Phylogeography and Phylodemography

INVITED ESSAY

New Frontiers in Bryology and Lichenology

Phylogeography and Phylodemography

A. JONATHAN SHAW AND STUART F. MCDANIEL
   Department of Biology, Duke University, Durham, North Carolina 27708, U.S.A.

OLAF WERNER AND ROSA M. ROS
   Departamento de Biología Vegetal, Facultad de Biología, Universidad de Murcia, Campus de Espinardo, 30100-Murcia, Spain

   Abstract.   *Phylogenetic analyses of infraspecific molecular data in relation to geographic and ecological information has come to be known as phylogeography. Bryophytes offer fertile material for such analyses, which can help clarify long standing biogeographic questions that were intractable before molecular data became available. In particular, molecular data can help distinguish between dispersal and fragmentation explanations for disjunct distributions that characterize many bryophytes at the specific as well as higher levels. Phylodemography is the application of molecular data and phylogenetic analyses to infer past changes in population size within species. Grounded in coalescence theory from population genetics, this new field could be fruitfully applied to bryophytes. Combining phylogeography and phylodemography yields a powerful strategy for elucidating evolutionary processes.*

Phylogenetic methods and perspectives have permeated biology. Taking phylogeny into account is often necessary to achieve statistical rigor in comparative studies because it reveals the nonindependence of events that are linked historically (Felsenstein 1985; Pagel 1997). Thus, today, phylogenetic methods play an important role in diverse fields ranging from human disease epidemiology (Grassly et al. 1999; Holmes et al. 1995; Ou et al. 1992) through functional genomics (Hasebe 1999; Sakakibara et al. 2001), to conservation biology (Faith 1992; Moritz & Faith 1998). The central role of phylogenetics in systematic biology per se is self-evident.

Phylogenetic methods provide powerful tools for studying patterns of biodiversity and biogeography, and for investigating evolutionary and demographic processes within species and populations. Bryophytes pose some exceptionally interesting biogeographic questions (Schofield & Crum 1972) and in the tradition of Darwin himself (1859), analyses of biogeography can yield important inferences about evolutionary processes (e.g., Crum 1972). Two aspects of bryophyte biogeography make them especially interesting subjects for analysis. One is that many bryophytes display the same patterns of disjunction that are characteristic of other plants, suggesting common historical explanations, and the other is that many bryophyte species are very widespread and are themselves disjunct across multiple continents. A fundamental biographic question is whether disjunct distributions so characteristic of bryophytes reflect fragmentation of widespread ancestral ranges (vicariance) or dispersal from one region to another. This review describes recent applications of molecular phylogenetics to bryophyte biogeography, evaluates available genomic regions for future studies, and introduces "phylodemography," the application of molecular phylogenetics to infer past changes in population size.

PHYLOGEOGRAPHY OF BRYOPHYTES: A HISTORICAL OVERVIEW

   Important studies during the 1970's through 1990's, mainly utilizing isozymes, first began to reveal infraspecific genetic structure in mosses and liverworts. These studies have been reviewed by Wyatt et al. (1989a), Stoneburner et al. (1991), and Bishler and Boisselier-Dubayle (1997). Isozyme analyses demonstrated that many or even most bryophytes contain relatively high levels of genetic diversity, higher than some had predicted based on the haploid-dominant life cycle, the propensity for asexual reproduction, and presumed high rates of

inbreeding. Some geographic patterns in genetic diversity, such as higher variation in North Carolina mountain than piedmont populations of *Plagiomnium ciliare* (C. Müll.) T. Kop. (Wyatt et al. 1989*b*), or different levels of genetic diversity among major geographic regions in *Scopelophila cataractae* (Mitt.) Broth. (Shaw 1995), were documented by isozyme studies. Perhaps most interestingly, numerous cases of cryptic speciation in liverworts and mosses have been revealed by isozyme studies (reviewed in Shaw 2001). Isozymes have clearly shown that morphological uniformity across huge geographic areas, in some cases at least, masks complex genetic patterns of "infraspecific" differentiation.

The use of allozymes in population genetic studies have been extended by the use of hypervariable DNA-based markers, including random amplified polymorphic DNA (RAPDs; Williams et al. 1990), amplified fragment length polymorphisms (AFLPs; Vos et al. 1995), and microsatellites (Provan et al. 2001; Van der Velde et al. 2000). Using RAPDs, Selkirk et al. (1997) and Skotnicki et al. (1998*a,b*, 2000) have documented surprisingly high levels of genetic diversity in asexual Antarctic populations of several mosses including *Bryum argenteum* Hedw., *B. pseudotriquetrum* (Hedw.) GMS, *Ceratodon purpureus* (Hedw.) Brid., and *Sarconeuron glaciale* (C. Müll.) Card & Bryhn. New Zealand and Australian plants of *B. argenteum* are more genetically similar to each other than either is to plants from Antarctica. Based on phenetic analyses of RAPD markers, Freitas and Brehm (2001) found that Macronesian populations of *Porella canariensis* (Web.) Und. clustered in two major groups: plants from Madeira were highly distinct from plants that originated in the Azores, Canary Islands, Cape Verde, and Portugal. An analysis of molecular variance (AMOVA) documented significant genetic heterogeneity among archipelagos, and Madeira plants contained a high level of genetic diversity relative to other island groups. While DNA fingerprinting methods are especially well suited for studies of population genetic structure, such markers may be too variable for phylogeographic inference, as marker homology can be difficult to assess, and are less amenable to phylogenetic analyses than are DNA sequences.

Bryophyte phylogenetics has undergone a renaissance in the past decade because of the increasing ease with which DNA sequence data can be acquired and analyzed, and a natural extension of this work is to use sequence data to examine classical biogeographical problems at the species level. The basic evidence needed to distinguish vicariance from dispersal, for example, comes from the topology of phylogenetic trees. Ancient fragmentation is expected to result in reciprocal monophyly between disjunct populations systems. When fragmentation of a continuous range first occurs, no such differentiation will of course exist. Over time, however, as independent neutral mutations accumulate in the isolated lineages and eventually go to fixation through genetic drift, reciprocal monophyly will ultimately occur. The speed at which reciprocal monophyly develops is dependent on a number of factors including population size, generation time, mutation rate, and sex ratios. Dispersal, in contrast, will lead to different phylogenetic patterns. Populations from a particular region may be "nested within" a clade whose basal members occur in the ancestral area. Populations from different, disjunct geographic ranges may exhibit sister group relationships. Tree topology can suggest whether dispersal has occurred multiple times, or just once.

