5-7-2015

Resolving Amphitropical Phylogeographic Histories in the Common Dung Moss Tetraplodon (Bryopsida: Splachnaceae).

Lily Roberta Lewis
University of Connecticut, Lilyrlewis@gmail.com

Resolving Amphitropical Phylogeographic Histories in the Common Dung Moss *Tetraplodon* (Bryopsida: Splachnaceae).

Lily Roberta Lewis, PhD

University of Connecticut, 2015

Many plants have geographic disjunctions, with one of the more rare, yet extreme being the amphitropical, or bipolar disjunction. Bryophytes (namely mosses and liverworts) exhibit this pattern more frequently relative to other groups of plants and typically at or below the level of species. The processes that have shaped the amphitropical disjunction have been infrequently investigated, with notably a near absence of studies focusing on mosses. This dissertation explores the amphitropical disjunction in the dung moss *Tetraplodon*, with a special emphasis on the origin of the southernmost South American endemic *T. fuegianus.* Chapter 1 delimits three major lineages within Tetraplodon with distinct yet overlapping geographic ranges, including an amphitropical lineage containing the southernmost South American endemic *T. fuegianus*. Based on molecular divergence date estimation and phylogenetic topology, the American amphitropical disjunction is traced to a single direct long-distance dispersal event across the tropics. Chapter 2 provides the first evidence supporting the role of migratory shore birds in dispersing bryophytes, as well as other plant, fungal, and algal diaspores across the tropics. Chapter 3 describes the complete chloroplast and mitochondrial genomes and nuclear ribosomal repeat across seven patches of the endemic *T. fuegianus*. Screening of variation within distinct patches of moss revealed inter-individual polymorphism within single patches of moss, and intra-individual variation in the nuclear ribosomal repeat. Chapter 4 employs a RAD-seq approach to sequence thousands of loci across the range of the amphitropical lineage inferred in Chapter 1 allowing for resolution of a monophyletic *T. fuegianus*, which shares an ancestor with populations from northwestern North America. Within the lineage, geographic structure is identified, suggesting a complex phylogeographic history for this group, likely shaped by Pleistocene glaciations in northwestern North America.

Resolving Amphitropical Phylogeographic Histories in the Common Dung Moss *Tetraplodon* (Bryopsida: Splachnaceae).

Lily Roberta Lewis

B.S., University of South Florida, 2007

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2015

APPROVAL PAGE

Doctor of Philosophy Dissertation

Resolving Amphitropical Phylogeographic Histories in the Common Dung Moss *Tetraplodon* (Bryopsida: Splachnaceae).

Presented by

Lily Roberta Lewis, B.S.

Major Advisor _____
Bernard Goffinet

Associate Advisor _____
Ricardo Rozzi

Associate Advisor _____
Paul Lewis

Associate Advisor _____
Kent Holsinger

University of Connecticut
2015

Resolving amphitropical phylogeographic histories in

the common dung moss *Tetraplodon* (Bryopsida: Splachnaceae).

Introduction

Intraspecific intercontinental disjunctions are common in bryophytes. Two major hypotheses stand to explain these disjunctions, vicariance and long distance dispersal (LDD). Various lines of evidence have corroborated a shift in acceptance from the vicariance to the LDD hypothesis (Shaw 2001). Transoceanic disjunctions within "morphological species" have historically been explained as resulting from vicariance based on the assumption that bryophytes species are ancient lineages (Herzog 1926; Schuster 1969). This view has been interpreted to suggest that bryophytes are evolutionarily stagnant, having not undergone allopatric speciation despite severe barriers to gene flow (Crum 1972). The spores of bryophytes are small, and become airborne easily. Analysis of particles in rain water revealed the presence of spores of exotic moss taxa (Petterson 1940), suggesting that bryophytes may not be dispersal limited. Experiments on spore viability after exposure to extreme conditions simulating LDD by wind, suggested that spores of moss species exhibiting transoceanic disjunctions resisted ultralow freezing and other harsh conditions, whereas those of continental endemic species died (van Zanten 1976; van Zanten 1978). Thus spore survival and not dispersal ability may be the limiting factor to long-range expansions. Correlations between floristic similarity for sporic plants (bryophytes and ferns) and wind connectivity among sub-Antarctic oceanic islands provided evidence that distribution patterns are primarily shaped by wind trajectories rather than geographic proximity (Muñoz et al. 2004),

supporting the hypothesis that if spores are resistant, LDD is feasible in the presence of a vector such as wind, which has prevailing west – east or east-west directionalities at equatorial and temperate/subpolar latitudes, respectively.

Molecular dating studies have shown that while bryophytes represent early land plant lineages (Qiu et al. 2006), extant taxa are young relative to the age of vicariant tectonic events (Devos & Vanderpoorten 2009; Shaw et al. 2010; Stenøien et al. 2011). Moreover, genetic diversity appears higher than would be expected from the presumed morphological stability of many bryophytes (Shaw 2000; McDaniel & Shaw 2003; Heinrichs et al. 2009; Feldberg et al. 2010; Kreier et al. 2010). This diversity is often partitioned along geographic gradients, refuting the hypothesis that allopatric speciation, albeit morphologically cryptic, is lacking in bryophytes (Shaw 2001; Heinrichs et al. 2009).

Experimental evidence, wind path and floristic correlations, and interpretation of molecular evidence weaken the basis for the vicariance hypothesis and have led to the current prevailing acceptance of the LDD via wind hypothesis for the explanation of intercontinental disjunctions in plants. Long distance dispersal via wind, however, becomes problematic when considering dispersal limited organisms and disjunctions between areas lacking wind connectivity (Du Rietz 1940; Schuster 1969). Here we challenge the hypothesis of LDD via wind to explain all disjunction patterns by focusing on putatively dispersal limited taxa with an extreme, yet recurrent, disjunction pattern notably lacking wind connectivity. The occurrence of disjunctions that lie outside the current paradigms of LDD suggest that the dispersal pathways and processes shaping our global floras are much more diverse than acknowledged by blanket hypotheses (Raven,

1963). LDD of putatively dispersal limited organisms, or across disjunctions for which dispersal vectors are not known, is arguably the scenario that most commonly leads to allopatric speciation and endemism, due to the presumed rareness of dispersal events, and may thus be significant in shaping microevolutionary processes (i.e., speciation) in seedless land plants.

Study system: Bipolar entomochorous mosses

The dung moss family Splachnaceae is unique among bryophytes in that some of its species rely on insects (entomochory) rather than wind (anemochory) to disperse their spores (Koponen 1990). The fly mediated spore dispersal syndrome includes deep purple or brightly colored capsules, a sticky spore mass, emission of volatile compounds mimicking carrion or feces, and a coprophilous habitat (Koponen 1990). Flies, seeking fresh decaying organic matter to feed, reproduce or lay their eggs on are attracted to the capsules and may inadvertently pick up spores. The moss offers no reward to the insect, and hence visitation time is short (Marino et al. 2009). When the fly lands on a piece of dung the spore may fall off and germinate if the substrate is suitable. The spores of entomochorous species are typically thin walled and green, promoting rapid germination upon fly mediated dispersal, but limiting their long term viability and resistance to ultraviolet radiation (van Zanten 1978). Entomochorous spores are well suited for local dispersal, but not for LDD via wind.

The strictly entomochorous genus *Tetraplodon* (Figure 1) is characterized by a geographic distribution spanning much of Laurasia with disjunctions in high altitude localities in the tropics (Northern Andes, East Africa and SE Asia) and in the southernmost regions of South America. The lack of wind connectivity between trans-

equatorial and particularly between extreme bipolar localities precludes the possibility of LDD via wind as an explanation for this pattern in *Tetraplodon*, as well as other bipolar disjunct taxa, including at least 60 species of moss (Norris et al. 1999; Ochyra & Buck 2003; Ochyra et al. 2008). Despite this limitation, gene flow between bipolar disjunct populations occurs across the tropics in various groups plants (Myllys et al. 2003; Gussarova et al. 2008; Wirtz et al. 2008; Kreier et al. 2010; Marcial et al. 2010; Popp et al. 2011; Piñeiro et al. 2012). Popp et al. (2011) argued that migratory birds would be the vector between antipodal populations, however, evidence to support this hypothesis, is limited, and comes primarily from evidence of dispersal over shorter distances (Bailey & James 1979).

Figure 1 *Tetraplodon* grows on carnivore dung. Upon dispersal by flies, the plants rapidly colonize the fresh substrate.

The following chapters address the history of amphitropicality in *Tetraplodon.* Chapter 1 delimits three major lineages with distinct yet overlapping geographic ranges, including an amphitropical lineage containing the southernmost South American endemic *T. fuegianus*. Based on molecular divergence date estimation and phylogenetic topology, the American amphitropical disjunction is traced to a single direct long-distance dispersal event across the tropics. Chapter 2 provides the first evidence supporting the role of migratory shore birds in dispersing bryophytes, as well as other plant, fungal, and algal diaspores across the tropics. Chapter 3 describes the complete chloroplast and mitochondrial genomes and nuclear ribosomal repeat across seven patches of the endemic *T. fuegianus*. Screening of variation within distinct patches of moss revealed inter-

individual polymorphism within single patches of moss, and intra-individual variation in the nuclear ribosomal repeat. Chapter 4 employs a restriction site associated DNA sequencing RAD-seq approach to identify structure within the amphitropical lineage inferred in Chapter 1 and identify northwestern North America as the source of the ancestor to *T. fuegianus*.

References

Bailey, R. & P. James. 1979. Birds and the dispersal of lichen propagules. *Lichenologist* 11: 105–106.

Crum H. A. 1972. The geographic origins of the mosses of North America's eastern deciduous forest. *Journal of the Hattori Botanical Laboratory* 35: 269–298.

Du Rietz G. E. 1940. Problems of bipolar plant distribution. *Acta Phytogeographica Suecica* 13: 215–282.

Devos N. & A. Vanderpoorten. 2009. Range disjunctions, speciation, and morphological transformation rates in the liverwort genus *Leptoscyphus*. *Evolution* 63: 779–92.

Gussarova, G., M, Popp, E. Vitek & C. Brochmann. 2008. Molecular phylogeny and biogeography of the bipolar *Euphrasia* (Orobanchaceae): recent radiations in an old genus. *Molecular Phylogenetics and Evoloution* 48: 444–460.

Heinrichs, J., J. Hentschel, K. Feldberg, A. Bombosch & H. Schneider. 2009. Phylogenetic biogeography and taxonomy of disjunctly distributed bryophytes. *Journal of Systematics and Evolution* 47: 497–508.

Herzog T. 1926. Geographie der Moose. Jena: Fischer.

Koponen, A. 1990. Entomophily in the Splachnaceae. *Botanical Journal of the Linnean Society* 104: 115–127.

Kreier, H. P., K. Feldberg, F. Mahr, A. Bombosch, A. R. Schmidt, R. L. Zhu, M. von Konrat, B. Shaw, A. J. Shaw & J. Heinrichs. 2010. Phylogeny of the leafy liverwort *Ptilidium*: Cryptic speciation and shared haplotypes between the Northern and Southern Hemispheres. *Molecular Phylogenetics and Evolution* 57: 1260–1267.

Marcial, E., V. Virginia, V. Pablo & L. Modesto. 2010. Bipolar disjunctions in *Carex*: long-distance dispersal, vicariance, or parallel evolution? *Flora* 205: 118–127.

Marino P., R. Raguso, B. Goffinet. 2009. The ecology and evolution of fly dispersed dung mosses (Family Splachnaceae): Manipulating insect behaviour through odour and visual cues. *Symbiosis* 47: 61–76.

Muñoz, J., A. M. Felicisimo, F. Cabezas, A. R. Burgaz & I. Martinez. 2004. Wind as a long-distance dispersal vehicle in the Southern hemisphere. *Science* 304: 1144–1147.

Myllys, L., S. Stenroos, A. Thell & T. Ahti. 2003. Phylogeny of bipolar *Cladonia arbuscula* and *Cladonia mitis* (Lecanorales, Euascomycetes). *Molecular Phylogenetics and Evolution* 27: 58–69.

Norris, D. H., T. Koponen & S. Piippo. 1999. Bryophyte flora of the Huon Peninsula, Papua New Guinea. LXVI Meesiaceae (Musci), with lists of boreal to temperate disjunct, bipolar, and widely spread species in New Guinea. *Annales Botanici Fennici* 36: 257–263.

Ochyra, R. & W. R. Buck. 2003. *Arctoa fulvella*, New to Tierra del Fuego, with notes on trans-American bipolar bryogeography. *Bryologist* 106: 532–538.

Ochyra R, L. Smith & H. Bednarek-Ochyra. 2008. The Illustrated Moss Flora of Antarctica. New York: Cambridge University Press.

Petterson B. 1940. Experimentelle Untersuchungen über die euanemochore Verbreitung der Sporenpflanzen. *Acta Bot. Fenn.* 25: 1–102.

Popp M., V. Mirré & C. Brochmann. 2011. A single Mid-Pleistocene long-distance dispersal by a bird can explain the extreme bipolar disjunction in crowberries (*Empetrum*). *Proceedings of the National Academy of Sciences of the United States of America* 108: 6520–6525.

Qiu Y-L, L. Li, B. Wang, Z. Chen, V. Knoop, M. Groth-Malonek, O. Dombrovska, J. Lee, L. Kent, J. Rest, et al. 2006. The deepest divergences in land plants inferred from phylogenomic evidence. *Proceedings of the National Academy of Sciences of the United States of America* 103: 15511–6.

Raven P. H. 1963. Amphitropical Relationships in the Floras of North and South America. *The Quarterly Review of Biology* 38: 151–177.

Du Rietz G. E. 1940. Problems of bipolar plant distribution. *Acta Phytogeographica Suecica* 13: 215–282.

Schuster R. M. 1969. Problems of antipodal distribution in lower land plants. *Taxon* 18: 46–91.

Shaw A. J., N. Devos, C. J. Cox, S. B. Boles, B. Shaw, A. M. Buchanan, L. Cave & R. Seppelt. 2010. Peatmoss (*Sphagnum*) diversification associated with Miocene Northern Hemisphere climatic cooling? *Molecular Phylogenetics and Evolution* 55: 1139–45.

Shaw A. J. 2001. Biogeographic patterns and cryptic speciation in bryophytes. *Journal of Biogeography* 28: 253–261.

Wirtz N., C. Printzen & H. T. Lumbsch. 2008. The delimitation of Antarctic and bipolar species of neuropogonoid *Usnea* (Ascomycota, Lecanorales): a cohesion approach of species recognition for the *Usnea perpusilla* complex. *Mycological Research* 112: 472–484.

van Zanten B. O. 1976. Preliminary report on germination experiments designed to estimate the survival chances of moss spores during aerial trans-oceanic long-range dispersal in the southern hemisphere, with particular reference to New Zealand. *Journal of the Hattori Botanical Laboratory* 41: 133–140.

van Zanten B. O. 1978. Experimental studies on trans-oceanic long-range dispersal of moss spores in the Southern Hemisphere. *Journal of the Hattori Botanical Laboratory* 44: 455–482.

Wiley Global Permissions <permissions@wiley.com> Fri, Apr 10, 2015 at 5:33 AM

To: Lily Lewis
Cc: Foster, Louise - Oxford

Dear Lily Lewis

Thank you for your request.

Permission is granted for you to use the material requested for your thesis/dissertation subject to the usual acknowledgements and on the understanding that you will reapply for permission if you wish to distribute or publish your thesis/dissertation commercially. You must also duplicate the copyright notice that appears in the Wiley publication in your use of the Material.

Permission is granted solely for use in conjunction with the thesis, and the material may not be posted online separately.

Any third party material is expressly excluded from this permission. If any of the material you wish to use appears within our work with credit to another source, authorisation from that source must be obtained.

Yours sincerely

Duncan James

Associate Permissions Manager

John Wiley & Sons Ltd

The Atrium

Southern Gate, Chichester

West Sussex, PO19 8SQ

UK

Journal of Biogeography (J. Biogeogr.) (2014)

ORIGINAL ARTICLE

Direct long-distance dispersal shapes a New World amphitropical disjunction in the dispersal-limited dung moss *Tetraplodon* (Bryopsida: Splachnaceae)

Lily R. Lewis[1]*, Ricardo Rozzi[2,3] and Bernard Goffinet[1]

[1]*Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269, USA,* [2]*Omora Ethnobotanical Park, Institute of Ecology and Biodiversity, and Universidad de Magallanes, Puerto Williams, Antarctic Province, Chile,* [3]*Department of Philosophy, University of North Texas, Denton, TX 76201, USA*

*Correspondence: Lily R. Lewis, Ecology and Evolutionary Biology, University of Connecticut, 75 N Eagleville Road, U-3043, Storrs, CT 06269, USA.
E-mail: Lily.Lewis@uconn.edu

ABSTRACT

Aim Many intercontinental disjunctions, especially among spore-producing plants, are shaped by long-distance dispersal (LDD) via wind currents. Amphitropical disjunctions are most commonly explained through LDD, but other vectors and dispersal scenarios must also be considered. To interpret the New World amphitropical disjunction in the dung-moss genus *Tetraplodon*, we compared stepwise migration along the Andes, direct LDD and ancient vicariance.

Location Global, specifically high-latitude and high-elevation localities, with a focus on the New World.

Methods Phylogenetic relationships were inferred from four loci sampled from 124 populations representing the global range of *Tetraplodon*, and analysed using maximum-likelihood and Bayesian optimality criteria, with divergence dates estimated in BEAST.

Results The monophyletic *T. mnioides* complex diversified between the early Miocene and early-to-mid Pliocene into three well-supported clades, each with a unique geographical distribution: Laurasian, primarily high-elevation tropical, and amphitropical. Populations from southernmost South American were reconstructed as a monophyletic lineage that diverged from high-latitude Northern Hemisphere populations around 8.63 Ma [95% highest posterior density (HPD) 3.07–10.11 Ma].

Main conclusions Direct LDD has resulted in the American amphitropical disjunction in *Tetraplodon*. A lack of modern or historical wind connectivity between polar regions and the poor resistance of *Tetraplodon* spores to the conditions associated with wind-dispersal suggest that bird-mediated LDD provides the best explanation for the establishment of amphitropicality.

Keywords
Amphitropical, bipolar, bryophyte, disjunctions, long-distance dispersal, migratory shorebird, New World, *Tetraplodon*.

INTRODUCTION

Intercontinental disjunctions are common among land plants and may be shaped by vicariance (Raven & Axelrod, 1974), dispersal (Nathan, 2006) or a combination of both processes (Cook & Crisp, 2005). Molecular tools and dating approaches have provided increasing support for the significance of dispersal in shaping modern disjunctions (de Queiroz, 2005; Heinrichs *et al.*, 2009). In both the Northern and Southern Hemispheres, wind serves as an important vector for long-distance dispersal (LDD; Muñoz *et al.*, 2004; Wilkinson *et al.*, 2012). Disjunctions between antipodal high latitudes (i.e. amphitropical disjunctions), and in some cases low-latitude, high-elevation localities, have also been largely shaped by LDD, according to inferences based on molecular phylogenetic topologies (see Wen & Ickert-Bond, 2009, for review) and dating (Gussarova *et al.*, 2008; Popp *et al.*, 2011; Fernández-Mendoza & Printzen, 2013). An absence of wind connectivity across the equator due to the Intertropical Convergence Zone (Hyeong *et al.*, 2005) has led to the proposal

http://wileyonlinelibrary.com/journal/jbi
doi:10.1111/jbi.12385

Journal of Biogeography

of alternative vectors to wind, such as migratory birds (Popp *et al.*, 2011), or complex dispersal scenarios involving stepwise migration along tropical highland bridges (Raven, 1963; Fernández-Mendoza & Printzen, 2013).

Du Rietz (1940) noted that amphitropicality is most common among bryophytes and lichens. Current estimates include at least 66 species of moss (Norris *et al.*, 1999; Ochyra & Buck, 2003; Ochyra & Bednarek-Ochyra, 2008), 24 species of liverwort (Schuster, 1983; Streimann, 1998; Norris *et al.*, 1999; Bednarek-Ochyra *et al.*, 2000) and 160 species of lichen (representing 41.4% of the lichen flora) from Antarctica and South Georgia (Øvstedal & Lewis Smith, 2001) as being disjunct across the tropics. Despite the frequency of this extreme pattern in bryophytes and lichens, relatively few studies have applied molecular phylogenetic tools to assess the processes that underlie the amphitropical distribution in bryophytes (Kreier *et al.*, 2010; Piñeiro *et al.*, 2012) and lichens (Myllys *et al.*, 2003; Wirtz *et al.*, 2008), with only a single study – on the lichen *Cetraria aculeata* – employing molecular dating approaches (Fernández-Mendoza & Printzen, 2013).

Unlike the vast majority of bryophytes, approximately half of the species in the dung-moss family Splachnaceae rely on insects (entomochory) rather than wind (anemochory) to disperse their spores (Koponen, 1990). The fly-mediated spore-dispersal syndrome includes deep purple or brightly coloured capsules, the emission of volatile compounds that mimic carrion or faeces, the production of sticky spores, and coprophily (Koponen, 1990). Flies, seeking fresh decaying organic matter on which to feed, reproduce or lay their eggs, are attracted to the capsules through olfactory and visual cues and may inadvertently pick up spores during their visit. The spores of entomochorous species rapidly germinate after they fall off the fly onto fresh dung. Although entomochorous species are well adapted to efficient dispersal locally and perhaps regionally via insect vectors (Marino, 1988a,b), their sticky spores may severely limit their ability to be dispersed by wind currents (Cameron & Wyatt, 1986). Because the spores are thin-walled, they are vulnerable to desiccation, freezing and UV radiation, and are thus unlikely to survive dispersal via high atmospheric winds (van Zanten, 1978).

Among entomochorous Splachnaceae, *Tetraplodon* Bruch & Schimp. (Fig. 1) displays the broadest geographical range and most extreme disjunctions. Its distribution spans Laurasia, with disjunct populations in Central Africa, Borneo, Papua New Guinea and, within the New World, in the northern Andes, south-eastern Brazil and southernmost South America. The populations from southernmost South America are accommodated under the putative endemic *Tetraplodon fuegianus* Besch., whereas those from south-eastern Brazil are treated as *Tetraplodon itatiaiae* Müll. Hal. All other tropical high-elevation populations are considered to belong to the widespread *Tetraplodon mnioides* (Hedw.) Bruch & Schimp. or *Tetraplodon urceolatus* (Hedw.) Bruch & Schimp. The identification of South American putative endemics is based on geography, with no known diagnostic morphological traits. The broad phenotypic variation in *T. mnioides* (Steere, 1977) and the uncertain status of *T. urceolatus* (Frisvoll, 1978) have confounded taxonomy within the genus, limiting evolutionary inferences within the group. In the absence of robust unambiguous morphological species, phylogenetic delimitation of lineages is necessary to reconstruct the phylogeographical history leading to the origin of the southernmost populations, and thus the New World disjunctions.

Several processes may have contributed to the amphitropical disjunctions observed within *Tetraplodon*. If vicariance resulting from the breakup of Pangaea shaped the disjunctions, the ages of lineages are expected to correlate with geological events, but only disjunctions above the generic level have to date been associated with ancient tectonic events (Mao *et al.*, 2012). If divergence times post-date continental movements, the New World amphitropical disjunction may result from either stepwise migration (i.e. a series of dependent dispersal events along the Andes) or direct LDD across the tropics (Nathan, 2006; Popp *et al.*, 2011). If stepwise dispersal has occurred from north to south via the Andes, we would expect northern Andean populations to mark intermediate dispersal events, and share a unique common ancestor with southern South American populations.

The detection of historical stepwise migration relies on the presence of species in the fossil record or extant flora of low-latitude intermediate regions. Without intermediate populations, it may not be possible to discriminate between stepwise and direct dispersal events. Fernández-Mendoza & Printzen (2013) suggested that stepwise migration along the Andes played a role in the trans-tropical range expansion of the lichen *Cetraria aculeata*. Gussarova *et al.* (2008) inferred stepping-stone migration across Malaysia as the process leading to amphitropical distributions in *Euphrasia.* Few studies addressing amphitropical disjunctions have focused on species with intermediate low-latitude populations, thus precluding differentiation between stepwise migration and direct LDD (Myllys *et al.*, 2003; Escudero *et al.*, 2010; Popp *et al.*, 2011), focusing rather on the rejection of ancient vicariance in favour of LDD *sensu lato*. Although the role of LDD *sensu lato* is being increasingly recognized as shaping amphitropical disjunctions at the infraspecific and infrageneric levels (see Wen & Ickert-Bond, 2009, for review), more studies are needed to disentangle stepwise and direct LDD. The New World distribution of *Tetraplodon*, with intermediate low-latitude, high-elevation populations in the northern Andes, suggests that stepwise migration may have played a role in the establishment of the New World amphitropical disjunction. *Tetraplodon* offers an opportunity to discriminate between the relative importance of stepwise and direct dispersal in establishing amphitropical disjunctions in mosses, a group of plants with many bipolar disjunct species but for which the phylogeographical history has never been reconstructed within an explicit time-calibrated evolutionary scenario.

Based on variation in four discrete loci, we seek (1) to define phylogenetic lineages within *Tetraplodon* and their

Figure 1 *Tetraplodon* (left centre) grows on dung, most commonly of carnivores, or on decaying carcasses, typically in open environments. Sporophytes (right) are a deep red–purple colour and produce sticky masses of bright yellow–green spores.

geographical ranges and (2) to estimate maximum divergence dates for major lineages and the New World amphitropical disjunction, in order (3) to assess the roles of stepwise and direct dispersal events, as well as ancient vicariance, in the origin of the New World amphitropical disjunction in *Tetraplodon*.

MATERIALS AND METHODS

Sampling, PCR amplification and sequencing

The samples we used represent the complete taxonomic and geographical ranges of *Tetraplodon* as well as an undescribed cleistocarpous taxon known only from Bhutan, included based on preliminary results suggesting affinities with *Tetraplodon*, and the genus *Voitia* Hornsch., which may be the sister genus to *Tetraplodon* (Goffinet *et al.*, 2004). Total genomic DNA was extracted from 128 accessions with Nucleospin Plant II Kit (Macherey-Nagel, Bethlehem, PA, USA) following the manufacturer's guidelines or using a modified CTAB protocol (Goffinet *et al.*, 1998). Three species of *Tayloria* subgenus *Orthodon* (R. Br.) Broth., and *Neomeesia paludella* (Besch.) Deguchi from the sister family Meesiaceae, were chosen as outgroups following Goffinet *et al.* (2004).

Three chloroplast loci (*atp*B–*rbc*L, *rps*4 and *trn*G) and one nuclear locus (ITS2) were targeted based on amplification success and informativeness at infraspecific and infrageneric levels (Stech & Quandt, 2010). Loci were sequenced on an ABI3100 Genetic Analyzer (Applied Biosystems, Grand Island, NY, USA), and manually edited in Sequencher 4.8 (Gene Codes Corporation, Ann Arbor, MI, USA). All sequences were deposited in GenBank (see Appendix S1 in Supporting Information). Sequences were aligned with Clustal W2 (http://www.ebi.ac.uk/) and manually edited in Mesquite 2.75 (Maddison & Maddison, 2011). Regions of ambiguous alignment for *trn*G, ITS2 and *atp*B–*rbc*L were identified and removed using the Gblocks server 0.91b, with settings allowing for smaller final blocks, gap positions within the final blocks, and less strict flanking positions (Castresana, 2000). Indels in the sequenced *atp*B–*rbc*L intergenic spacer were coded in SeqState 1.4.1 according to simple indel coding (SIC; Simmons & Ochoterena, 2000) after excluding ambiguous regions.

Phylogenetic analyses

Models of nucleotide evolution for each locus (Table 1) were chosen with jModelTest 2.1.3 (Posada, 2008) under the Bayesian information criterion (BIC) and corrected Akaike information criterion (AIC_c),with likelihood scores calculated from base trees optimized by maximum likelihood (ML) using the nearest-neighbour interchange search algorithm. The model of evolution for SIC indels followed Lewis (2001), implemented in garli 2.0 (Zwickl, 2006) as the standard variable model for ML analyses, and in MrBayes 3.2.1 (Ronquist & Huelsenbeck, 2003) as the standard discrete model with no transition-rate asymmetry across sites for Bayesian analyses. Indels were included only in single-locus *atp*B–*rbc*L analyses, with *atp*B–*rbc*L also analysed independently of coded indels.

Maximum-likelihood and bootstrap replicate analyses were performed in garli 2.0 (Zwickl, 2006). One hundred replicate searches were completed for each locus, with 2000 bootstrap replicates. Best trees with 50% majority-rule consensus bootstrap scores from each locus were visually inspected to identify conflicts in topology and bootstrap support values

Table 1 Sample size (*n*), total number of characters (chars) included in analyses, proportion of parsimony informative characters (PI chars), and model of molecular evolution used for each locus based on both corrected Akaike information criterion and Bayesian information criterion. Concatenated data sets were partitioned and modelled according to locus for final analyses. Samples represent the complete taxonomic and geographical ranges of *Tetraplodon*.

Locus	*n*	Chars (bp)	PI chars (%)	Model
*atp*B–*rbc*L	122	616	16	GTR + Γ
*rps*4	124	694	12	GTR + Γ
*trn*G	128	550	12	HKY + Γ
ITS2	113	417	14	HKY + Γ

between loci. No conflicts with higher than 70% bootstrap support were identified and loci were concatenated to form a 2277-bp final data set, excluding *atp*B–*rbc*L indels. The *atp*B–*rbc*L locus was analysed separately including coded indels to explore the phylogenetic significance of indels. Consistency indices based on *atp*B–*rbc*L, including coded indels, were calculated in PAUP* 4.0a129 (Swofford, 2003) under both ACCTRAN and DELTRAN parsimony optimization methods, based on the ML tree inferred from *atp*B–*rbc*L to identify non-homoplasious characters. The concatenated data set was partitioned by locus, with parameter estimates unlinked between partitions, and analysed as described above. Bootstrap values were mapped onto the best ML topology in the SumTrees program, part of the DendroPy-3.7.1 Python library (Sukumaran & Holder, 2010).

Bayesian analyses and calculation of clade posterior probabilities were carried out in MrBayes 3.2.1 (Ronquist & Huelsenbeck, 2003). Three chains were run independently twice for each of the four loci, and for *atp*B–*rbc*L with SIC indels, for 10^6 generations, discarding the first 10% of trees as burn-in. The 50% majority-rule consensus trees were visually inspected to identify conflicts in topology and posterior probabilities between loci. No conflicts with posterior probability greater than 0.95 were identified. The concatenated data set was partitioned by locus and run with three chains twice for 10^6 generations each, with parameter estimates unlinked between partitions. The effective sample sizes for all estimated parameters for all runs were visually checked in Tracer 1.5 (Rambaut & Drummond, 2007) to verify convergence of each analysis. Support values were compared across optimality criteria by mapping the ML bootstrap support values onto the 50% majority-rule consensus Bayesian topology in SumTrees (Fig. 2).

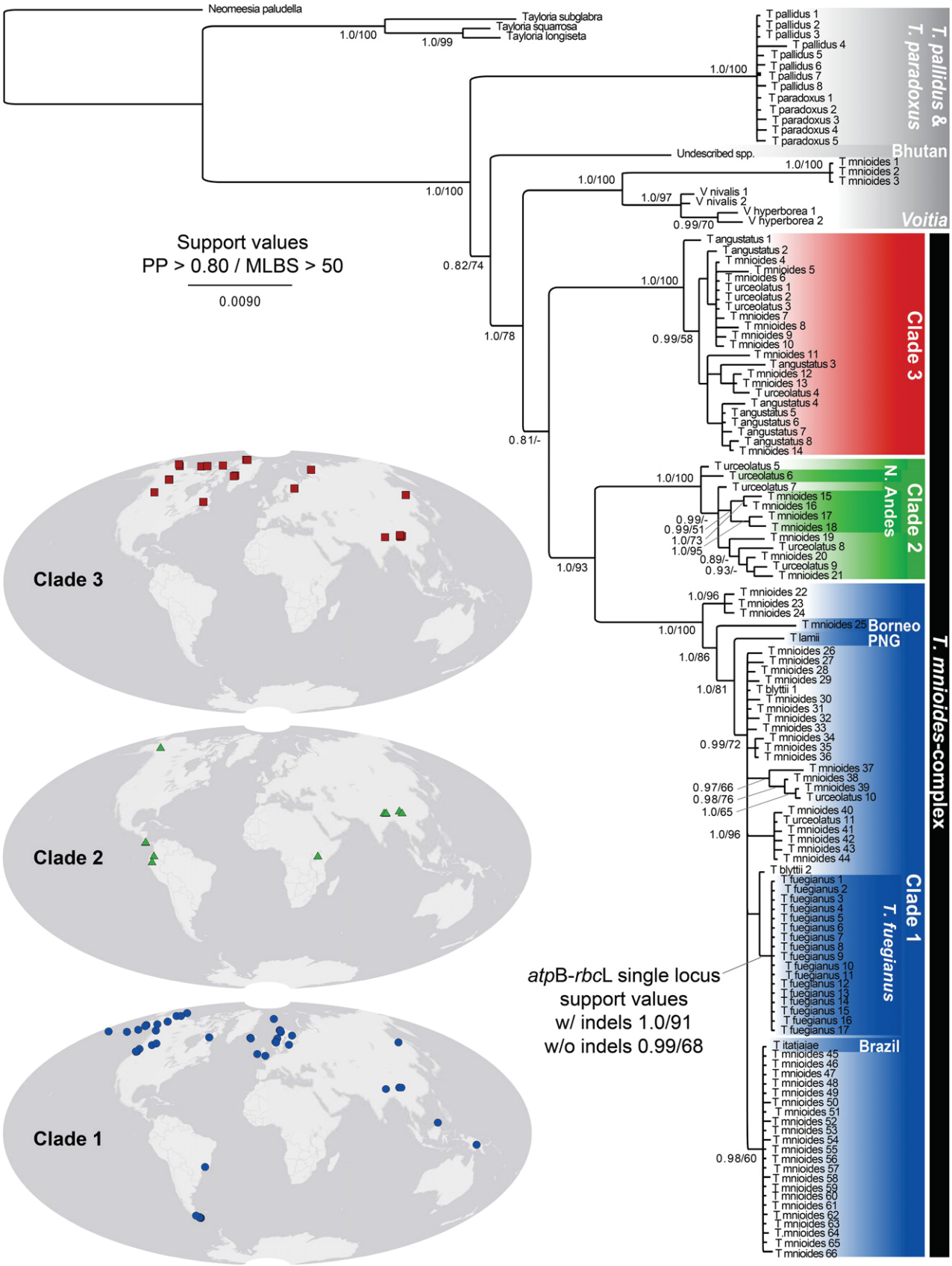
Molecular dating

beast 1.7.5 (Drummond *et al.*, 2012) was used to estimate divergence dates for *Tetraplodon* lineages based on chloroplast loci, excluding SIC indels. A lack of fossils or geological events that would be appropriate for use as calibration points within the Splachnaceae or Splachnales necessitates the use of a defined rate of substitution to tease apart substitution rates and time along branches. A mean chloroplast nucleotide substitution rate of 5×10^{-4} substitutions per site per million years has been used in previous molecular dating studies in mosses (Shaw *et al.*, 2010; Pokorny *et al.*, 2011) when fossil-based calibration was not possible or for testing alternative dating schemes. This rate is based on estimates derived from tracheophyte chloroplast coding regions (Sanderson, 2002), and it is well documented that substitution rates vary, sometimes dramatically, across lineages (Sanderson, 2002). Villarreal & Renner (2012) recovered the same rate *de novo* for hornwort *rbc*L sequences in a fossil-calibrated divergence dating analysis, suggesting that this rate may characterize a broader range of groups than simply tracheophytes. Furthermore, non-coding genomic regions typically undergo faster rates of substitution than coding regions. We expect the published estimate of chloroplast substitution rates based on tracheophyte exons and the *rbc*L region in hornworts to be much slower than the actual, and as yet unknown, rate of substitution in non-coding regions in *Tetraplodon*. Analyses using this rate are thus anticipated to inflate divergence-time estimates and will be used here to identify hypotheses for estimates of maximum divergence dates.

