a second set of analyses with these collection sites removed were run under the same parameters as the previous runs.

Sequence files used in the analyses can be found in Supplements 1 & 2.

BPP Analysis

After completing the SVDQ tree inference, the topology from the SVDQ analysis with the outgroups excluded was used as part of two Bayesian Phylogenetics and Phylogeography (BPP) age and population size inferences. BPP is a coalescent-based phylogenetics program that can be used for both species inference and for estimating population size (theta) and relative divergence times (tau; Rannala and Yang, 2017, 2003). Within the program, the A00 analysis attempts to infer these two parameters on a fixed tree using the multispecies coalescent under a general time reversible (GTR) substitution rate model with a strict molecular clock on neutrally evolving loci. The A00 analysis was run twice: once with all ingroup populations included, and once with the two putative hybrid collection sites excluded (FF290 & CR5205). The analyses were run on SAVIO, UC Berkeley's High Performance Computational Cluster, with each individual analysis using 20 cores. Due to the 72hr time-limit on SAVIO, the data had to be sub-sampled in the following way: include only four phased individuals per collection site, and 30 randomly selected loci from the 215 in the dataset,

which were selected at random using a custom R-script. Selected loci were kept consistent between both analyses. A00s were run with a 50,000 sample burnin, followed by a 150,000 generation long MCMC. Control file can be found in Appendix 6.

Once completed, the resulting phylogenetic trees had their root age scaled to 100 in FigTree in order to have their branch-length lables reflect the percentage of the total tree length that each internode represents in coalescence time.

Sequence and map files for the BPP analyses can be found in Supplements 3 – 6.

2.3 Results

Field Collections

The following specimens were collected from, or in close proximity to, the estimated type localities: FF286 for *I. orcuttii* and FF283 for *I. nuttallii*.

SVDQ results

In the Sites and Individuals Groupings of my SVDQ runs, the All Quartets Examined tree combined with the bootstrap analyses recovered nine clades with high bootstrap support. In addition, the alternate placement trees where the putative hybrid collection sites (FF290 and CR5205) were moved to alternate parts of the phylogeny regularly returned lower support values than those of the All Quartets Examined tree. For the the sites and individuals analysis, support within the five main clades with multiple collection sites (SC, SN, PR, CR, WV) was low, even when the support for the main clade was high. This indicates that there wasn't a particular topology that was well represented within each Major Clade, even if the Major Clades were well supported overall (Fig. 2.3, 2.4, 2.5). In addition, the species were not monophyletic in relation to my initial identifications based on the Flora of North America and the Jepson Manual (Taylor, 1993; Baldwin et al., 2012). Results will be discussed in the context of the Major Clades Grouping topologies and support values, though Site and Individual groupings will be discussed when relevant.

The Hog Lake (HL) clade, comprising samples from a single population at Hog Lake, Tehama County, was the most genetically distinct of all the plants included in this study aside from the outgroup samples (Fig. 2.5, to the degree that they were visually identifiable as distinctive from the rest of the sampled sequences when examined in AliView. While plants of this population was initially identified as *I. orcuttii* based on morphology and habitat, they are placed as sister to the rest of the ingroup, including the plants identified as *I. minima*, *I. nuttallii*, and all other plants identified as *I. orcuttii*.

The Minima (Min) clade comprised two individuals of *I. minima* from north east Oregon. These two individuals form a clade that is distinct from the rest of the PLC with 100% bootstrap support (Fig. 2.5), and are sister the rest of the PLC excluding the HL clade.

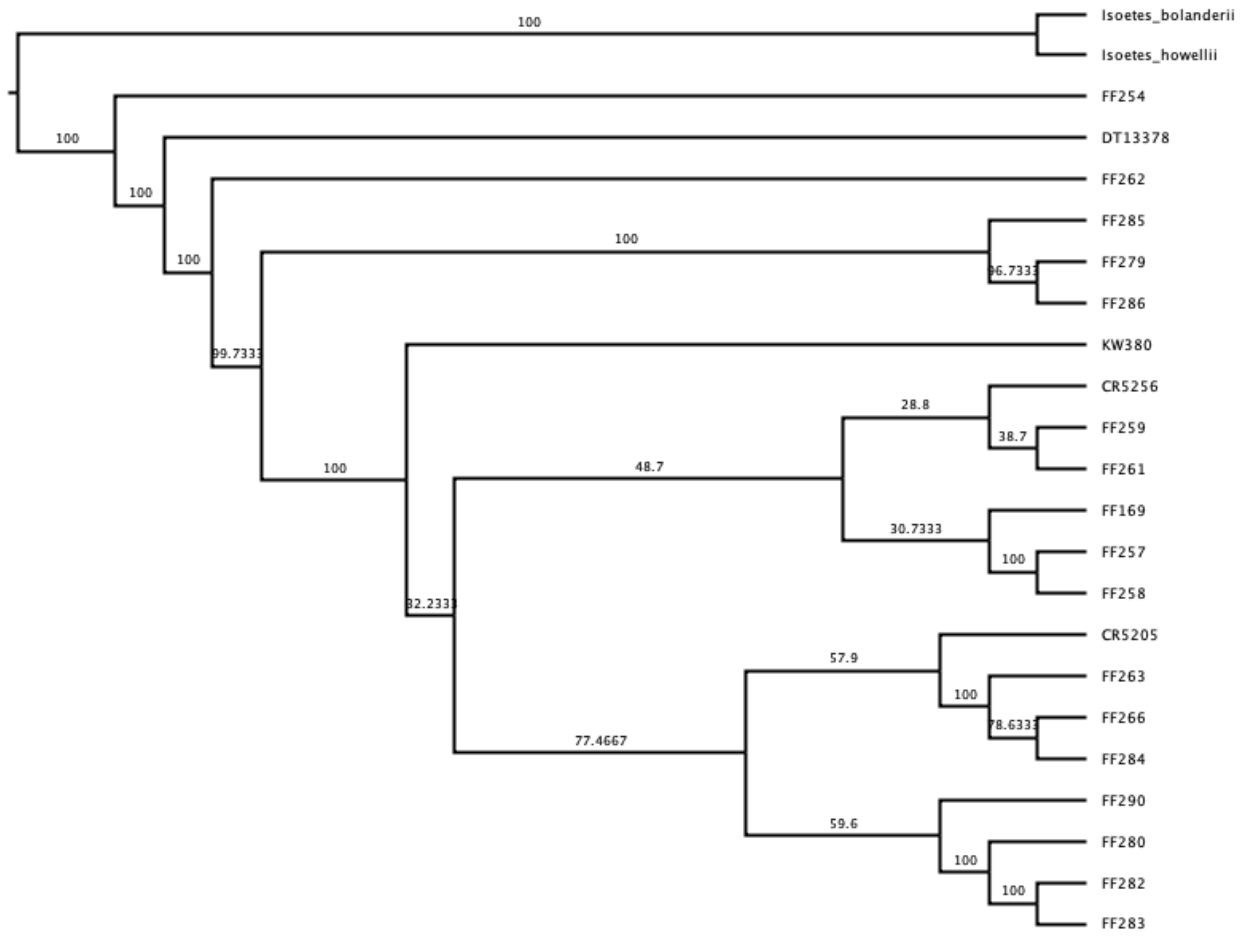

Figure 2.3: SVDQuartets cladogram of the Collection Sites Grouping coalescent tree. Branch labels show clade bootstrap support values.

The Jepson Prairie (JP) clade, so called because the single collection locality for the clade is the Jepson Prairie in Solano county (Fig. 2.5). Collections from this site were initially identified as *I. orcuttii* due to their habitat in the Jepson Prairie vernal pool, as well as the plants' size and spore characters. As with the Hog Lake and Minima clades, not much can be said about these plants due to a lack of sample localities outside the Jepson Prairie site. However, what is important to note is that it is phylogenetically distinct from the Hog Lake and Southern California clades, which also include plants identified as *I. orcuttii* using the existing taxonomic keys.

The Southern California (SC) clade comprises of three collection sites, two from south of the Transverse Ranges and west of the Peninsular Range (FF285 & 286), and one from Sonoma county in the Coast Ranges (FF279; Fig. 2.3, 2.5). Of these three localities, the

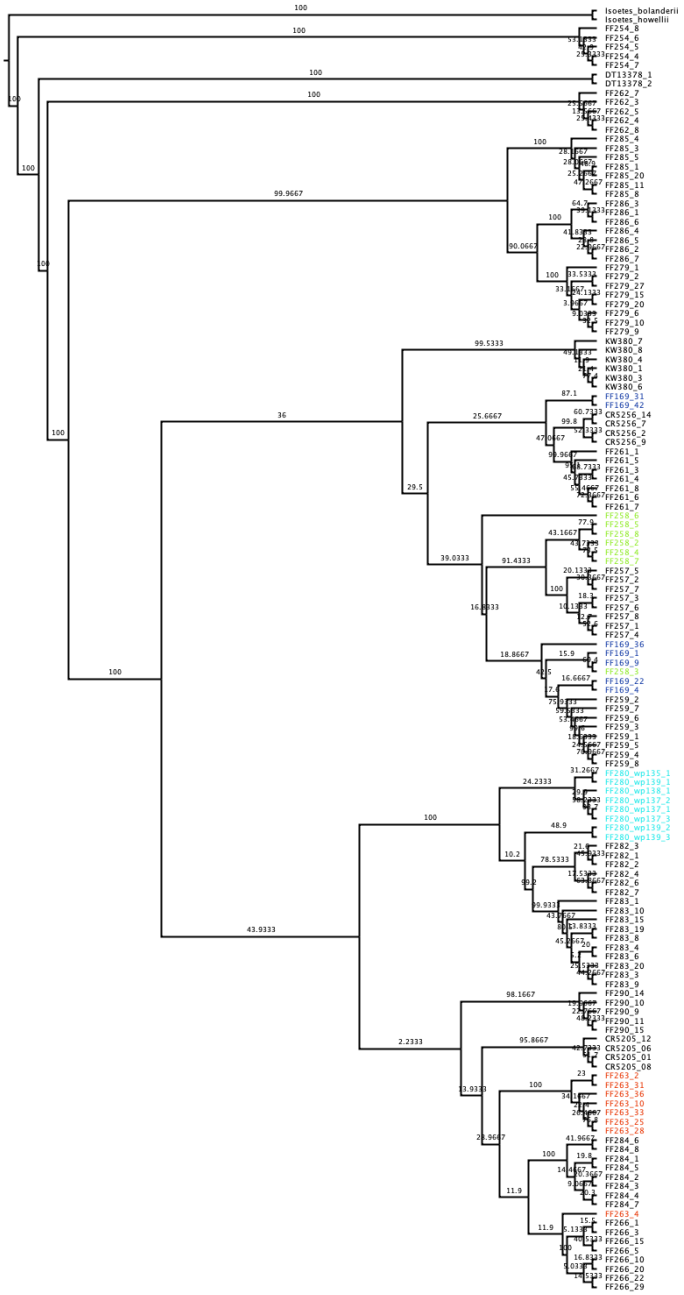

Figure 2.4: SVDQuartets cladogram of the Individual plants grouping coalescent tree. Branch labels show clade bootstrap support values. Samples that do not form monophyletic clades with other individuals collected from the same site are colored as follows: FF169 = Blue, FF258 = Green, FF263 = Red, FF280 = Teal.

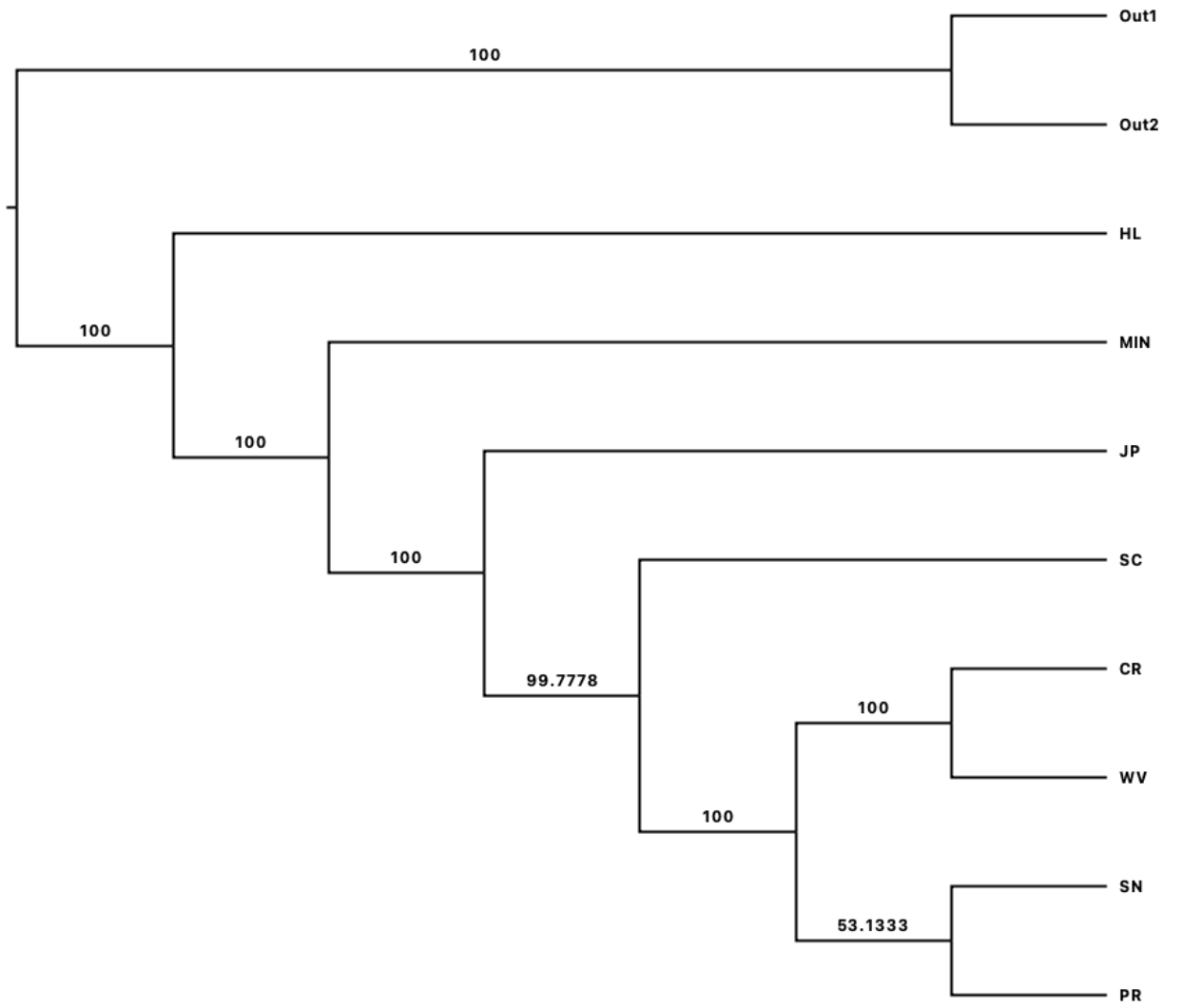

Figure 2.5: SVDQuartets cladogram of the Major Clades Grouping coalescent tree. Branch labels show clade bootstrap support values.

two sites from Southern California (FF 285 & 286) were identified as *I. orcuttii* by spore and leaf size characters. However, one was notable by being collected not from a vernal pool (as would be expected for *I. orcuttii*), but from a seep over a lithic outcrop, and occurred in shallow depressions in the rock surface, which could be vernal pool-like. However, the plants collected from Taylor Mountain Regional Park in Sonoma county (FF279) were initially identified as *I. nuttallii* by spore characters and habitat, which was a seep and wet-meadow region. Yet these plants are nested within the Southern Californian clade, as opposed to the Coast Range or Jepson Prairie clades, with which they shared greater geographic proximity.

The final clade consists of four subclades that are distinctly geographically structured.

The four will be referred to as the Peninsular Range (PR), Sierra Nevada (SN), Coast Range (CR), and Willamette Valley (WV) clades in the following discussion. When SVDQ is run with the Individuals and Major Clades Groupings, the four clades resolve as ((PR, SN),(CR, WV)); (Fig. 2.3, 2.5). However, when the Sites Grouping is used (samples clustered by collection sites), the topology changes to (PR (SN(CR, WV))), though this topology has low bootstrap support (Fig. 2.3).

While these four clades were always recovered, their bootstrap support values varied depending on the analysis. In the Individuals grouping analyses, bootstrap support for the trees would rarely rise above 50% for both tip and internal branches within the (PR, SN, CR, WV) portion of the tree, and would often fall well below 20% support when alternte topologies were proposed from the base All Quartets Examined tree. The support values for the remainder of the phylogeny was uniformly high, with all of the other clades with 100% bootstrap support for their placement and their own monophyly (Figs. 2.3–2.5).

When the Sites grouping is used, the bootstrap support for the backbone was improved, but there was still not good support among the "nuttallii" clades (SN, PR, CR, WV) clade, though the support for the major groupings was better (Fig. 2.3).

When the Major Clades Grouping is used, the bootstrap support for all of the clades goes up substantially, with (CR, WV) reaching 100% boostrap support as a clade, while (PR, SN) achieves ∼53% support (Fig. 2.5).

Through the above analyses, we identified three collection sites that had low placement support values, indicating that they were potentially hybrids: FF290, CR5205, and KW380. Of these, the former two occur in the (Coast Range, Willamette Valley) clade, while the latter migrates between being sister to (SN,CR,WV) or coming out as sister to SN. The placement of these collection sites in the initial, "all quartets" topologies regularly migrated depending on the analysis, though they each showed affinity for the Oregon (FF290) and Coast Range (CR5205) clades, respectively, while occasionally showing up as sister to one another and sister to both groups, or with FF290 as sister to the Coast Range group, with the latter being very poorly supported by the bootstrap analyses. Of the three, FF290 and CR5205 show strong morphological signs of either polyploid and/or hybrid origin: some to most plants from the collections at these sites have irregular or misshapen spores (Chapter 3). Conversely, despite the similarly irregular placement of KW380, the plants from this collection site were not excluded from these secondary analyses. While the plants did show uncertainty in their placement, they showed no evidence of hybrid origin when examined for irregularity in their spore size and overall morphology (i.e. lacking obvious deformity). As such, I am inclined to believe that the low support values for the placement of KW380 may be as a result of low sample numbers in relation to the rest of (PR,SN,CR,WV) clade, rather than hybrid origin, and therefore felt that there was not sufficient justification for removing these collections from the analyses. Given the possible hybrid nature of FF290 and CR5205, I performed a second set of SVDQ analyses with these two sites removed. As a consequence of this removal the resolution, and in some cases topology, of all three groupings (Sites, Individual, Ancestral) changed (Figs 2.6, 2.7, 2.8).

In the Major Clades Grouping topology with the putative hybrids removed both the

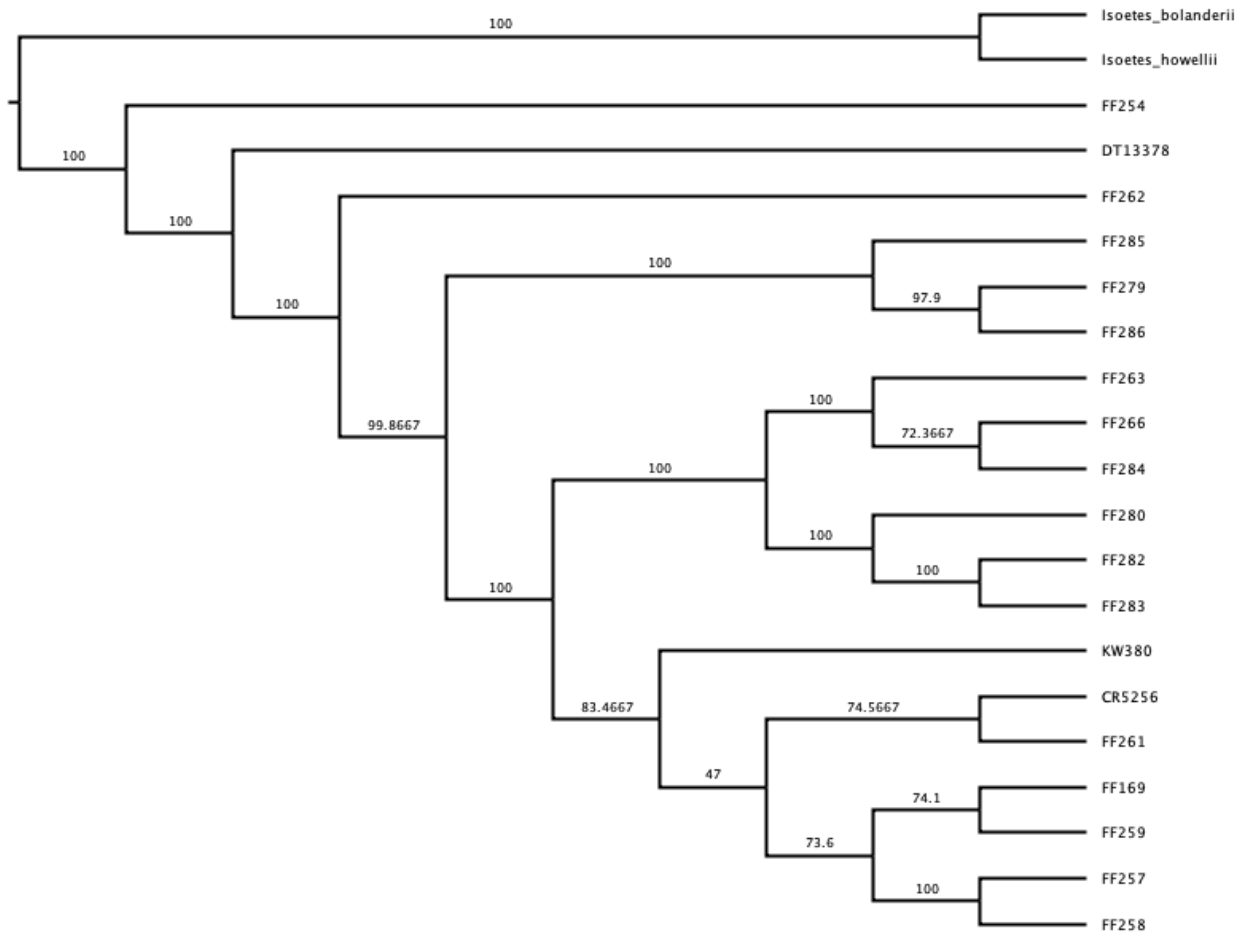

Figure 2.6: SVDQuartets cladogram of the Sites grouping coalescent tree excluding puatative hybrid collection sites. Branch labels show clade bootstrap support values.

Coast Range and Willamette Valley clades resolved with 100% bootstrap support as sister clades. There was also an effect on the resolution of the Peninsular Range site's placement when the putative hybrids were removed: the support for the PR plants resolving as sister to the SN clade rose greatly (69% boostrap support). Similarly, in the Sites grouping analysis, the removal of FF290 and CR5205 drastically improved the support for the Coast Range and Willamette Valley sites as monophyletic clades, and brought the support for the (PR, SN) clade up to ∼84%. As with the other analyses, the Individuals grouping had far lower bootstrap support within the (PR, SN, CR, WV) clade. The SVDQ All Qartets Analyses for the Individuals Grouping also placed the PR clade within the SN clade, though the boostrap support for this placement is very low at only ∼10%.

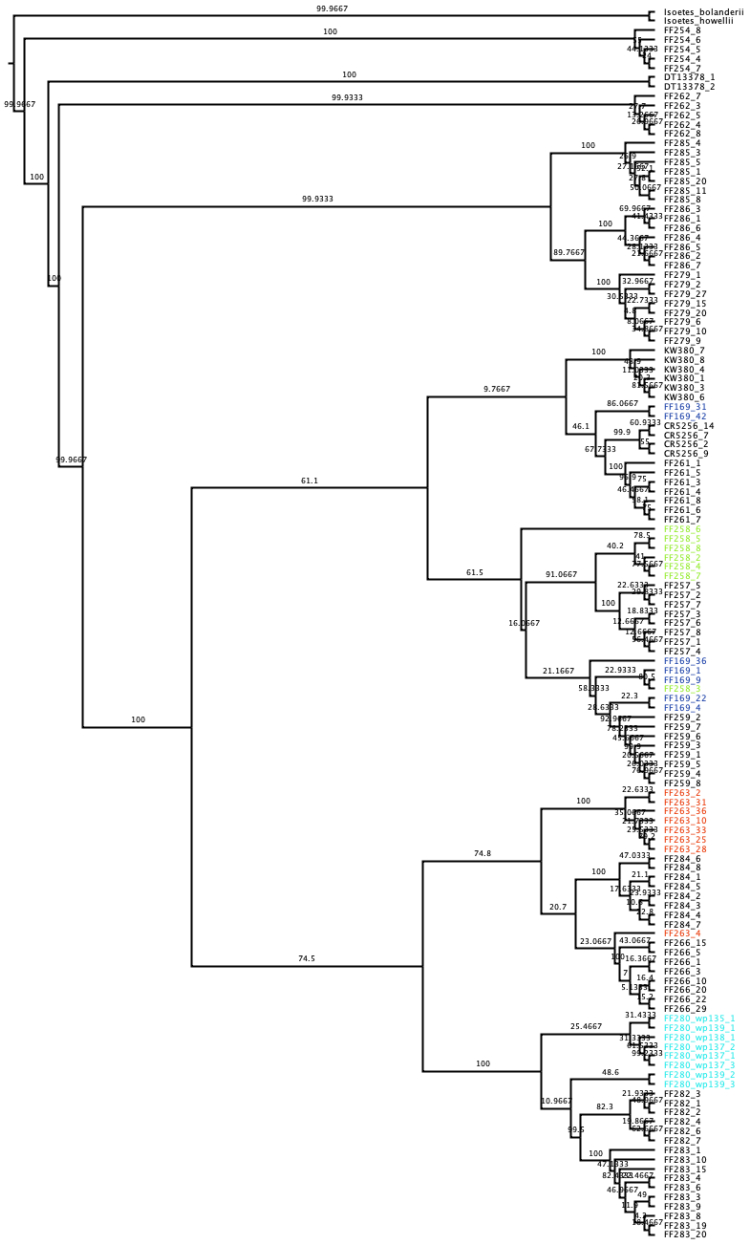

Figure 2.7: SVDQuartets cladogram of the Individual plants grouping coalescent tree excluding puatative hybrid collection sites. Branch labels show clade bootstrap support values. Samples that do not form monophyletic clades with other individuals collected from the same site are colored as follows: FF169 = Blue, FF258 = Green, FF263 = Red, FF280 = Teal.

Inferred relative ages

The inferred branch length of the A00 topologies varied depending on the inclusion or exclusion of the putative hybrid collection sites. From this point on, I will be treating the results

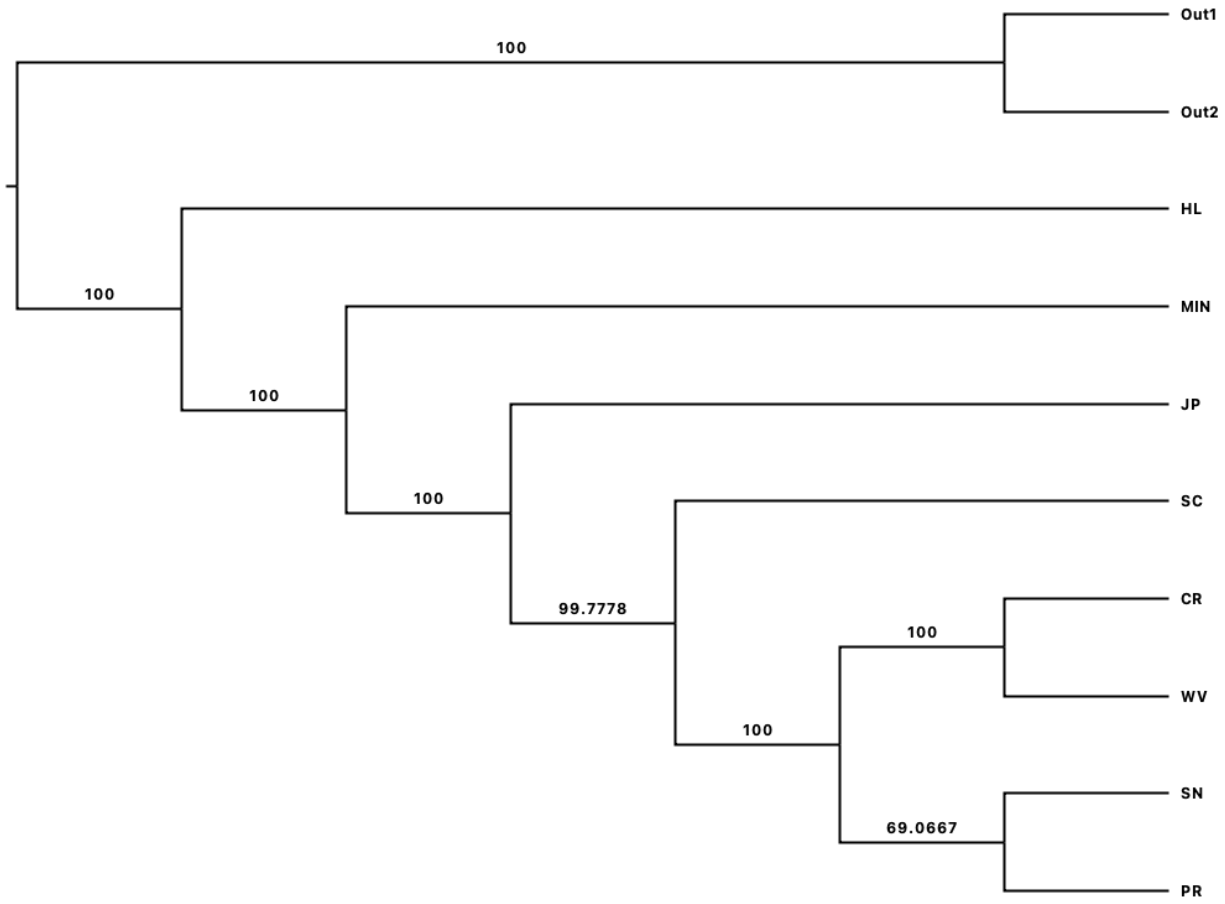

Figure 2.8: SVDQuartets cladogram of the Major Clades Grouping coalescent tree excluding puatative hybrid collection sites. Branch labels show clade bootstrap support values.

of the A00 analysis with the putative hybrid individuals removed as the primary analysis, and will only refer to the alternate analysis with the putative hybrids included when there is a noteworthy departure from the primary analysis. Calculated tau values can be found in Table 2.2.

Table 2.2: Results from the A00 analysis of the PLC, with putative hybrid collection sites removed from the sampling pool. Each sub-table shows the calculated values of theta (population size) and tau (divergence time). # of individuals is a count of the number of phased individuals included in the analysis, with at most 4 individuals sampled from a given collection site. Mean is the mean calculated value for theta or tau for the corresponding clade. Median shows the median calculated value for the clade. S.D. represents the standard deviation of the posterior distribution, as well as the minimum and maximum values for θ and τ. Next are 2.5% and 97.5%, which show the upper and lower limit of the 95% confidence interval on the MCMC distribution. 2.5%HPD and 97.5%HPD show the lower and upper bounds of the Highest Posterior Density (HPD) Interval for each clade's distribution. ESS* stands for Effective Sample Size, while Eff* shows the efficiency of the MCMC for a given clade. In both ESS* and Eff*, the larger the values, the better the mixing and the more reliable the estimates of the parameter values.

θ	# individuals	mean	median	S.D	min	max	2.50%	97.50%	2.5%HPD	97.5%HPD	ESS*	Eff*
HL	4	0.007219	0.007181	0.000674	0.004871	0.010506	0.005967	0.008611	0.005936	0.008565	2328.509867	0.015523
Min	2	0.01438	0.014232	0.001889	0.008325	0.026684	0.011115	0.018505	0.010832	0.018122	15048.0168	0.10032
JP	4	0.017196	0.017107	0.001561	0.011518	0.025155	0.014392	0.020518	0.014221	0.020283	8793.252134	0.058622
SC	12	0.050787	0.050683	0.003017	0.037991	0.066132	0.045168	0.057001	0.044911	0.056696	1936.727665	0.012912
PR	4	0.011857	0.011785	0.001144	0.007843	0.0183	0.009807	0.014296	0.009671	0.014128	3120.115698	0.020801
SN	24	0.112251	0.112094	0.005315	0.09127	0.139	0.10222	0.123062	0.102085	0.122908	1458.713896	0.009725
CR	12	0.067098	0.066914	0.004466	0.051615	0.092357	0.058895	0.076427	0.058639	0.076093	1972.692129	0.013151
WV	12	0.034491	0.034418	0.002036	0.026263	0.045619	0.030706	0.038668	0.030525	0.038458	1801.975519	0.012013
PR,SN	28	0.002776	0.001868	0.003688	0.000331	0.096015	0.0004	0.011001	0.000335	0.007877	189.475559	0.001263
CR,WV	24	0.002051	0.001191	0.003227	0.000389	0.093334	0.000429	0.009191	0.000393	0.006024	163.453466	0.00109
PR,SN,CR,WV	52	0.003768	0.002322	0.004658	0.000339	0.071788	0.000472	0.018439	0.000339	0.014196	49.628838	0.000331
SC,PR,SN,CR,WV	64	0.271462	0.268682	0.029948	0.180455	0.457275	0.220824	0.339214	0.216278	0.332151	234.205441	0.001561
JP,SC,PR,SN,CR,WV	68	0.096467	0.096177	0.009894	0.061077	0.154001	0.077918	0.116749	0.07704	0.115688	537.826974	0.003586
Min,JP,SC,PR,SN,CR,WV	70	0.017635	0.017539	0.001933	0.011458	0.029253	0.014125	0.021688	0.013965	0.02149	587.636133	0.003918
HL,Min,JP,SC,PR,SN,CR,WV	74	0.044614	0.044005	0.006456	0.024723	0.085227	0.033741	0.059021	0.032689	0.057496	1228.301508	0.008189

τ	# individuals	mean	median	S.D	min	max	2.50%	97.50%	2.5%HPD	97.5%HPD	ESS*	Eff*
PR,SN	28	0.002923	0.002921	0.000093	0.002559	0.003299	0.002746	0.003112	0.002743	0.003109	232.064804	0.001547
CR,WV	24	0.002919	0.002917	0.000093	0.002519	0.003301	0.00274	0.003109	0.002742	0.00311	238.528792	0.00159
PR,SN,CR,WV	52	0.002927	0.002925	0.000092	0.002562	0.003305	0.002751	0.003116	0.002752	0.003116	244.202194	0.001628
SC,PR,SN,CR,WV	64	0.002933	0.002931	0.000094	0.002565	0.003312	0.002754	0.003124	0.002751	0.003119	219.502468	0.001463
JP,SC,PR,SN,CR,WV	68	0.005168	0.005163	0.000224	0.004159	0.006089	0.004739	0.00562	0.004725	0.005605	332.085927	0.002214
Min,JP,SC,PR,SN,CR,WV	70	0.00889	0.008885	0.000256	0.007744	0.00999	0.008397	0.009401	0.008397	0.009401	418.398174	0.002789
HL,Min,JP,SC,PR,SN,CR,WV	74	0.015441	0.015436	0.000392	0.014002	0.017067	0.014687	0.016217	0.014675	0.016203	107.684673	0.000718

In the primary analysis, the main branch between the HL clade and the rest of the ingroup took up ∼42.4% of the total branch length of the tree (Fig. 2.9), while including the putative hybrid localities increased the percentage of this branch's length to ∼50.7% of the total tree length (Fig. 2.10). Conversely, the length of the branches leading to the node between the Minima and JP split, as well as the JP and SC splits, shrank when the putative hybrid sites were included, going from ∼24.1% to ∼19.3% for the former, and ∼14.5% to ∼13.7% in the latter. The remaining five sites (SC, SN, PR, CR, WV) effectively formed a polytomy, as the branch lengths of the internal branches between SC and (SN,PR,CR,WV) were effectively zero, though in the case where the putative hybrid localities were included, the branch between the node subtending (SN, PR, CR, WV) was slightly longer (0.1% of the total tree length) than when they were excluded ($8.9\text{x}10^{-2}$%).

In addition, the lengths of the branches leading to the tips in the (SC, SN, PR, CR, WV) portion of the tree were longer in the tree where the putative hybrid sites were excluded compared to the tree where they were included, with the branches taking up ∼16% in the latter, and ∼20% in the former.

Population genetics - Admixture plot

The admixture analysis gave us an optimal K-value of four (Fig. 2.11), which will be referred to as the Admixture Groups from this point on. Most collection sites came back with nearly 100% association with one of these four Admixture Groups. However, there were often either trace associations with other admixture groups, or larger associations in some cases. Because *I. minima* only had two individuals to represent their populations, the admixture plot that included them was not included in the results due to the effects such small sample sizes have on admixture plotting (Lawson et al., 2018).

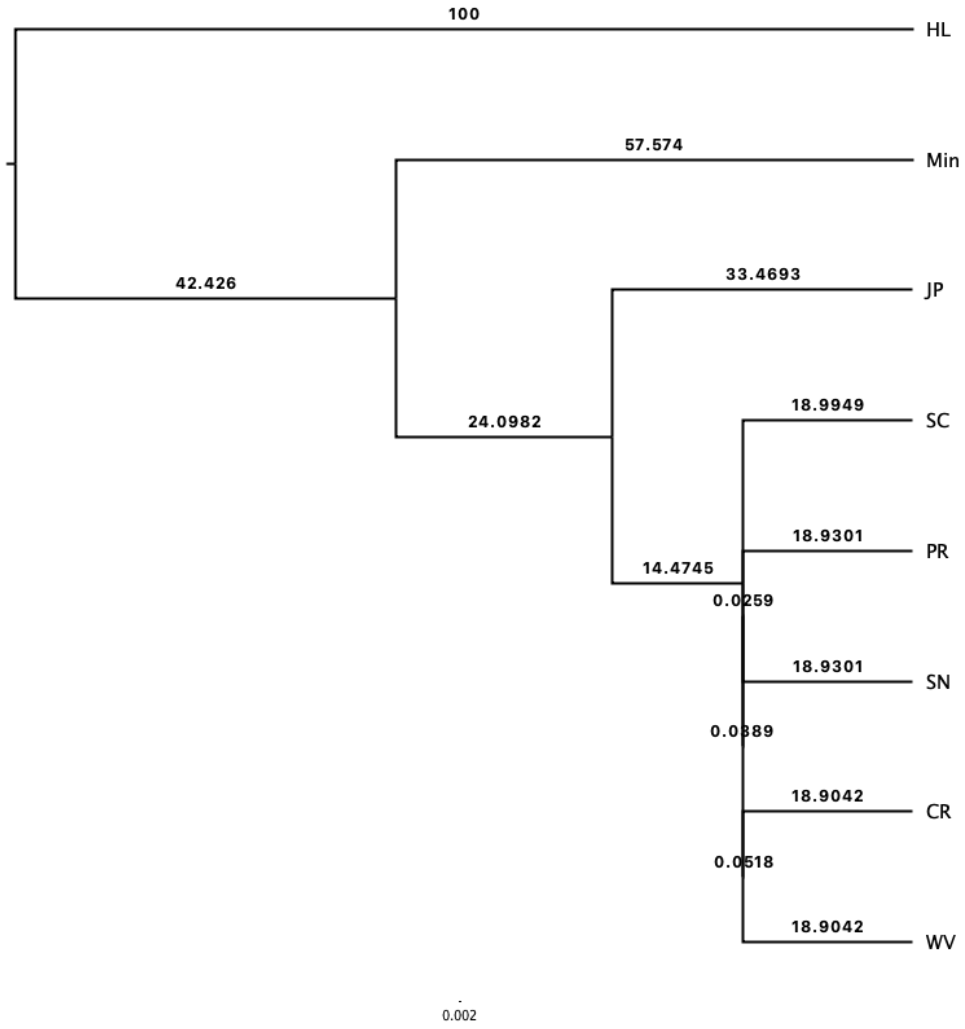

Figure 2.9: A00 analysis done on the PLC dataset excluding the putative hybrid collection sites. Branch lengths were scaled to a total length of 100 to present the percentage of the tree represented by the branch.