It is important to realize that reciprocal monophyly provides positive evidence of vicariance. The absence of reciprocal monophyly is negative evidence, and each case must be evaluated independently. A lack of reciprocal monophyly between disjunct population systems may reflect dispersal, but it could also reflect recent vicariance such that mutations have not had the time to reach fixation in sister group lineages. The genes being utilized for the study may simply not have the resolving power (due to low levels of variation) to demonstrate reciprocal monophyly. Disjunct populations may share alleles because they retain them from a common ancestor, rather than because of dispersal (gene flow) between them. The problem is analogous to using genetic markers for distinguishing species; reciprocal monophyly is evidence of different species, but shared markers may reflect hybridization or the retention of ancestral polymorphism. Molecular evidence, like any other, must be critically evaluated.

The gene regions that have been used for phylogeographic inference in bryophytes come from the single copy portion of the chloroplast genome, and nuclear ribosomal DNA. The chloroplast DNA region that has been most widely used for resolving patterns at the infraspecific level is the so-called *trnL* region, which includes coding and noncoding tRNA gene sequences. This genomic region actually includes an intron and short exon in the *trnL* (UAA) gene, and an intergenic spacer between the *trnL* and *trnF* (GAA) genes (Fig. 1). The *trnL* region, defined in our lab by the amplified sequence between the *trnC* and *trnF* primers of Taberlet et al. (1991), is usually about 450 bp (base pairs) in length. The chloroplast genome of plants is, in general, conservative at the nucleotide substitution level (Gaut et al. 1993; Palmer 1987), but it appears that the *trnL* region is generally one of the more
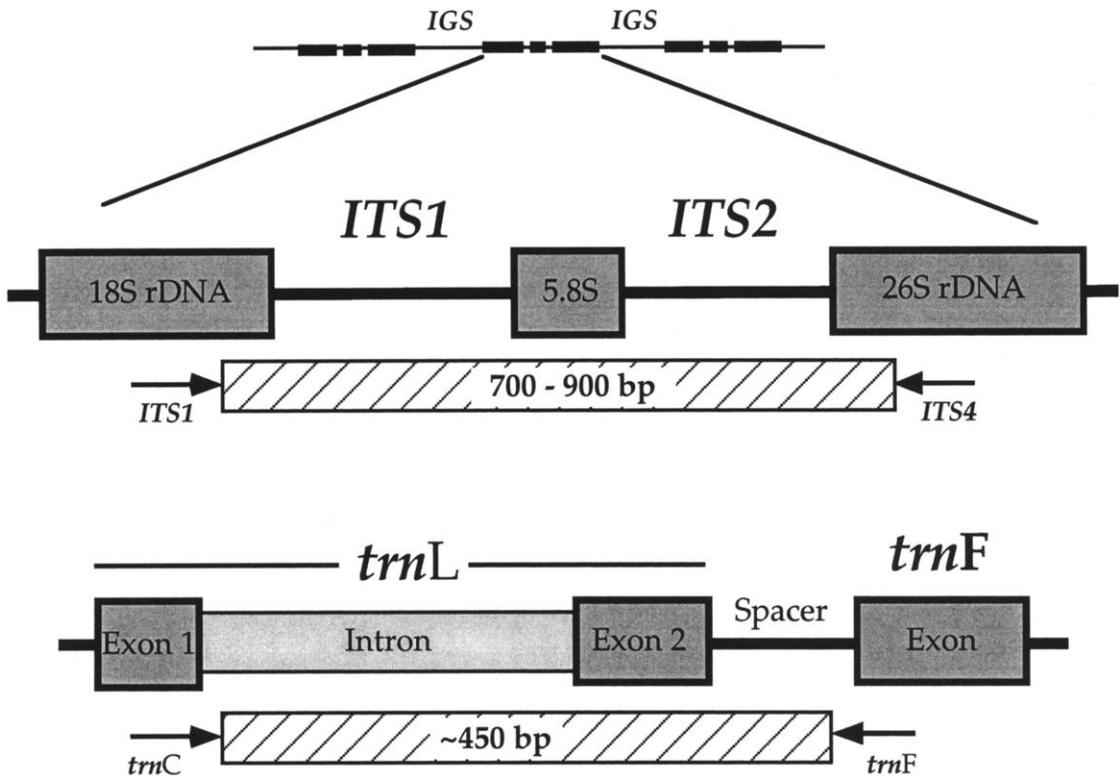
FIGURE 1.   Structure of the internal transcribed spacer (ITS) region of nuclear ribosomal gene arrays (above), and the trnL region of chloroplast DNA (below). The diagrams of ribosomal DNA show the relationship between individual arrays and IGS regions, and the structure of a single array consisting of three genes and spacers. See text for additional information.

variable, and therefore sometimes useful for specific and infraspecific studies.

Several studies have utilized trnL variation to investigate biogeographic patterns in Southern Hemisphere bryophytes. Pfeiffer (2000a) sequenced the *trnL* intron from thirteen plants of *Hymenophyton flabellatum* (Labill.) Trevis and *H. leptopodium* (Hook.f & Taylor) A. Evans from New Zealand and Tasmania and detected a limited number of substitutions and indels within species, although there were up to 11 substitutions and two indels distinguishing the two taxa. Pfeiffer (2000b) resolved what appear to be two taxa of *Hypopterygium* in New Zealand and Australia; a plant from southern Chile had an identical *trnL* intron sequence to one of the two New Zealand taxa. Similarly, Frey et al. (1999) found that *trnL* intron sequences from South American and New Zealand populations of *Lopidium concinnum* (Hook.) Wils. differed in only a single nucleotide substitution and a few indels. Pfeiffer (2000b) and Frey et al. (1999) interpreted the lack of differentiation between Old and New World populations of *Hypopterygium* and *Lopidium* as evidence of very slow evolution in these taxa ("stenospecies") because the authors assumed the

geographic disjunction reflects vicariance associated with the breakup up Gondwana. This would date the disjunction on the order of 60–80 million years old, and would indeed imply a remarkably low substitution rate. An alternative interpretation is that there has been long distance dispersal subsequent to continental drift.