For all beast analyses, clock and tree models were linked across partitions, and models of substitution (Table 1) were unlinked across loci. Strict and uncorrelated log-normal relaxed clocks (Drummond *et al.*, 2006) were both tested, with two speciation tree models: Yule and birth–death process (Gernhard, 2008). A normal distribution with a mean of 5×10^{-4} substitutions per site per million years, and a standard deviation of 1×10^{-4} (20% of the mean rate; Huttunen *et al.*, 2008) was set as the prior distribution of the substitution rate.

All analyses were run twice with three chains for 50,000,000 generations each. Parameter values were sampled

Figure 2 Bayesian consensus tree, showing posterior probabilities (PP) and maximum-likelihood bootstrap support values (MLBS) from a concatenated data set comprising the loci *atp*B–*rbc*L, *trn*G, *rps*4 and ITS2, representing the complete taxonomic range and global distribution of *Tetraplodon*. Support values are shown for nodes with values greater than 0.80 for PP and 50% for MLBS. Support values from *atp*B–*rbc*L single-locus analyses with and without coded indels included are also given for the *T. fuegianus* clade. Clades 1–3 are referred to as the *T. mnioides* complex. All Southern Hemisphere and tropical high-elevation accessions are highlighted and labelled (abbreviations: PNG, Papua New Guinea; N. Andes, northern Andes). The maps show the distinct global distributions for each clade in the *T. mnioides* complex, based on the sampling. Accession data can be found in Appendix S1 in Supporting Information.



every thousand generations, with the first 10% of trees discarded as burn-in. All analyses were run without data (data-free), sampling only from the prior, and posterior marginal densities were compared in Tracer 1.5 with those of analyses that included data, in order to determine the influence of the priors on posterior estimates. The effective sample sizes of all estimated parameters were checked in Tracer 1.5 to ensure values were greater than 200. Tree and log files from duplicate runs were combined using LogCombiner 1.7.5 (Drummond *et al.*, 2012) for converged runs; all duplicate runs converged. Bayes factors (Kass & Raftery, 1995) were calculated in Tracer 1.5 based on log-likelihood scores, to compare the fit of the different combinations of clock and tree models (Table 2). The sampled trees were summarized in TreeAnnotator 1.7.5 (Drummond *et al.*, 2012) and viewed in FigTree 1.3.1 (available at: http://tree.bio.ed.ac.uk/software/figtree/).

RESULTS

Three well-supported clades form what will be referred to as the *T. mnioides* complex (Fig. 2), and together include nearly all accessions that fall under the morphological or geographical species concepts of *T. mnioides*, *T. angustatus*, *T. urceolatus*, *T. blyttii*, *T. itatiaiae*, *T. lamii* and *T. fuegianus*. Each of the three clades in the *T. mnioides* complex has a different geographical distribution, with areas of sympatry in Alaska (clades 1–3), the Himalayas (clades 1–3) and northern Europe (clades 1 and 3) (Fig. 2). Clade 3 is strictly Laurasian, spanning northern high-latitude localities and southwards to the Rocky Mountains and Himalayas (Fig. 2). Clade 2 is primarily tropical high-elevation, with localities in the northern Andes, East Africa and the Himalayas, except for one sample collected near Denali, Alaska. Clade 1 is characterized globally by an amphitropical distribution, with low-latitude, high-elevation localities in Southeast Asia (Mount Kinabalu, Borneo, and Mount Giluwe, Papua New Guinea) and south-eastern Brazil. All populations sampled from narrow endemic taxa (i.e. *T. blyttii*, *T. fuegianus* and *T. itatiaiae*) belong to clade 1.

Tetraplodon fuegianus is reconstructed as a unique lineage nested within a well-supported amphitropical clade (Fig. 2: Clade 1). Single-locus analyses revealed that support for the monophyly of *T. fuegianus* is based on three non-homoplasious (consistency index of 1.0) apomorphic transversions (A to C) and two indels (single-A deletions) in the *atp*B–*rbc*L intergenic spacer. Support for its monophyly, however, is low based on the concatenated data set (posterior probability, PP 0.51; maximum-likelihood bootstrap, MLBS 28%), suggesting that homoplasious characters may be present in the other loci, and based on MLBS of the *atp*B–*rbc*L data set without indels (PP 0.99; MLBS 69%), a difference in support values probably associated with the reliance of this relationship on three non-homoplasious apomorphic transversions. The unambiguous placement of *T. fuegianus* within the amphitropical Clade 1, and its closer relationship to Northern Hemisphere populations than those from the northern Andes, which are resolved with high support in Clade 2, is robustly supported (Fig. 2). *Tetraplodon itatiaiae* is resolved within the amphitropical Clade 1, distinct from the *T. fuegianus* lineage (Fig. 2).

Table 2 Bayes factors ($\log_{10}$) calculated from log-likelihood scores in Tracer 1.5 were used to compare the fit of different clock (strict; relaxed) and tree (Yule; birth–death, BD) model combinations employed in beast 1.7.5 for the estimation of divergence times. The combination of a relaxed molecular clock with a Yule speciation prior provided the best fit to the concatenated data set (*atp*B–*rbc*L, *trn*G and *rps*4) representing the complete taxonomic and geographical range of *Tetraplodon*.

Model	Strict BD	Strict Yule	Relaxed BD
Relaxed Yule	6.384	18.87	1.945
Strict BD	—	12.486	−4.439
Strict Yule	—	—	−16.925

Accessions of *T. pallidus* and *T. paradoxus* are not differentiated from one another but form a robust clade sister to all other ingroup exemplars. *Voitia* is nested within *Tetraplodon* and sister to three interior Alaskan populations phenotypically aligned with the morphological species concept for *T. mnioides* (*T. mnioides* 1–3; Fig. 2, Appendix S1). Accessions sister to *Voitia* were resampled, extracted and sequenced in order to confirm this result. Soft incongruence in the topologies of both trees (Figs 2 & 3) reflect the unresolved placement of the *Voitia* lineage and the undescribed cleistocarpous taxon from Bhutan. *Tetraplodon*, *Voitia* and the sample from Bhutan compose a maximally supported monophyletic group. All major lineages are resolved with high support, although their relationships relative to each other remain ambiguous.

A relaxed clock with Yule speciation prior was the model that best fitted the data (Table 2). The lineages of the *T. mnioides* complex, along with the *Voitia* lineage, diverged from the ancestor to *T. paradoxus* and *T. pallidus* between the mid-Miocene and the early Oligocene, 11.2–33.4 million years ago (Ma) (average 26.6 Ma). The three major lineages of the *T. mnioides* complex (clades 1–3) began to diversify within largely overlapping time periods between the early Miocene and the early-to-mid Pliocene, with mean estimates of 12.6 Ma (95% highest posterior density, HPD, 4.7–17.7 Ma) for Clade 1, 7.9 Ma (95% HPD 3.6–13.8 Ma) for Clade 2, and 10.1 Ma (95% HPD 4.7–17.7 Ma) for Clade 3. The clade comprising *T. fuegianus* is estimated to have diverged from Laurasian populations between the late Miocene to Pliocene, with a mean estimate of 8.63 Ma (95% HPD 3.07–10.11 Ma) and to have diversified between the late Miocene and the late Pleistocene, with a mean estimate of 5.2 Ma (95% HPD 1.4 - 7.0 Ma) (Fig. 3).

DISCUSSION

Phylogenetic inferences from discrete loci reveal that the *Tetraplodon mnioides* complex comprises three robustly

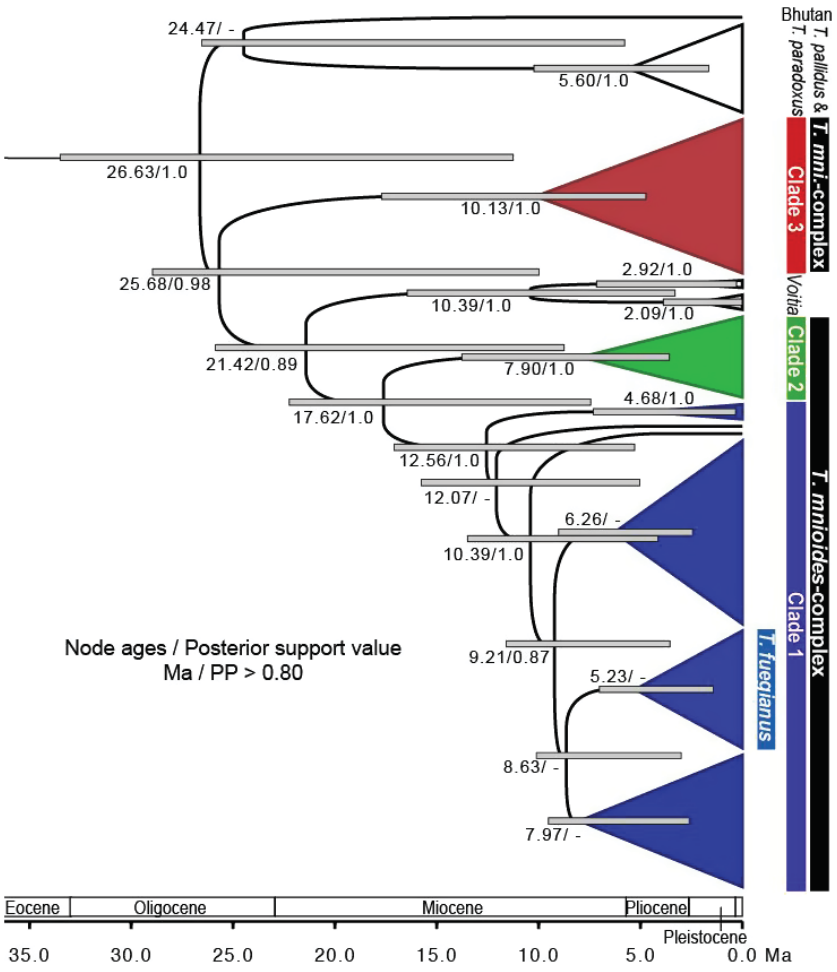


Figure 3 Chronogram showing mean divergence-date estimates of *Tetraplodon* in millions of years and posterior probabilities (node age/PP) for major lineages as estimated in BEAST 1.7.5 from a concatenated data set (*atp*B–*rbc*L, *trn*G and *rps*4) with a relaxed clock and Yule speciation prior with a fixed mean rate of substitution of 5×10^{-4} substitutions per site per million years. Samples represent the complete taxonomic and geographical ranges of *Tetraplodon*. Posterior support values are shown for nodes with values greater than 0.80. Node bars indicate 95% highest posterior density intervals around mean estimates. Labels used to identify clades correspond to those used in Fig. 1.

supported clades (Fig. 2) encompassing most of the diversity within the genus. These lineages display broad geographical distributions (Fig. 2): exclusively Laurasian (clade 3), low-latitude high-elevation with occurrences in Alaska and the Himalayas (clade 2), and Laurasian with disjunctions into southern South America, Brazil, Mount Kinabalu in Borneo and Mount Giluwe in the Southern Highlands of Papua New Guinea (amphitropical; clade 1 in Fig. 2). All three clades are sympatric in Alaska and in the Himalayas. Northern Andean and southern South American *T. fuegianus* populations are resolved within distinct lineages, with each being more closely related to phylogenetically distinct but sympatric Northern Hemisphere high-latitude populations than they are to each other. The Brazilian *T. itatiaiae* is a member of the amphitropical clade 1, is closely aligned with Laurasian *T. mnioides* populations, and does not share a unique common ancestor with populations of *T. fuegianus*. The clades in the *T. mnioides* complex do not match taxa under the morphological species concept (Lawton, 1971; Nyholm, 1975; Frisvoll, 1978; Chien & He, 1999; Marino, 2009), an incongruence which is unsurprising given the broad phenotypic variation in *T. mnioides* (Steere, 1977), *T. angustatus* (Frisvoll, 1978) and *T. urceolatus* (Frisvoll, 1978). Our phylogenetic results highlight the variability in specific traits used to distinguish species of *Tetraplodon* and the need for a critical taxonomic revision of this group.

The resolution of northern Andean and southernmost South American *T. fuegianus* populations in distinct clades (Fig. 2) is inconsistent with the hypothesis of stepwise migration along the Andes for the origin of the New World amphitropical disjunction observed in clade 1. Support for the monophyly of *T. fuegianus* from the *atp*B–*rbc*L locus, and the resolution of *T. itatiaiae*, with Laurasian populations based on posterior support from the concatenated data set, suggests that independent and direct dispersal events have led to the establishment of Brazilian and southernmost South American populations. A larger sampling of the rare southeast Brazilian populations should be included in future work to further test the phylogenetic affinities of *T. itatiaiae* relative to *T. fuegianus* and Laurasian populations.

An estimated Oligocene origin and Miocene diversification of *Tetraplodon* are incongruent with Late Jurassic to Early Cretaceous Pangaean tectonic events (160–138 Ma; Scotese, 2001; Rogers & Santosh, 2003; Smith *et al.*, 2004) accounting for intercontinental disjunctions within the genus. Divergence-time estimates suggest that the clades in the *T. mnioides* complex arose between the Oligocene and the mid-Miocene, or as late as the early Pliocene for clade 1, and

diversified in the mid-Miocene to late Pliocene. Southernmost South American populations diverged from Laurasian populations 8.63 Ma (95% HPD 3.07–10.11 Ma), consistent with estimates for the origin of amphitropical disjunctions in seed plants (Wen & Ickert-Bond, 2009). Divergence times, based on a conservative absolute rate, thus strongly support direct LDD as the mechanism behind the amphitropical disjunctions in clade 1.

Bipolar LDD events must be explained in the absence of a continuous wind path given the presence of the Intertropical Convergence Zone during the diversification of the clades of the *T. mnioides* complex (Hyeong *et al.*, 2005). In this case, the vulnerability of *Tetraplodon* spores to the extreme conditions associated with high-altitude wind currents (van Zanten, 1978) must also be considered. Bird-mediated dispersal provides the most likely scenario for direct LDD across the tropics (Popp *et al.*, 2011). Birds can disperse diaspores internally (endozoochory) or externally (ectozoochory). Endozoochory requires that birds consume and then retain diaspores throughout their migration and that diaspores remain viable. Spores of the liverwort *Riella* were found to be viable following a 30-minute passage through the gut of domesticated mallard ducks (*Anas platyrhynchos*; Proctor, 1961), and viable spores and vegetative fragments have been recovered from the dung of slugs (*Arion vulgaris*, *Arion rufus* and *Limax cinereoniger*; Boch *et al.*, 2013) and spectacled flying fox (*Pteropus conspicillatus*; Parsons *et al.*, 2007). These studies suggest that bryophytes may be resistant to ingestion, although studies on seed plants have shown that resistance varies according to both vector and plant (Traveset *et al.*, 2001) and that viability decreases in general with increasing time inside a bird (van Leeuwen *et al.*, 2012). Although bird-mediated endozoochory is a more effective means of dispersal than ectozoochory in terms both of number and of diversity for seeds (Brochet *et al.*, 2010; Costa *et al.*, 2014) and aquatic invertebrates (Sánchez *et al.*, 2012), the relative significance of these two modes of zoochory has not been estimated for sporic plants. Migratory shorebirds of the order Charadriiformes are the most likely candidate vectors, as a number of species occupy and provide a direct path between the suitable but disjunct habitats of amphitropical bryophyte species (Pyle, 2008). Vegetative fragments of mosses have been recovered from the plumage of transtropical migrant birds (order Charadriiformes), demonstrating that suitable vectors do pick up diaspores (Lewis *et al.*, 2014). Vegetative diaspores of mosses and fungi have been also recovered from the coats and hooves of roe deer (*Capreolus capreolus*), wild boar (*Sus scrofa*) (Heinken *et al.*, 2001), domesticated sheep (Pauliuk *et al.*, 2011) and the feet of albatrosses (Bailey & James, 1979), suggesting that they are able to adhere to animals' bodies despite a lack of features to facilitate adhesion.

Effective dispersal ultimately depends on establishment in a suitable habitat and substrate, growth and reproduction. Although *Tetraplodon* is specialized for substrates associated with carnivores and carrion, it may not be strictly confined to growth on these substrates (Koponen, 1990), and may be most commonly associated with dung and carrion as a result of fly-mediated dispersal (Cameron & Wyatt, 1986) or its ability to withstand high nutrient concentrations better than other mosses (Fischer, 1936; Koponen, 1990). LDD events may lead to the establishment of sexual populations if the individuals are monoecious, as in the case of the angiosperm genus *Empetrum* (Popp *et al.*, 2011) or *T. mnioides* (Lawton, 1971), and self-compatible, a trait which has been assumed but not yet tested in *Tetraplodon*. Bird-mediated LDD events followed by establishment are assumed to be extremely rare events, but the large size of migrating bird populations and a period of 12.6 Myr since clade 1 began diversifying, may have provided sufficient opportunities for effective LDD in *Tetraplodon* (Nathan, 2006; Gillespie *et al.*, 2012).

CONCLUSIONS

The analysis of discrete loci resolved well-supported lineages within *Tetraplodon*. Three major lineages diversified between the early Miocene and early-to-mid Pliocene and contain the majority of the diversity within the genus, each displaying a different but overlapping geographical distribution. Southernmost South American populations were reconstructed as a monophyletic lineage that diverged from high-latitude Northern Hemisphere populations around 8.63 Ma. The recovery of northern Andean and southernmost South American *T. fuegianus* in distinct clades supports the establishment of southernmost South American populations via direct dispersal rather than stepwise migration along the Andes. Divergence-date estimates and phylogenetic inferences support direct LDD as the mechanism behind the amphitropical disjunction in *Tetraplodon*. Given a lack of wind connectivity between the Northern and the Southern Hemisphere, we propose that migratory birds of the order Charadriiformes are the most likely dispersal vectors.

ACKNOWLEDGEMENTS

This study was supported by National Science Foundation (NSF) Graduate Research Fellowship to L.R.L. and NSF grants DEB-1311405 to L.R.L. and DEB-1212505 to B.G., UConn EEB Ronald Bamford Fund (L.R.L.), UCONN Center for Conservation and Biodiversity Silander Award (L.R.L.), and generous donations from Trudy Gerlach. Fieldwork in the Cape Horn Archipelago, Chile, was made possible by Ricardo Rozzi and Omora Ethnobotanical Park, and by a National Science Foundation grant (DEB-0948380) to William R. Buck. For assistance with Alaskan fieldwork, we thank K. Heather Sinclair, Diana Thomas and Jack Sinclair. For assistance with various aspects of the project, we thank Jon Shaw, Blanka Shaw, Kristian Hassel, Tommy Prestø, Yang Liu, Laura Forest, Juan Carlos Villarreal, Jessica Budke and Rafael Medina. For comments on earlier drafts, we

thank Elizabeth Jockush, Paul Lewis and Kent Holsinger. We thank Heidi Golden, Shawnie Vedder, Tomas Pócs, Alain Vanderpoorten and NYBG for contributing samples. We thank three anonymous referees for their comments on an earlier version of this manuscript.

REFERENCES

Bailey, R.H. & James, P.W. (1979) Birds and the dispersal of lichen propagules. *Lichenologist*, **11**, 105.

Bednarek-Ochyra, H., Váňa, J., Ochyra, R. & Lewis Smith, R.I. (2000) *The liverwort flora of Antarctica*. Polish Academy of Sciences, Krakow.

Boch, S., Berlinger, M., Fischer, M., Knop, E., Nentwig, W., Türke, M. & Prati, D. (2013) Fern and bryophyte endozoochory by slugs. *Oecologia*, **172**, 817–822.

Brochet, A.L., Guillemain, M., Fritz, H., Gauthier-Clerc, M. & Green, A.J. (2010) Plant dispersal by teal (*Anas crecca*) in the Camargue: duck guts are more important than their feet. *Freshwater Biology*, **55**, 1262–1273.

Cameron, R.G. & Wyatt, R. (1986) Substrate restriction in entomophilous Splachnaceae: role of spore dispersal. *Bryologist*, **89**, 279–284.

Castresana, J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, **17**, 540–552.

Chien, G. & He, S. (1999) Splachnaceae. *Moss flora of China*, English Version Volume 3. *Grimmiaceae–Tetraphidaceae* (ed. by G. Chien, M.R. Crosby and S. He), pp. 101–119. Science Press, Beijing.

Cook, L.G. & Crisp, M.D. (2005) Not so ancient: the extant crown group of *Nothofagus* represents a post-Gondwanan radiation. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 2535–2544.

Costa, J.M., Ramos, J.A., da Silva, L.P., Timoteo, S., Araújo, P.M., Felgueiras, M.S., Rosa, A., Matos, C., Encarnação, P., Tenreiro, P.Q. & Heleno, R.H. (2014) Endozoochory largely outweighs epizoochory in migrating passerines. *Journal of Avian Biology*, **45**, 59–64.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**, e88.

Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Du Rietz, G.E. (1940) Problems of bipolar plant distribution. *Acta Phytogeographica Suecica*, **13**, 215–282.

Escudero, M., Valcárel, V., Vargas, P. & Luceño, M. (2010) Bipolar disjunctions in *Carex*: long-distance dispersal, vicariance, or parallel evolution? *Flora*, **205**, 118–127.

Fernández-Mendoza, F. & Printzen, C. (2013) Pleistocene expansion of the bipolar lichen *Cetraria aculeata* into the Southern hemisphere. *Molecular Ecology*, **22**, 1961–1983.

Fischer, J. (1936) Zur Verbreitungsökologie von *Splachnum sphaericum* (L. fil.) Swartz. *Beihefte zum Botanischen Centralblatt Abteilung*, **55**, 122–144.

Frisvoll, A.A. (1978) The genus *Tetraplodon* in Norway. A taxonomic revision. *Lindbergia*, **4**, 225–246.

Gernhard, T. (2008) The conditioned reconstructed process. *Journal of Theoretical Biology*, **253**, 769–778.

Gillespie, R.G., Baldwin, B.G., Waters, J.M., Fraser, C.I., Nikula, R. & Roderick, G.K. (2012) Long-distance dispersal: a framework for hypothesis testing. *Trends in Ecology and Evolution*, **27**, 47–56.

Goffinet, B., Bayer, R.J. & Vitt, D.H. (1998) Circumscription and phylogeny of the Orthotrichales (Bryopsida) inferred from *rbc*L sequence analyses. *American Journal of Botany*, **85**, 1324–1337.

Goffinet, B., Shaw, A.J. & Cox, C.J. (2004) Phylogenetic inferences in the dung-moss family Splachnaceae from analyses of cpDNA sequence data and implications for the evolution of entomophily. *American Journal of Botany*, **91**, 748–759.

Gussarova, G., Popp, M., Vitek, E. & Brochmann, C. (2008) Molecular phylogeny and biogeography of the bipolar *Euphrasia* (Orobanchaceae): recent radiations in an old genus. *Molecular Phylogenetics and Evolution*, **48**, 444–460.

Heinken, T., Lees, R., Raudnitschka, D. & Runge, S. (2001) Epizoochorous dispersal of bryophyte stem fragments by roe deer (*Capreolus capreolus*) and wild boar (*Sus scrofa*). *Journal of Bryology*, **23**, 293–300.

Heinrichs, J., Hentschel, J., Feldberg, K., Bombosch, A. & Schneider, H. (2009) Phylogenetic biogeography and taxonomy of disjunctly distributed bryophytes. *Journal of Systematics and Evolution*, **47**, 497–508.

Huttunen, S., Hedenäs, L., Ignatov, M.S., Devos, N. & Vanderpoorten, A. (2008) Origin and evolution of the Northern Hemisphere disjunction in the moss genus *Homalothecium* (Brachytheciaceae). *American Journal of Botany*, **95**, 720–730.

Hyeong, K-S., Park, S-H., Yoo, C.M. & Kim, K-H. (2005) Mineralogical and geochemical compositions of the eolian dust from the northeast equatorial Pacific and their implications on paleolocation of the Intertropical Convergence Zone. *Paleoceanography*, **20**, PA1010.

Kass, R.E. & Raftery, A.E. (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.

Koponen, A. (1990) Entomophily in the Splachnaceae. *Botanical Journal of the Linnean Society*, **104**, 115–127.

Kreier, H.-P., Feldberg, K., Mahr, F., Bombosch, A., Schmidt, A.R., Zhu, R.-L., von Konrat, M., Shaw, B., Shaw, A.J. & Heinrichs, J. (2010) Phylogeny of the leafy liverwort *Ptilidium*: cryptic speciation and shared haplotypes between the Northern and Southern Hemispheres. *Molecular Phylogenetics and Evolution*, **57**, 1260–1267.

Lawton, E. (1971) *Moss flora of the Pacific Northwest*, 1st edn. Hattori Botanical Laboratory, Nichinan, Miyazaki, Japan.

van Leeuwen, C.H.A., van der Velde, G., van Groenendael, J.M. & Klaassen, M. (2012) Gut travellers: internal dispersal of aquatic organisms by waterfowl. *Journal of Biogeography*, **39**, 2031–2040.

Lewis, P.O. (2001) A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, **50**, 913–925.

Lewis, L.R., Behling, E., Gousse, H., Qian, E., Elphick, C.S., Lamarre, J.F., Bêty, J., Liebezeit, J., Rozzi, R. & Goffinet, B. (2014) First evidence of bryophyte diaspores in the plumage of transequatorial migrant birds. *PeerJ*, **2**, 424.

Maddison, W.P. & Maddison, D.R. (2011) *Mesquite: a modular system for evolutionary analysis*. Version 2.75. Available at: http://mesquiteproject.org/.

Mao, K.-S., Milne, R.I., Zhang, L.-B., Peng, Y.-L., Liu, J.-Q., Thomas, P., Mill, R.R. & Renner, S.S. (2012) Distribution of living Cupressaceae reflects the breakup of Pangea. *Proceedings of the National Academy of Sciences USA*, **109**, 7793–7798.

Marino, P.C. (1988a) Coexistence on divided habitats: mosses in the family Splachnaceae. *Annales Zoologici Fennici*, **25**, 89–98.

Marino, P.C. (1988b) The North American distributions of the circumboreal species of *Splachnum* and *Tetraplodon*. *Bryologist*, **91**, 161–166.

Marino, P.C. (2009) Splachnaceae Greville & Arnott. *Bryophyte flora of North America,* Volume 2 (ed. by R. Zander). Missouri Botanical Garden, St Louis, MO. Available at: http://www.mobot.org/plantscience/bfna/V2/SplaSplachnaceae.htm.

Muñoz, J., Felicísimo, Á.M., Cabezas, F., Burgaz, A.R. & Martínez, I. (2004) Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science*, **304**, 1144–1147.

Myllys, L., Stenroos, S., Thell, A. & Ahti, T. (2003) Phylogeny of bipolar *Cladonia arbuscula* and *Cladonia mitis* (Lecanorales, Euascomycetes). *Molecular Phylogenetics and Evolution*, **27**, 58–69.

Nathan, R. (2006) Long-distance dispersal of plants. *Science*, **313**, 786–788.

Norris, D.H., Koponen, T. & Piippo, S. (1999) Bryophyte flora of the Huon Peninsula, Papua New Guinea. LXVI. Meesiaceae (Musci), with lists of boreal to temperate disjunct, bipolar, and widely spread species in New Guinea. *Annales Botanici Fennici*, **36**, 257–263.

Nyholm, E. (1975) *Illustrated moss flora of Fennoscandia*, 2nd edn. Swedish Research Council, Stockholm.

Ochyra, R. & Bednarek-Ochyra, H. (2008) New and rare species from the Antarctic. *Nova Hedwigia*, **87**, 457–477.

Ochyra, R. & Buck, W.R. (2003) *Arctoa fulvella*, new to Tierra del Fuego, with notes on trans-American bipolar bryogeography. *Bryologist*, **106**, 532–538.

Øvstedal, D.O. & Lewis Smith, R.I. (2001) *Lichens of Antarctica and South Georgia: a guide to their identification and ecology*. Cambridge University Press, Cambridge, UK.

Parsons, J.G., Cairns, A., Johnson, C.N., Robson, S.K.A., Shilton, L.A. & Westcot, D.A. (2007) Bryophyte dispersal by flying foxes: a novel discovery. *Oecologia*, **152**, 112–114.

Pauliuk, F., Müller, J. & Heinken, T. (2011) Bryophyte dispersal by sheep on dry grassland. *Nova Hedwigia*, **92**, 327–341.

Piñeiro, R., Popp, M., Hassel, K., Listl, D., Westergaard, K.B., Flatberg, K.I., Stenøien, H.K. & Brochmann, C. (2012) Circumarctic dispersal and long-distance colonization of South America: the moss genus *Cinclidium*. *Journal of Biogeography*, **39**, 2041–2051.

Pokorny, L., Oliván, G. & Shaw, A.J. (2011) Phylogeographic patterns in two Southern Hemisphere species of *Calyptrochaeta* (Daltoniaceae, Bryophyta). *Systematic Botany*, **36**, 542–553.

Popp, M., Mirré, V. & Brochmann, C. (2011) A single Mid-Pleistocene long-distance dispersal by a bird can explain the extreme bipolar disjunction in crowberries (*Empetrum*). *Proceedings of the National Academy of Sciences USA*, **108**, 6520–6525.

Posada, D. (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Proctor, V.W. (1961) Dispersal of *Riella* spores by waterfowl. *Bryologist*, **64**, 58–61.

Pyle, P. (2008) *Identification guide to North American birds, Part II*. Slate Creek Press, Point Reyes Station, CA.

de Queiroz, A. (2005) The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology and Evolution*, **20**, 68–73.

Rambaut, A. & Drummond, A.J. (2007) *Tracer v1.4*. University of Edinburgh, Edinburgh, UK. Available at: http://beast.bio.ed.ac.uk/Tracer.

Raven, P.H. (1963) Amphitropical relationships in the floras of North and South America. *Quarterly Review of Biology*, **38**, 151–177.

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

Rogers, J.J.W. & Santosh, M. (2003) Supercontinents in Earth history. *Gondwana Research*, **6**, 357–368.

Ronquist, F. & Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Sánchez, M.I., Hortas, F., Figuerola, J. & Green, A.J. (2012) Comparing the potential for dispersal via waterbirds of a native and an invasive brine shrimp. *Freshwater Biology*, **57**, 1896–1903.

Sanderson, M.J. (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution*, **19**, 101–109.

Schuster, R.M. (1983) Phytogeography of the Bryophyta. *New manual of bryology* (ed. by R.M. Schuster), pp. 463–626. Hattori Botanical Laboratory, Nichinan, Miyazaki, Japan.

Scotese, C.R. (2001) *Atlas of Earth history*. University of Texas, Arlington, TX.

Shaw, A.J., Devos, N., Cox, C.J., Boles, S.B., Shaw, B., Buchanan, A.M., Cave, L. & Seppelt, R. (2010) Peatmoss (*Sphagnum*) diversification associated with Miocene Northern Hemisphere climatic cooling? *Molecular Phylogenetics and Evolution*, **55**, 1139–1145.

Simmons, M.P. & Ochoterena, H. (2000) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology*, **49**, 369–381.

Smith, A.G., Smith, D.G. & Funnell, B.M. (2004) *Atlas of Cenozoic and Mesozoic coastlines*. Cambridge University Press, Cambridge, UK.

Stech, M. & Quandt, D. (2010) 20,000 species and five key markers: the status of molecular bryophyte phylogenetics. *Phytotaxa*, **9**, 196–228.

Steere, W.C. (1977) *Tetraplodon paradoxus* and *T. pallidus* (Musci: Splachnaceae) in northern North America. *Brittonia*, **29**, 353–367.

Streimann, H. (1998) Bryological comparisons between Australia and Northern Europe. *Folia Cryptogamica Estonica*, **32**, 97–105.

Sukumaran, J. & Holder, M.T. (2010) A Python library for phylogenetic computing. *Bioinformatics*, **26**, 1569–1571.

Swofford, D.L. (2003) *Phylogenetic analysis using parsimony (*and other methods)*. Version 4.0a129. Sinauer Associates, Sunderland, MA.

Traveset, A., Riera, N. & Mas, R.E. (2001) Passage through bird guts causes interspecific differences in seed germination characteristics. *Functional Ecology*, **15**, 669–675.

Villarreal, J.C. & Renner, S.S. (2012) Hornwort pyrenoids, carbon-concentrating structures, evolved and were lost at least five times during the last 100 million years. *Proceedings of the National Academy of Sciences USA*, **109**, 18873–18878.

Wen, J. & Ickert-Bond, S.M. (2009) Evolution of the Madrean–Tethyan disjunctions and the North and South American amphitropical disjunctions in plants. *Journal of Systematics and Evolution*, **47**, 331–348.

Wilkinson, D.M., Koumoutsaris, S., Mitchell, E.A.D. & Bey, I. (2012) Modelling the effect of size on the aerial dispersal of microorganisms. *Journal of Biogeography*, **39**, 89–97.

Wirtz, N., Printzen, C. & Lumbsch, H.T. (2008) The delimitation of Antarctic and bipolar species of neuropogonoid *Usnea* (Ascomycota, Lecanorales): a cohesion approach of species recognition for the *Usnea perpusilla* complex. *Mycological Research*, **112**, 472–484.

van Zanten, B.O. (1978) Experimental studies on trans-oceanic long-range dispersal of moss spores in the Southern Hemisphere. *Journal of the Hattori Botanical Laboratory*, **44**, 455–482.

Zwickl, D.J. (2006) *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion*. PhD Thesis, University of Texas at Austin, Austin, TX.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Specimen data.

BIOSKETCHES

Lily R. Lewis is a PhD candidate in the Department of Ecology and Evolutionary Biology at the University of Connecticut. Her dissertation research focuses on the phylogenetics and population genetics of *Tetraplodon* with an emphasis on the southern South American *T. fuegianus* and the inclusion of bryophyte diversity in conservation and ecotourism in Southern Chile.

Bernard Goffinet is a professor at the University of Connecticut, where he teaches bryophyte and lichen biology and the evolution of green plants. His research focuses on the systematics and diversification of the Bryophyta (mosses) and lichen-forming fungi, as well as the evolution of organellar genome structure in bryophytes.

Ricardo Rozzi is a professor of philosophy and religious studies and director of the Sub-Antarctic Biocultural Conservation Program at the University of North Texas, and co-founder and director of the Omora Ethnobotanical Park, Chile. His expertise is in ecology and environmental philosophy.