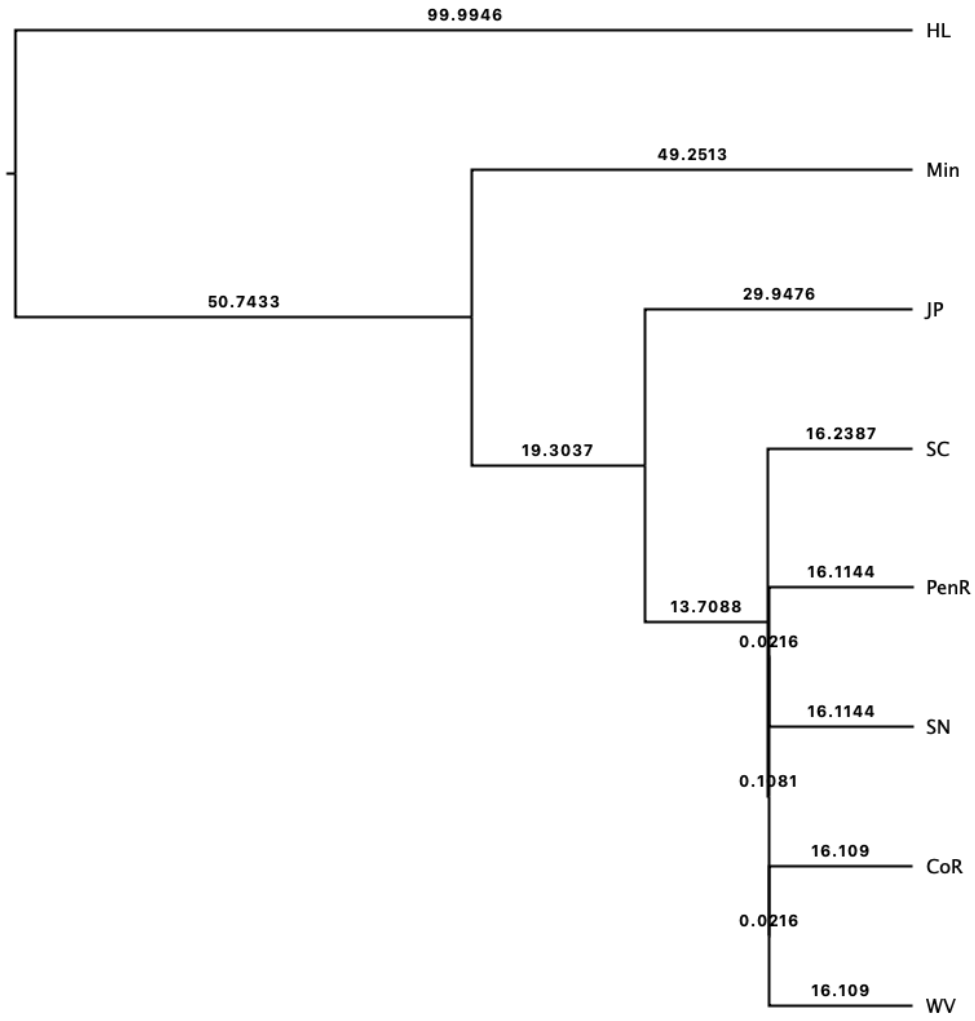

Figure 2.10: A00 analysis done on the PLC dataset including the putative hybrid collection sites. Branch lengths were scaled to a total length of 100 to present the percentage of the tree represented by the branch.

Figure 2.11: Admixture plot for the Pacific Laurasian Clade, excluding *Isoëtes minima* for the optimal K-value of 4. Names adjacent to colored bars indicated which sampled individual the bar represents, while the color of the bar indicates the percentage association each individual has to one of the four populations.

Admixture Group 1 – Hog Lake consists entirely of plants from the Hog Lake Ancestral group (FF254), as well as low levels of presence in the other populations.

Admixture Group 2 – Vernal is associated with plants from the Jepson Prairie (FF262) and the Southern California (FF279, 285, 286) Ancestral Groups. Of note is that FF286 shows a ∼22% affinity for Group 3, while still being predominantly associated with Group 2.

Admixture Group 3 – Montane is by far the largest of the four, and consists of all the plants found in the Sierra Nevada ancestral group (FF169, 257, 258, 259, 261 and CR5256). Of the collections that show mixed ancestry, those predominantly associated with Group 3 are the Peninsular Range site (KW380), with about 25% identity with Group 2 and both CR sites, with CR5205 showing around 50% association between Groups 3 and 4, while and CR5256 shows around 25% identity with Group 2.

Admixture Group 4 – Coastal is associated with the Willamete Valley collections (FF280, 282, 283) and Coast Range group (FF263, 266). Samples that were predominantly associated with Group 4, but one sample (FF284), showed a degree of introgression with another Admixture Group (25% identity with Vernal).

Potential Hog Lake Polyploids

When assembling the sequences for phylogenetic analysis, I discovered two individuals that had unusually high numbers of recovered paralogs within the HypPiper assembly. One of these individuals, FF254_2, recovered 28 paralogs from 220 assembled loci. This is unusually high for my sampling, with most plants having 5 to 9 paralog warnings across their assemblies. When examining and aligning the exonic sequence assemblies for each of the genes in AliView, the sequences didn't appear any different from the general sequence properties of any of the other individuals collected at Hog Lake, which, as noted above in the SVDQ results section, were visually distinct from the rest of the PLC and the outgroup, yet still more like the former than the latter. All of this excess paralogy in the recovered sequences without the sequences themselves looking all that different from the rest of the Hog Lake specimens could indicate that individual FF254_2 is be an autopolyploid (Eriksson et al., 2017; Soltis et al., 2007).

The more unusual individual that was recovered with high paralog numbers was FF254_3, with 139 paralogs warnings for 221 recovered assemblies. However, unlike FF254_3, the genes recovered from this individual could be visually identified as both Pacific Laurasian Clade and American Clade when inspected in AliView. The presence of genes from both clades suggests that this plant was a hybrid (potentially allopolyploid) between the Hog Lake plants and an American Clade species such *I. howellii*, which occurs in the field adjacent to and along the water's edge of the Hog Lake vernal pool.

However, spores sampled from these two individuals did not appear aberrant like those of FF290, and did not have the same size disparity range as CR5205. There were, however, two spore morphologies found in the soil on FF254_3 and FF254_6: the typical Hog Lake morphology, and a dramatically different one that was both larger, less glossy, and had

distinctly tuberculate ornamentation (Fig. 2.12). Why I suspect that these alternate spores come from the surrounding *I. howellii* plants and not the Hog Lake PLC plants is two-fold: the morphology of these spores is like that of *I. howellii* instead of the morphology observed for the rest of the HL plants, and FF254_6 shows no sign of polyploidy in its sequence reconstruction.

Figure 2.12: Examples of the two spore morphoolgies found at the Hog Lake site. The smaller, smooth white spores represent the typical spore morphology for the Hog Lake plants, while the larger, tubercled, beige spore is suspected to be a contaminant from the *I. howellii* plants growing in the adjacent field and along the shore of Hog Lake

2.4 Discussion

For the purposes of discussing the results of this study, we are going to focus on the SVDQ and BPP analyses that omit the two putative hybrid collection sites (FF290 and CR5205), but will address some of the effects their inclusion had on the trees when relevant, and possible explanations, in a separate section below.

Interpreting the short internal branches for the crown clade

Of my results, the short internal / long terminal branch lengths of the (SC,((PR,SN),(CR,WV)) portion of the final A00 analysis (Fig. 2.9) was one of the most unexpected. Coupling these results with the strong topological distinctiveness of each Major Clade within the Sites and Major Clades SVDQ analysis (Fig. 2.6 & 2.8), I speculate that the five Major Clades involved were once part of a larger, panmictic population that experienced a relatively sudden cessation of geneflow between the different elements of their range. Given that I was unable to produce an fully parameterized, time-calibrated A00 tree, the exact age when these breaks occurred can't be determined in anything other than relative time. However, I am willing to propose a possible explanation. Considering the distribution of these plants (Fig. 2.2), the pattern of divergences in the phylogenetic tree (Fig. 2.10), and the long branches leading to the terminals, I suspect that the populations were subdivided due the onset of glaciation in the Sierra Nevada (Birkeland, 1964; Moore and Moring, 2013; Bowerman and Clark, 2011), and Coast Ranges (Holway, 1911; Kuchta and TAN, 2005; Eckert et al., 2008) during one of the region's glacial periods. Periods of intense glaciation have the effect of subdividing formerly widespread taxa and forcing them into refugia, both within the West Coast region (Liu et al., 2011; Kuchta and TAN, 2005; Stone and Wolfe, 2021; Eckert et al., 2008), and outside it (Chung et al., 2017; Hewitt, 2004, 2000; Dumolin-Lapegue et al., 1997; Magri et al., 2006), cutting them off from other populations of their species (Aradhya et al., 2017; Kuchta and TAN, 2005). For plants like *Isoëtes*, which already have a patchy distribution, I suspect that it wouldn't take much to eliminate geneflow between the populations. There's an already well documented paucity in geneflow between vernal pools in different regions (Freeland et al., 2012; Sloop et al., 2012; Elam, 1998; Gordon et al., 2012) and certain alpine meadows species (Matter et al., 2013; Hirao and Kudo, 2008; Kitamoto et al., 2006; Pauli, 1994; Westerbergh and Saura, 1994; Ennos, 1994). A disruption between this already low connectivity between the populations in different regions, particularly repeated disruptions like those found in the glacial - interglacial cycle of the West Coast, could potentially create the sort of pattern I observe in the divergence times.

Phylogenetic and pop-gen results framed by the preliminary identifications.

Taxonomically, the results of our phylogenetic analysis does not support two of the currently described species as natural clades. In the case of *I. orcutii*, plants identified as this species (FF254, FF262, FF285, FF286 and KW 380) are distinctly polyphyletic in the phylogenetic trees (Figs. 2.6, 2.7, 2.8). Not only is *I. orcuttii* polyphyletic, it appears to be spread across four distinct clades within the phylogeny. Of these four, the Hog Lake collection site is the most distinctive, in that it is sister to the rest of the PLC, rather than part of a grade like the rest of the *I. orcuttii* specimens. This distinctness isn't just in it's placement in relation to the rest of the plants identified as *I. orcuttii*, but in terms of it's placement in the admixture plot, where it forms its own distinct Admixture Group. Even when looking at the aligned

sequences, the Hog Lake individuals stand out as easily identifiable from the rest of the PLC.

The rest of the Major Clades primarily identified as *I. orcuttii* based on their morphological characters are the Jepson Prairie, Southern California, and Penninsular Range.

Of these three, the Penninsular Range plants were the most unlike the rest of the plants identified as *I. orcuttii* in terms of habitat preference, as they were found growing in mossy depressions on top of a granite outcrop, rather than a vernal pool. However, the mossy bank where they were collected was in many ways vernal pool like, in that it was within a depression on the rock's surface where water accumulated before flowing out of the depression, rather than along the edge of the outcrop as with most of the other observed mossy banks. The majority of their observable morphology, namely their size and spore characters when I ran them through the taxonomic keys led to me classifying the collections as *I. orcuttii* instead of *I. nuttallii* (See Chapter 3 for details). However, there were certain qualities that I cannot properly articulate that made me question whether or not they were *I. orcuttii*. This uncertainty appears to have been merited, as the PR site was not placed definitively as part of the rest of the plants identified as *I. orcuttii*. Instead, its placement was most strongly associated with *I. nuttallii*, and specifically as sister to the Sierra Nevada clade. This placement had only 83% support, with the branch leading to the SN clade having only a 47% support in the Sites grouping tree (Fig 2.6. This uncertainty is also reflected in the Major Clades grouping, which had ∼70% bootstrap support for the placement of the PR clade as sister to the SN clade. The Individual Grouping tree placed it as nested within the SN clade, though this placement had less than single digit support (Fig. 2.7) for the sister relationship between the PR plants and the sub-portion of the SN clade SVDQ placed them as sister to. This uncertainty of placement, yet strong affinity for the SN clade is further corroborated when examining the admixture plot. In the admixture plot, the plants from the PR collection site (KW380) show ca. 75% affinity with Admixture Group 3 - Montane, which is associated with the plants collected in the Sierra Nevada, and that make up the SN Clade. Yet, they also have a 25% association with Admixture Group 2 - Venral, which includes both the SC and JP clades. I do not believe this is a case of introgression, as the plants do not show spore morphology consistent with such a phenomenon (see Chapter 3).

Of the two remaining clades, the JP plants most closely resembled the description of *I. orcuttii*. What was unexpected was it's placement as sister to the clade that contained the rest of *I. orcuttii* and all collection sites identified as *I. nuttallii*. However, much like the PR and HL clade plants, the JP clade does not form a monophylletic unit with the rest of the plants identified as *I. orcuttii*. This site does come out as strongly part of Admixture Group 2 in the admixture plot, which would indicate that it is at least associated with the SC clade, which includes the bulk of the collection localities identified as *I. orcuttii*. I suspect that any other collections from the central valley would very likely come out as part of this particular clade, especially considering that plants identified as *I. orcuttii* are generally absent from the southern part of the valley proper (Fig. 2.2).

Finally, the SC clade was unexpectedly non-monophyletic in terms of my initial identifications, in that it includes plants identified as both *I. orcuttii* (FF285 & FF286) and *I. nuttallii* (FF279). What is important to note is that the type locality of *I. orcuttii* is in San

Diego, and potentially from within proximity to one of the two localities sampled. These plants identified as *I. nuttallii* were found in Sonoma County, where as the other two collection sites were in southern California in Riverside and San Diego counties. The placement of the FF279 in the SC Clade was a surprise, as there were other *Isoëtes* collections (FF266) made within 6km of the site, and on an adjacent hill chain. When originally collected, I expected FF279 and FF266 to be each-other's closest relatives based on geographic proximity, but this was not the case in both the SVDQ and Admixture Plot. Instead, they came out as part of distinctly different clades, and showed minimal amounts of admixture between their admixture groups (Fig. 2.11) Given the geographic distribution of the three localities, and the presence of other plants identified as *I. orcutii* to the west of the Southern Coast Ranges, I suspect that *I. orcuttii* type plants west of the Southern Coast Range mountains may be part of the SC clade, which if I'm right about sampling in proximity to or at the type locality, would indicate that this unit represents *I. orcuttii* s.s.

Isoëtes nuttallii (CR5256, FF169, FF257, FF258, FF259, FF261, FF263, FF266, FF279, FF280, FF282, FF283, FF284) is also non-monophyletic, though not to the same degree as *I. orcuttii*, in that they formed a mostly coherent clade in the crown of the phylogeny (Fig. 2.8).

Of the three Major Clades that were recovered in the Site and Individual Grouping phylogenies, WV is the most likely to be *I. nuttallii* s.s., as it comes from the region in closest proximity to the type locality for the species. If WV is treated as *I. nuttallii* s.s., if I wanted to be conservative about my interpretation of the phylogeny, the argument could be made to consider both WV and CR as *I. nuttallii*, as they form a well supported (CR,WV) clade. The admixture plot also appears to support this, as both the WV and CR collection sites belong to Admixture Group 4. They are, however, not intergraded with one another, but instead form distinct clades with 100% bootstrap support for the splits (Fig. 2.6). This, coupled with the long terminal branches in the A00 analysis, including those for the CR and WV clades, makes me disinclined towards treating both Major Clades as part of a continuous metapopulation. Instead, I believe that they should be treated as separate units, and therefore separate species, or possibly subspecies.

The plants in the Sierra Nevada Major Clade, while identified as *I. nuttallii* initially, are both phylogenetically distinct from the Willamette Valley / Coast Range plants, and form a unique Admixture Group in the admixture plot. While an argument could be made for these plants to also be a sub-species or variety of *I. nuttallii*, there are two pieces of evidence that make me believe that it is not part of the same species as CR and/or WV. The first and more direct is the admixture plot for both of the putative hybrid collection sites. When looking at the admixture plots for both FF290 and CR5205, they are shown to have identity with both Admixture Group 3 and 4 (Fig. 2.11), with a small amount of Group 2 in FF290. Were this simply a case of sub-populations introgressing, it could be seen as evidence for the unity of SN, CR and WV as a single unit of *I. nuttallii*. However, the fact that both of these sites show the kind of morphology associated with interspecies hybrids serves as evidence to the contrary. If the admixture plot does indicate that there was introgression between the two populations, and that introgression resulted in a hybrid/polyploid, I would

consider that evidence to sufficient incompatibility between the two lineages that they are no longer the same species. The second is that the most well supported phylogeny doesn't place SN as directly sister to CR and WV, but rather, is sister to PR, with the resulting (SN,PR) clade sister to (CR,WV). As addressed above, the PR clade plants are noticeably morphologically distinct from all three of the other clades in question, something that led me to initially identifying it as a different species entirely. Given all of the above, I feel confident in concluding that the SN clade is not part of the same species as CR and/or WV.

Isoëtes minima appears to be the only member of the PLC that has phylogenetic support for the taxonomy. However, as I only have two samples from a single locality, I am not as confident of this assessment as I am with the other Major Clades. As with HL and JP, I believe more samples need to be added to the dataset from other localities in order to fully and properly assess the strength of the taxon as a biological lineage.

Potential Hog Lake Polyploids

The putative polyploid members of the Hog Lake site represent a potentially major discovery not only within *Isoëtes*, but in plant systematics as a whole.

Isoetes are well known for forming polyploids, with an estimated 49 species of 105 with known chromosome numbers being triploid or higher (Hoot et al., 2004, 2006; Troia et al., 2016). However, if verified as being a genuine polyploid/hybrid, the FF254_3 would be the first documented case of a hybrid forming between the major *Isoëtes* clades. Furthermore, it would mean that two species who's most recent common ancestor was estimated to have diverged ca. 103 Million Years Ago, as estimated by Larsén and Rydin (2016) successfully formed a hybrid. This would put the estimated age of divergence between these two clades in the late Albian period of the Cretacious. Other major evolutionary events in plants that were occuring at this time include the split between Gnetales Blume and Welwitschiales Skottsb. ex Reveal (Lu et al., 2014), as well as the diversification of the following major groups of plants: Asteridae Takht., Rosidae Takht. (Wikström et al., 2001; Bell et al., 2010), and the Eupolypod 1 ferns (Rothfels et al., 2015; Schneider et al., 2004a).

If FF254_3 do indeed represent hybrids between the AC and PLC, it would also represent a introgression event that exceeds the age of *xCystocarpium roskamianum* Fraser-Jenk., which currently stands as the oldest known hybridization event among plants, and possibly eukaryotes (Rothfels et al., 2015). Furthermore, the spores of these two individuals do not show the same type of deformity that other *Isoëtes* show when introgressing with other, more closely related species (Musselman et al., 1996, 1997). This lack of obvious deformity of the spores that is so often associated with these sort of hybrid events raises some concerns about the results.

One possibility for this is on-site cross contamination, as there were American clade plants growing in the adjacent pasture by the Hog Lake site, and there was a morphologically unusual spore was found among the other spores in the soil at the base of one of the individuals in question (FF254_3). In addition, no American Clade plants were present in the extractions for that particular day, and it was only one individual within the whole plate

to show this sort of mixed ploidy. Further examination of the Hog Lake site and plants is needed to determine if there is indeed introgression between the PLC and AC occurring in the area, or if the odd sequences were due to some form of contamination either in the pre- or post-extraction stage.

2.5 Conclusions

From the results of both the phylogenetic analysis and population genetics in combination with the original species designations based on the taxonomic keys from both the Flora of North America and Jepson Manual, it is clear that the current species designations for the PLC do not reflect monophyletic groups, and therefore do not fit the current accepted concept of a species as they are circumscribed. As such, the status of the species as currently circumscribed should be evaluated in a morphological sense to determine if there are any traits (morphological, ecological, etc) that can be used to distinguish between the clades identified by the phylogeny.

Furthermore, the combination of the SVDQ and A00 analysis, particularly the tau values for each branch tip generated in the latter, indicates that the SC, SN, PR, CR and WV clades were all at one point a single, panmictic population, only for this larger unit to later become suddenly subdivided and begin to evolve along it's own evolutionary trajectory.

Chapter 3

Taxonomy and Systematics of the Pacific Laurasian Clade of *Isoëtes*

3.1 Introduction

One of the core goals of biology is to identify, describe, and study the diversity of life on Earth. In its earliest days, most European and Middle Eastern biology was framed in the context of examining life in order to understand the divine plan or cosmic order. However, this paradigm was eroded as the explanation for life's existence and diversity shifted away from a divine source and to a material one. This shift was ultimately synthesized in Darwin and Wallace's seminal works on the theory of evolution by natural selection (Darwin et al., 1858; Darwin, 1859; Darwin et al., 1958; Wallace, 1876, 2016). These works gave us a purely material, observable, and testable framework through which the lens of inquiry shifted from simply observing and cataloging to understanding the nature and evolution of life in a purely physical sense. Yet, the impetus of biologists to catalog and understand the myriad forms life takes has remained a cornerstone of the science, regardless of the how that biological diversity came to exist in the form that it does today. The focus has merely shifted from just cataloging to also studying the evolutionary process, while said descriptions have moved from simply things that share superficial phenotypic similarity (i.e. all things with yellow flowers are put into one catagory, or all ground dwelling birds) to one based on inferred evolutionary common descent. With the development of molecular phylogenetic methods, greater sampling of both contemporary and prehistoric organisms, and increasingly powerful inferential tools, great strides have been made in this effort to understand the evolution and diversity of life.

Molecular systematics have continued to refine our earlier concepts of organismal diversity, as well as identified areas where previous proposed relationships proved to be untrue. This has led to the revision of many long-used classifications, such as Scrophulariaceae (originally classified based on shared ancestral, or symplesiomorphic, floral morphology; Garnock-Jones, 1993; Olmstead et al., 2001; Oxelman et al., 2005), Polypodiaceae and

Grammitidaceae (the latter being nested within the former) (Schneider et al., 2004b), and the various genera within Pteridaceae and Phrymaceae that have proven to be para- and polyphyletic, such as *Mimulus* or *Notholaena* (Fraga, 2012; Barker et al., 2012; Rothfels et al., 2008; Yatskievych et al., 2008). Higher level taxonomy has also been drastically revised in many ways, introducing new kingdoms to properly classify non-animal life into monophyletic clades, including a whole new rank, the Domain, to catagorize the three highest levels of organismal diversity (Archaea, Prokaryotes and Eukaryotes). And, in addition to identifying where taxonomic revisions are needed, molecular systematics has also led to recognition of another phenomenon: cryptic lineages.

Cryptic lineages are those that are difficult to recognize, either visually or systematically, though the nature of this crypsis can vary greatly. In some cases, the lineage is considered cryptic because it is difficult to find in its natural environment, either due to small size, habitat (living in inaccessible areas, underground, or as a symbiont in another organism), or because they're camouflaged. Other cryptic lineages are not necessarily difficult to find in their environment, but are difficult to tell apart from other lineages due to an absence of diagnostic morphological features (potentially as a result of convergent evolution or conserved, shared ancestral morphology; Yoder et al., 2000; Moura et al., 2011; Sheets et al., 2018; Ladner and Palumbi, 2012; Geml et al., 2006; Carlsen et al., 2011; Barker et al., 2012) or morphological differences being unrecognized due to an assumption of morphological differences between sites simply being local variation rather than something taxonomically significant (Dauphin et al., 2014; Stoughton et al., 2017, 2018; Carlsen et al., 2011). In all of the above cases, crypsis leads to errors in our taxonomy, and therefore our understanding of the evolutionary history of these organisms.

Among vascular plants, the genus *Isoëtes* L. is one that is both visually and taxonomically cryptic. These plants are small, perennial, semi-woody geophytes that occupy mesic to aquatic habitats on every continent other than Antarctica, as well as several oceanic islands (Hoot et al., 2006; Pfeiffer, 1922; Larsén and Rydin, 2016; Troia et al., 2019). Macromorphologically, the above-ground portion of these plants are inconspicuous and "unremarkable" in appearance, consisting of one, or rarely more, crowns of linear leaves with distinct spiraled phyllotaxis. Aside from their phyllotaxis, they have few obvious traits that are visible without excavating the plants from their substrate. As such, they are often and easily mistaken for grasses, rushes, or sedges (Freund, pers obs of herbarium and iNaturalist.org records). However, *Isoëtes* are not seed plants. Instead, they're one of the three extant types of lycophyte, the sister group to the euphyllophytes, which consists of the spermatophytes (extant seed plants) and monilophytes (eusporangiate and leptosporangiate ferns, horsetails, whisk-ferns). At our best estimate, euphyllophytes and lycophytes diverged from back in the Devonian, ca. 400mya Larsén and Rydin (2016); Taylor et al. (2009); Gensel and Berry (2001).

Historically, species delimitation in *Isoëtes* has relied on the few morphological or ecological traits that are present and variable within the genus: megaspore size, ornamentation, and color; presence of leaf scales/phyllopodia; root branching; habitat preference; velum (a membrane that covers the sporangium on the adaxial face of their leaves) coverage; leaf

size and stiffness; stomatal presence; or more rarely things like microspore ornamentation and ligule shape (Pfeiffer, 1922; Engelmann, 1882; Taylor et al., 2009). These traits can be highly plastic within a given species or convergent between taxa, such as spore ornamentation, habitat preference, or leaf length. In other cases, characters used to identify species can be misinterpreted, such as incorrectly identifying spore ornamentation (Freund pers. obs., Fig. 3.1).

Thanks to the previously discussed complications, molecular taxonomy of *Isoëtes* might seem like it would be more reliable than morphology. While molecular systematics have helped infer much about the macro-level evolution of the genus, there are still hurdles that must be overcome when working at the species level. Species identification can be difficult in *Isoëtes*, leading to collections being misidentified when uploaded to online database like genbank or vouchered for herbarium collections (Freund, pers. obs.). *Isoëtes* are also notorious for a propensity to form allopolyploid lineages through hybridization, with ca. 50% of the known species being known polyploids of tetraploid (4N) or higher (Hoot et al., 2004, 2006; Larsén and Rydin, 2016; Troia et al., 2016). Until recently, *Isoëtes* molecular systematics was dependent on a small number of chloroplast regions, the internal transcribed spacer (ITS), and the LFY-homolog (Larsén and Rydin, 2016; Rydin and Wikström, 2002; Hoot et al., 2004, 2006; Hoot and Taylor, 2001). Chloroplast genes, while abundant and easily sequenced due to "universal primers" are more slowly evolving than most nuclear genes, and are also strictly materially inherited. As such, while one of the strengths of chloroplast genes is that they're good for identifying deeply diverging lineages, they can lack the signal needed when working with recently diversified ones. They are also ineffective when for inferring the parentage of polyploids, since chloroplasts only reflect one maternal history. This means that only the direct maternal ancestry can be recovered in the polyploid, which can be problematic when trying to understand the entirety of the polyploid's evolutionary history, particularly when the lineage is the result of multiple rounds of introgressive polyploidization (Rydin and Wikström, 2002). Similarly, ITS, while easily amplified and having readily available primers, have similar problems. For example, although the intergenic spacers (18s & 28s) used in most ITS analysis are sufficiently variable for younger ages than most chloroplast genes, they are still fairly slowly evolving, meaning very recent radiations cannot be resolved with confidence (Larsén and Rydin, 2016; Hoot et al., 2006, 2004). There is also an unusual phenomenon in ITS where the gene is rapidly homogenized between different versions through concerted evolution to only have a single version within the nuclear genome. Like with the chloroplast, this results in all but one version of ITS inherited in a hybridization event rapidly being lost (Feliner and Rosselló, 2007).

Low- or single-copy nuclear genes, like the LFY-homolog, have the potential to solve some of the problems. However, they also have their own share of complication when it comes to their use. The first is logistical. Unlike ITS or chloroplast genes, which have "universal" primers that can generate sequence for a wide breadth of organisms across the tree of life, the process to develop sequencing primers for low- or single-copy nuclear genes is both expensive and time consuming in non-model organisms. Even after identifying low- or single-copy genes, there is the potential concern that an inference based on a single gene tree may not be

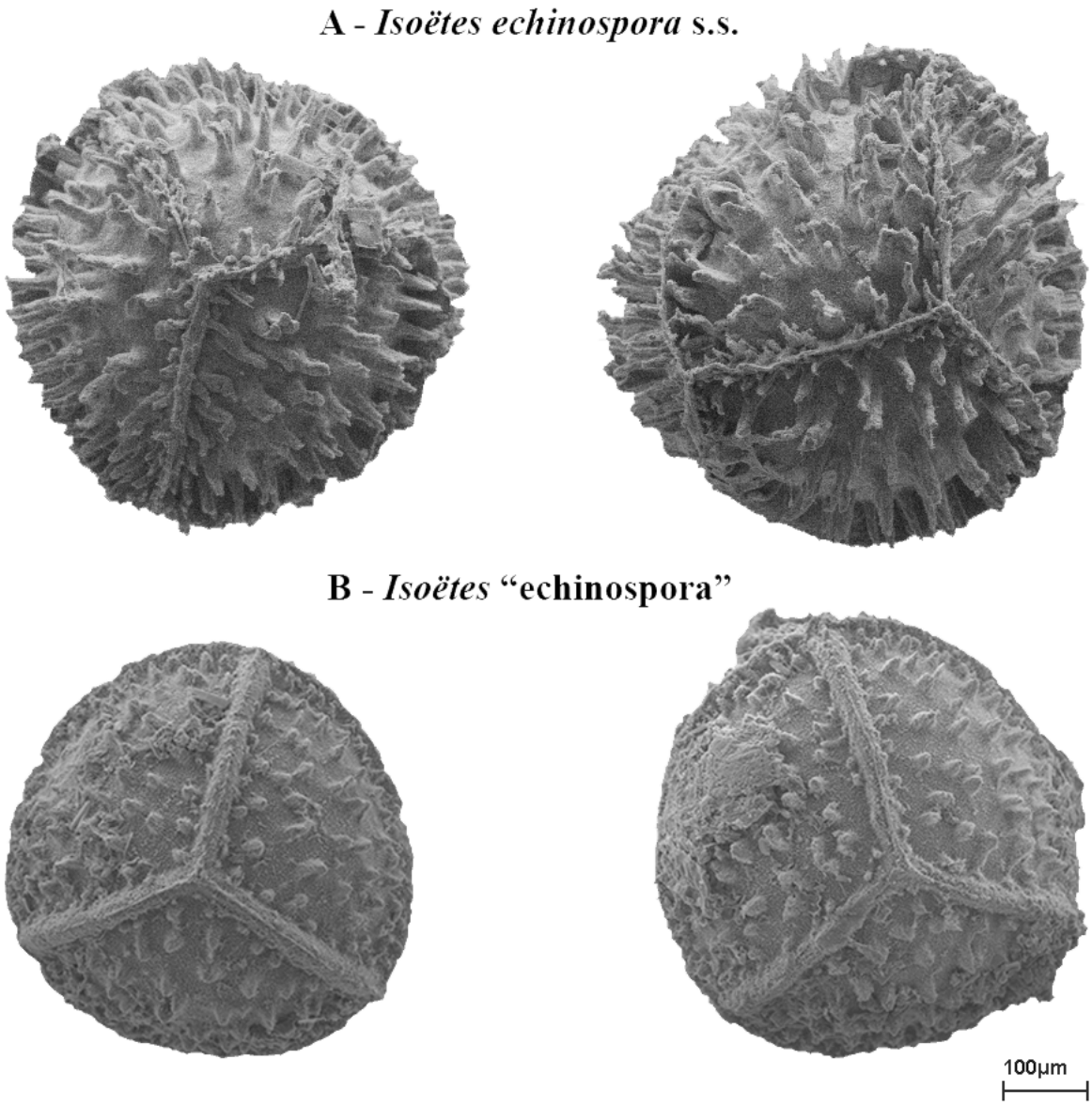

Figure 3.1: Comparison between the spores of *Isoëtes echinospora* S.S. with *I.* "echinospora" from Snag Lake in the Sierra Nevada mountains of California.

representative of the entirety of the lineage's phylogenetic history (i.e., the species tree). This is not unique to low- or single-copy genes, and can in fact be just as problematic, if not more so, in plastid and ITS phylogenies. This phenomenon can arise from two sources: incomplete lineage sorting (ILS), or introgression. In the former, the process of coalescence, or fixation of genes unique to one lineage, hasn't had enough time to occur. This can lead to one or more gene-trees that conflict with the encapsulating species-tree (Degnan and Rosenberg, 2006; Liu et al., 2009; Zhang et al., 2013; O'Meara, 2010). In single-gene phylogenies, ILS can lead to errors in our interpretation of the evolutionary history of the organisms being examined, since the gene-trees do not represent the actual pattern of divergence between metapopulations in time. Similarly, gene capture via horizontal gene transfer or introgression from hybridization can lead to errors in the phylogeny for similar reasons, and can result in lineages that are not each other's closest relatives being clustered together (Xi et al., 2012; Davis and Xi, 2015; Xi et al., 2013; Wickell and Li, 2020).

The obvious solution to the problem of gene-tree / species-tree incongruence is to use multiple low-copy nuclear loci in the analysis. While a conceptually simple solution, it is far more difficult in practice. To begin, generating sequences for multiple low- or single-copy nuclear genes compounds the time and expense in generating the primers/probes compared to a single gene. The next issue is that of the actual analysis. The least difficult method is to concatenate all of the genes together and perform a single phylogenetic analysis on the concatenation. The hope in this case is that the total phylogenetic signal will overcome the conflicting signal from ILS / introgression. However, this isn't necessarily advisable: concatenation assumes that each of the genes in the alignment is evolving at the same rate unless it is properly partitioned or uses a +G model, or as a single unit, which isn't necessarily true outside the chloroplast or genes known to be in linkage disequilibrium. Furthermore, in the process of inferring a single gene tree from the concatenation, any information on the underlying gene-trees, including which genes are under ILS / introgression, is lost. Even if the data is partitioned to allow for different models of substitution to be used, the inferential tools will treat the entire concatenated sequence as a single evolutionary unit. As a result, if there is a significant amount of conflicting signal, the models may not be able to resolve the correct relationships within the tree, or have such conflicting signal that the inference returns a polytomy or support for the wrong relationships among the sampled organisms.

The use of multiple low-copy multiple nuclear genes is still our best option when trying to infer lineage history, but to do so, we need methods and models specifically designed to handle the kind of data we are now capable of generating. One such method is the multi-species coalescent (MSC). The idea of the multi-species coalescent is to use gene trees from multiple loci, then infer which species tree is most likely to have produced the provided gene-trees. While it is conceptually simple, the implementation is far more complex. Most non-Bayesian MSC methods, and some of the more prominent ones, such as ASTRAL (Mirarab et al., 2014) treat the underlying trees as fixed rather than as something with uncertainty, meaning errors in the inference of the underlying trees will propagate downstream (Rannala and Yang, 2017). These limitations can be mitigated, however, if one can jointly estimate both the gene and species-trees. However, such methods are computationally expensive and can

have difficulty reaching convergence, especially in datasets with a large number of samples or ¿100 loci (Rannala and Yang, 2003). In particular, the primary constraints placed on the species-tree by the underlying gene-trees mean that the timing of any split in the gene-trees must predate the splits in the species-tree. This constraint means that proposals to the species-tree topology may be rarely accepted and can result in poor mixing of the MCMC. While still computationally expensive, such a model allows for better mixing of the MCMC and therefore more robust estimation of the coalescent and species tree Rannala and Yang (2003, 2017).

In the last decade, there has been increased interest in free-sporing plants of all kinds, which has led to a flush of newly available genome wide low-copy nuclear data for not only monilophytes and bryophytes (mosses, liverworts and hornworts), but lycophytes, including *Isoëtes* (GoFlag project and 1KP project) (Leebens-Mack et al., 2019; Carpenter et al., 2019; Breinholt et al., 2021). From this newly acquired genome wide low-copy nuclear gene resources, we have a wide berth of data available to implement the multi-species coalescent that couldn't previously be implemented because of the limited genomic data available, including methods specifically built to account for incomplete lineage sorting (Chifman and Kubatko, 2014; Rannala and Yang, 2003, 2017). These new methods and resources can greatly improve our ability to not only resolve the relationships within *Isoëtes*, but also identify cryptic lineages.

Here, I use the results of my multi-species coalescent based phylogenetic research and population structuring of the Pacific-Laurasian Clade (PLC) of *Isoëtes* (Chapter 2) to address the taxonomy of the PLC. Unlike most other *Isoëtes* found in the Americas, which belong to the larger, more species-rich American Clade, the PLC comprises only three described species (*I. nuttallii* A. Braun ex Engelm., *I. orcuttii* A.A. Eaton, and *I. minima* A.A. Eaton) that are endemic to the west coast of North America and are nested deep within the Laurasian Clade whichis otherwise found in Europe, North Africa, and India (Larsén and Rydin, 2016). This disjunction between the PLC and it's nearest sister lineage has been inferred to have occurred around 38 MYA in the Eocene, though there is a high degree of uncertainty in this estimate (Larsén and Rydin, 2016). Plants in the PLC can be easily distinguished from their AC cousins by the presence of a consistently trilobate corm as opposed to a bilobate one (Freund et al., 2018).

Of the three described species, the most widespread is *I. nuttallii*, which is found from southern British Columbia to Northern Baja, west of the Cascades with a small number of localities to the east, as well as the Sierra Nevada and Peninsular Range of California. It inhabits a variety of wet environments, including meadows, seeps, and moss-banks on boulder outcrops with water flowing over their surface (Fig. 3.2, 3.3). In addition to it's habitat preferences, *I. nuttallii* is distinguished from other members of the PLC by their leaf size (8–17cm), as well as megaspore ornamentation (tuberculate), size (400-528mm diameter) and color (white to light grey when dry or wet) (Table 3.1.

Isoëtes minima, often considered the rarest of the three, is found in similar habitats to *I. nuttallii* in the montane areas in and around the Columbia Plateau and eastern edge of the Great Basin Desert (Fig. 3.2). Currently, these plants are only known from eleven localities

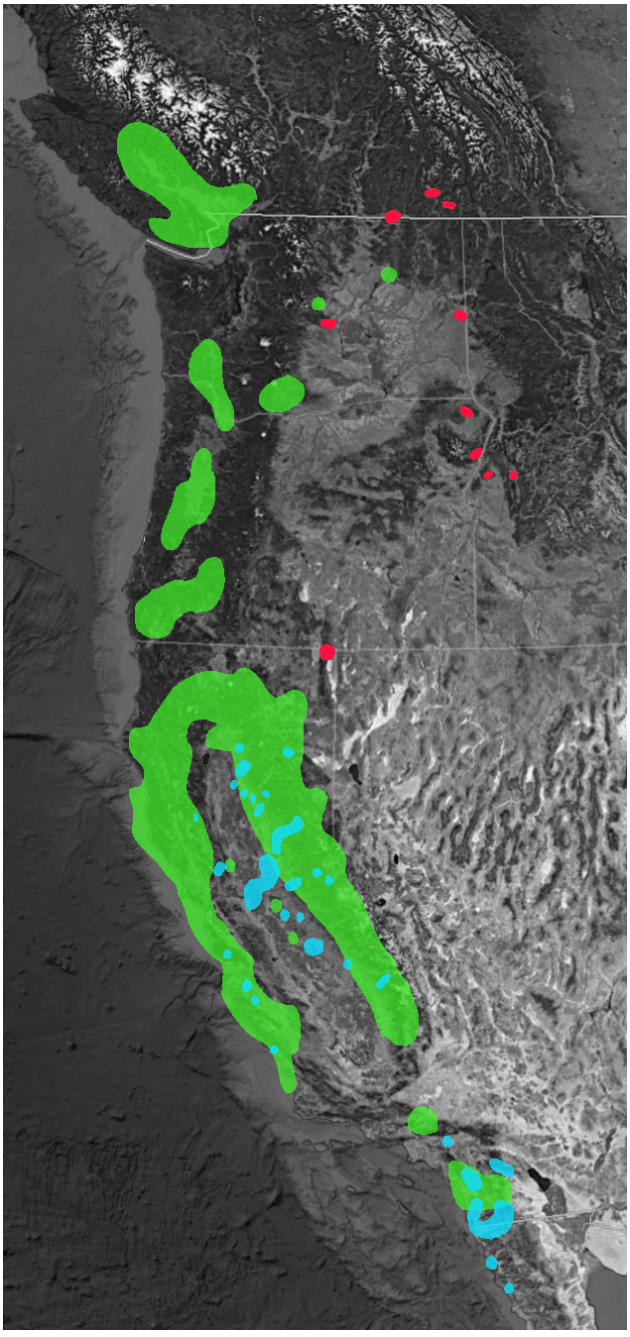

Figure 3.2: Known distribution of *Isoëtes* species from the Pacific Mediterranean Clade based on historical and contemporary collections. Range data gathered from Consortium of California Herbaria, Oregon State University and University of Washington. Green - *Isoëtes nuttallii*, Blue - *Isoëtes orcuttii*, Red - *Isoëtes minima*.

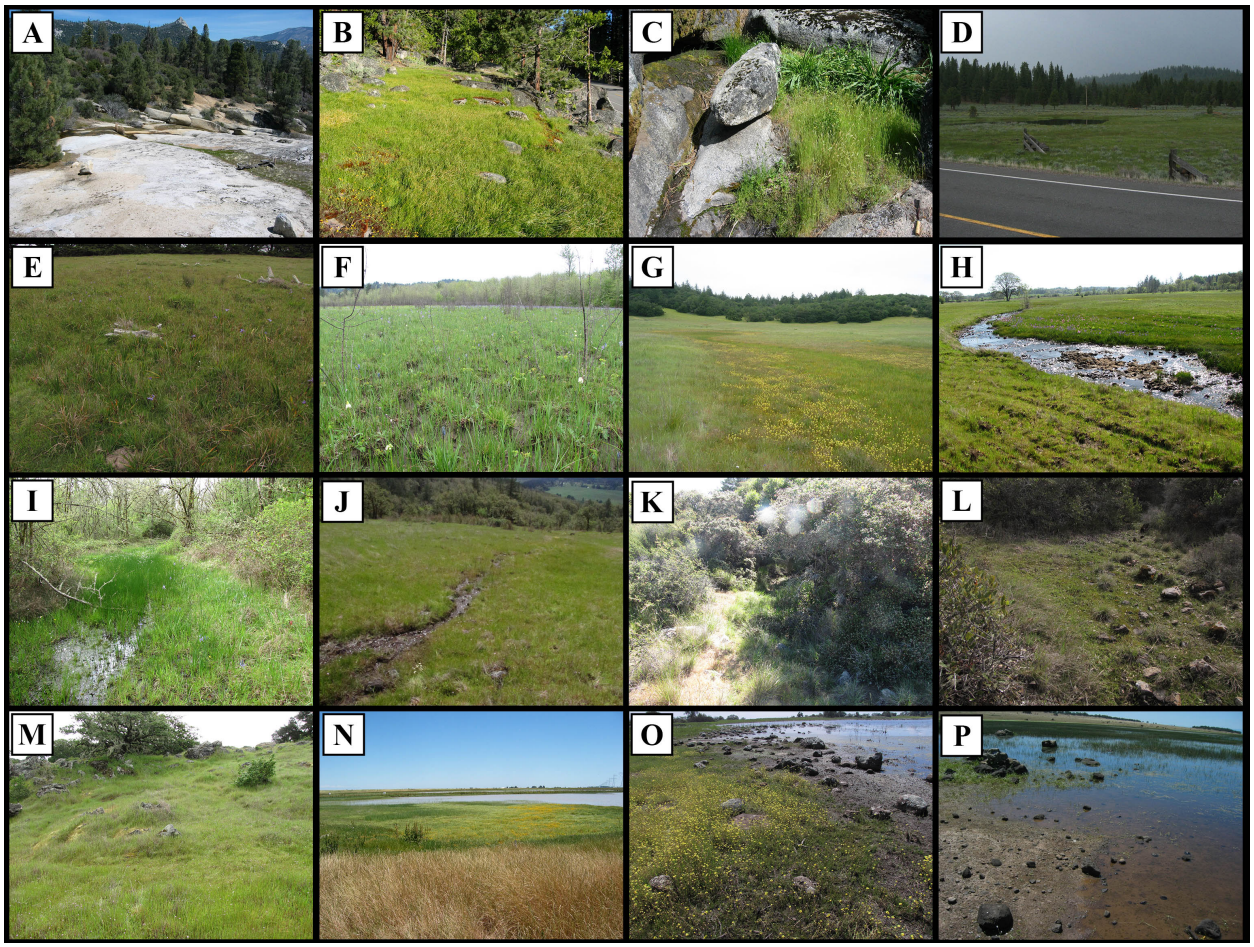

Figure 3.3: Habitats where Pacific Laurasian Clade plants can be found growing. A–C: Mossy banks; D–G: Wet meadows; H–J: Ripparian; J–M: Wet hillsides; N–P: Vernal Pools

within their range, according to records in the Consortium of Pacific Northwest Herbaria online database (https://www.pnwherbaria.org/), and provided by Duncan Thomas at Oregon State University. As suggested by their specific epithet, *I. minima* is described as a small species, with leaves that range from 2-4 cm in length. This is, however, contested by recent publications by the Committee on the status of Endangered Wildlife in Canada (Canada and Change, 2020), who claim that the small size of the plants is due to their being observed at an early stage in the growing season, and that they can reach larger sizes as the season progresses. *Isoëtes minima* is also distinct from it's nearest relatives by having an incomplete velum (the membrane that covers the sporangium imbedded in the adaxial surface of each sporophyll), as well as their short, echinate megaspore ornamentation, size (290-350mm) and toothed equatorial girdle (Table 3.1).

