American Journal of Botany

RESEARCH ARTICLE

Peripatric speciation associated with genome expansion and female-biased sex ratios in the moss genus *Ceratodon*

Marta Nieto-Lugilde[1,5], Olaf Werner[1], Stuart F. McDaniel[2], Petr Koutecký[3], Jan Kučera[3], Samah Mohamed Rizk[4], and Rosa M. Ros[1]

Manuscript received 21 November 2017; revision accepted 21 March 2018.

[1] Departamento de Biología Vegetal, Facultad de Biología, Universidad de Murcia, Campus de Espinardo, 30100 Murcia, Spain

[2] Biology Department, University of Florida, Gainesville, Florida 32611, USA

[3] Faculty of Science, University of South Bohemia, Branišovská 1760, CZ-370 05 České Budějovice, Czech Republic

[4] Genetics Department, Faculty of Agriculture, Ain Shams University, 68 Hadayek Shubra, 11241 Cairo, Egypt

[5] Author for correspondence (e-mail: manilu@um.es)

Citation: Nieto-Lugilde, M., O. Werner, S. F. McDaniel, P. Koutecký, J. Kučera, S. M. Rizk, and R. M. Ros. 2018. Peripatric speciation associated with genome expansion and female-biased sex ratios in the moss genus *Ceratodon*. *American Journal of Botany* 105(6): 1–12.

doi:10.1002/ajb2.1107

PREMISE OF THE STUDY: A period of allopatry is widely believed to be essential for the evolution of reproductive isolation. However, strict allopatry may be difficult to achieve in some cosmopolitan, spore-dispersed groups, like mosses. We examined the genetic and genome size diversity in Mediterranean populations of the moss *Ceratodon purpureus* s.l. to evaluate the role of allopatry and ploidy change in population divergence.

METHODS: We sampled populations of the genus *Ceratodon* from mountainous areas and lowlands of the Mediterranean region, and from Western and Central Europe. We performed phylogenetic and coalescent analyses on sequences from five nuclear introns and a chloroplast locus to reconstruct their evolutionary history. We also estimated genome size using flow cytometry (employing propidium iodide) and determined the sex of samples using a sex-linked PCR marker.

KEY RESULTS: Two well-differentiated clades were resolved, discriminating two homogeneous groups: the widespread *C. purpureus* and a local group mostly restricted to the mountains in Southern Spain. The latter also possessed a genome size 25% larger than the widespread *C. purpureus*, and the samples of this group consist entirely of females. We also found hybrids, and some of them had a genome size equivalent to the sum of the *C. purpureus* and Spanish genome, suggesting that they arose by allopolyploidy.

CONCLUSIONS: These data suggest that a new species of *Ceratodon* arose via peripatric speciation, potentially involving a genome size change and a strong female-biased sex ratio. The new species has hybridized in the past with *C. purpureus*.

KEY WORDS allopolyploidy; coalescent analysis; cosmopolitan bryophytes; flow cytometry; hybridization; nuclear and chloroplast DNA sequencing; phylogenetic data; recombinants; sex determination; Spanish Sierra Nevada.

The origin of new species represents a major unsolved problem in evolutionary biology (Rieseberg and Willis, 2007; Seehausen et al., 2014; Dev, 2015). Theory shows that the simplest mechanism for generating new species is through allopatric speciation, in which some portion of a species' range becomes geographically isolated, allowing natural selection or genetic drift to drive allele frequency changes that ultimately generate additional reproductive barriers (Mayr, 1963; Barraclough and Vogler, 2000; Coyne and Orr, 2004). This is because even modest levels of gene flow can homogenize allele frequencies between populations, retarding divergence (Wright, 1931). While local adaptation can drive peripatric or sympatric divergence in cases where the immigrant rate is less than the intensity of selection (Lenormand, 2002), most empirical studies cannot exclude the possibility that speciation was preceded by a period of allopatry (Nadachowska-Brzyska et al., 2013; Shaner et al., 2015). This presents a paradox in species-rich groups like mosses, where long-distance migration appears to be common, but speciation and diversification have occurred in spite of the fact that geographic barriers may not cause a long-term impediment to gene flow (Shaw et al., 2003; Piñeiro et al., 2012; Lewis et al., 2014a; Szövényi et al., 2014; Barbé et al., 2016).

One potential resolution of this paradox is sympatric speciation through polyploidy, which is frequent in flowering plants (Ramsey and Schemske, 1998; Mallet, 2005) and potentially in mosses (Crosby, 1980; Kuta and Przywara, 1997; Såstad, 2005; McDaniel et al., 2010; Rensing et al., 2013). Polyploidy generates a strong reproductive barrier in a single mutational event (Ramsey and Schemske, 1998; Madlung, 2013). However, the homogeneity in bryophyte genome sizes (Voglmayr, 2000) raises the possibility that the role played by

polyploidy in moss speciation may be small in relation to other speciation mechanisms. The nature of the genomic, demographic, or ecological factors (beyond geographic isolation and polyploidy) that generate reproductive barriers between nascent species of mosses remain poorly characterized (McDaniel et al., 2010; Yousefi et al., 2017).

Within mosses, the genetic basis of reproductive barriers is best characterized among populations of *Ceratodon purpureus* (Hedw.) Brid. (Ditrichaceae) (McDaniel et al., 2007, 2008). Moreover, the developing genomic and laboratory tools make this species a promising model for further ecological genomic study (McDaniel et al., 2016). *Ceratodon purpureus* is abundant on every continent and grows on a wide variety of substrates (Crum, 1973). Molecular population genetic analyses indicated that gene flow among Northern Hemisphere and even Southern Hemisphere populations was frequent but that tropical populations were more genetically isolated (McDaniel and Shaw, 2005). These observations suggest that the current level of sampling may be insufficient to detect the full scope of population structure among populations in this taxon. Indeed, partial hybrid breakdown was clearly evident in crosses between a temperate and a tropical population, suggesting that reproductive barriers may be in the process of evolving between ecologically distinct regions of the distribution of *C. purpureus* (McDaniel et al., 2007, 2008). These barriers did not involve ploidy differences. However, the genome size of *C. purpureus* is well characterized in only a modest number of European samples (mean ± SD = 0.39 ± 0.0046 pg, n = 10; Voglmayr, 2000), leaving open the possibility that polyploidy contributes to reproductive isolation among isolates from other parts of its broad cosmopolitan distribution.