Editor: Christine Maggs

Journal of Biogeography

SUPPORTING INFORMATION

Direct long-distance dispersal shapes New World amphitropical disjunction in the dispersal-limited dung moss *Tetraplodon* (Bryopsida: Splachnaceae)

Lily R. Lewis, Ricardo Rozzi and Bernard Goffinet

APPENDIX S1 Specimen data.

Table S1 Accessions used in analyses. All accessions are represented in Figs 2 & 3. Herbaria include University of Connecticut (CONN), New York Botanical Gardens (NY), Duke University (DUKE), Finnish Museum of Natural History at University of Helsinki (H), University of Alberta (ALTA) and the Norwegian University of Science and Technology (TRH). Only the generic names *Tetraplodon* (*T.*) and *Voitia* (*V.*) are abbreviated. Sample numbers correspond to the sample numbers appended to the end of taxon names in Figs 2 & 3.

Taxon	Collector & collection number	Herbarium	Country	GenBank accession numbers			
				*atp*B–*rbc*L	*rps*4	*trn*G	ITS2
Ingroups							
T. angustatus 1	Pujmanova	NY	Russia	KJ488308	KJ488067	KJ488427	KJ488192
T. angustatus 2	Tan 97-402	NY	Russia	—	—	KJ488428	—
T. angustatus 3	Long 21054	CONN	Nepal	KJ488305	KJ488065	KJ488424	KJ488189
T. angustatus 4	Buck 27855	NY	Canada	KJ488303	KJ488063	KJ488422	KJ488187
T. angustatus 5	Marino 2734	CONN	Canada	KJ488304	KJ488064	KJ488423	KJ488188
T. angustatus 6	Talbot T6400x	NY	Canada	KJ488307	KJ488163	KJ488426	KJ488191
T. angustatus 7	Vanderpoorten 4923	CONN	Alaska, USA	KJ488300	KJ488060	—	KJ488184

Taxon	Collector & collection number	Herbarium	Country	GenBank accession numbers			
				*atp*B–*rbc*L	*rps*4	*trn*G	ITS2
T. angustatus 8	Scotter 22318	NY	Canada	KJ488306	KJ488066	KJ488425	KJ488190
T. blyttii 1	Holten 164466	TRH	Norway	KJ488309	KJ488068	KJ488430	KJ488193
T. blyttii 2	Frisvoll 73525	TRH	Norway	—	KJ488069	KJ488431	KJ488194
T. fuegianus 1	Goffinet 9588	CONN	Chile	KJ488323	—	KJ488445	KJ488209
T. fuegianus 2	Goffinet 9584	CONN	Chile	KJ488324	KJ488084	KJ488446	KJ488210
T. fuegianus 3	Lewis 936	CONN	Chile	KJ488315	KJ488075	KJ488437	KJ488201
T. fuegianus 4	Lewis 952	CONN	Chile	KJ488319	KJ488079	KJ488441	KJ488205
T. fuegianus 5	Lewis 948	CONN	Chile	KJ488317	KJ488077	KJ488439	KJ488203
T. fuegianus 6	Lewis 943	CONN	Chile	KJ488316	KJ488076	KJ488438	KJ488202
T. fuegianus 7	Larrain 34895a	CONN	Chile	KJ488321	KJ488081	KJ488443	KJ488207
T. fuegianus 8	Long 1987	CONN	Chile	KJ488326	KJ488086	KJ488448	KJ488212
T. fuegianus 9	Lewis 950	CONN	Chile	KJ488318	KJ488078	KJ488440	KJ488204
T. fuegianus 10	Goffinet 10507	CONN	Chile	KJ488322	—	KJ488444	KJ488208
T. fuegianus 11	Goffinet 7028	CONN	Chile	KJ488310	KJ488070	KJ488432	KJ488195
T. fuegianus 12	Goffinet 9589	CONN	Chile	KJ488325	KJ488085	KJ488447	KJ488211
T. fuegianus 13	Lewis 998	CONN	Chile	KJ488320	KJ488080	KJ488442	KJ488206
T. fuegianus 14	Lewis 516	CONN	Chile	KJ488313	KJ488073	KJ488435	KJ488198
T. fuegianus 15	Lewis 507	CONN	Chile	KJ488312	KJ488072	KJ488434	KJ488197
T. fuegianus 16	Lewis 530	CONN	Chile	KJ488311	KJ488071	KJ488433	KJ488196
T. fuegianus 17	Lewis 513	CONN	Chile	KJ488314	KJ488074	KJ488436	KJ488199
T. itatiaiae	Vital & Buck 19941	NY	Brazil	KJ488327	KJ488087	KJ488449	KJ488213
T. lamii	De Sloover 43076	H	Papua New Guinea	KJ488328	KJ488088	KJ488450	KJ488214
T. mnioides 1	Goffinet 9371	CONN	Alaska, USA	KJ488361	KJ488121	KJ488484	KJ488245

Taxon	Collector & collection number	Herbarium	Country	GenBank accession numbers			
				*atp*B–*rbc*L	*rps*4	*trn*G	ITS2
T. mnioides 2	Goffinet 9405	CONN	Alaska, USA	KJ488360	KJ488120	KJ488483	KJ488244
T. mnioides 3	Goffinet 9362	CONN	Alaska, USA	KJ488362	KJ488122	KJ488485	KJ488246
T. mnioides 4	Shaw 9053	CONN	Canada	KJ488380	KJ488140	KJ488504	KJ488262
T. mnioides 5	Gillett 18810	NY	Canada	KJ488382	KJ488142	KJ488506	KJ488264
T. mnioides 6	Shaw 9046a	CONN	Canada	KJ488381	KJ488141	KJ488505	KJ488263
T. mnioides 7	Steere 62-285	NY	Greenland	KJ488367	KJ488127	KJ488490	KJ488251
T. mnioides 8	Buck 8867	NY	Alaska, USA	KJ488335	KJ488095	KJ488457	KJ488220
T. mnioides 9	Ireland 16503	NY	Canada	KJ488383	KJ488143	—	KJ488265
T. mnioides 10	Steere 62-460	NY	Greenland	KJ488368	KJ488128	KJ488491	KJ488252
T. mnioides 11	Shevock 31055	NY	China	KJ488345	KJ488104	KJ488467	KJ488230
T. mnioides 12	Shevock 30926	NY	China	KJ488346	KJ488105	KJ488468	KJ488231
T. mnioides 13	Goffinet 10682	CONN	China	KJ488350	KJ488109	KJ488472	KJ488235
T. mnioides 14	Marino 2733	CONN	Canada	KJ488342	KJ488115	KJ488478	KJ488239
T. mnioides 15	Holz CR 00-0713	CONN	Costa Rica	KJ488352	KJ488111	KJ488474	KJ488236
T. mnioides 16	Holz CR 99-0666	CONN	Costa Rica	KJ488353	KJ488112	KJ488475	KJ488237
T. mnioides 17	Shaw 11228	NY	Ecuador	KJ488357	KJ488117	KJ488480	KJ488241
T. mnioides 18	Ramfrez 4737	NY	Columbia	KJ488351	KJ488110	KJ488473	—
T. mnioides 19	LaFarge 9179	ALTA	Uganda	KJ488387	KJ488147	KJ488510	KJ488269
T. mnioides 20	Starling & Brickell A.G.S.E.S. 147	NY	India	KJ488370	KJ488130	KJ488496	KJ488253
T. mnioides 21	Ma 12-3661	CONN	China	KJ488349	KJ488108	KJ488471	KJ488234
T. mnioides 22	Shaw 13457	CONN	Alaska, USA	KJ488337	KJ488097	KJ488459	KJ488223
T. mnioides 23	Golden DM7	CONN	Alaska, USA	KJ488302	KJ488062	KJ488421	KJ488186
T. mnioides 24	Golden DM2	CONN	Alaska, USA	KJ488301	KJ488061	KJ488420	KJ488185

Taxon	Collector & collection number	Herbarium	Country	GenBank accession numbers			
				*atp*B–*rbc*L	*rps*4	*trn*G	ITS2
T. mnioides 25	Tan 89-627	NY	Malaysian Borneo	KJ488341	KJ488101	KJ488464	KJ488227
T. mnioides 26	Lewis 7	CONN	Alaska, USA	KJ488333	KJ488093	KJ488455	KJ488218
T. mnioides 27	Shevrock 23351	CONN	China	KJ488344	KJ488103	KJ488466	KJ488229
T. mnioides 28	Lewis 410	CONN	Alaska, USA	KJ488331	KJ488091	KJ488453	KJ488216
T. mnioides 29	Björk 28592	CONN	Canada	KJ488384	KJ488144	KJ488507	KJ488266
T. mnioides 30	Goffinet 9418	CONN	Alaska, USA	KJ488355	KJ488114	KJ488477	KJ488222
T. mnioides 31	Lewis 515	CONN	Alaska, USA	KJ488329	KJ488089	KJ488451	—
T. mnioides 32	Long 22884	CONN	India	KJ488371	KJ488131	KJ488494	—
T. mnioides 33	Tan 97-276	NY	Russia	KJ488386	KJ488146	KJ488509	—
T. mnioides 34	Goffinet 9414	CONN	Alaska, USA	KJ488354	KJ488113	KJ488476	KJ488238
T. mnioides 35	Lewis 5	CONN	Alaska, USA	KJ488332	KJ488092	KJ488454	KJ488217
T. mnioides 36	Lewis 500	CONN	Alaska, USA	KJ488330	KJ488090	KJ488452	KJ488215
T. mnioides 37	Shevock 23293	CONN	China	KJ488343	KJ488102	KJ488465	KJ488228
T. mnioides 38	Goffinet 10801	CONN	China	KJ488348	KJ488107	KJ488470	KJ488233
T. mnioides 39	Goffinet 10802	CONN	China	KJ488347	KJ488106	KJ488469	KJ488232
T. mnioides 40	Taylor 96-12	CONN	Canada	KJ488340	KJ488100	KJ488463	KJ488226
T. mnioides 41	Norris 89585	NY	USA	KJ488388	KJ488148	KJ488511	KJ488270
T. mnioides 42	Vedder OLY18	CONN	Washington, USA	KJ488389	KJ488149	KJ488512	KJ488271
T. mnioides 43	Talbot AML002-X-11	CONN	Alaska, USA	KJ488336	KJ488096	KJ488458	KJ488221
T. mnioides 44	Looman & Talbot B161740	NY	Alaska, USA	KJ488334	KJ488094	KJ488456	KJ488219
T. mnioides 45	Goffinet 9372	CONN	Alaska, USA	KJ488359	KJ488119	KJ488482	KJ488243
T. mnioides 46	Goffinet 9386	CONN	Alaska, USA	KJ488358	KJ488118	KJ488481	KJ488242
T. mnioides 47	Pujamanova	NY	Russia	KJ488385	KJ488145	KJ488508	KJ488268

Taxon	Collector & collection number	Herbarium	Country	GenBank accession numbers			
				*atp*B–*rbc*L	*rps*4	*trn*G	ITS2
T. mnioides 48	Goffinet 9383	CONN	Alaska, USA	KJ488365	KJ488125	KJ488488	KJ488249
T. mnioides 49	Goffinet 9387	CONN	Alaska, USA	KJ488364	KJ488124	KJ488487	KJ488248
T. mnioides 50	Goffinet 9377	CONN	Alaska, USA	KJ488363	KJ488123	KJ488486	KJ488247
T. mnioides 51	Hyvönen 6210	CONN	Finland	KJ488366	KJ488126	KJ488489	KJ488250
T. mnioides 52	Goffinet 1687	CONN	UK	KJ488356	KJ488116	KJ488479	KJ488240
T. mnioides 53	Holyoak 03-277	CONN	Ireland	KJ488372	KJ488132	KJ488495	KJ488254
T. mnioides 54	Goffinet 3501	CONN	Canada	KJ488338	KJ488098	KJ488460	KJ488224
T. mnioides 55	Lewis 1004	CONN	Norway	KJ488378	KJ488138	KJ488502	KJ488260
T. mnioides 56	Lewis 1029	CONN	Norway	KJ488375	KJ488135	KJ488499	KJ488257
T. mnioides 57	Lewis 1013	CONN	Norway	KJ488374	KJ488134	KJ488498	KJ488256
T. mnioides 58	Lewis 1025	CONN	Norway	KJ488376	KJ488136	KJ488500	KJ488258
T. mnioides 59	Lewis 1010	CONN	Norway	KJ488379	KJ488139	KJ488503	KJ488261
T. mnioides 60	Lewis 1022	CONN	Norway	KJ488373	KJ488133	KJ488497	KJ488255
T. mnioides 61	Lewis 1033	CONN	Norway	KJ488377	KJ488137	KJ488501	KJ488259
T. mnioides 62	Buck 27878	NY	Canada	—	KJ488099	KJ488461	KJ488225
T. mnioides 63	Schofield, Godfrey & Goward No: 98231	NY	Canada	KJ488339	—	KJ488462	—
T. mnioides 64	Aptroot 4988	NY	Iceland	KJ488369	—	KJ488493	—
T. mnioides 65	Aptroot 7-1980	NY	Iceland	—	KJ488129	KJ488492	—
T. mnioides 66	Pujmanova 16.8.1986	NY	Russia	KJ488299	KJ488059	KJ488429	KJ488183
T. pallidus 1	Scotter 26302	CONN	Canada	KJ488395	KJ488154	KJ488517	KJ488276
T. pallidus 2	Golden DM10	CONN	Alaska, USA	KJ488392	KJ488152	KJ488515	KJ488273
T. pallidus 3	Golden DM8	CONN	Alaska, USA	KJ488393	KJ488153	KJ488516	KJ488274

Taxon	Collector & collection number	Herbarium	Country	GenBank accession numbers			
				*atp*B–*rbc*L	*rps*4	*trn*G	ITS2
T. pallidus 4	Shaw 9050	DUKE	Canada	KJ488397	KJ488157	KJ488519	KJ488277
T. pallidus 5	Hedderson15594	CONN	Norway	KJ488394	KJ488158	KJ488520	KJ488275
T. pallidus 6	Scotter 45607	CONN	Canada	KJ488396	KJ488156	KJ488518	—
T. pallidus 7	Vitt 15936	CONN	Canada	—	—	KJ488521	—
T. pallidus 8	Isoviita 4.8.1951	H	Russia	KJ488391	KJ488151	KJ488514	KJ488272
T. paradoxus 1	Hedderson 15624	CONN	Norway	KJ488402	KJ488164	KJ488527	KJ488279
T. paradoxus 2	Murray 77-840A	CONN	Alaska, USA	KJ488399	KJ488160	KJ488523	—
T. paradoxus 3	Steere 25101	CONN	Alaska, USA	KJ488400	KJ488161	KJ488524	—
T. paradoxus 4	Batten 340	CONN	Canada	KJ488401	KJ488162	KJ488526	—
T. paradoxus 5	Scotter 76373	CONN	Canada	—	—	KJ488525	—
T. urceolatus 1	Shaw 9108	CONN	Canada	KJ488412	KJ488174	KJ488537	KJ488288
T. urceolatus 2	Scotter 28242	NY	Canada	KJ488410	KJ488172	KJ488535	KJ488286
T. urceolatus 3	Scotter 28333	NY	Canada	KJ488411	KJ488173	KJ488536	KJ488287
T. urceolatus 4	Goffinet 10803	CONN	China	KJ488404	KJ488166	KJ488529	—
T. urceolatus 5	Goffinet 9417	CONN	Alaska, USA	KJ488407	KJ488169	KJ488532	KJ488283
T. urceolatus 6	Crosby 3887	CONN	Costa Rica	KJ488406	KJ488168	KJ488531	KJ488282
T. urceolatus 7	Long 24045	CONN	China	KJ488405	KJ488167	KJ488530	KJ488281
T. urceolatus 8	Long 22624	CONN	India	KJ488408	KJ488170	KJ488533	KJ488284
T. urceolatus 9	Long 20942	CONN	Nepal	KJ488409	KJ488171	KJ488534	KJ488285
T. urceolatus 10	Goffinet 10804	CONN	China	KJ488403	KJ488165	KJ488528	KJ488280
T. urceolatus 11	Wittmann B111786	CONN	Washington, USA	KJ488413	KJ488175	KJ488538	KJ488289
Undescribed sp.	Miehe 26 Sept. 2000	CONN	Bhutan	KJ488294	—	KJ488415	KJ488178
Voitia hyperborea 1	Scotter 7/27/1990	ALTA	Canada	KJ488414	KJ488176	KJ488539	KJ488290

Taxon	Collector & collection number	Herbarium	Country	GenBank accession numbers			
				*atp*B–*rbc*L	*rps*4	*trn*G	ITS2
V. hyperborea 2	Buck 8901	CONN	Alaska, USA	—	KJ488177	KJ488540	KJ488291
V. nivalis 1	Long 26833	CONN	China	—	—	KJ488542	KJ488292
V. nivalis 2	Buck 8827	CONN	Alaska, USA	—	—	KJ488541	—
Outgroups							
Neomeesia paludella	Goffinet 5862	CONN	Chile	KJ488295	KJ488055	KJ488416	KJ488179
Tayloria longiseta	Touw 15895	H	New Guinea	KJ488296	AY499631	KJ488417	KJ488180
Tayloria squarrosa	Long 21285	CONN	Nepal	KJ488297	AY039056	KJ488418	KJ488181
Tayloria subglabra	Tan 86-3-3	H	Philippines	KJ488298	AY499632	KJ488419	KJ488182

PeerJ

First evidence of bryophyte diaspores in the plumage of transequatorial migrant birds

Lily R. Lewis[1], Emily Behling[1], Hannah Gousse[1], Emily Qian[1], Chris S. Elphick[1], Jean-François Lamarre[2], Joël Bêty[2], Joe Liebezeit[3], Ricardo Rozzi[4,5] and Bernard Goffinet[1]

[1] Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT, USA
[2] Centre d'Études Nordiques et Département de Biologie, Université du Québec à Rimouski, Québec, Canada
[3] Audubon Society of Portland, OR, USA
[4] Omora Ethnobotanical Park, Puerto Williams, Antarctic Province, Chile
[5] Department of Philosophy, University of North Texas, Denton, TX, USA

ABSTRACT

Correlations between transequatorial migratory bird routes and bipolar biogeographic disjunctions in bryophytes suggest that disjunctions between northern and southern high latitude regions may result from bird-mediated dispersal; supporting evidence is, however, exclusively circumstantial. Birds disperse plant units (diaspores) internally via ingestion (endozoochory) or externally by the attachment of diaspores to the body (ectozoochory). Endozoochory is known to be the primary means of bird-mediated dispersal for seeds and invertebrates at local, regional, and continental scales. Data supporting the role of bird-mediated endozoochory or ectozoochory in the long distance dispersal of bryophytes remain sparse, however, despite the large number of bryophytes displaying bipolar disjunctions. To determine if transequatorial migrant shorebirds may play a role in the ectozoochory of bryophyte diaspores, we developed a method for screening feathers of wild birds. We provide the first evidence of microscopic bryophyte diaspores, as well as those from non-bryophyte lineages, embedded in the plumage of long distance transequatorial migrant birds captured in their arctic breeding grounds. The number of diaspores recovered suggests that entire migratory populations may be departing their northern breeding grounds laden with potentially viable plant parts and that they could thereby play significant roles in bipolar range expansions of lineages previously ignored in the migrant bird dispersal literature.

Subjects Biogeography, Ecology
Keywords Bryophyte, Bipolar, Ectozoochory, Long-distance dispersal, Diaspore, Endozoochory, Shorebirds, Sporic, Transequatorial

Submitted 8 April 2014
Accepted 20 May 2014
Published 12 June 2014

Corresponding author
Lily R. Lewis, Lily.Lewis@uconn.edu

Academic editor
Anna Traveset

Additional Information and Declarations can be found on page 9

DOI **10.7717/peerj.424**

INTRODUCTION

Climate, geological processes, and long-distance dispersal shape global species distributions. Although wind is the primary vector for long-distance dispersal (LDD) of microscopic diaspores in the Northern (*Wilkinson et al., 2012*) and Southern (*Muñoz et al., 2004*) hemispheres, it is an unlikely candidate for explaining bipolar disjunctions

(i.e., transequatorial distributions between high latitude areas), which characterize many sporic plants, including at least 60 species of moss (*Ochyra, 1992*; *Ochyra & Buck, 2003*; *Ochyra, Smith & Bednarek-Ochyra, 2008*), 17 liverwort species (*Schuster, 1983*; *Bednarek-Ochyra et al., 2000*), and 160 lichen species (*Øvstedal & Lewis Smith, 2001*). Bipolar disjunctions at the species or infrageneric levels largely originated in the Miocene through the Pleistocene (*Wen & Ickert-Bond, 2009*; *Popp, Mirré & Brochmann, 2011*; *Fernández-Mendoza & Printzen, 2013*) and thus correspond with the continued presence of the Intertropical Convergence Zone (*Hyeong et al., 2005*), which produces a barrier to wind dispersal across low latitudes. This time frame predates human activities, which in modern times have greatly expanded dispersal opportunities (*Wilkinson, 2010*). Inferences based on molecular data have largely supported direct LDD across the tropics as the process shaping bipolar disjunctions in plants and lichens (*Wirtz, Printzen & Lumbsch, 2008*; *Wen & Ickert-Bond, 2009*; *Kreier et al., 2010*; *Popp, Mirré & Brochmann, 2011*; *Piñeiro et al., 2012*), with only a few examples supporting ancient vicariance (*Mao et al., 2012*) or stepwise migration across the tropics (*Fernández-Mendoza & Printzen, 2013*). Bird migration routes between boreal and austral regions provide a direct link between antipodal populations, and hence migratory birds are routinely invoked as the dispersal vectors that account for bipolar disjunctions (*Du Rietz, 1940*; *Wen & Ickert-Bond, 2009*; *Popp, Mirré & Brochmann, 2011*). Evidence of birds dispersing sporic plant or lichen units (i.e., diaspores), as well as those of other lineages, across the tropics, however, is exclusively circumstantial.

Birds may disperse diaspores internally (endozoochory) or externally (ectozoochory; *Ridley, 1930*; *Carlquist, 1974*). The feasibility of avian mediated endozoochory in the LDD of aquatic invertebrates and seed plants is well supported by diaspore retention times in captive birds and the survival of ingested diaspores (*Proctor, 1968*; *Figuerola & Green, 2002*) as well as the recovery of viable seeds from the dung of wild birds (*Bruun, Lundgren & Philipp, 2008*). Based on experimentally derived retention times and viability estimates paired with migratory movement data, modeling of potential dispersal distances supports intercontinental-scale movement of aquatic organisms and seeds (*Viana et al., 2013*). The effects of avian ingestion on the spores of the liverwort *Riella* was explored by *Proctor (1961)* who demonstrated that viable spores could be recovered from the dung of domesticated mallard ducks (*Anas platyrhynchos*) after approximately 30 min. Spores of slime molds have been recovered from the dung of migratory songbirds (*Suthers, 1985*). Furthermore, bryophyte spores and vegetative diaspores have been shown to withstand ingestion by slugs (*Arion vulgaris*, *A. rufus*, and *Limax cinereoniger*; *Boch et al., 2013*), or by the spectacled flying fox (*Pteropus conspicillatus*; *Parsons et al., 2007*), respectively. Studies on seed plants have shown that the effects of internal passage vary according to bird species and seed (*Traveset, Riera & Mas, 2001*), but in general, diaspore viability decreases with increasing exposure to a bird's digestive tract (*van Leeuwen et al., 2012*).

Alternatively, ectozoochory requires that (1) diaspores become attached to the exterior of a bird prior to migration, and (2) remain on the bird over the course of the journey (*Figuerola & Green, 2002*). Comparisons of endozoochory and ectozoochory of aquatic invertebrates and seed plants suggest a secondary role of ectozoochory in terms of number

and diversity of diaspores dispersed (*Brochet et al., 2010*; *Sánchez et al., 2012*; *Costa et al., 2014*). Several studies, however, suggest that ectozoochory may provide an important means of dispersal for fungi or sporic plants. Viable fungal spores were recovered from the feathers of wild birds up to 45 days after inoculation (*Warner & French, 1970*) and diaspores of lichen-forming fungi can become attached to the feet of albatross (*Bailey & James, 1979*). Screening of the coats and hooves of roe deer (*Capreolus capreolus*), wild boar (*Sus scrofa*) (*Heinken et al., 2001*) and domesticated sheep (*Pauliuk, Muller & Heinken, 2011*) showed that mosses, most notably pleurocarpous mosses, are commonly carried by mammals, despite lacking any specialized means of adhering to them (*Sorensen, 1986*). Although birds are known to actively transport bryophytes locally for use as nesting material (*Osorio-Zuñiga, Fontúrbel & Håkan, 2014*), evidence for avian ectozoochory playing a role in the regional or global dispersal of bryophytes is lacking. To test this hypothesis we sought to assess whether the first condition for ectozoochory of bryophytes would be met, namely do potential vectors between high latitude ecosystems harbor bryophyte diaspores.

To evaluate the first condition for ectozoochory by transequatorial migrant birds, we developed a method for microscopically screening feathers for diaspores. We present the results from the screening of 23 individual birds, representing eight transequatorial long-distance migrant species in the order Charadriiformes (shorebirds) captured in their arctic breeding grounds. We provide empirical evidence demonstrating that the first condition for ectozoochorous LDD is met, and with a frequency that suggests that birds are active carriers of diaspores and thus may play a critical role in shaping global species distributions.

METHODS

Feather samples were collected from transequatorial migrant birds in their breeding ranges at a site along the Ikpikpuk River, U.S.A. (approximate location: 70.55343°N, 154.69750°W; United States Department of the Interior, permit #23566) and Bylot Island, Nunavut Territory, Canada (approximate locations: 73.15623°N, 79.97065°W; and 72.89216°N, 79.90510°W; Comité de protection des animaux de l'Université du Québec à Rimouski, permit # CPA-42-10-77 - R1) (Fig. 1) between late May and July of 2008 through 2013 by members of the Arctic Shorebird Demographic Network. Target bird species were selected based on migration paths connecting the Northern and Southern Hemispheres, having breeding and wintering ranges in habitats where bryophytes are abundant, and all represent members of the order Charadriiformes. Feather sampling times correspond with the availability of mature bryophyte spores, and preceded the commencement of fall migration. Individual birds were captured using bow nets positioned on nests. When triggered, the bow nets formed a dome over the nest so birds were not immobilized or pressed against the ground, thus lowering the risk that diaspores were picked up due to extraordinary circumstances. Birds were held by grasping the back with clean hands or using nitrile gloves. Feathers were sampled from the base of the breastbone, to minimize invasiveness of feather removal, and based on our prediction that the ventral surface of birds is most likely to make contact with mature bryophyte sporophytes, which are

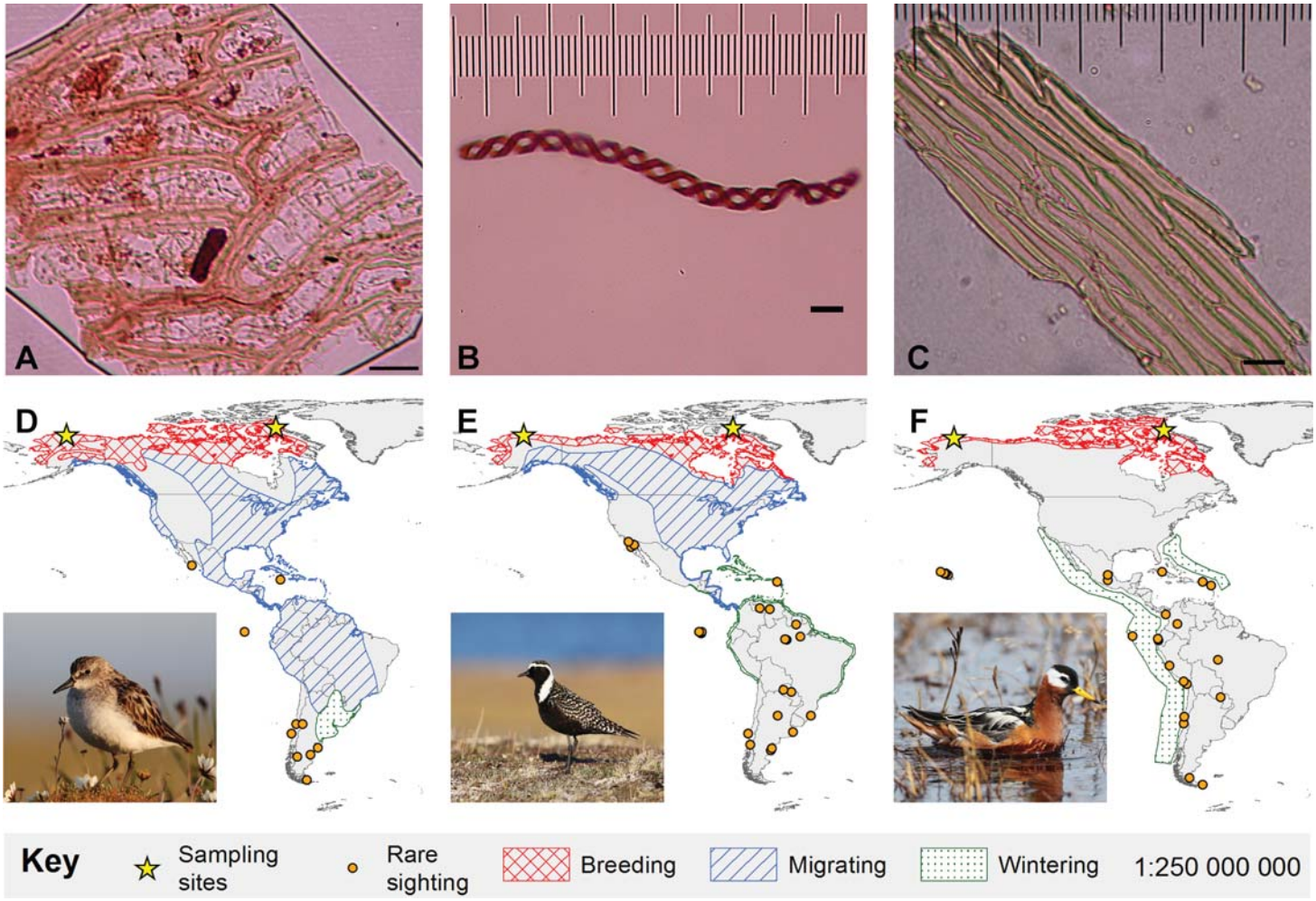


Figure 1 Bryophyte diaspores and their vectors. Three bryophyte diaspores, (A) *Sphagnum* leaf fragment, (B) liverwort elater, and (C) Bryopsid moss leaf fragment recovered from (D) semipalmated sandpiper (Alaska-6-July-2013), (E) American golden-plover (Canada-30-June-2011) and (F) red phalarope (Alaska-22-June-2013), respectively. 5 μm horizontal scale bars are in the lower right corner of each diaspore image. Maps show Western Hemisphere breeding, migratory, and wintering distributions as well as rare sightings for each bird species (*Ridgeley et al., 2012*). The migratory and wintering range for red phalaropes (F) overlaps. Bird photo credits: Cameron Rutt.

typically erect, or vegetative parts of bryophyte mats. Three to six contour and undercoat feathers were collected using tweezers from each bird and immediately placed into a clean paper envelope. If prebasic molt had begun, older-generation feathers were selected. Following collection, feathers were stored in sealed paper envelopes at room temperature. Contour and undercoat feathers collected from a single individual on the same date were pooled for screening. Feathers from 23 individual birds, representing 8 species, American golden-plover (*Pluvialis dominica*; $n = 11$ individuals), red phalarope (*Phalaropus fulicarius*; $n = 3$), red-necked phalarope (*Phalaropus lobatus*; $n = 2$), ruddy turnstone (*Arenaria interpres*; $n = 1$), dunlin (*Calidris alpina*; $n = 1$), Baird's sandpiper (*Calidris bairdii*; $n = 1$), white-rumped sandpiper (*Calidris fuscicollis*; $n = 1$), and semipalmated sandpiper (*Calidris pusilla*; $n = 3$) were screened (Table 1).

Table 1 Feather screening results. Bird species screened, total number of individuals (and feathers) screened, total number of vectors detected (individual bird carrying diaspores) per species, and individual vector identities (reported as location and date of sampling) and number of diaspores recovered per individual. Collection localities are shown in Figs. 1D–1F. Recovered diaspores are shown in Figs. 1A–1C and Figs. 2A–2T. Thirteen diaspores were recovered from red phalarope Alaska-22-June-2013. This bird showed no signs of sickness and did not exhibit any peculiar behaviors. Seven individuals representing three species were found to be vectors out of a total of 23 individuals representing 8 species screened. Small sample sizes likely account for the absence of diaspores in some species.

Bird species	Total individuals screened (# feathers)	Total # vectors	Total # diaspores recovered	Vector ID	# Diaspores per vector	Figures(s)
American golden-plover	11 (23)	3	6	Canada-30-June-2011	2	Figs. 1B and 2P
				Canada-7-July-2011	2	Figs. 2I and 2N
				Canada-13-July-2013	2	Figs. 2B and 2C
Semipalmated sandpiper	3 (21)	2	3	Alaska-5-July-2013	2	Figs. 2D and 2T
				Alaska-6-July-2013	1	Figs. 1A
Red phalarope	3 (14)	2	14	Alaska-22-June-2013	13[a]	Figs. 1C, 2A, 2F–2H, 2J–2M, 2O and 2Q–2S
				Alaska-29-June-2013	1	2E
Red-necked phalarope	2 (8)	0	–	–	–	–
Ruddy turnstone	1 (3)	0	–	–	–	–
Dunlin	1 (2)	0	–	–	–	–
Baird's sandpiper	1 (1)	0	–	–	–	–
White-rumped sandpiper	1 (1)	0	–	–	–	–
TOTAL	**23 (73)**	**7**	**23**	–	–	–

Notes.

[a] Red phalarope Alaska-22-June-2013 harbored more diaspores than could be reliably counted, thus 13 representative diaspores were photo-documented.

All screening was performed in a laminar flow hood in which the surface of the work area and all materials were sterilized with 10% bleach, 70% ethanol and 15 min of ultraviolet light exposure. The laminar flow hood was located in a laboratory where bryophytes are not handled. Feathers were placed in a sterile 1.5 mL centrifuge tube with 350 µL of autoclaved distilled water, and vortexed for 5 min at 2,500 rpm to dislodge any particulate matter attached to the feather. The sample was then centrifuged for 5 min at 14,000 g to collect the particulate matter in the base of the microcentrifuge tube. The feather was removed from the tube and placed in a second sterile 1.5 mL microcentrifuge tube for drying. The wash water was then re-centrifuged for 3 min in case the pellet was disturbed when the feather was removed from the tube.

Three hundred µL of supernatant were pipetted off and discarded following the second centrifuging to reduce the sample size for efficient screening. The tube with the remaining 50 µL of the wash water was vortexed to redistribute the pellet, and the solution divided into two 25 µL-samples for microscopic examination. Samples were screened at 40X magnification. The presence of any putative diaspores was recorded and photographed with a Nikon Coolpix E995 camera.