The final member of the clade, *I. orcuttii*, is a California floristic province endemic that is found exclusively in vernal pools west of the Sierra Nevada foothills and Peninsular Ranges

Table 3.1: Morphological characters of the three described species of *Isoëtes* in the Pacific Laurasian Clade.

Character	*Isoëtes nuttallii*	*Isoëtes orcuttii*	*Isoëtes minima*
Habit	Terrestrial to aquatic	amphibious	Terrestrial to amphibious
Habitat	Damp prairie, wet meadows, seeps, mossy banks of creeks flowing over granite, submergent aquatics in shallow creeks and waterways.	vernal pools	Damp prairie, wet meadows, seeps and vernal pools.
Latitudinal range	32.721500° : 49.950000°	31.541700° : 42.453409°	41.986479° : 49.300000°
Longitudinal range	-123.729700° : -116.561900°	-122.961900° : -116.453333°	-120.176320° : -116.249435°
Leaf number	13-60	6-14	6-12
Leaf length	8-17 cm	2-6.5 cm (rarely 10cm)	2-4cm
Leaf diameter	>1 mm	<1mm	.67-.74mm
Leaf description	rigid, ± brittle, tapered to tip	long, fine, errect, soft, flexible, tapered to tip	long, round, slender
Leaf scales	present	present	?
Sporangium length	4-7mm	2-5mm	4mm
Sporangium coloration	unpigmented	unpigmented	unspeckled
Megaspore ornamentation	± tubercled, ± shiny	small, indistinct papillae or smooth	short, slender, blunt, distinct spinules
Megaspore size	400-528 µm	216-360 µm	290-350 µm
Megaspore color	white to light grey	grey at maturity, brown/black when wet	grey to white
Microspore ornamentation	papillose	spinulose to smooth	papillose or sparsingly spinulose
Microspore length	25-20µm	21-27µm	26-31µm
Microspore color	brown	brown	white
Velum coverage	complete	complete	66-75%
Corm length	?	?	2-3mm
Corm width	?	?	3-4mm
Stomata	numerous	present	present
Bast bundles	3	0-2	4
Ligule	small, triangular	triangular	triangular, slightly elongated
sources	Pfeiffer 1922, Frye and Jackson 1913, Jepson Online, Flora of North America, personal observations	Pfeiffer 1922, Jepson Online, Flora of North America, personal observations	Pfeiffer 1922, Eaton 1898, Frye and Jackson 1913, WNHP 2005

(Fig. 3.2). Like *I. minima*, *I. orcuttii* is characterized as a smaller species, with leaves ranging from 2-6.5cm, though they are noted to reach up to 10cm on occasion. They are also noted for having either smooth or indistinctly papillate spores that range from 216-360mm in diameter, and are grey when dry and dark grey/black when wet at maturity (Table 3.1).

However, as was shown in my previous chapter, neither *I. nuttallii* nor *I. orcuttii*, as typically circumscribed and identified using the published keys (Baldwin et al., 2012; Taylor, 1993), are monophyletic, and therefore do not meet the criteria for a species. There is much debate about how one defines a species (Mayr, 1976; De Queiroz, 2007), or if the concept itself is even useful within a ranked taxonomic system (Cantino et al., 2020; Mishler, 2010; Cellinese et al., 2012). In this study, I have chosen to use the conceptual framework of the unified species concept of de Queiroz (2005); De Queiroz (2007): a species is a monophyletic, meta-populational evolutionary unit of organisms that can be identified via genetic, morphological/anatomical, ecological, and/or geographic features. Using this framing, I will use statistical analysis (MANOVA/ANOVA and Principal Components Analysis), as well as observations of the plant's habitats and qualitative morphology in order to argue for a redefinition of the taxonomic units within the PLC, and find diagnosable morphological features

of the Major Clades as described in Chapter 2.

3.2 Methods

Type Locality Identification

Type localities for two of the three taxa, *I. nuttallii* and *I. orcuttii*, were identified from herbarium records, Tropicos.org and the Pfeiffer monograph (Pfeiffer, 1922). For *I. nuttallii*, the type specimen is Nuttall s.n., 183- from the Columbia River region of Oregon / Washington. The holotype is housed at the Missouri Botanic Garden (MO-2140273). For *I. orcuttii*, the type specimen is Orcutt 1242, June 7th, 1884, from the San Diego region of California. The hologype is at the Missouri Botanic Garden (MO200567), while an isotype is at the United States National Herbarium (USNH431328).

Habitat Designations

Collections were assigned to one of five habitat types, based on my own observation and descriptions of where the plants were growing: mossy bank, wet meadow, riparian, wet hillside, vernal pool. Mossy banks are the thin band of moss covered soil that occurs along the edges of rocky outcrops with water flowing over the stone's surface (Fig. 3.3 A–C). Wet meadows are flatlands with wet to waterlogged soils, and may or may not include some form of standing or running water (Fig. 3.3 D–G). Riparian habitats are those found along creeks, streams or rivers (Fig. 3.3 H–J). Wet hillsides include habitats such as seeps that occur on a slope and are either permanently or seasonally wet, though not fully inundated (Fig. 3.3 J–M). Vernal pools are any habitat that form temporary pools over hardpan or rocky surfaces that dry over the course of the year (Fig. 3.3 N–P).

In addition to the habitat types, the plants were designated as terrestrial, amphibious, or aquatic. Terrestrial plants are those where the substrate did not appear to be submerged in water at any point of the year. Amphibious plants were those observed in conditions were their substrate was submerged for part of the year, but would become completely exposed later on in the growing season. Finally, aquatics were those that appeared to be growing in substrates that were permanently submerged.

Leaf blade length measurements and observations:

For collections made at UC Berkeley, where it was possible to examine fresh, living plants, each specimen was cleaned of any encrusting soil by gentle washing with deionized water. For plants cleaned in the lab, the wash water and soil were allowed to air dry. Once dried, the soil was collected and placed into zip-top storage bags, for later examination for shed megaspores. After the soil was cleaned from the plants, all individuals from a given collection locality were photographed using either the UC Berkeley and Jepson Herbarium's specimen

digitization stations or a Canon PowerShot SX20 IS attached to a photography stand or tripod if still in the field. In all cases, the photos were taken with scale bars and color palettes for digital measurements.

For the small number of sites where I collected prior to my time at UC Berkeley, the above images and soil resources could not be gathered as their procurement was not part of my collection protocol at the time of collection.

The imaged plants were then measured digitally in ImageJ (Collins, 2007). Scale was set with the in-image scale bars, and leaf lengths were measured using the the line segment tool. On each plant, all visible leaves were measured from the distal-most point to the base of the blade just above the sporangium. In the case of collection sites where the majority of the leaves were obviously damaged by herbivory, leaves were counted, but excluded from downstream length analysis.

Measurements were compiled into an Microsoft Excel spreadsheet, with each collection site given its own sub-sheet within the main excel file. For each plant observed, the shortest and longest leaf lengths were recorded, while the average length was calculated using all observed leaves on each individual (Appendix 3.1).

Spore extraction, scales/phyllopodia, and sex determination:

For each collection, spores from ca. 8 individual plants were examined with a Zeiss Stemi 508 stereoscope and photographed with a Canon E0S M3 digital camera. Fresh plant materials did not require any special processing and were observed directly. For dried plant material, leaf blades above the sporangia were carefully cut off and put back into silica storage. The remaining corm and leaf bases were rehydrated by placing each one in a separate petri dish and covering it with enough deionized water to completely submerge all tissue. Corms and leaf bases were allowed to rehydrate for ca. 10–20 minutes, after which the leaf bases were removed by carefully cutting them free with a fine-pointed scalpel blade. Before removing leaf bases for examination, corms were observed for the presence of scales, phyllopodia, and megaspores / microspores from the previous year (Bray et al., 2018). Corms were photographed at this stage to document the presence of scales and/or phyllopodia.

Rosettes were examined for the presence of either mega- or microsporophylls. If spore-bearing leaves were found, one was carefully removed with a pair of needle-nose forceps and a scalpel. Velum coverage of the sporangium was observed and recorded. If the removed sporangium had mature megaspores, it was moved to a petri dish, where it was cut open with a sharp pointed scalpel, and flushed using a thin-stream spray bottle to extract the spores. If the megaspores were not mature in the current year's sporangia, any mature megaspores found on the corm within the previous year's leaves were carefully removed through a combination of gentle prodding with a dissecting needle, needle-nosed forceps, and spraying with a thin-stream water bottle. If surface ornamentation was obstructed by soil, spores were placed into a 2µl round bottomed tube with 1µl of deionized water, and cleaned via 1-2 seconds of sonication in a eye-glass cleaner. Once all spores were free of the sporangia and any obstructing soil and debris, the contents of the container were passed

through filter paper under light vacuum to remove the water. Filter paper was then dried overnight in the lab fume hood. Dried megaspores were placed into labeled 1ml tubes for archival storage.

Spore measurements:

Preliminary spore photos were imported into ImageJ. Once imported, the "set scale" function was used to determine the number of pixels per millimeter in the image. In each image, the diameter of up to 20 spores from a given individual were measured. Spores from the two putatively hybrid sites (FF290 and CR5205) were not included in the measurements and downstream analysis.

Spore morphological examination:

Six to twenty megaspores from one to four individuals at each collection site were selected to showcase the morphological variation within a given site. Spores were chosen from across the range of size and ornamentation, as well as to showcase morphological uniformity or deformation. Selected spores were photographed using a Zeiss Stemi 508 stereoscope and photographed with a Canon E0S M3 digital camera. Photographs were taken at three set magnifications: 6.5x, 20x and 50x. At the higher magnifications, multiple images were taken across the focal plane to capture the entire depth of field. Once imaging was completed, spore images for a given magnification and depth of field series were imported into Adobe Photoshop using the "Load Files Into Stack" script and auto-aligned. Once auto-aligning was completed, the "auto-blend layer" function was used to merge the depth-of-field images and bring all of the spore ornamentation into focus.

Using the spore morphology descriptions in Lellinger and Taylor (1997) and Hickey (1986a), each of the spores was described based on their main ornamentation, as well as their girdle morphology and their color when wet and dry. For the Hog Lake collection, there were two spore morphologies observed. One was only seen in two spores extracted from the soil on the base of the corms, and differed greatly from that of the spores extracted from within the sporangia, in that they had highly distinct and different ornamentation and color and were much larger in size. As there were plants from the American clade (identified as *Isoëtes howellii* Engelm. using the same keys used to identify the study species) growing in the adjacent field, these anomalous spores were not considered as part of the assessment due to the high likelihood that they were from this species.

Morphological assessment of Major Clades by leaf and spore characters:

Using the Major Clades identified in the SVDQuartets analysis from Ch. 2 [Hog Lake (HL), Minima (Min), Jepson Prairie (JP), Southern California (SC), Penninsular Range (PR), Coast Range (CR), Willamette Valley (WV)] , I performed MANOVA analysis following leaf

and spore features: shortest, longest, and average leaf length for each individual, as well as the average number of leaves per individual, and smallest, largest, and average spore diameters for each individual plant. Four collection sites were omitted from the morphological assessment: FF169 because leaf measurements could not be made on fresh materials due to only dried materials being available and no photos of fresh materials having been taken with scale bars at the time of collection; FF284, because the older leaves were heavily impacted by herbivores, and therefore it was impossible to get accurate measurements of all but the newest leaves; and both FF290 and CR5205, due to the potentially hybrid/polyploid origin of the plants and the uncertainty of their placement within the Major Clades.

Tukey post-hoc tests were performed on each morphological feature that returned statistically significant differences to determine if there were statistically significant differences between the Major Clades. After the statistical analys was completed, bar plots were generated for each statistical test with pairwise comparison lines that show the Tukey post-hoc test p-values between pairs of Major Clades that were found to be significantly different. All analyses were done in R using using a custom script with the following packages:dplyr, readr, tidyverse, ggpubr, rstatix, car and broom (R Core Team, 2021).

Concurrent to the MANOVA/ANOVA analyses, principal components analyses on the above leaf and spore characters based on the Major Clades were conducted using a custom R-script with the following packages: factoextra, ape, tools, stringr, readr, missMDA. The leaf PCA included the following variables: average leaf length, shortest leaf, and longest leaf, aggregated by individuals in a given population; the spore PCA was done with the average, smallest, and largest spore diameters of each individual with observed spores in a given population.

R-scripts for the MANOVA and Tukey Post-Hoc tests can be found in Appendices 7 – 8. R-scripts for the PCA analyses can be found in Appendices 9 – 10. Leaf and spore measurements can be found in Supplements 7 – 8

3.3 Results

Habitats, identifications, and phylogenetic placement

In most cases, habitat preferences followed those of the described taxonomy. However, in four cases, the habitat, species identification and phylogenetic placement did not agree with the published descriptions: FF276, a collection from a wet hillside habitat initially identified as *I. nuttallii* was nested within the Southern Californian clade, which is the clade that potentially includes the type locality for *I. orcuttii*. This was an unusual finding, as *I. orcuttii* is described as a vernal pool species, and therefore wouldn't be expected to occur in a wet hillside habitat. FF285 was found on a wet hillside, though one that was a somewhat vernal pool like habitat in that the plants were on a rocky outcrop in thin soil in depressions on the rock surface, which is similar to a vernal pool habitat in many ways. These plants also were placed in the Southern California group. The third is KW380, the only representative

of the Peninsular Range group, which was identified as *I. orcuttii* based on leaf and spore morphology, yet occurred in the mossy bank habitat type, and was one of the units that had uncertain placement in the phylogenies (see Chapter 2), though with most of its support placing it as sister to or nested within the Sierra Nevada group. Finally, plants found in Willamette Valley (FF280, 282 and 283) and identifed as *I. nuttallii* were found growing terrestrially and as emergent and submergent aquatics. Much like FF276 above, this habitat occurrence is contradictory to the species descriptions; *I. nuttallii* is a terrestrial species, and is not known to grow with its corm completely submerged in such a way.

Leaf Length and Spore Morphology

I observed the following leaf and spore characters of the Major Clades identified in chapter 2. Measured values will be reported in a "minimum size"; "average"; "maximum size" format (Table 3.2). In light of the phylogeny, spores identified to *I. orcuttii* were compared to scanning electron microscope images taken of spores taken from the type specimen (USNH431328) of *I. orcuttii* at the U.S. National Herbarium (Fig. 3.4).

Table 3.2: Measurements of average, longest, and shortest leaf lengths in mm; average, smallest, and largest spores sizes in mm; spore coloration when dry and wet; ornamentation; and other observations made about the spores.

	Leaf			Spore						
Major Clade	Average	Shortest	Longest	Average	Smallest	Largest	Ornamentation	Color when dry	Color when wet	Other features
HL	26.05 mm	4.07 mm	74.66 mm	0.35 mm	0.28 mm	0.41 mm	Smooth	White	Bronze	Glossy. Strong points at junction of apical sutures and equatorial ridges
MIN	42.97 mm	12.61 mm	63.62 mm	0.36 mm	0.3 mm	0.42 mm	Short, dense echinate	White to light grey	Dark grey	Dentate equatorial girdle
JP	22.49 mm	4.08 mm	61.38 mm	0.29 mm	0.22 mm	0.35 mm	Smooth to sub-tuberculate	Light grey	Dark grey	
SC	37.15 mm	4.2 mm	111.23 mm	0.41 mm	0.25 mm	0.53 mm	Densely to sparsely sub-tuberculate	Light grey	Dark grey	Tubercles are so tightly packed that they begin to appear to be rugate on some spores
PR	34.2 mm	5.8 mm	97.45 mm	0.41 mm	0.3 mm	0.48 mm	Strongly tuberculate, becoming echinate	White to light grey	Dark grey	Smooth to dentate girdle
SN	53.82 mm	3.74 mm	149.03 mm	0.39 mm	0.21 mm	0.54 mm	Tuberculate	White to light-grey	White to light-grey	Apical ridges and girdle are inconsistent, with some being obscure, while others having prominent ridges and girdles
CR	87 mm	14.53 mm	206.56 mm	0.49 mm	0.37 mm	0.64 mm	Smooth to sub-echinate to strongly tuberculate	White	White	
WV	71.22 mm	9.22 mm	186.26 mm	0.41 mm	0.26 mm	0.55 mm	Strongly tuberculate to sub-echinate	White to light-grey	White to light-grey	

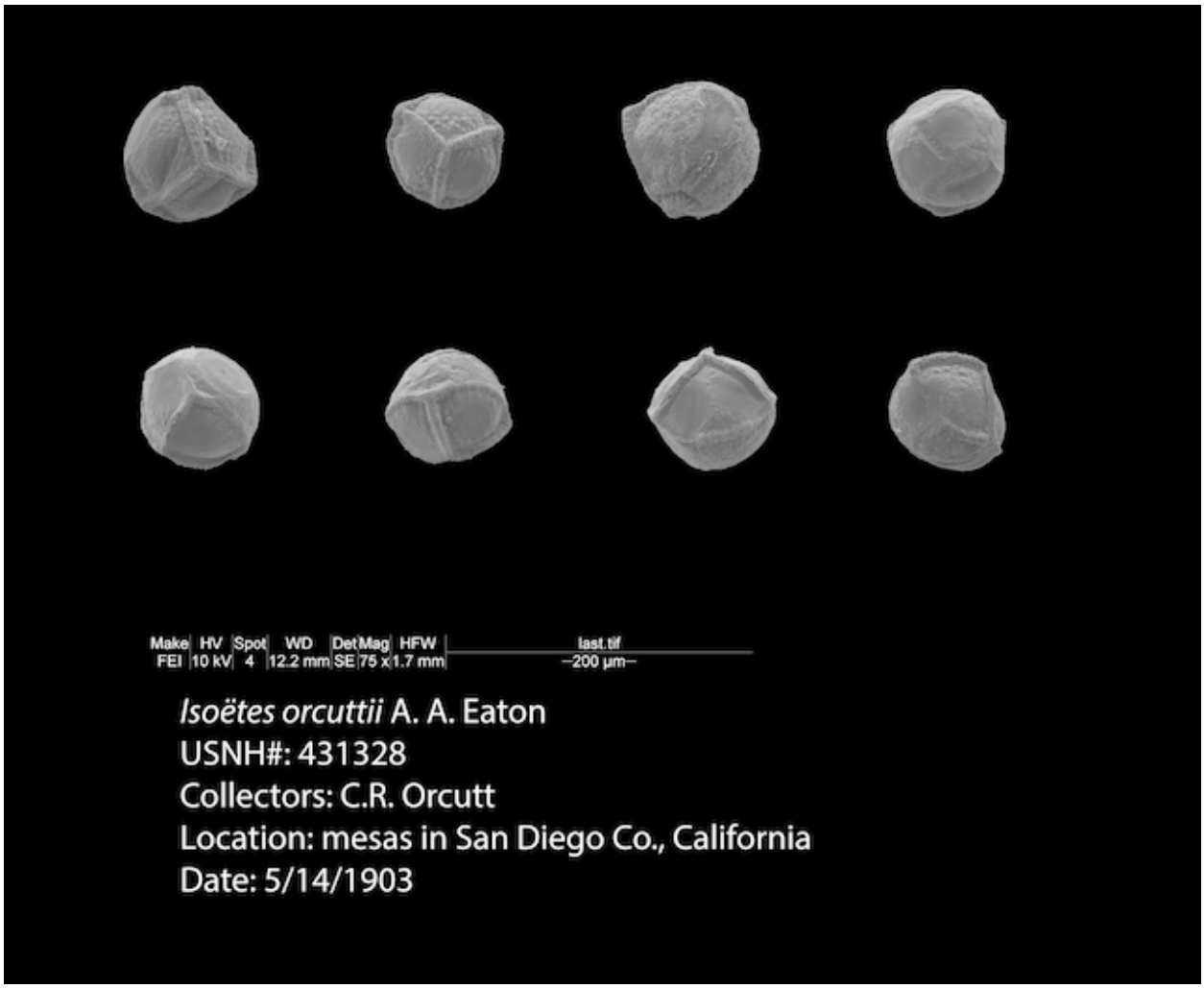

Figure 3.4: Spores from the type specimen of *I. orcuttii* (USNH431328), imaged at the U.S. National Herbarium.

HL plants have spores values of: 0.22mm ; 0.29mm ; 0.35mm, while their leaves are 4.07mm ; 26.05mm ; 74.66mm. Among all the clades observed, HL was the only one that had strongly linear leaves, most likely due to their aquatic habitat (Freund, pers obs). What sets the HL spores apart from the others in the PLC is their ornamentation, which is almost entirely smooth, with very minor tubercles in a small number of cases. Further, the joints of the apical and equatorial sutures have strongly pronounced points, giving the spores an almost triangular shape when viewed from above (Fig 3.5). In addition, the apical ridges appear to be taller and thinner than most of the other specimens observed. Also, unlike most of the other plants identified as *I. orcuttii*, the spores of these plants are white when dry, and an almost bronze grey when wet.

The spores of the Minima Major Clade, which represented the *I. minima* plants in the analysis, were 0.30mm ; 0.36mm ; 0.42mm in diameter. Min leaves have the following

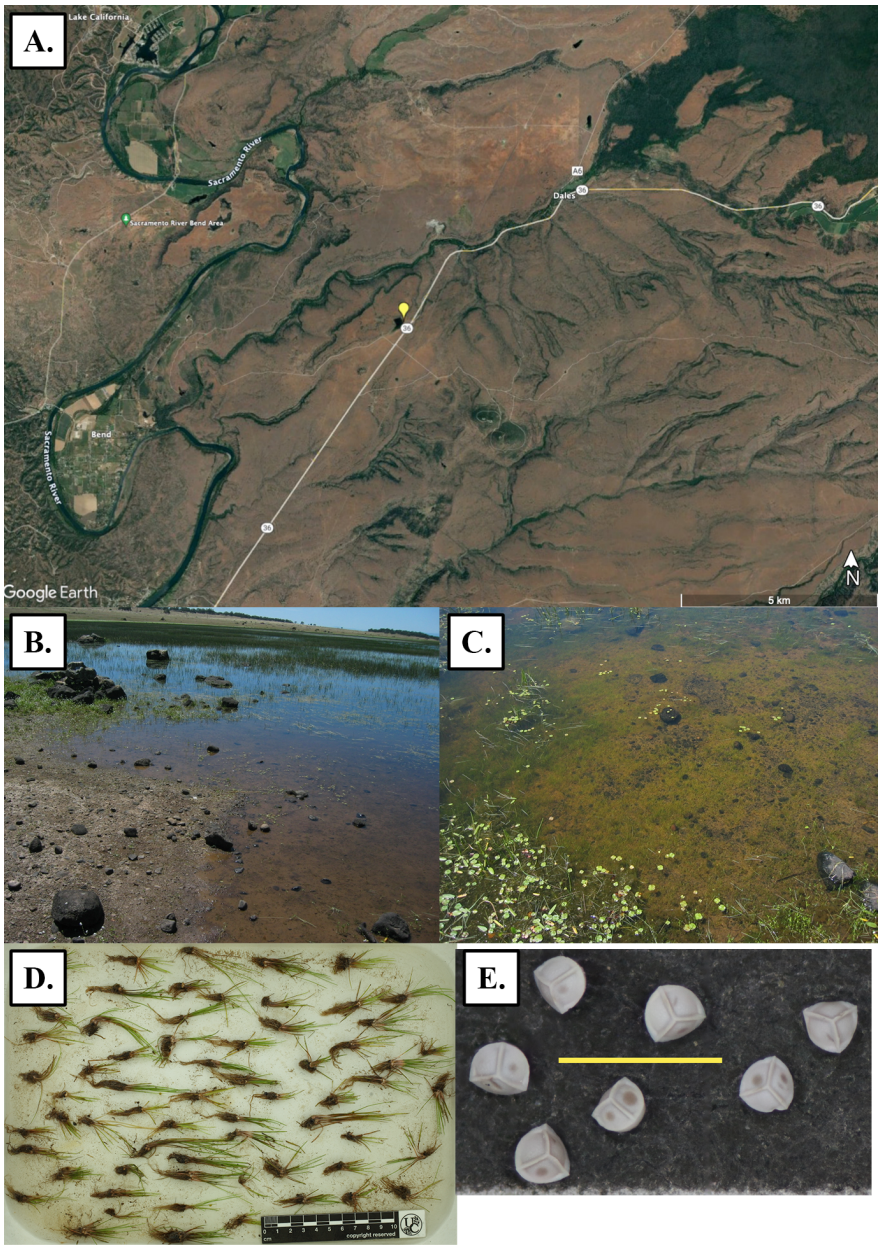

Figure 3.5: Hog Lake Major Clade collection 1/1 - FF254. A. Locality of the collection site (Lat: 40.284446°, Long: -122.119057°); B. Wide shot of northern end of the Hog Lake pool where the plants were collected.; C. Plants *in situ*. The light green patch in the center of the photo are the Hog Lake plants, with the adjacent darker green patches a mixture of *Eleocharis* sp. and other water plants.; D. Plants after collection and cleaning. Scale bar is in centimeters; E. Merged images of spores. Yellow scale bar = 1mm.

values: 12.61mm ; 42.97mm ; 63.62mm. Ornamentation wise, they had short, dense echinate ornamentation with a distinctive, dentate equatorial girdle. The spores themselves were light grey in color when dry, and dark grey when wet (Fig. 3.6).

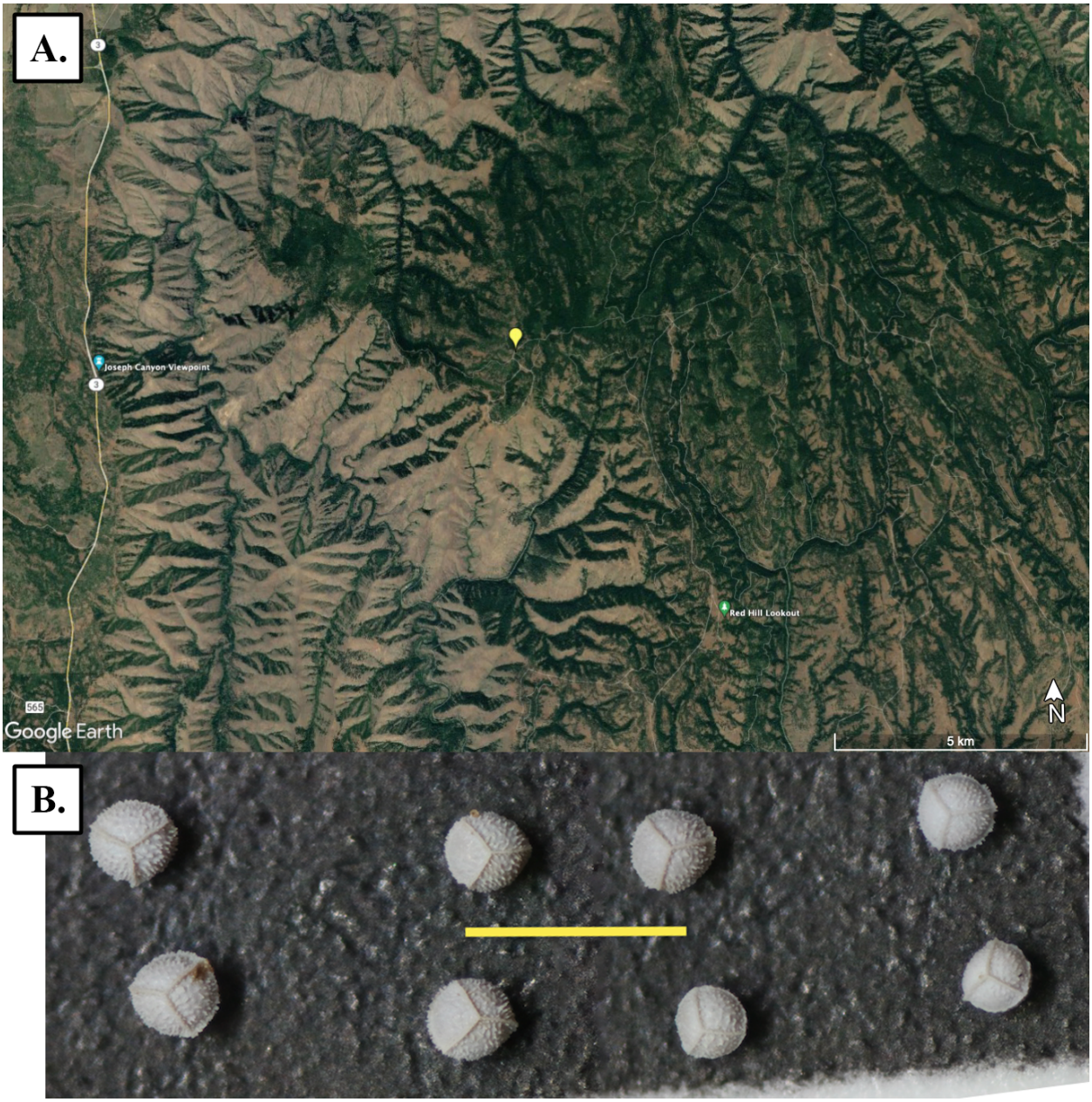

Figure 3.6: Minima Major Clade collection 1/1 - DT13378. A. Locality of the collection site (Lat: 45.83542°, Long: -117.15324°); B. Merged images of spores. Yellow scale bar = 1mm.

The Jepson Prairie plants have both smaller spores and smaller overall leaf lengths compared to the rest of the plants in the study, with the measured values of the former 0.22mm ;

0.29mm ; 0.35mm, and the latter 4.08mm ; 22.49mm ; 61.38mm. Their spore ornamentation is smooth to sub-tuberculate, with the spores themselves light grey in color when dry, and dark grey when wet. One sporangium had spores that were dark brown in color, regardless of moisture, and were glossy instead of dull, but was not unusual in terms of its placement in the phylogeny (Fig. 3.7

Spores from the Southern California group had the following min/average/max sizes: 0.25mm ; 0.41mm ; 0.53mm. Leaves observed had the following measurements: 4.20mm ; 37.15mm ; 111.23mm. Ornamentation is densely to sparsely sub-tuberculate, with many small tubercles densely packed on their surface. In some cases, the tubercles are so tightly packed that they begin to appear to be rugate, and are white to very light grey when dry, and dark grey when wet (Figs. 3.8—3.10)

Plants in the Peninsular Range group have spores with a sizes of 0.30mm ; 0.41mm ; 0.48mm in size, while their leaf measurements of 5.80mm ; 34.20mm ; 97.45mm. The ornamentation is strongly tuberculate, becoming echinate, with a small number of spores having reticulate ornamentation, while the girdle is smooth to dentate, though the latter is not as common as in *I. minima*. Their coloration was white to light grey when dry, and dark grey when wet (Fig. 3.11)

Sierra Nevada plants have spores measurement values of 0.21mm ; 0.39mm ; 0.54mm and leaf values of 3.74mm ; 53.82mm ; 149.03mm. Spore ornamentation for these plants is tuberculate, with a small number of spores tuberculate-sub-echinate, and mostly white to light-grey in color, but not appearing much darker when wet, while the apical ridges and girdle are inconsistent, with some being obscure, while others having prominent ridges and girdles. (Figs. 3.12–3.16)

The spores from the Coast Range plants are 0.37mm ; 0.49mm ; 0.64mm in size, with their leaves 14.53mm ; 87.00mm ; 206.56mm. The ornamentation of these spores is smooth to sub-echinate to strongly tuberculate, and they are white in color when wet and dry. The apical ridges and equatorial girdles are inconsistent across samples, with some being broad and rounded, while others are narrower sharp. (Figs. 3.17–3.19)

Finally, the Willamette Valley spores are 0.26mm ; 0.41mm ; 0.55mm in size, and leaves with a range of 9.22mm ; 71.22mm ; 186.26mm. The spores are white to light-grey in color both when dry and when wet, and are finely to strongly tuberculate with a few becoming sub-echinate. Equatorial ridges and apical sutures are inconsistent in the same way as the Coast Range plants described above (Figs. 3.20–3.22).

Spores from the puatative hybrid sites, FF290 and CR5205, both showed obvious signs of abnormality. Of the two, CR 5205 had less obvious deformity of it's spores. Rather, they had considerable and obvious disparity in their sizes (Fig. 3.23). Conversely, FF290 had highly deformed spores in addition to the size disparities, with the most extreme deformities consisting of fused or oblong spores(Fig. 3.24). In *Isoëtes*, either of these morphologies are typical of hybrids or polyploids (Musselman et al., 1996, 1997; Taylor et al., 1985).

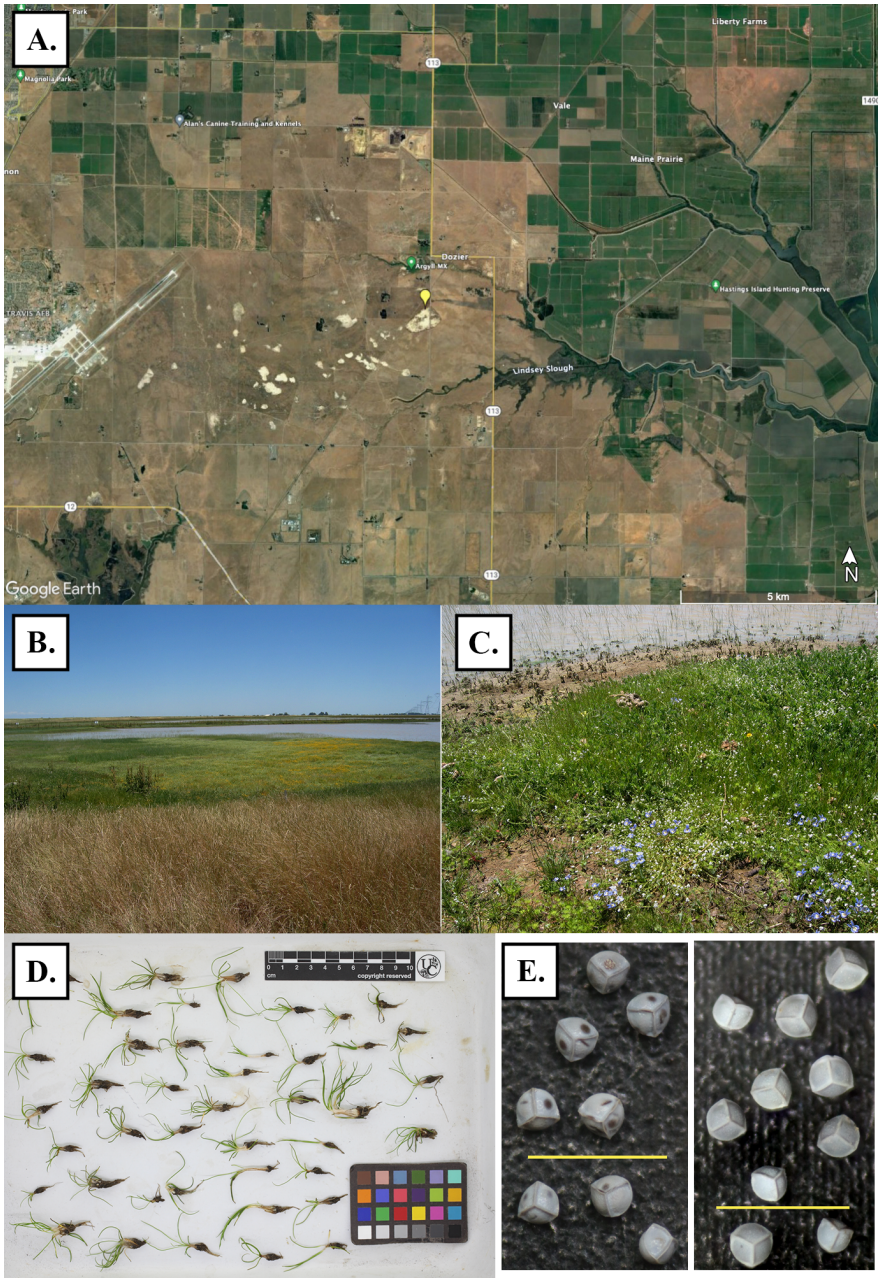

Figure 3.7: Jepson Prairie Major Clade collection 1/1 - FF262. A. Locality of the collection site (Lat: 38.27240°, Long: -121.82690°; B. Wide shot of northern end of the large Jepson Prairie vernal pool where the plants were collected; C. Close up of the habitat where the plants were collected. Plants are not visible due to being interspersed within the existing vegetation. D. Plants after collection and cleaning. Scale bar is in centimeters.; E. Merged images of spores. Yellow scale bar = 1mm.

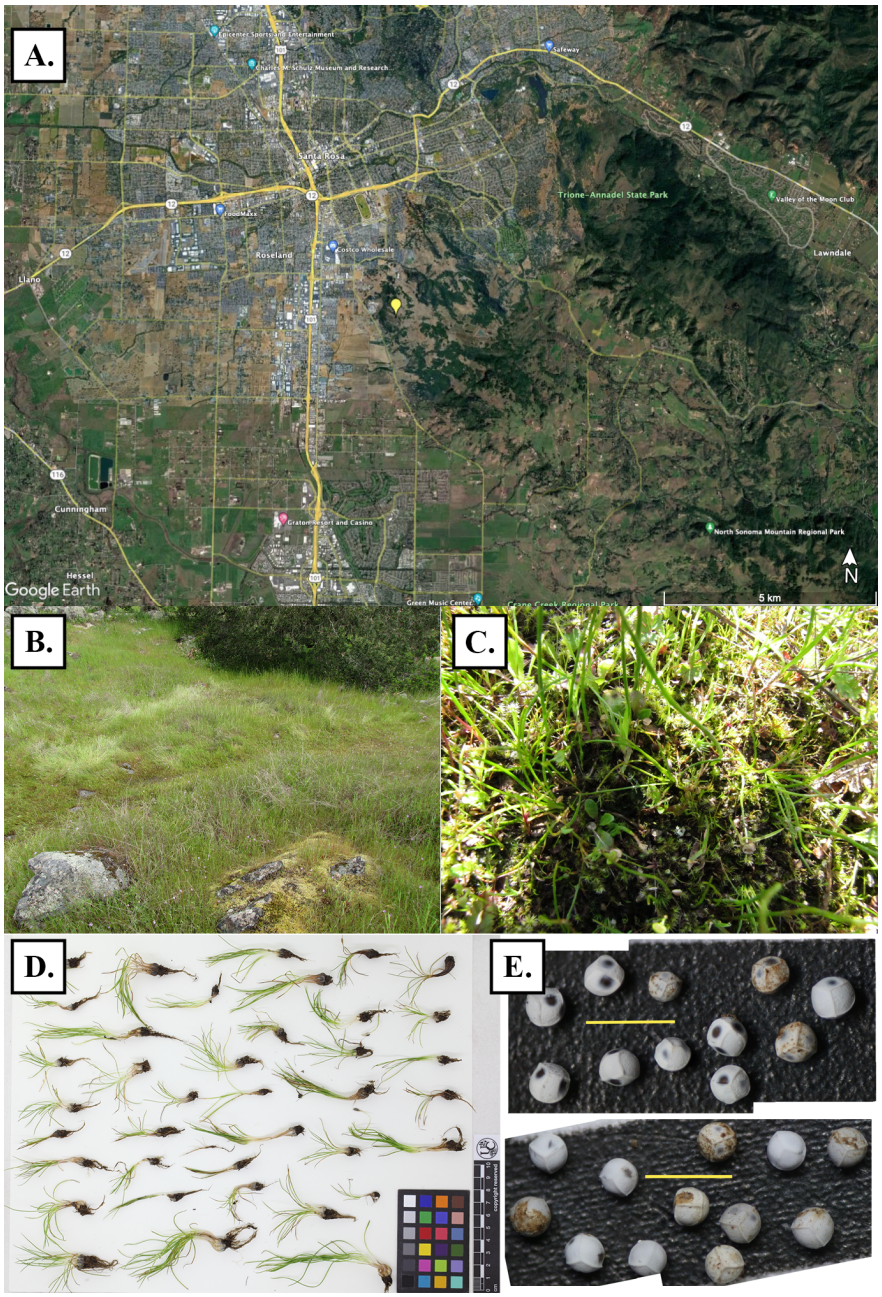

Figure 3.8: Southern California Major Clade collection 1/3 - FF279. A. Locality of the collection site (Lat: 38.40514°, Long: -122.69225°); B. Image of the slight depression where most of the plants were collected; C. Close up of the habitat where the plants were collected. Plants are not visible due to being interspersed within the existing vegetation. D. Plants after collection and cleaning. Scale bar is in centimeters.; E. Merged images of spores. Yellow scale bar = 1mm.