In contrast to the examples described above, more substantial differentiation in *trnL* sequences have been documented for South American versus New Zealand populations of *Monoclea* (Meissner et al. 1998). These correspond to two species that can be distinguished morphologically. McDaniel and Shaw (unpublished) obtained sequences from three chloroplast regions (*atpB-rbcL* spacer, *rps4* gene, *trnL* intron and spacer) of *Pyrrhobryum mnioides* (Hook.) Manuel and found that New and Old World populations formed mutually monophyletic clades (Fig. 2). Furthermore, plants from southern South America were readily distinguishable from northern South American plants. Using the analytical method of nested cladistic analysis (Templeton et al. 1987) and the inference key provided in Templeton (1998), our results suggest a). allopatric fragmentation between New and Old
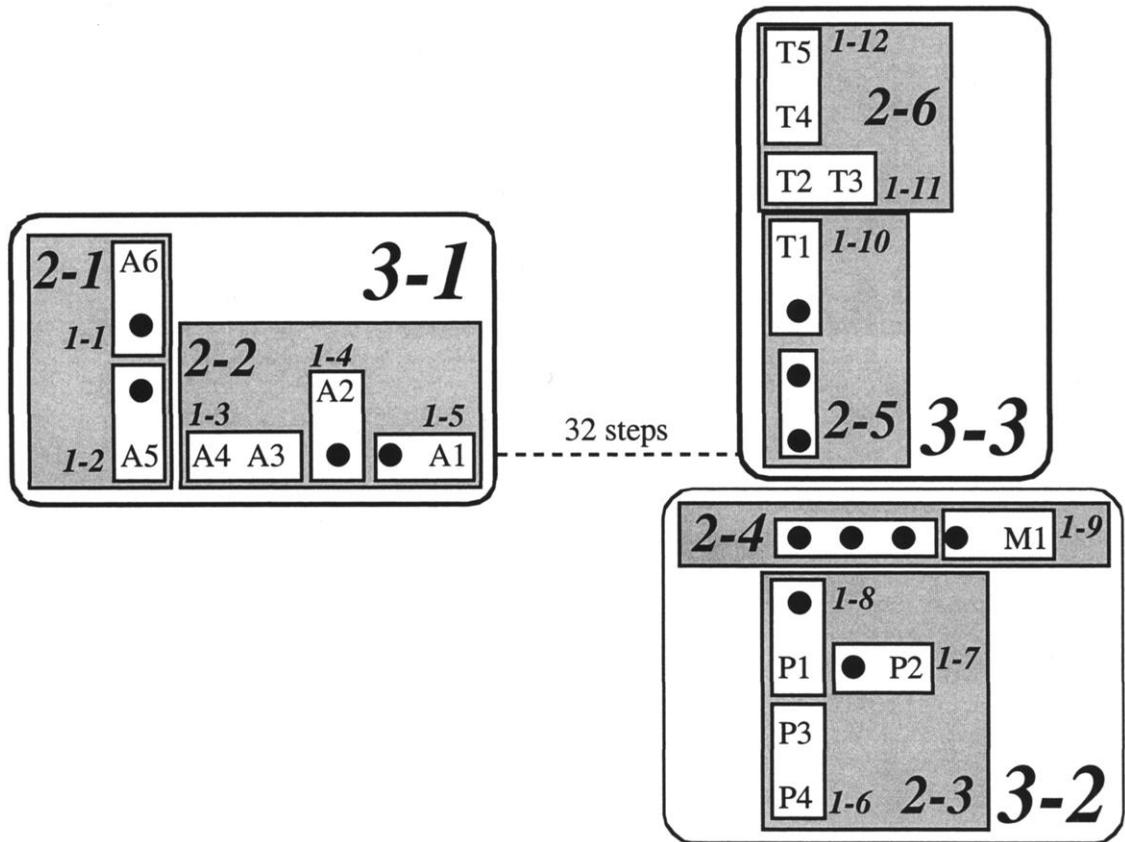
FIGURE 2.   Haplotype network for 31 populations of *Pyrrhobryum mnioides*; such unrooted networks provide an alternative way of presenting genealogical relationships sequences. Populations came from Australasia (A), including Australia, New Zealand, Tasmania and Macquarie Is., Patagonia (P), Magellanes (M), and the montane Neotropics (N), based on sequences of the *trnL, rps4,* and *atpB-rbcL* spacer (chloroplast DNA). Sampled haplotypes are labeled by location and number (e.g., A6 = haplotype 6, from Australasia). Unsampled (therefore hypothetical) haplotypes are represented by a dot. The smallest rectangles enclose pairs of haplotypes that differ by one mutation; 1-step clades (e.g., A4, A3). These 1-step clades are labeled 1–1 through 1–12, following Templeton et al. (1987). The next larger rectangles enclose 1-step clades that are distinguished by two mutations; 2-step clades. These are labeled 2–1 through 2–6. Two step clades are in turn nested within 3-step clades (3–1 through 3–3). This nesting procedure is followed until all clades are nested (three levels in the case of this data set). So-called nested cladistic analysis has been widely used to infer population history from infraspecific molecular data. The nesting procedure allows one to identify more precisely, associations between geography (or phenotypic traits) and genealogy.

World populations, as well as between Patagonian and Neotropical populations, b). isolation by distance between Australia and New Zealand, and c). either isolation by distance or allopatric fragmentation between populations in Magellanes and Patagonia. A date of 14 million years ago (mya) was used to calibrate the disjunction between Patagonian and Neotropical populations (corresponding to the origin of the Atacama Desert). Extrapolating from that date, the disjunction between Patagonian and Australian populations was estimated at 80 mya (42–143 mya., 95% confidence interval), which is consistent with Gonwanan vicariance.

Although still limited, DNA sequence data seem to show a remarkable correspondence to the experimental results of van Zanten (1978). *Pyrrhobryum*

mnioides showed poor germination following van Zanten's desiccation treatments, suggesting limited potential for long distance dispersal. In contrast, *Leptotheca gaudichaudii* Schwaegr. and *Ceratodon purpureus* both exhibited germination after all treatments, and *rps4* and *trnL* sequences from New and Old World populations of both species show no differentiation (McDaniel, unpubl.), consistent with recent dispersal. *Lopidium concinnum*, which also exhibits little differentiation between Old and New World plants, did not germinate under any experimental conditions, so van Zanten's experiments are uninformative. If van Zanten's (1978) experimental results indeed reflect the potential for species to undergo successful long distance dispersal, we can identify several additional taxa from his ta-

bles that look especially promising for studies of Gondwanic vicariance. *Acrocladium auriculatum* (Mont.) Mitt., *Philonotis scabrifolia* (Hook.f. & Wils.) Braithw., *Polytrichadelphus magellanicus* (Hedw.) Mitt., *Ptychomnion aciculare* (Brid.) Mitt., and *Weymouthia mollis* (Hedw.) Broth. all contain disjunctions between Old World-New World populations, but exhibit a patterns in which spore germination dropped off substantially with longer storage or increasingly severe desiccation/freezing treatments. This prediction can be readily tested using nucleotide sequence data.