In a previous phylogeographic analysis (McDaniel and Shaw, 2005), the Mediterranean region was found to contain several rare haplotypes that were distantly related to the common haplotypes found throughout the range of *C. purpureus*. In the present study, we tested for the existence of any relationship between the genetic diversity and DNA content found in the Mediterranean area in the moss genus *Ceratodon*. McDaniel and Shaw (2005) argued that frequent gene flow maintained the genetic homogeneity of the species, at least among the temperate Northern Hemisphere populations, but that the divergent populations were simply outside the main area of spore rain, and therefore had not yet been homogenized. Alternatively, these isolated populations could represent cryptic species, and reproductive isolation evolved in spite of this gene flow (McDaniel et al., 2007, 2008). To distinguish between these alternatives, we evaluated the patterns of polymorphism in five nuclear introns and a single chloroplast locus in plants sampled from mountainous areas of the Mediterranean region and other mountain regions and lowlands, mostly from Southern Europe. We also estimated the genome size of these isolates using flow cytometry. These data clearly show that a new species has evolved within the genus *Ceratodon*, accompanied both by large non-polyploid and allopolyploid changes in genome size and, potentially, by major changes in sexual system. These insights also highlight the complexity of peripatric speciation mechanisms in bryophytes.

MATERIALS AND METHODS

Plant material

For this study, we generated genetic data for a total of 93 samples, 71 (76.4%) from Mediterranean mountain areas (47 from the Spanish Sierra Nevada, 19 from the Spanish central mountain ranges, three from the Spanish southeastern mountains, and two from Sicilian Mount Etna). Of the remaining 22 samples, 11 (11.8%) were from other European mountainous systems (eight from the Alps and three from the Pyrenees) and 11 specimens (11.8%) were from lowlands (three from Czech Republic, two from Germany, two from Sweden, two from United Kingdom, and two from South Africa). Mainly between April and November 2011–2014 (for more detailed information, see Appendix 1), we collected 84 new samples for this study, all of which are deposited at MUB (Herbarium of the University of Murcia, Spain); nine samples were loaned from herbaria, including BOL (Bolus Herbarium, University of Cape Town, South Africa), CBFS (University of South Bohemia, Czech Republic), and S (Herbarium of the Swedish Museum of Natural History, Sweden); and two samples were donated from Laura Forrest (at Royal Botanic Garden Edinburgh, United Kingdom). We sequenced four specimens of *Cheilothela chloropus* (Brid.) Lindb. to use as an outgroup (voucher information and GenBank accession numbers are listed in Appendix 1).

DNA sequencing

To examine the genealogical relationships among the 93 isolates, we sequenced five nuclear exon-primed intron-spanning loci, including *rpL23A* and *TRc1b3.05* (McDaniel et al., 2013b; referenced by EST accessions AW086590 and AW098560), *hp23.9*, *PPR*, and *TBP* (McDaniel et al., 2013a, b), and a single chloroplast locus (*trn*L). We amplified all loci from all individuals in 20 μL polymerase chain reaction (PCR) using DreamTaq DNA Polymerase (Thermo Fisher Scientific, Waltham, Massachusetts, USA). The cycling conditions were 94°C for 2 min, then 10 cycles of 94°C for 15 s, an annealing temperature of 65°C that dropped one degree each cycle, and 72°C for 1 min, followed by 20 cycles of 94°C for 15 s, 56°C for 30 s, and 72°C for 1 min, and terminating with 72°C for 7 min (McDaniel et al., 2013a). To make the resulting PCR products ready for sequencing, we removed unincorporated primers and inactivated unincorporated nucleotides using PCR cleanup reaction with Exo I and FastAP Alkaline Phosphatase enzymes (Thermo Fisher Scientific). Both enzymes were heat inactivated by maintaining the mixture at 85°C for 15 min. Sequencing was accomplished on an ABI3730XL DNA Analyzer, Applied Biosystems (Macrogen Europe, Amsterdam, The Netherlands).

Cloning of DNA sequences

In samples where we observed double peaks in the chromatograms, we cloned all loci. PCR products were isolated from agarose gels, and cloned using the CloneJet PCR Cloning Kit (Thermo Fisher Scientific). Cloning efficiency and accuracy were checked using PCR reactions. Successful clones then were sequenced using an ABI3730XL DNA Analyzer (Macrogen).

Phylogenetic analyses

We aligned the DNA sequences using CLUSTALW (Larkin et al., 2007) as implemented in Bioedit (Hall, 1999) and manually resolved inconsistencies in the resulting alignment. DnaSP version 5 (Librado and Rozas, 2009) was used to observe characteristics such as total length with and without gaps, number of constant positions, and number of parsimony-informative variable positions

about all loci. We coded gaps as informative with a simple indel coding strategy (Simmons and Ochoterena, 2000) implemented in SeqState (Müller, 2005). We performed phylogenetic analyses using MrBayes version 3.2 (Ronquist et al., 2012). The need for a priori model testing was removed using the substitution model space in the Bayesian MCMC analysis itself (Huelsenbeck et al., 2004) with the option nst=mixed. The sequence and indel data were treated as separate and unlinked partitions. The a priori probabilities supplied were those specified in the default settings of the program. Posterior probability distributions of trees were generated using the Metropolis-coupled Markov chain Monte Carlo (MCMCMC) method. To search for convergence in the phylogenetic analyses, we used two runs with different settings for some of the loci. For *hp23.9*, *TBP*, and *trn*L, four chains with 1×10^7 generations were run simultaneously, with the temperature of the single heated chain set to the default in MrBayes. Nevertheless, eight chains with 1×10^6 generations each were run, changing the temperature of the single heated chain set to 2 (*PPR*), 3 (*TRc1b3.05*), and 6 (*rpL23A*), because with the default temperature setting, convergence was not reached in initial runs. Chains were sampled every 1000 generations and the respective trees were written into a tree file. The first 25% of the total sampled trees of each run were discarded as burn-in. Consensus trees and posterior probabilities of clades were calculated by combining the two runs and using the trees sampled after the chains converged and had become stationary. The sump command of MrBayes was used to check whether an appropriate sample from the posterior was obtained. To do so, we first inspected visually the log likelihood plot, which should not show tendencies to decrease or increase over time; the different runs should show similar values. Then we checked that the effective sampling size (ESS) values for all parameters reached ≥500 and that the potential scale reduction factor (PSRF) for each parameter was ~1.00. The genealogies were rooted with sequences from *Cheilothela chloropus*. The final trees were edited with TreeGraph2 (Stöver and Müller, 2010). We performed phylogenetic analyses using the same setting as before, combining the new sequences generated here with other sequences for the *TBP* locus available on GenBank from Antarctica (1), Australia (1), and eastern North America (54), which were previously reported by McDaniel et al. (2013a).

Low resolution in phylogenetic reconstructions can sometimes be caused by incongruence or conflicts in the molecular datasets that lead to different, equally possible, solutions (Huson and Bryant, 2006; Draper et al., 2015). To evaluate this possibility, we reconstructed a phylogenetic network based on the neighbor-net method (Bryant and Moulton, 2004) using the program SplitsTree4 version 4.13.1 (Huson and Bryant, 2006) for the six concatenated loci. The calculations were based on uncorrected *p*-distances. To test the hypothesis of recombination in each graph, a pairwise homoplasy index (phi-test) was calculated, which is a robust and reliable statistic to detect recombination. This estimates the mean refined incompatibility score from nearby sites. Under the null hypothesis of no recombination, the genealogical correlation of adjacent sites is invariant to permutations of the sites because all sites have the same history. The order of the sites is important when levels of recombination are finite, because distant sites will tend to have less genealogical correlation than adjacent sites (Bruen et al., 2006). The significance is then tested using a permutation test by default. In accordance with Bruen et al. (2006) for the phi-test of recombination, $P < 0.05$ indicates the presence of recombination signal.