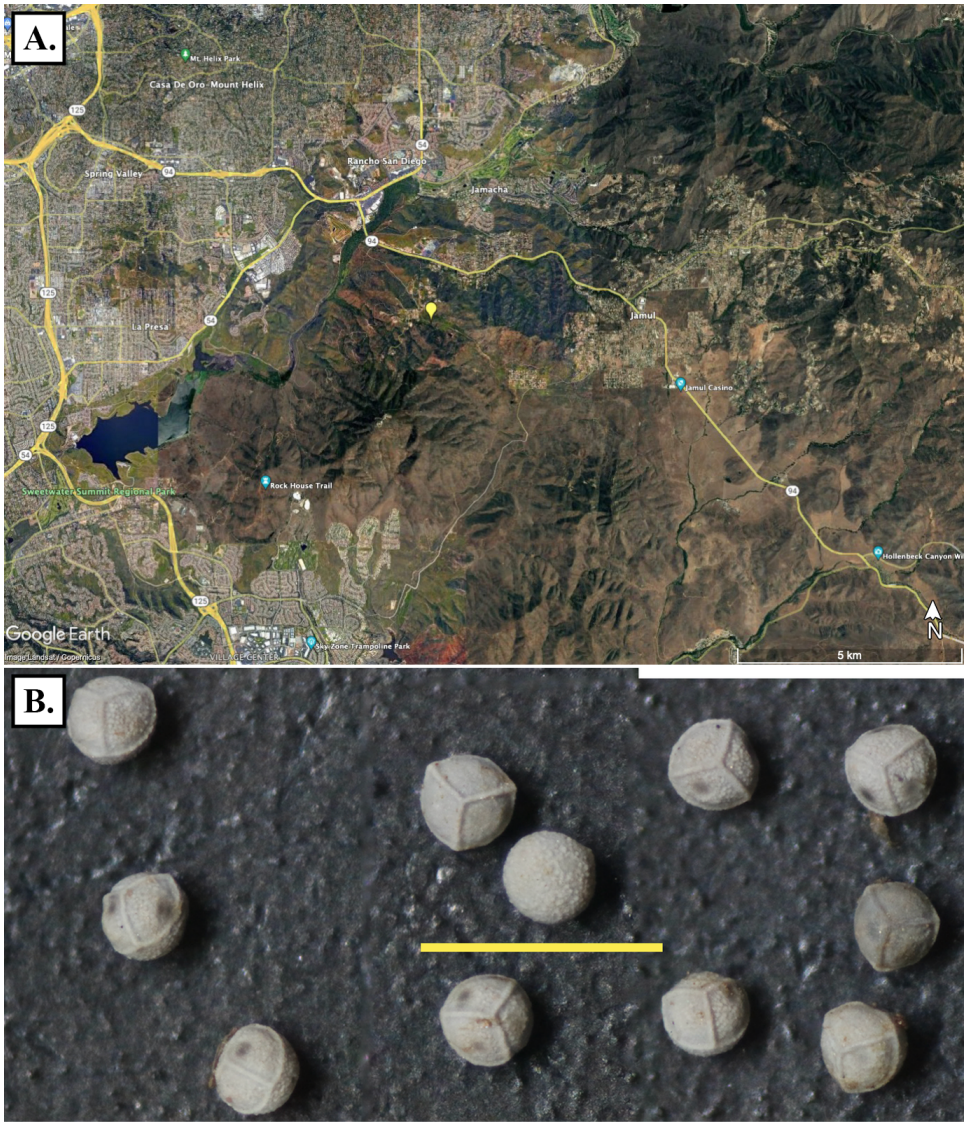

Figure 3.9: Southern California Major Clade collection 2/3 - FF285. A. Locality of the collection site (Lat:32.71539° Long: -116.92503°); B. Merged images of spores.

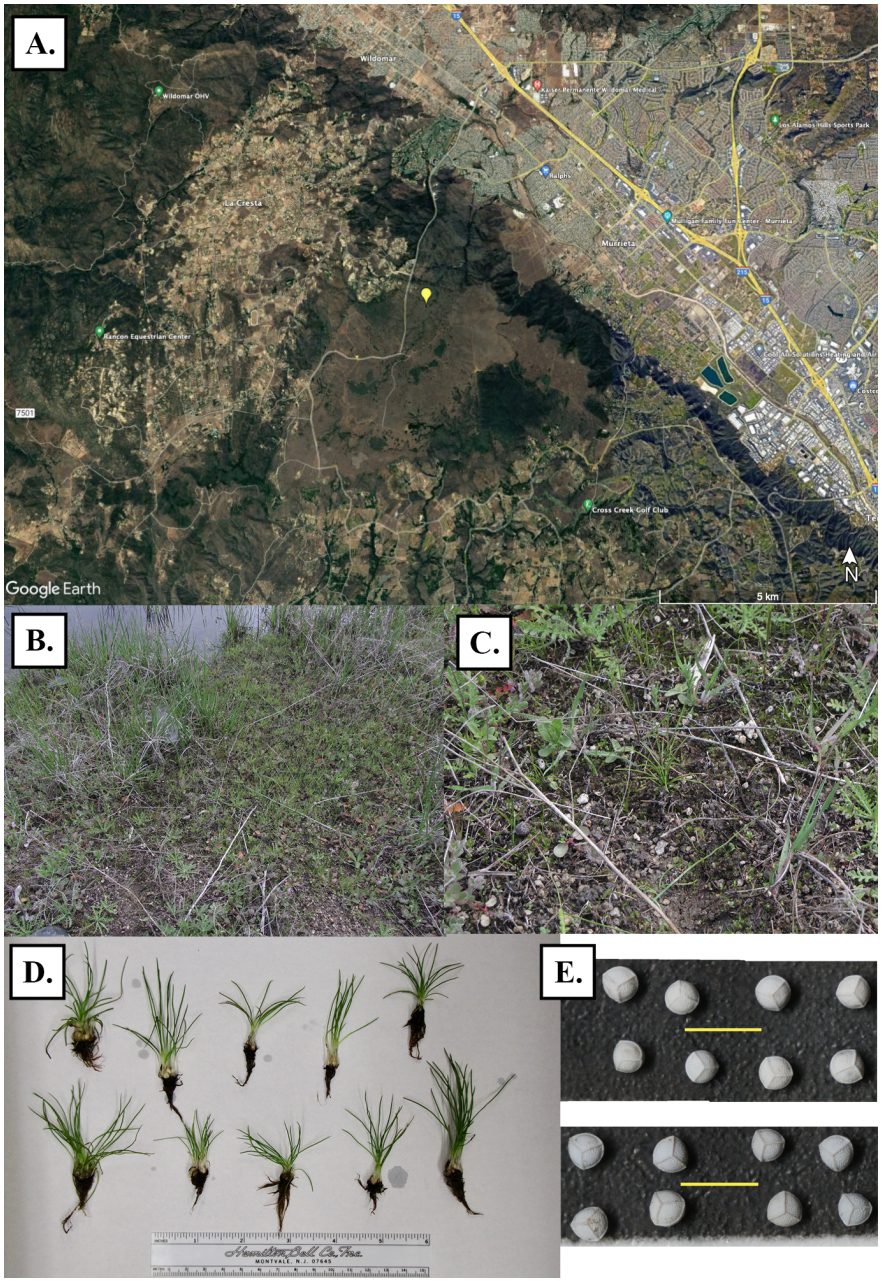

Figure 3.10: Southern California Major Clade collection 3/3 - FF286. A. Locality of the collection site (Lat: 33.53882°, Long: -117.266284°); B. Image of the pond edge where most of the plants were collected; C. Close up of the habitat where the plants were collected. D. Plants after collection and cleaning. Scale bar is in centimeters.; E. Merged images of spores. Yellow scale bar = 1mm.

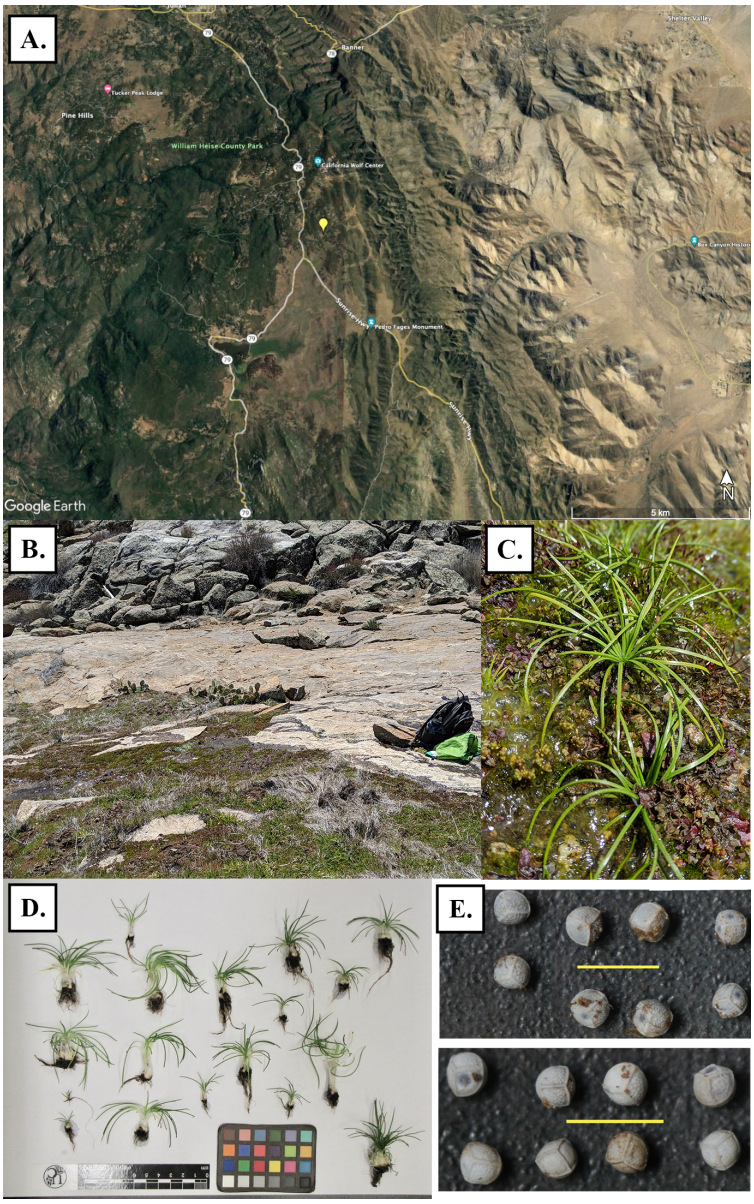

Figure 3.11: Peninsular Range Major Clade collection 1/1 – KW380. A. Locality of the collection site (Lat: 33.53882°, Long: -117.266284°); B. Image of the site habitat. Plants were found in the dark, moss covered soil in the foreground. C. Close up of the plants *in situ*. D. Plants after collection and cleaning. Scale bar is in centimeters.; E. Merged images of spores. Yellow scale bar = 1mm. Photos B. and C. taken by Dr. Keir Wefferling and used with permission.

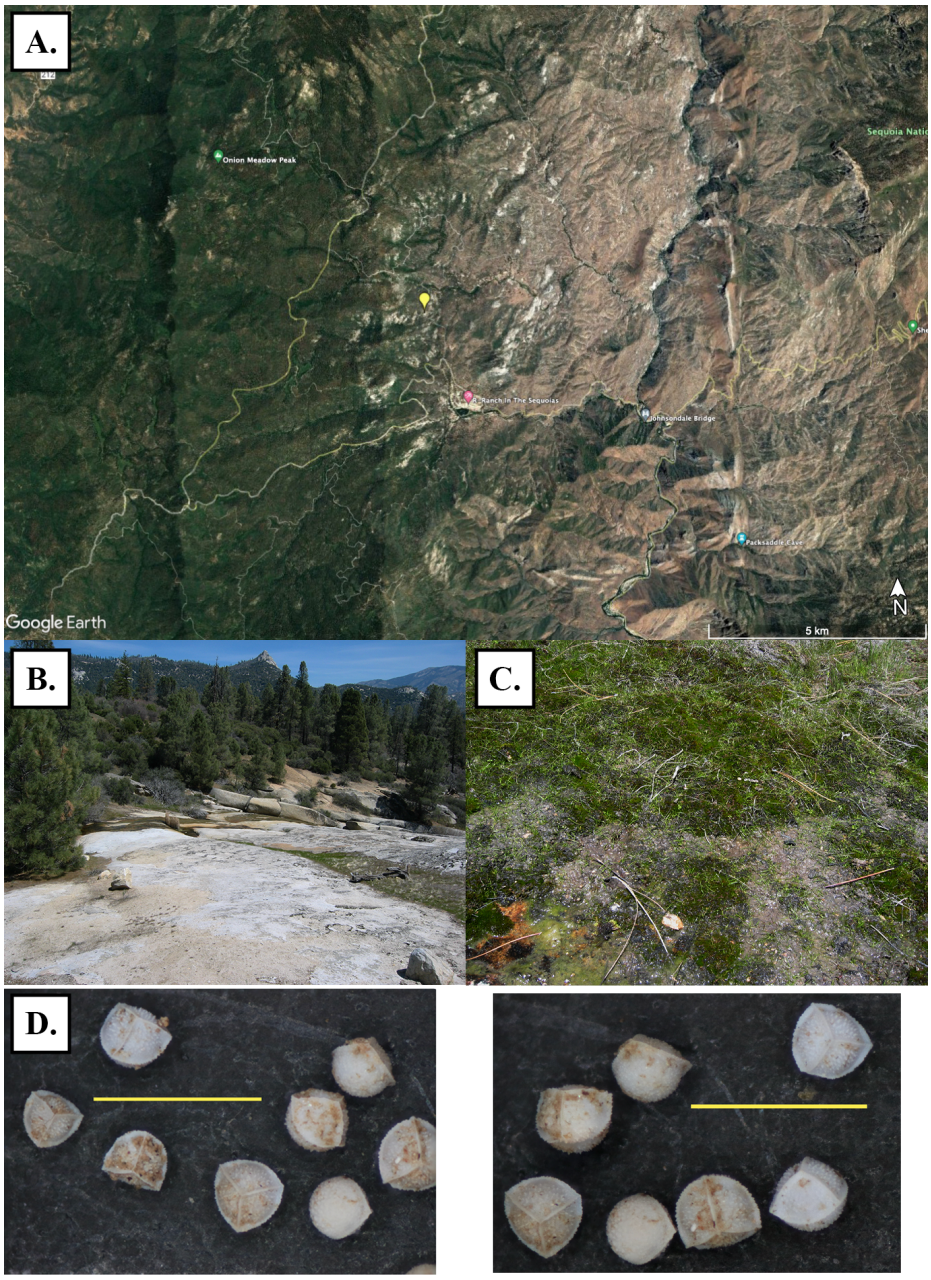

Figure 3.12: Sierra Nevada Major Clade collection 1/5 - FF169. A. Locality of the collection site (Lat: 35.99341°, Long: -118.54799°); B. Image of the granite outcrop where the plants were found. The habbitat where the plants occur is the thin band of green along the edge of the exposed rock by the water; C. Close up of the habitat where the plants were collected. The *Isoëtes* are the lighter green in the dark green moss matrix. D. Merged images of spores. Yellow scale bar = 1mm.

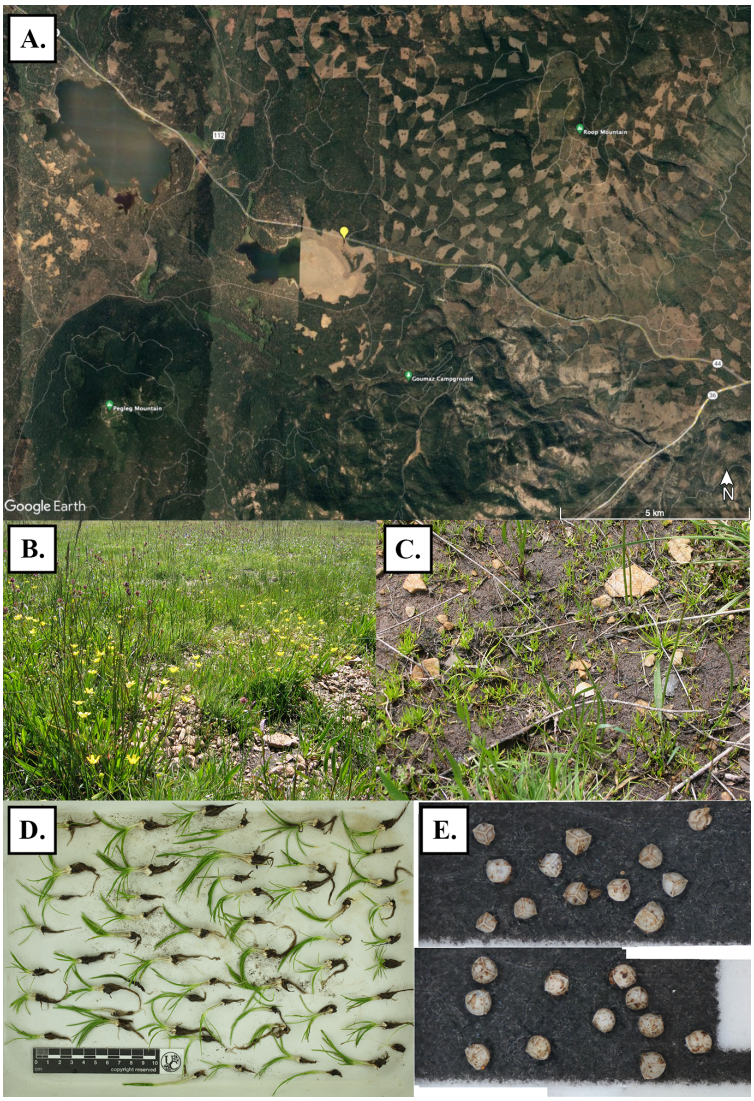

Figure 3.13: Sierra Nevada Major Clade collection 2/5 - FF257. A. Locality of the collection site (Lat: 40.4444°, Long: -120.881551°); B. Image of the region of the meadow where the plants were collected.; C. Close up of the habitat where the plants were collected. D. Plants after collection and cleaning. Scale bar is in centimeters.; E. Merged images of spores. Yellow scale bar = 1mm.

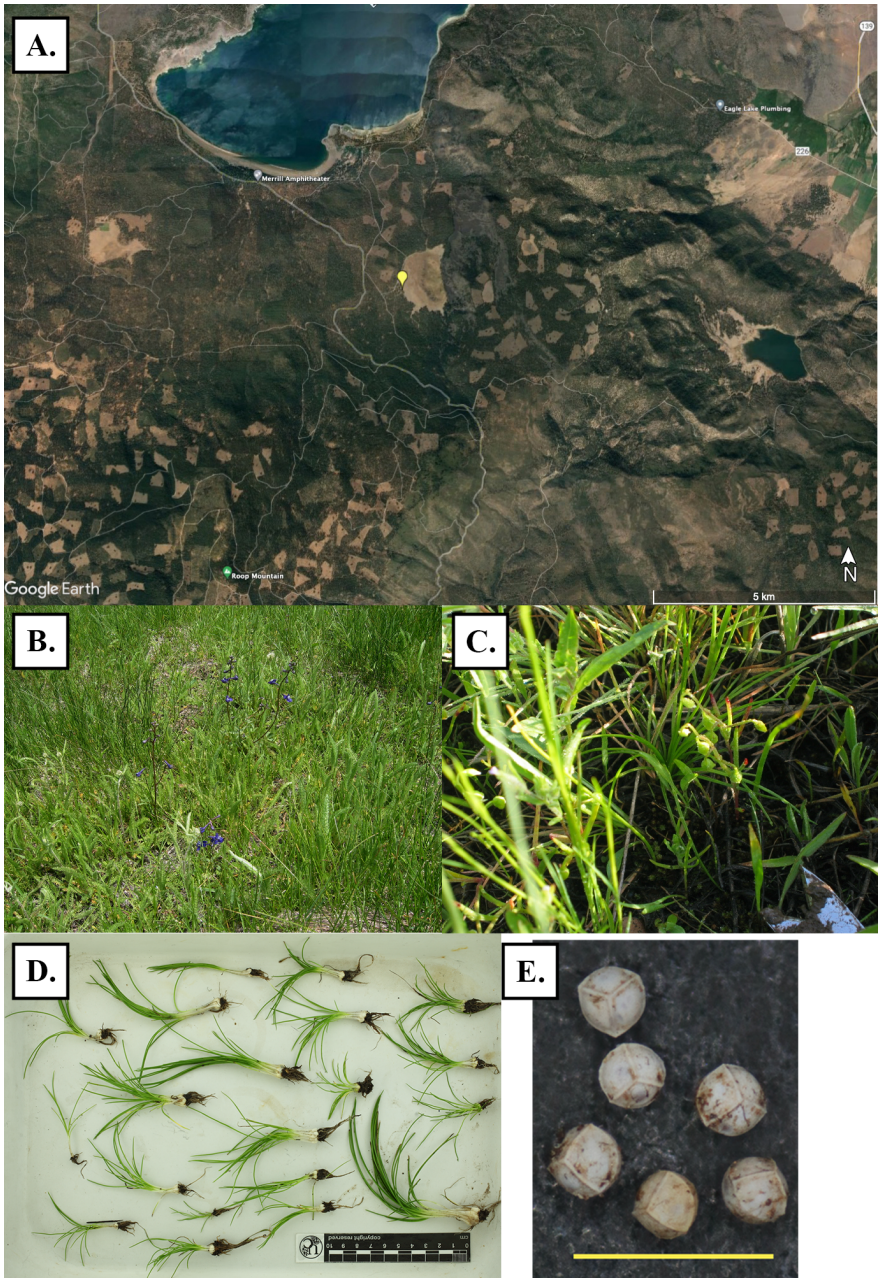

Figure 3.14: Sierra Nevada Major Clade collection 3/5 - FF258. A. Locality of the collection site (Lat: 40.525134°, Long: -120.766309°); B. Image of the region of the meadow where the plants were collected.; C. Close up of the habitat where the plants were collected. D. Plants after collection and cleaning. Scale bar is in centimeters.; E. Merged images of spores. Yellow scale bar = 1mm.

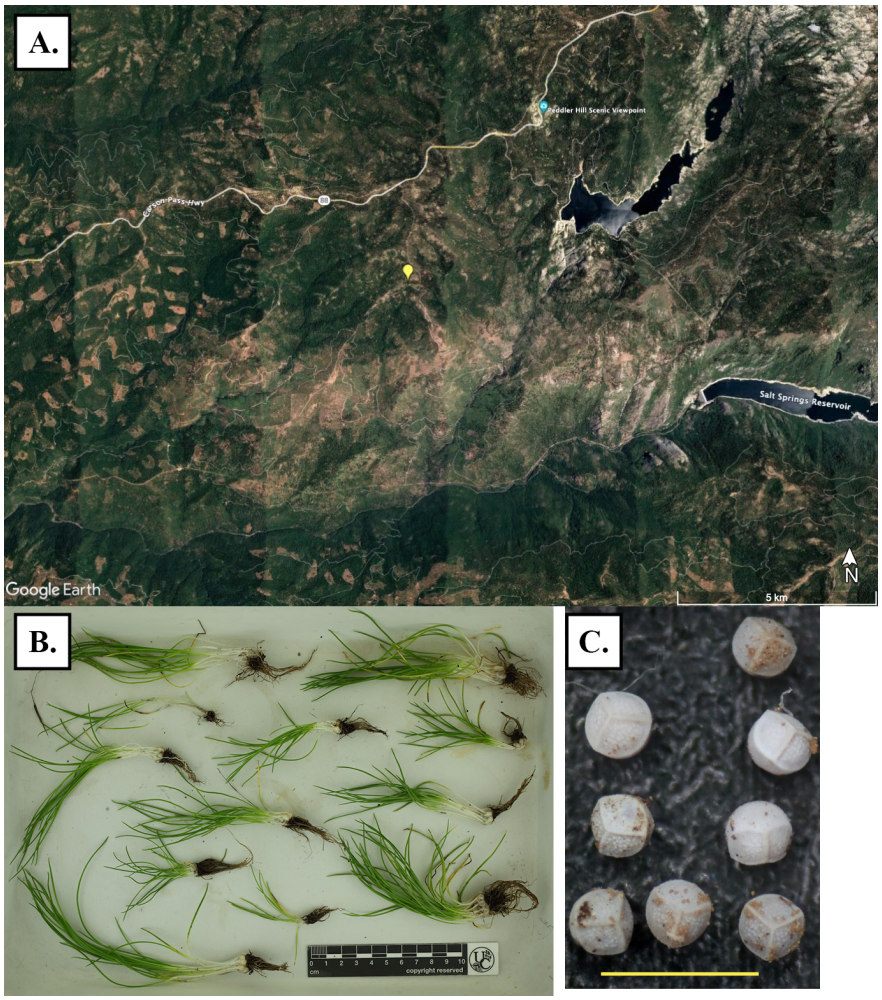

Figure 3.15: Sierra Nevada Major Clade collection 4/5 - FF261. A. Locality of the collection site (Lat: 38.525331°, Long: -120.304478°); B. Plants after collection and cleaning. Scale bar is in centimeters.; C. Merged images of spores. Yellow scale bar = 1mm.

Velum coverage, phyllopodia, and scales

I. minima plants and the Willamette Valley clade have incomplete vela. All other plants had complete vela. Of the samples observed, only DT13378 (Min) lacked scales, while only the following samples had phyllopodia: FF279 (SC), FF280(WV), and FF285 (SC).

Significant statistical differences in leaf and spore sizes

All metrics in the MANOVA analyses of leaves and spores came back as highly significant at the $p \leq 0.05$ threshold (Table 3.3-3.4).

Plants of the morphotype identified in the field as *I. orcuttii* (Hog Lake, Southern California, Jepson Prairie, and Peninsular Range), were not significantly different from each other

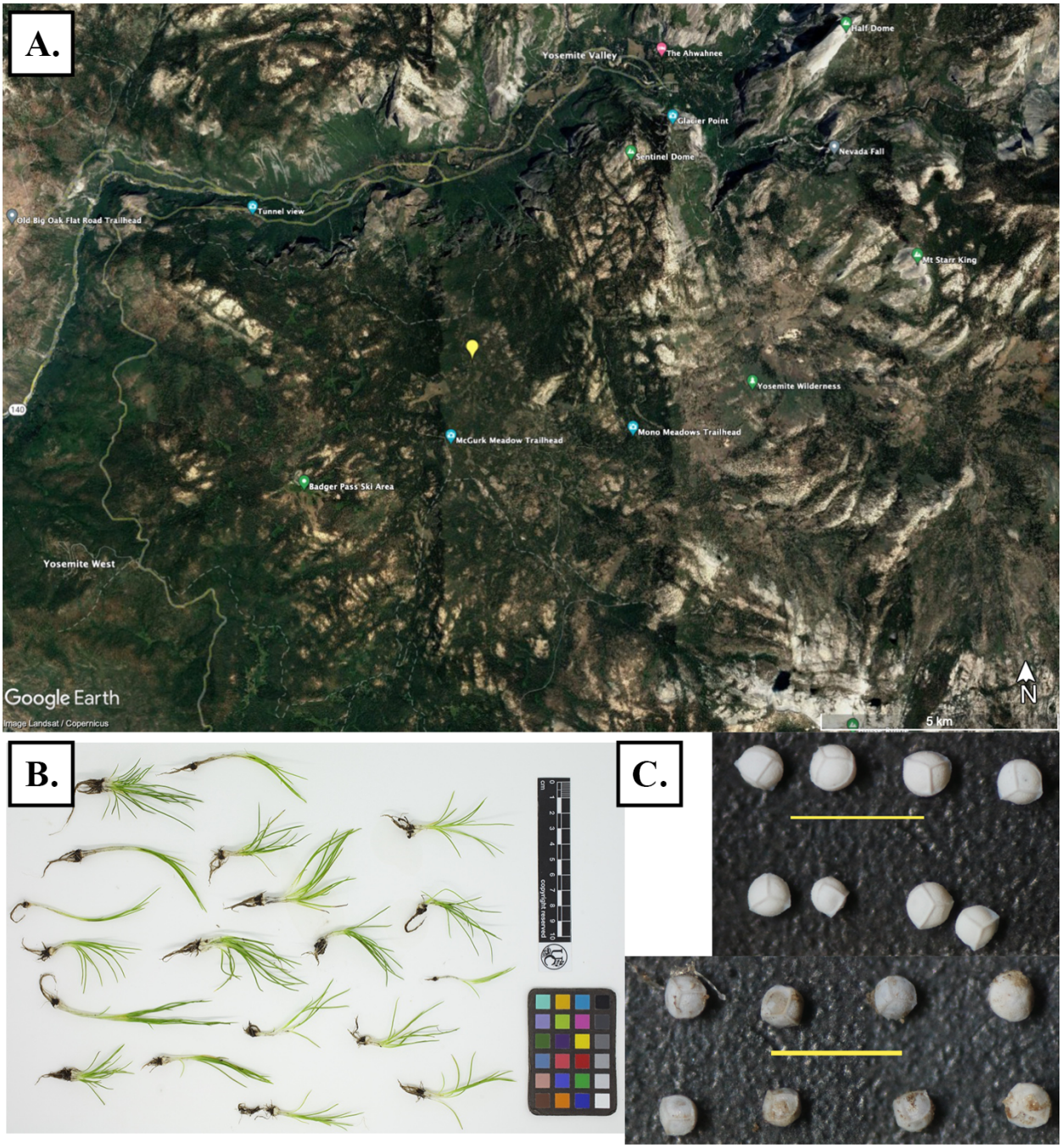

Figure 3.16: Sierra Nevada Major Clade collection 5/5 - CR5256. A. Locality of the collection site (Lat: 37.68443°, Long: -119.62164°); B. Close up of the habitat where the plants were collected. The *Isoëtes* are the lighter green in the dark green moss matrix. C. Merged images of spores. Yellow scale bar = 1mm.

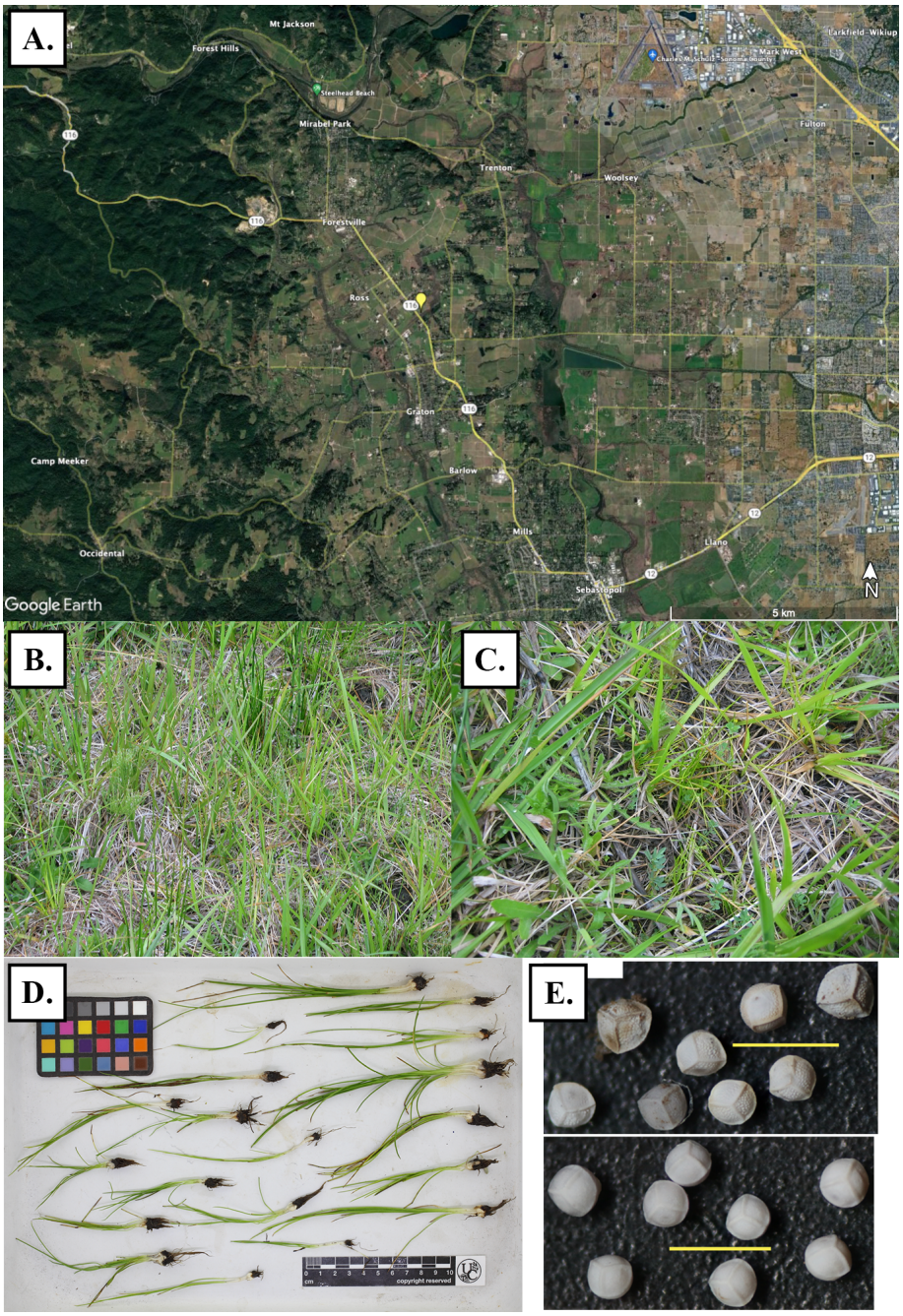

Figure 3.17: Coast Range Major Clade collection 1/3 - FF263. A. Locality of the collection site (Lat: 38.45578°, Long: -122.8703°); B. Image of the region of the meadow where the plants were collected.; C. Close up of the habitat where the plants were collected. D. Plants after collection and cleaning. Scale bar is in centimeters.; E. Merged images of spores. Yellow scale bar = 1mm.

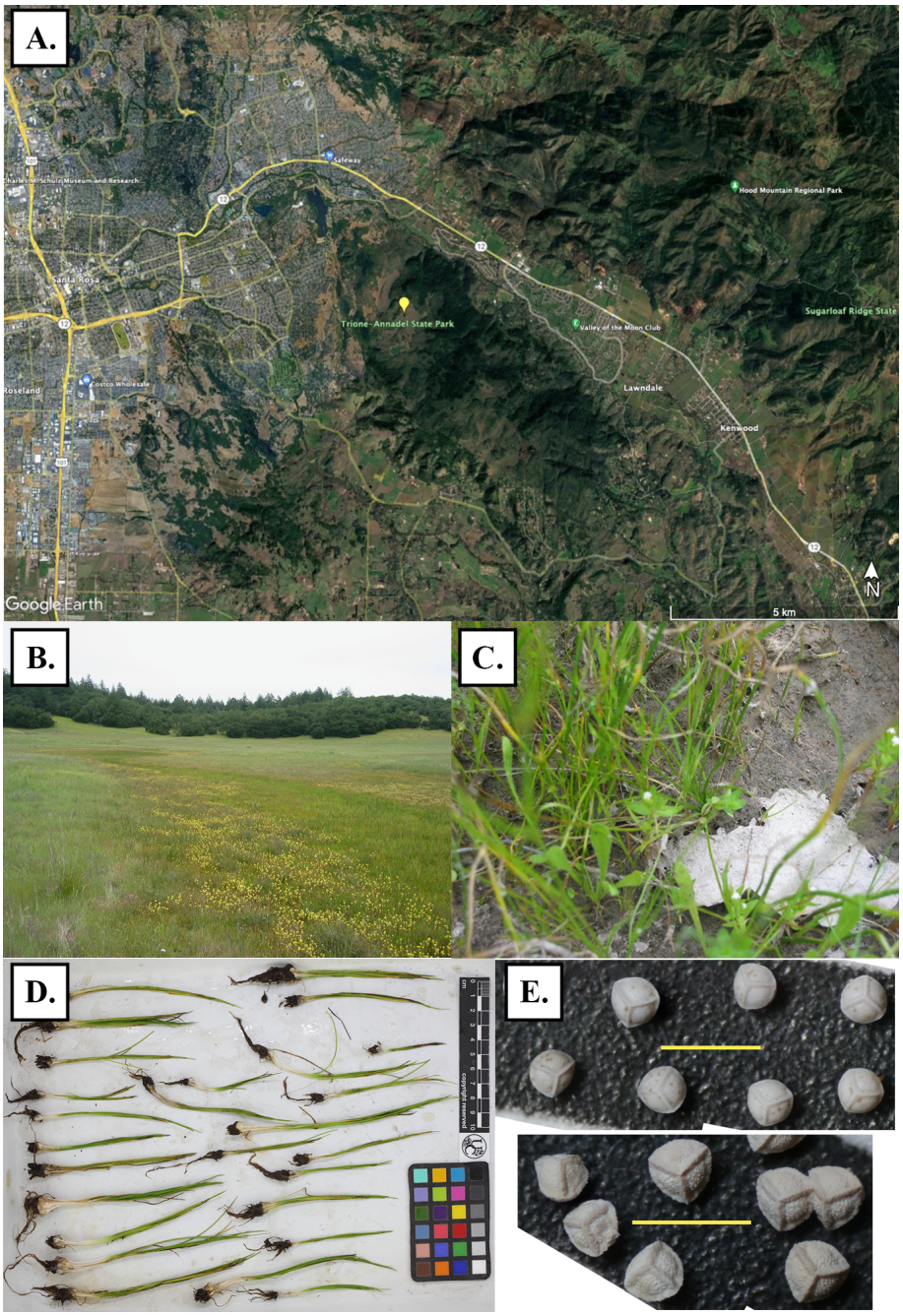

Figure 3.18: Coast Range Major Clade collection 2/3 - FF266. A. Locality of the collection site (Lat: 38.43426°, Long: -122.6309°); B. Image of the region of the meadow where the plants were collected.; C. Close up of the habitat where the plants were collected. D. Plants after collection and cleaning. Scale bar is in centimeters.; E. Merged images of spores. Yellow scale bar = 1mm.

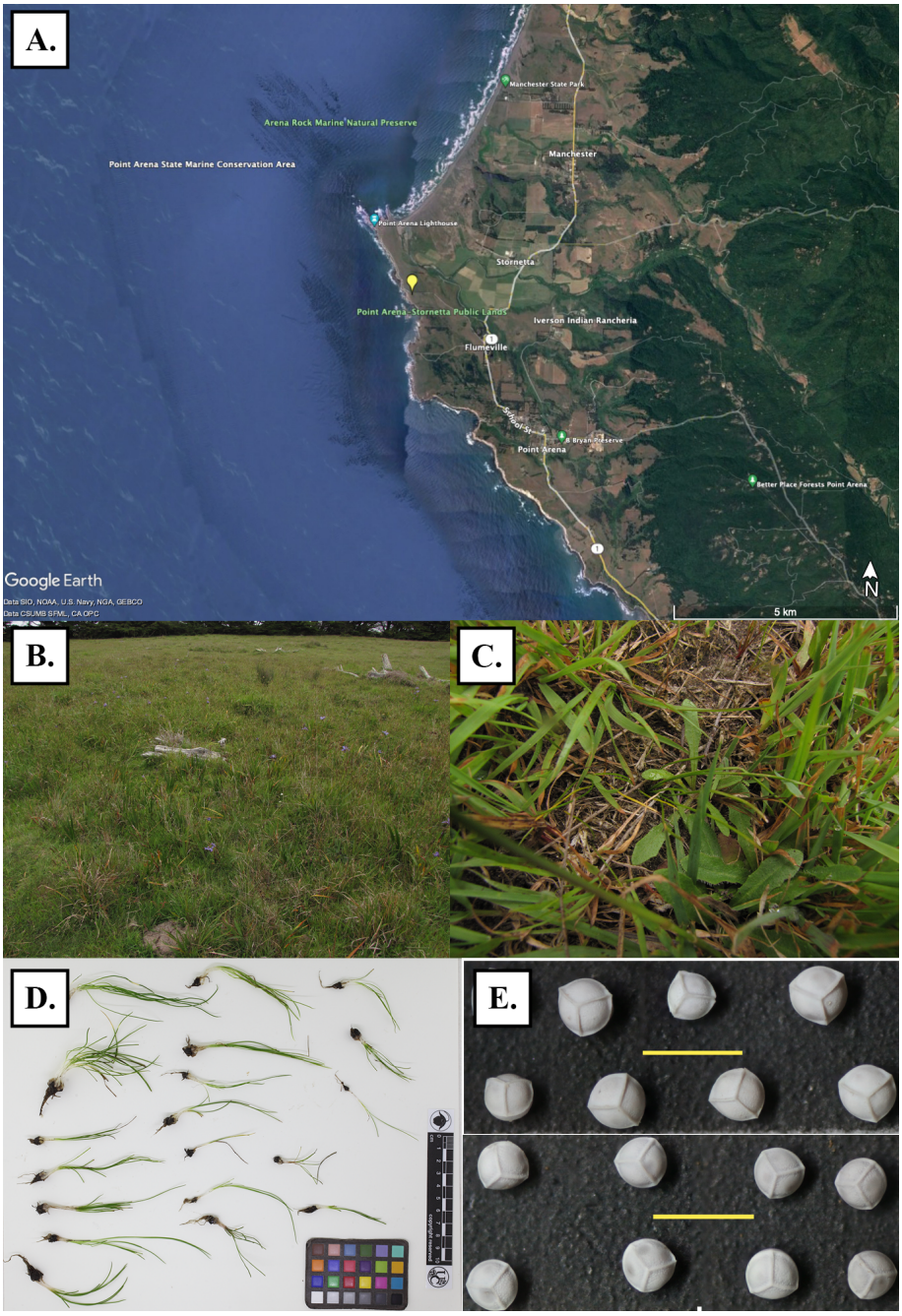

Figure 3.19: Coast Range Major Clade collection 3/3 - FF284. A. Locality of the collection site (Lat: 38.93970833°, Long: -123.7286111°); B. Image of the region of the meadow where the plants were collected.; C. Close up of the habitat where the plants were collected.; D. Plants after collection and cleaning. Scale bar is in centimeters.; E. Merged images of spores. Yellow scale bar = 1mm.