RESULTS & DISCUSSION

We recovered 23 structures representing bryophytes (mosses and liverworts), cyanobacteria, algae and fungi from the small down and contour breast feathers of two semipalmated

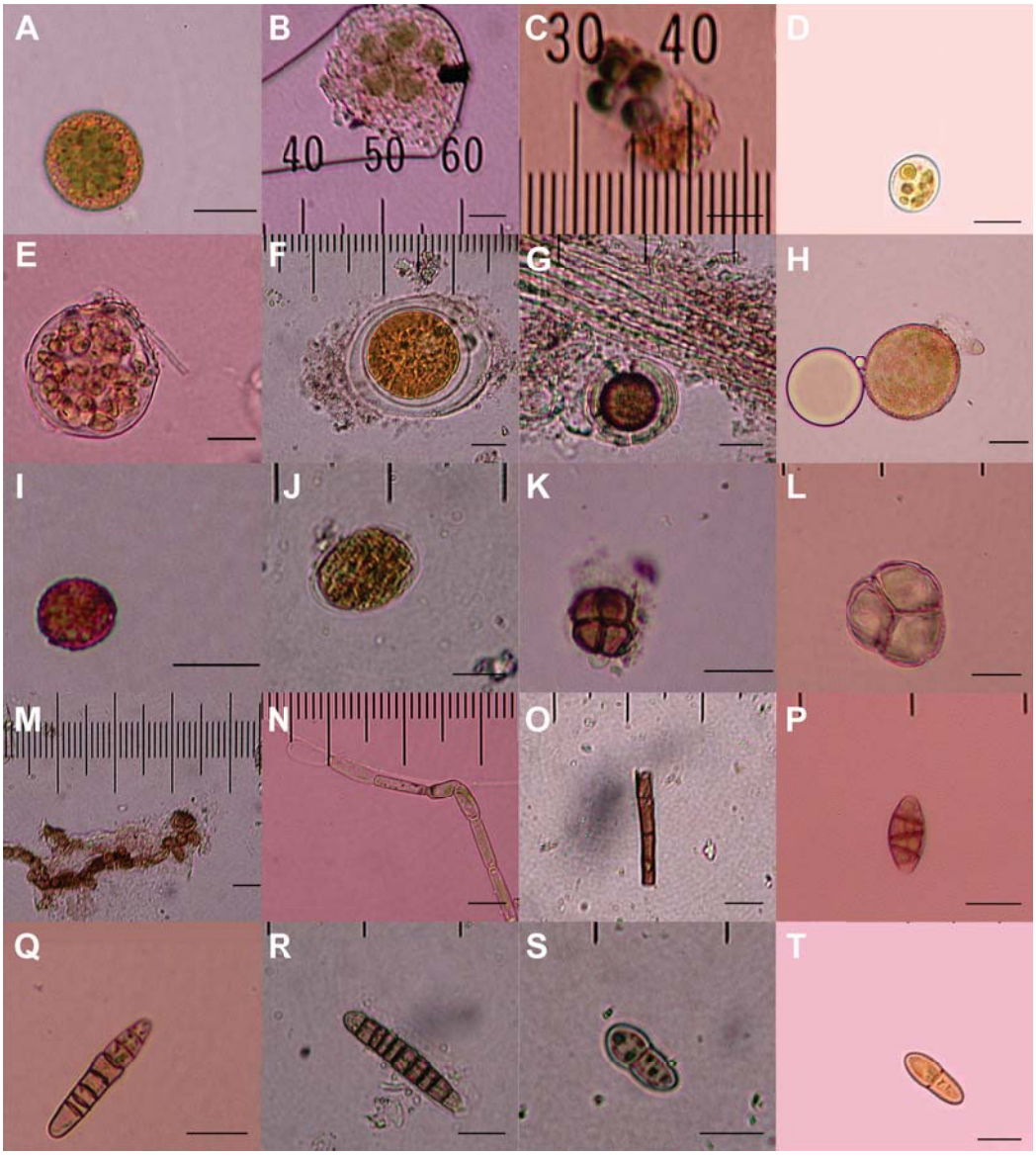


Figure 2 Diaspores recovered from the feathers of 23 birds. Twenty of a total of 23 putative diaspores recovered from breast feathers of migratory shorebirds in their breeding ranges. A–G, K, and M are believed to represent green algae or cyanobacteria; H & I meiotic spores, with L representing an immature meiotic product; J, N, and O are multicellular plant fragments; P–T are fungal spores. 5 µm horizontal scale bars are in the lower right corner of each image. Bryophyte diaspores are shown in Fig. 2 with their vectors. Vectors for each diaspore are listed in Table 1.

sandpipers, two red phalaropes, and three American golden-plovers (Table 1; Figs. 1 and 2). Two of the recovered structures are moss leaf fragments. The first is composed of dimorphic cells, with relatively thin chlorophyllose cells alternating with wide hyaline cells, characteristic of *Sphagnum* leaves (*Goffinet, Buck & Shaw, 2009*; Fig. 1A). The second has smooth elongate cells, most likely representing a moss from the class Bryopsida, potentially belonging to the family Bryaceae s. lat. or to the the pleurocarpous mosses (superorder

Hypnanae; *Goffinet, Buck & Shaw, 2009*; Fig. 1C). The third bryophyte structure is a liverwort elater, an elongated cell with two helical thickenings (*Crandall-Stotler, Stotler & Long, 2009*; Fig. 1B). Elaters are dead, hygroscopic structures in the capsule that aid in the dispersal of the spores. The presence of an elater in the plumage highlights that even rare structures, available during limited periods of the year may be picked-up by birds. The recovery of two vegetative diaspores suggests that unspecialized diaspores, which are available consistently throughout the year, may play an important role in ectozoochory of bryophytes.

The behavior of these transequatorial migrant birds in their northern breeding grounds likely promotes their inadvertent acquisition of diaspores. American golden-plovers, semipalmated sandpipers, and red phalaropes all breed in coastal tundra (*Tracy, Schamel & Dale, 2002*; *Hicklin, Gratto-Trevor & Poole, 2010*; *Johnson & Connors, 2010*), where bryophytes are common. Shallow nests are constructed by scraping depressions into the ground with breast, feet and beaks, and are commonly lined with plant materials (*Tracy, Schamel & Dale, 2002*; *Hicklin, Gratto-Trevor & Poole, 2010*; *Johnson & Connors, 2010*). Timing of molt and migratory behavior will affect the likelihood of attached diaspores being dispersed across the birds' migratory range. Molt in American golden-plovers and semipalmated sandpipers takes place primarily after concluding southward migration, but may occasionally commence before they leave the Arctic (*Pyle, 2008*; *Hicklin, Gratto-Trevor & Poole, 2010*; *Johnson & Connors, 2010*). Red phalaropes replace most body feathers prior to, and flight feathers after, southward migration (*Tracy, Schamel & Dale, 2002*; *Pyle, 2008*). American golden-plovers and semipalmated sandpipers make only terrestrial stopovers throughout their migration to South America (*Hicklin, Gratto-Trevor & Poole, 2010*; *Johnson & Connors, 2010*), while red phalaropes travel largely over the ocean and spend the non-breeding period offshore (*Tracy, Schamel & Dale, 2002*). The individuals screened for this study were sampled between June and July (with one exception being a ruddy turnstone sampled on May 31st). Fall migrations in the species found to be vectors commence as early as July for semipalmated sandpipers and red phalaropes, and August for American golden-plovers (*Tracy, Schamel & Dale, 2002*; *Hicklin, Gratto-Trevor & Poole, 2010*; *Johnson & Connors, 2010*), which overlaps with the production of spores by many arctic bryophyte species. Vegetative fragments are constantly available and do not require any temporal correlations between vectors and diaspores. The post-migratory molt and terrestrial destinations of American golden-plovers and semipalmated sandpipers are compatible with the requirements for dispersal across the equator and subsequent establishment of diaspores. The majority of migratory shorebirds with non-breeding grounds in the Southern Hemisphere provide similar opportunities for diaspore dispersal, with the molt typically occurring on the southern non-breeding grounds (*Pyle, 2008*).

Prior evidence of ectozoochory by birds has shown that this process plays a role in the dispersal of aquatic invertebrates (*Dundee, Phillips & Newsom, 1967*; *Green & Figuerola, 2005*; *Sánchez et al., 2012*), seed plants (*Brochet et al., 2010*; *Costa et al., 2014*), and fungi (*Warner & French, 1970*). Sporic plants commonly exhibit broad intercontinental

disjunctions, even across disjunctions that lack wind connectivity. Despite this situation, little work has been done to investigate the possibility of long distance ectozoochory of sporic plants by birds. Our observations provide unequivocal evidence of bryophyte, fungal, and protist diaspores, embedded within the plumage of transequatorial migrant birds, demonstrating that the first condition for bird-mediated ectozoochory is met. Furthermore, our study suggests that vegetative fragments may be significant dispersal units for ectozoochory of bryophytes, supported more broadly by the screening of non-avian vectors (*Heinken et al., 2001*; *Pauliuk, Muller & Heinken, 2011*), relative to anemochory whereby spores may be the primary dispersal units. Experimental studies comparing the ability of spores and vegetative fragments to attach and be carried by animal vectors will be necessary to explore this hypothesis further.

The potential of diaspores to establish new populations depends on their sustained viability over the course of dispersal. In bryophytes, resistance of spores to the extreme conditions associated with LDD by wind is a determinant of a given species' potential for wind mediated range expansions, with continental endemics displaying greater vulnerability during transport than transoceanic disjunct species (*van Zanten, 1978*). The conditions experienced by diaspores trapped in the plumage of a bird are plausibly less severe than those associated with high atmospheric wind dispersal. Bryophytes are well known for their physiological drought tolerance and the totipotency of their vegetative tissues, which can develop mature plants even after severe grinding (*Shaw, 1986*), rapid passage through a mammalian digestive tract (*Parsons et al., 2007*) and even after being frozen under a glacier for 400 years (*La Farge, Williams & England, 2013*). Future work will explicitly address the second condition for successful ectozoochory by employing culturing techniques to assess viability of recovered diaspores and DNA barcoding for diaspore identification. Based on the general literature supporting the resilience of vegetative bryophyte diaspores, however, resistance to dispersal conditions is unlikely a strong selective force governing the effectiveness of bird-mediated ectozoochory and associated range expansions (*Sánchez et al., 2012*).

Establishment of a new population is a stochastic event following any mode of dispersal, and may pose the greatest challenge toward range expansions. Considering that 23 diaspores were recovered from seven out of 23 birds sampled, the frequency with which birds may externally transport such structures over time may be sufficiently high to explain bipolar distributions. Extrapolated to entire migratory populations, which range in size in North America from 500,000 individuals for American golden-plovers to 2,260,000 individuals for semipalmated sandpipers (*Andres et al., 2012*), hundreds of thousands of diaspores may be transported annually across the Equator, dramatically increasing the probability of successful dispersal, establishment and thereby range extension.

ACKNOWLEDGEMENTS

Louise Lewis provided access to a laminar flow hood, Richard Lanctot helped in the submission of this project to the Arctic Shorebird Demographic Network, the Arctic Shorebird Demographic Network provided support and collaboration, Margaret Rubega

and Matthew Chmielewski provided advice on project development, Susan Hochgraf assisted with preliminary protocol testing, Rafael Medina provided laboratory assistance, and Cameron Rutt contributed bird photos. Andy Green, Casper van Leeuwen, and one anonymous reviewer provided valuable feedback on the initial submission of this work.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Funding for this project comes from the University of Connecticut Katie Bu Memorial Fund awarded to EQ, Switzer Environmental Fellowship awarded to LRL, and National Science Foundation grants DEB 1311405 and 1212505 to LRL and BG, respectively. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
University of Connecticut Katie Bu Memorial Fund.
Switzer Environmental Fellowship.
National Science Foundation grants: DEB 1311405, 1212505.

Competing Interests

Chris Elphick is an Academic Editor for PeerJ. Joe Liebezeit is an employee of the Audubon Society of Portland and Ricardo Rozzi is an employee of the Omora Ethnobotanical Park. All other authors have no competing interests.

Author Contributions

- Lily R. Lewis and Bernard Goffinet conceived and designed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Emily Behling and Emily Qian conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, reviewed drafts of the paper.
- Hannah Gousse performed the experiments, analyzed the data, wrote the paper, reviewed drafts of the paper.
- Chris S. Elphick conceived and designed the experiments, analyzed the data, wrote the paper, reviewed drafts of the paper.
- Jean-François Lamarre analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Joël Bêty and Joe Liebezeit analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, reviewed drafts of the paper.
- Ricardo Rozzi analyzed the data, wrote the paper, reviewed drafts of the paper.

Field Study Permissions

The following information was supplied relating to field study approvals (i.e., approving body and any reference numbers):

Canadian samples:

Comité de protection des animaux de l'Université du Québec à Rimouski (CPA-UQAR), permit # CPA-42-10-77 (R1).

United States (Alaska) samples:

United States Department of the Interior, U.S. Geological Survey

Patuxent Wildlife Research Center

Bird Banding Laboratory

Federal Bird Banding Permit, Permit # 23566.

REFERENCES

Andres BA, Smith PA, Morrison RIG, Gratto-Trevor CL, Brown SC, Friis CA. 2012. Population estimates of North American shorebirds. *Wader Study Group Bulletin* **119**:178–194.

Bailey R, James P. 1979. Birds and the dispersal of lichen propagules. *Lichenologist* **11**:105–106 DOI 10.1017/S0024282979000141.

Bednarek-Ochyra H, Vana J, Ochyra R, Lewis Smith RI. 2000. *The liverwort flora of Antarctica*. Poland: W. Szafer Institute of Botany, Polish Academy of Sciences.

Boch S, Berlinger M, Fischer M, Knop E, Nentwig W, Türke M, Prati D. 2013. Fern and bryophyte endozoochory by slugs. *Oecologia* **172**:817–822 DOI 10.1007/s00442-012-2536-0.

Brochet AL, Guillemain M, Fritz H, Gauthier-Clerc M, Green AJ. 2010. Plant dispersal by teal (*Anas crecca*) in the Camargue: duck guts are more important than their feet. *Freshwater Biology* **55**:1262–1273 DOI 10.1111/j.1365-2427.2009.02350.x.

Bruun HH, Lundgren R, Philipp M. 2008. Enhancement of local species richness in tundra by seed dispersal through guts of muskox and barnacle goose. *Oecologia* **155**:101–110 DOI 10.1007/s00442-007-0892-y.

Carlquist S. 1974. *Island biology*. New York & London: Columbia University Press.

Costa MJ, Ramos JA, da Silva LP, Timoteo S, Araújo PM, Felgueiras MS, Rosa A, Matos C, Encarnação P, Tenreiro PQ, Heleno RH. 2014. Endozoochory largely outweights epizoochory in migrating passerines. *Journal of Avian Biology* **45**:59–64 DOI 10.1111/j.1600-048X.2013.00271.x.

Crandall-Stotler B, Stotler RE, Long DG. 2009. Morphology and classification of the Marchantiophyta. In: Goffinet B, Shaw AJ, eds. *Bryophyte biology*. 2nd edition. Cambridge: Cambridge University Press, 1–54.

Du Rietz GE. 1940. Problems of bipolar plant distribution. *Acta Phytogeographica Suecica* **13**:215–282.

Dundee DS, Phillips PH, Newsom JD. 1967. Snails on migratory birds. *Nautilus* **80**:89–91.

Fernández-Mendoza F, Printzen C. 2013. Pleistocene expansion of the bipolar lichen *Cetraria aculeata* into the Southern hemisphere. *Molecular Ecology* **22**:1961–1983 DOI 10.1111/mec.12210.

Figuerola J, Green AJ. 2002. Dispersal of aquatic organisms by waterbirds: a review of past research and priorities for future studies. *Freshwater Biology* **47**:483–494 DOI 10.1046/j.1365-2427.2002.00829.x.

Goffinet B, Buck WR, Shaw AJ. 2009. Morphology and classification of Bryophyta. In: Goffinet B, Shaw AJ, eds. *Bryophyte biology*. 2nd edition. Cambridge: Cambridge University Press, 55–138.

Green AJ, Figuerola J. 2005. Recent advances in the study of long-distance dispersal of aquatic invertebrates via birds. *Diversity and Distributions* **11**:149–156 DOI 10.1111/j.1366-9516.2005.00147.x.

Hicklin P, Gratto-Trevor CL. 2010. Semipalmated sandpiper (*Calidris pusilla*). In: Poole A, ed. *The birds of North America online*. Ithaca: Cornell Lab of Ornithology.

Heinken T, Lees R, Raudnitschka D, Runge S. 2001. Epizoochorous dispersal of bryophyte stem fragments by roe deer (*Capreolus capreolus*) and wild boar (*Sus scrofa*). *Journal of Bryology* **23**:293–300 DOI 10.1179/jbr.2001.23.4.293.

Hyeong K, Park S, Yoo CH, Kim K. 2005. Mineralogical and geochemical compositions of the eolian dust from the northeast equatorial Pacific and their implications on paleolocation of the Intertropical Convergence Zone. *Paleoceanography* **20**:PA1010 DOI 10.1029/2004PA001053.

Johnson OW, Connors PG. 2010. American Golden-Plover (*Pluvialis dominica*). In: Poole A, ed. *The birds of North America online*. Ithaca: Cornell Lab of Ornithology.

Kreier HP, Feldberg K, Mahr F, Bombosch A, Schmidt AR, Zhu RL, von Konrat M, Shaw B, Shaw AJ, Heinrichs J. 2010. Phylogeny of the leafy liverwort *Ptilidium*: cryptic speciation and shared haplotypes between the Northern and Southern Hemispheres. *Molecular Phylogenetics and Evolution* **57**:1260–1267 DOI 10.1016/j.ympev.2010.10.002.

La Farge C, Williams KH, England JH. 2013. Regeneration of Little Ice Age bryophytes emerging from a polar glacier with implications of totipotency in extreme environments. *Proceedings of the National Academy of Sciences of the United States of America* **110**:9839–9844 DOI 10.1073/pnas.1304199110.

Mao K, Milne RI, Zhang L, Peng Y, Liu J, Thomas P, Mill RR, Renner SS. 2012. Distribution of living Cupressaceae reflects the breakup of Pangea. *Proceedings of the National Academy of Sciences of the United States of America* **109**:7793–7798 DOI 10.1073/pnas.1114319109.

Muñoz J, Felicísimo AM, Cabezas F, Burgaz AR, Martínez I. 2004. Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science* **304**:1144–1147 DOI 10.1126/science.1095210.

Ochyra R. 1992. *Amblyodon dealbatus* (Musci, Meesiaceae)—a bipolar disjunct. *Fragmenta Floristica et Geobotanica* **37**:251–259.

Ochyra R, Smith L, Bednarek-Ochyra H. 2008. *The illustrated moss flora of Antarctica*. New York: Cambridge University Press.

Ochyra R, Buck WR. 2003. *Arctoa fulvella*, new to Tierra del Fuego, with notes on trans-American bipolar bryogeography. *Bryologist* **106**:532–538 DOI 10.1639/0007-2745(2003)106[532:AFNTTD]2.0.CO;2.

Osorio-Zuñiga F, Fontúrbel FE, Håkan R. 2014. Evidence of mutualistic synzoochory between cryptogams and hummingbirds. *Oikos* **123**:553–558 DOI 10.1111/j.1600-0706.2013.01027.x.

Øvstedal DO, Lewis Smith RI. 2001. Lichens of Antarctica and South Georgia: a guide to their identification and ecology. In: *Studies in polar research*. Cambridge: Cambridge University Press.

Parsons JG, Cairns A, Johnson CN, Robson SKA, Shilton LA, Westcot DA. 2007. Bryophyte dispersal by flying foxes: a novel discovery. *Oecologia* **152**:112–114 DOI 10.1007/s00442-006-0639-1.

Pauliuk F, Muller J, Heinken T. 2011. Bryophyte dispersal by sheep on dry grassland. *Nova Hedwigia* **92**:327–341 DOI 10.1127/0029-5035/2011/0092-0327.

Piñeiro R, Popp M, Hassel K, Listl D, Westergaard KB, Flatberg KI, Stenøien HK, Brochmann C. 2012. Circumarctic dispersal and long-distance colonization of South America: the moss genus *Cinclidium*. *Journal of Biogeography* **39**:2041–2051 DOI 10.1111/j.1365-2699.2012.02765.x.

Popp M, Mirré V, Brochmann C. 2011. A single Mid-Pleistocene long-distance dispersal by a bird can explain the extreme bipolar disjunction in crowberries (*Empetrum*). *Proceedings of the National Academy of Sciences of the United States of America* **108**:6520–6525 DOI 10.1073/pnas.1012249108.

Proctor VW. 1961. Dispersal of *Riella* spores by waterfowl. *Bryologist* **64**:58–61 DOI 10.1639/0007-2745(1961)64[58:DORSBW]2.0.CO;2.

Proctor VW. 1968. Long-distance dispersal of seeds by retention in digestive tract of birds. *Science* **160**:321–322 DOI 10.1126/science.160.3825.321.

Pyle P. 2008. *Identification guide to North American birds, part II*. Point Reyes Station, CA: Slate Creek Press.

Ridgeley RS, Allnutt TF, Brooks T, NcNicol DK, Mehlman DW, Young BE, Zook JR, BirdLife International. 2012. Digital distribution maps of the birds of the Western Hemisphere, version 5.0. In: *BirdLife International & NatureServe Bird species distribution maps of the world*, Version 2.0. Arlington: BirdLife International, Cambridge, UK and NatureServe.

Ridley HN. 1930. *The dispersal of plants throughout the world*. Ashford: L. Reeve & Co. Ltd.

Sánchez MI, Hortas F, Figuerola J, Green AJ. 2012. Comparing the potential for dispersal via waterbirds of a native and an invasive brine shrimp. *Freshwater Biology* **57**:1896–1903 DOI 10.1111/j.1365-2427.2012.02852.x.

Schuster RM. 1983. Phytogeography of the bryophyta. In: Schuster RM, ed. *New manual of bryology*. Nichinan, Miyazaki: Hattori Botanical Laboratory. Chapter 10.

Shaw AJ. 1986. A new approach to the experimental propagation of bryophytes. *Taxon* **35**:671–675 DOI 10.2307/1221609.

Sorensen AE. 1986. Seed dispersal by adhesion. *Annual Review of Ecology and Systematics* **17**:443–463 DOI 10.1146/annurev.es.17.110186.002303.

Suthers HB. 1985. Ground-feeding migratory songbirds as cellular slime mold distribution vectors. *Oecologia* **65**:526–530 DOI 10.1007/BF00379667.

Tracy DM, Schamel D, Dale J. 2002. Red Phalarope (*Phalaropus fulicarius*). In: Poole A, ed. *The birds of North America online*. Ithaca: Cornell Lab of Ornithology.

Traveset A, Riera N, Mas RE. 2001. Passage through bird guts causes interspecific differences in seed germination characteristics. *Functional Ecology* **15**:669–675 DOI 10.1046/j.0269-8463.2001.00561.x.

van Leeuwen CHA, van der Velde G, van Groenendael JM, Klaassen M. 2012. Gut travellers: internal dispersal of aquatic organisms by waterfowl. *Journal of Biogeography* **39**:2031–2040 DOI 10.1111/jbi.12004.

Viana DS, Santamaria L, Michot TC, Figuerola J. 2013. Migratory strategies of waterbirds shape the continental-scale dispersal of aquatic organisms. *Ecography* **36**:430–438 DOI 10.1111/j.1600-0587.2012.07588.x.

van Zanten BO. 1978. Experimental studies on trans-oceanic long-range dispersal of moss spores in the Southern Hemisphere. *Journal of the Hattori Botanical Laboratory* **44**:455–482.

Warner GM, French DW. 1970. Dissemination of fungi by migratory birds: survival and recovery of fungi from birds. *Canadian Journal of Botany* **48**:907–910 DOI 10.1139/b70-127.

Wen J, Ickert-Bond SM. 2009. Evolution of the Madrean-Tethyan disjunctions and the North and South American amphitropical disjunctions in plants. *Journal of Systematics and Evolution* **47**:331–348 DOI 10.1111/j.1759-6831.2009.00054.x.

Wilkinson DM. 2010. Have we underestimated the importance of humans in the biogeography of free-living terrestrial microorganisms? *Journal of Biogeography* **37**:393–397 DOI 10.1111/j.1365-2699.2009.02236.x.

Wilkinson DM, Koumoutsaris S, Mitchell EAD, Bey I. 2012. Modelling the effect of size on the aerial dispersal of microorganisms. *Journal of Biogeography* **39**:89–97 DOI 10.1111/j.1365-2699.2011.02569.x.

Wirtz N, Printzen C, Lumbsch HT. 2008. The delimitation of Antarctic and bipolar species of neuropogonoid *Usnea* (Ascomycota, Lecanorales): a cohesion approach of species recognition for the *Usnea perpusilla* complex. *Mycological Research* **112**:472–484 DOI 10.1016/j.mycres.2007.05.006.

Chapter 3: Complete plastid and mitochondrial genomes and nuclear ribosomal repeat of the Chilean sub-Antarctic endemic *Tetraplodon fuegianus* (Bryophyta) with insights into intraspecific polymorphism.

Abstract

The Bryophyta (mosses) are the second largest lineage of land plants and are ecologically and evolutionarily significant, yet genomic resources for this group are limited. Complete annotated chloroplast and mitochondrial genomes and the nuclear ribosomal repeat have only been sequenced for the model species *Physcomitrella patens*. As next generation sequencing approaches become increasingly utilized in the study of early plant lineages, the small size of individuals and ambiguity in defining an individual may present a sampling challenge with potential downstream implications. Here we present the complete annotated chloroplast and mitochondrial genomes and nuclear ribosomal repeat for seven patches of the southernmost South American endemic dung moss *Tetraplodon fuegianus*. We show that discrete patches of *T. fuegianus* represent more than a single genetic individual, with chloroplast heterogeneity reflecting inter-individual variation, and nuclear ribosomal heterogeneity reflecting intra-individual variation. The discovery of haplotype heterogeneity within single patches of moss has implications for both DNA sampling and downstream analyses, as well as the monitoring of number of individuals for conservation assessments, such as those used by the IUCN, where patches are treated as individuals.

Background

The Bryophyta (mosses) comprise about 12,000 species (Crosby et al. 1999), occur on all continents and in all biomes (Vanderpoorten & Goffinet 2009) and compose a

lineage whose origin predates that of vascular plants (Wickett et al. 2014). Despite their diversity, and their ecological and evolutionary significance, genomic resources for the bryophyta are limited relative to other groups of plants such as angiosperms (Wu, Tembrock, & Ge 2015). Methodological advances have simplified the sequencing of complete chloroplast (CP) and mitochondrial (MT) genomes and the nuclear ribosomal repeat (NRR) (Liu et al. 2013), but the number of assembled and annotated genomes remains low for the mosses, especially for the CP, as well as the NRR. To date, three annotated moss CP genomes have been published: the model species *Physcomitrella patens* (Sugiura et al. 2003), *Syntrichia ruralis* (Oliver et al. 2010; published under the synonym *Tortula ruralis*), and *Tetraphis pellucida* (Bell et al. 2014). The complete NRR has only been sequenced and annotated for two closely related mosses, *Funaria hygrometrica* (Capesius 1997) and *Enthostodon obtusus* (Liu et al. 2013) both from the Funariaceae. Sixteen moss MT genomes have been published (Liu, Medina, & Goffinet 2014; Sawicki et al. 2014; Bell et al. 2014; Liu et al. 2011; Terasawa et al. 2007), with the majority being recently contributed by Liu, Medina, and Goffinet (2014). The complete CP, MT, and NRR are relatively easy to sequence using basic next generation sequencing approaches due to their high copy numbers, and serve as significant resources as they are the source of the majority of phylogenetic markers currently used to delimit plant species (Stech & Quandt 2010; Álvarez & Wendel 2003).

One of the challenges associated with genomic studies in bryophytes, and other small non-model organisms is acquiring sufficient DNA, especially to facilitate next generation sequencing approaches. In DNA based analyses of small non-model organisms, such as most non-vascular plants, adjacent wild collected plant material is typically pooled to

yield sufficient DNA. Unlike pooling of other small organisms, such as insects (Emerson et al. 2010), in bryophytes it is typically unknown if a single or multiple genets and ramets are sampled, as individual diaspores may yield several individuals and discrete patches may arise from several spores (Mägdefrau 1982). Discrete bryophyte patches are regarded as individuals in assessments of IUCN red-list species status for bryophytes (Hallingback et al. 1998). However, sequencing of the *Syntrichia ruralis* CP genome from multiple cultured spores taken from distinct parents of the same patch of moss revealed 27 polymorphisms, representing inter-individual variation (Oliver et al. 2010). If multiple spores of the same species germinate on a substrate, especially for monoecious species, which represent about one half of all bryophyte species (Wyatt 1982), it may be impossible to distinguish how many individuals (i.e. origin from the same spore) are present.. As next generation sequencing approaches become increasingly common, it is important to explore the effects of this anonymous pool sampling approach on intra-sample variability, especially for studies that extend beyond descriptive genomic reports, such as phylogenetic and phylogeographic studies. More fundamentally, information related to the composition of discrete patches in bryophytes is scant, and is necessary to better understand how to conduct analyses of high throughput bryophyte DNA data assessment of population sizes or numbers in conservation planning.

Here we present the complete CP and MT genomes and NRR sequence of the monoecious endemic dung moss *Tetraplodon fuegianus* Besch. *Tetraplodon fuegianus* represents a monophyletic lineage within the Splachnaceae found only in southernmost South America (Lewis, Rozzi, & Goffinet 2014). *T. fuegianus* is the second bryophyte, after the model species *Physcomitrella patens*, for which the complete CP and MT

genomes and NRR have been sequenced assembled and annotated, and the first to have all three sequences reported at once. Screening of all three genomic compartments within and between 7 discrete patches revealed (a) 16 polymorphic sites in the CP, 1 in the MT, and 21 in the NRR; re-sequencing of selected polymorphic sites from individual gametophyte stems with Sanger technology revealed (b) that polymorphisms within samples are due either to paralogous copies in the case of NRR or to CP haplotype heterogeneity within discrete patches of *T. fuegianus*. These results are the first detailed description of variation within multiple discrete patches of a bryophyte species.

Results & Discussion

Chloroplast genome

The *Tetraplodon fuegianus* chloroplast genome is the fourth completely sequenced and annotated moss CP genome. It is 123,664 – 123,675 bp long, with a quadripartite architecture. The genome is composed of an 84,946 bp large single copy (LSC) and 18,691 bp small single copy (SSC) region with a 10,016 bp inverted repeat (IR; Average sizes reported; Table 1). The GC content of 28.7% is similar to that of *Physcomitrella patens* (28.5%; Sugiura et al. 2003) and *Tetraphis pellucida* (29.4%; Bell et al. 2014). The genome comprises 82 known protein coding genes, 33 tRNA genes and four rRNA genes (with genes duplicated in the IR counted once). A 210 codon open reading frame (ORF) located between the *psb*J and *pet*A genes, appears to be homologous to ORF 197 reported in *Syntricha ruralis* (NC_012052; Oliver et al. 2010) and *P. patens* (Sugiura et al. 2003). All rRNA genes are located in the IR, with *trn*V-GAC and *trn*A-GTT as the terminal genes of each IR. The chloroplast genome of *T. fuegianus* is collinear with *S.*

ruralis (Oliver et al. 2010) and *T. pellucida* (Bell et al. 2014), including the absence of the *pet*N gene, which is present in *P. patens* (Sugiura et al. 2003), and the 71Kb inversion in the LSC diagnostic of the Funariales and Encalyptales (Goffinet et al. 2007). The *rpo*A gene is also absent in *T. fuegianus*, consistent with the loss of this gene in all arthrodontous moss lineages surveyed (Goffinet et al. 2005).

Chloroplast polymorphism

Screening and comparing the complete chloroplast genome for seven discrete patches revealed 16 polymorphic sites (Table 2). Three sites had alleles fixed within patches but variable among patches, whereas the remaining 13 sites were variable within at least one patch. Across all patches, variable positions represent 0.013% of the *T. fuegianus* CP genome. Seven within patch polymorphisms were unique to a single patch, four polymorphisms were shared by two patches, and two polymorphisms were shared by all patches. An average of 4.29 variants (2–6) were detected within each patch. Overall, polymorphisms were represented by 12 single nucleotide substitutions, including 8 transitions and four transversions, and four indels of one to four nucleotides in intergenic regions. Five single nucleotide substitutions including two transitions and two transversions were located in coding sequences, resulting in four non-synonymous and one synonymous substitution. The most polymorphic region was the *psb*J – *pet*A intergenic spacer, with two polymorphisms detected within all sampled patches, and one detected within patches 1 and 2 (Table 2).

Sequencing of the A/G variant in the *rps*12 – *trn*V-GAC intergenic region located near the junction of the LSC and IRA (position 84,936) across eight individual gametophyte stems taken from a single patch (sample 2) recovered the A state in two

individuals stems, and the G state in the remaining six. Both forward and reverse sequences were used to confirm the character state in individuals, providing strong support for interindividual variation within discrete patches of *T. fuegianus* (Figure 2). Sampling of single *T. fuegianus* stems provides sufficient DNA for Sanger sequencing, but not for next generation sequencing approaches. Sampling of single stems and re-sequencing of putative variants allowed for distinction between intraindividual and interindividual variation. Sanger sequence chromatograms did not show evidence of double peaks at polypormphic positions, suggesting that variations not due to heteroplasmy. Interindividual variation has also been recovered across the CP of *Syntrichia* (Oliver et al. 2010), but is here confirmed with Sanger sequencing to rule out sequencing error (Wu, Tembrock, & Ge 2015) and intra-individual variation (heteroplasmy; Wolfe & Randle 2004). These results are congruent with the assumption that intraindividual variation is not a major concern when utilizing plastid markers, due to the single copy nature of the CP genome (Stech & Quandt 2010) and monoplastidic inheritance of the CP genome (Renzaglia et al. 1994; Brown & Lemmon 1990).

Throughout the *T. fuegianus* CP genome, polymorphism both within and between patches is scattered (Figure 1). Screening for variable plastid loci at the intraspecific level by targeting individual loci may prove an inefficient means of discovering polymorphism for population genetic studies of closely related samples. Variable positions in *T. fuegianus* CP genome are not located in commonly sequenced regions (Stech & Quandt 2010) for which primers are available for a variety of taxa. Comparison of complete genomes allows for the discovery of variable positions that can subsequently be targeted

directly to infer fine scale phylogenetic and population relationships (Capella-Gutierrez, Kauff, & Gabaldón 2014).

Mitochondrial genome & polymorphism

The *T. fuegianus* MT genome represents the first completely sequenced and annotated moss MT genome within the Splachnales. It is 104,742 bp long, comprising 40 protein-coding genes, 24 tRNA genes and three rRNA genes, with a GC content of 40%. The gene order and content is identical to that reported for *Anomodon attenuatus* (NC_021931) as well as all the mosses recently sampled across the moss macroevolutionary tree by Liu et al. 2014. A single polymorphic site (C/T transition) resulting in a nonsynonymous amino acid change in the second exon of the *nad*1 gene (position 72,549) within patches 2 and 3 was detected.