Coalescent stochasticity analyses

Individual gene trees often differ from each other and from the species tree (Rosenberg, 2002; Mao et al., 2014). To check whether the differentiation we found between the Sierra Nevada (SN) and Worldwide (Ww) clades was a good fit to the multispecies coalescent model (MSCM), we employed an approach based on posterior predictive simulation implemented in the R package "P2C2M" (Gruenstaeudl et al., 2016). In this approach, a posterior distribution of gene genealogies estimated from empirical data is compared to a posterior predictive distribution of genealogies simulated under a model of interest. We first used *BEAST (Heled and Drummond, 2010) to infer genealogies and species trees, and the simulation of genealogies under the MSCM with ms program (Hudson, 2002) under a JC +I +G nucleotide substitution model selected as the most probable in all loci using jModelTest (Posada, 2008). Gaps were included as a character state in these analyses. The run was conducted assuming a strict clock for each locus. We selected "Yule Model" as the species tree prior, and employed "Piecewise linear and constant root" as the population size model. Finally, the default values for MCMC analysis were used. We compared the genealogies from the posterior distribution to the species trees and compared the genealogies from the posterior predictive distribution to the species tree using two descriptive summary statistics: *lcwt* (likelihood of the coalescent waiting times) and *ndc* (number of deep coalescences). When samples are drawn from data with a good fit to the MSCM, the summary statistics from each distribution should be approximately equal and the expected difference between the two is zero (Reid et al., 2014). Data that are a poor fit to the MSCM are indicated by a deviation from the expectation of a difference distribution that is centered on zero is encountered above a specified quantile level (Gruenstaeudl et al., 2016).

Genome size determination

We used flow cytometry (FCM) technology for 75 specimens to estimate nuclear DNA content. One shoot of each sample was chopped with a razor blade together with the internal standard *Carex acutiformis* Ehrh. 1C = 0.41 pg (Lipnerová et al., 2012) or *Bellis perennis* L. (1C = 1.56 pg; our own calibration against *Carex acutiformis*) in 1 mL of LB01 buffer (Doležel et al., 1989). The fluorochrome propidium iodide and RNase IIa (both at final concentration 50 µg/mL) were added immediately; the samples were stained for ≥10 min. The samples were analyzed using a Partec CyFlow SL flow cytometer equipped with a 532 nm (green) diode-pumped solid-state laser (100 mW output); the fluorescence intensity of 12,000 particles was recorded. When possible, we used in vitro cultivated fresh material, but for 47 samples that did not grow satisfactorily in vitro, we used dry material collected in the years 2009–2014. The fluorescence histograms were processed using FlowJo version 10.2 (https://www.flowjo.com/).

Sex determination

To determine sex, one plant per sample was employed. We amplified the *rpS15A* sex-linked locus by PCR and digested the product with HindIII. An intron in the *rpS15A* amplicon contains a cut-site difference between the male and female products (Norrell et al., 2014) that is clearly observable in the banding patterns, which were visualized after electrophoresis in an agarose gel and

scored by hand. We identified the sex of 82 samples, 88.17% of the total, which were from the Spanish Sierra Nevada (42), Spanish central mountain ranges (16), Spanish southeastern mountains (3), Sicilian Mount Etna (2), Alps (7), Pyrenees (3), South Africa (2), Germany (2), Czech Republic (3), and Sweden (2). For the remaining samples, we could not unambiguously interpret the pattern in the restriction-site fragment length polymorphism in the *rpS15A* amplicon. We express the results as a proportion of males and computed the 95% confidence interval (CI) for this estimate with the "dbinom" function in R (R Development Core Team, 2017).

Calculation of the binomial proportion confidence interval

If the total number of experiments and the number of positive outcomes of a success-failure experiment are known, it is possible to calculate the CIs for the probability of success. As a consequence, the CI for the proportion of males (or females) in a population (success) based on the results of sex-determination of a given number of individuals can be calculated. We used the "Hmisc" package (Harrell, 2018) in R3.4.3 (R Development Core Team, 2017), with the options "Wilson" and "Exact" (= Clopper-Pearson) to calculate CIs (presented in parentheses below).

RESULTS

Phylogenetic analyses

The sequence alignments varied in total length between 207 (215 with coded gaps) and 848 (891 with coded gaps) positions for *hp23.9* and *rpL23A*, respectively. The number of constant positions was between 186 and 715 for the above-mentioned loci, and the parsimony-informative variable positions differed between 5 and 95 for *trn*L and *rpL23A*, respectively (Table 1). The loci *PPR*, *TBP*, *rpL23A*, and *TRc1b3.05* showed two well-differentiated clades with support of 0.87 (0.85–0.88) to 0.77 (0.75–0.79) posterior probability (pp), 0.96 (0.95–0.96) to 1.00 (1.00–1.00) pp, 1.00 (1.00–1.00) to 1.00 (1.00–1.00) pp, and 1.00 (1.00–1.00) to 1.00 (1.00–1.00) pp, respectively (Fig. 1; see also the Supplemental Data with this article, Appendices S1, S2, S3). In the case of *rpL23A*, sequences of *Cheilothela chloropus* were not obtained for use as outgroup, but again two clades were resolved. The *hp23.9* locus had support for one clade of 1.00 (1.00–1.00) pp, but the other clade had a value of 0.55 (0.45–0.61) pp (Appendix S4). In all five nuclear loci studied, one of the clades was formed always by 34 Sierra Nevada samples and one of the Spanish southeastern mountains (hereafter "SN group"). The second clade consistently included 42 specimens coming from the rest of the sampled areas, including one from the Sierra Nevada