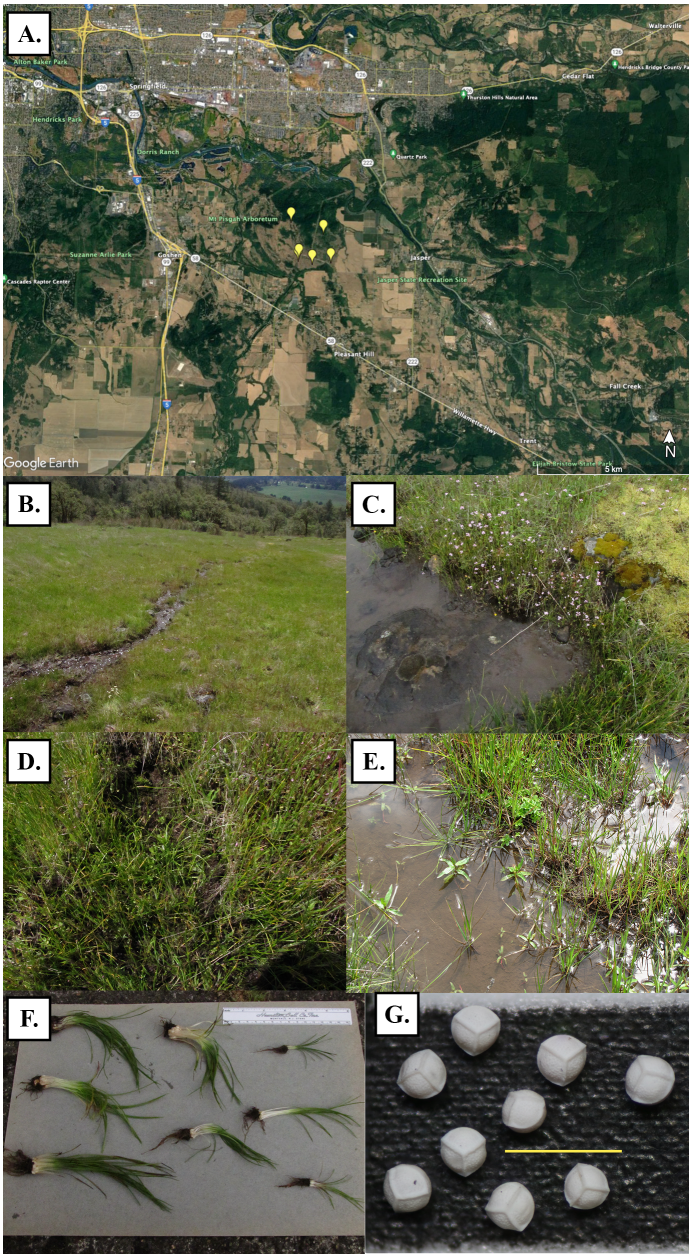

Figure 3.20: Willamette Valley Major Clade collection 1/3 - FF280. A. Locality of the collection site (Lat: 43.994397°, Long: -122.957115°); B. The hillside spring where some of the collections were made.; C. The creek at the base of Mt. Pisgah were other collections were made.; D. Close up of the habitat in B where the plants were collected.; E. Close up of the habitat in C where the plants were collected.; F. Plants after collection and cleaning. Scale bar is in centimeters.; G. Merged images of spores. Yellow scale bar = 1mm.

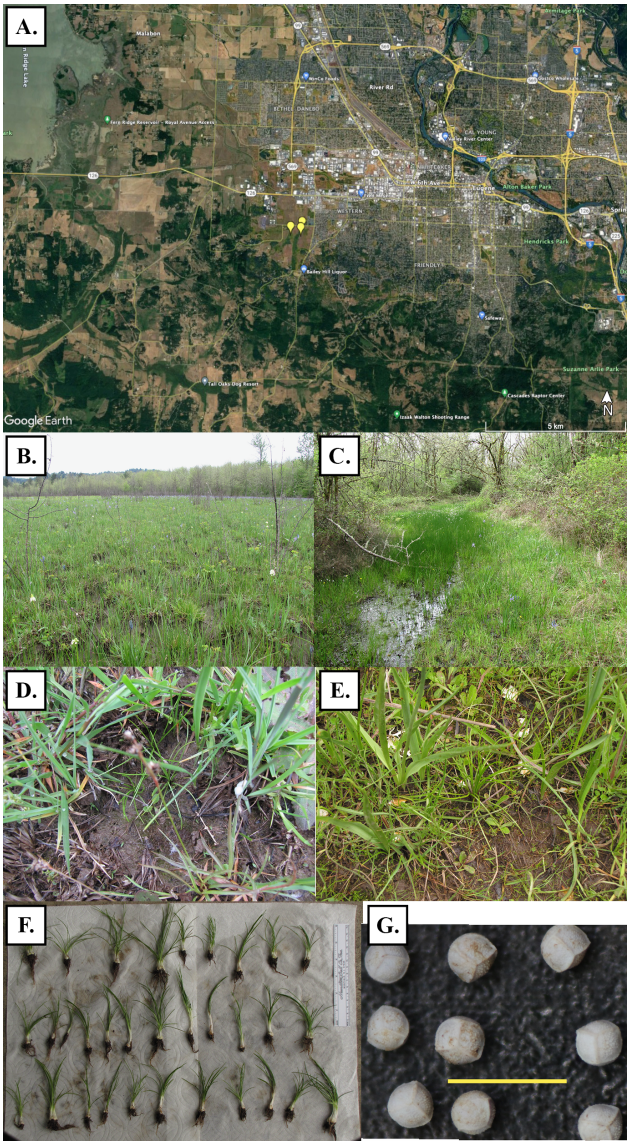

Figure 3.21: Willamette Valley Major Clade collection 2/3 - FF282. A. Locality of the collection site (Lat: 44.77682°, Long: -122.745669°); B. Wet meadow where some of the collections were made. This area had signs of recent burning, with the *Isoëtes* being found growing in the gaps between the grass hummocks; C. Flooded region off the main trail. Plants were found growing in and around the water.; D. Close up of the habitat in B where the plants were collected.; E. Close up of the habitat in C where the plants were collected.; F. Plants after collection and cleaning. Scale bar is in centimeters.; G. Merged images of spores. Yellow scale bar = 1mm.

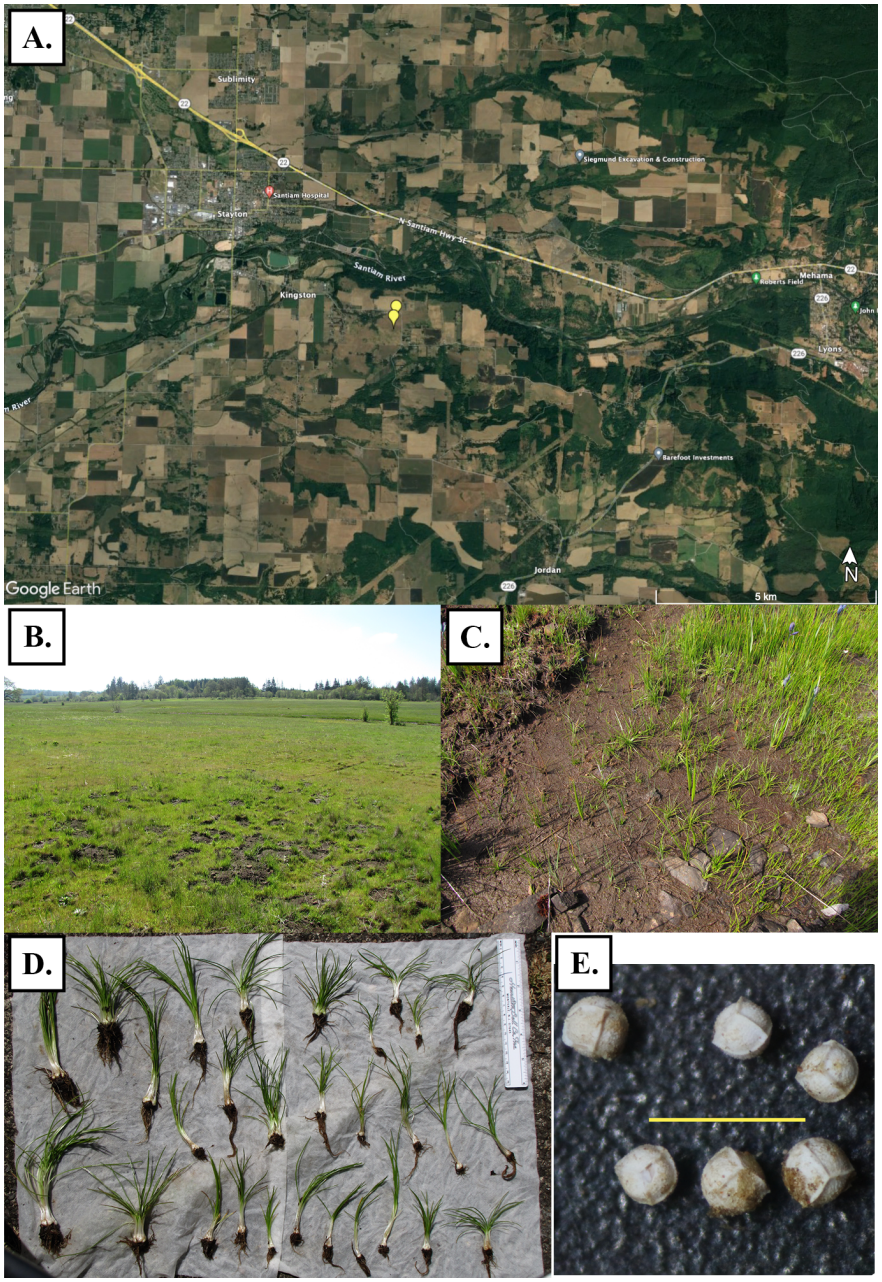

Figure 3.22: Willamette Valley Major Clade collection 3/3 - FF283. A. Locality of the collection site (Lat: 44.77682°, Long: -122.74567°); B. Image of the region of the meadow where the plants were collected. Plants were found throughout the meadow, including in the creek running through it; C. Close up of the habitat where the plants were collected. Appeared to be an overflow tributary of the creek.; D. Plants after collection and cleaning. Scale bar is in centimeters.; E. Merged images of spores. Yellow scale bar = 1mm.

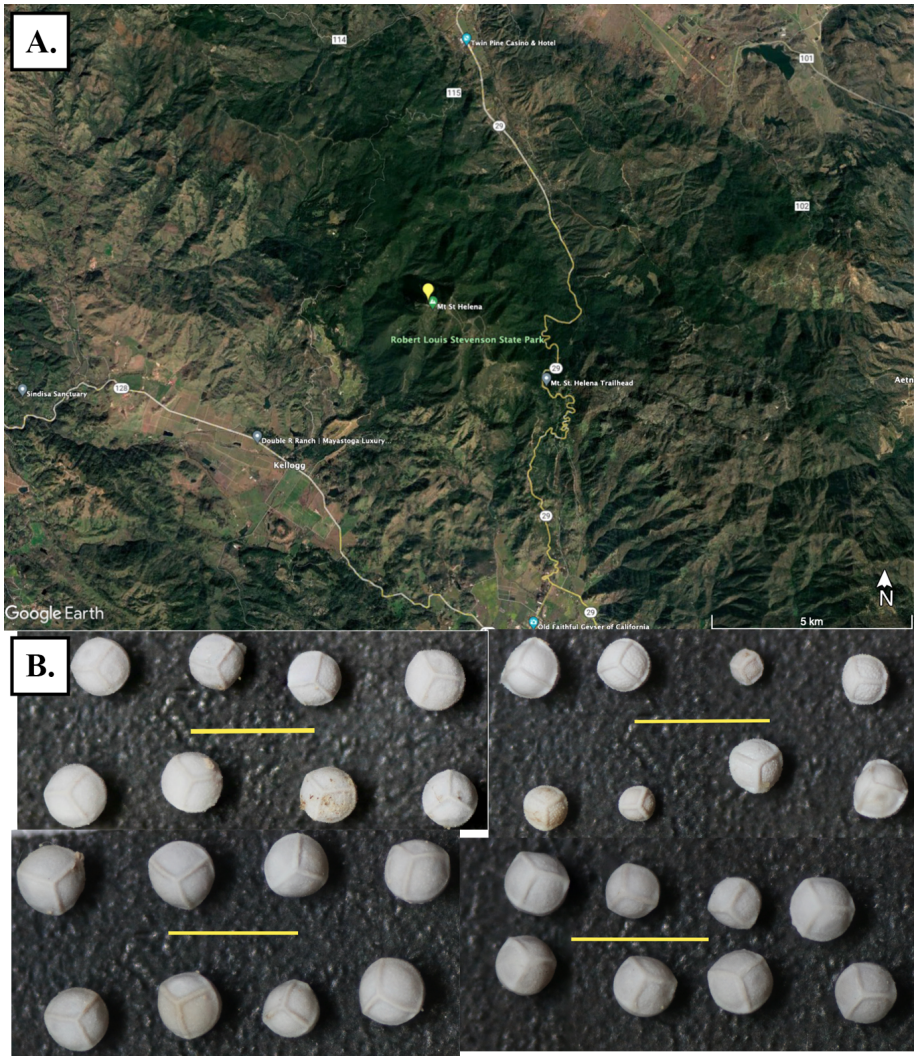

Figure 3.23: Rogue collection site 2/2 - CR5205. A. Locality of the collection site (Lat: 38.64959699°, Long: -122.579892°); B. Merged images of spores. Scale bars = 1mm in length.

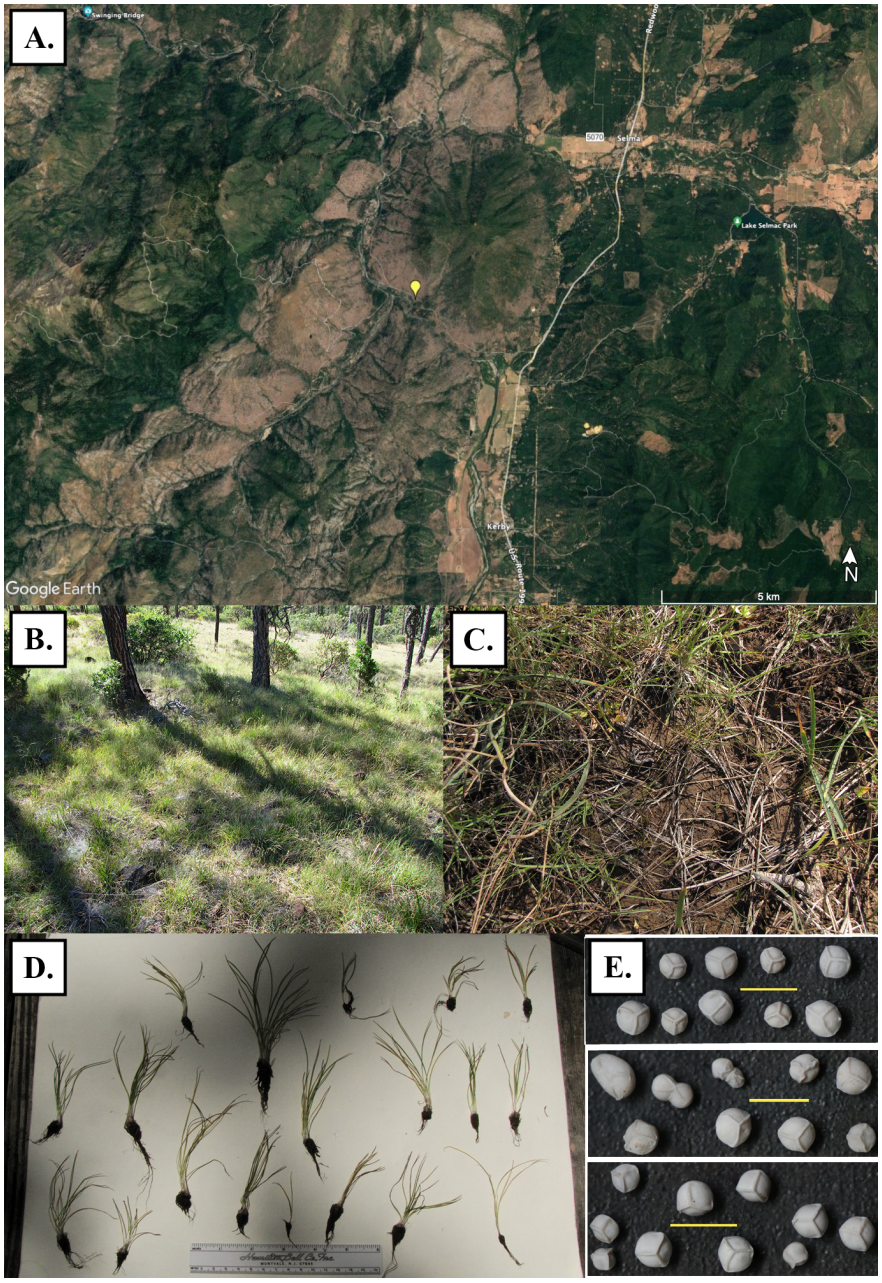

Figure 3.24: Rogue collection site 1/2 - FF290. A. Locality of the collection site (Lat: 42.242627°, Long: -123.678025°); B. Hillside seep where the plants were collected. Plants were found in a swale between the trees; C. Close up of the habitat where the plants were collected. Appeared to be an overflow tributary of the creek.; D. Plants after collection and cleaning. Scale bar is in centimeters.; E. Merged images of spores. Yellow scale bar = 1mm.

Table 3.3: MANOVA calculations for number of leaves, longest leaf, shortest leaf, and average leaf by Major Clade. Df = degrees of freedom, Sum Sq = Sum of Squares, Mean Sq = Mean Squared, F value = Variance between populations, Pr(>F) = the significance of F, Sig = degree of significance.

Response Number_of_leaves :						
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Sig
Group	7	4480.9	640.12	23.378	<2.2e-16	***
Residuals	639	17497	27.38			

Response longest_leaf_length :						
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Sig
Group	7	425481	60783	96.001	<2.2e-16	***
Residuals	639	404581	633			

Response shortest_leaf_length :						
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Sig
Group	7	73865	10552.2	41.125	<2.2e-16	***
Residuals	639	163958	256.6			

Response Average_leaf_length :						
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Sig
Group	7	237780	33969	90.304	<2.2e-16	***
Residuals	639	240365	376			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

based on leaf characters, but did differ significantly from the *I. nuttallii* morphotypes (Sierra Nevada, Coast Range and Willamette Valley). Interestingly, *I. minima* was significantly different from everything other than plants from the Sierra Nevada and Willamette Valley groups, while Coast Range differed from everything but Willamette Valley (Fig. 3.25–3.28, Tables 3.5–3.8).

For spores, there were few plants that had significantly different sized spores. The exceptions were Jepson Prairie, which had spores significantly smaller than every other collection site except Hog Lake and Minima, while Coast Range had spores that were significantly larger than any other Major Clade (Fig. 3.29–3.31, Tables 3.9–3.11).

PCA morphological clusters and loadings

For leaf qualities, the two components with the largest loading values were average leaf length and number of leaves (Table 3.12). Plants that belonged to the *I. orcuttii* and *I. minima*

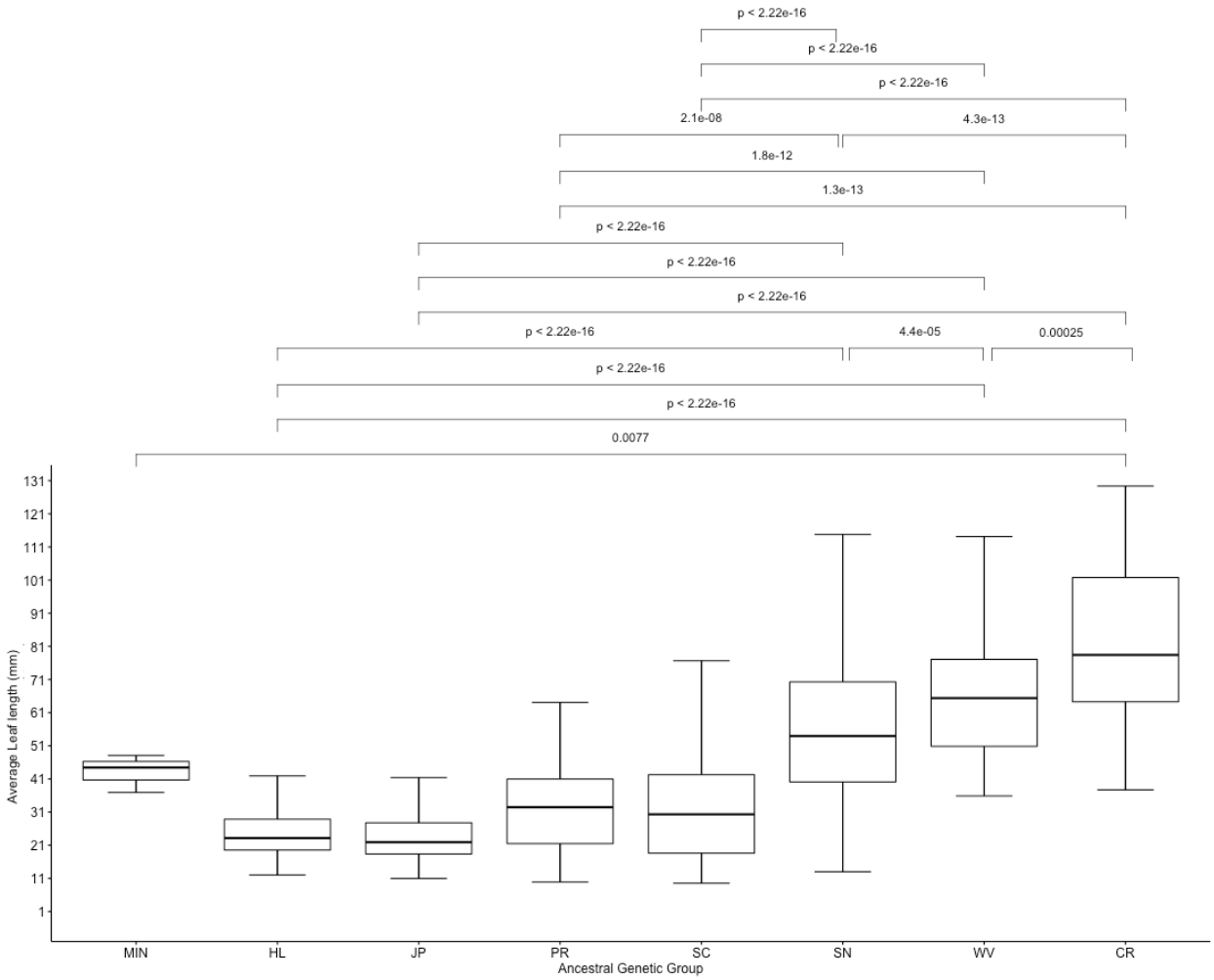

Figure 3.25: Box-plot of the range of values for average leaf size for each Major Clade. Pairwise comparison bars show Major Clade pairs that are statistically significantly different according to the Tukey-post hoc tests at the 95% confidence level, and the p-value for each pairwise comparison.

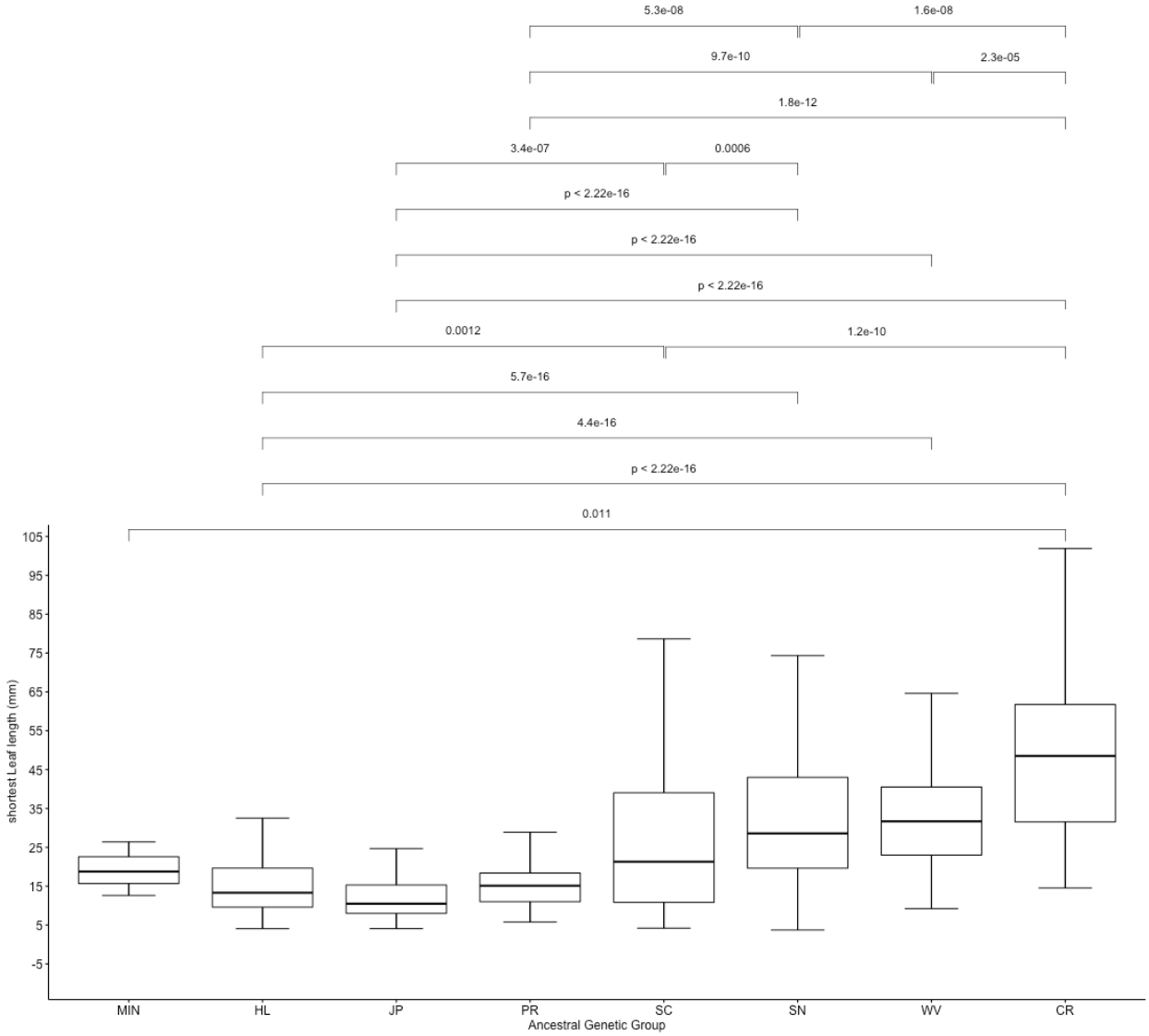

Figure 3.26: Box-plot of the range of values for shortest leaf size for each Major Clade. Pairwise comparison bars show Major Clade pairs that are statistically significantly different according to the Tukey-post hoc tests at the 95% confidence level, and the p-value for each pairwise comparison.

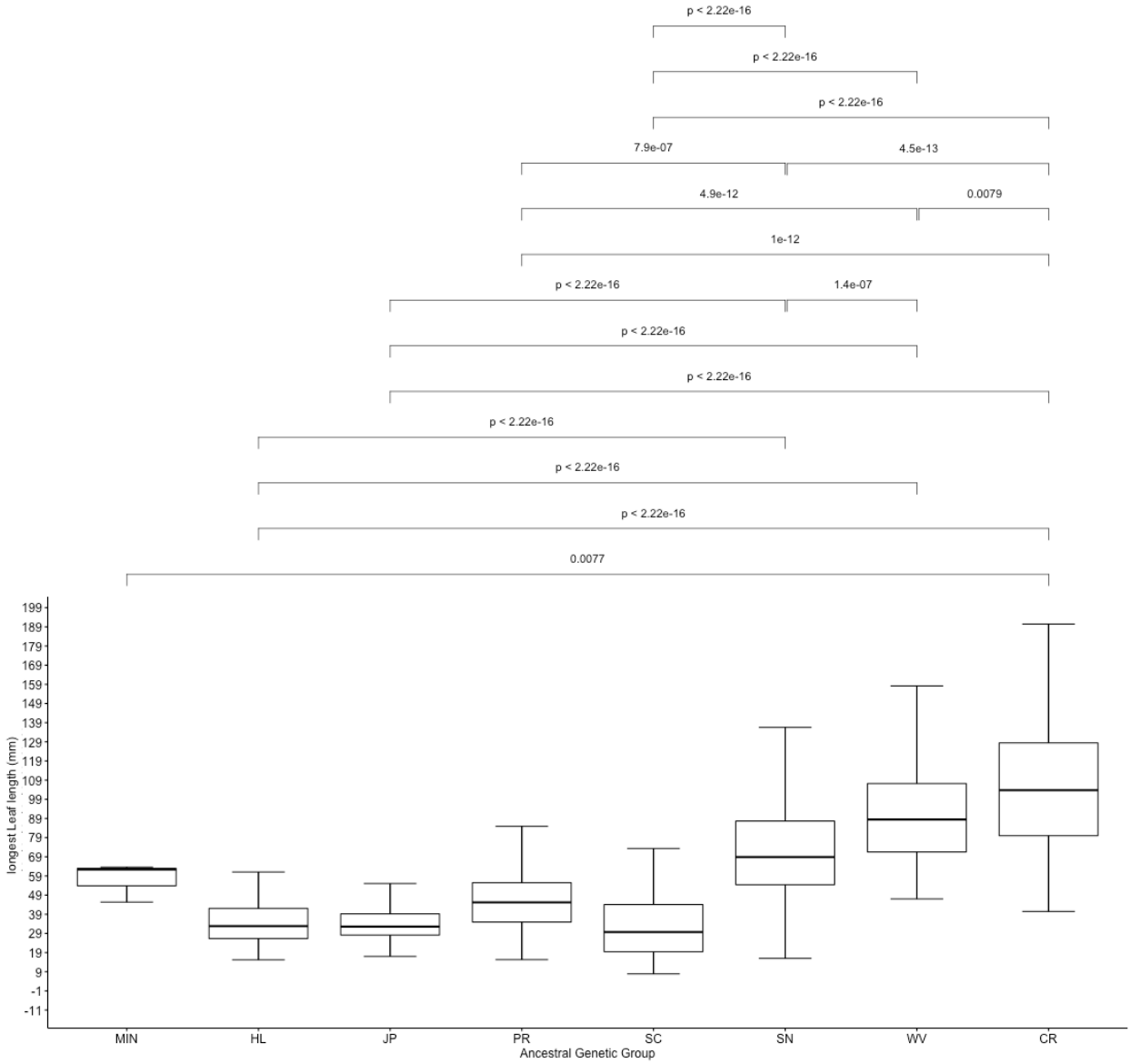

Figure 3.27: Box-plot of the range of values for longest leaf size for each Major Clade. Pairwise comparison bars show Major Clade pairs that are statistically significantly different according to the Tukey-post hoc tests at the 95% confidence level, and the p-value for each pairwise comparison.

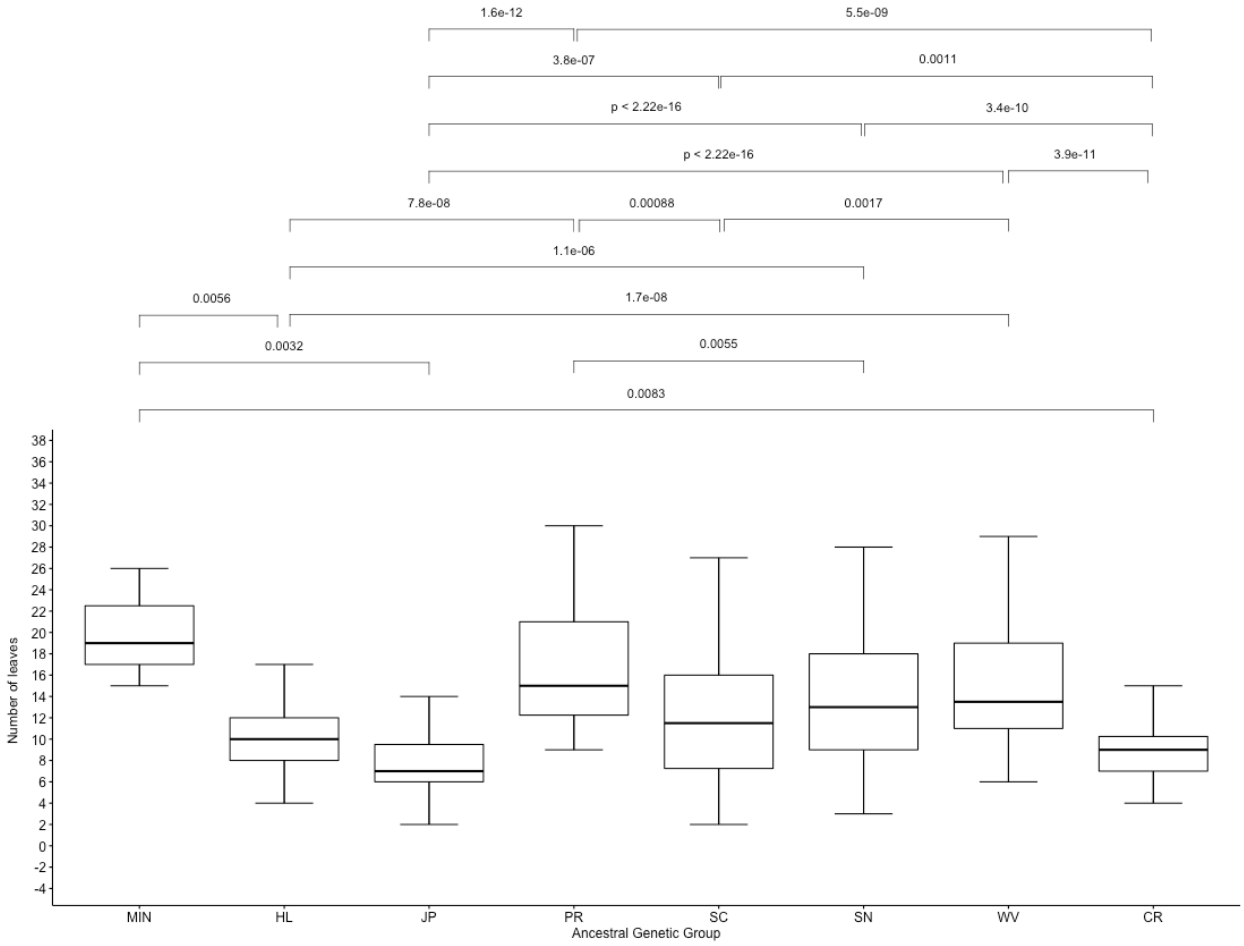

Figure 3.28: Box-plot of the range of values for longest leaf size for each Major Clade. Pairwise comparison bars show Major Clade pairs that are statistically significantly different according to the Tukey-post hoc tests at the 95% confidence level, and the p-value for each pairwise comparison.

Table 3.4: MANOVA calculations for average, smallest and largest spore sizes by Major Clade. Df = degrees of freedom, Sum Sq = Sum of Squares, Mean Sq = Mean Squared, F value = Variance between populations, Pr(>F) = the significance of F, Sig = degree of significance.

Response Average Spore Size:						
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Sig
Group	7	0.2032	0.029028	13.255	1.78E-11	***
Residuals	86	0.18834	0.00219			

Response Smallest Spore Size:						
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Sig
Group	7	0.16989	0.0242704	9.5902	9.28E-09	***
Residuals	86	0.21764	0.0025307			

Response Largest Spore Size:						
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Sig
Group	7	0.2354	0.033628	13.659	9.36E-12	***
Residuals	86	0.21173	0.002462			

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

morphotypes form a loose cluster together, while plants with the *I. nuttallii* morphotype formed another loose cluster (Fig. 3.32).

The two spore qualities that had the highest loading values were average size and smallest size. When the PCA was plotted for the first two components, there were three distinct clusters. The first included only one of the Major Clades: Jepson Prairie. The second included Hog Lake, Minima, Southern California, Peninsular Range and Sierra Nevada. The final cluster consisted of the Coast Range and Willamette Valley samples (Fig. 3.33).

3.4 Discussion

Taxonomic Re-evaluation in Light of the Placement of the Type Localities and Existing Species Descriptions

Initial identifications of the samples to the three described species (*I. orcuttii*, *I. minima* and *I. nuttallii*), did not return monophyletic groups, particularly when put into the context of the collections made near our best estimates for the type localities of *I. orcuttii* (FF285) and *I. nuttallii* (FF283).

The first and most obvious disagreement is the Hog Lake site. While consistent with *I. orcuttii* in its leaf morphology and habitat, the spores are distinctive both from the spores taken from the type specimen (Fig. 3.4) and the other plants identified as *I. orcuttii*, being

Table 3.5: Tukey post-hoc tests for pairwise comparisons of average leaf length.

	Group.diff	Group.lwr	Group.upr	Group.p.adj	significant
HL-JP	1.86	-8.10	11.82	1.00	
SC-JP	8.93	-0.07	17.94	0.05	
PR-JP	9.44	-3.20	22.09	0.31	
MIN-JP	19.29	-15.40	53.98	0.69	
SN-JP	32.66	24.98	40.33	0	*
WV-JP	44.64	35.03	54.25	0	*
CR-JP	58.67	49.06	68.28	0	*
SC-HL	7.07	-2.53	16.67	0.33	
PR-HL	7.58	-5.50	20.67	0.65	
MIN-HL	17.43	-17.43	52.28	0.80	
SN-HL	30.80	22.42	39.17	0	*
WV-HL	42.77	32.60	52.95	0	*
CR-HL	56.81	46.64	66.99	0	*
PR-SC	0.51	-11.86	12.88	1.00	
MIN-SC	10.36	-24.24	44.95	0.99	
SN-SC	23.72	16.52	30.93	0	*
WV-SC	35.70	26.46	44.94	0	*
CR-SC	49.74	40.50	58.98	0	*
MIN-PR	9.84	-25.87	45.56	0.99	
SN-PR	23.21	11.78	34.65	$3.33\text{x}10^{-08}$	*
WV-PR	35.19	22.37	48.01	0	*
CR-PR	49.23	36.41	62.05	0	*
SN-MIN	13.37	-20.90	47.64	0.94	
WV-MIN	25.35	-9.41	60.10	0.34	
CR-MIN	39.38	4.63	74.14	$1.40\text{x}10^{-02}$	*
WV-SN	11.98	4.03	19.93	$1.49\text{x}10^{-04}$	*
CR-SN	26.02	18.07	33.96	0	*
CR-WV	14.04	4.21	23.87	$4.33\text{x}10^{-04}$	*

far smoother, non-grey in color, and having sharper points on the junctions of their apical sutures and girdle. Combined with its placement in the phylogeny, and uniqueness in the admixture plot, I feel safe in concluding that the plants from the Hog Lake site do not correspond with *I. orcuttii* s.s., and represent an undescribed species (Fig. 3.5).

The Minima clade, which consists of the single *Isoetes minima* site, was the most consistent with its descriptions in the literature and my observations of the plants, especially in terms of their spores. In addition, they had 100% bootstrap support in phylogenetic placement in the SVDQ tree (Fig. 2.6). While their leaves were slightly larger than literature descriptions (Pfeiffer, 1922; Frye and Jackson, 1913; Eaton, 1898), *I. minima* plants can

Table 3.6: Tukey post-hoc tests for pairwise comparisons of longest leaf length.

	Group.diff	Group.lwr	Group.upr	Group.p.adj	significant
SC-JP	0.97	-10.71	12.65	1.00	
HL-JP	1.47	-11.45	14.40	1.00	
PR-JP	13.59	-2.82	30.00	0.19	
MIN-JP	22.92	-22.09	67.93	0.78	
SN-JP	38.78	28.82	48.74	0	*
WV-JP	59.04	46.57	71.51	0	*
CR-JP	73.61	61.14	86.08	0	*
HL-SC	0.50	-11.96	12.96	1.00	
PR-SC	12.61	-3.43	28.66	0.25	
MIN-SC	21.95	-22.93	66.82	0.81	
SN-SC	37.81	28.46	47.15	0	*
WV-SC	58.07	46.09	70.05	0	*
CR-SC	72.64	60.65	84.62	0	*
PR-HL	12.11	-4.86	29.09	0.37	
MIN-HL	21.45	-23.77	66.66	0.84	
SN-HL	37.31	26.45	48.17	0	*
WV-HL	57.57	44.37	70.77	0	*
CR-HL	72.14	58.94	85.34	0	*
MIN-PR	9.33	-37.00	55.67	1.00	
SN-PR	25.19	10.36	40.03	$8.95x10^{-06}$	*
WV-PR	45.46	28.83	62.08	0	*
CR-PR	60.03	43.40	76.65	0	*
SN-MIN	15.86	-28.60	60.32	0.96	
WV-MIN	36.12	-8.97	81.21	0.23	
CR-MIN	50.69	5.60	95.78	$1.53x10^{-02}$	*
WV-SN	20.26	9.95	30.57	$1.07x10^{-07}$	*
CR-SN	34.83	24.52	45.14	0	*
CR-WV	14.57	1.82	27.32	$1.27x10^{-02}$	*

grow larger than the commonly published values according to the Committee on the status of Endangered Wildlife in Canada Canada and Change (2020). However, as there was only a single site with two individuals included in the analyses, I do not believe that there is enough information to properly assess the morphology or taxonomy of this species at this time, especially considering that they are highly fragmented in their distribution, and are also quite rare. Further collections from other localities are needed to better characterize and support their status as a distinct species (Fig. 3.6).