Nuclear Ribosomal Repeat

The *T. fuegianus* nuclear ribosomal repeat (NRR) is the first sequenced from the order Splachnales, and the first moss outside the Funariaceae. The NRR is 10,394-10,398 bp with a GC content of 54%. The sequence includes the 5S, splitting the IGS into IGS1 and IGS2 congruent with the L-type organization, i.e. the small ribosomal subunit, 5.8S, large ribosomal subunit and 5S region occurring in tandem, inferred for mosses by Wicke et al. (2011). The complete sequence is comprised of four genes coding for rRNA subunits, two internal transcribed spacers, and two intergenic spacers arranged in the following order: 18S (1,822 bp), ITS1 (231 bp), 5.8S (157 bp), ITS2 (264 bp), 26S (3,423 bp), IGS1 (543 bp), 5S (120 bp) and IGS2 (3,833 – 3,837 bp). The sequence is collinear

with that of *Funaria hygrometrica* ([X80212; Capesius 1997] & [JQ736823;(Liu et al. 2013]) and *Entosthodon obtusus* (JQ736824; Liu et al. 2013).

Nuclear ribosomal repeat polymorphism

A total of 21 polymorphisms were detected in the NRR, with five fixed within patches but variable between patches and 16 variable within patches. One polymorphism was represented within a single patch, one was shared by two patches, eight were represented within three to six patches, and six were represented within all patches. Polymorphic positions represented approximately 0.21% of the NRR sequence. An average of 11.4 polymorphisms (10 to 14) were found within patches. Overall, 15 polymorphisms were represented by single nucleotide substitutions, including nine transitions and six transversions. Indels represent six polymorphisms between patches, including five single nucleotide indels and a single two-bp indel. Variation in the length of the repeat across samples is a result of indel variation in the IGS2. Seventeen polymorphisms were observed in the second Intergenic Spacer (IGS2), three in IGS1, and one in ITS2.

Sequencing of the A/G polymorphism in ITS2 for 2 individual gametophyte stems from patches 1 and 2 recovered signals for both A and G character states (i.e. double peaks) at the same position within each single gametophyte stem DNA extract at the putative variant site (Figure 2). Thus the variation recovered in the NRR is the result of fixed heterozygosity. ITS2 is one of the most widely used markers in plant phylogenetics (Stech & Quandt 2010; Álvarez & Wendel 2003) and paralogy, or similarly, fixed heterozygosity at this locus has been previously discussed (Poczai & Hyvönen 2010;

Vanderpoorten, Goffinet, & Quandt 2006; Álvarez & Wendel 2003). ITS paralogy and fixed heterozygosity is widely recognized as a potential source of error in phylogenetic analyses, however, concerted evolution is widely believed to eliminate intra-genomic heterogeneity (Hillis & Dixon 1991). In bryophytes intra-individual ITS variation has been demonstrated in *Tortula muralis* (Košnar et al. 2012) and *Plagiomnium* (Harris 2008). Intra-individual variation in the NRR of *T. fuegianus* is likely due to a lag in concerted evolution relative to the mutation rate in the variable regions, as proposed for *T. muralis*. For the commonly used marker ITS2, ambiguity at a single site should not pose significant issues for phylogenetic analysis (Lewis, Rozzi, & Goffinet 2014; Hillis & Dixon 1991). Intra-individual heterogenity may however also be useful in the reconstruction of evolutionary relationships (Razafimandimbison, Kellogg, & Bremer 2004). Whether the number of intra-genomic variant sites in the IGS could be a challenge or inform for phylogenetic analyses is beyond the scope of this study, however further should be explored in the future.

Conclusion

We have assembled and annotated the complete chloroplast and mitochondrial genomes and nuclear ribosomal repeat across seven discrete patches of the Southern South American endemic dung moss *Tetraplodon fuegianus* from high-throughput Illumina shotgun sequence data. Gene content is conserved and collinear for all sequences relative to *Syntrichia ruralis* chloroplast genome (NC_012052), *Anomodon attenuates* mitochondrial genome (NC_021931) and *Funaria hygrometrica* nuclear ribosomal repeat (X80212). Variation has been detected both within and between patches

in each of the genomic compartments, though with variable frequencies consistent with relative rates of evolution between CP, MT, and Nuclear genomic compartments (Wolfe, Li, and Sharp 1987). Sanger sequencing of select CP and NRR polymorphisms from individual stem DNA extracts confirms inter-individual variation as the source of CP polymorphism within patches, and intra-individual variation (fixed heterozygosity) as the source of NRR polymorphism detected within patches.

Pooling of individuals is often necessary for genetic studies of small non-model organisms. In bryophytes it may be impossible to distinguish if a discrete patch or mat of moss represents a single or multiple individuals, and thus it is unclear if pooling multiple stems or sporophytes equates to pooling individuals. Here we show that a minimum of two haplotypes are present in all patches of *Tetraplodon fuegianus*, which were sampled by pooling stems and or sporophytes within discrete patches of moss. Due to the method of pooling, it is unclear how many haplotypes are present in each sampled patch of *T. fuegianus*, however it is clear that there is more than one present in each patch. Our results suggest that where the lines between individuals are indiscernible, measures should be taken to identify and account for variation within samples, and choose analyses appropriate for the data (Gompert et al. 2010). In organismal groups where this may be the case, we propose that the most efficient means of accounting for this variation is through bioinformatic processing of samples prior to cross sample analyses (Emerson et al. 2010).

Methods

Sampling

Discrete patches of moss were collected in the Magellanes region of Chile, seven from the south western coast of Isla Grande de Tierra del Fuego in the Antarctic province and one from the Brunswick Peninsula in the Magellanes province (Table 3). DNA was extracted from multiple leafy gametophytes and / or sporophytes from each patch of moss. Libraries were prepared according to TruSeq protocol, with DNA from each patch barcoded, pooled, and sequenced on the Illumina HiSeq2000 platform (samples 1-7), following Liu, Medina, and Goffinet (2014). An additional patch (sample 8) was sequenced on the 454 platform following Liu et al. (2013) to assist with assembly of chloroplast and mitochondrial genome assemblies, but was not used for subsequent analyses due to poor coverage. The previous sections discussed samples 1 – 7 only. Sample 8 is discussed only in the methods where used to facilitate assemblies.

De novo assembly from shotgun data

Paired-end reads were joined in CLC genomics workbench v. 6.5, with quality scores retained. Unpaired and failed reads were discarded. Paired reads for each sample were de novo assembled with reads mapped back to initial contigs. Parameter settings required 80% similarity across at least 50% of reads mapped back to contigs, with minimum contig length of 1000, mismatch cost of 2, and insertion and deletion costs of 3. De novo contigs were blasted (blastn – somewhat similar setting) against a custom database including chloroplast genomes (CP) from *Syntrichia ruralis* (NC_012052) and

Physcomitrella patens (NC_005087), mitochondrial genomes (MT) from *Anomodon attenuatus* (NC_021931) and *Physcomitrella patens* (NC_007945), and the complete *Funaria hygrometrica* nuclear ribosomal repeat (NRR) sequence (X80212).

Assembly of NRR, MT genome, and CP sequences

Blast identified MT and NRR contigs for all patches were combined and de novo assembled in Geneious v. 7.0.4 and aligned to the *Anomodon attenuatus* MT (NC_021931) and *Funaria* NRR (X80212) sequences, respectively. The consensus pre-draft *T. fuegianus* MT and NRR were annotated based on the references, and manually edited using ExPASy translate tool for coding regions (Gasteiger et al. 2003) and tRNAscan SE 1.21 webserver for tRNAs (Lowe & Eddy 1997).

De novo assemblies for each sample did not generate significant CP genome coverage based on alignment to *Physcomitrella patens* (NC_005087) and *Syntrichia ruralis* (NC_012052). CP assembly was accomplished by pooling reads across 4 samples (1, 2, 7, and 454 sequenced sample). Pooled reads were de novo assembled according to parameters used for MT and NRR assembly and contigs were blasted against *Physcomitrella* and *Syntrichia* CP genomes to identify CP contigs. CP contigs were de novo assembled in Geneious 7.0.4, and aligned to *Physcomitrella* and *Syntrichia* CP genomes with the second inverted repeat removed to assess success of *de novo* CP genome assembly. Junctions between the large single copy (LSC), inverted repeats A and B (IRA; IRB), and small single copy (SSC) were sequenced on an ABI3100 Genetic Analyzer (Applied Biosystems, Grand Island, NY, USA), to confirm the pre-draft assembly of the *T. fuegianus* CP genome. The CP pre-draft sequence was annotated

based on the complete *Syntrichia* CP genome and manually edited as described above.

Paired reads from each sample were mapped randomly (random mapping allows for mapping of repeat sequences, i.e. the inverted repeats) to the pre-drafts (similarity fraction of 0.95 over at least 0.90 of the read length with equal mismatch, insertion and deletion costs), and draft consensus sequences were generated for NRR, CP, and MT sequences for each samples 1 - 7. Sample 8, sequenced on the 454 platform, was not used in in further analyses due to poor coverage.

Variant detection

Paired reads were mapped randomly back to the draft, using the same parameters as were used in mapping reads to the pre-draft, and a final reads consensus was generated with ambiguity codes denoting variant sites. Ambiguity codes were used to denote variants in sites only if there was a minimum depth of 100x quality filtered reads. Bases with a frequency lower than 0.30 were filtered out as noise and only variants represented by a minimum nucleotide count of 30 were considered in the use of ambiguity codes. Final consensus sequences were aligned in Geneious 7.1.2, using the progressive Mauve algorithm (Darling et al. 2004), part of the Mauve genome alignment plugin. In order to confirm select variants and distinguish between sequencing error (Wu, Tembrock, & Ge 2015) inter-individual and intra-individual variation (i.e. paralogy), DNA was re-extracted, but from individual gametophyte stem tips sampled randomly across a sample. Sanger sequencing following Lewis, Rozzi, and Goffinet (2014) was used to sequence selected variants. Variant positions were checked for double chromatogram peaks and variation across individuals from the same sample.

References

Álvarez, I., and J. F. Wendel. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29: 417–434.

Bell, N. E., J. L. Boore, B. D. Mishler, and J. Hyvönen. 2014. Organellar genomes of the Four-Toothed moss, *Tetraphis Pellucida*. *BMC Genomics* 15: 383.

Brown, R. C., and B. E. Lemmon. 1990. Monoplastidic cell division in lower land plants. *American Journal of Botany* 77: 559.

Capella-Gutierrez, S., F. Kauff, and T. Gabaldón. 2014. A phylogenomics approach for selecting robust sets of phylogenetic markers. *Nucleic Acids Research* 42: e54.

Capesius, I. 1997. Analysis of the ribosomal RNA gene repeat from the moss *Funaria Hygrometrica*. *Plant Molecular Biology* 33: 559–564.

Crosby, M. R., R. E. Magill, B. Allen, and S. He. 1999. A checklist of the mosses. St. Louis, Missouri, USA.

Darling, A. C. E., B. Mau, F. R. Blattner, and N. T. Perna. 2004. Multiple alignment of Cconserved genomic sequence with rearrangements. *Genome Research* 14: 1394–1403.

Emerson, K. J., C. R. Merz, J. M Catchen, P. A. Hohenlohe, W. A. Cresko, W. E. Bradshaw, and C. M. Holzapfel. 2010. Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences of the United States of America* 107: 16196-16200.

Gasteiger, E., A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, and A. Bairoch. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* 31: 3784–3788.

Goffinet, B., N. J. Wickett, A. J. Shaw, and C. J. Cox. 2005. Phylogenetic significance of the *rpo*A loss in the chloroplast genome of mosses. *Taxon* 54: 353.

Goffinet, B., N. J. Wickett, O. Werner, R. M. Ros, A. J. Shaw, and C. J. Cox. 2007. Distribution and phylogenetic significance of the 71-Kb inversion in the plastid genome in Funariidae (Bryophyta). *Annals of Botany* 99: 747–753.

Gompert, Z., M. L. Forister, J. A. Fordyce, C. C. Nice, R. J. Williamson, and A. Buerkle. 2010. Bayesian analysis of molecular variance in pyrosequences quantifies population genetic structure across the genome of *Lycaeides* butterflies. Molecular Ecology 19: 2455-2473.

Hallingback, T., N. Hodgetts, G. Raeymaekers, R. Schumacker, L. Soderstrom, N. Stewart, and J. Vana. 1998. Guidelines for application of the revised IUCN threat categories to bryophytes. *Lindbergia* 23: 6–12.

Harris, E. S. J. 2008. paraphyly and multiple causes of phylogenetic incongruence in the moss genus *Plagiomnium* (Mniaceae). *Taxon* 57: 417–433.

Hillis, D. M, and M. T Dixon. 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. *The Quarterly Review of Biology* 66: 411–453.

Košnar, J., M. Herbstová, F. Kolář, P. Koutecký, and J. Kučera. 2012. A case study of intragenomic ITS variation in bryophytes : assessment of gene flow and role of polyploidy in the origin of european taxa of the *Tortula Muralis* (Musci : Pottiaceae) Complex. *Taxon* 61: 709–720.

Lewis, L. R., R. Rozzi, and B. Goffinet. 2014. Direct long-distance dispersal shapes a new world amphitropical disjunction in the dispersal-limited dung moss *Tetraplodon* (Bryopsida: Splachnaceae). *Journal of Biogeography* 41: 2385–2395.

Liu, Y., L. L. Forrest, J. D. Bainard, J. M. Budke, and B. Goffinet. 2013. Organellar genome, nuclear ribosomal DNA repeat unit, and microsatellites isolated from a small-scale of 454 GS FLX sequencing on two mosses. *Molecular Phylogenetics and Evolution* 66: 1089–1094.

Liu, Y., R. Medina, and B. Goffinet. 2014. 350 Million years of mitochondrial genome stasis in mosses, an early land plant lineage. *Molecular Biology and Evolution* 31: 8–13.

Liu, Y., J. Y. Xue, B. Wang, L. Li, and Y. Long Qiu. 2011. The mitochondrial genomes of the early land plants *Treubia Lacunosa* and *Anomodon Rugelii*: dynamic and conservative evolution. *PLoS ONE* 6: e25836.

Lowe, T. M., and S. R. Eddy. 1997. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* 25: 955–964.

Mägdefrau, K. 1982. Life-Forms of Bryophytes. In *Bryophyte Ecology*, edited by A. J. E. Smith. Springer Netherlands.

Oliver, M. J., A. G. Murdock, B. D. Mishler, J. V. Kuehl, J. L. Boore, D. F. Mandoli, K. D. E. Everett, P. G. Wolf, A. M. Duffy, and K. G. Karol. 2010. Chloroplast genome sequence of the moss *Tortula Ruralis*: gene content, polymorphism, and structural arrangement relative to other green plant chloroplast genomes. *BMC Genomics* 11: 143.

Poczai, P., and J. Hyvönen. 2010. Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects. *Molecular Biology Reports* 37: 1897–912.

Razafimandimbison, S. G., E. A. Kellogg, and B. Bremer. 2004. Recent origin and phylogenetic utility of divergent ITS putative pseudogenes: a case study from *Naucleeae* (Rubiaceae). *Systematic Biology* 53: 177–192.

Renzaglia, K. S., R. C. Brown, B. E. Lemmon, J. G. Duckett, and R. Ligrone. 1994. Occurrence and phylogenetic significance of monoplastidic meiosis in liverworts. *Canadian Journal of Botany* 72: 65–72.

Sawicki, J., K. Wo, H. Bednarek-Ochyra, R. Ochyra, and M. Szczecin. 2014. The complete mitochondrial genome of the South American endemic moss *Codriophorus Laevigatus* (Grimmiaceae). *Mitochondrial DNA* 1736: 1–2.

Stech, M., and D. Quandt. 2010. 20,000 Species and five key markers: the status of molecular bryophyte phylogenetics. *Phytotaxa* 9: 196–228.

Sugiura, C., Y. Kobayashi, S. Aoki, C. Sugita, and M. Sugita. 2003. Complete chloroplast DNA sequence of the moss *Physcomitrella Patens*: evidence for the loss and relocation of *rpo*A from the chloroplast to the nucleus. *Nucleic Acids Research* 31: 5324–5331.

Terasawa, K., M. Odahara, Y. Kabeya, T. Kikugawa, Y. Sekine, M. Fujiwara, and N. Sato. 2007. The mitochondrial genome of the moss *Physcomitrella Patens* sheds new light on mitochondrial evolution in land plants. *Molecular Biology and Evolution* 24: 699–709.

Vanderpoorten, A., and B. Goffinet. 2009. *Introduction to Bryology*. Cambridge UK: Cambridge University Press.

Vanderpoorten, A., B. Goffinet, and D. Quandt. 2006. Utility of the internal transcribed spacers of the 18S-5.8S-26S nuclear ribosomal DNA in land plant systematics, with special emphasis on bryophytes. In *Plant Genome: Biodiversity and Evolution, Vol. 2, Part B.*, edited by A.K. Sharma and A. Sharma, 385–407. Enfield, New Hampshire: Science Publishers.

Wicke, S., A. Costa, J. Muñoz, and D. Quandt. 2011. Restless 5S: the re-Arrangement(s) and evolution of the nuclear ribosomal DNA in land plants. *Molecular Phylogenetics and Evolution* 61: 321–32.

Wickett, N. J., S. Mirarab, N. Nguyen, T. Warnow, E. Carpenter, N. Matasci, S. Ayyampalayam, et al. 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences of the United States of America* 111: E4859–E4868.

Wolfe, A. D., and C. P. Randle. 2004. Recombination, heteroplasmy, haplotype polymorphism, and paralogy in plastid genes: implications for plant molecular systematics plastid genes. *Systematic Botany* 29: 1011–1020.

Wolfe, K. H., W. H. Li, and P. M. Sharp. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of America* 84: 9054–8.

Wu, Z., L. R. Tembrock, and S. Ge. 2015. Are differences in genomic data sets due to true biological variants or errors in genome assembly: an example from two chloroplast genomes. *Plos One* 10: e0118019.

Wyatt, R. 1982. Population Ecology of bryophytes. *Journal of the Hattori Botanical Laboratory* 52: 179–198.

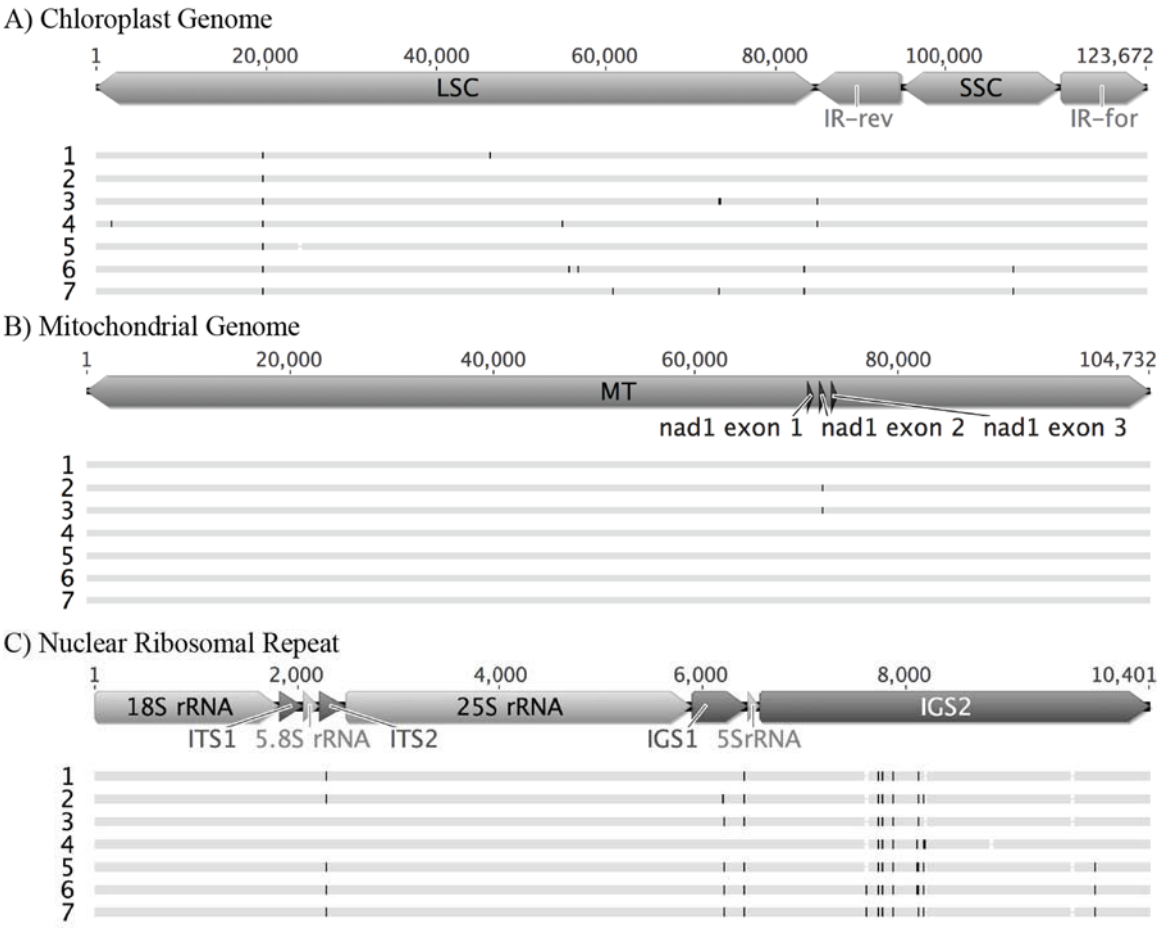


Figure 1. (a) Chloroplast genome, (b) nuclear ribosomal repeat, and (c) mitochondrial genome alignments showing within patch variants. Partial sequence annotations are shown above the Chloroplast and nuclear ribosomal repeat alignments for reference. Consensus sequences are ordered from top to bottom 1 to 7 for each alignment, and within patch variants are shown by dark marks in the sequence, and indels are shown by white marks. For a detailed list of putative variants see Table 2.

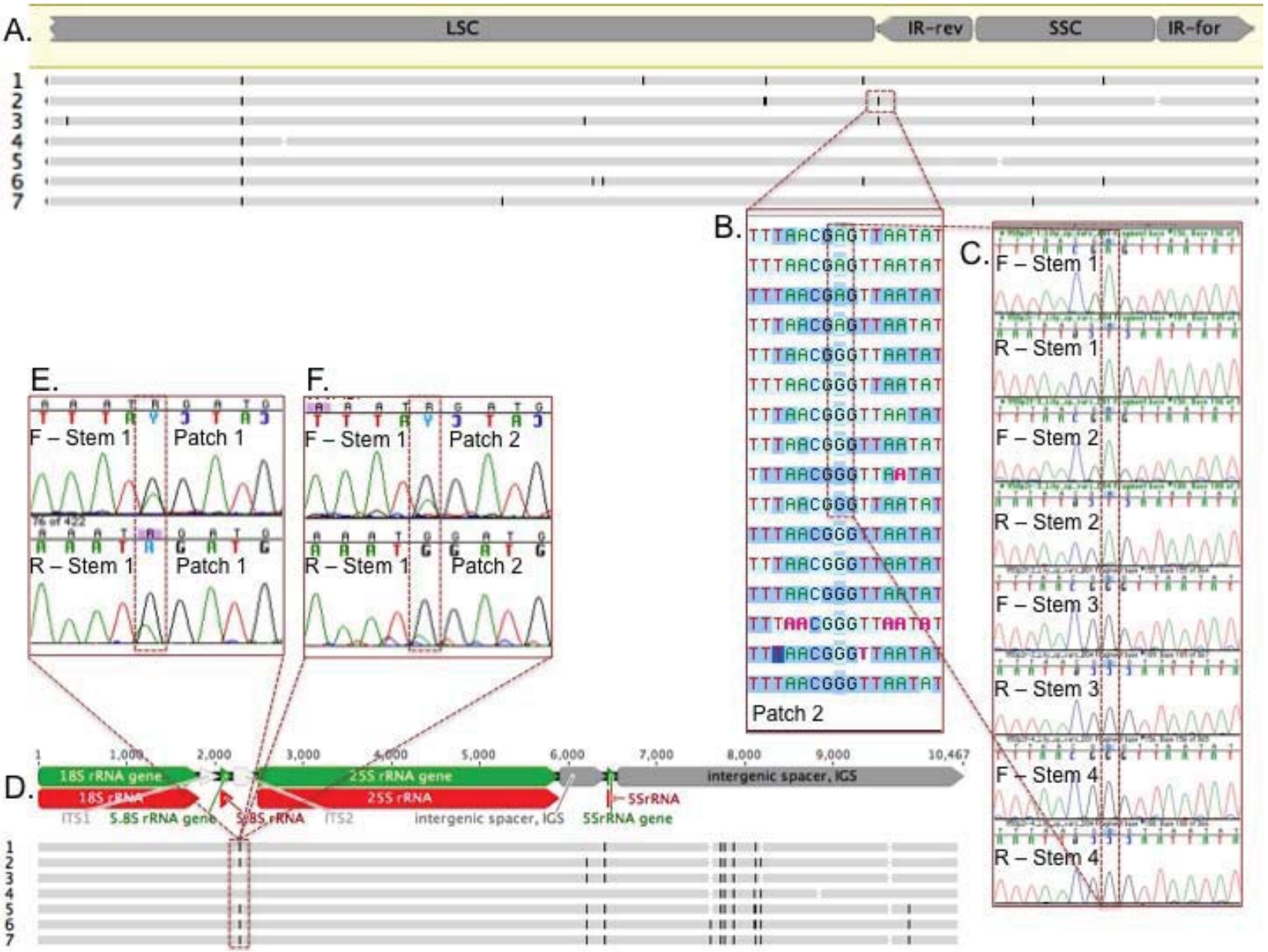

A.
LSC
IR-rev
SSC
IR-for
B.
Patch 2
C.
F – Stem 1
R – Stem 1
F – Stem 2
R – Stem 2
F – Stem 3
R – Stem 3
F – Stem 4
R – Stem 4
E.
F – Stem 1
Patch 1
R – Stem 1
Patch 1
F.
F – Stem 1
Patch 2
R – Stem 1
Patch 2
D.
1
1,000
2,000
3,000
4,000
5,000
6,000
7,000
8,000
9,000
10,467
18S rRNA gene
25S rRNA gene
intergenic spacer, IGS
18S rRNA
25S rRNA
5SrRNA
ITS1
5.8S rRNA gene
5.8S rRNA
ITS2
intergenic spacer, IGS
5SrRNA gene

Figure 2. Following identification of putative variants, single gametophyte stems were re-sampled, and select variants were sequened from single stem DNA extracts to differentiate between inter-individual and intra-individual variation. Forward (F) and reverse (R) reads are shown for all Sanger sequence chromatograms. Distinctions are made between distinct gametophyte stems (i.e. Stem 1, Stem 2) and distinct patches of moss. Numbers used to distinguish patches of moss correspond to patch numbers in sequence alignments in (D.) and (A.) and Figure 1. **(**A.) Chloroplast genome alignment across seven patches of moss showing sanger sequencing results from single gametophyte stems in (B.) and (C.). (D.) Nuclear ribosomal repeat alignment across seven patches of moss with sanger sequencing results from single gametophyte stems from two distinct patches of moss (E.) and (F.). Partial sequence annotations are shown above the Chloroplast and nuclear ribosomal repeat alignments for reference. Consensus sequences are ordered from top to bottom 1 to 7 for each alignment, and within patch variants are shown by dark marks in the sequence, and indels are shown by white marks.

	Chloroplast				Mitochondrial	NRR
	LSC	**SSC**	**IR (x2)**	**Total**		
Size (bp)	84,946	18,691	10,016	123,669	104,741	10,401
Protein	68/ 53,833 / 63.4	14 / 15,336 / 82.1	0	82 / 65,169 / 52.7	40 / 32,143 / 30.7	0
tRNA	26 / 1,987 / 2.3	2 / 154 / 0.8	5 / 363 / 3.62	33 / 2,867 / 2.3	24 / 1,801 / 1.7	0
rRNA	0	0	4 / 4,522 / 45.1	4 / 9,044 / 7.3	3 / 5,232 / 5.0	4 / 5522 / 53.1
G/C (%)	26	25.3	43.3	28.7	40	54

Table 1. Gene content, size, and GC content of chloroplast and mitochondrial genomes and nuclear ribosomal repeat for *T. fuegianus*. Inverted repeat (IR) genes and sequence length are only counted once. ORF 210 is not counted in number of protein coding genes. Protein, tRNA and rRNA genes include number of genes / total base pairs / percent of total size for each sequence. The information on protein coding genes includes coding regions only, thus introns are not included. Abbreviations used are: Large single copy (LSC); small single copy (SSC); inverted repeat (IR); transfer RNA (tRNA); ribosomal RNA (rRNA); nuclear ribosomal repeat (NRR).

Chloroplast Genome 123,664 – 123,675 bp					
Position	**Location**	**Length (bp)**	**Variant**	**Amino Acid**	**Within patches**
1,784	rpl2 (LSC)	1	R (A or G) TS	Q	3
19,649	psbJ-petA (LSC)	1	W (A or T) TV	n/a	All
19,654		1	Y (C or T) TS	n/a	All
19,660		1	W (A or T) TV	n/a	1, 2
24,035	ycf4-psal (LSC)	2	AT deletion	n/a	4
46,280	trnT(GGU)-trnG(UUC) (LSC)	1	Y (C or T) TS	n/a	7
47,662	ycf2	1	G/T TV	H / N	Inter-patch
54,850	psbA-trnK(UUU) IGR (LSC)	1	R (A or G) TS	V	3
55,639	matK-trnK(UUU) (LSC)	1	R (A or G) TS	G / S	6
56,686		1	Y (C or T) TS	H / Y	6
60,827	psbI-trnS(GCU) (LSC)	1	R (A or G) TS	n/a	1
61,851	ycf12-trnG(UCC) (LSC)	1	A / G TS	n/a	Inter-patch
69,665	rpoC2 (LSC)	1	A / T TV	N / I	Inter-patch
83,334	ndhB-rps7 (LSC)	1	W (A or T) TV	n/a	1, 6
84,936	rps12-trnV(GAC) (IRA) *	1	R (A or G) TS	n/a	2*, 3
108,033	psaC-ndhD (SSC)	4	WWWW (A or T) TV	n/a	1, 6
Mitochondrial Genome 104,732 bp					
Position	**Location**	**Length (bp)**	**Variant**	**Amino Acid**	**Within patches**
72,549	nad1 exon 2	1	Y (C or T) TS	P / L	3, 2
Nuclear Ribosomal Repeat 10,394-10,398 bp					
Position	**Location**	**Length (bp)**	**Variant**	**Amino Acid**	**Within patches**
2,289	ITS2	1	R (A or G) TS	n/a	1*, 2*, 5, 6, 7
6,202	IGS1	1	R (A or G) TS	n/a	2, 3, 5, 6, 7

6,408		1	M (A or C) TV	n/a	1, 2, 3, 5, 6, 7
6,410		1	R (A or G) TS	n/a	1, 2, 3, 5, 6, 7
7,602	IGS2	1	W (A or T) TV	n/a	6, 7
7,604		1	T deletion	n/a	Inter-patch
7,606		2	AT deletion	n/a	Inter-patch
7,612		1	G deletion	n/a	Inter-patch
7,720		1	M (A or C) TV	n/a	All
7,728		1	M (A or C) TV	n/a	All
7,763		1	Y (C or T) TS	n/a	All
7,770		1	R (A or G) TS	n/a	All
7,865		1	Y (C or T) TS	n/a	All
7,871		1	Y (C or T) TS	n/a	All
8,101		1	M (A or C) TV	n/a	4,5,6
8,120		1	Y (C or T) TS	n/a	1, 2,3, 5, 6, 7
8,174		1	K (G or T) TV	n/a	2, 4, 5, 6, 7
8,181		1	R insertion	n/a	4
8,840		1	T deletion	n/a	Inter-patch
9,640		1	C insertion	n/a	Inter-patch
9,853		1	R (A or G) TS	n/a	4, 6, 7

Table 2. Polymorphic positions in the *T. fuegianus* chloroplast and mitochondrial genomes and nuclear ribosomal repeat recovered from within and across patch screening. For multi-nucleotide variants, the starting position is listed. Positions and patches marked with an "*" were re-sequenced with Sanger Sequencing. Abbreviations used are: Large single copy (LSC); small single copy (SSC); inverted repeat (IR); transition (TS); transversion (TV); not applicable (n/a). Amino acid abbreviations follow standard genetic code.

Sample (Patch)	Collector & Collection #	Latitude	Longitude	Locality details
1	Lewis 998	-53.39414	-71.2416	XII region, Prov. Magellanes, Brunswick peninsula, outside of Laguna Parilliar, km marker 20. Growing in empetrum heath, magellanic moarland.
2	Lewis 950	-54.77608	-69.59472	XII region, Prov. Antarctica Chilena, Comuna de Cabo de Hornos, Darwin Cordillera. Parque Nacional Alberto de Agostini, Isla Grande de Tierra del Fuego, extreme SW end of Bahia romanche. Open forest with gunnera understory, between rocks.
3	Lewis 952	-54.77559	-69.59238	XII region, Prov. Antarctica Chilena, Comuna de Cabo de Hornos, Darwin Cordillera. Parque Nacional Alberto de Agostini, Isla Grande de Tierra del Fuego, extreme SW end of Bahia romanche. Small spit of land between two arms of Ventisquero Romanche.
4	Lewis 948	-54.77633	-69.59486	XII region, Prov. Antarctica Chilena, Comuna de Cabo de Hornos, Darwin Cordillera. Parque Nacional Alberto de Agostini, Isla Grande de Tierra del Fuego, extreme SW end of Bahia romanche. Rocky shore at interface between open forest and shoreline rocks.
5	Larrain 34895 (a)	-54.879306	-69.397778	XII region, Prov. Antarctica Chilena, Comuna de Cabo de Hornos, Darwin Cordillera. Parque Nacional Alberto de Agostin, Isla Grande de Tierra del Fuego, NE arm of fiordo Pia, Bahia Terra Australis, near glacier Ventisquero Alemania.
6	Larrain 34895 (b)	-54.879306	-69.397778	XII region, Prov. Antarctica Chilena, Comuna de Cabo de Hornos, Darwin Cordillera. Parque Nacional Alberto de Agostin, Isla Grande de Tierra del Fuego, NE arm of fiordo Pia, Bahia Terra Australis, near glacier Ventisquero Alemania.
7	Lewis 936	-54.9493	-69.12909	XII region, Prov. Antarctica Chilena, Comuna de Cabo de Hornos, Darwin Cordillera. Parque Nacional Alberto de Agostin, Isla Grande de Tierra del Fuego. W. coast of Caleta Olla, at E end of Brazo Noroesta of beagle channel. Graminoid tundra.
8	Goffinet 10507	-54.77559	-69.59238	XII region, Prov. Antarctica Chilena, Comuna de Cabo de Hornos, Darwin Cordillera. Glacier Pia; subantarctic vegetation with scattered trees, and abundant bryo-lichen vegetattion over rocks. Forming large cushions over rock in moss vegetation.

Table 3. *T. fuegianus* samples included in study. Sample #8 was sequenced on the 454 platform and was only used to facilitate assemblies, all others were sequenced on the Illumina HiSeq Platform and were used in variant analyses.

Chapter 4: Resolving amphitropical phylogeographic histories in the dung moss *Tetraplodon* (Bryopsida: Splachnaceae): A RAD-seq approach.