While chloroplastic regions are variable enough to distinguish ancient vicariance from dispersal, they generally do not contain sufficient variation for infraspecific phylogeographic analyses. The more variable internal transcribed spacer (ITS) region of nuclear ribosomal RNA genes has been a workhorse for molecular studies at the specific and sometimes infraspecific levels (Baldwin 1992; Baldwin et al. 1995). Ribosomal RNA genes occur as arrays of tandem repeats that are dispersed in a variable number of locations across the genome (Fig. 1). Repeat units, each consisting of the 18S ribosomal RNA gene, ITS-1, the 5.8S rRNA gene, ITS-2, and the 26S rRNA gene, are separated by nontranscribed intergenic spacers (IGS). The IGS is said to often be even more variable than ITS-1 and ITS-2. Capesius (1997) reported an IGS sequence from *Funaria hygrometrica* Hedw., but we have been unsuccessful in amplifying the region in *Sphagnum* or other bryophytes at Duke. ITS-1, which is generally 300–600 bp in length (in mosses), is frequently more variable than ITS-2, which is typically 150–300 bp. The 5.8S ribosomal RNA gene, typically amplified along with ITS-1 and 2, is highly conserved, and therefore generally constant (or nearly so) within species. The 5.8S gene is more or less uniformly 158 bp in length.

It appears that in most organisms, concerted evolution results in the homogenization of ITS repeats such that a given plant has many copies of a single ITS haplotype (Baldwin 1992; Baldwin et al. 1995). Nevertheless, concerted evolution may not always completely homogenize ITS haplotypes within a genome, and multiple ITS types may persist within single individuals (Buckler et al. 1997; Buckler & Holtsford 1996; O'Donnell & Cigelnik 1997; Sang et al. 1995). Inadvertently comparing paralogous (rather than homologous) ITS haplotypes across organisms can lead to spurious phylogenetic inferences. Whether this is a problem in a particular study can generally be determined empirically. Multiple persistent ITS haplotypes in a single individual result in "messy" sequence chromatograms in which polymorphism appears to occur at particular nucleotide positions. Cloning and sequencing multiple ITS haplotypes from a given DNA extraction often, in such cases, demonstrates multiple different haplotypes. This is rarely a serious problem in practice; we have found that in most cases among the mosses we have studied using ITS sequences, a single haplotype was generated from each plant. In the relatively few cases where multiple ITS types can be demonstrated within individual plants, the haplotypes differ in just one or a few nucleotides and the variation has little effect on the analysis of biogeographic patterns. Sometimes the existence of multiple ITS haplotypes within an individual can itself provide relevant geographic or evolutionary information if, for example, the two haplotypes are characteristic of two different species (Sang et al. 1995) or otherwise occur in allopatric plants. The former may suggest interspecific hybridization and the latter could reflect previous contact and interbreeding, or the persistence of ancestral polymorphism.

The level of variation observed in ITS sequences varies tremendously among bryophyte genera. Longton and Hedderson (2000) found high levels of variation among populations of *Bryum argenteum* in a worldwide survey of populations. ITS sequences distinguish cryptic lineages within the rare moss, *Mielichhoferia elongata* (Funck) Loeske (Shaw 2000). One lineage occurs in both North America and Europe while the other appears to be limited to western and northern North America. The sister species to *M. elongata*, *M. mielichoferiana* (Hoppe & Hornsch.) Hornsch., is characterized by much lower levels of nucleotide variation, paralleling levels of isozyme variation (Shaw & Schneider 1995). Consistency in this pattern among data sources (various isozymes, ITS sequences) suggest that genome-wide evolutionary processes have resulted in higher genetic diversity in *M. elongata* than in *M. mielichoferiana*.

ITS sequences, along with chloroplast DNA polymorphism, similarly document cryptic speciation in *Fontinalis*. North American and European populations of *F. antipyretica* Hedw. are clearly differentiated in ITS sequences despite an absence of morphological differentiation (Shaw & Allen 2000). Analyses of ITS variation among disjunct populations of *Dicranoweisia cirrata* (Hedw.) Milde, *Scleropodium touretii* (Brid.) L. Koch, and *Claopodium whippleanum* (Sull.) Ren. & Card. in Europe and North America revealed subtle differentiation at the sequence level that did not correspond at all to morphological patterns (Shaw, Werner, & Ros, unpublished). Each of the three species contained higher nucleotide variation in North American than European plants. The disjunct distributions of these species, between southern Europe and western North America, would appear to

be classic examples of the so-called Madrean-Tethyan pattern thought to reflect range fragmentation dating some 25 million or more years into the past (Axelrod 1975). However, considering the minimal degree of sequence differentiation between intercontinental disjuncts, no plausible mutation rate could yield such an old age for the geographic discontinuity and the Madrean-Tethyan hypothesis was rejected in favor of more recent dispersal.

Levels of ITS variation in many acrocarpous genera of mosses are so high that sequences from different congeneric species (e.g., *Pohlia*) cannot be aligned because of length variation. New versus Old World populations of *Pyrrhobryum mnioides* are so divergent they cannot be aligned (McDaniel & Shaw, unpublished). In contrast, within the Amblystegiaceae, and even across other related families of pleurocarpous mosses including the Leskeaceae and Brachytheciaceae, ITS sequences can be readily aligned (Vanderpoorten et al. 2001; in press). The relatively low level of sequence variation in ITS across genera of pleurocarps parallels similarity in other nuclear and chloroplast genes in these taxa, and may reflect a recent radiation of pleurocarps. However, Chiang and Schaal (1999) found sufficient variation in the ITS-2 region of *Hylocomium splendens* (Hedw.) Schimp to resolve geographic patterns in this circumboreal species. The potential utility of the ITS region is an empirical question, and each taxon of interest needs to be investigated for levels of variation.