Plants from the Jepson Prairie site closely conform to the morphology and habitat of *I. orcuttii*. And, in the admixture groupings from the Chapter 2 (Fig. 2.11), they fall out with

Table 3.7: Tukey post-hoc tests for pairwise comparisons of shortest leaf length.

	Group.diff	Group.lwr	Group.upr	Group.p.adj	significant
HL-JP	3.13	-5.89	12.14	0.97	
PR-JP	4.19	-7.25	15.64	0.95	
MIN-JP	6.89	-24.51	38.28	1.00	
SC-JP	13.61	5.46	21.75	1.36E-05	*
SN-JP	20.14	13.19	27.08	0	*
WV-JP	21.24	12.54	29.93	0	*
CR-JP	45.91	37.70	54.11	0	*
PR-HL	1.07	-10.77	12.91	1.00	
MIN-HL	3.76	-27.78	35.30	1.00	
SC-HL	10.48	1.79	19.17	6.42×10^{-3}	*
SN-HL	17.01	9.43	24.59	0	*
WV-HL	18.11	8.90	27.32	1.01×10^{-7}	*
CR-HL	42.78	34.03	51.53	0	*
MIN-PR	2.70	-29.62	35.02	1.00	
SC-PR	9.41	-1.78	20.61	0.17	
SN-PR	15.94	5.59	26.29	9.28×10^{-5}	*
WV-PR	17.04	5.45	28.64	2.48×10^{-4}	*
CR-PR	41.71	30.48	52.95	0	*
SC-MIN	6.72	-24.59	38.02	1.00	
SN-MIN	13.25	-17.77	44.26	0.90	
WV-MIN	14.35	-17.10	45.80	0.86	
CR-MIN	39.02	7.70	70.34	4.10×10^{-3}	*
SN-SC	6.53	0.01	13.05	4.93×10^{-2}	*
WV-SC	7.63	-0.73	15.99	0.10	
CR-SC	32.30	24.45	40.15	0	*
WV-SN	1.10	-6.09	8.29	1.00	
CR-SN	25.77	19.18	32.36	0	*
CR-WV	24.67	16.25	33.09	0	*

some of the other plants identified as *I. orcuttii*. However, their placement in the SVDQ phylogeny puts them as sister to the rest of the *I. orcuttii* plants, which are intermixed with plants identified as *I. nuttallii*. Morphologically, the Jepson Prairie plants are the smallest of the collected PLC plants overall (Figs. 3.12, 3.13). Much like the Hog Lake site, I believe that this may mean that Jepson Prairie represents an undescribed species (Fig. 3.7); I suspect that all plants that have been ascribed to *I. orcuttii* in the Central Valley may be part of this undescribed species, though more collections are needed to confirm this hypothesis.

The final part of the tree consists of the Southern California (Figs. 3.8-3.10), Peninsular Range (Fig. 3.11), Sierra Nevada (Figs. 3.12-3.16), Coast Range (Figs. 3.17-3.19)

Table 3.8: Tukey post-hoc tests for pairwise comparisons of leaf counts.

	Group.diff	Group.lwr	Group.upr	Group.p.adj	significant
CR-JP	1.44	-1.00	3.87	0.62	
HL-JP	2.24	-0.43	4.91	0.18	
SC-JP	4.60	2.18	7.01	$3.11\text{x}10^{-7}$	*
SN-JP	5.85	3.79	7.91	0	*
WV-JP	7.58	5.00	10.16	0	*
PR-JP	9.26	5.87	12.66	0	*
MIN-JP	12.06	2.76	21.37	$2.28\text{x}10^{-3}$	*
HL-CR	0.80	-1.79	3.39	0.98	
SC-CR	3.16	0.83	5.49	$1.08\text{x}10^{-3}$	*
SN-CR	4.41	2.46	6.37	0	*
WV-CR	6.14	3.64	8.64	0	*
PR-CR	7.83	4.50	11.16	0	*
MIN-CR	10.63	1.34	19.91	$1.24\text{x}10^{-2}$	*
SC-HL	2.36	-0.22	4.93	0.10	
SN-HL	3.61	1.37	5.86	$3.47\text{x}10^{-5}$	*
WV-HL	5.34	2.61	8.07	$1.23\text{x}10^{-7}$	*
PR-HL	7.03	3.52	10.54	$5.42\text{x}10^{-8}$	*
MIN-HL	9.83	0.47	19.18	$3.15\text{x}10^{-2}$	*
SN-SC	1.25	-0.68	3.19	0.50	
WV-SC	2.98	0.50	5.46	$6.62\text{x}10^{-3}$	*
PR-SC	4.67	1.35	7.99	$5.70\text{x}10^{-4}$	*
MIN-SC	7.47	-1.81	16.75	0.22	
WV-SN	1.73	-0.40	3.86	0.21	
PR-SN	3.41	0.35	6.48	$1.73\text{x}10^{-2}$	*
MIN-SN	6.21	-2.98	15.41	0.45	
PR-WV	1.69	-1.75	5.12	0.81	
MIN-WV	4.49	-4.84	13.81	0.83	
MIN-PR	2.80	-6.78	12.38	0.99	

and Willamette Valley (Figs. 3.20-3.22) plants. It is in this portion of the sampling that things become less clear in terms of the phylogenetic, pop-gen, morphological, and habitat characters of the plants. This portion of the phylogeny is also important because it includes the regions where the type localities for both *I. orcuttii* and *I. nuttallii* occur, the former being included in the Southern California Major Clade, while the latter is included in the Willamette Valley Major Clade. At my best estimate, the two sites that are the closest in proximity to the type locallities are FF285 (*I. orcuttii*), and FF283 (*I. nuttallii*).

In terms of leaf size, Southern California and Peninsular Range both fall within the published values for *I. orcuttii* (Table 3.1, 3.2), and do not significant differ from one another

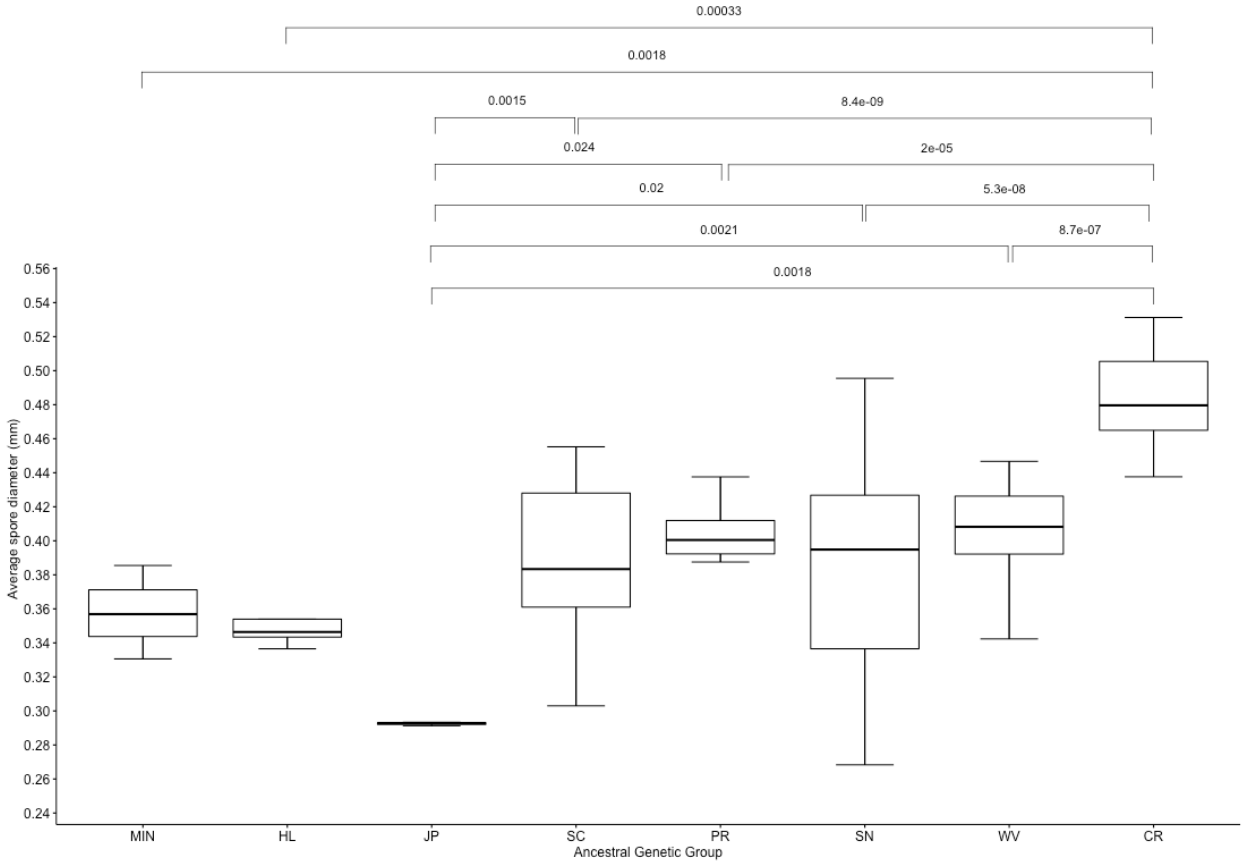

Figure 3.29: Box-plot of the range of values for the average diameter of the spores collected from each plant in each Major Clade. Pairwise comparison bars show Major Clade pairs that are statistically significantly different according to the Tukey-post hoc tests at the 95% confidence level, and the p-value for each pairwise comparison.

in any of their leaf length metrics (Tables 3.5–3.11, Figs. 3.34–3.28). *Isoëtes orcuttii*'s established leaf length range is 2–6.5cm, though this can reach up to 10cm (Baldwin et al., 2012; Taylor, 1993). The Southern California plants have an average leaf length of 3.7cm, with the longest ones reaching 11cm, which are within but slightly exceed the published values for the species, while the Penninsular Range plants have an average leaf length of 3.4cm with their longest leaves at 9.7cm. However, neither species has an average spore diameter that falls within the range of the published values for *I. orcuttii*, which are 216–360µm. Southern California plants have an average diameter of 410µm, and their smallest and largest spores averaging to 250µm and 530µm respectively. The Peninsula Range plants have an identical average of 410µm, while the averaged smallest and largest spores measuring 300µm and 480µm. In terms of spore ornamentation and coloration, both have light grey spores that turn dark-grey, almost black, when wet, which is consistent with *I. orcuttii*. As for spore ornamentation, both plants have tuberculate spores. For the Southern California

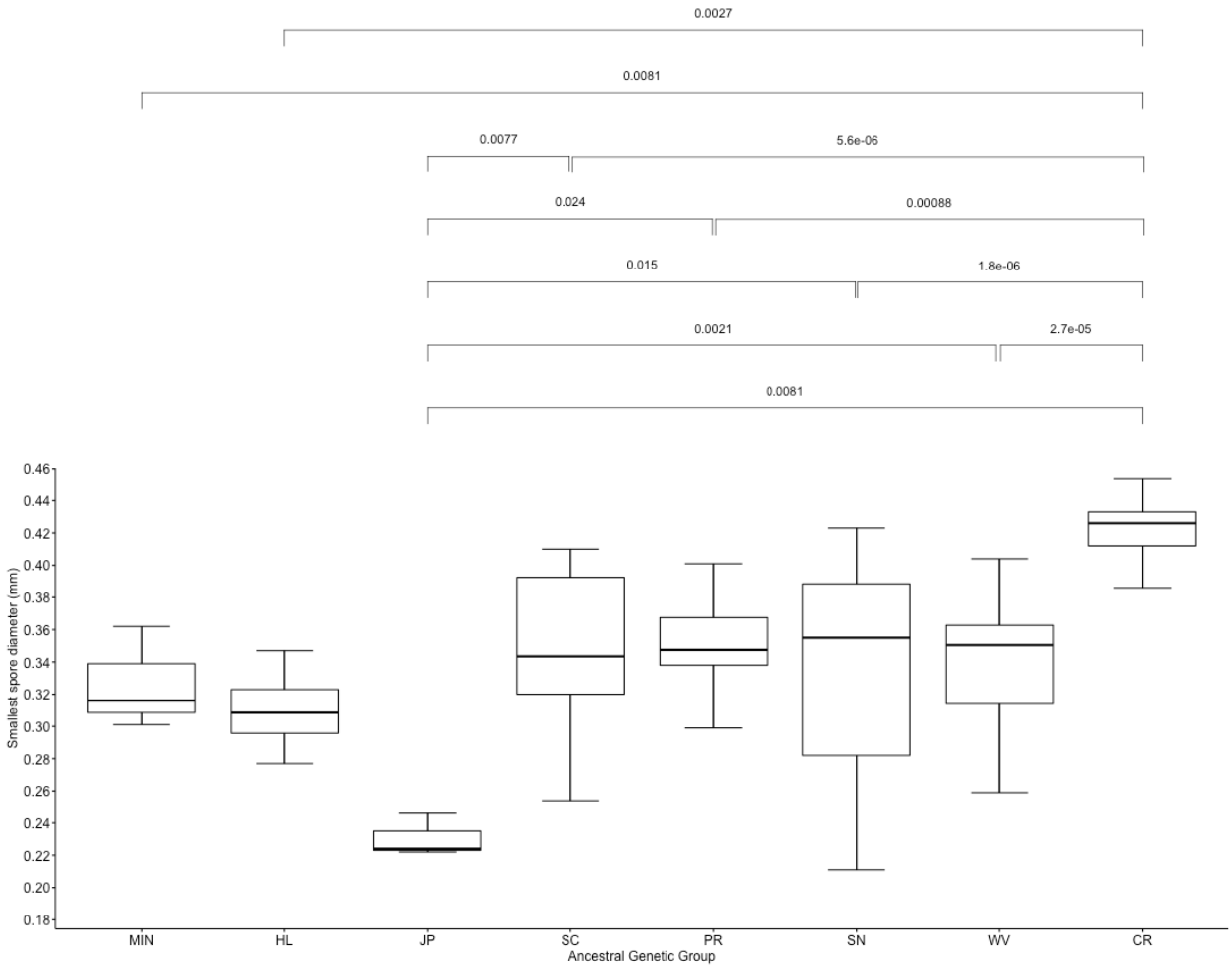

Figure 3.30: Box-plot of the range of values for the diameter of the smallest spore collected from each plant in each Major Clade. Pairwise comparison bars show Major Clade pairs that are statistically significantly different according to the Tukey-post hoc tests at the 95% confidence level, and the p-value for each pairwise comparison.

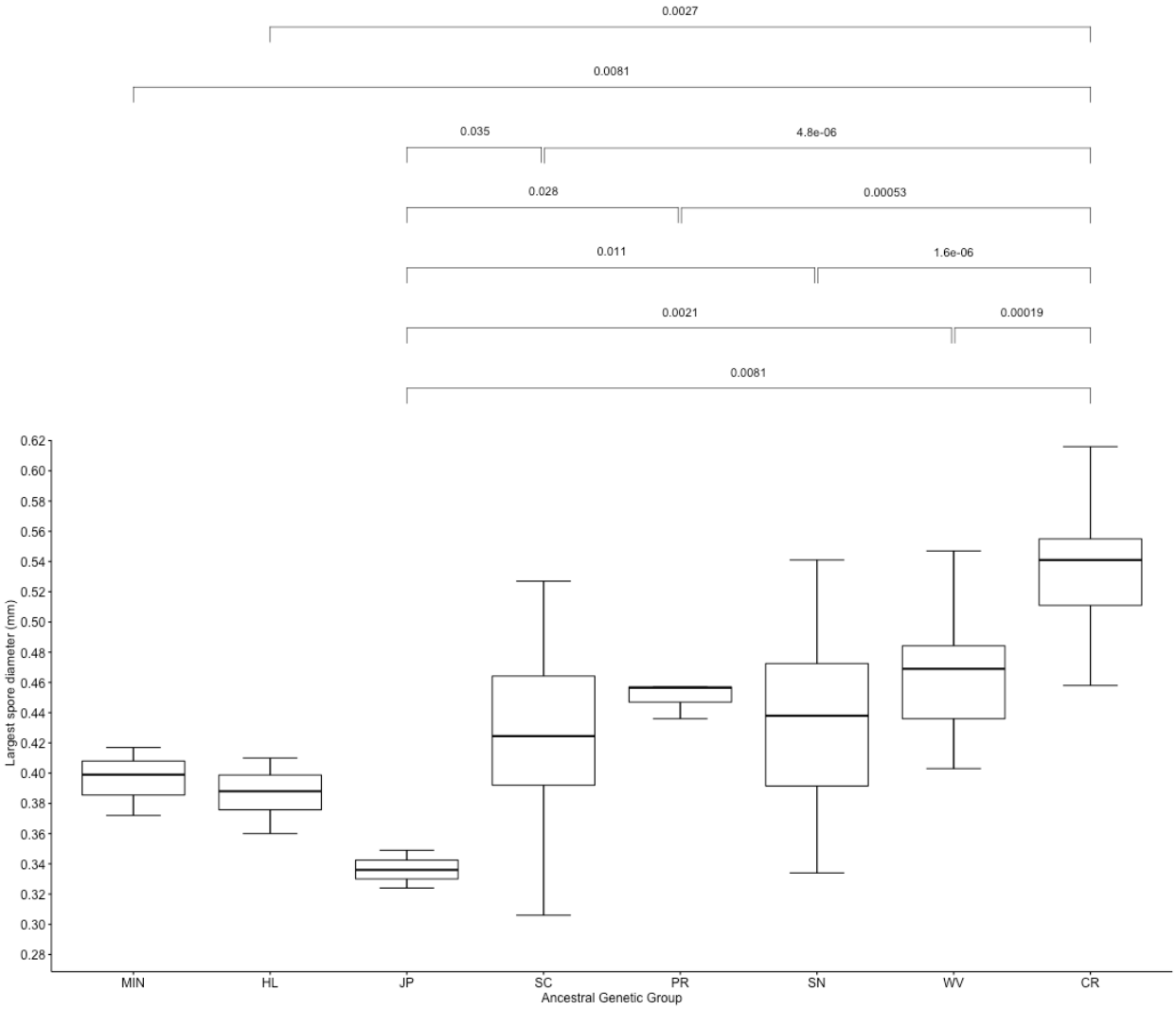

Figure 3.31: Box-plot of the range of values for the diameter of the largest spore collected from each plant in each Major Clade. Pairwise comparison bars show Major Clade pairs that are statistically significantly different according to the Tukey-post hoc tests at the 95% confidence level, and the p-value for each pairwise comparison.

Table 3.9: Tukey post-hoc tests for pairwise comparisons of average spore diameters.

	Group.diff	Group.lwr	Group.upr	Group.p.adj	significant
HL-JP	0.06	-0.05	0.17	0.73	
MIN-JP	0.07	-0.05	0.18	0.68	
SN-JP	0.09	0.00	0.18	3.39E-02	*
OR-JP	0.10	0.01	0.19	2.51E-02	*
PR-JP	0.11	0.01	0.22	2.14E-02	*
WV-JP	0.11	0.02	0.21	4.84E-03	*
CR-JP	0.20	0.11	0.29	2.97E-08	*
MIN-HL	0.01	-0.10	0.12	1.00	
SN-HL	0.03	-0.04	0.11	0.87	
OR-HL	0.04	-0.04	0.12	0.79	
PR-HL	0.05	-0.04	0.15	0.62	
WV-HL	0.06	-0.03	0.14	0.40	
CR-HL	0.14	0.06	0.22	1.19E-05	*
SN-MIN	0.03	-0.06	0.12	0.98	
OR-MIN	0.03	-0.06	0.12	0.95	
PR-MIN	0.05	-0.06	0.15	0.84	
WV-MIN	0.05	-0.04	0.14	0.71	
CR-MIN	0.14	0.04	0.23	3.41E-04	*
OR-SN	0.01	-0.04	0.05	1.00	
PR-SN	0.02	-0.05	0.09	0.98	
WV-SN	0.02	-0.02	0.07	0.82	
CR-SN	0.11	0.06	0.15	2.01E-09	*
PR-OR	0.01	-0.05	0.08	1.00	
WV-OR	0.02	-0.03	0.07	0.97	
CR-OR	0.10	0.05	0.15	1.43E-07	*
WV-PR	0.00	-0.07	0.07	1.00	
CR-PR	0.09	0.02	0.16	3.74E-03	*
CR-WV	0.09	0.04	0.14	2.35E-05	*

plants, these tubercles are subtle, which matches the descriptions in Baldwin et al. (2012) and Taylor (1993), as well as resembling the spores from the type specimen 3.4. Conversely, the Peninsular Range plant's spores are more obviously ornamented, with their tubercles becoming sub-echinate, and are far more pronounced than the SC plants. In a few of the specimens, they become sub-reticulate (Fig. 3.11). Yet, the most striking feature of the spores of the Peninsular Range plants would be their equitorial girdles, which are dentate like those of *I. minima* (Table 3.1, Fig. 3.6). Because this feature is considered so diagnostic of *I. minima*, it's presence in the PR clade plants could explain the records for *I. minima* in San Diego and Baja California included by Preiffer Pfeiffer (1922) in her monograph.

Table 3.10: Tukey post-hoc tests for pairwise comparisons of largest spore diameters.

	Group.diff	Group.lwr	Group.upr	Group.p.adj	significant
HL-JP	0.06	-0.05	0.17	0.73	
MIN-JP	0.07	-0.05	0.18	0.68	
SN-JP	0.09	0.00	0.18	$3.39x10^{-2}$	*
OR-JP	0.10	0.01	0.19	$2.51x10^{-2}$	*
PR-JP	0.11	0.01	0.22	$2.14x10^{-2}$	*
WV-JP	0.11	0.02	0.21	$4.84x10^{-3}$	*
CR-JP	0.20	0.11	0.29	$2.97x10^{-8}$	*
MIN-HL	0.01	-0.10	0.12	1.00	
SN-HL	0.03	-0.04	0.11	0.87	
OR-HL	0.04	-0.04	0.12	0.79	
PR-HL	0.05	-0.04	0.15	0.62	
WV-HL	0.06	-0.03	0.14	0.40	
CR-HL	0.14	0.06	0.22	$1.19x10^{-5}$	*
SN-MIN	0.03	-0.06	0.12	0.98	
OR-MIN	0.03	-0.06	0.12	0.95	
PR-MIN	0.05	-0.06	0.15	0.84	
WV-MIN	0.05	-0.04	0.14	0.71	
CR-MIN	0.14	0.04	0.23	$3.41x10^{-4}$	*
OR-SN	0.01	-0.04	0.05	1.00	
PR-SN	0.02	-0.05	0.09	0.98	
WV-SN	0.02	-0.02	0.07	0.82	
CR-SN	0.11	0.06	0.15	$2.01x10^{-9}$	*
PR-OR	0.01	-0.05	0.08	1.00	
WV-OR	0.02	-0.03	0.07	0.97	
CR-OR	0.10	0.05	0.15	$1.43x10^{-7}$	*
WV-PR	0.00	-0.07	0.07	1	
CR-PR	0.09	0.02	0.16	$3.74x10^{-3}$	*
CR-WV	0.09	0.04	0.14	$2.35x10^{-5}$	*

As for the rest of the Major Clades in this portion of the phylogeny, the Sierra Nevada, Coast Range, and Willamette Valley clades mostly conform to *I. nuttallii*, though each of the three differs significantly in terms of their leaf sizes, with SN being the smallest of the three, and CR being the largest (Tables. 3.5–3.11, Figs. 3.31–3.34). All three sites fall within the range of both the leaf lengths (8-17cm) and spore diameters (400-528µm) of *I. nuttallii*. However, the Sierra Nevada plants fall outside the lower end of these values, while the Coast Range has an upper end outside of said range (Table 3.1). Spore morphology is also within the broad description of *I. nuttallii*'s morphology: ± tuberculate and ± shiny. In the case of the Coast Range plants and Willamette Valley, their tuberculation tends to

Table 3.11: Tukey post-hoc tests for pairwise comparisons of smallest spore diameters.

	Group.diff	Group.lwr	Group.upr	Group.p.adj	significant
HL-JP	0.08	-0.04	0.20	0.44	
MIN-JP	0.10	-0.03	0.22	0.29	
SN-JP	0.11	0.01	0.20	$1.94\text{x}10^{-2}$	*
WV-JP	0.11	0.02	0.21	$1.27\text{x}10^{-2}$	*
PR-JP	0.12	0.01	0.23	$2.37\text{x}10^{-2}$	*
OR-JP	0.12	0.02	0.22	$5.57\text{x}10^{-3}$	*
CR-JP	0.20	0.10	0.30	$2.22\text{x}10^{-7}$	*
MIN-HL	0.02	-0.10	0.14	1.00	
SN-HL	0.03	-0.06	0.11	0.98	
WV-HL	0.03	-0.05	0.12	0.93	
PR-HL	0.04	-0.06	0.14	0.92	
OR-HL	0.04	-0.05	0.13	0.82	
CR-HL	0.12	0.04	0.21	$9.10\text{x}10^{-4}$	*
SN-MIN	0.01	-0.09	0.10	1.00	
WV-MIN	0.02	-0.08	0.12	1.00	
PR-MIN	0.02	-0.09	0.13	1.00	
OR-MIN	0.02	-0.07	0.12	0.99	
CR-MIN	0.11	0.01	0.20	$2.48\text{x}10^{-2}$	*
WV-SN	0.01	-0.04	0.06	1.00	
PR-SN	0.01	-0.06	0.09	1.00	
OR-SN	0.02	-0.03	0.06	0.98	
CR-SN	0.10	0.05	0.14	$5.93\text{x}10^{-7}$	*
PR-WV	0.01	-0.07	0.08	1.00	
OR-WV	0.01	-0.05	0.06	1.00	
CR-WV	0.09	0.03	0.14	$7.08\text{x}10^{-5}$	*
OR-PR	0.00	-0.07	0.07	1.00	
CR-PR	0.08	0.01	0.16	$2.13\text{x}10^{-2}$	*
CR-OR	0.08	0.03	0.13	$1.99\text{x}10^{-4}$	*

Table 3.12: Percentage contribution of each variable in the leaf principal component analysis across all variable dimensions.

	Dim.1	Dim.2	Dim.3	Dim.4
Number of leaves	6.10	90.48	3.37	$4.81\text{x}10^{-2}$
Longest leaf length	26.43	8.28	59.12	6.17
Shortest leaf length	32.04	0.23	34.08	33.65
Average leaf length	35.42	1.02	3.43	60.13

Figure 3.32: Leaf feature principal components analysis graph of the first two principle components by Major Clades. Ellipses encompas 95% of the measured values around the centroid.

Table 3.13: Percentage contribution of each variable in the spore principal component analysis across all variable dimensions.

	Dim.1	Dim.2	Dim.3
Average diameter	35.71	1.22	63.07
Smallest diameter	31.46	58.56	9.98
Largest diameter	32.82	40.22	26.95

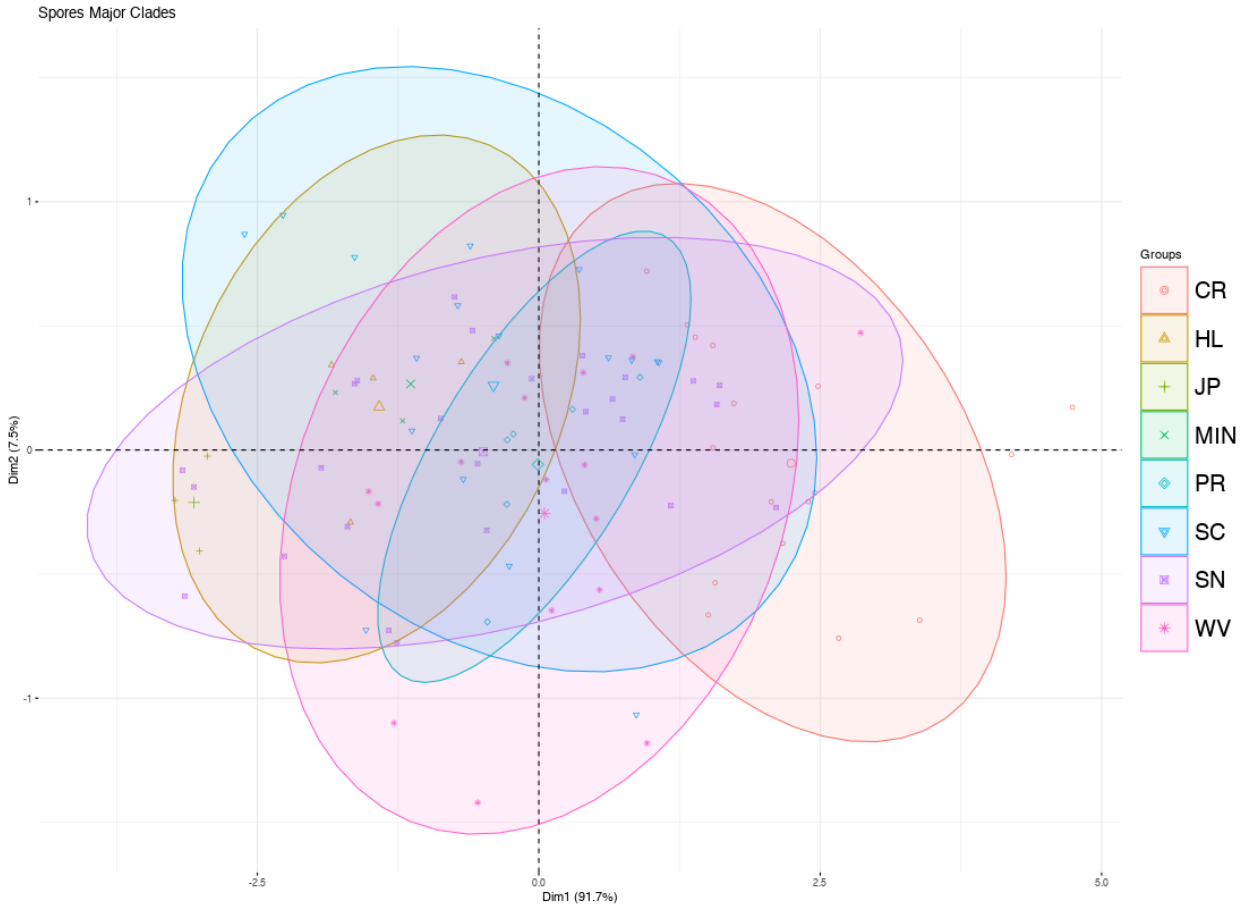

Figure 3.33: Spore feature principal components analysis graph of the first two principal components by Major Clades. Ellipses encompas 95% of the measured values around the centroid.

be more pronounced and dense than that of the Sierra Nevada plants (Figs. 3.12–3.22.

The SVDQ analysis and admixture plotting subdivides this portion of the PLC. In both the phylogeny and admixture plot, the Coast Range and Willamette Valley form a single unit, while Sierra Nevada and Peninsular Range form another; the SVDQ phylogeny further subdivides them with high support into their respective Major Clades, and mostly conform morphologically to *I. nuttallii* (Table 3.1). A third Admixture Group, comprised of SC and JP is also included in this portion of the SVDQ tree, but does not form a monophylletic clade. When examining the A00 analysis's inferred branch lengths on the SVDQ topology, (SC,SN,PR,CR,WV) effectively form a polytomy (Fig. 2.9). However, the length of the terminal branches leading to each of the Major Clades in the polytomy takes up around 20% of the total tree length, indicating that while their divergence from one another may have been sudden, they have been separated from one another for an appreciable period of time. Furthermore, this polytomy includes the Southern California plants, which were identified as *I. orcuttii)*, along with Jepson Prairie and Peninsular Range localities, which is

Figure 3.34: Tukey Post Hoc comparisons between Major Clades, with red boxes indicating a statistically significant differences between the pair. The top grid shows the pairwise leaf comparisons, while the bottom grid shows the pairwise spore comparisons.

non-monophyletic based on both the admixture plot and SVDQ phylogeny.

SC and JP do not differ significantly from one another in terms of their leaf sizes, though JP is still smaller than SC on average. However, JP does have statistically smaller spores than SC (Table 3.8–3.11, Figs. 3.34, 3.29–3.31). Conversely, SN and PR show significant differences in all of their leaf length metrics, but no significant differences in their spore sizes (Tables 3.5–3.8, Figs. 3.25–3.28). Yet, while there is no significant difference in the size of their spores, there are obvious differences in the ornamentation and color: SN spores are white to light grey when wet or dry, have tuberculate ornamentation and a smooth equitorial girdle, while the PR plants have spores that are light-grey when dry and dark grey when wet, with strongly tuberculate, almost echinate ornamentation, and smooth to dentate equatorial girdles (Table 3.2). Finally, CR and WV show significant differences in all of their leaf length, leaf number, and spore size metrics. Regarding their spores, both have tuberculate ornamentation, though the CR plants had a wide range of morphologies, going from smooth to very strongly tuberculate, while WV were moderately to strongly tuberculate. WV plants were also notable in that they occur not only terrestrially, but aquatically, with the largest plants found growing under water in what appear to be semi-permanent to permanent creeks, brooks, and woodland marshes. They were also distinct from the rest of the plants identified as *I. nuttallii* in that they did not have a complete velum. Instead, the velum was incomplete, and only covered 70% of the sporangia in the observed leaves. The only other species in the

PLC that have incomplete vela are *I. minima*. Outside the PLC, there are many species of *Isoëtes* with incomplete vela, including *I. howellii* and *I. bolanderi*. However, a complete velum is supposed to be diagnositic of *I. nuttallii*.

The two unusual sites – CR5205 (Fig. 3.23) and FF290 (Fig. 3.24) – had molecular and morphological characteristics suggesting that they were hybrids, and potentially allopolyploids. For the former, their placement in the phylogenies always had low support, and both showed signs of admixture between other genetic groups (See Chapter 2). Furthermore, when they were removed from the phylogeny, there was a marked increase in the bootstrap support values for not only the Coast Range and Willamette Valley clades, but for the placement of the Peninsular Range clade as sister to the Sierra Nevada clade. While on its own this could just indicated introgression between different populations, the aberrant spore morphology of these two populations is typically a sign of hybridization or polyploidy in *Isoëtes* (Barrington et al., 1986; Britton and Brunton, 1995; Pfeiffer, 1922; Bagella et al., 2011). While the spores of CR5205 are not especially deformed, they vary greatly in size, while the spores of FF290 are highly deformed, with malformed apical sutures among the least extreme of their deformities, with the more extreme resulting in incomplete fission of a pair of spores. Both types of deformity are common in hybrid *Isoëtes* (Musselman et al., 1996, 1997; Taylor et al., 1985) Both deformation and size disparity, coupled with their uncertain placement in the phylogeny, are good preliminary evidence that some or most of the population is an introgression or of hybrid origin between two distinct lineages in the PLC.

Recommended changes to Species

Given the results of the admixture plotting, phylogenetic analysis, morphological analysis and traits, as well as habitat, I believe that my initial hypothesis that the three existing species descriptions do not fully reflect the taxonomic diversity within the PLC is supported, and that there are in fact several cryptic lineages within the PLC. Given this evidence, I would argue that there are up to eight distinct lineages that could be recognized, depending on how the results are weighed.

The easiest case to argue is that of the Hog Lake site as distinctive from *I. orcutii*, as the weight of the evidence of it as distinctive is compelling. As seen in the previous chapter, it is distinct in both the admixture plot and in the phylogenetic analysis (Figs. 2.6, 2.7). Their spores are also morphologically distinctive from other plants initially identified as *I. orcuttii*. In fact, they are also distinct from the spores of the type specimen (Figs. 3.4 & 3.5), with one of the most distinct characters of the HL site being the sharp points at the junctions of their equitorial girdles and lack tubercles, as well as their spore coloration, which is white when dry, and bronzy when wet, as opposed to light grey and dark grey to black, respectively. In light of this, I believe that the plants in Hog Lake should be designated as a new species, and that further research be done on the site and surrounding region in order to better document this new species.

Isoëtes minima is a similar case to the Hog Lake plants, in that it is morphologically and phylogenetically distinct from the rest of the PLC. As such, I would consider this a good

taxon and see no immediate need for revision to its circumscription at this time. However, I believe that more sampling of this species from across its range is needed to properly define the species.

The Jepson Prairie site is the one that conforms most strongly to the published description of *I. orcutii* in terms of its size, habitat and morphology. In the admixture plot, the site also falls into the same population grouping as Southern California, the Major Clade that includes the plants collected from near the type locality for *I. orcuttii* (Fig. 2.11). However, Jepson Prairie does not form a monophyletic clade with Southern California (Fig. 2.8, . The JP plants also have significantly smaller spores than those of the Southern California group (Table 3.9 - 3.11), while their coloration and ornamentation is primarily differentiated by JP having glossy spores, while SC has duller spores with more obvious tubercles. Like the Hog Lake site, I believe that the Jepson Prairie site represents an undescribed, cryptic species. And, like the Hog Lake plants, I believe that further work is needed to fully identify the range and properties of this new taxa.

Where the designations become slightly less clear cut is in the (SC,PR,SN,CR,WV) portion of the phylogeny.

In the most conservative interpretation of this part of the phylogeny, all five would be considered a single species, as there is very short divergence times between the five Major Clades. If we were to accept this grouping, it would necessitate the sublimation of *I. orcuttii* into *I. nuttallii*, as the types for both species occur in this portion of the phylogeny and the latter has nomenclatural priority (McNeill et al., 2012). However, in doing so, we would have to ignore the differences in the size of the plants' spores and leaves (Tables 3.5–3.11), spore ornamentation (Table 3.2), and the results of the admixture plot (see Chapter 2). Furthermore, the SVDQuartets Sites and Individuals phylogenies shows that there is strong internal consistency for each of the Major Clades, with all members of a given Major Clade forming distinct units without any intermixing. In light of this, I do not believe there is justification for this lumping of all five Major Clades into a single taxon.

The next possibility is to break the taxonomic designations up based on the admixture plot, and denote SC as its own species, with (SN + PR) as one taxon, and (CR + WV) as another. However, like the above, the issue of the A00's terminal branch-lengths, and the clustering of the collection sites in the SVDQ analysis does not support this designation. Nor does the morphology. All five of the clades in question show significant differences in their spore and leaf sizes (Tables 3.5 - 3.11) and spore ornamentation (Table 3.2) between members of the Major Clade pairings. Specifically, PR is significantly smaller than SN in terms of leaf size, and WV is significantly smaller than CR, while both WV and CR are significantly larger than either PR or SN. Similarly, the ornamentation of the spores between these four Major Clades would not match up with that of the (SN + PR) (CR + WV). While CR and WV have similar ornamentation of their spores, they are significantly different in size (WV has smaller spores and leaves than CR). In addition to the size differences between WV and CR, there are several other traits that tare distinct between the two, and from *I. nuttallii* in general. The first and most obvious is that plants from the WV Major Clade can be found in aquatic habits, something that is not observed in any of the other

populations identified as *I. nuttallii*. All observed plants in California that fit the *I. nuttallii* morphospecies are terrestrial, or amphibious at most. The vela of the WV plants were also distinct from the other "*I. nuttallii*" specimens in that they were incomplete, and covered only 60-70% of the sporangia. This lower coverage is distinct from the rest of the plants in the collections other than *I. minima*, which all have complete vela (Pfeiffer, 1922; Engelmann, 1882; Eaton, 1898; Baldwin et al., 2012; Taylor, 1993). In regards to the possibility of an (SN + PR) grouping, the same arguments as above apply: significant differences in size and distinct spore ornamentation between the two Major Clades, and the strong support for their distinctiveness in all three of the SVDQ analyses. As such, I do not believe that this taxonomic designation is justified either.

This leaves one final designation: all Major Clades represent distinct species / evolutionary lineages. While this is the conclusion that I am in support of, I do not believe that at this time I have enough information to justify a full description of these new species. First and foremost, the phylogenetic trees infered by SVDQ strongly support the Major Clades as being distinct from one another. There are also statistically significant morphological differences between the Major Clades, as well as observable qualitative differences in their morphologies. All of the above, combined with the long terminal branches in the phylogenetic trees (Chapter 2), indicates that there is deep divergences between the Major Clades, even if there was rapid diversification between the clades themselves.

Future Work

I believe that there are several things that must occur to resolve the questions of taxonomy within the PLC.

First, we need to get representative accessions from the areas that were not fully sampled in this study, which include most of the Coast Range north and south of the bay area, central valley, and Sierra Nevada foothills. In addition, there needs to be further sampling in Oregon, Washington, southern British Columbia and Vancouver Island, as well as the Peninsular Range and Baja California (Fig. 3.35. These collections would need to provide good documentation of their population level morphology (both leaf and spore), and habitat, and comparable sequencing approaches as in this thesis in order to allow for integration with these results.