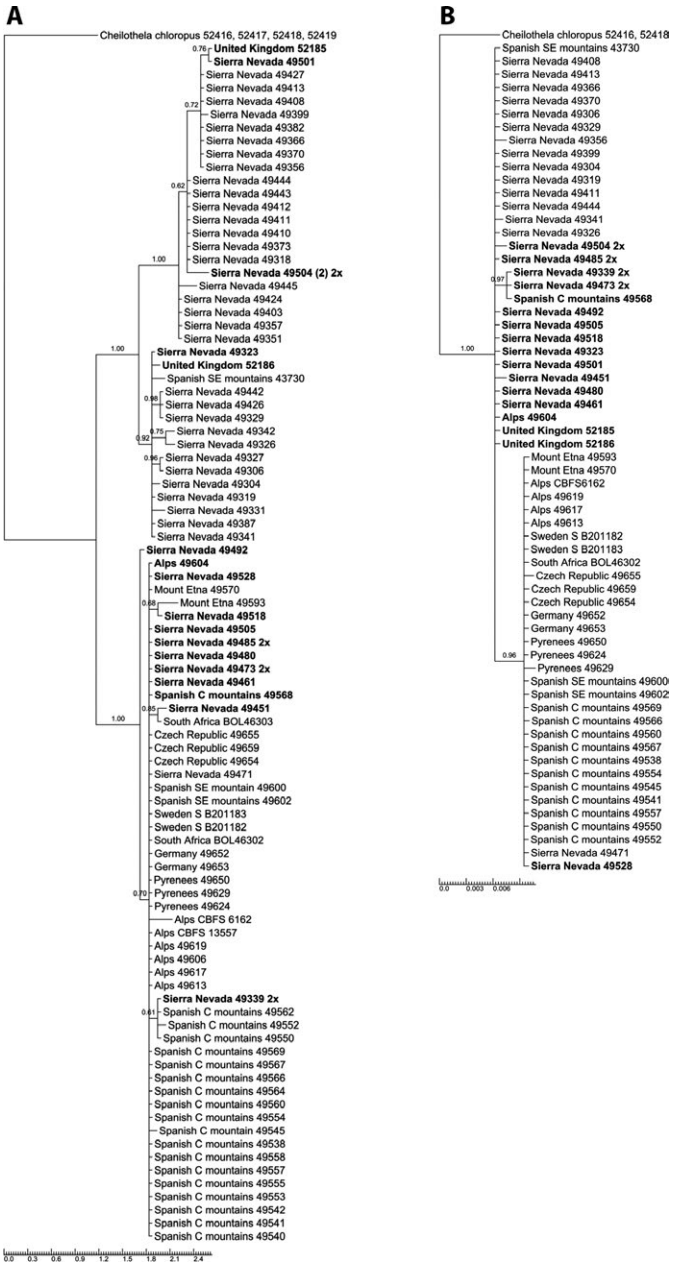

FIGURE 1. Phylogenetic trees inferred from two of the studied loci: (A) nuclear *TRc*1b3.05 and (B) chloroplast *trn*L. For each tip in the trees, geographic origin and herbarium number are given (numbers without letters are from MUB); "2x" indicates diploid samples; the number of sequences obtained by cloning is in parentheses if there was more than one; and bold letters indicate recombinant samples.

TABLE 1. Characteristics of the loci used for molecular evolutionary analyses. The genomic location "nuclear–putative autosomal" is based on unpublished data.

Locus	Genomic location	Sequence length (with gaps)	Invariant sites	Parsimony-informative sites
hp23.9	Nuclear–autosomal	207 (215)	186	15
PPR	Nuclear–U/V	331 (334)	309	8
rpL23A	Nuclear–putative autosomal	848 (891)	715	95
TBP	Nuclear–autosomal	365 (365)	337	11
TRc1b3.05	Nuclear–putative autosomal	402 (417)	362	28
*trn*L	Chloroplast	320 (320)	311	5

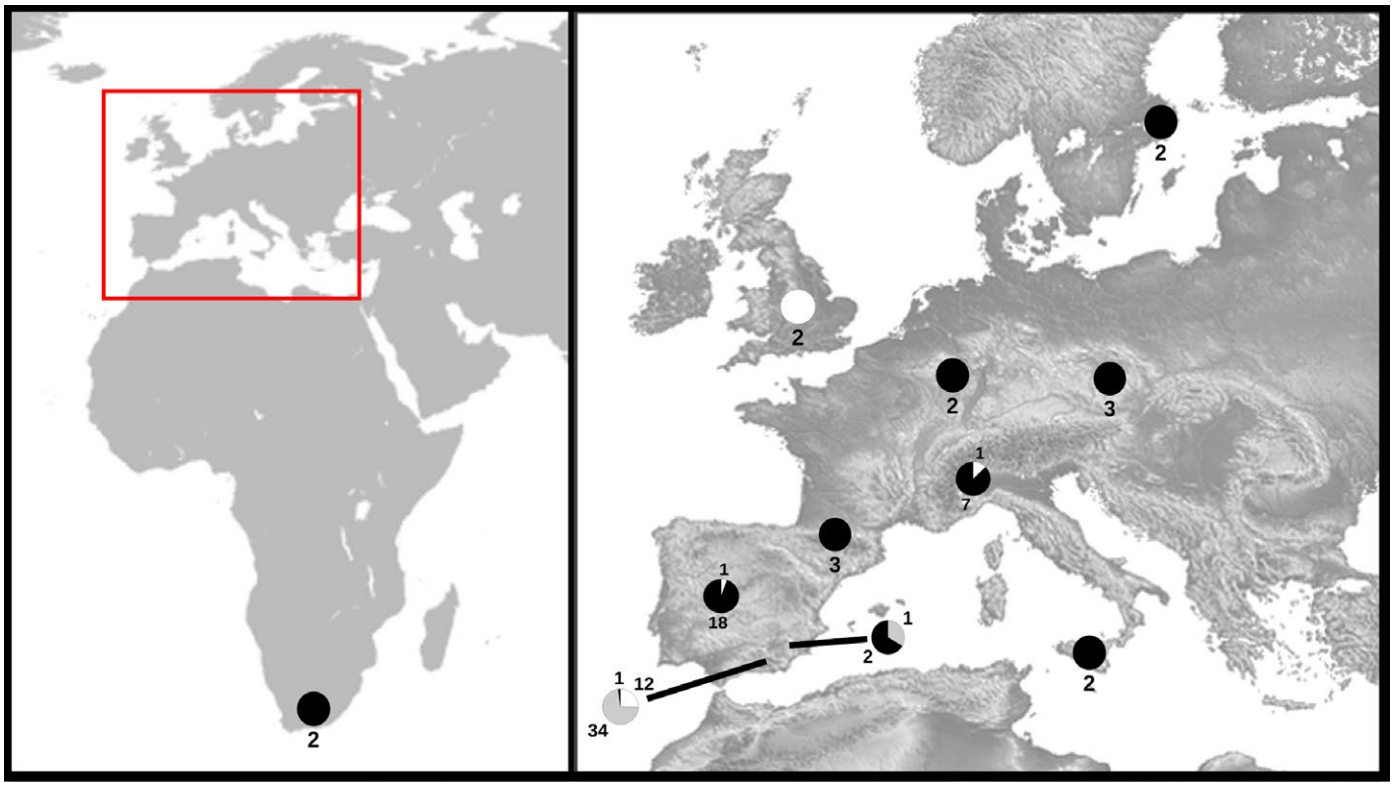

FIGURE 2. Geographic location of *Ceratodon* samples included in this study. Pie charts indicate proportion of samples of each genomic group by area (black: Ww group; gray: SN group; white: recombinant samples), with number of samples for each.