THE FUTURE: PHYLODEMOGRAPHY AND NEW
GENOMIC REGIONS

Demography is concerned with numbers of individuals in populations and the factors that determine population size. Phylodemography utilizes gene trees to infer past demographic changes. The conceptual framework for making such inferences from phylogenetic trees comes from coalescence theory (Hudson 1990; Kingman 1982). A coalescent tree describes the ancestor-descendent relationships among a set of DNA sequences sampled from a large population. Coalescence models have a long history in population genetics, but have become central to phylogenetic analyses of infraspecific sequence data over the last decade as such data became readily available. Basically, coalescence models turn "traditional" approaches to population genetics on their head by looking backward in time rather than looking forward; e.g., charting expected times to the points at which lineages converge or coalesce to a common ancestor, rather than making predictions about time to fixation of alleles in one lineage or another. Just as the time to fixation of a neutral allele is related to population size, the tim-

ing of coalescence events in a gene genealogy contains information about past population sizes. Coalescence models assume no recombination within the locus being sampled, phylogenetically random sampling of sequences, Wright-Fisher patterns of reproduction (nonoverlapping generations, parentage of offspring random, and independent of previous generations), and that the number of sequences is much smaller than the number of individuals in the population (Pybus et al. 1999).

A number of approaches have been developed for making demographic inferences from DNA sequence data. These approaches were recently reviewed by Emerson et al. (2001), so only a brief overview is provided here, hopefully to stimulate interest and research. The earliest methods for making demographic inferences were based on descriptors of nucleotide variation within a set of aligned sequences from some set of individuals within a species with no reference to the genealogical relationships among those sequences. These include statistics such as the number of segregating sites or the distribution of pairwise differences among sequences, both of which can be used to estimate the parameter theta (the product of genetic effective population size and mutation rate) (Slatkin & Hudson 1991; Watterson 1975). Tajima's D (Tajima 1989) compares two estimates of theta to test whether selection or a previous population bottleneck has influenced the amount of sequence variation at the locus. If a significant value of D is found in a multiple unlinked loci, demographic rather than selective explanations are favored.

Other methods for reconstructing population demographic patterns consider genealogical relationships among DNA sequences, and specifically, *gene tree shape*. Felsenstein (1992) showed that taking phylogeny into account permits more reliable estimates of demographic parameters such as population size and growth rate. Tree shape is a measure of the frequency of coalescence events over time, with time measured in terms of mutations. The only other parameter of relevance in determining tree shape is the effective population size. The neutral coalescent model describes the predicted shape of a phylogenetic tree under neutrality in a randomly mating population. Departures from that shape suggest additional evolutionary factors, and it turns out that changes in population size result in predictable changes in tree shape. In populations that have been stable in size, most coalescence events are near the tips (Fig. 3). In a population that has been growing, coalescence events occur closer to the base of the tree. An exponentially shrinking population would have the coalescence events even more strongly clustered toward the tips relative to stable populations (Fig. 3). If

Exponentially growing

Constant

Exponentially declining

Number of lineages (log transformed)

Time

Time

Epidemic transformation
(m(t) - ln{[n(0)-n(t)]/n(0)n(t)})

Endemic transformation
(T x n(n-1)/2)

exponential decelerating

linear growth

Number of lineages

exponential constant

constant size

exponential accerating

declining

Time

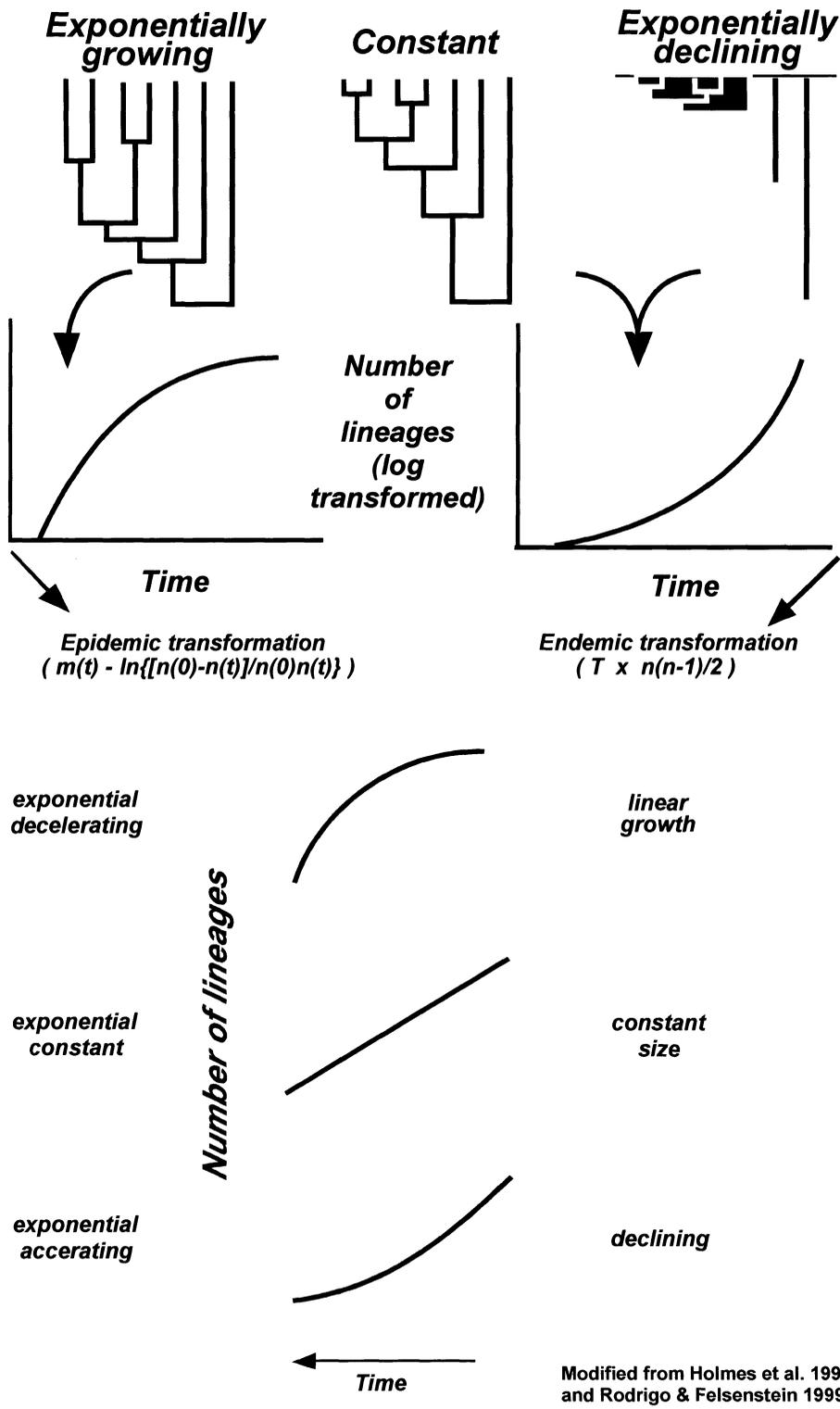Modified from Holmes et al. 1999
and Rodrigo & Felsenstein 1999

FIGURE 3. Lineage through time (LTT) plots under three contrasting demographic scenarios. Such plots are just one way of inferring population demographic history; see text for additional information. Log transformed LTT plots are first constructed. Utilizing either a so-called epidemic transformation, for samples that conform to exponentially growing populations, or an endemic transformation for stable or declining populations, permit further resolution of demographic features (e.g., populations growing at a constant exponential rate, decelerating exponential rate, etc.).

constant population size can be rejected, the population growth rate (r) can be estimated. When N_0 (current population size) and r are both large, tree shape resembles a "star" phylogeny, with almost all coalescence events at or near the base of the tree. Based on similar reasoning, tree shape has been used to infer patterns of diversification from higher level phylogenies of groups of organisms (Nee et al. 1992).