However, even with these preliminary results, it is clear that there is more taxonomic work to be done in the PLC. At the very least, the species designations as they stand do not appear to reflect the lineages within the group. *Isoëtes orcuttii* is obviously polyphyletic given the phylogenetic trees and popgen results. *Isoëtes nuttallii* appears somewhat more coherent, with the exception of the *I. orcuttii* clade mixed in with them. In addition, there are discernible subgroups within the clade that could represent varieties/sub-species, or possibly full species, depending on the degree of infertility between the genotypes, were it not for the possible paraphylly this would create due to the Peninsular Range Major Clade, identified as *I. orcuttii*, which resolves as sister to the Sierra Nevada Major Clade.

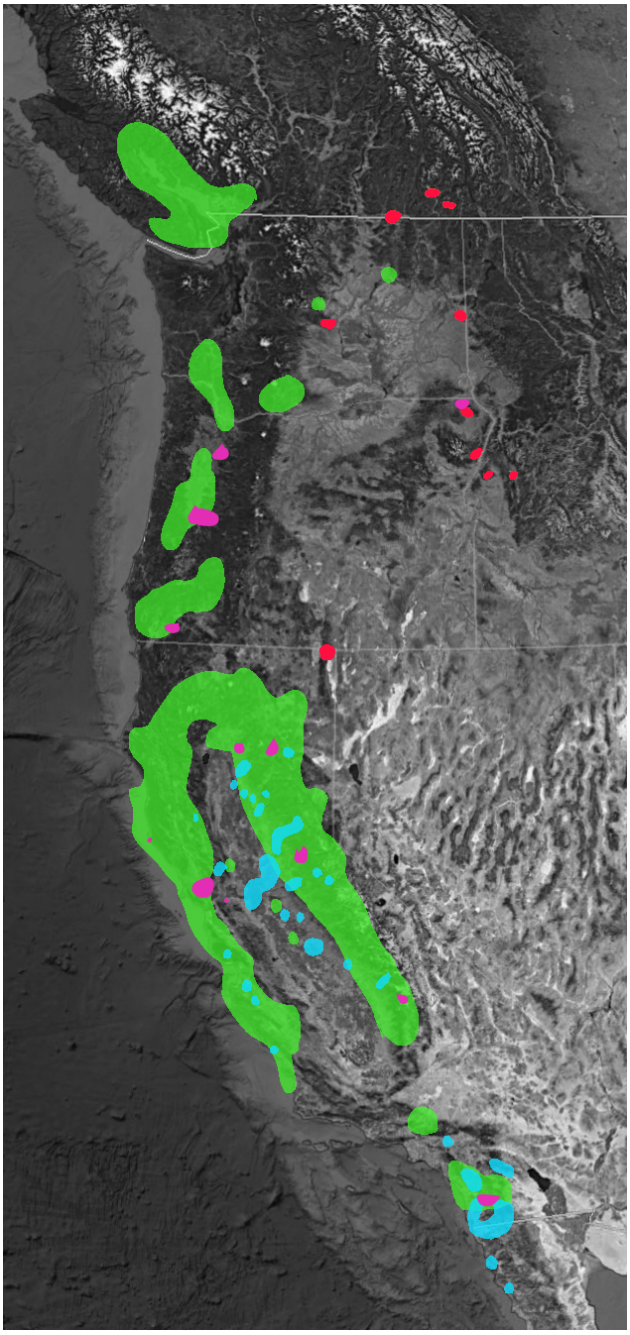

Figure 3.35: Known distribution of *Isoëtes* species from the Pacific Mediterranean Clade based on historical and contemporary collections. Range data gathered from Consortium of California Herbaria, Oregon State University and University of Washington. Green - *Isoëtes nuttallii*, Blue - *Isoëtes orcuttii*, Red - *Isoëtes minima*, Magenta - Regions sampled for this study.

There is also the question of the status of FF290 and CR5205. While the plants show signs that they may be hybrid or polyploid in nature, at this time it is unclear which they are, or if they represent something else entirely. In order to address this, I believe the first step should be to do introgression tests such as ABBA-BABA (Goulet et al., 2017; Martin et al., 2015; Springer and Gatesy, 2019) on the existing genomic dataset, with the expectation that FF290 represents either a hybrid between Coast Range and Willamette Valley, or (Coast Range, Willamette Valley) and Sierra Nevada. Given that the admixture plot (Chapter 2) shows FF290 with a 50/50 admixture identity between Groups 3 and 4, which correspond to the Sierra Nevada and Coast Range + Wilamette Valley groups, respectively, I would expect the FF290 site to be recovered as a hybrid of the Sierra Nevada and Willamette Valley clade. The Klamath Mountains, where FF290 was collected, is known for being a hotspot of hybridity Stone and Wolfe (2021); Ertter and Lewis (2016); Gugger et al. (2010); Patterson and Givnish (2004), particularly between species from the Cascades, which are adjacent to Willamette Valley, and the Sierra Nevada. If these two collections do show evidence of hybridity in the ABBA-BABA results, it could further evidence for treating one or more of the potentially new species identified above as distinct taxa.

As noted prior, *Isoëtes* are prone to forming hybrids (Taylor et al., 1985; Hoot et al., 2004; Musselman et al., 1996, 1997), indicating that the prezygotic barriers to hybrid formation are weak within the genus. Yet, hybrid individuals are almost always sterile due to not producing viable spores unless they undergo genome doubling to form polyploid gametophytes that will restore chromosome pairing if they mate with another, similarly polyploid spore (Dobzhansky, 1934; Boom, 1980; Soltis and Soltis, 1999; Taylor et al., 1985; Hickey et al., 1989). Because these hybrid offspring are sterile without a non-disjunction event during meiosis, I do not consider the presence of hybrids to indicated that the lineages are part of the same species. Rather, I consider them to be evidence for there being different species, because the hybrids do not act as a bridge for gene-flow between the two parental lines. Instead, they act as either a gate (sterile hybrids) or as the genesis point for a new, independent polyploid lineage (new allopolyploids) (Soltis et al., 2004, 2014; Mallet, 2007; Rieseberg and Willis, 2007).

3.5 Conclusions

I believe that I have shown that the taxonomic designations ascribed to *Isoëtes* species in the Pacific Laurasian Clade do not reflect their true evolutionary history.

The *I. orcuttii* morphospecies is the most obvious case of non-monophyly within the PLC. While they are not statistically different in their spore and leaf morphology, phylogenetically the morphospeices comes out as four distinct clades, which are intergraded with other identified species. Morphologically, there are few traits that can be used to identify the four different Major Clades: Hog Lake has smoother spores with sharp points on the junctions of their apical sutures, and are white/bronze when dry/wet, as opposed to grey/black. Jepson Prairie are overall slightly smaller than the others, and have darker spores, even when dry.

Southern California, as the clade that includes the type locality, are fairly consistent with the description, though there were some individuals found that were not in the typical vernal pool habitat that the plants are supposed to occupy. Peninsular Range plants have distinct spores that are more similar to *I. minima* due to their inconsistently dentate equatorial girdles and sub-echinate spore ornamentation.

Similarly, *I. nuttallii* is non-monphylletic, and includes three different Major Clades that I believe to represent distinct species: Sierra Nevada plants are significantly smaller than the other plants identified as *I. nuttallii*, in both spores and leaves. Willamette Valley plants are found both aquatically and terrestrially, and are the only plants that do not have a complete velum. Coast Range plants have the largest leaves and spores of the three. In addition to their morphological differences between the Major Clades, the presence of the putative hybrids, particularly FF290 adds weight to splitting the Sierra Nevada from (Willamette Valley + Coast Range). In addition to the extreme deformity of the spores, and it the site's uncertain placement in the phylogeny, the admixture plot shows a ∼50/50 split between Admixture Groups 3 and 4, which correspond to SN and (WV+CR), respectively. And, given the depth of the divergence time between the three, as well as the Penninsular Range and Southern California clades (Chapter 2), I believe that there is justification for splitting all of them into separate species. However, further research is needed to determine exactly where the splits belong.

Bibliography

Aars, J. and Ims, R. A. (2000). Population dynamic and genetic consequences of spatial density-dependent dispersal in patchy populations. *The American Naturalist*, 155(2):252–265.

Aguilar, R., Quesada, M., Ashworth, L., Herrerias-Diego, Y., and Lobo, J. (2008). Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Molecular ecology*, 17(24):5177–5188.

Amstutz, E. (1957). *Stylites*, a new genus of Isoetaceae. *Annals of the Missouri Botanical Garden*, 44(1):121–123.

Aradhya, M., Velasco, D., Ibrahimov, Z., Toktoraliev, B., Maghradze, D., Musayev, M., Bobokashvili, Z., and Preece, J. E. (2017). Genetic and ecological insights into glacial refugia of walnut (*Juglans regia* L.). *PloS one*, 12(10):e0185974.

Ash, S. R. and Pigg, K. B. (1991). A new Jurassic *Isoetites* (Isoetales) from the Wallowa Terrane in Hells Canyon, Oregon and Idaho. *American Journal of Botany*, 78(12):1636–1642.

Bagella, S., Caria, M. C., Molins, A., and Rosselló, J. A. (2011). Different spore structures in sympatric *Isoëtes histrix* populations and their relationship with gross morphology, chromosome number, and ribosomal nuclear ITS sequences. *Flora*, 206(5):451–457.

Bagella, S., Peruzzi, L., Caria, M. C., and Filigheddu, R. (2015). Unraveling the taxonomy and nomenclature of the *Isoëtes histrix* Bory species complex (Isoetaceae, Lycopodiidae). *Turkish Journal of Botany*, 39:383–387.

Baldwin, B. G., Goldman, D. H., Keil, D. J., Patterson, R., Rosatti, T. J., and Vorobik, L. A. (2012). *The Jepson manual: vascular plants of California*. Univ of California Press.

Barker, W., Nesom, G., Beardsley, P. M., and Fraga, N. S. (2012). A taxonomic conspectus of Phrymaceae: a narrowed circumscription for *Mimulus*, new and resurrected genera, and new names and combinations. *Phytoneuron*, 39:1–60.

Barrington, D. S., Paris, C. A., and Ranker, T. A. (1986). Systematic inferences from spore and stomate size in the ferns. *American Fern Journal*, 76(3):149–159.

Bauder, E. T. and McMillan, S. (1998). Current distribution and historical extent of vernal pools in southern California and northern Baja California, Mexico. In *Ecology, Conservation and Management of Vernal Pool Ecosystems–Proceedings from a 1996 Conference. California Native Plant Society, Sacramento, CA*, pages 56–70.

Bawa, K. S. (1980). Evolution of dioecy in flowering plants. *Annual review of ecology and systematics*, 11(1):15–39.

Bell, C. D., Soltis, D. E., and Soltis, P. S. (2010). The age and diversification of the angiosperms re-revisited. *American journal of botany*, 97(8):1296–1303.

Belletti, P., Monteleone, I., and Ferrazzini, D. (2008). A population genetic study in a scattered forest species, wild service tree [*Sorbus torminalis* (L.) Crantz], using RAPD markers. *European Journal of Forest Research*, 127(2):103–114.

Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., et al. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *nature*, 456(7218):53–59.

Bhambie, S. (1963). Studies in pteridophytes. *Proceedings of the National Academy of Sciences, India Section B*, 58(3):153–164.

Birkeland, P. W. (1964). Pleistocene glaciation of the northern Sierra Nevada, north of Lake Tahoe, California. *The Journal of Geology*, 72(6):810–825.

Bollmer, J. L., Whiteman, N. K., Cannon, M. D., Bednarz, J. C., Vries, T. d., and Parker, P. G. (2005). Population genetics of the Galápagos hawk (*Buteo galapagoensis*): genetic monomorphism within isolated populations. *The Auk*, 122(4):1210–1224.

Boom, B. M. (1980). Intersectional hybrids in *Isoëtes*. *American Fern Journal*, pages 1–4.

Bovee, K. M., Merriam, K. E., and Gosejohan, M. C. (2018). Livestock grazing affects vernal pool specialists more than habitat generalists in montane vernal pools. *Applied Vegetation Science*, 21(1):12–20.

Bowerman, N. D. and Clark, D. H. (2011). Holocene glaciation of the central Sierra Nevada, California. *Quaternary Science Reviews*, 30(9-10):1067–1085.

Bray, R. D., Schafran, P. W., and Musselman, L. J. (2018). Interesting, Provocative, and Enigmatic: Morphological Observations on Southeastern Quillworts (*Isoëtes* Isoetaceae, Lycopodiophyta). *Castanea*, 83(2):263–269.

Breinholt, J. W., Carey, S. B., Tiley, G. P., Davis, E. C., Endara, L., McDaniel, S. F., Neves, L. G., Sessa, E. B., von Konrat, M., Chantanaorrapint, S., et al. (2021). A target enrichment probe set for resolving the flagellate land plant tree of life. *Applications in plant sciences*, 9(1):e11406.

Britton, D. M. and Brunton, D. F. (1995). *Isoëtes*× *marensis*, a new interspecific hybrid from western Canada. *Canadian Journal of Botany*, 73(9):1345–1353.

Brunton, D. F. and Britton, D. M. (1997). Appalachian quillwort (*Isoetes appalachiana*, sp. nov.; Isoetaceae), a new pteridophyte from the eastern United States. *Rhodora*, pages 118–133.

Budke, J. M., Hickey, R. J., and Heafner, K. D. (2005). Analysis of morphological and anatomical characteristics of *Isoëtes* using *Isoëtes tennesseensis*. *Brittonia*, 57(2):167–182.

Canada, E. and Change, C. (2020). Government of Canada.

Cantino, P. D., De Queiroz, K., et al. (2020). *PhyloCode: a phylogenetic code of biological nomenclature*. CRC Press.

Carlsen, T., Engh, I. B., Decock, C., Rajchenberg, M., and Kauserud, H. (2011). Multiple cryptic species with divergent substrate affinities in the *Serpula himantioides* species complex. *Fungal Biology*, 115(1):54–61.

Carpenter, E. J., Matasci, N., Ayyampalayam, S., Wu, S., Sun, J., Yu, J., Jimenez Vieira, F. R., Bowler, C., Dorrell, R. G., Gitzendanner, M. A., et al. (2019). Access to RNA-sequencing data from 1,173 plant species: The 1000 Plant transcriptomes initiative (1KP). *GigaScience*, 8(10):giz126.

Cellinese, N., Baum, D. A., and Mishler, B. D. (2012). Species and phylogenetic nomenclature. *Systematic Biology*, 61(5):885–891.

Cesca, G. and Peruzzi, L. (2001). *Isoëtes* (Lycophytina, Isoetaceae) with terrestrial habitat in Calabria (Italy). New karyological and taxonomical data. *Flora Mediterranea*, 11:303–309.

Chen, Y.-Y., Liao, L., Li, W., and Li, Z.-Z. (2010). Genetic diversity and population structure of the endangered alpine quillwort *Isoëtes hypsophila* Hand.- Mazz. revealed by AFLP markers. *Plant Systematics and Evolution*, 290(1):127–139.

Chiang, S.-H. T. (1976). Growth cycle of cambium and the structure of the vascular tissue in the corm of *Isoëtes taiwanensis*. *Taiwania*.

Chifman, J. and Kubatko, L. (2014). Quartet inference from SNP data under the coalescent model. *Bioinformatics*, 30(23):3317–3324.

Chifman, J. and Kubatko, L. (2015). Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes, site-specific rate variation, and invariable sites. *Journal of theoretical biology*, 374:35–47.

Choi, H.-K., Jung, J., and Kim, C. (2008). Two new species of *Isoëtes* (Isoetaceae) from Jeju Island, South Korea. *Journal of Plant Biology*, 51(5):354–358.

Chung, M. Y., López-Pujol, J., and Chung, M. G. (2017). The role of the Baekdudaegan (Korean Peninsula) as a major glacial refugium for plant species: A priority for conservation. *Biological Conservation*, 206:236–248.

Collins, T. J. (2007). ImageJ for microscopy. *Biotechniques*, 43(S1):S25–S30.

Cook, C. D. K. (2004). Aquatic and wetland plants of southern Africa.

Cox, P. A. and Hickey, R. J. (1984). Convergent megaspore evolution and *Isoëtes*. *American Naturalist*, 124(3):437–441.

Croft, J. (1980). A taxonomic revision of *Isoëtes* L.(Isoetaceae) in Papuasia. *Blumea: Biodiversity, Evolution and Biogeography of Plants*, 26(1):177–190.

Crouch, N. R. (2012). *Ferns of Southern Africa: a comprehensive guide*. Penguin Random House South Africa.

Darwin, C. (1859). *The origin of species by means of natural selection*. Pub One Info.

Darwin, C., Wallace, A. R., et al. (1958). Evolution by natural selection. *Evolution by natural selection.*

Darwin, C., Wallace, A. R., Lyell, S. C., and Hooker, J. D. (1858). On the tendency of species to form varieties: and on the perpetuation of varieties and species by natural means of selection. Linnean Society of London.

Dauphin, B., Vieu, J., and Grant, J. R. (2014). Molecular phylogenetics supports widespread cryptic species in moonworts (*Botrychium* s.s., Ophioglossaceae). *American Journal of Botany*, 101(1):128–140.

Davis, C. C. and Xi, Z. (2015). Horizontal gene transfer in parasitic plants. *Current Opinion in Plant Biology*, 26:14–19.

de Queiroz, K. (2005). A unified concept of species and its consequences for the future of taxonomy. *Proceedings of the California Academy of Sciences*.

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic biology*, 56(6):879–886.

de Vos, J. M., Keller, B., Isham, S. T., Kelso, S., and Conti, E. (2012). Reproductive implications of herkogamy in homostylous primroses: variation during anthesis and reproductive assurance in alpine environments. *Functional Ecology*, 26(4):854–865.

Degnan, J. H. and Rosenberg, N. A. (2006). Discordance of species trees with their most likely gene trees. *PLoS genetics*, 2(5):e68.

Dobzhansky, T. (1934). Studies on hybrid sterility. *Zeitschrift für Zellforschung und mikroskopische Anatomie*, 21(2):169–223.

Dumolin-Lapegue, S., Demesure, B., Fineschi, S., Le Come, V., and Petit, R. (1997). Phylogeographic structure of white oaks throughout the European continent. *Genetics*, 146(4):1475–1487.

Eaton, A. A. (1898). *Isoëtes minima*. *Fern Bull*, 6:30.

Eckert, A. J., Tearse, B. R., and Hall, B. D. (2008). A phylogeographical analysis of the range disjunction for foxtail pine (*Pinus balfouriana*, Pinaceae): the role of Pleistocene glaciation. *Molecular Ecology*, 17(8):1983–1997.

Edgar, R. C. (2004). Muscle: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5):1792–1797.

Elam, D. (1998). Population genetics of vernal pool plants: theory, data and conservation implications. In Ornduff, R., Witham, C. W., Bauder, E. T., Belk, D., and Ferren Jr, W. R., editors, *Ecology, Conservation, and Management of Vernal Pool Ecosystems Proceedings from a Conference.*, pages 180–189, Sacramento, CA. California Native Plants Society.

Engelmann, G. (1882). *The genus Isoëtes in North America*. éditeur non identifié.

Ennos, R. (1994). Estimating the relative rates of pollen and seed migration among plant populations. *Heredity*, 72(3):250–259.

Eriksson, J., Blanco-Pastor, J., Sousa, F., Bertrand, Y., and Pfeil, B. (2017). A cryptic species produced by autopolyploidy and subsequent introgression involving *Medicago prostrata* (Fabaceae). *Molecular Phylogenetics and Evolution*, 107:367–381.

Ertter, B. and Lewis, W. H. (2016). Relationships, infrataxa, and hybrids of *Rosa gymnocarpa* (Rosaceae). *Madroño*, 63(3):268–280.

Feliner, G. N. and Rosselló, J. A. (2007). Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular phylogenetics and evolution*, 44(2):911–919.

Finger, A. J., Parmenter, S., and May, B. P. (2013). Conservation of the Owens pupfish: genetic effects of multiple translocations and extirpations. *Transactions of the American Fisheries Society*, 142(5):1430–1443.

Fraga, N. S. (2012). A revision of *Erythranthe montioides* and *Erythranthe palmeri* (Phrymaceae), with descriptions of five new species from California and Nevada, USA. *Aliso: A Journal of Systematic and Evolutionary Botany*, 30(1):49–68.

Freeland, J. R., Biss, P., and Silvertown, J. (2012). Contrasting patterns of pollen and seed flow influence the spatial genetic structure of sweet vernal grass (*Anthoxanthum odoratum*) populations. *Journal of Heredity*, 103(1):28–35.

Freund, F. D. (2016). Characterizing quantitative variation in the glossopodia of three western North American *Isoëtes* species. *American Fern Journal*, 106(2):87–115.

Freund, F. D., Freyman, W. A., and Rothfels, C. J. (2018). Inferring the evolutionary reduction of corm lobation in *Isoëtes* using Bayesian model-averaged ancestral state reconstruction. *American Journal of Botany*, 105(2):275–286.

Freyman, W. A. and Höhna, S. (2018). Cladogenetic and anagenetic models of chromosome number evolution: a Bayesian model averaging approach. *Systematic biology*, 67(2):195–215.

Frye, T. C. and Jackson, M. M. (1913). The Ferns of Washington. *American Fern Journal*, 3(3):65–83.

Garnock-Jones, P. (1993). Phylogeny of the *Hebe* complex (Scrophulariaceae: Veroniceae). *Australian systematic botany*, 6(5):457–479.

Gaudeul, M., Taberlet, P., and Till-Bottraud, I. (2000). Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L.(Apiaceae), inferred from amplified fragment length polymorphism markers. *Molecular ecology*, 9(10):1625–1637.

Gehring, J. L. and Delph, L. F. (1999). Fine-scale genetic structure and clinal variation in *Silene acaulis* despite high gene flow. *Heredity*, 82(6):628–637.

Geml, J., Laursen, G., O'NEILL, K., Nusbaum, H. C., and Taylor, D. L. (2006). Beringian origins and cryptic speciation events in the fly agaric (*Amanita muscaria*). *Molecular Ecology*, 15(1):225–239.

Gensel, P. G. and Berry, C. M. (2001). Early lycophyte evolution. *American Fern Journal*, 91(3):74–98.

Gensel, P. G. and Pigg, K. B. (2010). An arborescent lycopsid from the Lower Carboniferous Price Formation, southwestern Virginia, USA and the problem of species delimitation. *International Journal of Coal Geology*, 83(2-3):132–145.

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E. M., Brockman, W., Fennell, T., Giannoukos, G., Fisher, S., Russ, C., et al. (2009). Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature biotechnology*, 27(2):182–189.

Gordon, S. P., Sloop, C. M., Davis, H. G., and Cushman, J. H. (2012). Population genetic diversity and structure of two rare vernal pool grasses in central California. *Conservation Genetics*, 13(1):117–130.

Goulet, B. E., Roda, F., and Hopkins, R. (2017). Hybridization in plants: old ideas, new techniques. *Plant physiology*, 173(1):65–78.

Grauvogel-Stamm, L. and Lugardon, B. (2001). The Triassic Lycopsids *Pleuromeia* and *Annalepis* Relationships, Evolution, and Origin. *American Fern Journal*, 91(3):115–149.

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4):711–732.

Grivet, D., ROBLEDO-ARNUNCIO, J. J., Smouse, P. E., and Sork, V. L. (2009). Relative contribution of contemporary pollen and seed dispersal to the effective parental size of seedling population of California valley oak (*Quercus lobata*, Née). *Molecular Ecology*, 18(19):3967–3979.

Gugger, P. F., Sugita, S., and Cavender-Bares, J. (2010). Phylogeography of Douglas-fir based on mitochondrial and chloroplast DNA sequences: testing hypotheses from the fossil record. *Molecular Ecology*, 19(9):1877–1897.

Hedrick, P. W. (2019). Galapagos islands endemic vertebrates: a population genetics perspective. *Journal of Heredity*, 110(2):137–157.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405(6789):907–913.

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1442):183–195.

Hickey, R. J. (1986a). *Isoëtes* megaspore surface morphology: nomenclature, variation, and systematic importance. *American Fern Journal*, pages 1–16.

Hickey, R. J. (1986b). The early evolutionary and morphological diversity of *Isoëtes*, with descriptions of two new Neotropical species. *Systematic Botany*, pages 309–321.

Hickey, R. J., Macluf, C., and Taylor, W. C. (2003). A re-evaluation of *Isoëtes savatieri* Franchet in Argentina and Chile. *American Fern Journal*, 93(3):126–136.

Hickey, R. J., Macluf, C. C., and Link-Pérez, M. (2009). *Isoëtes maxima*, a new species from Brazil. *American Fern Journal*, pages 194–199.

Hickey, R. J., Taylor, W. C., and Luebke, N. T. (1989). The species concept in Pteridophyta with special reference to *Isoëtes*. *American Fern Journal*, pages 78–89.

Hirao, A. and Kudo, G. (2008). The effect of segregation of flowering time on fine-scale spatial genetic structure in an alpine-snowbed herb *Primula cuneifolia*. *Heredity*, 100(4):424–430.

Höhna, S., Landis, M. J., Heath, T. A., Boussau, B., Lartillot, N., Moore, B. R., Huelsenbeck, J. P., and Ronquist, F. (2016). RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Systematic biology*, 65(4):726–736.

Holway, R. S. (1911). An extension of the known area of Pleistocene glaciation to the Coast Ranges of California. *Bulletin of the American Geographical Society*, 43(3):161–170.

Hoot, S. B., Napier, N. S., and Taylor, W. C. (2004). Revealing unknown or extinct lineages within *Isoëtes* (Isoetaceae) using DNA sequencesfrom hybrids. *American Journal of Botany*, 91(6):899–904.

Hoot, S. B. and Taylor, W. C. (2001). The utility of nuclear ITS, a LEAFY homolog intron, and chloroplast atpB-rbcL spacer region data in phylogenetic analyses and species delimitation in *Isoëtes*. *American Fern Journal*, 91(3):166–177.

Hoot, S. B., Taylor, W. C., and Napier, N. S. (2006). Phylogeny and biogeography of *Isoëtes* (Isoëtaceae) based on nuclear and chloroplast DNA sequence data. *Systematic Botany*, 31(3):449–460.

Hou, Y. and Lou, A. (2011). Population genetic diversity and structure of a naturally isolated plant species, *Rhodiola dumulosa* (Crassulaceae). *PLoS One*, 6(9):e24497.

Hudson, R. R. et al. (1990). Gene genealogies and the coalescent process. *Oxford surveys in evolutionary biology*, 7(1):44.

Huelsenbeck, J. P., Larget, B., and Alfaro, M. E. (2004). Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. *Molecular biology and evolution*, 21(6):1123–1133.

Ibrahim, K. M., Nichols, R. A., and Hewitt, G. M. (1996). Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, 77(3):282–291.

Jennings, J. R., Karrfalt, E. E., and Rothwell, G. W. (1983). Structure and affinities of *Protostigmaria eggertiana*. *American Journal of Botany*, 70(7):963–974.

Jeong, Y.-H. and Choe, H.-G. (1986). *Isoëtes coreana*, a new species from Korea. *Korean Journal of Plant Taxonomy*, 16(1):1–1.

Johnson, M. G., Gardner, E. M., Liu, Y., Medina, R., Goffinet, B., Shaw, A. J., Zerega, N. J., and Wickett, N. J. (2016). HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in plant sciences*, 4(7):1600016.

Jordano, P. et al. (2010). Pollen, seeds and genes: the movement ecology of plants. *Heredity*, 105(4):329–330.

Kalisz, S., Randle, A., Chaiffetz, D., Faigeles, M., Butera, A., and Beight, C. (2012). Dichogamy correlates with outcrossing rate and defines the selfing syndrome in the mixed-mating genus *Collinsia*. *Annals of Botany*, 109(3):571–582.

Karrfalt, E. (1984a). Further observations on *Nathorstiana* (Isoetaceae). *American journal of Botany*, 71(8):1023–1030.

Karrfalt, E. (1984b). The origin and early development of the root-producing meristem of *Isoëtes andicola* LD Gomez. *Botanical Gazette*, 145(3):372–377.

Karrfalt, E. E. and Eggert, D. A. (1977a). The comparative morphology and development of *Isoëtes* L. II. Branching of the base of the corm in *I. tuckermanii* A. Br. and *I. nuttallii* A. Br. *Botanical Gazette*, 138(3):357–368.

Karrfalt, E. E. and Eggert, D. A. (1977b). The comparative morphology and development of *Isoëtes* L. Lobe and furrow development in *I. tuckermanii* A. Br. *Botanical Gazette*, 138(2):236–247.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the american statistical association*, 90(430):773–795.

Kates, H. R., Johnson, M. G., Gardner, E. M., Zerega, N. J., and Wickett, N. J. (2018). Allele phasing has minimal impact on phylogenetic reconstruction from targeted nuclear gene sequences in a case study of *Artocarpus*. *American Journal of Botany*, 105(3):404–416.

Kim, C., Bounphanmy, S., Sun, B.-Y., and Choi, H.-K. (2010). *Isoëtes laosiensis*, a new species from Lao PDR. *American Fern Journal*, 100(1):45–53.

Kingman, J. F. C. (1982). The coalescent. *Stochastic processes and their applications*, 13(3):235–248.

Kitamoto, N., Ueno, S., Takenaka, A., Tsumura, Y., Washitani, I., and Ohsawa, R. (2006). Effect of flowering phenology on pollen flow distance and the consequences for spatial genetic structure within a population of *Primula sieboldii* (Primulaceae). *American Journal of Botany*, 93(2):226–233.

Klopfstein, S., Vilhelmsen, L., and Ronquist, F. (2015). A nonstationary Markov model detects directional evolution in hymenopteran morphology. *Systematic biology*, 64(6):1089–1103.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., and Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular ecology resources*, 15(5):1179–1191.

Korneliussen, T. S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: analysis of next generation sequencing data. *BMC bioinformatics*, 15(1):1–13.

Kuchta, S. R. and TAN, A.-M. (2005). Isolation by distance and post-glacial range expansion in the rough-skinned newt, *Taricha granulosa*. *Molecular Ecology*, 14(1):225–244.

Kuss, P., Pluess, A. R., Ægisdóttir, H. H., and Stöcklin, J. (2008). Spatial isolation and genetic differentiation in naturally fragmented plant populations of the Swiss Alps. *Journal of Plant Ecology*, 1(3):149–159.

Ladner, J. T. and Palumbi, S. R. (2012). Extensive sympatry, cryptic diversity and introgression throughout the geographic distribution of two coral species complexes. *Molecular Ecology*, 21(9):2224–2238.

Larsén, E. and Rydin, C. (2016). Disentangling the phylogeny of *Isoëtes* (isoetales), using nuclear and plastid data. *International Journal of Plant Sciences*, 177(2):157–174.

Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, 30(22):3276–3278.

Lawson, D. J., Van Dorp, L., and Falush, D. (2018). A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nature communications*, 9(1):1–11.

Leebens-Mack, J. H., Barker, M. S., Carpenter, E. J., Deyholos, M. K., Gitzendanner, M. A., Graham, S. W., Grosse, I., Li, Z., Melkonian, M., Mirarab, S., et al. (2019). One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*, 574(7780):679–685.

Lellinger, D. and Taylor, W. (1997). A classification of spore ornamentation in the Pteridophyta. *Pteridology in Perspective. Royal Botanic Gardens, Kew, Reino Unido*, pages 33–42.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Liu, H., Wang, Q. F., and Taylor, W. C. (2006). Morphological and anatomical variation in sporophylls of *Isoëtes sinensis* Palmer (Isoetaceae), an endangered quillwort in China. *American Fern Journal*, 96(3):67–74.

Liu, J.-X., Tatarenkov, A., Beacham, T. D., Gorbachev, V., Wildes, S., and Avise, J. C. (2011). Effects of Pleistocene climatic fluctuations on the phylogeographic and demographic histories of Pacific herring (*Clupea pallasii*). *Molecular ecology*, 20(18):3879–3893.

Liu, L., Yu, L., Kubatko, L., Pearl, D. K., and Edwards, S. V. (2009). Coalescent methods for estimating phylogenetic trees. *Molecular Phylogenetics and Evolution*, 53(1):320–328.

Lu, Y., Ran, J.-H., Guo, D.-M., Yang, Z.-Y., and Wang, X.-Q. (2014). Phylogeny and divergence times of gymnosperms inferred from single-copy nuclear genes. *PloS one*, 9(9):e107679.

MacDonald, Z. G., Anderson, I. D., Acorn, J. H., and Nielsen, S. E. (2018). Decoupling habitat fragmentation from habitat loss: butterfly species mobility obscures fragmentation effects in a naturally fragmented landscape of lake islands. *Oecologia*, 186(1):11–27.

Macluf, C., Meza Torres, E. I., and Solís, S. M. (2010). *Isoëtes pedersenii*, a new species from Southern South America. *Anais da Academia Brasileira de Ciências*, 82:353–359.

Macluf, C. C. and Hickey, R. J. (2007). *Isoëtes araucaniana*, a new species from southern South America. *American Fern Journal*, pages 220–224.

Madigan, D. and Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window. *Journal of the American Statistical Association*, 89(428):1535–1546.

Mägdefrau, K. (1932). Über einige bohrgänge aus dem unteren muschelkalk von jena. *Paläontologische Zeitschrift*, 14(3):150–160.

Magri, D., Vendramin, G. G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D., Latałowa, M., Litt, T., Paule, L., Roure, J. M., et al. (2006). A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New phytologist*, 171(1):199–221.

Mallet, J. (2007). Hybrid speciation. *Nature*, 446(7133):279–283.

Martin, S. H., Davey, J. W., and Jiggins, C. D. (2015). Evaluating the use of ABBA–BABA statistics to locate introgressed loci. *Molecular biology and evolution*, 32(1):244–257.

Matesanz, S., Rubio Teso, M. L., García-Fernández, A., and Escudero, A. (2017). Habitat fragmentation differentially affects genetic variation, phenotypic plasticity and survival in populations of a gypsum endemic. *Frontiers in Plant Science*, 8:843.

Matter, P., Kettle, C. J., Ghazoul, J., and Pluess, A. R. (2013). Extensive contemporary pollen-mediated gene flow in two herb species, *Ranunculus bulbosus* and *Trifolium montanum*, along an altitudinal gradient in a meadow landscape. *Annals of Botany*, 111(4):611–621.

Mayr, E. (1976). Species concepts and definitions. In *Topics in the Philosophy of Biology*, pages 353–371. Springer.

McCraney, W. T., Goldsmith, G., Jacobs, D. K., and Kinziger, A. P. (2010). Rampant drift in artificially fragmented populations of the endangered tidewater goby (*Eucyclogobius newberryi*). *Molecular Ecology*, 19(16):3315–3327.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, 20(9):1297–1303.

McNeill, J., Barrie, F., Buck, W., Demoulin, V., Greuter, W. e., Hawksworth, D., Herendeen, P., Knapp, S., Marhold, K., Prado, J., et al. (2012). *International Code of Nomenclature for algae, fungi and plants (Melbourne Code)*, volume 154. Koeltz Scientific Books Königstein.

Merrill, E. D. and Perry, L. (1940). A new Philippine *Isoëtes*. *American Fern Journal*, 30(1):18–20.

Mirarab, S., Reaz, R., Bayzid, M. S., Zimmermann, T., Swenson, M. S., and Warnow, T. (2014). ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics*, 30(17):i541–i548.

Mishler, B. D. (2010). Species are not uniquely real biological entities. *Contemporary debates in philosophy of biology*, 110:122.

Moore, J. G. and Moring, B. C. (2013). Rangewide glaciation in the Sierra Nevada, California. *Geosphere*, 9(6):1804–1818.

Moura, C. J., Cunha, M. R., Porteiro, F. M., and Rogers, A. D. (2011). The use of the DNA barcode gene 16S mRNA for the clarification of taxonomic problems within the family Sertulariidae (Cnidaria, Hydrozoa). *Zoologica Scripta*, 40(5):520–537.

Musselman, L. J., Bray, R. D., and Knepper, D. A. (1996). *Isoëtes*× *bruntonii* (*Isoetes engelmannii*× *I. hyemalis*), a new hybrid quillwort from Virginia. *American Fern Journal*, pages 8–15.

Musselman, L. J., Bray, R. D., and Knepper, D. A. (1997). *Isoëtes x carltaylorii* (*Isoëtes acadiensis* x *Isoëtes engelmannii*), a new interspecific quillwort hybrid from the Chesapeake Bay. *Canadian journal of botany*, 75(2):301–309.

Nordborg, M. (2019). Coalescent theory. *Handbook of Statistical Genomics: Two Volume Set*, pages 145–30.

Olmstead, R. G., de Pamphilis, C. W., Wolfe, A. D., Young, N. D., Elisons, W. J., and Reeves, P. A. (2001). Disintegration of the Scrophulariaceae. *American journal of Botany*, 88(2):348–361.

O'Meara, B. C. (2010). New heuristic methods for joint species delimitation and species tree inference. *Systematic Biology*, 59(1):59–73.

Osborn, T. (1922). Some observations on *Isoëtes drummondii*, A. Br. *Annals of Botany*, 36(141):41–54.

Oxelman, B., Kornhall, P., Olmstead, R. G., and Bremer, B. (2005). Further disintegration of Scrophulariaceae. *Taxon*, 54(2):411–425.

Pant, D. D. and Srivastava, G. K. (1962). The genus *Isoëtes* in India. In *Proc Natl Inst Sci India*, volume 28, pages 242–270.

Patterson, T. B. and Givnish, T. J. (2004). Geographic cohesion, chromosomal evolution, parallel adaptive radiations, and consequent floral adaptations in *Calochortus* (Calochortaceae): evidence from a cpDNA phylogeny. *New Phytologist*, 161(1):253–264.

Pauli, G. G. M. G. H. (1994). Climate effects on mountain plants. *Nature*, 369:448.

Pereira, J. and Labiak, P. H. (2013). A new species of *Isoëtes* with tuberculate spores from Southeastern Brazil (Isoetaceae). *Systematic Botany*, 38(4):869–874.

Peter, B. M. (2016). Admixture, population structure, and F-statistics. *Genetics*, 202(4):1485–1501.

Pfeiffer, N. E. (1922). Monograph of the Isoetaceae. *Annals of the Missouri Botanical Garden*, 9(2):79–233.

Pigg, K. B. (1992). Evolution of isoetalean lycopsids. *Annals of the Missouri Botanical Garden*, pages 589–612.

Pigg, K. B. and Rothwell, G. (1985). Cortical development in *Chaloneria cormosa* (Isoetales), and the biological derivation of compressed lycophyte decortication taxa. *Palaeontology*, 28(3):545–553.

Pither, R., Shore, J., and Kellman, M. (2003). Genetic diversity of the tropical tree *Terminalia amazonia* (Combretaceae) in naturally fragmented populations. *Heredity*, 91(3):307–313.

Prada, C. and Rolleri, C. H. (2005). A new species of *Isoëtes* (Isoetaceae) from Turkey, with a study of microphyll intercellular pectic protuberances and their potential taxonomic value. *Botanical Journal of the Linnean Society*, 147(2):213–228.

Pryer, K. M., Schuettpelz, E., and Rothfels, C. J. (2021). Fern* Labs Database. https://fernlab.biology.duke.edu/.

Qing-Feng, W., Xing, L., Taylor, W. C., and Zhao-Rong, H. (2002). *Isoëtes yunguiensis* (Isoetaceae), a new basic diploid quillwort from China. *Novon*, pages 587–591.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rannala, B. and Yang, Z. (2003). Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, 164(4):1645–1656.

Rannala, B. and Yang, Z. (2017). Efficient Bayesian species tree inference under the multispecies coalescent. *Systematic biology*, 66(5):823–842.

Rao, L. N. (1944). A new species of *Isoëtes* from Bangalore, Mysore State. *Current Science*, 13(11):286–287.

Rasmussen, I. and Brødsgaard, B. (1992). Gene flow inferred from seed dispersal and pollinator behaviour compared to DNA analysis of restriction site variation in a patchy population of *Lotus corniculatus* L. *Oecologia*, 89(2):277–283.

Richter, P. (1910). *Nathorstiana* P. Richter und *Cylindrites spongioides* Goepp. *Zeitschrift der Deutschen Geologischen Gesellschaft*, pages 278–284.

Rieseberg, L. H. and Willis, J. H. (2007). Plant speciation. *science*, 317(5840):910–914.

Roland, J., Keyghobadi, N., and Fownes, S. (2000). Alpine *Parnassius* butterfly dispersal: effects of landscape and population size. *Ecology*, 81(6):1642–1653.

Rothfels, C. J., Johnson, A. K., Hovenkamp, P. H., Swofford, D. L., Roskam, H. C., Fraser-Jenkins, C. R., Windham, M. D., and Pryer, K. M. (2015). Natural hybridization between genera that diverged from each other approximately 60 million years ago. *The American Naturalist*, 185(3):433–442.

Rothfels, C. J., Windham, M. D., Grusz, A. L., Gastony, G. J., and Pryer, K. M. (2008). Toward a monophyletic *Notholaena* (Pteridaceae): Resolving patterns of evolutionary convergence in xeric-adapted ferns. *Taxon*, 57(3):712–724.

Rothwell, G. W. (1984). The apex of *Stigmaria* (Lycopsida), rooting organ of Lepidodendrales. *American Journal of Botany*, 71(8):1031–1034.

Rothwell, G. W. and Erwin, D. M. (1985). The rhizomorph apex of *Paurodendron*; implications for homologies among the rooting organs of Lycopsida. *American Journal of Botany*, 72(1):86–98.

Roux, J., Hopper, S., and Smith, R. (2009). *Isoëtes eludens* (Isoetaceae), a new endemic species from the Kamiesberg, Northern Cape, South Africa. *Kew Bulletin*, 64(1):123–128.

Rury, P. M. (1978). A new and unique, mat-forming Merlin's-grass (*Isoëtes*) from Georgia. *American Fern Journal*, pages 99–108.

Rydin, C. and Wikström, N. (2002). Phylogeny of *Isoëtes* (Lycopsida): resolving basal relationships using rbcL sequences. *Taxon*, 51(1):83–89.