and two from the Spanish southeastern mountains (hereafter "Ww group"). For one marker (*TBP*) we added sequences available at GenBank, including samples from Antarctica, Australia, and North America. The resulting tree topology shows that our samples give a reasonably good representation of the Ww group and that none of these additional sequences is closely related to the SN samples (Appendix S5). The remaining 17 sequenced samples were strongly resolved in either the SN clade or the Ww clade, depending on the studied locus (they did not present intermediate sequences between both clades; Appendix S6); we considered these samples recombinants. The term "hybrid," when applied to bryophytes, should strictly be used only for the sporophytic hybrids (2n) (Anderson, 1980); for their gametophytic progeny (n) showing combination of parental alleles after meiosis, "recombinant" should be used (Shaw, 1994, 1998) in order not to confuse these with hybrids observed among vascular plants. The recombinants were derived mainly from the Sierra Nevada, but also from the Spanish central mountain ranges, the Alps, and the lowlands of the United Kingdom (Fig. 2). The chloroplast locus showed one well-supported clade 0.96 (0.95–0.98) pp, and all remaining samples with deeper coalescent events (Fig. 1). All the samples considered as recombinants based on the nuclear markers were closely related and sister to the rest of the SN samples, with the only exception of one specimen from the Sierra Nevada (MUB 49528), which is a recombinant and belongs to the Ww chloroplast clade.

The apparently uncertain position of some individuals is clarified by the result of the Neighbor-Net network (Fig. 3). Moreover, for the phi-test when the six loci were studied together, a highly significant *P* value (0.0) was obtained, confirming the presence of recombination signal. Graphically, two extreme groups can be observed, the SN group and the Ww group, with some individuals in intermediate positions.

Cloning DNA sequences

Cloning of loci confirmed that diploid specimens (see below) present two different copies of the same loci in most cases. The loci *TRc1b3.05*, *PPR*, and *rpL23A* presented predominantly a single copy, although some individuals presented the two copies in other loci (Appendix S6). Some haploid individuals presented two

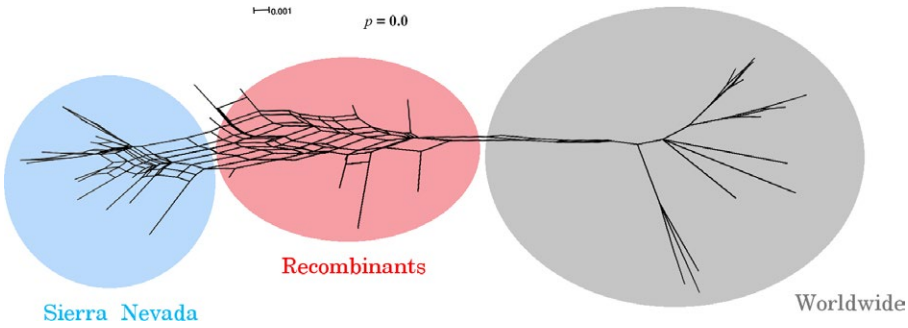

FIGURE 3. Neighbor-Net network to test signals of reticulate evolution between the samples. The main groups are highlighted by the colored circles matched to their names. The *P* value from the phi-test of recombination is indicated.

different copies of a locus. This may be due to the possibility of gene redundancy, which can result from unequal crossing over, retroposition, or chromosomal (or genome) duplication (Magadum et al., 2013).

Coalescent stochasticity analyses

Although our data suggested the existence of recombinants between the two groups, incomplete lineage sorting and hybridization may result in similar molecular signals. Nevertheless, the two summary statistics *ndc* and *lcwt*, employed for the comparison of the genealogies from the posterior distribution to the species trees and from the posterior predictive distribution to the species trees, show significant ($P < 0.05$) differences between our data with respect to MSCM (Table 2), indicating that incomplete lineage sorting (coalescent model) alone cannot explain the different tree topologies.

Flow cytometry analyses

We obtained three clearly differentiated groups of cytotypes for both fresh and dry material (Table 3 and Fig. 4). Measurements from dry material gave higher values (by 18% on average) than those from fresh material, and therefore a conversion factor (1/1.18 = 0.85) was applied to the former. When fresh and dry materials were considered together, the first cytotype had a mean value of 1C = 0.37 pg, and the second one showed 25.4% more DNA content (1C = 0.46 pg). The third cytotype had 1C = 0.82 pg mean value of DNA content. All specimens of the Ww group belonged to the smallest cytotype, while all those of the SN group were categorized in the second cytotype, and the recombinant specimens were found in both the second and the third cytotype (Appendix S6).

TABLE 2. Results of P2C2M analysis in which *lcwt* and *ndc* descriptive summary statistics are shown for each DNA locus analyzed. All loci under study are of nuclear origin, except *trn*L. Asterisk indicates poor model fit at a probability level of 0.05; n.a. = not applicable.

Locus	*lcwt*	*ndc*
hp23.9	510.16 ± 117.86*	−57.46 ± 29.93*
PPR	530.67 ± 121.31*	−59.88 ± 29.87*
rpL23A	462.75 ± 110.74*	−53.04 ± 29.16*
TBP	525.94 ± 120.55*	−59.56 ± 29.77*
TRc1b3.05	516.43 ± 121.96*	−58.31 ± 29.96*
*trn*L	n.a.	−65.33 ± 29.26*
Sum of all genes	n.a.	−353.59 ± 105.62*

Sex determination

All samples from the SN group (29) and all the recombinant samples (15) were females, while the Ww group (38) consisted mainly of females and two males (one from the Sierra Nevada and the other from the Alps; see Appendix 1). The high proportion of females in the Ww samples may be due to a collection bias, given that we preferentially chose moss cushions with sporophytes, because sporophyte morphology is one of the few characters that enabled us to make a clear distinction in the field between *Ceratodon* and other morphologically similar genera (e.g., in the Bryales Limpr. and Pottiales M. Fleisch.). The presence of sporophytes, however, indicates that males must have been present. Male buds (perigonia) are deciduous and may not be produced each season, which means that sterile plants may have been males. In the Sierra Nevada, we never observed sporophytes (the identity of all samples was verified in the laboratory by examining microscopic gametophytic characters). Similarly, Rams et al. (2014) reported finding no sporophytes in deep sampling carried out in the Sierra Nevada from early spring to autumn from 2002 to 2004, and García-Zamora et al. (1998) found only one fructified specimen identified as *C. purpureus* in a survey of a zone close to the Sierra Nevada in 1990–1991. Moreover, none of the *Ceratodon* samples from southeastern Spain in the MUB and GDA/GDAC (University of Granada, Spain) herbaria showed sporophytes. If we exclude a possible bias in the case of the Sierra Nevada samples, we can conclude, based on the binominal distribution, that the 95% CI for the probability of encountering males in the SN-type populations lies in the range of $p = 0.00$ for the lower limit and $p = 0.12$ for the upper limit for both tested methods ("Wilson" and "Clopper-Pearson"), which means that males might even be completely absent.