Abstract: Amphitropical disjunctions exist across a diversity of organisms, but are particularly prevalent and common at or below the level of species in mosses. *Tetraplodon* is widespread in northern Laurasia and disjunct in high elevation tropical regions and in southern South America, following dispersal events in the last xxx my. The crown group of *Tetraplodon* exhibits an amphitropical distribution, characterized by three disjunctions. Variation in discrete genetic loci was insufficient to resolve phylogenetic and hence geographic structure within this lineage. Here we apply a RAD-seq approach to generate thousands of loci for samples across the geographic range of the amphitropical lineage in order to confirm the monophyly of southernmost South American endemic *T. fuegianus*, to estimate the geographic origin of the ancestor to *T. fuegianus*, and to infer overall geographic structure in the lineage. Analysis of RAD-loci using maximum likelihood, species tree estimation, and individual assignment allowed for resolution of a monophyletic *T. fuegianus*, which shares an ancestor with populations from northwestern North America. Within the lineage, geographic structure is identified. Incongruence across some aspects of the results suggests a complex phylogeographic history for this group, likely shaped by Pleistocene glaciations in northwestern North America. Challenges and future directions are discussed, however the power of the RAD-seq approach seems promising.

Introduction

The amphitropical disjunction is known from organisms across the tree of life including bacteria (Sul et al. 2013), animals (Crame 1993), flowering plants (Popp, Mirré, & Brochmann 2011), lichens (Fernández-Mendoza & Printzen 2013), and bryophytes (Lewis, Rozzi, & Goffinet 2014; Piñeiro et al. 2012; Kreier et al. 2010). This extreme biogeographic pattern has captivated scientists, particularly botanists, throughout the last century (Donoghue 2011; Raven 1963; Du Rietz 1940). Bipolar plant distributions have primarily been attributed to Miocene to Pleistocene long distance dispersal based on phylogenetic analyses of standard genetic markers (see Wen & Ickert-Bond 2009 for review) and molecular divergence dating (Lewis, Rozzi, & Goffinet 2014; Fernández-Mendoza & Printzen 2013; Popp, Mirré, & Brochmann 2011; Gussarova et al. 2008). Among the various groups that exhibit the amphitropical distribution, bryophytes and lichens are the most widely represented (Lewis, Rozzi, & Goffinet 2014; Du Rietz 1940).

Tetraplodon Bruch & Schimp. is the only moss for which phylogenetic analysis and divergence date estimation have been employed to infer the origin and timing of an amphitropical disjunction, with one other study drawing congruent conclusions for the moss *Cinclidium stygium* but from topology alone (Piñeiro et al. 2012). Lewis et al. (2014) estimated that southernmost South American populations, *Tetraplodon fuegianus* Besch., diverged from a northern hemisphere ancestor approximately 8.63 Ma [95% highest posterior density (HPD) 3.07–10.11 Ma]. The monophyly of *T. fuegianus* was supported by three non-homoplasious apomorphic transversions and two indels (single-A deletions) in the *atp*B–*rbc*L intergenic spacer, however three other loci surveyed were unable to support or refute the monophyly of *T. fuegianus*. Based on the conservative

divergence time estimate and the phylogenetic topology recovered from four discrete loci and 128 accessions spanning the geographic and taxonomic range of the genus, Lewis et al. (2014) inferred that the putatively monophyletic *T. fuegianus* is the result of a single direct long distance dispersal event across the tropics. The geographic range of the amphitropical lineage containing *T. fuegianus* was delimited, with a broad Laurasian range and additional disjunctions into the highlands of Papua New Guinea and the Himalayas (Figure 1). The loci used to infer these aspects of the history of *T. fuegianus* and the range of the amphitropical lineage were, however, insufficient for resolving additional structure within the amphitropical lineage. Additional characters were deemed necessary to resolve the phylogeographic history of the lineage.

As next generation sequencing approaches have become reasonably affordable for research on non-model organisms, a number of techniques have been developed that allow for mining of thousands of loci suitable for phylogenetic and population genetic studies (Davey et al. 2011; Holsinger 2010). One of the most promising new techniques is restriction site associated DNA sequencing (RAD-seq; Rowe, Renaut, & Guggisberg 2011), first described by Baird et al. (2008). The approach involves cutting up complete genomic DNA with a restriction enzyme and sequencing the regions flanking the restriction sites. If samples share restriction sites, sequencing and genotyping of thousands of homologous loci across a large numbers of samples is possible. Simulation studies have shown that RAD-seq loci can be used to reliably reconstruct phylogenies in closely related groups (Cariou, Duret, & Charlat 2013; Rubin, Ree, & Moreau 2012). RAD-seq loci have been used to reconstruct phylogenies in a number of difficult-to-resolve groups, with results that correspond to geography (Emerson et al. 2010; Pante et

al. 2014) or morphology (Wagner et al. 2012), and even in groups that are deeply divergent (Hipp et al. 2014). To date there are no published studies applying RAD-seq to the study of bryophytes or toward the understanding of amphitropicality in any organism.

Here we apply a RAD-seq approach toward understanding the relationships and structure within the amphitropical *Tetraplodon* lineage (Lewis, Rozzi, & Goffinet 2014) with an emphasis on identifying the geographic source of the LDD event that gave rise to *T. fuegianus*. With increased locus sampling, the following questions are addressed: (i) Does *T. fuegianus* form a monophyletic lineage, consistent with previously identified support from the *atp*B–*rbc*L intergenic spacer or rather a geographically disjunct population of ubiquitous Northern Hemisphere *T. mnioides*? (ii) what is the geographic origin of the ancestor of *T. fuegianus*? (iii) Is there geographic structure across the Laurasian range of the amphitropical *Tetraplodon* lineage?

Methods

Sampling

Eighty-one samples were collected across the range of the amphitropical *Tetraplodon* lineage inferred by Lewis et al. (2014; Figure 1). Membership in the amphitropical clade was confirmed using an rps4 barcoding approach based on previously described incongruence between morphological species concepts and the molecular phylogenetic results (Lewis et al. 2014). Analysis of rps4 data for all samples was done under maximum likelihood optimality criteria as implemented in Garli v. 2.0 (Zwickl 2006) alongside the genus wide sampling and following the methods of Lewis et al. (2014; Figure S1). DNA was extracted following the anonymous pool sampling approach

discussed by Lewis et al. (chapter 3), where gametophyte stems and sporophytes from a discrete patch are pooled to provide sufficient DNA yields for next generation sequencing techniques. Plant material was ground with liquid nitrogen and DNA was extracted using Nucleospin Plant II Midi kits (Macherey-Nagel, Bethlehem, PA, USA) following the manufacturer's guidelines.

Library preparation

Restriction site associated DNA libraries were prepared according to the protocol described by Etter and Johnson (2012) and Etter et al. (2011) with minor modifications. Wild collected *Tetraplodon fuegianus* was cultured in the lab, and used for flow cytometry genome size estimates with CyStain PI Absolute P DNA Staining Kit for Plant Genome Size (Partec Inc.) following manufactures guidelines. The GC content of the *Tetraplodon* genome is unknown, and thus estimates were taken from the published genome for the model moss species *Physcomitrella patens* (Rensing et al. 2008). Estimated genome size and GC content were used to optimize the choice of restriction enzyme. DNA was digested with SbfI and ligated to (P1) barcoded modified Solexa© adapters (2006 Illumina, Inc.; Etter et al. 2011). Eight bp barcodes were designed with three differences between barcodes. Barcoded samples were pooled and sheared to an average size of 400 bp using an M220 Focused-ultrasonicator™ (Covaris, Inc.) following the manufacture's guidelines. DNA fragments between 400 and 600 bp were selected using Agencourt AMPure XP beads (Beckman Coulter Inc.) at a 0.8:1.0 beads to library ratio, and ligated to a second (P2) modified Solexa© adapter. DNA fragments with both P1 and P2 adapters were PCR amplified using the primers specified by Etter et al. (2011).

Complete P1 barcoded and P2 adapter and PCR oligo sequences are listed in Appendix 2. Libraries were sequenced on an Illumina© MiSeq with 600 cycle v.3 chemistry (2015 Illumina, Inc.).

Bioinformatics processing

Bioinformatics processing was conducted in PyRAD v. 2.17 (Eaton 2014). PyRAD processing is broken into seven steps (Eaton 2014), which are here briefly described along with the parameter settings used. Explanation of particular steps, including clustering, ploidy and heterozygosity estimation, and final dataset generation are discussed in greater detail in the subsequent paragraphs. (1) Reads were de-multiplexed, allowing a maximum of two mismatches in each eight base-pair barcode. (2) Reads were quality filtered, with base calls having phred quality scores of <20 changed to "N" (undetermined) and reads discarded if they contained more than 10 bps with a Phred quality score of <20. Restriction enzyme cut sites and adapter sequences were trimmed from all reads. (3) Reads were clustered within samples. Reads were first de-replicated with number of replicate read occurrences recorded. De-replicated reads were clustered at an 88% similarity threshold within samples (see discussion below). Clusters with a depth greater than the mean depth of all within sample clusters plus two standard deviations of that mean depth, or a depth > 500 were excluded as putative assembled paralogs. (4) Maximum likelihood error rate estimation was done with expected heterozygosity set to zero (haploid; see discussion below) using the maximum likelihood equation described by Lynch (2008). (5) Consensus sequences were generated within each sample for each cluster based on error rate estimations from step four. Only within sample clusters with a

read depth of 5x or greater were kept. This is the minimum depth recommended in the PyRAD full tutorial (http://dereneaton.com/software/pyrad/) for implementation of the statistical base calls method for consensus generation. While greater read depth increases confidence in nucleotide base calls, informativeness of a dataset given sequencing resources is optimized by lowering sequencing depth to as low as 1x and increasing number of samples and loci sequenced (Buerkle & Gompert 2013). Loci were also excluded if they contained more than 5 undetermined bases (N), or contained more than one allele, thus only loci with alleles fixed within samples were retained (see discussion below). (6) Clustering of consensus sequences across samples was done at an 88% similarity threshold. (7) Loci were aligned and datasets were generated (see discussion below; Table 1 & 2).

During data exploration, within and across sample clustering was done at both 85% and 88% similarity. Approximately 150 putative loci were lost with the 3% increase in similarity threshold, suggesting that increasing the similarity threshold did not result in splitting of variable homologous loci (i.e. over splitting), but rather removed loci that may have passed depth requirements through assembly of non-homologous loci. Application of similarity thresholds >88% presented computational challenges, and may not be warranted since oversplitting of loci by excessively stringent similarity thresholds has been shown to be more of problem for RAD-seq datasets of closely related samples than under-splitting (Harvey et al. 2015), especially if false heterozygous loci are filtered out due to ploidy (Ilut, Nydam, & Hare 2014). All analyses are based on 88% similarity clustering, as datasets clustered at different thresholds may not allow for meaningful comparisons (Harvey et al. 2015).

Tetraplodon is haploid, however Lewis et al. (chapter 3) have shown that the anonymous pool sampling approach used for DNA extraction introduces intra-sample variation. The goal was to identify polymorphism between samples, rather than at the individual level within samples. Within sample heterozygous loci were thus excluded. This approach likely results in the exclusion of rare haplotypes, and thus could present issues associated with ascertainment bias if the objective of this study were diversity estimation (Helyar et al. 2011). The pooling and consensus sequence approach has, however, been shown to allow for powerful estimation of broad scale phylogeographic histories in other groups where pooling of individuals was necessary to acquire sufficient DNA yields (Emerson et al. 2010; Gompert et al. 2010). Ultimately, the exclusion of heterozygous loci within samples allows for the identification of alleles fixed within samples, which is suitable for the goals of this study.

Multiple datasets were generated in the final step of the PyRAD pipeline (Eaton 2014). Loci for which at least a minimum number of the total samples had data, i.e. min taxa datasets, were produced for min taxa 60, 50, 40, and 20. For example, a min taxa 20 dataset included all loci for which at least 20 samples had data. Loci were either concatenated to form supermatrices or mined for single nucleotide polymorphisms (SNPs). For min taxa SNP datasets, one SNP was randomly selected from each locus in the associated supermatrix to produce matrices of putatively unlinked SNPs. SNP datasets used in individual assignment analyses were drawn from min taxa 60 and min taxa 20 supermatrices. Datasets were generated for two sample partitions, the first including all samples (i.e. all localities; All samples included, i.e. N=81) and the second for a subset of samples collected across Laurasia (N=65). The Laurasian subset datasets

were produced with higher stringency, including min taxa 60 and min taxa 20 full locus and SNP datasets, out of a total of 65 samples. Min taxa 20 (i.e. the largest) supermatrices for both sample partitions (N=81 & N=65) were blasted against the complete chloroplast and mitochondrial genomes and nuclear ribosomal repeat of *Tetraplodon fuegianus* (Lewis et al. chapter 3).

Maximum likelihood phylogenetic analysis

Maximum likelihood phylogenetic analyses were performed with RAxML v. 8.1.3 (Stamatakis 2006; Stamatakis, Hoover, & Rougemont 2008) for all min taxa supermatrices for both the all localities sampling (i.e. min taxa 60, 50, 40, & 20 for N=81) and the Laurasian subset (i.e. min taxa 60 & 20 for N=65). The GTR-CAT model approximation (Stamatakis 2006) was used to complete single full maximum likelihood tree searches, with 100 bootstrap replicates using the rapid bootstrapping algorithm (Stamatakis, Hoover, & Rougemont 2008).

Species and lineage tree analyses

SVDquartets (Chifman & Kubatko 2014), as implemented in PAUP* v. 4.0a142 (Swofford 2003), was used to infer species and lineage trees for min taxa 60 and 20 supermatrices for both sample partitions (i.e. N=81 & N=65). SVDquartets estimates species level phylogenetic relationships by inferring relationships between quartets of samples under the coalescent model directly from multilocus sequence data using algebraic statistical techniques (Chifman & Kubatko 2014). Misleading phylogenetic results due to incomplete lineage sorting may not be detectable from the analysis of

concatenated supermatrices alone, and thus SVDquartets was employed alongside RAxML analyses for comparison of resolved relationships. This method is relatively fast compared to methods that estimate the posterior distribution of species trees using a Markov Chain Monte Carlo (MCMC) algorithm, and does not rely on summary statistics, thus utilizing the full information present in the data (Chifman & Kubatko 2014). Species trees were inferred for N=81 (all localities) min taxa 60 and 20 supermatrices by evaluating all possible quartets with 100 bootstrap replicates using the multispecies coalescent tree model with samples restrained under two sample different sample partitions, (1) based on collection locality (Figure 1; Appendix 1) and (2) based on RAxML inferred lineages (Figure 2A, C, and D). The main difference between the two partitions is the addition of a distinct lineage including samples from Alaska, Norway, and Labrador, referred to here as the "Laurasia mix" in the RAxML sample partition. Lineage trees were inferred without sample partition information, evaluating all possible quartets under the multispecies coalescent tree model. Bootstrapping was not accomplished during lineage tree estimation due to computational constraints, which could not be overcome under any attempted variation in number of possible quartets evaluated or bootstrap replicates.

Individual assignment

STUCTURE v. 2.3.4 (Pritchard, Stephens, & Donnelly 2000; Falush, Stephens, & Pritchard 2003; Hubisz et al. 2009) was used for individual assignment analyses for min taxa 60 and min taxa 20 unlinked SNP datasets for the all localities sampling (N=81) and the Laurasian subset sampling (N=65). All localities and Laurasian sample sets were both

run under the admixture model with alpha inferred. Sample locations were not used to inform the analyses. Allele frequencies were treated as independent with lambda set to 1.0 for the all localities (N=81) min taxa 20 and min taxa 60 datasets. Allele frequencies were modeled as correlated, with lambda set to 1.0 for Laurasian (N=65) min taxa 60 and min taxa 20 datasets. Preliminary trials suggested that the range of reasonable K values for all datasets was K=2 through K=7 for each sample partition. Each dataset was run with a burnin period of 10,000 MCMC reps, followed by 1,000,000 MCMC reps after burnin for five independent run at each K value. Results from each run were compiled using the CLUMPAK (Cluster Markov Packager Across K) server (Kopelman et al. 2015), which calls on CLUMPP (Jakobsson & Rosenberg 2007) and DISTRUCT (Rosenberg 2004), and calculates optimal K values according to peaks in delta K values as described by Evanno et al. (Evanno, Regnaut, & Goudet 2005) and highest mean Prob(K) value as described by Pritchard et al (Pritchard, Stephens, & Donnelly 2000).

Results

Sequencing & bioinformatics processing

The size of the nuclear genome of *Tetraplodon fuegianus* was estimated to be approximately 642.48 Mbp. Sequencing on the Illumina MiSeq, followed by demultiplexing and quality filtering yielded 20,834,429 total reads, with an average of 257,215 reads per sample. Reads were not evenly distributed across samples, ranging from 47,730 to 665,070 reads per sample. An average of 6,065 loci, ranging from 842 to 16,625 loci, were recovered per sample after filtering for minimum depth (5x) and putative assembled paralogs. After final alignment and filtering, including removal of

loci represented in only one sample (i.e. singletons), an average of 4,137 loci (range 556—8,266) were recovered per sample. A total of 40,174 loci were represented in at least two samples (i.e. min taxa 2) out of 81 total samples, with an average size of 279.5 bp per locus. Composition of min taxa supermatrices used in RAxML and SVDquartets analyses are listed in Table 1, and SNP matrices analyzed in STRUCTURE are described in Table 2. Loci in the all localities supermatrices had an average of 6.5 (6 – 7) parsimony informative sites per locus, 4.33 times more than the Laurasian subset, which had an average of 1.5 (1 – 2) parsimony informative sites per locus. Missing data ranged from 18.46% to 37.08% in the all localities supermatrices, and 6.43% and 35.47% in Laurasian subset supermatrices, with the lowest percentages of missing data in the min taxa 60 supermatrices. The Laurasian subset min taxa 60 dataset, which by definition had the least amount of missing data, was the smallest supermatrix and had the lowest percentage of parsimony informative sites. Although the min taxa 60 supermatrices were the most complete in terms of having the lowest percentages of missing data, simulation studies (Huang & Knowles 2014) and non-simultion studies (Wagner et al. 2012) have shown that larger RAD locus supermatrices with more parsimony informative sites, despite having more missing data, perform better than smaller supermatrices with fewer parsimony informative sites for phylogenetic inference and individual assignment analyses (Chattopadhyay, Garg, & Ramakrishnan 2014). The min taxa 20 all localities supermatrix was the largest supermatrix with the highest percentage and number of parsimony informative sites, but also with the highest percentage of missing data. Blasting loci from min taxa 20 supermatrices for both all localities and Laurasian subset sample partitions to the *T. fuegianus* complete chloroplast and mitochondrial genomes

and nuclear ribosomal repeat showed that only 0.25% of loci sampled here belonged to these regions, with one locus from the chloroplast, three from the mitochondrion, and six from the nuclear ribosomal repeat. Blast searches identified the same loci in both of the min taxa 20 datasets.

Maximum likelihood phylogenetic analysis

Min taxa 60, 50, 40, and 20 supermatrices for the all localities (N=81) sampling were analyzed in RAxML (Figure 2). All supermatrices maximally support Papua New Guinea (PNG; N=2) and Nepal (N=2) samples as composing distinct monophyletic lineages. RAxML trees are rooted with PNG samples based on results from the *rps*4 sample confirmation analysis (Figure S1) and the topology reported by Lewis et al. (2014). Min taxa 60, 40, and 20 supermatrices maximally support the sample from Western Arctic Canada (W. Canada) as sister to all other lineages. All supermatrices maximally support the monophyly of Chilean *Tetraplodon fuegianus* samples (N=11). The sample from Washington State, U.S.A (WA; N=1) was inferred as the sister to Chilean samples with high support (BS 90 or higher) in analyses of min taxa 60 and 20 supermatrices, and with lower support (BS 68) by the min taxa 40 supermatrix. Min taxa 60, 40, and 20 supermatrices reconstruct identical geographically structured lineages for Norway and Sweden, Labrador, and Alaska, and two lineages with a mixed Laurasian membership, collectively referred to here as the "Laurasia mix". All trees infer the same branching pattern, with high support from min taxa 60, 40, and 20 supermatrices, for the geographically structured Norway and Sweden, Labrador, and Alaska lineages, with

Alaska (+ 1 Norway sample) sister to Norway and Sweden, and Labrador clades. The lineages inferred from the min taxa 50 dataset are nearly all poorly supported (i.e. BS < 90; except for Chilean lineage and the AK + Labrador lineage of the Laurasia mix) and vary slightly from those inferred from the other three supermatrices. The estimated min taxa 60 and 40 tree topologies are nearly identical with differences only in branch lengths and BS support values. The overall topology inferred from the min taxa 20 supermatrix differs from the min taxa 60 and 40 topology in the position of the Chile + WA and Laurasia mix lineages. The min taxa 20 topology highly supports (BS 92) Chile + WA as sister to all other Laurasian samples, including both the Laurasia mix and the geographically structured Norway and Sweden, Labrador, and Alaska lineages. The min taxa 60 and 40 topology resolves with lower support (BS 67 and 84 respectively) the Laurasia mix as sister to Chile + WA and the geographically structured Norway and Sweden, Labrador, and Alaska lineages. (i.e. Chile + WA nested between the Laurasia mix and geographically structured Laurasia clades; Figure 2).

Min taxa 60 and 20 Laurasian sample subset supermatrices were also analyzed in RAxML (Figure 3). All lineages, except for the Alaska + Labrador Laurasia Mix clade are poorly supported in the min taxa 60 tree. The min taxa 20 topology is highly supported at all internal nodes, except for that inferring Washington State, U.S.A (WA; N=1) as sister to the Alaska + Labrador Laurasia mix lineage. The min taxa 20 supermatrix provides high support for the relationships recovered by the all localities min taxa 60, 40, and 20 supermatrices (Figure 2) for the geographically structured Laurasian lineages, with Alaska (+ 1 Norway sample) sister to the Norway and Sweden, and Labrador lineages. Both min taxa 60 and min taxa 20 analyses recover the Laurasia mix

clades, with the sole difference being the placement of the WA sample within the min taxa 20 Laurasia mix group, albeit with poor support.

Species and lineage tree analyses

Lineage tree topologies were inferred in SVDquartets for all localities (N=81) min taxa 20 and min taxa 60 supermatrices. The major clades inferred by both lineage trees are congruent, but the topologies differ (Figure 4). Congruent with all RAxML analyses of the N=81 data set, Papua New Guinea and Nepal are sister to all other samples (Figure 4) and Chile is resolved as monophyletic in both lineage trees. A Laurasia mix group is inferred in both topologies, in addition to geographically structured Norway and Sweden, Labrador, and Alaska clades, largely congruent with RAxML topologies. Congruent with only the RAxML min taxa 50 topology, the sample from Western Arctic Canada is inferred within the Laurasia mix clade, rather than sister to Chile and the rest of Laurasia as inferred from RAxML min taxa 60, 40, and 20 analyses. The overall lineage tree topology inferred from the min taxa 20 supermatrix is congruent with the RAxML min taxa 20 topology, with the Laurasia mix clade sister to Chile + WA and the geographically structured Laurasian clades (i.e. Laurasia lineages split by the Chile + WA group; Figure 4), however as pointed out above, the placement of W. Canada is inconsistent. The min taxa 60 topology infers WA as sister to all other lineages, and Chile sister to all Laurasian lineages. This topology is most congruent with the RAxML min taxa 50 topology with the poorly supported internal nodes of that topology collapsed.

Species trees with bootstrap support values were inferred in SVDquartets for the all localities sampling (N=81) under two *a priori* species group partitions, (a) with samples

grouped according to collection locality, and (b) according to the RAxML clades from min taxa 60, 40, and 20 all localities (N=81) topologies. The difference between the two sample partitions was the inclusion of a Laurasia mix group in addition to the geographically structured Laurasian clades in the RAxML ID partition. All species trees for both sample partitions and min taxa supermatrices (60 & 20) resolved Papua New Guinea, Nepal, and W. Canada (respectively) as a grade sister to all other groups (Figure 5). In the collection locality ID partition for min taxa 60 (Figure 5A), WA is sister to all Chile and Laurasia samples with maximal support, and Chile is sister to all Laurasia samples, which form three maximally supported groups, with lower support (BS 64). The min taxa 20 topology resolves WA and Chile as a monophyletic group (BS 78) sister to the Alaska, Labrador, and Norway + Sweden groups. In the RAxML ID species trees the Laurasia mix clades were sister to all others with maximal support, Washington and Chile are monophyletic but with different support values, BS 52 for min taxa 60, and BS 87 for min taxa 20. The three geographically structured Laurasian clades were resolved as a monophyletic Alaska and Labrador (min taxa 60 BS 66; min taxa 20 BS 72) sister to Norway + Sweden under the RAxML species ID partition for both min taxa 60 and min taxa 20. This relationship is incongruent with all RAxML trees and SVDquartets species trees under the collection locality ID partition, but congruent with the both SVDquartets lineage trees.

Individual assignment

STRUCTURE analyses were run for both sample subsets (N=81 & N=65) using min taxa 60 and min taxa 20 SNP datasets for K values 2 through 7. Results were largely

consistent across min taxa datasets within each sample subset, and are here discussed jointly except where indicated. The optimal K value for the all localities sample subset (N=81) under the criteria of Evanno et al. (2005) was K=3, and K=5 under the criteria of Pritchard et al. (2000) for both min taxa datasets (Figure 6). Between K=3 and K=5, Nepal and Papua New Guinea (PNG) are moved into separate clusters. Based on RAxML and SVDquartets phylogenetic results, which consistently resolved these two localities as distinct, as well as the clear distinction of Nepal and PNG in K=4 through K=7, K=3 does not accurately describe the data by grouping the two localities. K=5 has a significantly higher Prob(K) than K=4, the clusters inferred between K=4 and K=5 differ little, with a small proportions of the W. Canada sample and two Alaskan samples assigned to the additional cluster. Thus, K=4 is proposed to be the optimal K value for both min taxa datasets. The sample from W. Canada primarily grouped with the rest of the Laurasian samples. The sample from WA primarily clusters with the rest of the Laurasian samples, but also with the Chile cluster, with an average of 0.259 percent identity with the Chile cluster in the min taxa 60 data set, and an average of 0.370 cluster identity with Chile in the min taxa 20 dataset. For K=7 for min taxa 60 and K=5 for min taxa 20, the Chile cluster shares some cluster identity with the Laurasian samples (Figure 6). Analysis of the min taxa 20 dataset at K=6 and K=7, two modes were recovered. The alternative modes (Figure 6G) resolve only three distinct clusters, PNG, Nepal, and Laurasia, with all additional clusters are composed of only small proportions of the W. Canada sample and two Alaskan samples.

The optimal K value for the Laurasian subset sampling (N=65) min taxa 60 dataset according to the Evanno et al. (2005) criteria was K=4, and according to the Pritchard et

al. (2000) criteria was K=6 (Figure 7). An additional K value of eight was included for the min taxa 60 dataset. Between K=4 and K=6, Norway and Sweden are resolved as a distinct cluster, suggesting that K=4 is too stringent. Between K=5 and K=6, Labrador begins to cluster into a new, yet mixed group, which is developed most clearly at K=7. At K=7, however some clusters are composed only of small proportions of samples. The addition of new groups (i.e K=8) does not contribute to overall structure between the existing clusters. K=5 appears to fit the data best, as the clusters are most clearly resolved, i.e. Norway & Sweden, Alaska plus Labrador, Washington (WA), and two Laurasia mix groups. For the min taxa 20 dataset the Evanno et al. (2005) criteria suggests K=3, and for the Pritchard et al. (2000) criteria K=4 (Figure 7). Between K=3 and K=4 Norway and Sweden emerge as a distinct cluster with affinities to Labrador. At K=5 (Figure 7F.), the sample from WA forms a distinct cluster, with affinities to Labrador and Alaska and the "Laurasia mix". Three out of five independent runs at K=5 inferred and alternative mode with Norway & Sweden, Alaska, and Labrador as belonging to the same cluster (Figure 7G.). K=5 best fits the min taxa 20 dataset, as it allowed for clear resolution of geographically structured clusters. STRUCTURE analyses of both datasets cluster the "Laurasia mix" samples into two groups, consistent with the RAxML phylogenetic reconstructions of this clade. The sample from WA is largely distinct from all other clusters under the optimal K value, K=5. (Figure 7).

Discussion

Summary of topologies

Three primary topologies emerge from the RAxML and SVDquartets lineage and species tree analyses (Figure 8) consistently resolving populations from Papua New Guinea, Nepal, and W. Canada forming a "basal" grade, with W. Canada sister to a clade comprising the remaining samples. Within the latter, the Chilean populations compose a monophyletic group, which is either sister to WA or a clade with all other populations except WA, which is then sister to this combined clade. The topology most frequently recovered (4xs; Figure 8C) resolves WA as sister to Chile, subtended by the Laurasia Mix, with the geographically structured Alaska, Norway and Sweden, and Labrador samples forming a monophyletic crown group. The SVDquartets min taxa 20 lineage tree (Figure 4B) is congruent with the most widely recovered topology (Figure 8C), except that for W. Canada sample resolved within the Laurasia mix group. The third topology, recovered two times (Figure 8B), also resolved WA as sister to Chile, with all Alaska, Labrador, and Scandinavian taxa forming an apical monophyletic group. The final topology, also recovered twice (Figure 8A), resolves WA and Chile as non-sister lineage, with WA sister to a lineage combining Chile and all Laurasian samples.

Refugia and relict haplotypes

Particular patterns arise from the various trees recovered, supporting phylogeographic hypotheses relevant to the objectives of this study. All analyses support the monophyly of Chilean *Tetraplodon fuegianus*, with *T. fuegianus* being notably distinct from all other samples, congruent with the results of Lewis et al. (2014). Chilean *T. fuegianus* is most frequently and with moderate to high support resolved as sister to WA, or alternatively nested between WA and other samples from western North America, suggesting that the

ancestor to *T. fuegianus* dispersed south from western North America. Dispersal into southernmost South America occurred between the late Miocene to Pliocene, with a mean estimate of 8.63 Ma (95% HPD 3.07–10.11 Ma), an estimate that is likely inflated, pushing estimates further back due to the limitations of molecular divergence date estimates on taxa lacking a fossil record (Lewis, Rozzi, & Goffinet 2014).

Western North America has a complex phylogeographic history due to the region's topology and multiple glacial refugia that persisted throughout the Pleistocene glacial cycles (Shafer et al. 2010). A large body of evidence, which has been extensively reviewed three times over the last two decades supports the presence of multiple Pleistocene refugia, including the Beringian refugium in Alaska and a series of southern refugia south of the Cordilleran and Laurentide ice sheets, (Shafer et al. 2010; Soltis et al. 1997; Brunsfeld et al. 2000). The Olympic Peninsula and nearby archipelagos, as well as the Beringian refugium, harbored plants, mammals and birds. *Tetraplodon* comprises strictly coprophilous species, and diversified from an ancestor that predates the Pleistocene glaciations (Lewis, Rozzi, & Goffinet 2014). *Tetraplodon*'s entomochorous dispersal syndrome allows for efficient local dispersal to new substrates, specifically fecal matter or carrion, via a fly vector. The spores of *Tetraplodon* are green allowing them to rapidly germinate, quickly colonizing new substrates. *Tetraplodon* may not be strictly confined to growth on carrion or dung (Koponen 1990), and may be most commonly associated with these substrates as a result of fly-mediated dispersal (Cameron & Wyatt 1986) or its ability to withstand high nutrient concentrations better than other mosses (Fischer 1936; Koponen 1990). *Tetraplodon*'s coprophilous habit and entomochorous dispersal syndrome may have been an effective competitive strategy in

refugial communities where real-estate was scarce, yet available as a result of the fauna known to have survived in the refugia (Shafer et al. 2010).

Trends in the phylogeographic histories of species that survived in western North American refugia suggest that habitat specialists were most likely to reside in a single refugium, whereas generalists and species with large contemporary ranges and high dispersal ability were likely to have survived in multiple refugia (Shafer et al. 2010). *Tetraplodon* is a specialist in terms of habitat (Koponen 1990), however its preferred habitat was present in refugia through the PNW, and its long distance dispersal abilities have been assumed to be poor; however support for colonization via extreme long distance dispersal suggests that this may not always be the case (Lewis, Rozzi, & Goffinet 2014). The dispersal abilities of *Tetraplodon* may be best reflected by its broad range throughout the boreal and arctic regions, with disjunctions into high elevation tropical and southern hemisphere temperate localities (Lewis, Rozzi, & Goffinet 2014). Dispite habitat specificity, these features match the profile of a species that was capable of persisting in multiple refugia. Furthermore, the idea of what constitutes a refugium for a moss has recently shifted with the recovery of live mosses from the retreating edges of glaciers (La Farge, Williams, & England 2013). The relationship recovered here between Chile and WA suggests that the ancestral haplotype that dispersed into southernmost South America was preserved in part as a refugial relic in WA. Despite STRUCTURE and phylogenetic evidence that the WA sample is not completely isolated from other Laurasian populations, gene flow between WA populations must be low enough to prevent complete homogenization of refugia haplotypes.

Contemporary populations of *Tetraplodon* are increasingly common heading north from WA into the Arctic. Following glacial retreat populations may have migrated north, allowing for mixing of haplotypes previously isolated in multiple refugia. The greatest haplotype mixing would be expected at the leading edge of recolonization, with southern refugia haplotypes remaining relatively isolated at the lower latitude. This may explain why the WA sample retains the putative genetic signature of the ancestral haplotype that dispersed south, colonizing sub-Antarctic Chile. Preservation of Pleistocene refugia haplotypes in Western Arctic Canada may similarly explain the "basal" position of this taxon in the majority of resolved topologies. Future work to test these hypotheses will rely on expanding sampling in putative refugial relict populations in the Pacific Northwest and Western Arctic Canada.

RAD-tag loci have allowed for the resolution of distinct geographic structure in the amphitropical *Tetraplodon* clade. Geographically disjunct populations are distinct from Laurasian populations and show no evidence that gene flow is ongoing among disjunct populations or between the disjunct populations and Laurasia. This supports the hypothesis that disjunct populations of *Tetraplodon* result from rare long distance dispersal events (Lewis, Rozzi, & Goffinet 2014). All analyses except for the SVDquartets species tree analyses with samples partitioned according to collection locality (i.e where Alaska, Scandinavia, and Labrador were restrained to being respectively monophyletic) resolve a "Laurasia mix" group composed primarily of Alaskan samples, as well as samples from Norway and Labrador. The presence of two distinct Laurasian groups, the mix group, and the geographically structured group, is also consistent with previously observed phylogeographic trends for taxa that survived in

multiple refugia in the PNW (Shafer et al. 2010). Alternate species trees resolving with high support both a monophyletic and polyphyletic Laurasia group, as well as inconsistent branching orders among the geographically structured Laurasian groups suggests however that there may be incomplete lineage sorting among the Laurasian samples. This would also be consistent with survival in multiple refugia during the Pleistocene followed by recent reintroduction during recolonization of previously glaciated lands. The difficulty in resolving a tidy Laurasian group reflects the recent and complex history of this group, and suggests that *Tetraplodon* represents a great study system for understanding the postglacial history of a moss.