It is one thing to theorize about tree shape under different demographic models, and quite another to assess whether a given tree conforms to one shape rather than another. Two general approaches have been taken. Some assume that the phylogeny is known, and utilize lineages through time (LTT) (Holmes at al. 1995) or skyline plots (Pybus et al. 2000) to describe tree shape and compare that shape to alternative hypothetical shapes reflecting different demographic scenarios. A simple approach to quantifying tree shape is counting the number of coalescence events that occur between the root of the tree and midway to the tips; it turns out that this parameter sufficiently captures tree shape and distinguishes exponentially growing from stable populations (Pybus et al. 1999). Another set of methods do not take a particular phylogenetic tree as correct, but derive demographic parameters by sampling over all possible genealogies. Recent studies take into account factors such as population subdivision before demographic expansions, which would affect the chronology of subsequent coalescence events (Knowles 2001; Wakeley 2000).

Phylodemographic studies are by necessity, sampling-intensive. While phylogenetic studies at the generic, familial, or deeper levels generally utilize a single accession to represent a species (or sometime a genus), phylogeographic and phylodemographic studies obviously require population sampling. Intensive sampling is especially important if demographic inferences are to be made; studies should aim for a minimum of about 30 samples to achieve any degree of statistical rigor (Holmes et al. 1995; Pybus et al. 2000). In addition, phylodemographic analyses require highly variable genomic regions, which is one of the reasons these methods have been applied to the hypervariable genomes of HIV strains (Holmes et al. 1999; Rodrigo & Felsenstein 1999). Studies of ancient human demographic patterns have generally been based on mitochondrial DNA sequences, which are often highly variable (Harpending et al. 1993, 1998; Hawks et al. 2000). Nevertheless, the disadvantage of mitochondrial DNA studies is that genealogical patterns might reflect evolutionary forces (e.g., selection) acting specifically on that genome (which is effectively a single linked gene) rather than more

general demographic properties. Patterns that are reflected in multigenic, multigenomic data, almost certainly reflect processes related to population size rather than to selection on specific loci.

Clearly for these methods to be implemented in bryophytes, other variable gene regions must be explored. The most promising gene regions are intron-containing nuclear gene regions. These regions tend to be variable, and primers can be designed for more conserved flanking exons. Thus far it has been difficult to sequence such regions for exactly the reason that they are desirable: the rapid sequence evolution makes designing primers for more than a particular species or group of species difficult. However two strategies have provided promising results. The increasing availability of genomic sequence data from a number of plant species, including the mosses *Physcomitrella patens* (Hedw.) B.S.G. and *Ceratodon purpureus*, on public databases like GenBank, has made it possible to design primers for additional nuclear genes that might be used for phylogeographic analyses. Sequences from distantly related taxa can be aligned to find conserved regions useful as priming sites.

Individual plants of *Amblystegium* species proved to be highly polymorphic for adensine kinase (ADK) sequences (Vanderpoorten et al., unpubl.), indicating gene duplication and divergence among the paralogs. Indeed, reconciling ADK patterns with a genealogy independently derived from ITS and chloroplast DNA *atpB-rbcL* sequences, suggested numerous duplications of the ancestral ADK locus. The situation in *Amblystegium* is so complex that the locus is clearly not well suited for phylogenetic studies despite high levels of variation. (The evolution of ADK itself warrants further study.) ADK may exist as multiple paralogs in *Ceratodon purpureus* as well (McDaniel, unpublished). Genes for which homology across organisms is ambiguous should be utilized with care. Unpublished work (McDaniel) on phylogeographic patterns in *Ceratodon* suggests that the nuclear genes, DES-6 (Δ^6-fatty acyl acetylenase/desaturase; Sperling et al. 2000) and PHY (phytochrome; Pasentsis et al. 1998) are variable among infraspecific populations.

Another approach to finding novel gene regions is to clone and sequence selected RAPD markers. The following general protocol has yielded three highly variable regions that have proven useful in inter- as well as intraspecific phylogeographic studies in *Sphagnum*. A selection of eight relatively divergent species of *Sphagnum* were screened for RAPD markers using a single primer-pair. Fragments present in all species were selected and each band was cut out of the gel, cloned, and sequenced. (RAPD fragments that are monomorphic across a selection of species in terms of size are not neces-

sarily monomorphic at the nucleotide level.). Specific primers were designed from these cloned sequences and large scale comparative population surveys were conducted. BLAST searches are conducted to identify the anonymous sequences with GenBank accessions; in the case of *Sphagnum*, BLAST searches proved unsuccessful and the three regions remain anonymous for the present. As markers of phylogenetic relationship, this anonymity is not a problem, and all three regions are variable at the specific level.

We believe that these approaches offer exciting opportunities for research on bryophytes. The sorts of changes in population size that are detectable using phylogenetic methods depends on the mutation rate of the gene and generation time of the organism; even relatively variable genes in long-lived organisms may not reflect very recent changes in population size on an ecological time scale, but rather older changes on more of an evolutionary time scale (Lavery et al. 1996). These methods might be especially useful for investigating events and processes such as post-glacial recolonization, colonization of oceanic island systems, or long distance dispersal between continents and across latitudes (e.g., tropical outliers of temperate taxa). Ideally, phylogeographic and phylodemographic analyses can be done in concert to elucidate population histories on both geographic and local scales.

LITERATURE CITED

AXELROD, D. I. 1975. Evolution and biogeography of Madrean-Tethyan sclerophyll vegetation. Annals of the Missouri Botanical Garden 62: 280–334.

BALDWIN, B. G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. Molecular Phylogenetics and Evolution 1: 3–16.