DISCUSSION

In most major models of speciation, a period of allopatry is essential to the evolution of reproductive isolation (Coyne and Orr, 2004). However, in many cosmopolitan species, including many mosses and ferns, the entire habitable range of the species is within the range of the dispersal distance of its spores (Muñoz et al., 2004;

TABLE 3. Nuclear DNA content as measured by flow cytometry. Cytotypes considered, number of samples used in the analyses (*n*), mean value of DNA, standard deviation, and range of values obtained for each cytotype are given (asterisk indicates conversion factor of 0.85 applied to dry material when fresh and dry material are combined).

	Cytotype	*n*	Mean (pg)	Standard deviation	Minimum (pg)	Maximum (pg)
Fresh material	a	5	0.36	<0.01	0.36	0.37
	b	20	0.46	0.01	0.45	0.48
	c	3	0.81	0.01	0.81	0.82
Dry material	a*	25	0.44	0.01	0.41	0.45
	b*	21	0.54	0.01	0.52	0.57
	c*	1	0.97	–	–	–
Fresh + dry material (*)	a+a*	30	0.37	0.01	0.35	0.38
	b+b*	41	0.46	0.01	0.44	0.48
	c+c*	4	0.82	0.01	0.81	0.82
		75				

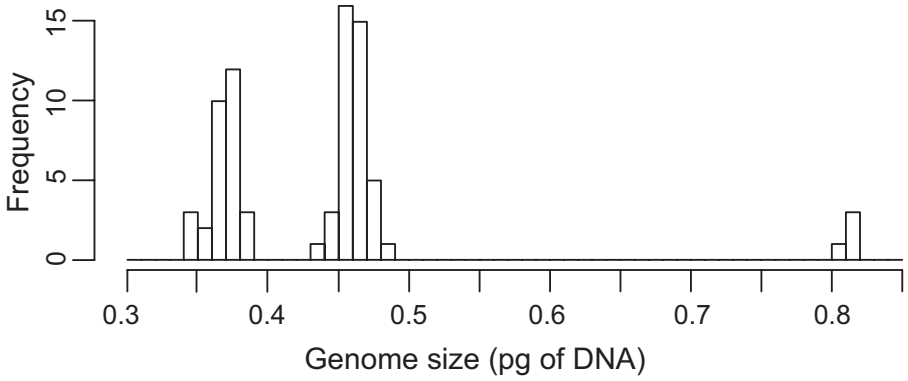

FIGURE 4. Histogram of genome sizes of representative samples of *Ceratodon* generated by flow cytometry. A conversion factor of 0.85 was applied to the data obtained from dry material.

Frahm, 2007; Pisa et al., 2013), making strict allopatry unlikely. Therefore, it is reasonable to propose that speciation mechanisms that either occur in sympatry or accommodate some gene flow contribute to generating the extant diversity in such groups. The two best-studied sympatric speciation mechanisms in plants are polyploidy and the evolution of self-fertilization (Barringer, 2007). Here, we have shown that the evolution of a new species, closely related to the cosmopolitan *Ceratodon purpureus*, was associated with a 25% increase in genome size and a significant decrease in frequency of males (Nieto-Lugilde et al., 2018). Surprisingly, although we have found neither males nor evidence of recent sexual reproduction (i.e., sporophytes) in the new species, the genetic diversity among members of this species is relatively high. Despite the long period of isolation suggested by the sequence divergence between *C. purpureus* and the new species, we have found evidence of interspecific hybridization, suggesting that the new species apparently has retained the capacity for sexual reproduction. In a separate paper, we discuss the taxonomic implications of this discovery (Nieto-Lugilde et al., 2018). Here, we use genealogical and genome size data to make inferences regarding the genetic architecture of speciation, and the demographic parameters that permit such divergence.

Taxonomists have struggled with species delimitation in the genus *Ceratodon* since the description of the genus. Burley and Pritchard (1990) found references for nearly 50 specific or subspecific taxa within *Ceratodon*, but—based on an extensive survey of herbarium specimens—recognized only four species: *C. antarcticus* Cardot., *C. conicus* (Hampe) Lindb., *C. heterophyllus* Kindb., and *C. purpureus*, including three infraspecific taxa (subsp. *convolutus* (Reichardt) Burley, subsp. *purpureus*, and subsp. *stenocarpus* (Bruch & Schimp.) Dixon). Previous molecular population genetic analyses indicated that disjunct populations of *C. purpureus* were sometimes very closely related, clearly showing that long-distance dispersal, even among continents, was frequent enough to erase any signal of strong population structure (McDaniel and Shaw, 2005). However, these data did not provide strong genealogical support either for or against the existence of distinct species other than *C. purpureus*. Subsequent classical genetic analyses showed that geographically and ecologically distant populations were partially reproductively isolated from one another (McDaniel et al., 2007, 2008), but these appeared to be somewhat porous reproductive barriers, and it was unclear that the populations represented different species.

McDaniel and Shaw (2005) did find some isolates of *C. purpureus* that were genetically distant from the more common haplotypes found in northern temperate regions. Here, we found strong evidence that haplotypes that are distantly related to the typical *C. purpureus* haplotypes are locally abundant in the Sierra Nevada of southern Spain. We also found populations containing SN haplotypes and recombinants, together with some rare samples with exclusively the typical *C. purpureus* haplotypes. To evaluate the possibility that the segregation of these divergent haplotypes in the SN populations represents the retention of ancestral variation in the species (i.e., coalescent stochasticity causing incomplete lineage sorting), we generated coalescent simulations using *BEAST and P2C2M. These analyses showed that the divergence between these two haplotypic classes was too great to be explained by coalescent stochasticity. The fact that this polymorphism is found in all the nuclear loci that we sampled, and that it is geographically concentrated in the Sierra Nevada region, suggests that balancing selection is also an unlikely explanation. Collectively, these data suggest that the SN haplotypes constitute a rare species, sister to and partially reproductively isolated from the cosmopolitan *C. purpureus*.

The default mode for the evolution of reproductive isolation is allopatric speciation. The sympatric occurrence of typical *C. purpureus* haplotypes and SN haplotypes, even at modest frequencies, contradicts the suggestion by McDaniel and Shaw (2005) that the Mediterranean populations were isolated from the rest of the species as a result of decreased spore rain in peripheral populations separated by prevailing global wind patterns. If we assume that the current dispersal capabilities of *C. purpureus* represent the ancestral condition, this suggests that geography may not have been the primary isolating mechanism between the nascent species. Morphological analyses of plants of both species grown in a common garden (as well as putative recombinants between them; Nieto-Lugilde et al., 2018) indicate that members of the Ww group can be distinguished morphologically from the SN group on the basis of multivariate biometrical evaluation of microscopic features of the caulidia and phyllidia (stem length, presence or absence of apical comal tuft, leaf size and shape, leaf costa width at base of lamina, and leaf costa excurrence). Nevertheless, we were unable to distinguish between the SN group plants and recombinants in field samples, which suggests that the environment influences the variance in taxonomically important characters. It is certainly possible that an extrinsic factor, like a habitat preference, isolated the two species.