Data and locus coverage

RAD-sequencing presents a powerful tool for rapid and relatively low cost sequencing and genotyping of homologous loci across many samples (Davey et al. 2010). It is however a new technique with much work to be done in terms in refining technical and analytical methods (Davey et al. 2012). The datasets generated in this study highlight a few of the challenges. Factors associated with library preparation, including GC bias during PCR and potential loss of smaller restriction fragments during shearing and size selection, may be the reason for variation in sample coverage across loci, although the latter concern is expected to be less of an issue for studies using infrequent cutters, such as SbfI, as used here (Davey et al. 2013). Another factor that could result in uneven coverage across loci is differences in sequencing across samples due to variability in input DNAs (Davey et al. 2013). One of the most significant challenges specific to RAD-seq data is the potential for systematic allele dropout due to mutations in restrictions sites

among particular lineages or populations (Gautier et al. 2013; Davey et al. 2013; Arnold et al. 2013). Given the strong differentiation of Papua New Guinea and Nepal from the rest of the clades in all inferred topologies and STRUCTURE results, the potential role of allele dropout should be considered for these divergent groups. If the highly divergent samples have allele dropout issues, then they may have fewer alignable loci, and thus higher percentages of missing data in final datasets. This however is not the case. For the N=81 min taxa 60 dataset the two PNG samples have 18.75% and 17.92% missing data, which is very close to the average 18.46% missing data for the complete dataset (Table 1), showing that these samples do not have extraordinary proportions of missing data that would signal allele dropout. In the N=81 min taxa 20 dataset PNG samples have 33.55% and 31.54% missing data, again very close to the average percent of missing data, 37.07%, in the entire dataset (Table 1). For Nepal, one of the samples has a high percentage of missing data relative to the overall missing data, whereas the other has average percentages of missing data. The sample with high missing data also had 127,351 less quality filtered reads than the overall average. The Nepal sample with average percentage of missing data had 271,443 more quality filtered reads than the overall average. Thus the high missing data associated with one of the Nepal samples is more likely due to poor sequencing of this sample rather than systematic allele dropout. Overall uneven coverage across loci likely stems from factors associated with library preparation and sequencing, rather than to systematic error in the case of *Tetraplodon*. RAD-seq is still a young technique, and there is work to be done to optimize the ways in which the data are generated, processed, and used, however the promise of this technique even as it

is still developing is very high (Davey et al. 2013; Davey et al. 2010; Davey et al. 2011) and has here resulted in the largest population level dataset for any bryophyte to date.

Future data exploration

The drop in BS support values between RAxML and SVD species tree analyses, suggests that there are suites of loci that may support alternate topologies, i.e. "suboptimal" trees that RAxML does not report. When new loci are introduced in increasingly large supermatrices, loci supporting suboptimal trees may be getting included at varying proportions, resulting in incongruence between min taxa datasets. This may be the reason for the drop in support values and different topology inferred via RAxML analysis of the min taxa 50 topology (Figure 2B). However, with the inclusion of additional loci (i.e. min taxa 40, and 20), the loci supporting the "sub-optimal" topology appear to have been swamped out by the loci that favored the "optimal tree" (i.e. the topology reported by RAxML) and BS support values increased with increasing "optimal tree" loci. Future work will aim at identifying "sub-optimal" topologies in combination with species tree analyses in order to better understand the phylogeographic history in this group (Hipp et al. 2014).

Simulation studies by Maddison & Knowles (2006) have shown that even in cases of incomplete lineage sorting, phylogenies may be reliably reconstructed given sufficient locus and taxon sampling, with shallow divergences benefiting most from increased taxon sampling. At the time of this simulation study the number of loci feasibly attainable for a non-model organism were low relative to what is now possible using next generation sequencing methods. Maddison and Knowles (2006) found that with the

highest number of loci considered as reasonably attainable in their study, 54 loci, the benefit of increasing loci was equal to the benefit of increasing sampling. This suggests that with the datasets informed by thousands of loci, in addition to a large taxon sampling, incomplete lineage sorting may not completely hinder accurate phylogenetic inference. Oaks (*Quercus*, Fagaceae; Hipp et al. 2014) and the Lake Victoria cichlids (Wagner et al. 2013), provide two examples supporting the power of extensive genome wide locus sampling in recovering accurate phylogenies in groups with incomplete lineage sorting and hybridization. As analytical techniques suited to RAD-tag loci are increasingly developed, it will be critical to focus on computationally feasible methods for exploring the effects of incomplete lineage sorting and presence of optimal and sub-optimal topologies (Hipp et al. 2014) in order to resolve species level phylogenies and phylogeographic histories, despite complex evolutionary histories.

Conclusion

To date, this is the largest population scale dataset generated for any amphitropical or widespread Laurasian organism. By increasing locus sampling relative to previous work, we provide unambiguous support for the monophyly of *Tetraplodon fuegianus*. We also identify northwestern North America as the source of the ancestor to *T. fuegianus*, and recover geographic structure within the amphitropical lineage and in particular within the Laurasian portion of the lineage's range. The latter two features of the history of this group are likely closely integrated with the complex glacial history of northwest North America. The framework for the phylogeographic history of the amphitropical *Tetraplodon* lineage presented here can be further tested by increasing sampling in

putative Pleistocene refugial relict populations in Washington, USA and Western Arctic Canada. Future analyses will focus further on using the coalescent model for species tree estimation, as well as on the exploration of suboptimal topologies to identity potential suites of loci supporting alternative phylogeographic scenarios (Hipp et al. 2014).

References

Arnold, B., R. B. Corbett-Detig, D. Hartl, and K. Bomblies. 2013. RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Molecular Ecology* 22:3179-3190.

Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, W. A. Cresko, and E. A. Johnson. 2008. Rapid SNP discovery and genetic mapping using mequenced RAD Markers. *PloS One* 3: e3376.

Bescherelle, E. 1885. Mousses Nouvelles de l'Amerique Australe. *Bulletin De La Societe Botanique De France.* 32: (iIv – Ixix).

Brunsfeld, S., J. Sullivan, D. Soltis, and P. Soltis. 2001. Comparative phylogeography of northwestern North America: A synthesis. In (J. Silvertown and J. Antonovics, eds.) Integrating ecological and evolutionary processes in a spatial context. Pp. 319-339. Blackwell Science, Oxford.

Buerkle, A. C., and Z. Gompert. 2013. Population genomics based on low coverage sequencing: how low should we go? *Molecular Ecology* 22: 3028–3035.

Cameron, R. G., and R. Wyatt. 1986. Substrate restriction in entomophilous Splachnaceae: role of spore dispersal. *The Bryologist* 89: 279–284.

Cariou, M., L. Duret, and S. Charlat. 2013. Is RAD-Seq suitable for phylogenetic inference? An *in silico* assessment and optimization. *Ecology and Evolution* 3: 846–852.

Chattopadhyay, B., K. M. Garg, and U. Ramakrishnan. 2014. Effect of diversity and missing data on genetic assignment with RAD-Seq markers. *BMC Research Notes* 7: 841.

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

Crame, J. A. 1993. Bipolar molluscs and their evolutionary implications. *Journal of Biogeography* 20: 145–161.

Davey, J. W., T. Cezard, P. Fuentes-Utrilla, C. Eland, K. Gharbi, and M. L. Blaxter. 2012. Special features of RAD sequencing data: implications for genotyping. *Molecular Ecology* 22: 3151-3164.

Davey, J. W., J. L. Davey, M. L. Blaxter, and M. W. Blaxter. 2010. RAD-Seq: next-generation population genetics. *Briefings in Functional Genomics* 9: 416–423.

Davey, John W, Paul a Hohenlohe, Paul D Etter, Jason Q Boone, Julian M Catchen, and Mark L Blaxter. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* 12: 499–510.

Donoghue, M. J. 2011. Bipolar biogeography. *Proceedings of the National Academy of Sciences of the United States of America* 108: 6341–6342.

Du Rietz, G. E. 1940. Problems of bipolar plant distribution. *Acta Phytogeographica Suecica* 13: 215–282.

Eaton, D. A. R. 2014. PyRAD: Assembly of *de novo* RAD-seq loci for phylogenetic analyses. *Bioinformatics* 30: 1844–1849.

Emerson, K. J., C. R. Merz, J. M. Catchen, P. A. Hohenlohe, W. A. Cresko, W. E. Bradshaw, and C. M. Holzapfel. 2010. Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences of the United States of America* 107: 16196-16200.

Etter, P. D., and E. Johnson. 2012. RAD paired-end sequencing for local *de Novo* assembly and SNP discovery in non-model organisms. In *Data Production and Analysis in Population Genomics: Methods and Protocols, Methods in Molecular Biology, Vol. 888.*, (edited by F. Pompanon and A. Bonin) 888:135–151. New York: Springer Science+Business Media.

Etter, P. D., J. L. Preston, S. Bassham, W. A. Cresko, and E. A. Johnson. 2011. Local *de novo* assembly of RAD paired-end contigs using short sequencing reads. *PloS One* 6: e18561.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A Simulation Study. *Molecular Ecology* 14: 2611-2620.

Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.

Fernández-Mendoza, F., and C. Printzen. 2013. Pleistocene expansion of the bipolar lichen *Cetraria aculeata* into the Southern Hemisphere. *Molecular Ecology* 22: 1961–1983.

Fischer, J. 1936. Zur Verbreitungsökologie von *Splachnum sphaericum* (L. Fil.) Swartz. *Beihefte zum Botanischen Centralblatt Abteilung* 55: 122–144.

Gautier, M., K. Gharbi, T. Cezard, and J. Foucaud. 2013. The effect of RAD allele dropout on the estimation of genetic variation within and between populations. *Molecular Ecology* 22: 3165–3178.

Gompert, Z., M. L. Forister, J. A. Fordyce, C. C. Nice, R. J. Williamson, and A. Buerkle. 2010. Bayesian analysis of molecular variance in pyrosequences quantifies population genetic structure across the genome of *Lycaeides* butterflies. Molecular Ecology 19: 2455-2473.

Gussarova, G., M. Popp, E. Vitek, and C. Brochmann. 2008. Molecular phylogeny and biogeography of the bipolar *Euphrasia* (Orobanchaceae): recent radiations in an old genus. *Molecular Phylogenetics and Evolution* 48: 444–460.

Harvey, M. G., C. Duffie-Judy, G. F. Seeholzer, J. M. Maley, G. R. Graves, and R. T. Brumfield. 2015. Similarity threshholds used in short read assembly reduce the comparability of population histories across species. *PeerJ* PrePrints: 1–34.

Helyar, S. J., J. Hemmer-Hansen, D. Bekkevold, M. I. Taylor, R. Ogden, M. T. Limborg, A. Cariani, G. E. Maes, E. Diopere, G. R. Carvalho, E. E. Nielsen. 2011. Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Molecular Ecology Resources* 11 Suppl 1: 123–36.

Hipp, A. L., D. A. R. Eaton, J. Cavender-Bares, E. Fitzek, R. Nipper, and P. S. Manos. 2014. A framework phylogeny of the American oak clade based on sequenced RAD data. *PLoS ONE* 9: e102272.

Holsinger, K. E. 2010. Next generation population genetics and phylogeography. *Molecular Ecology* 19: 2361–3.

Huang, H., and L. L. Knowles. 2014. Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Systematic Biology* vol? pages? Doi:10.1093/sysbio/syu046.

Hubisz, M. J., D. Falush, M. Stephens, and J. K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* 9: 1322–1332.

Ilut, D. C., M. L. Nydam, and M. P. Hare. 2014. Defining loci in restriction-based reduced representation genomic data from nonmodel species: sources of bias and diagnostics for optimal clustering. BioMed Research International Article ID 67515.

Jakobsson, M., and N. A. Rosenberg. 2007. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801–1806.

Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg, and I. Mayrose. 2015. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* doi: 10.1111/1755-0998.12387.

Koponen, A. 1990. Entomophily in the Splachnaceae. *Botanical Journal of the Linnean Society* 104: 115–127.

Kreier, H., K. Feldberg, F. Mahr, A. Bombosch, A. R. Schmidt, R. Zhu, M. von Konrat, B. Shaw, A. J. Shaw, and J. Heinrichs. 2010. Phylogeny of the leafy liverwort *Ptilidium*: cryptic speciation and shared haplotypes between the Northern and Southern Hemispheres. *Molecular Phylogenetics and Evolution* 57: 1260–1267.

La Farge, C., K. H. Williams, and J. H. England. 2013. Regeneration of Little Ice Age bryophytes emerging from a polar glacier with implications of totipotency in extreme environments. *Proceedings of the National Academy of Sciences of the United States of America* 110: 9839-9844.

Lewis, L. R., R. Rozzi, and B. Goffinet. 2014. Direct long-distance dispersal shapes a new world amphitropical disjunction in the dispersal-limited dung moss *Tetraplodon* (Bryopsida: Splachnaceae). *Journal of Biogeography* 41: 2385–2395.

Lynch, M. 2008. Estimation of nucleotide diversity, disequilibrium coefficients, and mutation rates from high-coverage genome-sequencing projects. *Molecular Biology and Evolution* 25: 2409–2419.

Maddison, W. P., and L. L. Knowles. 2006. Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* 55: 21–30.

Pante, E., J. Abdelkrim, A. Viricel, D. Gey, S. C. France, M. C. Boisselier, and S. Samadi. 2014. Use of RAD Sequencing for delimiting species. *Heredity*: 1–10.

Piñeiro, R., M. Popp, K. Hassel, D. Listl, K. B. Westergaard, K. I. Flatberg, H. K. Stenøien, and C. Brochmann. 2012. Circumarctic dispersal and long-distance colonization of South America: the moss genus *Cinclidium*. *Journal of Biogeography* 39: 2041–2051.

Popp, M., V. Mirré, and C. Brochmann. 2011. A single mid-Pleistocene long-distance dispersal by a bird can explain the extreme bipolar disjunction in crowberries (*Empetrum*). *Proceedings of the National Academy of Sciences of the United States of America* 108: 6520–6525.

Pritchard, J., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

Raven, P. H. 1963. Amphitropical relationships in the floras of North and South America. *The Quarterly Review of Biology* 38: 151–177.

Rensing, S. A., D. Lang, A. D. Zimmer, A. Terry, A. Salamov, H. Shapiro, T. Nishiyama, et al. 2008. The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319: 64–9.

Rosenberg, N. A. 2004. DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes* 4: 137–138.

Rowe, H. C., S. Renaut, and A. Guggisberg. 2011. RAD in the realm of next-generation sequencing technologies. *Molecular Ecology* 20: 3499–502.

Rubin, B. E. R., R. H. Ree, and C. S. Moreau. 2012. Inferring phylogenies from RAD sequence data. *PloS One* 7: e33394.

Shafer, A. B. A., C. I. Cullingham, S. D. Côté, and D. W. Coltman. 2010. Of glaciers and refugia: A decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology* 19: 4589–4621.

Soltis, D. E., M. A. Gitzendanner, D. D. Strenge, and P. S. Soltis. 1997. Chloroplast DNA intraspecific phylogeography of plants from the pacific northwest of North America. *Plant Systematics and Evolution* 206: 353–373.

Stamatakis, A. 2006. Phylogenetic models of rate heterogeneity: A high performance computing perspective. In *Proceedings of the IPDPS2006, Rhodos, Greece.*

Stamatakis, A., P. Hoover, and J. Rougemont. 2008. A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology* 57: 758–771.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–90.

Sul, W. J., T. A. Oliver, H. W. Ducklow, L. A. Amaral-Zettler, and M. L. Sogin. 2013. Marine bacteria exhibit a bipolar distribution. *Proceedings of the National Academy of Sciences of the United States of America* 110: 2342–2347.

Swofford, D. L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sunderland, Massachusetts: Sinauer Associates.

Wagner, C. E., I, Keller, S. Wittwer, O. M. Selz, S. Mwaiko, L. Greuter, A. Sivasundar, and O. Seehausen. 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology* 22: 787–798.

Wen, J., and S. M. Ickert-Bond. 2009. Evolution of the Madrean-Tethyan disjunctions and the North and South American amphitropical disjunctions in plants. *Journal of Systematics and Evolution* 47: 331–348.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. The University of Texas at Austin. https://www.nescent.org/wg_garli/Main_Page.

List of figures, tables, and appendices:

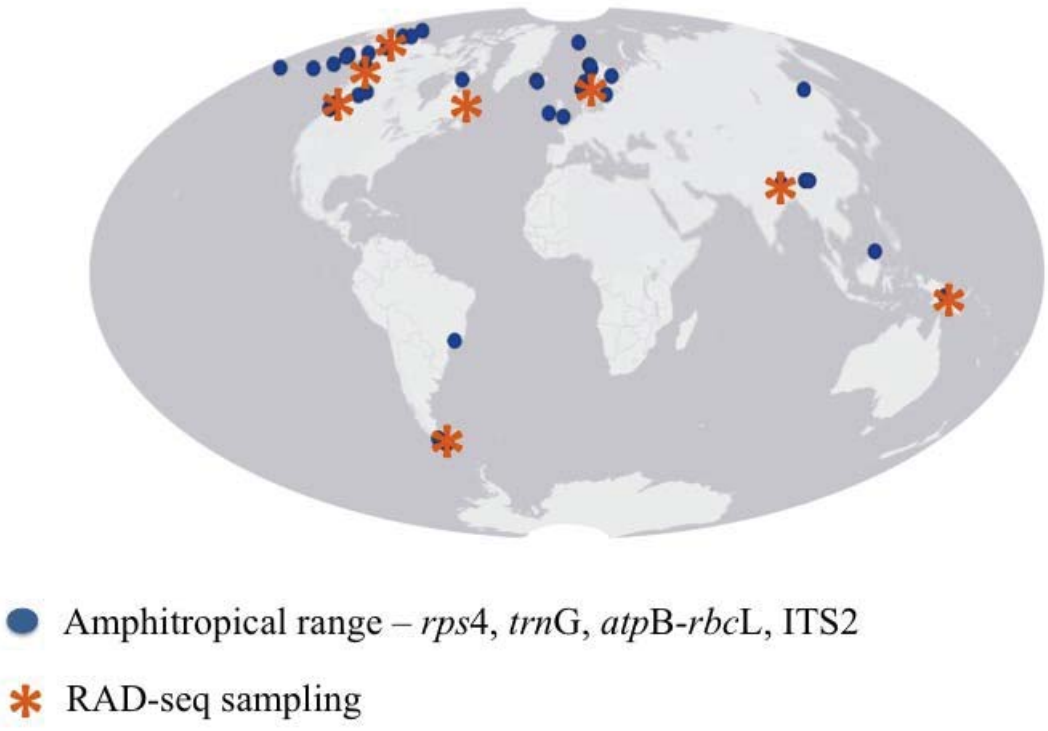
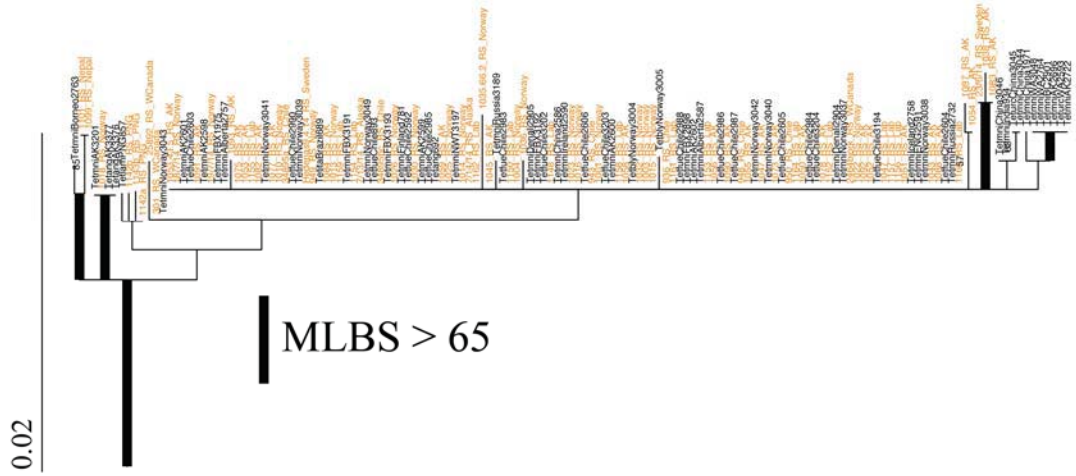


Figure 1. Global range of the *Tetraplodon* amphitropical lineage resolved by Lewis et al. (2014), based on sequencing of four loci (*rps*4, *trn*G, *atp*B-*rbc*L, and ITS2) across the taxonomic and geographic range of the genus (blue dots). Localities sampled for RAD-seq libraries are marked by an orange asterisk, and represent nearly the complete range of the Amphitropical lineage.



SI Figure 1. Portion of the *rps*4 maximum likelihood phylogeny showing the amphitropical clade with sampling from Lewis et al. (2014) in black and samples used for RAD-seq library preparation in orange. Branches supported by a bootstrap value of more than 65 are indicated with bold branches.

Table 1						
All Localities (N = 81)						
Min taxa / locus	**# loci**	**bp**	**# PI**	**% PI**	**Ave PI sites / locus**	**% Missing data**
60	1,407	407,604	8,942	2.194	6	18.460
50	2,377	688,890	15,644	2.271	7	24.228
40	3,148	912,453	20,319	2.227	6	29.224
20	4,077	1,189,336	27,247	2.291	7	37.075
Laurasia (N = 65)						
60	73	21,089	98	0.465	1	6.430
20	3,880	1,126,370	9,437	0.838	2	35.470

Table 1. Composition of min taxa supermatrices used in RAxML and SVDquartets analyses. Abbreviations used include parsimony informative (PI) and base-pair (bp).

Table 2			
Sample subset	**Min taxa / locus**	**# SNPs**	**% Missing data**
N=81 (All Localities)	***60***	1,397	19.731
	20	4,020	37.871
N=65 (Laurasia)	***60***	63	10.452
	20	3,508	36.947

Table 2. Composition of single nucleotide polymorphism (SNP) data sets analyzed in STRUCTURE.

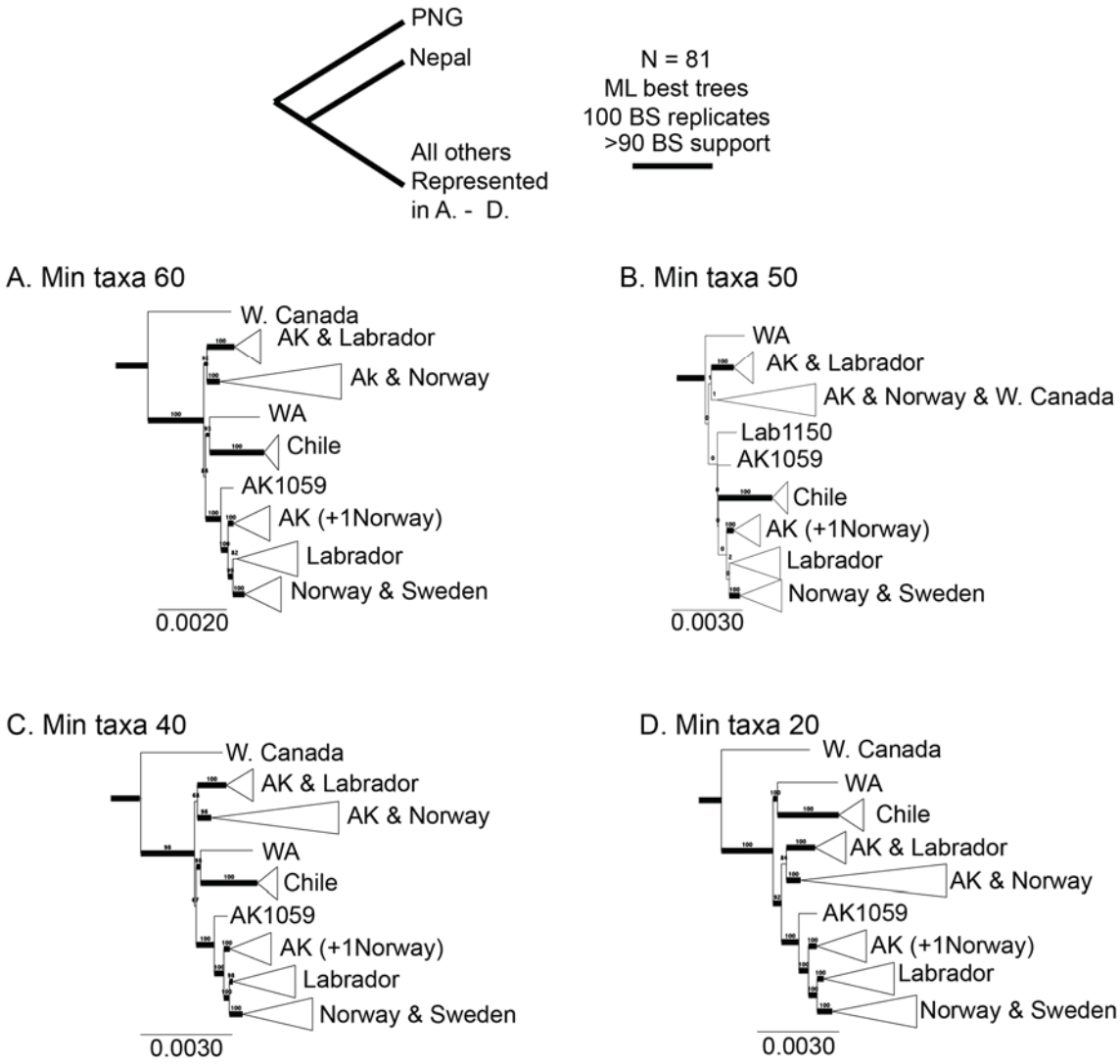


Figure 2. Topologies with maximum likelihood (ML) bootstrap (BS) support values indicated by branch thickness for the all localities (N=81) sampling inferred for min taxa 60, 50, 40, and 20 datasets (A. – D.) in RAxML. Topologies shown in parts A. through D. are all rooted with samples from Papua New Guinea (PNG). Other locality abbreviations used include, Alaska (AK), Washington (WA), and West Canada (W. Canada). "AK & Labrador" and "AK & Norway" clades collectively comprise the "Laurasia mix" lineage.

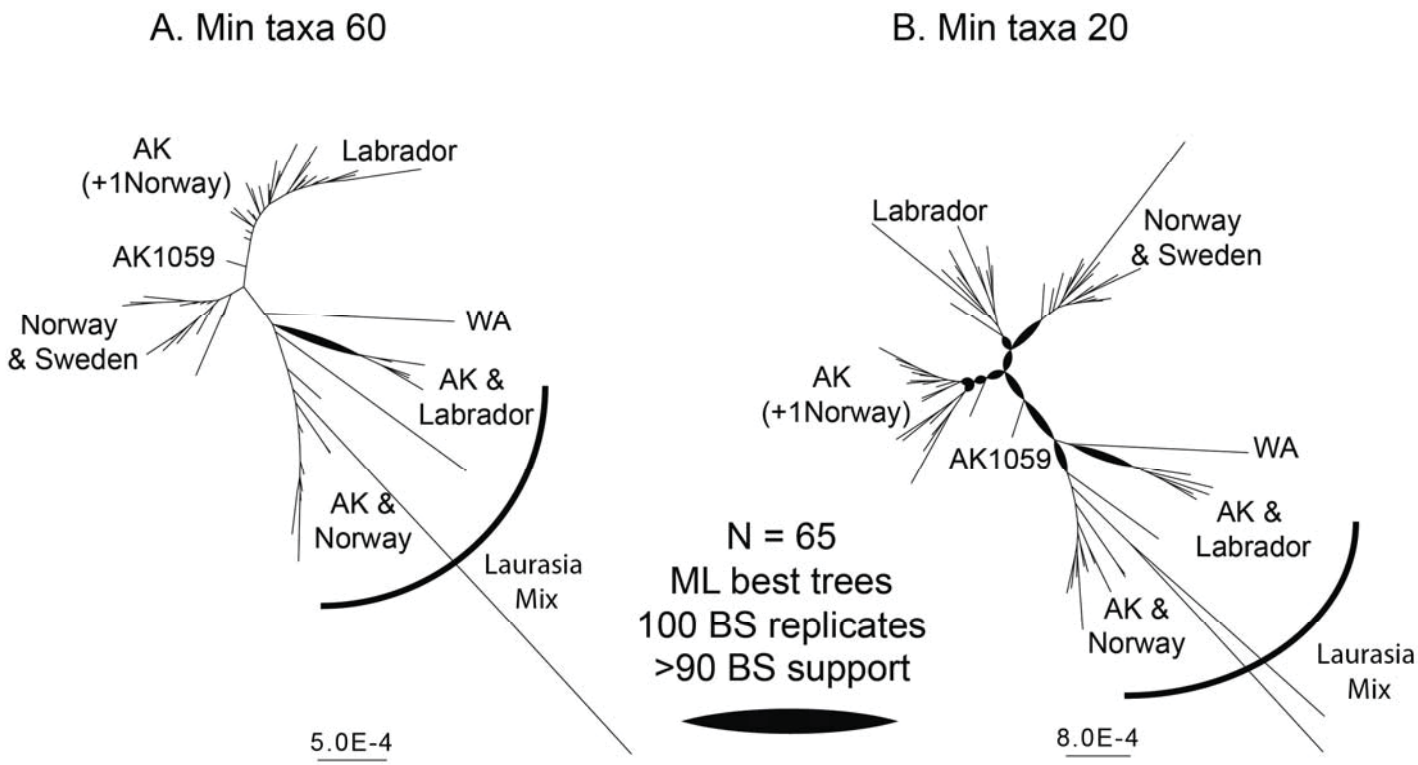


Figure 3. Unrooted topologies with maximum likelihood (ML) bootstrap (BS) support values indicated by branch thickness for the Laurasia subset (N=65) sampling inferred for min taxa 60 (A.) and 20 (B.) datasets in RAxML. Locality abbreviations used include, Alaska (AK) and Washington (WA),. "AK & Labrador" and "AK & Norway" clades collectively comprise the "Laurasia mix" lineage.

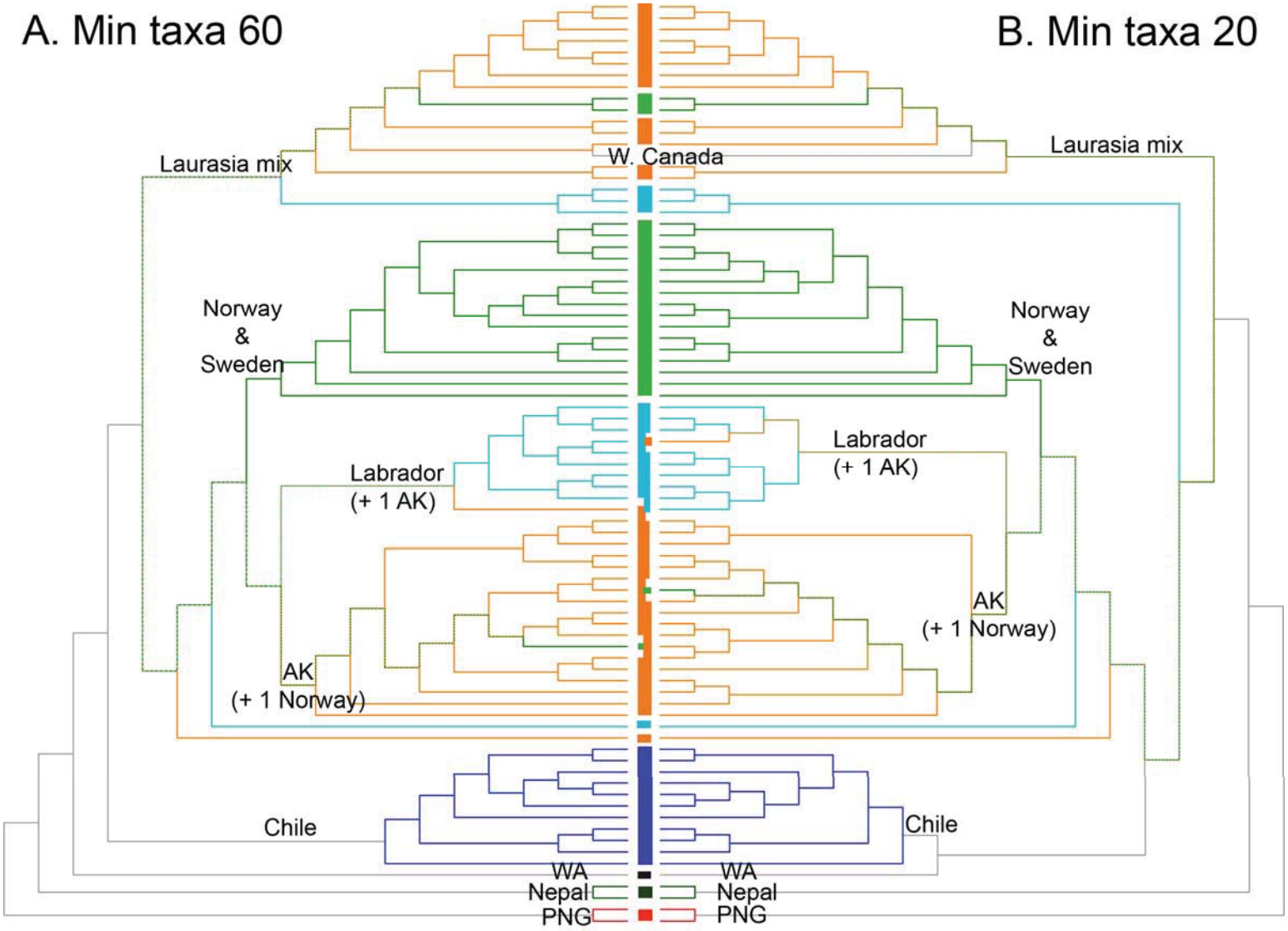

A. Min taxa 60
B. Min taxa 20
Laurasia mix
W. Canada
Laurasia mix
Norway
&
Sweden
Norway
&
Sweden
Labrador
(+ 1 AK)
Labrador
(+ 1 AK)
AK
(+ 1 Norway)
AK
(+ 1 Norway)
Chile
Chile
WA
Nepal
PNG
WA
Nepal
PNG

Figure 4. SVDquartets Lineage tree topologies for the all localities (N=81) sampling inferred for min taxa 60 and 20 datasets (A. & B.). Both trees are rooted with samples from Papua New Guinea (PNG). Other locality abbreviations used include, Alaska (AK), Washington (WA), and West Canada (W. Canada). A. infers Chile as sister to Laurasia, and WA as sister to Chile and Laurasia. B. infers the "Laurasia mix" as sister to Chile & rest of laurasia, with WA as sister to Chile. Colors correspond to collection locality with Alaska - orange, Norway & Sweden - green, Labrador - light blue, Chile - dark blue, Washington – black, Nepal - dark green, and Papua New Guinea – red. Maximum likelihood support values were not inferred due to computational constraints.