———, M. J. SANDERSON, J. M. PORTER, M. F. WOJCIECHOWSKI, C. S. CAMPBELL & M. J. DONOGHUE. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Annals of the Missouri Botanical Garden 82: 247–277.

BISCHLER, H. & M. C. BOISSELIER-DUBAYLE. 1997. Population genetics and variation in liverworts. Advances in Bryology 6: 1–34.

BUCKLER, E. S. & T. P. HOLTSFORD. 1996. *Zea* ribosomal repeat evolution and substitution patterns. Molecular Biology and Evolution 13: 623–632.

———, A. IPPOLITO & T. P. HOLTSFORD. 1997. The evolution of ribosomal DNA: divergent paralogs and phylogenetic implications. Genetics 145: 821–832.

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

CHIANG, T. Y. & B. A. SCHAAL. 1999. Phylogeography of North American populations of the moss species *Hylocomium splendens* based on the nucleotide sequence of internal transcribed spacer 2 of nuclear ribosomal DNA. Molecular Ecology 8: 1037–1042.

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

DARWIN, C. 1859. On the Origin of Species. Facsimile of the First Edition (1964). Harvard University Press, Cambridge.

EMERSON, B. C., E. PARADIS & C. THÉBAUD. 2001. Revealing the demographic histories of species using DNA sequences. Trends in Ecology and Evolution 16: 707–716.

FAITH, D. P. 1992. Conservation evaluation and phylogenetic diversity. Biological Conservation 61: 1–10.

FELSENSTEIN, J. 1985. Phylogenies and the comparative method. American Naturalist 125: 1–15.

———. 1992. Estimating effective population size from samples of sequences: inefficiency of pairwise and segregating sites as compared to phylogenetic estimates. Genetical Research, Cambridge 59: 139–147.

FREITAS, H. & A. BREHM. 2001. Genetic diversity of the Macronesian leafy liverwort *Porella canariensis* inferred from RAPD markers. Journal of Heredity 92: 339–345.

FREY, W., M. STECH & K. MEISSNER. 1999. Chloroplast DNA-relationship in paleoaustral *Lopidium concinnum* (Hypopterygiaceae, Musci). An example of stenoevolution in mosses. Studies in austral temperate rain forest bryophytes 2. Plant Systematics and Evolution 218: 67–75.

GAUT, B. S., S. V. MUSE & M. T. CLEGG. 1993. Relative rates of nucleotide substitution in the chloroplast genome. Molecular Phylogenetics and Evolution 2: 89–96.

GRASSLY, N. C., P. H. HARVEY & E. C. HOLMES. 1999. Population dynamics of HIV-1 inferred from gene frequencies. Genetics 151: 427–438.

HARPENDING, H. C., S. T. SHERRY, A. R. ROGERS & M. STONEKING. 1993. The genetic structure of ancient human populations. Current Anthropology 34: 483–496.

———, M. A BATZER, M. GURVENS, L. B. JORDE, A. R. ROGERS & S. T. SHERRY. 1998. Genetic traces of ancient demography. Proceedings of the National Academy of Sciences, U.S.A. 95: 1961–1967.

HASEBE, M. 1999. Evolution of the reproductive organs in land plants. Journal of Plant Research 112: 463–474.

HAWKS, J., K. HUNLEY, S. H. LEE & M. WOLPOFF. 2000. Population bottlenecks and Pleistocene human evolution. Molecular Biology and Evolution 17: 2–22.

HOLMES, E. C., S. NEE, A. RAMBAUT, G. GARNETT & P. HARVEY. 1995. Revealing the history of infectious disease epidemics through phylogenetic trees. Philosophical Transactions of the Royal Society of London, Biology 349: 33–40.

———, O. G. PYBUS & P. H. HARVEY. 1999. The molecular population dynamics of HIV-1, pp. 177–207. *In* K. A. Crandall (ed.), The Evolution of HIV. Johns Hopkins University Press, Baltimore and London.

HUDSON, R. R. 1990. Gene genealogies and the coalescent process. Oxford Surveys in Evolutionary Biology 7: 1–44.

KINGMAN, J. F. C. 1982. On the genealogy of large populations. Journal of Applied Probability 19A: 27–43.

KNOWLES, L. L. 2001. Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in

montane grasshoppers. Molecular Ecology 10: 691–701.

LAVERY, S., C. MORITZ & D. R. FIEDLER. 1996. Genetic patterns suggest exponential growth in a declining species. Molecular Biology and Evolution 13: 1106–1113.

LONGTON, R. E. & T. A. HEDDERSON. 2000. What are rare species and why conserve them? Lindbergia 25: 53–61.

MEISSNER, K., J.-P. FRAHM, M. STECH & W. FREY. 1998. Molecular divergence patterns and infraspecific relationship of Monoclea (Monocleales, Hepaticae). Nova Hedwigia 67: 289–302.

MORITZ, C. & D. P. FAITH. 1998. Comparative phylogeography and the identification of genetically divergent areas for conservation. Molecular Ecology 7: 419–429.

NEE, S., A. Ø. MOOERS & P. H. HARVEY. 1992. Tempo and mode of evolution revealed from molecular phylogenies. Proceedings of the National Academy of Sciences, U.S.A. 89: 8322–8326.

O'DONNELL, K. & E. CIGELNIK. 1997. Two divergent intragenomic rDNA types within a monophyletic lineage of the fungus Fusarium are nonorthologous. Molecular Phylogenetics and Evolution 7: 103–116.

OU, C.-Y., C. A. CIESIELSKI G. MEYERS et al. (13 additional authors). 1992. Molecular epidemiology of HIV transmission in a dental practice. Science 256: 1165–1171.

PAGEL, M. 1997. Inferring evolutionary processes from phylogenies. Zoologica Scripta 26: 331–348.

PALMER, J. 1987. Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. American Naturalist 130: S6–S29.

PASENTSIS, K., N. PAULO, P. ALGARRA, P. DITTRICH & F. THÜMMLER. 1998. Characterization and expression of the phytochrome gene family in the moss Ceratodon purpureus. The Plant Journal 13: 51–61.

PFEIFFER, T. 2000a. Molecular relationship of Hymenophyton species (Metzgeriidae, Hepaticophytina) in New Zealand and Tasmania. Studies in austral temperate rain forest bryophytes 5. New Zealand Journal of Botany 38: 415–423.

———. 2000b. Relationships and divergence patterns in Hypopterygium 'rotundum' s. l. (Hypopterygiaceae, Bryopsida) inferred from trnL intron sequences. Studies in austral temperate rain forest bryophytes 7. Edinburgh Journal of Botany 57: 291–307.