It is also possible that an intrinsic factor isolated the two species. Remarkably, however, we detected only females in the SN species, which suggests that male lethality could contribute to isolating the two species. Sex in dioecious bryophytes like *C. purpureus* is determined at meiosis, by the segregation of a UV chromosome pair, meaning that ~50% of the spores produced in a population should be males. Some meiotic sex ratio variation has been observed in this species in natural populations (overall mean of proportion of males was 0.41 [95% CI: 0.17–0.72]; Norrell et al., 2014) and artificial crosses (male-biased sex ratio = 60%; McDaniel et al., 2008). Even given our sample size (n = 29, with no males), we can conclude that the percentage of males in the SN populations is much lower (95%

CI included 0–12%; additional samples not included in this study lower the 95% CI to a range of 0–6.7%). We do not know whether the decrease in males coincided with the speciation event or occurred subsequent to the evolution of reproductive isolation. The evolution of apomixis or obligate selfing from historically outcrossing lineages is a well-documented route to the evolution of new species in plants (Stebbins, 1974; Barrett, 2010; Wright et al., 2013), and parthenogenetic lineages associated with the loss of males are frequent in some animal lineages (insects: Hagimori et al., 2006; Montelongo and Gómez-Zurita, 2015; vertebrates: Neaves and Baumann, 2011; Gutekunst et al., 2018). However, we know of no other cases where the loss of males has been associated with speciation in bryophytes.

The presence of recombinants containing both typical *C. purpureus* alleles and alleles from the SN species indicated that rare interspecies hybridization has occurred between individuals of the two species. Most of the recombinants possessed the SN chloroplast type, based on the *trnL* sequence data, which suggests that this species was more often the maternal parent (consistent with the rarity of males). We found one instance of a recombinant plant that had a typical *C. purpureus trn*L sequence, but we cannot determine whether this was a rare case of a hybridization involving an SN male (i.e., a cross in the opposite direction) or whether this resulted from a backcross of a male recombinant to a typical *C. purpureus* female. Intrinsic genetic incompatibilities are often manifest as Dobzhansky-Muller interactions, which result in asymmetric introgression patterns at the causative loci (McDaniel et al., 2008) due to the death of incompatible multilocus genotypes. Although we sampled only six loci across the genome, the recombinants tended to have the SN alleles at the *TBP* and *rpL23A* loci. We are currently examining the frequency of polymorphism across the genome of the SN and recombinant genotypes to distinguish among forms of extrinsic and intrinsic isolation between the SN and typical *C. purpureus* populations.

The flow cytometric data also showed that members of the SN species had a genome ~25% larger than typical members of *C. purpureus*. It is possible that the speciation involved a whole-genome duplication event followed by rapid genome reduction, the duplication of large chromosomes (Inoue et al., 2015; Panchy et al., 2016), or the accumulation of transposable elements (TEs), which contribute to the extraordinary variation in genome size within even closely related species in angiosperms (Vitte and Bennetzen, 2006). Although the current data represent the most comprehensive sampling of variation in genome size in *Ceratodon*, we still lack cytological data to determine whether variation in nuclear DNA content is due to an increase in the size of chromosomes or to the increase in the number of chromosomes. The variance in genome size is almost equal between the two groups, which suggests that the SN species is fixed for whatever loci underlie the genome size change. Additionally, recombinants between the two groups have the genome size of SN species, not an intermediate value, suggesting that the increase in genome size may come from a single genomic change, rather than many small changes across genome. One hypothesis is that these plants have gained DNA on the sex chromosome, which comprises nearly one-third of the genome (Heitz, 1932; Jachimsky, 1935; McDaniel et al., 2007). Sex chromosomes in other organisms are known to accumulate genomic material rapidly, sometimes in large translocations, potentially generating pronounced evolutionary and ecological consequences (Tennesse et al., 2017). We are now attempting to generate artificial crosses to evaluate the genetic basis of the genome size difference.

We also found a third rare cytotype with a genome size approximately twice that of either SN plants or typical *C. purpureus* plants. These isolates all had mixed haplotypes (i.e., gene sequences from both the SN and typical *C. purpureus* clades) and a genome size very close to the sum of the SN group and Ww group (~1.2% smaller than the sum of the group means), suggesting that they arose from an allopolyploid event. Without more sequence or cytological data, we cannot formally eliminate the possibility that the larger cytotype arose from autopolyploidy followed by hybridization, although this would require the gain of ~10% or loss (~12%) of the genomic DNA. Additionally, allopolyploidy is a widely observed mechanism to restore the fertility of F_1 hybrids between partially reproductively isolated species with karyotypic differences and exhibit meiotic abnormalities (De Storme and Mason, 2014).

Finally, the new SN species apparently maintains levels of genetic diversity nearly equivalent to typical populations of its sister species *C. purpureus* without obviously undergoing sexual reproduction. Sexual reproduction in mosses occurs when males and females grow in close proximity and sperm cells disperse, typically in humid conditions, from male to female plants, producing sporophytes, a very common observation in most populations of *C. purpureus*. Given the complete absence of sporophytes in observed Sierra Nevada samples, reproduction seems to be predominantly by fragmentation of the gametophores. Moss gametophores can persist for many years, even in relatively stressful conditions, and can easily spread clonally by fragmentation. In some cases, such fragments may be dispersed a considerable distance (Frahm, 2007; Lewis et al., 2014b). It is clear that spatially heterogeneous selection (Vrijenhoek, 1978) or frequency-dependent selection (Weeks and Hoffmann, 2008) can maintain high genetic diversity in clonal organisms. Antarctic populations of *C. purpureus*, which similarly lack any sexual reproduction, were also quite variable, although less polymorphic than was observed in the closely related nearby populations from Australia (Clarke et al., 2009). Also similar to the Antarctic studies, we found polymorphic nuclear ITS sequences between samples collected a few meters apart (M. Nieto-Lugilde, O. Werner, S. F. McDaniel, R. M. Ros, unpublished data), indicating that these localities were colonized several times independently. However, unlike the Antarctic case, the SN isolates are genetically distinct from any known spore source. It is possible that sexual reproduction in the SN species generated the current variation under a past climate regime, or in undetected localities, although it is clearly far rarer than in *C. purpureus*. Further analyses of the evolutionary history of the SN population are likely to produce a better understanding of the phenomena that generate new species in cosmopolitan taxa.