Schafran, P. W. (2019). *Molecular systematics of Isoëtes (Isoetaceae) in eastern North America*. PhD thesis, Old Dominion University.

Schmitt, J. (1980). Pollinator foraging behavior and gene dispersal in *Senecio* (Compositae). *Evolution*, pages 934–943.

Schneider, H., Schuettpelz, E., Pryer, K. M., Cranfill, R., Magallón, S., and Lupia, R. (2004a). Ferns diversified in the shadow of angiosperms. *Nature*, 428(6982):553–557.

Schneider, H., Smith, A. R., Cranfill, R., Hildebrand, T. J., Haufler, C. H., and Ranker, T. A. (2004b). Unraveling the phylogeny of polygrammoid ferns (Polypodiaceae and Grammitidaceae): exploring aspects of the diversification of epiphytic plants. *Molecular phylogenetics and evolution*, 31(3):1041–1063.

Schrinner, S. D., Mari, R. S., Ebler, J., Rautiainen, M., Seillier, L., Reimer, J. J., Usadel, B., Marschall, T., and Klau, G. W. (2020). Haplotype threading: accurate polyploid phasing from long reads. *Genome biology*, 21(1):1–22.

Schuettpelz, E. and Hoot, S. B. (2006). Inferring the root of *Isoëtes*: exploring alternatives in the absence of an acceptable outgroup. *Systematic Botany*, 31(2):258–270.

Scobie, A. and Wilcock, C. (2009). Limited mate availability decreases reproductive success of fragmented populations of *Linnaea borealis*, a rare, clonal self-incompatible plant. *Annals of Botany*, 103(6):835–846.

Scott, D. H. and Hill, T. G. (1900). The structure of *Isoëtes hystrix*. *Annals of Botany*, 14(55):413–454.

Segelbacher, G., Höglund, J., and Storch, I. (2003). From connectivity to isolation: genetic consequences of population fragmentation in capercaillie across Europe. *Molecular Ecology*, 12(7):1773–1780.

Sharma, B. and Singh, R. (1984). The ligule in *Isoëtes*. *American fern journal*, pages 22–28.

Shaw, S. W. and Hickey, R. J. (2005). Comparative morphology of the glossopodia of three North American *Isoëtes* ligules. *American Fern Journal*, pages 94–114.

Sheets, E. A., Warner, P. A., and Palumbi, S. R. (2018). Accurate population genetic measurements require cryptic species identification in corals. *Coral Reefs*, 37(2):549–563.

Singh, S. K., Srivastava, G., and Shukla, P. (2010). Comparative morphology of ligules of three Indian species of *Selaginella*. *American Fern Journal*, pages 71–79.

Skotte, L., Korneliussen, T. S., and Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics*, 195(3):693–702.

Slimp, M., Williams, L. D., Hale, H., and Johnson, M. G. (2021). On the potential of Angiosperms353 for population genomic studies. *Applications in plant sciences*, 9(7).

Sloop, C. M., Eberl, R., and Ayres, D. R. (2012). Genetic diversity and structure in the annual vernal pool endemic *Limnanthes vinculans* Ornduff (Limnanthaceae): implications of breeding system and restoration practices. *Conservation Genetics*, 13(5):1365–1379.

Soltis, D. E. and Soltis, P. S. (1999). Polyploidy: recurrent formation and genome evolution. *Trends in ecology & evolution*, 14(9):348–352.

Soltis, D. E., Soltis, P. S., Schemske, D. W., Hancock, J. F., Thompson, J. N., Husband, B. C., and Judd, W. S. (2007). Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon*, 56(1):13–30.

Soltis, D. E., Soltis, P. S., and Tate, J. A. (2004). Advances in the study of polyploidy since plant speciation. *New phytologist*, 161(1):173–191.

Soltis, D. E., Visger, C. J., and Soltis, P. S. (2014). The polyploidy revolution then... and now: Stebbins revisited. *American journal of botany*, 101(7):1057–1078.

Springer, M. S. and Gatesy, J. (2019). An ABBA-BABA test for introgression using retroposon insertion data. *bioRxiv*, page 709477.

Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics*, 24(5):637–644.

Stewart, W. N. (1947). A comparative study of stigmarian appendages and *Isoëtes* roots. *American Journal of Botany*, pages 315–324.

Stokey, A. G. (1909). The anatomy of *Isoëtes*. *Botanical Gazette*, 47(4):311–335.

Stone, B. W. and Wolfe, A. D. (2021). Phylogeographic analysis of shrubby beardtongues reveals range expansions during the Last Glacial Maximum and implicates the Klamath Mountains as a hotspot for hybridization. *Molecular Ecology*, 30(15):3826–3839.

Stoughton, T. R., Jolles, D. D., and O'Quinn, R. L. (2017). The western spring beauties, *Claytonia lanceolata* (Montiaceae): a review and revised taxonomy for California. *Systematic Botany*, 42(2):283–300.

Stoughton, T. R., Kriebel, R., Jolles, D. D., and O'Quinn, R. L. (2018). Next-generation lineage discovery: A case study of tuberous *Claytonia* L. *American Journal of Botany*, 105(3):536–548.

Takamiya, M., Watanabe, M., and Ono, K. (1998). Biosystematic Studies on the Genus *Isoëtes* (Isoetaceae) in Japan. IV.: Morphology and Anatomy of Sporophytes, Phytogeography and Taxonomy. *Acta Phytotaxonomica et Geobotanica*, 48(2):89–121.

Takebayashi, N., Wolf, D., and Delph, L. (2006). Effect of variation in herkogamy on outcrossing within a population of *Gilia achilleifolia*. *Heredity*, 96(2):159–165.

Taylor, E. L., Taylor, T. N., and Krings, M. (2009). *Paleobotany: the biology and evolution of fossil plants*. Academic Press.

Taylor, W. (1993). Isoetaceae Reichenbach. *Flora of North america*, 2:64–75.

Taylor, W. C. and Hickey, R. J. (1992). Habitat, evolution, and speciation in *Isoëtes*. *Annals of the Missouri Botanical Garden*, pages 613–622.

Taylor, W. C., Luebke, N. T., and Smith, M. B. (1985). Speciation and hybridisation in North American quillworts. *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences*, 86:259–263.

Thomson, J. D. and Barrett, S. C. (1981). Selection for outcrossing, sexual selection, and the evolution of dioecy in plants. *The American Naturalist*, 118(3):443–449.

Troia, A. and Greuter, W. (2014). A critical conspectus of Italian *Isoëtes* (Isoetaceae). *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology*, 148(1):13–20.

Troia, A., Johnson, G., and Taylor, W. C. (2019). A contribution to the phylogeny and biogeography of the genus *Isoëtes* (Isoetaceae, Lycopodiidae) in the Mediterranean region. *Phytotaxa*, 395(3):168–178.

Troia, A., Pereira, J. B. S., Kim, C., and Taylor, W. C. (2016). The genus *Isoëtes* (Isoetaceae): a provisional checklist of the accepted and unresolved taxa. *Phytotaxa*, 277(2):101–145.

Tryon, R. M. (1994). Pteridophyta of Peru. Part VI. 2. Marsileaceae - 28. Isoetaceae. *Fieldiana Bot NS*, 34:1–123.

Underwood, L. M. (1888). The distribution of *Isoëtes*. *Botanical Gazette*, 13(4):89–94.

Van Rossum, F. and Triest, L. (2006). Fine-scale genetic structure of the common *Primula elatior* (Primulaceae) at an early stage of population fragmentation. *American Journal of Botany*, 93(9):1281–1288.

Wallace, A. R. (1876). *The geographical distribution of animals*. Harper & Brothers.

Wallace, A. R. (2016). *Evolution and character*. Read Books Ltd.

Weber, U. (1922). Zur anatomie und systematik der gattung *Isoëtes* L. *Nova Hedwigia*, 63:219–262.

Westerbergh, A. and Saura, A. (1994). Gene flow and pollinator behaviour in *Silene dioica* populations. *Oikos*, pages 215–224.

Whiteman, N. K., Kimball, R. T., and Parker, P. G. (2007). Co-phylogeography and comparative population genetics of the threatened Galápagos hawk and three ectoparasite species: ecology shapes population histories within parasite communities. *Molecular Ecology*, 16(22):4759–4773.

Wickell, D. A. and Li, F.-W. (2020). On the evolutionary significance of horizontal gene transfers in plants. *New Phytologist*, 225(1):113–117.

Wikström, N., Savolainen, V., and Chase, M. W. (2001). Evolution of the angiosperms: calibrating the family tree. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1482):2211–2220.

Williams, S. (1944). I.—On *Isoëtes australis* S. Williams, a new species from Western Australia. Part I. General morphology. *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences*, 62(1):1–8.

Wood, D. P., Olofsson, J. K., McKenzie, S. W., and Dunning, L. T. (2018). Contrasting phylogeographic structures between freshwater lycopods and angiosperms in the British Isles. *Botany Letters*, 165(3-4):476–486.

Xi, Z., Bradley, R. K., Wurdack, K. J., Wong, K., Sugumaran, M., Bomblies, K., Rest, J. S., and Davis, C. C. (2012). Horizontal transfer of expressed genes in a parasitic flowering plant. *BMC genomics*, 13(1):1–8.

Xi, Z., Wang, Y., Bradley, R. K., Sugumaran, M., Marx, C. J., Rest, J. S., and Davis, C. C. (2013). Massive mitochondrial gene transfer in a parasitic flowering plant clade. *PLoS genetics*, 9(2):e1003265.

Yatskievych, G. et al. (2008). A new species and three generic transfers in the fern genus *Notholaena* (Pteridaceae). *Novon: A Journal for Botanical Nomenclature*, 18(1):120–124.

Yi, S.-Y. and Kato, M. (2001). Basal meristem and root development in *Isoëtes asiatica* and *Isoëtes japonica*. *International Journal of Plant Sciences*, 162(6):1225–1235.

Yoder, A. D., Rasoloarison, R. M., Goodman, S. M., Irwin, J. A., Atsalis, S., Ravosa, M. J., and Ganzhorn, J. U. (2000). Remarkable species diversity in Malagasy mouse lemurs (Primates, *Microcebus*). *Proceedings of the National Academy of Sciences*, 97(21):11325–11330.

Zhang, J., Kapli, P., Pavlidis, P., and Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, 29(22):2869–2876.

Zhu, Y., Geng, Y., Tersing, T., Liu, N., Wang, Q., and Zhong, Y. (2009). High genetic differentiation and low genetic diversity in *Incarvillea younghusbandii*, an endemic plant of Qinghai-Tibetan Plateau, revealed by AFLP markers. *Biochemical Systematics and Ecology*, 37(5):589–596.

APPENDIX 1

Vouchers examined for morphological determinations in this study. Isoëtes bolanderi: RSA (811637, 811638, 811639, 811643). I. howellii: RSA (796366, 796367, 796368, 796369, 796370, 796371), UC (F. Freund 255, 266.1). I. nuttallii: RSA (796374, 796375, 796376), UC (F. Freund 256, 257, 258, 259, 261, 263, 264, 265, 266). I. occidentalis: RSA (811640, 811641, 811642). I. orcuttii: UC (F. Freund 254, 262). I. storkii: GH00021453, CM0102, NY00144272 (virtual herbarium sheets). For herbarium acronyms, see Index Herbariorum (http://sweetgum.nybg.org/science/ih/).

APPENDIX 2

Table .14: Voucher information for samples included in Ch. 2 & 3

Species ID from Keys	Collection number	Institution	Asscesion Number	Fern DB #
I. minima	D.Thomas 13378	N/A	N/A	N/A
I. nuttallii	C.J.Rothfels5205	UC	N/A	12321
I. nuttallii	C.J.Rothfels5256	UC	UC2080835	13842
I. nuttallii	F.Freund169	RSA	RSA796374	11519
I. nuttallii	F.Freund257	UC	UC2062876	11532
I. nuttallii	F.Freund258	UC	UC2062873	11533
I. nuttallii	F.Freund259	UC	UC2062871	11535
I. nuttallii	F.Freund261	UC	UC2062872	11536
I. nuttallii	F.Freund263	UC	UC2079857	13542
I. nuttallii	F.Freund266	UC	UC2079858	13545
I. nuttallii	F.Freund279	UC	N/A	13550
I. nuttallii	F.Freund280	UC	N/A	13551
I. nuttallii	F.Freund282	UC	N/A	13552
I. nuttallii	F.Freund283	UC	N/A	13554
I. nuttallii	F.Freund284	UC	N/A	13555
I. nuttallii	F.Freund289	UC	N/A	13849
I. nuttallii	F.Freund290	UC	N/A	13850
I. orcuttii	F.Freund254	UC	UC2062874	11529
I. orcuttii	F.Freund262	UC	UC2079855	13541
I. orcuttii	F.Freund285	UC	N/A	13841
I. orcuttii	F.Freund286	UC	N/A	13845
I. orcuttii	K.Wefferling380	UC	N/A	14588

Appendix 3

```
# R-script to convert .nexus files with excluded loci
# to .fasta files.

library(ape)
library(tools)

# the directories
setwd(# path to location of source_dir)
source_dir <- # path to source directory
out_dir    <- # path to write directory

# create the output folder
dir.create(out_dir, recursive = TRUE)

# list all nexus files
nexus_files <- list.files(source_dir, full.names = TRUE)
nexus_files <- nexus_files[file_ext(nexus_files) %in%
c("nexus", "nex")]

# read all the nexus files
alns <- vector("list", length(nexus_files))
for(i in 1:length(alns)) {

    # get the file
    this_file <- nexus_files[i]

    cat("Processing file ", this_file, "\n")

    # read the alignment
    this_aln <- try(read.nexus.data(this_file))

    if ( class(this_aln) != "try-error" ) {

        # get the new filename
        new_file <- gsub(source_dir, out_dir, this_file)
        new_file <- gsub(".nexus", ".fasta", new_file)
        new_file <- gsub(".nex", ".fasta", new_file)

        # drop excluded characters
        aln_lines <- readLines(this_file)
```

```
        exclude_line <- grep("EXSET", aln_lines, value = TRUE)
        exclude_line <- strsplit(exclude_line, "=")[[1]][2]
        exclude_line <- strsplit(exclude_line, " ")[[1]]
        exclude_line <- gsub(";", "", exclude_line)
        exclude_line <- exclude_line[exclude_line != ""]
        if ( length(exclude_line) > 0 ) {
            excludes <- unlist(lapply(exclude_line, function(x)
            {
                tmp <- as.numeric(strsplit(x,"-")[[1]])
                tmp <- seq(head(tmp, 1), tail(tmp, 1))
                return(tmp)
            }))

            for(j in 1:length(this_aln)) {
                this_aln[[j]] <- this_aln[[j]][-excludes]
            }
        }

        # reformat
        this_aln <- as.DNAbin(this_aln)

        # write the alignment
        write.dna(this_aln, file = new_file, format = "fasta",
        nbcol = -1, colsep="")

    } else {
        cat("Problem with aln ", this_file, "\n")
    }

}
```

Appendix 4

```
# R–Script to generate SVDQuartets sequence and control files

library(ape)
library(tools)
library(stringr)

# point to the delimitation file and set parameters of the model

#ancestral
# delimitation_file <- # path to csv file
# outgroup_population <- c(# names of the taxa
that will be treated as the outgroup)
# num_quartets <- 25000000
# num_bootstrap_reps <- 3000
# num_quartets_bootstrap <- 100000

#sites
# delimitation_file <- # path to csv file
# outgroup_population <- c(# names of the taxa
that will be treated as the outgroup)
# num_quartets <- 50000000
# num_bootstrap_reps <- 1000
# num_quartets_bootstrap <- 570000

#Individuals
delimitation_file <- # path to csv file
outgroup_population <- c(# names of the taxa
that will be treated as the outgroup)
num_quartets <- 50000000
num_bootstrap_reps <- 1000
num_quartets_bootstrap <- 570000
#

# specify the samples to exclude
samples_to_exclude <- c(# names of the taxa
in the dataset to be excluded)

# point to the dataset directory
dir <- # path to source directory
# dir <- "Phased_data_nex/"
```

```
# analysis settings
model <- "lset nst=6 rmatrix=estimate
basefreq=empirical rates=gamma ncat=4 shape=estimate
pinvar=estimate;"

num_tree_runs <- 1
num_bootstrap_runs <- 1

# create the new directory
new_dir <- file_path_sans_ext(delimitation_file)
dir.create(new_dir, showWarnings = FALSE)

# read the delimitation file
delimitation <- read.csv(delimitation_file)

# list all the files there
if ( any(grepl("nexus", list.files(dir))) ) {
  format <- "nexus"
  phased_alignments <- list.files(dir, pattern = "phased.nexus",
  full.names = TRUE)
} else {
  format <- "fasta"
  phased_alignments <- list.files(dir, pattern = "phased.fasta",
  full.names = TRUE)
}

# create the alignment
alignments <- vector("list", length(phased_alignments))
for(i in 1:length(phased_alignments)) {

  cat(i,"\n")

  # read the alignment
  if ( format == "nexus" ) {

    # read the nexus file
    aln <- try(read.nexus.data(phased_alignments[i]))
    if (class(aln) == "try-error") {
      warning("Could not read alignment ",
      phased_alignments[i], ". Make sure it is a nexus file")
      next
```

```
    }

    # drop excluded characters
    aln_lines <- readLines(phased_alignments[i])
    exclude_line <- grep("EXSET", aln_lines, value = TRUE)
    exclude_line <- strsplit(exclude_line, "=")[[1]][2]
    exclude_line <- strsplit(exclude_line, " ")[[1]]
    exclude_line <- gsub(";", "", exclude_line)
    exclude_line <- exclude_line[exclude_line != ""]
    if ( length(exclude_line) > 0 ) {
      excludes <- unlist(lapply(exclude_line, function(x) {
        tmp <- as.numeric(strsplit(x,"-")[[1]])
        tmp <- seq(head(tmp, 1), tail(tmp, 1))
        return(tmp)
      }))

      for(j in 1:length(aln)) {
        aln[[j]] <- aln[[j]][-excludes]
      }
    }

    aln <- as.DNAbin(aln)

  } else {
    aln <- read.FASTA(phased_alignments[i])
  }

  # remove samples that are to be excluded
  these_samples_to_exclude <- unlist(sapply(samples_to_exclude,
  function(x) grep(x, names(aln), value = TRUE)))
  aln <- aln[names(aln) %in% these_samples_to_exclude == FALSE]

  # get rid of spaces in sequence name
  names(aln) <- sapply(strsplit(names(aln), " "), head, n=1)

  # drop the gene name
  aln <- as.matrix(aln)
  old_names <- rownames(aln)
  new_names <- sapply(strsplit(old_names, "-"), head, n=1)
  new_names <- paste0(new_names, ifelse(grepl("_h1", old_names),
  ".A", ".B"))
  rownames(aln) <- new_names
```

```
  # record the alignment
  alignments[[i]] <- aln

}

# remove nulls
alignments <- alignments[sapply(alignments, is.null) == FALSE]

# concatenate the alignments
bind_aln <- function(...) cbind.DNAbin(..., fill.with.gaps = TRUE)
concatenated_aln <- do.call(bind_aln, alignments)

# relabel the taxa
old_names <- rownames(concatenated_aln)
new_names <- paste0("^", old_names)
rownames(concatenated_aln) <- new_names

# get the groups
groups <- unique(delimitation\$Assignment.to.species)
ngroups <- length(groups)
tips_per_group <- vector("list", ngroups)
names(tips_per_group) <- groups
for(i in 1:ngroups) {

  # get the group
  this_group <- groups[i]

  # find all sequences in this group
  these_sequences <- delimitation
  [delimitation\$Assignment.to.species == this_group,]$Specimen
  these_sequences <- as.vector(outer(c("A","B"), these_sequences,
  function(x, y) paste0("^",y, ".", x) ))

  # store
  tips_per_group[[i]] <- these_sequences

}

# remove empty groups
tips_per_group <- tips_per_group[sapply(tips_per_group,
is.null) == FALSE]
```

```
groups <- names(tips_per_group)
ngroups <- length(groups)

# get all the samples
all_samples <- unlist(tips_per_group)

# write the alignment
write.nexus.data(concatenated_aln,
                 file = paste0(new_dir, "/seq.nex"),
                 interleaved = FALSE)

# write the SVDQ nexus file
svdq_file <- paste0(new_dir, "/svdq.nex")
cat("#NEXUS\n\n", file = svdq_file, sep = "")

cat("begin paup;
  cd *;
  set hashcomment=y;
end;\n\n", file = svdq_file, sep = "", append = TRUE)

cat("execute seq.nex;\n\n", file = svdq_file, sep = "",
append = TRUE)

cat("begin sets;\n", file = svdq_file, sep = "", append = TRUE)
cat("taxpartition species =\n", file = svdq_file, sep = "",
append = TRUE)
for(i in 1:length(tips_per_group)) {
  this_group    <- groups[i]
  these_samples <- tips_per_group[[i]]
  these_samples   <- paste0(these_samples, collapse = " ")
  cat("\t", this_group, " : ", these_samples,
  file = svdq_file, sep = "", append = TRUE)
  if ( i == length(tips_per_group) ) {
    cat(";\n\n", file = svdq_file, sep = "", append = TRUE)
  } else {
    cat(",\n", file = svdq_file, sep = "", append = TRUE)
  }
}

cat("charpartition lociset =\n", file = svdq_file, sep = "",
append = TRUE)
current_position <- 0
```

```
for(i in 1:length(alignments)) {

  # get the locus
  this_aln <- alignments[[i]]

  # get the positions
  start <- 1 + current_position
  end   <- ncol(this_aln) + current_position

  # add the info
  cat(i, ": ", start, "-", end, file = svdq_file, sep = "",
  append = TRUE)
  if ( i == length(alignments) ) {
    cat(";\n", file = svdq_file, sep = "", append = TRUE)
  } else {
    cat(",\n", file = svdq_file, sep = "", append = TRUE)
  }

  # increment the positions
  current_position <- end

}
cat("end;\n\n", file = svdq_file, sep = "", append = TRUE)

# specify the outgroup
outgroup <- unlist(tips_per_group[outgroup_population])
cat("outgroup ", paste0(outgroup, collapse=" "), ";\n",
file = svdq_file, sep = "", append = TRUE)

# specify the model
cat(model, "\n\n", file = svdq_file, sep = "", append = TRUE)

# replicates
for(i in 1:num_tree_runs) {
  cat("svdq taxpartition=species evalQuartets=all showScores=no
  replace=yes;\n", file = svdq_file, sep = "", append = TRUE)
  cat("savetrees brlens file=MyTree", i, ".tre replace;\n\n",
  file = svdq_file, sep = "", append = TRUE)
}

for(i in 1:num_bootstrap_runs) {
  cat(paste0("svdq taxpartition=species nquartets=",
```

```
  num_quartets_bootstrap, " showScores=no bootstrap nreps=",
  num_bootstrap_reps, " treeFile=mybootstraptrees",i,
  ".tre replace=yes;\n"), file = svdq_file, sep = "",
  append = TRUE)
  cat("savetrees brlens file=MyBSTree", i,
  ".tre replace;\n\n", file = svdq_file, sep = "",
  append = TRUE)
}

cat("quit;\n", file = svdq_file, sep = "", append = TRUE)
```

Appendix 5

```
# Generate bootstrap support values from multiple SVDQbootstrap
# runs.
library(ape)

# filenames
target_tree      <- "MyTree.tre"
bootstrap_trees <- "AllBootTrees.tre"
output_tree      <- "MyTree_Bootstraps.tre"
consensus_tree <- "ConsensusTree.tre"

# read the files
tree      <- read.nexus(target_tree)
bs_trees <- read.nexus(bootstrap_trees)

# compute the bootstraps
bs_scores <- 100 * prop.clades(tree, bs_trees) /
length(bs_trees)
tree$node.label <- bs_scores

# write the tree
write.tree(tree, file = output_tree)


#create consensus tree
cons_tree <- consensus(bs_trees, p=0.70,
check.labels = TRUE, rooted = FALSE)
write.tree(cons_tree, file = consensus_tree)
```

Appendix 6
*Control file for execution of BPP A00 analysis.
seed = -1 /p
seqfile = bpp.txt
Imapfile = bppmap.txt
outfile = bpp_out_A00_18.txt
mcmcfile = mcmc_A00_18.txt
speciesdelimitation = 0
speciestree = 0
species&tree = 8 HL Min JP SC PenR SN CoR WV
60 60 60 60 60 60 60 60
HL,(Min,(JP,(SC,((PenR,SN),(CoR,WV))))));
phase = 0 0 0 0 0 0 0 0
cleandata = 0
usedata = 1 * 0: no data (prior); 1:seq like
nloci = 30 * number of data sets in seqfile
model = gtr
alphaprior = 1 1 4
thetaprior = 3 0.002 e * uncomment this to estimate thetas (pop sizes)
tauprior = 3 0.002 * invgamma(a, b) for root tau & Dirichlet(a) for other tau's
locusrate = 1 0 0 2 dir * (1) turn on/off variation (2) a_mubar (3) b_mubar (4) a_mui (5) optional prior
clock = 1 * Mike said start simple (this specifies strict clock)
finetune = 1: 5 0.001 0.001 0.001 0.3 0.33 1.0 * finetune for GBtj, GBspr, theta, tau, mix, locusrate, seqerr
print = 1 0 0 0 * MCMC samples, locusrate, hereditystalars, Genetrees
burnin = 50000
sampfreq = 2
nsample = 150000
Threads = 20 1 1

Appendix 7

```
#MANOVA analysis of leaf characters

#set working directory
setwd(#path to working directory)

#libraries needed
library(dplyr)
library(readr)
library(tidyverse)
library(ggpubr)
library(rstatix)
library(car)
library(broom)

#importing data
#Averaged leaf values
Leaf_measurments <- read_csv(#path to .csv file with
# leaf measurements)

#Manova test for averaged leaf values
res.man <- manova(cbind(Number_of_leaves, longest_leaf_length,
shortest_leaf_length, Average_leaf_length) ~ Group,
data = Leaf_measurments)
summary(res.man)

#Examine residuals
summary <- summary.aov(res.man)
summary


# Summary Stats
Post_hoc_Pillai <- summary.manova(res.man, test = c("Pillai"),
intercept = FALSE)
Post_hoc_Pillai
Post_hoc_Wilks <- summary.manova(res.man, test = c("Wilks"),
intercept = FALSE)
Post_hoc_Wilks


#leaf average
```

```
png(filename= "Leaf_Average_MANOVA.png",width = 1100,
height = 400, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
ggboxplot(Leaf_measurments,
          x = "Group",
          y = c("Average_leaf_length"
          ),
          merge = TRUE,
          bxp.errorbar = TRUE,
          notch = FALSE,
          width = 0.75,
          outlier.shape = NA
)
dev.off()

#leaf shortest
png(filename= "Leaf_shortest_MANOVA.png",width = 1100,
height = 400, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
ggboxplot(Leaf_measurments,
          x = "Group",
          y = c("shortest_leaf_length"
          ),
          merge = TRUE,
          bxp.errorbar = TRUE,
          notch = FALSE,
          width = 0.75,
          outlier.shape = NA
)
dev.off()

#leaf longest
png(filename= "Leaf_longest_MANOVA.png",width = 1100,
height = 400, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
ggboxplot(Leaf_measurments,
          x = "Group",
          y = c("longest_leaf_length"
                      ),
```

```
            merge = TRUE,
            bxp.errorbar = TRUE,
            notch = FALSE,
            width = 0.75,
            outlier.shape = NA
)
dev.off()

#number of leaves
png(filename= "Leaf_count_MANOVA.png",width = 1100,
height = 400, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
ggboxplot(Leaf_measurments,
            x = "Group",
            y = c("Number_of_leaves"
            ),
            merge = TRUE,
            bxp.errorbar = TRUE,
            notch = FALSE,
            width = 0.75,
            outlier.shape = NA
)
dev.off()

#Tukey HSD
LeafAverageTukey <- TukeyHSD(aov(Average_leaf_length ~ Group,
Leaf_measurments), ordered = TRUE)
LeafAverageTukeyData <- as.data.frame(LeafAverageTukey[1:1])
write.csv(LeafAverageTukeyData, file="LeafAverageTukey.csv",
row.names = TRUE)

LeafSmallestTukey <- TukeyHSD(aov(shortest_leaf_length ~ Group,
Leaf_measurments), ordered = TRUE)
LeafSmallestTukeyData <- as.data.frame(LeafSmallestTukey[1:1])
write.csv(LeafSmallestTukeyData, file="LeafSmallestTukey.csv",
row.names = TRUE)

LeafLargestTukey <- TukeyHSD(aov(longest_leaf_length ~ Group,
Leaf_measurments), ordered = TRUE)
LeafLargestTukeyData <- as.data.frame(LeafLargestTukey[1:1])
write.csv(LeafLargestTukeyData, file="LeafLargestTukey.csv",
```

```
row.names = TRUE)

LeafNumberTukey <- TukeyHSD(aov(Number_of_leaves ~ Group,
Leaf_measurments), ordered = TRUE)
LeafNumberTukeyData <- as.data.frame(LeafNumberTukey[1:1])
write.csv(LeafNumberTukeyData, file="LeafNumberTukey.csv",
row.names = TRUE)
```

Appendix 8

```
#MANOVA analysis of spore characters

#set working directory
setwd(#path to working directory)

#libraries needed
library(dplyr)
library(readr)
library(tidyverse)
library(ggpubr)
library(rstatix)
library(car)
library(broom)

#importing data
#Averaged leaf values
Spore_measurments <- read_csv(#path to .csv file with
# spore measurements)

#Manova test for averaged leaf values
res.man <- manova(cbind(Average, Smallest, Largest) ~ Group,
data = Spore_measurments)
summary(res.man)

#Examine residuals
summary <- summary.aov(res.man)
summary


# Summary Stats
Post_hoc_Pillai <- summary.manova(res.man, test = c("Pillai"),
intercept = FALSE)
Post_hoc_Pillai
Post_hoc_Wilks <- summary.manova(res.man, test = c("Wilks"),
intercept = FALSE)
Post_hoc_Wilks

#Tukey HSD
SporeAverageTukey <- TukeyHSD(aov(Average ~ Group,
Spore_measurments), ordered = TRUE)
```

```
SporeAverageTukeyData <- as.data.frame(SporeAverageTukey[1:1])
#write.csv(SporeAverageTukeyData, file="SporeAverageTukey.csv",
row.names = TRUE)

SporeSmallestTukey <- TukeyHSD(aov(Smallest ~ Group,
Spore_measurments), ordered = TRUE)
SporeSmallestTukeyData <- as.data.frame(SporeSmallestTukey[1:1])
#write.csv(SporeSmallestTukeyData, file="SporeSmallestTukey.csv",
row.names = TRUE)

SporeLargestTukey <- TukeyHSD(aov(Average ~ Group,
Spore_measurments), ordered = TRUE)
SporeLargestTukeyData <- as.data.frame(SporeLargestTukey[1:1])
#write.csv(SporeLargestTukeyData, file="SporeLargestTukey.csv",
row.names = TRUE)


#spore average measurements
png(filename= "Spore_MANOVA.png",width = 1100,
height = 800, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
ggboxplot(Spore_measurments,
          x = "Group",
          y = "Average",
          merge = TRUE,
          bxp.errorbar = TRUE,
          notch = FALSE,
          width = 0.75,
          outlier.shape = NA
)
dev.off()

#spore largest measurements
png(filename= "Spore_largest_MANOVA.png",width = 1100,
height = 800, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
ggboxplot(Spore_measurments,
          x = "Group",
          y = "Largest"
          ,
```

```
            merge = TRUE,
            bxp.errorbar = TRUE,
            notch = FALSE,
            width = 0.75,
            outlier.shape = NA
)
dev.off()

#spore Smallest
png(filename= "Spore_smallest_MANOVA.png",width = 1100,
height = 800, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
ggboxplot(Spore_measurments,
            x = "Group",
            y = "Smallest",
            merge = TRUE,
            bxp.errorbar = TRUE,
            notch = FALSE,
            width = 0.75,
            outlier.shape = NA
)
dev.off()
```

Appendix 9

```
#PCA of Isoetes leaf data


library(factoextra)
library(ape)
library(tools)
library(stringr)
library(readr)
library(missMDA)

#Load in Data into R
setwd(#path to working directory)

PCA_Data <- read_csv(#path to CSV file with the leaf data,

                     col_types = cols(Plant_number = col_number(),
                         Number_of_leaves = col_number(),
                         longest_leaf_length = col_number(),
                         shortest_leaf_length = col_number(),
                         Average_leaf_length = col_number()))

PCA_Data
#Set the columns and rows to analyze, where x & y are the range of
#rows and columns to be analyzed in 'i:j' format, where i is the
#first row, and j is the last.
PCA_data.active <- PCA_Data[1:647, 5:8]
PCA_data.active

#Perform the principal components analysis
res.pca <- prcomp(PCA_data.active, scale = TRUE)

# Eigenvalues
eig.val <- get_eigenvalue(res.pca)
write.csv(eig.val, file = "PCA_Eigen_Leaf.csv", row.names = TRUE)

# Results for Variables
res.var <- get_pca_var(res.pca)
write.csv(res.var$coord, file = "PCA_Leaf_vars_coord.csv",
row.names = TRUE) # Coordinates
write.csv(res.var$contrib, file = "PCA_Leaf_vars_contrib.csv",
```

```
row.names = TRUE) # Contributions to the PCs
write.csv(res.var$cos2, file = "PCA_Leaf_vars_qual.csv",
row.names = TRUE) # Quality of representation

# Results for individuals
res.ind <- get_pca_ind(res.pca)
write.csv(res.ind$coord, file = "PCA_Leaf_ind_coord.csv",
row.names = TRUE) # Coordinates
write.csv(res.ind$contrib, file = "PCA_Leaf_ind_contrib.csv",
row.names = TRUE) # Contributions to the PCs
write.csv(res.ind$cos2, file = "PCA_Leaf_ind_qual.csv",
row.names = TRUE) # Quality of representation

#Graph Eigenvalues
png(filename= "Leaf_PCA_Eigen.png",width = 1100,
height = 800, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
fviz_eig(res.pca)
dev.off()

#plot cummulative variance explained per componenet
png(filename = "leaf_cum_variance.png", width = 1100,
height = 800)
d <- get_eigenvalue(res.pca)
#plot the cummultative variance of the first 6 components
barplot(d$cumulative.variance.percent[1:6],
names.arg =row.names(d)[1:6])
dev.off()


#plot cummulative variance explained per componenet
png(filename = "cum_variance_leaf.png", width = 1100,
height = 800)
d <- get_eigenvalue(res.pca)
#plot the cummultative variance of the first 6 components
barplot(d$cumulative.variance.percent[1:6],
names.arg =row.names(d)[1:6])
dev.off()


#Graph Individuals
```

```
png(filename= "Leaf_PCA_Individuals.png",width = 1100,
height = 800, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
fviz_pca_ind(res.pca,
             col.ind = "cos2", # Color by the quality of
             representation
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE,     # Avoid text overlapping
             label = "none"
)
dev.off()

#Graph Variables
png(filename= "Leaf_PCA_Variables.png",width = 1100,
height = 800, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
fviz_pca_var(res.pca,
             col.var = "contrib", # Color by contributions
             to the PC
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE     # Avoid text overlapping
)
dev.off()

#Graph Quantatitive / Catagorical variables
#By Gene group
Gene_group <- as.factor(PCA_Data$Group[1:647])
png(filename= "Leaf_PCA_ANC_GR.png",width = 1100,
height = 800, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
fviz_pca_ind(res.pca,
             col.ind = Gene_group, # color by groups
             addEllipses = TRUE, # Concentration ellipses
             ellipse.type = "t",
             legend.title = "Groups",
             repel = TRUE,
             label = "none",
             title = "Major Clades"
) +
```

```
    theme(legend.key.size = unit(1.5, "cm"),
          legend.text = element_text(size=20))
dev.off()

 #By habitat
 Habitat <- as.factor(PCA_Data$Habitat[1:666])
 png(filename= "Leaf_PCA_Habitat.png",width = 1000,
height = 600, units = "px", pointsize = 12,
     bg = "white",  res = NA,
     type = c("cairo", "cairo-png", "Xlib", "quartz"))
 fviz_pca_ind(res.pca,
              col.ind = Habitat,  color by groups
              addEllipses = TRUE,  Concentration ellipses
              ellipse.type = "confidence",
              legend.title = "Groups",
              repel = TRUE,
              label = "none",
              title = "PCA_Habitat"
 )
 dev.off()

 #Graph Quantatitive / Catagorical variables
 #By Collection
 Collections <- as.factor(PCA_Data$Collection[1:666])
 png(filename= "Leaf_PCA_collections.png",width = 1000,
height = 600, units = "px", pointsize = 12,
     bg = "white",  res = NA,
     type = c("cairo", "cairo-png", "Xlib", "quartz"))
 fviz_pca_ind(res.pca,
              col.ind = Collections,  color by groups
              addEllipses = TRUE,  Concentration ellipses
              ellipse.type = "confidence",
              legend.title = "Groups",
              repel = TRUE,
              label = "none",
              title = "PCA_Collections"
 )
 dev.off()
```

Appendix 10

```
#PCA of Isoetes spore data


library(factoextra)
library(ape)
library(tools)
library(stringr)
library(readr)
library(missMDA)

#Load in Data into R
setwd(#Path to working directory)

PCA_Data <- read_csv(#path to CSV file with the spore data,
             col_types = cols(Plant = col_number(),
                              'Average' = col_number(),
                              'Smallest' = col_number(),
                              'Largest' = col_number()))

PCA_Data
#Set the columns and rows to analyze, where x & y are the range
#of rows and columns to be analyzed in 'i:j' format, where i is
#the first row, and j is the last.
PCA_data.active <- PCA_Data[1:94, 5:7]
PCA_data.active

#Perform the principal components analysis
res.pca <- prcomp(PCA_data.active, scale = TRUE)

# Eigenvalues
eig.val <- get_eigenvalue(res.pca)
write.csv(eig.val, file = "PCA_Eigen.csv", row.names = TRUE)

# Results for Variables
res.var <- get_pca_var(res.pca)
write.csv(res.var$coord, file = "PCA_Spores_vars_coord.csv",
row.names = TRUE) # Coordinates
write.csv(res.var$contrib, file = "PCA_Spores_vars_contrib.csv",
row.names = TRUE) # Contributions to the PCs
write.csv(res.var$cos2, file = "PCA_Spores_vars_qual.csv",
```

```
row.names = TRUE) # Quality of representation

# Results for individuals
res.ind <- get_pca_ind(res.pca)
write.csv(res.ind$coord, file = "PCA_Spores_ind_coord.csv",
row.names = TRUE) # Coordinates
write.csv(res.ind$contrib, file = "PCA_Spores_ind_contrib.csv",
row.names = TRUE) # Contributions to the PCs
write.csv(res.ind$cos2, file = "PCA_Spores_ind_qual.csv",
row.names = TRUE) # Quality of representation

#Graph Eigenvalues
png(filename= "Spore_PCA_Eigen.png",width = 1100,
height = 800, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
fviz_eig(res.pca)
dev.off()

#plot cummulative variance explained per componenet
png(filename = "Cum_variance_spore.png", width = 1100,
height = 800)
d <- get_eigenvalue(res.pca)
#plot the cummultative variance of the first 6 components
barplot(d$cumulative.variance.percent[1:6],
names.arg =row.names(d)[1:6])
dev.off()

#Graph Individuals
png(filename= "Spore_PCA_Individuals.png",width = 1100,
height = 800, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
fviz_pca_ind(res.pca,
             col.ind = "cos2", # Color by the
             #quality of representation
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE,     # Avoid text overlapping
             label = "none"
)
dev.off()
```

```
#Graph Variables
png(filename= "Spore_PCA_Variables.png",width = 1100,
height = 800, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
fviz_pca_var(res.pca,
             col.var = "contrib", # Color by contributions
             #to the PC
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
             repel = TRUE     # Avoid text overlapping
)
dev.off()

#Graph Quantatitive / Catagorical variables
#By Gene group
Gene_group <- as.factor(PCA_Data$Group[1:94])
png(filename= "Spore_PCA_ANC_GR.png",width = 1100,
height = 800, units = "px", pointsize = 10,
    bg = "white",  res = NA,
    type = c("cairo", "cairo-png", "Xlib", "quartz"))
fviz_pca_ind(res.pca,
             col.ind = Gene_group, # color by groups
             addEllipses = TRUE, # Concentration ellipses
             ellipse.type = "t",
             legend.title = "Groups",
             repel = TRUE,
             label = "none",
             title = "Spores Major Clades"
) +
    theme(legend.key.size = unit(1.5, "cm"),
          legend.text = element_text(size=20))
dev.off()

 #By habitat
Habitat <- as.factor(PCA_Data$Habitat[1:94])
 png(filename= "Spore_PCA_Habitat.png",width = 1000,
height = 600, units = "px", pointsize = 12,
     bg = "white",  res = NA,
     type = c("cairo", "cairo-png", "Xlib", "quartz"))
 fviz_pca_ind(res.pca,
              col.ind = Habitat,  color by groups
              addEllipses = TRUE,  Concentration ellipses
```

```
                ellipse.type = "t",
                legend.title = "Groups",
                repel = TRUE,
                label = "none",
                title = "Spores Habitat"
 )
 dev.off()

 #By Collection
 Collections <- as.factor(PCA_Data$Collection[1:94])
 png(filename= "Spore_PCA_collections.png",width = 1000,
height = 600, units = "px", pointsize = 12,
     bg = "white",  res = NA,
     type = c("cairo", "cairo-png", "Xlib", "quartz"))
 fviz_pca_ind(res.pca,
                col.ind = Collections, # color by groups
                addEllipses = TRUE, # Concentration ellipses
                ellipse.type = "t",
                legend.title = "Groups",
                repel = TRUE,
                label = "none",
                title = "Spores Collections"
 )
 dev.off()
```