PROVAN, J., W. POWELL & P. M. HOLLINGSWORTH. 2001. Chloroplast microsatellites: new tools for studies in plant ecology and evolution. Trends in Ecology and Evolution 16: 142–147.

PYBUS, O. G., E. C. HOLMES & P. H. HARVEY. 1999. The mid-depth method and HIV-1: a practical approach for testing hypotheses of viral epidemic history. Molecular Biology and Evolution 16: 953–959.

———, A. RAMBAUT & P. H. HARVEY. 2000. An integrated framework for the inference of viral population history from reconstructed genealogies. Genetics 155: 1429–1437.

RODRIGO, A. G. & J. FELSENSTEIN. 1999. Coalescent approaches to HIV populations genetics, pp. 233–274. In K. A. Crandall (ed.), The Evolution of HIV. Johns Hopkins University Press, Baltimore and London.

SAKAKIBARA, K., T. NISHIYAMA, M. KATO & M. HASEBE. 2001. Isolation of homeodomain-leucine zipper genes from the moss Physcomitrella patens and the evolution of homeodomain-leucine zipper genes in land plants. Molecular Biology and Evolution 18: 491–502.

SANG, T., D. J. CRAWFORD & T. F. STEUSSY. 1995. Documentation of reticulate evolution in peonies (Paeonia) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. Proceedings of the National Academy of Sciences, U.S.A. 92: 6813–6817.

SCHOFIELD, W. B. & H. A. CRUM. 1972. Disjunctions in bryophytes. Annals of the Missouri Botanical Garden 59: 174–202.

SELKIRK, P. M., M. SKOTNICKI, K. D. ADAM, M. B. CONNETT, T. DALE, T. W. JOE & J. ARMSTRONG. 1997. Genetic variation in Antarctic populations of the moss Sarconeuron. Polar Biology 18: 344–350.

SHAW, A. J. 1995. Genetic biogeography of the rare "copper moss", Scopelophila cataractae. Plant Systematics and Evolution 197: 43–58.

———. 2000. Molecular phylogeography and cryptic speciation in the mosses, Mielichhoferia elongata and M. mielichhoferiana (Bryaceae). Molecular Ecology 9: 595–608.

———. 2001. Biogeographic patterns and cryptic speciation in bryophytes. Journal of Biogeography 28: 253–261.

——— & B. H. ALLEN. 2000. Phylogenetic relationships, morphological incongruence, and geographic speciation in the Fontinalaceae (Bryophyta). Molecular Phylogenetics and Evolution 16: 225–237.

——— & R. E. SCHNEIDER. 1995. Genetic biogeography of the rare "copper moss," Mielichhoferia elongata (Bryaceae). American Journal of Botany 82: 8–17.

SKOTNICKI, M. L., J. A. NINHAM & P. M. SELKIRK. 2000. Genetic diversity, mutagenesis and dispersal of Antarctic mosses: a review of progress with molecular studies. Antarctic Science 12: 363–373.

———, ——— & ———. 1998a. High levels of RAPD diversity in the moss Bryum argenteum in Australia, New Zealand, and Antarctica. Antarctic Science 10: 423–430.

———, P. M. SELKIRK & C. BEARD. 1998b. RAPD profiling of genetic diversity in two populations of the moss Ceratodon purpureus in Victoria Land, Antarctica. Polar Biology 19: 172–176.

SLATKIN, M. & R. R. HUDSON. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129: 555–562.

SPERLING, P., M. LEE, T. GIRKE, U. ZÄHLINGER, S. STYMNE & E. HEINZ. 2000. A bifunctional Δ^6-fatty acyl acetylenase/desaturase from the moss Ceratodon purpureus. European Journal of Biochemistry 267: 3801–3811.

STONEBURNER, A. R. WYATT & I. J. ODRZKOSKI. 1991. Applications of enzyme electrophoresis to bryophyte systematics and population biology. Advances in Bryology 4: 1–27.

TABERLET, P., L. GIELLY, G. PAUTOU & J. BOUVET. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Molecular Biology 17: 1105–1109.

TAJIMA, F. 1989. Statistical properties for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585–595.

TEMPLETON, A. R. 1998. Nested cladistic analyses of phylogeographic data: testing hypotheses about gene flow and population history. Molecular Ecology 7: 381–397.

———, E. BOERWINKLE & C. F. SING. 1987. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Ba-

sic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. Genetics 117: 343–351.

VANDERPOORTEN, A., A. J. SHAW & B. GOFFINET. 2001. Testing controversial alignments in *Amblystegium* and related genera (Amblystegiaceae: Bryopsida). Evidence from rDNA ITS sequences. Systematic Botany 26: 470–479.

———, L. HEDENÄS, C. J. COX & A. J. SHAW. 2002. Phylogeny and morphological evolution in the Amblystegiaceae (Bryopsida). Molecular Phylogenetics and Evolution. 23: 1–21.

VAN DER VELDE, M., H. J. VAN DER STRATE, L. VAN DE ZANDE & R. BIJLSMA. 2000. Isolation and characterization of microsatellites in the moss species *Polytrichum formosum*. Molecular Ecology 9: 1678–1680.

VAN ZANTEN, B. O. 1978. Experimental studies on transoceanic long-range dispersal of moss spores in the Southern Hemisphere. Journal of the Hattori Botanical Laboratory 44: 455–482.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAM DE LEE, M. HORNES, A. FRJTERS, J. POT, J. PELEMAN, M. KUIPER & M. ZABEAU. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407–4414.

WAKELEY, J. 2000. The effects of subdivision on the genetic divergence of populations and species. Evolution 54: 1092–1101.

WATTERSON, G. A. 1975. On the number of segregating sites in genetic models without recombination. Theoretical and Applied Biology 7: 256–276.

WILLIAMS, J. G. K., A. R. KUBELIK, K. J. LIVAK, R. J. ANTONI & S. V. TINGEY. 1990. DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 15: 6531–6535.

WYATT, R., A. STONEBURNER & I. J. ODRZYKOSKI. 1989a. Bryophyte isozymes: systematics and evolutionary implications, pp. 221–240. *In* D. E. Soltis & P. S. Soltis (eds.), Isozymes in Plant Biology. Dioscorides Press, Portland.

———, I. J. ODRZYKOSKI & A. STONEBURNER. 1989b. High levels of genetic variability in the haploid mosses, *Plagiomnium ciliare*. Evolution 43: 1085–1096.