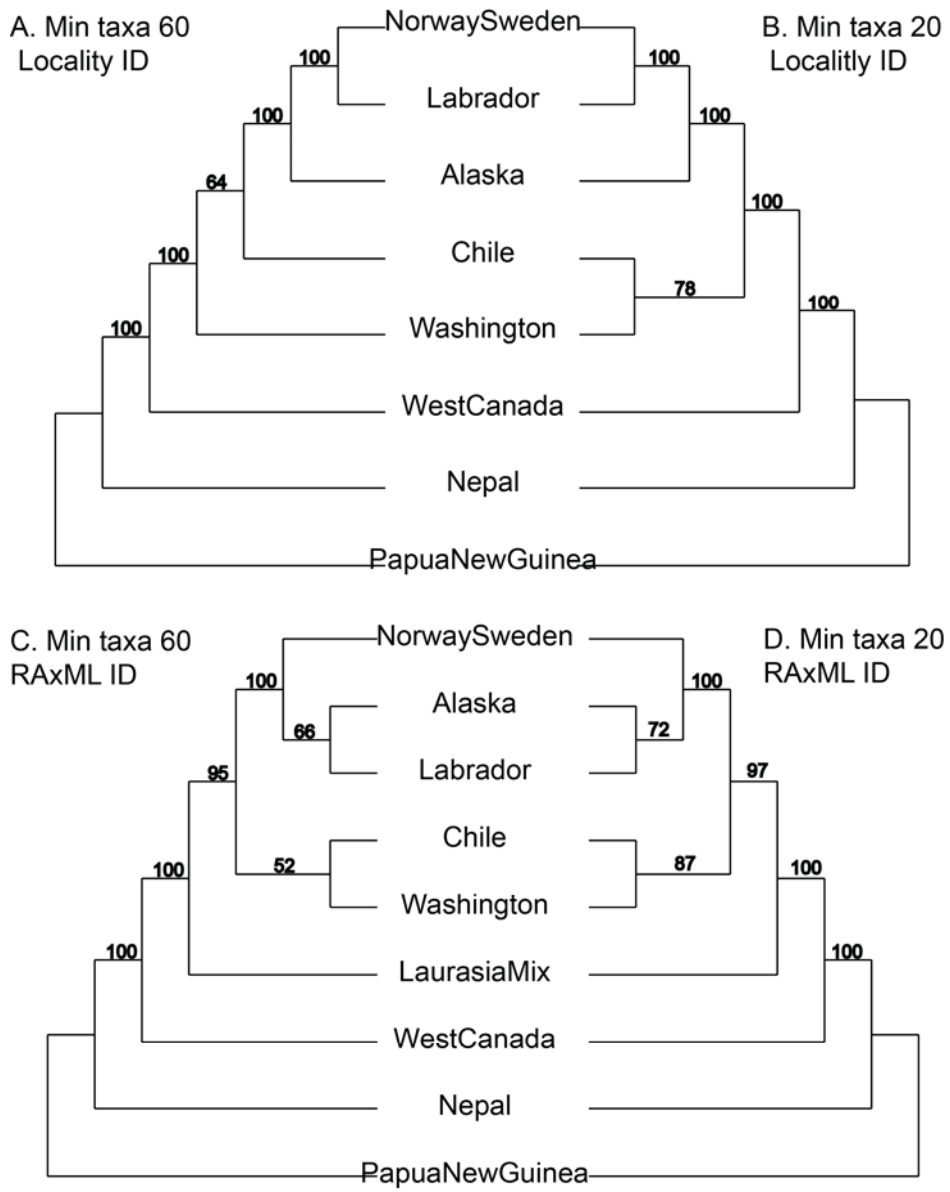


Figure 5: SVDquartets species tree topologies with maximum likelihood (ML) bootstrap (BS) support values for the all localities (N=81) sampling inferred for min taxa 60 (A. & C.) and 20 (B. & D.) datasets with species defined *a priori* under two different sample partitions, based on collection locality (Locality ID: A. & B.) and clades inferred from RAxML analyses (RAxML ID: C. & D.). Trees are rooted with samples from Papua New Guinea.

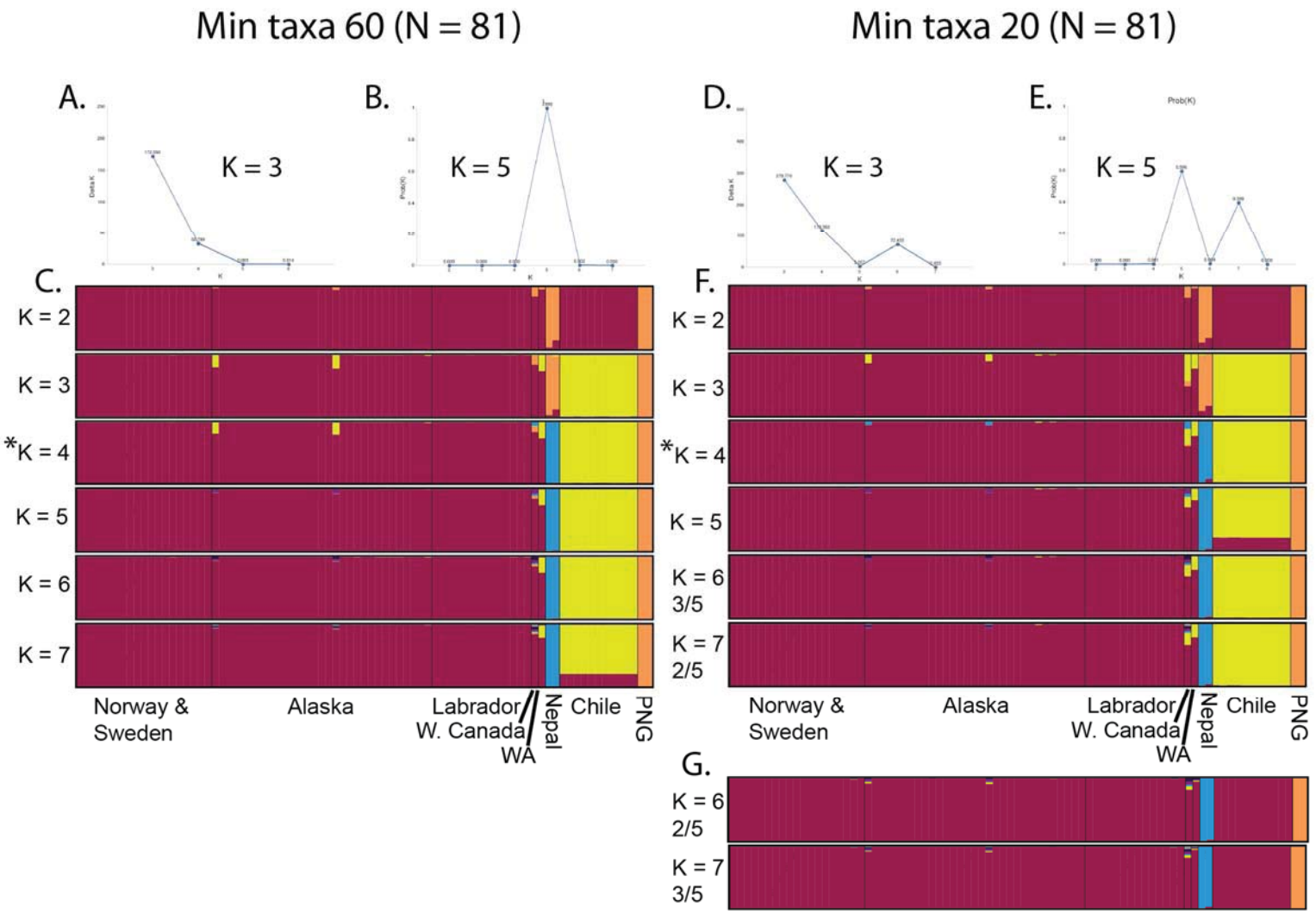

Min taxa 60 (N = 81)
Min taxa 20 (N = 81)
A.
B.
K = 3
K = 5
C.
K = 2
K = 3
*K = 4
K = 5
K = 6
K = 7
Norway & Sweden
Alaska
Labrador W. Canada WA
Nepal
Chile
PNG
D.
E.
K = 3
K = 5
F.
K = 2
K = 3
*K = 4
K = 5
K = 6 3/5
K = 7 2/5
Norway & Sweden
Alaska
Labrador W. Canada WA
Nepal
Chile
PNG
G.
K = 6 2/5
K = 7 3/5

Figure 6: STRUCTURE results for K values 2 through 7 and optimal K value estimates under the criteria of Evanno et al. (2005) and Pritchard et al. (2000) for the all localities sampling (N=81) min taxa 60 (A. – C.) and min taxa 20 (D. – G.) SNP datasets. Optimal K values under each criterion are listed. An asterisk indicates the hypothesized optimal K values for each dataset (C. & F.), K=4, based on the considerations discussed in the text. Briefly, K=5 has a significantly higher Prob(K) than K=4 (B. & E.) for both data sets, however, there was little overall difference in the clusters inferred between K=4 and K=5, i.e. Laurasia, Nepal, Chile, and Papua New Guinea (PNG), with only small proportions of the West Canada (W. Canada) sample and two Alaskan samples assigned to the additional cluster. The sample from Washington is labeled WA. Thus, K=4 is proposed to be the optimal K value. Minor modes were inferred for min taxa 20 K=6 and K=7. Where two modes were inferred, the fraction of runs supporting each mode is indicated below the respective K value (F. & G., K=6 & K=7).

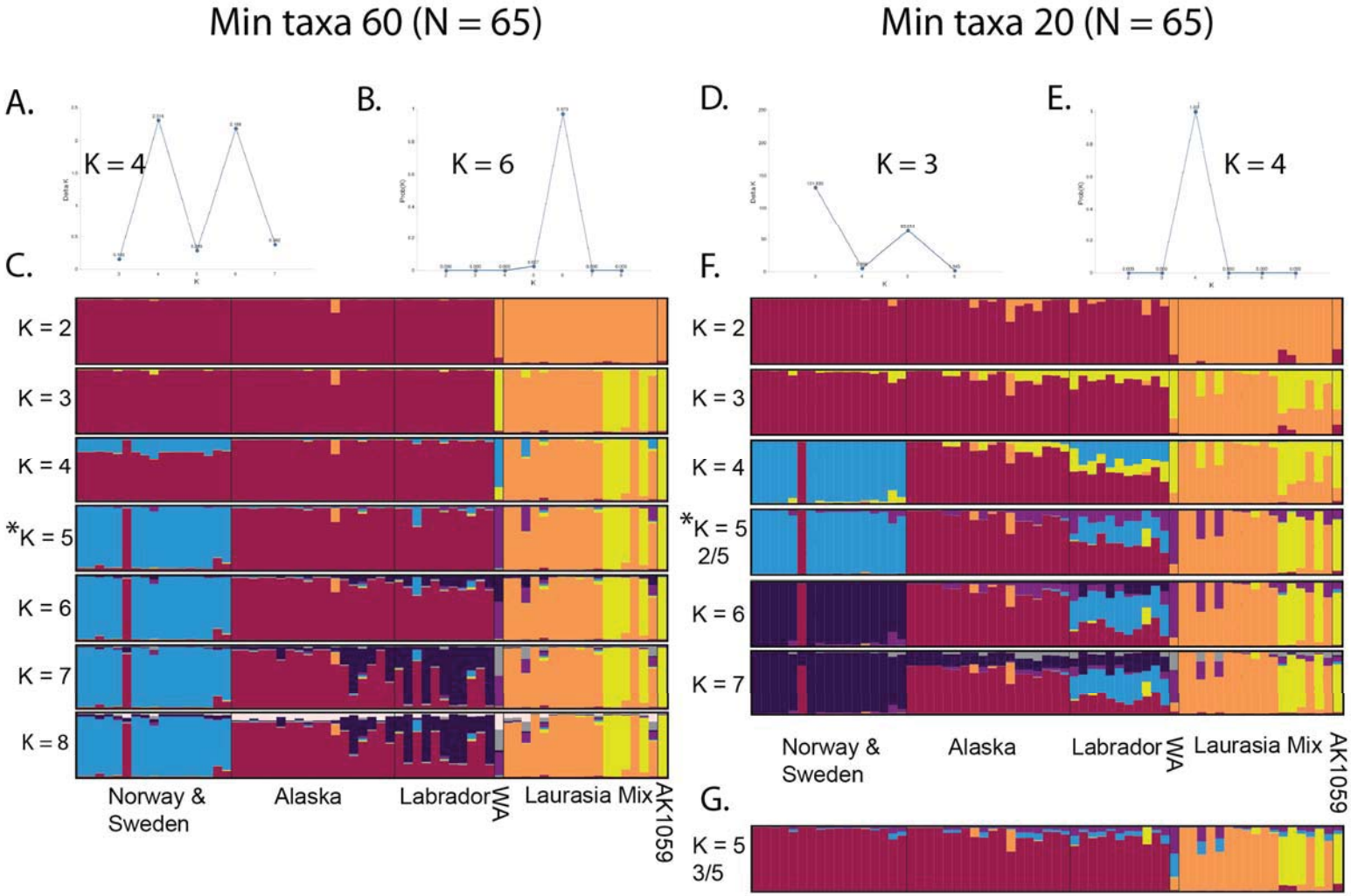
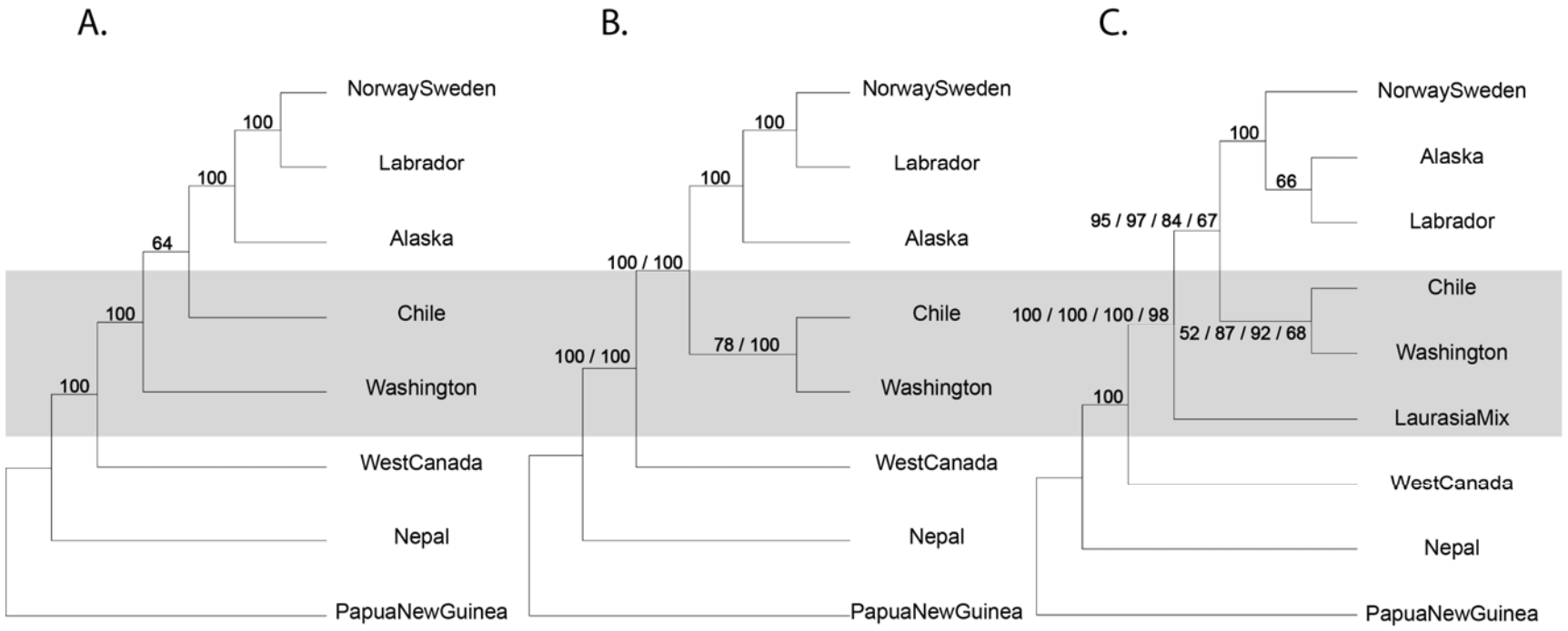

Min taxa 60 (N = 65)
Min taxa 20 (N = 65)
A.
K = 4
B.
K = 6
D.
K = 3
E.
K = 4
C.
F.
K = 2
K = 3
K = 4
*K = 5
K = 6
K = 7
K = 8
*K = 5 2/5
Norway & Sweden
Alaska
Labrador
WA
Laurasia Mix
AK1059
G.
K = 5 3/5

Figure 7: STRUCTURE results for K values 2 through 8 (Min taxa 60; C.) and 2 through 7 (Min taxa 20; F. & G.) and optimal K value estimates under the criteria of Evanno et al. (2005) and Pritchard et al. (2000) for the Laurasian subset sampling (N=65) min taxa 60 (A. – C.) and min taxa 20 (D. – G.) SNP datasets. Optimal K values under each criterion are listed. An asterisk indicates the hypothesized optimal K values for each dataset (C. & F.), K=5, based on the considerations discussed in the text. The sample from Locality abbreviations are used for Alaska (AK) and Washington (WA). Minor modes were inferred for min taxa 20 K=5. Where two modes were inferred, the fraction of runs supporting each mode is indicated below the respective K value (F. & G., K=5).



Supporting topologies:
1) Min taxa 60 Loc ID spp tree (SVD)
2) Min taxa 60 Lineage tree (SVD)
Support Values: 1

Supporting topologies:
1) Min taxa 20 Loc ID spp tree (SVD)
2) Min taxa 20 (RAxML)
Support Values: 1 / 2

Supporting topologies:
1) Min taxa 60 RAx ID spp tree (SVD)
2) Min taxa 20 RAx ID spp tree (SVD)
3) Min taxa 60 (RAxML)
4) Min taxa 40 (RAxML)
5) Min taxa 20 Lineage tree (SVD)*
Support Values: 1 / 2 / 3 / 4
*except for placemment of W. Canada

Figure 8: Summary of the three main topolologies recovered from RAxML and SVDquartets analyses. Across topologies the relationship between Chile and Washington, as well as the Laurasia mix (C.) is shaded grey. Topologies shown are those inferred directly from SVDquartets species tree analyses, with all other congruent topologies listed below. Maximum likelihood (ML) bootstrap (BS) Support values from RAxML and SVDquartets (SVD) analyses are presented at each node in the order listed below each topology. Additional abbreviations used to describe congruent topologies include Locality ID for SVDquartets species trees (Loc ID) and RAxML ID for SVDquartets species trees (RAx ID).

Appendix S1 Specimen data.

Table S1 Accessions used in analyses. All accessions are represented in the all localities sampling (N=81). Accessions included in the Laurasia subset sampling (N=65) are indicated. All accessions are held at the University of Connecticut (CONN). The generic name *Tetraplodon* (*T.*) is abbreviated.

Taxon	Collection #	Collector	Country (& state)	Herbarium	Inclusion in Laurasia subset sampling (N = 65)
T. fuegianus	513	L. R. Lewis	Chile	CONN	N
T. fuegianus	530	L. R. Lewis	Chile	CONN	N
T. fuegianus	936	L. R. Lewis	Chile	CONN	N
T. fuegianus	943	L. R. Lewis	Chile	CONN	N
T. fuegianus	948	L. R. Lewis	Chile	CONN	N
T. fuegianus	949	L. R. Lewis	Chile	CONN	N
T. fuegianus	950	L. R. Lewis	Chile	CONN	N
T. fuegianus	952	L. R. Lewis	Chile	CONN	N
T. fuegianus	998	L. R. Lewis	Chile	CONN	N
T. fuegianus	34895a	J. Larrain	Chile	CONN	N
T. fuegianus	34895bc	J. Larrain	Chile	CONN	N
T. lamii	1142b	L. R. Lewis	Papua New Guinea	CONN	N
T. lamii	1143b	L. R. Lewis	Papua New Guinea	CONN	N
T. mnioides	410	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	500	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	515	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1001	L. R. Lewis	Norway	CONN	Y
T. mnioides	1003	L. R. Lewis	Norway	CONN	Y
T. mnioides	1004	L. R. Lewis	Norway	CONN	Y

T. mnioides	1010	L. R. Lewis	Norway	CONN	Y
T. mnioides	1013	L. R. Lewis	Norway	CONN	Y
T. mnioides	1014	L. R. Lewis	Norway	CONN	Y
T. mnioides	1015	L. R. Lewis	Norway	CONN	Y
T. mnioides	1022	L. R. Lewis	Norway	CONN	Y
T. mnioides	1023	L. R. Lewis	Norway	CONN	Y
T. mnioides	1025	L. R. Lewis	Norway	CONN	Y
T. mnioides	1026	L. R. Lewis	Norway	CONN	Y
T. mnioides	1030	L. R. Lewis	Norway	CONN	Y
T. mnioides	1032	L. R. Lewis	Norway	CONN	Y
T. mnioides	1035	L. R. Lewis	Norway	CONN	Y
T. mnioides	1036	L. R. Lewis	Norway	CONN	Y
T. mnioides	1038	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1043	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1045	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1047	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1050	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1051	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1052	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1053	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1054	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1055	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1056	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1058	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1059	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1073	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1077	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1078	L. R. Lewis	U.S.A., AK	CONN	Y

T. mnioides	1082	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1083	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1084	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1087	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1088	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1090	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1091	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1093	L. R. Lewis	U.S.A., AK	CONN	Y
T. mnioides	1099	L. R. Lewis	Nepal	CONN	N
T. mnioides	1101	L. R. Lewis	Nepal	CONN	N
T. mnioides	1147	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	1148	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	1150	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	1152	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	1153	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	1154	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	1155	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	1156	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	1157	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	1158	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	1159	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	1160	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	1162	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	1164	L. R. Lewis	Canada, Labrador	CONN	Y
T. mnioides	9377	B. Goffinet	U.S.A., AK	CONN	Y
T. mnioides	9383	B. Goffinet	U.S.A., AK	CONN	Y
T. mnioides	9387	B. Goffinet	U.S.A., AK	CONN	Y
T. mnioides	28592	C. Björk	Canada, Nunavut	CONN	N

T. mnioides	7.14.13	H. Golden	U.S.A., AK	CONN	Y
T. mnioides	18.7.14	L. Hedenäs	Norway	CONN	Y
T. mnioides	19.7.14	L. Hedenäs	Norway	CONN	Y
T. mnioides	23.6.14	L. Hedenäs	Sweden	CONN	Y
T. mnioides	4.9.14	L. Hedenäs	Sweden	CONN	Y
T. mnioides	Oly 18	Vedder	U.S.A. WA	CONN	Y

Appendix 2 RAD-seq adapter & primer sequences.

Etter, P. D, Preston, J. L., Bassham, S., Cresko, W. A., and Johnson, E. A.. 2011. Local de Novo Assembly of RAD Paired-End Contigs Using Short Sequencing Reads. PloS One 6: e18561.

Oligo specifications:
[Phos] denotes phosphate group
* denotes phosphorothioate bond
Prepared with IDT TrueGrade Purification
Prepared at ~ 10μM concentration.

P2 Adapter Oligos
PE-P2_Forward
[Phos]GATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCAGAACAA
PE-P2_Reverse
CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCG
ATC*T

PCR Primers
Long-P1-Forward_PCRprimer:
[Phos]AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCG
ATC*T
P2-Reverse_PCRprimer:
5'-CAAGCAGAAGACGGCATACG*A-3'

P1 Adapter Oligos
(P1_[For/Rev] _8bpIndex-Oligo) Complimentary oligos are listed in the same order.
P1_For_AACCAACG-ACACTCTTTCCCTACACGACGCTCTTCCGATCTaaccaacgTGC*A
P1_For_AACCGCAA-ACACTCTTTCCCTACACGACGCTCTTCCGATCTaaccgcaaTGC*A
P1_For_AATATCAT-ACACTCTTTCCCTACACGACGCTCTTCCGATCTaatatcatTGC*A
P1_For_ACCGTCCA-ACACTCTTTCCCTACACGACGCTCTTCCGATCTaccgtccaTGC*A
P1_For_ACCTGCTT-ACACTCTTTCCCTACACGACGCTCTTCCGATCTacctgcttTGC*A
P1_For_ACGCGAAG-ACACTCTTTCCCTACACGACGCTCTTCCGATCTacgcgaagTGC*A
P1_For_AGAGTCTT-ACACTCTTTCCCTACACGACGCTCTTCCGATCTagagtcttTGC*A
P1_For_AGCTTGCG-ACACTCTTTCCCTACACGACGCTCTTCCGATCTagcttgcgTGC*A
P1_For_AGTATGGA-ACACTCTTTCCCTACACGACGCTCTTCCGATCTagtatggaTGC*A
P1_For_ATACGAGC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTatacgagcTGC*A
P1_For_ATGATTAA-ACACTCTTTCCCTACACGACGCTCTTCCGATCTatgattaaTGC*A
P1_For_ATTAGCTA-ACACTCTTTCCCTACACGACGCTCTTCCGATCTattagctaTGC*A
P1_For_ATTCATTG-ACACTCTTTCCCTACACGACGCTCTTCCGATCTattcattgTGC*A
P1_For_CAAGTCAA-ACACTCTTTCCCTACACGACGCTCTTCCGATCTcaagtcaaTGC*A
P1_For_CAATTATC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTcaattatcTGC*A
P1_For_CAGATTCC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTcagattccTGC*A
P1_For_CATTCTAA-ACACTCTTTCCCTACACGACGCTCTTCCGATCTcattctaaTGC*A
P1_For_CCGCCATT-ACACTCTTTCCCTACACGACGCTCTTCCGATCTccgccattTGC*A

P1_For_CCGGTAAC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTccggtaacTGC*A
P1_For_CGCAAGGT-ACACTCTTTCCCTACACGACGCTCTTCCGATCTcgcaaggtTGC*A
P1_For_CGCCGAGG-ACACTCTTTCCCTACACGACGCTCTTCCGATCTcgccgaggTGC*A
P1_For_CGCGATAC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTcgcgatacTGC*A
P1_For_CGTCAGCC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTcgtcagccTGC*A
P1_For_CTCAGGTC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTctcaggtcTGC*A
P1_For_CTTCCAAG-ACACTCTTTCCCTACACGACGCTCTTCCGATCTcttccaagTGC*A
P1_For_GAAGTTGC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTgaagttgcTGC*A
P1_For_GACTGCGC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTgactgcgcTGC*A
P1_For_GATGCCAG-ACACTCTTTCCCTACACGACGCTCTTCCGATCTgatgccagTGC*A
P1_For_GCAGCTTG-ACACTCTTTCCCTACACGACGCTCTTCCGATCTgcagcttgTGC*A
P1_For_GCAGGAAT-ACACTCTTTCCCTACACGACGCTCTTCCGATCTgcaggaatTGC*A
P1_For_GCATATAA-ACACTCTTTCCCTACACGACGCTCTTCCGATCTgcatataaTGC*A
P1_For_GCATGGCG-ACACTCTTTCCCTACACGACGCTCTTCCGATCTgcatggcgTGC*A
P1_For_GCCTCGAC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTgcctcgacTGC*A
P1_For_GCTGCGGT-ACACTCTTTCCCTACACGACGCTCTTCCGATCTgctgcggtTGC*A
P1_For_GGTACTCC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTggtactccTGC*A
P1_For_GGTCAAGT-ACACTCTTTCCCTACACGACGCTCTTCCGATCTggtcaagtTGC*A
P1_For_GGTTCGTA-ACACTCTTTCCCTACACGACGCTCTTCCGATCTggttcgtaTGC*A
P1_For_GTAGACCT-ACACTCTTTCCCTACACGACGCTCTTCCGATCTgtagacctTGC*A
P1_For_GTCAACGG-ACACTCTTTCCCTACACGACGCTCTTCCGATCTgtcaacggTGC*A
P1_For_TAATTCGG-ACACTCTTTCCCTACACGACGCTCTTCCGATCTtaattcggTGC*A
P1_For_TAGTAATT-ACACTCTTTCCCTACACGACGCTCTTCCGATCTtagtaattTGC*A
P1_For_TCGGATGC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTtcggatgcTGC*A
P1_For_TCTTCATC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTtcttcatcTGC*A
P1_For_TGGAATAG-ACACTCTTTCCCTACACGACGCTCTTCCGATCTtggaatagTGC*A
P1_For_TGGAGGCC-ACACTCTTTCCCTACACGACGCTCTTCCGATCTtggaggccTGC*A
P1_For_TTACCGGT-ACACTCTTTCCCTACACGACGCTCTTCCGATCTttaccggtTGC*A
P1_For_TTCTGGCT-ACACTCTTTCCCTACACGACGCTCTTCCGATCTttctggctTGC*A
P1_For_TTGAAGGA-ACACTCTTTCCCTACACGACGCTCTTCCGATCTttgaaggaTGC*A
P1_For_TTGCGTCA-ACACTCTTTCCCTACACGACGCTCTTCCGATCTttgcgtcaTGC*A
P1_For_TTGGCATG-ACACTCTTTCCCTACACGACGCTCTTCCGATCTttggcatgTGC*A
--
P1_Rev_AACCAACG-[Phos]cgttggttAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_AACCGCAA-[Phos]ttgcggttAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_AATATCAT-[Phos]atgatattAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_ACCGTCCA-[Phos]tggacggtAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_ACCTGCTT-[Phos]aagcaggtAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_ACGCGAAG-[Phos]cttcgcgtAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_AGAGTCTT-[Phos]aagactctAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_AGCTTGCG-[Phos]cgcaagctAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_AGTATGGA-[Phos]tccatactAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_ATACGAGC-[Phos]gctcgtatAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_ATGATTAA-[Phos]ttaatcatAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_ATTAGCTA-[Phos]tagctaatAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_ATTCATTG-[Phos]caatgaatAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_CAAGTCAA-[Phos]ttgacttgAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_CAATTATC-[Phos]gataattgAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_CAGATTCC-[Phos]ggaatctgAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_CATTCTAA-[Phos]ttagaatgAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_CCGCCATT-[Phos]aatggcggAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

P1_Rev_CCGGTAAC-[Phos]gttaccggAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_CGCAAGGT-[Phos]accttgcgAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_CGCCGAGG-[Phos]cctcggcgAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_CGCGATAC-[Phos]gtatcgcgAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_CGTCAGCC-[Phos]ggctgacgAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_CTCAGGTC-[Phos]gacctgagAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_CTTCCAAG-[Phos]cttggaagAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GAAGTTGC-[Phos]gcaacttcAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GACTGCGC-[Phos]gcgcagtcAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GATGCCAG-[Phos]ctggcatcAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GCAGCTTG-[Phos]caagctgcAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GCAGGAAT-[Phos]attcctgcAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GCATATAA-[Phos]ttatatgcAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GCATGGCG-[Phos]cgccatgcAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GCCTCGAC-[Phos]gtcgaggcAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GCTGCGGT-[Phos]accgcagcAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GGTACTCC-[Phos]ggagtaccAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GGTCAAGT-[Phos]acttgaccAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GGTTCGTA-[Phos]tacgaaccAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GTAGACCT-[Phos]aggtctacAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_GTCAACGG-[Phos]ccgttgacAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_TAATTCGG-[Phos]ccgaattaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_TAGTAATT-[Phos]aattactaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_TCGGATGC-[Phos]gcatccgaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_TCTTCATC-[Phos]gatgaagaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_TGGAATAG-[Phos]ctattccaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_TGGAGGCC-[Phos]ggcctccaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_TTACCGGT-[Phos]accggtaaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_TTCTGGCT-[Phos]agccagaaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_TTGAAGGA-[Phos]tccttcaaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_TTGCGTCA-[Phos]tgacgcaaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
P1_Rev_TTGGCATG-[Phos]catgccaaAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

Concluding remarks and conservation

As high latitude regions become increasingly threatened by industrial development and climate change, understanding the processes by which the floras of these regions are shaped becomes increasingly pertinent. Patterns of global gene flow and biodiversity will inform our estimates of how regional floras will respond to change. The dissertation research presented here has incorporated newly developed technologies and techniques to make novel contributions to the fields of phylogeography and botany. Bryophyte evolution is dynamic, and a traditionally prevailing paradigm that saw bryophytes as ancient, evolutionary static organisms is rapidly changing. The growing body molecular phylogenetic data is largely responsible for this paradigm shift. This dissertation contributes to this trend, and adds detailed evidence for a mechanism of long distance dispersal that had not been previously inferred for any bryophyte taxon.

An integral part of my dissertation work has been my collaboration with an innovative biocultural conservation program in Southernmost Chile. The Miniature Forests of Cape Horn Program (Goffinet et al. 2012) is a place based educational and conservation program that is successfully promoting local and global awareness of the great diversity of non-vascular plants found in the region. The approach involves changing lenses to better understand high latitude environments, shifting from a charismatic mega-fauna and mega-flora centric concept of biodiversity to place based concepts that emphasize the unique attributes of given regions (Rozzi et al. 2008). In the

Cape Horn Region, > 0.05% of the world's bryophyte species can be found on < 0.01% of the world's land surface (Rozzi et al. 2008). When compared to the five species of trees found in this region, it is clear that there is a wealth of bryophyte diversity in Cape Horn bryophyte flora, which includes *Tetraplodon fuegianus* as one of the species endemic to southern South America.

Southern South America has been an ideal site for investigations of bipolar bryophytes given the conservation dialogue surrounding the unique bryoflora, which has contributed significantly to regional conservation efforts, most namely the designation of the Cape Horn Biosphere Reserve (CHBR), and education and outreach programs (Rozzi et al. 2008; Rozzi et al. 2010; Goffinet et al. 2012; Rozzi et al. 2012). Work contributing to the understanding of regional floristic evolution has a unique venue for communication of research results to a global audience through the Ecotourism with a Hand Lens program (Rozzi et al. 2012; Goffinet et al. 2012). The natural history of dung mosses, their showy sporophytes, and occurrence of endemic species have lent them to public outreach and education. Public education makes it possible to bring bryophytes into the conservation dialogue, and the success of this strategy in the CHBR will serve as a model for conservation programs in other high latitude regions where biodiversity is highest among less widely recognized groups of organisms. In high latitude regions, where floristic biodiversity may appear low when described through vascular plant surveys, recognition of bryophytes, with their far greater abundance and diversity, could shift our perceptions of high latitude biodiversity (Rozzi et al. 2008). Future work will aim to further collaborate for the integration of innovative bryological research into conservation programs.

References:

Goffinet, B., R. Rozzi, L. Lewis, W. Buck, and F. Massardo. 2012. *Miniature Forests of Cape Horn: Ecotourism with a Hand Lens*. University of North Texas Press.

Rozzi, R., and C. B. Anderson, J. C. Pizarro, F. Massardo, Y. Medina, A. O. Mansilla, J. H. Kennedy, J. Ojeda, T. Contador, V. Morales, K. Moses, A. Poole, J. J. Armesto & M. T. Kalin. 2010. Field environmental philosophy and biocultural conservation at the Omora Ethnobotanical Park: methodological approaches to broaden the ways of integrating the social component ("S") in Long-Term Socio-Ecological Research Sites (LTSER). *Revista Chilena de Historia Natural* 83: 27 – 68.

Rozzi, R., J. J. Armesto, B. Goffinet, W. Buck, F, Massardo, J, Silander, M. T. K. Arroyo, S. Russell, C. B. Anderson, L. A. Cavires & J. B. Callicott. 2008. Changing Lenses to Assess Biodiversity: Patterns of Species Richness in Sub-Antarctic Plants and Implications for Global Conservation. *Frontiers in Ecology and the Environment* 6: 131–137.

Rozzi, R., L. Lewis, F. Massardo, Y. Medina, K. Moses, M. Mendez, L. Sancho, P. Vezzani, S. Russell, and B. Goffinet. 2012. *Ecotourism with a Hand Lens at Omora Park*. University of North Texas Press & Ediciones Universidad de Magellanes.