ACKNOWLEDGEMENTS

The authors thank the curators of herbaria BOL, GDA/GDAC, and S for providing plant material; L. Forrest of Royal Botanic Garden Edinburgh, United Kingdom, for donating samples; the management group of the National Park of Sierra Nevada Natural Space for allowing us to collect samples in the National and Natural Parks; and two anonymous reviewers for their valuable comments, which improved the manuscript. This study was supported financially by the Spanish "Ministerio de Ciencia e Innovación" (project CGL2011-22936/BOS), by "Ministerio de

Economía y Competitividad" (project CGL2014-52579-R), co-financed by ERDF of the European Union, and by a U.S. National Science Foundation grant (DEB 1541005) to S.F.M. M.N.-L. also thanks the Spanish "Ministerio de Economía y Competitividad" for a "Formación de Personal Investigador" fellowship (FPI program 2012; reference BES-2012-056799).

DATA ACCESSIBILITY

The aligned sequences and trees are available at TreeBase: http://purl.org/phylo/treebase/phylows/study/TB2:S22472

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

LITERATURE CITED

Anderson, L. E. 1980. Cytology and reproductive biology of mosses. *In* R. J. Taylor and L. E. Levotin [eds.], The mosses of North America, 37–76. California Academy of Sciences, San Francisco, California, USA.

Barbé, M., N. J. Fenton, and Y. Bergeron. 2016. So close and yet so far away: Long-distance dispersal events govern bryophyte metacommunity reassembly. *Journal of Ecology* 104: 1707–1719.

Barraclough, T. G., and A. P. Vogler. 2000. Detecting the geographical pattern of speciation from species-level phylogenies. *American Naturalist* 155: 419–434.

Barrett, S. C. H. 2010. Understanding plant reproductive diversity. *Philosophical Transactions of the Royal Society B* 365: 99–109.

Barringer, B. C. 2007. Polyploidy and self-fertilization in flowering plants. *American Journal of Botany* 94: 1527–1533.

Bruen, T. C., H. Philippe, and D. Bryant. 2006. A simple and robust statistical test for detecting the presence of recombination. *Genetics* 172: 2665–2681.

Bryant, D., and V. Moulton. 2004. Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution* 21: 255–265.

Burley, J. S., and N. M. Pritchard. 1990. Revision of the genus *Ceratodon* (Bryophyta). *Harvard Papers in Botany* 1: 17–76.

Clarke, L. J., D. J. Ayre, and S. A. Robinson. 2009. Genetic structure of East Antarctic populations of the moss *Ceratodon purpureus*. *Antarctic Science* 21: 51–58.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland, Massachusetts, USA.

Crosby, M. R. 1980. Polyploidy in bryophytes with special emphasis on mosses. *In* W. H. Lewis [ed.], Polyploidy: Biological relevance, 93–198. Plenum Press, New York, New York, USA.

Crum, H. A. 1973. Mosses of the Great Lakes forest. University of Michigan Herbarium, Ann Arbor, Michigan, USA.

De Storme, N., and A. Mason. 2014. Plant speciation through chromosome instability and ploidy change: Cellular mechanisms, molecular factors and evolutionary relevance. *Current Plant Biology* 1: 10–33.

Dev, S. B. 2015. Unsolved problems in biology—The state of current thinking. *Progress in Biophysics and Molecular Biology* 117: 232–239.

Doležel, J., P. Binárová, and S. Lucretti. 1989. Analysis of nuclear DNA content in plant cells by flow cytometry. *Biologia Plantarum* 31: 113–120.

Draper, I., L. Hedenäs, M. Stech, J. Patiño, O. Werner, J. M. González-Mancebo, M. Sim-Sim, et al. 2015. How many species of *Isothecium* (Lembophyllaceae, Bryophyta) are there in Macaronesia? A survey using integrative taxonomy. *Botanical Journal of the Linnean Society* 177: 418–438.

Frahm, J. P. 2007. Diversity, dispersal and biogeography of bryophytes (mosses). *In* W. Foissner and D. L. Hawksworth [eds.], Protist diversity and geographical distribution. vol 8, 43–50. Springer, Dordrecht, The Netherlands.

García-Zamora, P., R. M. Ros, and J. Guerra, J. 1998. Bryophyte flora of the Sierras de Filabres, Cabrera, Alhamilla and Cabo de Gata (Almería, SE Spain). *Journal of Bryology* 20: 461–493.

Gruenstaeudl, M., N. M. Reid, G. L. Wheeler, and B. C. Carstens. 2016. Posterior predictive checks of coalescent models: P2C2M, an R package. *Molecular Ecology Resources* 16: 193–205.

Gutekunst, J., R. Andriantsoa, C. Falckenhayn, K. Hanna, W. Stein, J. Rasamy, and F. Lyko. 2018. Clonal genome evolution and rapid invasive spread of the marbled crayfish. *Nature Ecology & Evolution* 1.

Hagimori, T., Y. Abe, S. Date, and K. Miura. 2006. The first finding of a *Rickettsia* bacterium associated with parthenogenesis induction among insects. *Current Microbiology* 52: 97–101.

Hall, T. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.

Harrell, F. E. 2018. Hmisc: Harrell Miscellaneous. R package version 4.1-1. https://CRAN.R-project.org/package=Hmisc.

Heitz, E. 1932. Geschlectschromosomen bei einem Laubmoss. *Berichte der Deutschen Botanischen Gesellschaft* 30: 204–206.

Heled, J., and A. J. Drummond. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* 27: 570–580.

Hudson, R. R. 2002. Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics* 18: 337–338.

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

Huson, D. H., and D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23: 254–267.

Inoue, J., Y. Sato, R. Sinclair, K. Tsukamoto, and M. Nishida. 2015. Rapid genome reshaping by multiple-gene loss after whole-genome duplication in teleost fish suggested by mathematical modeling. *Proceedings of the National Academy of Sciences USA* 112: 14918–14923.

Jachimsky, H. 1935. Beitrag zur Kenntnis von Geschlechtschromosomen und Heterochromatin bei Moosen. *Jahrbücher für Wissenschaftliche Botanik* 81: 203–238.

Kuta, E., and L. Przywara. 1997. Polyploidy in mosses. *Acta Biologica Cracoviensia* 39: 17–26.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, et al. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.

Lenormand, T. 2002. Gene flow and the limits to natural selection. *Trends in Ecology & Evolution* 17: 183–189.

Lewis, L. R., E. Behling, H. Gousse, E. Qian, C. S. Elphick, J.-F. Lamarre, J. Bêty, et al. 2014b. First evidence of bryophyte diaspores in the plumage of transequatorial migrant birds. *PeerJ* 2: e424.

Lewis, L. R., R. Rozzi, and B. Goffinet. 2014a. 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.

Librado, P., and J. Rozas. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.

Lipnerová, I., P. Bureš, L. Horová, and P. Šmarda. 2012. Evolution of genome size in *Carex* (Cyperaceae) in relation to chromosome number and genomic base composition. *Annals of Botany* 111: 79–94.

Madlung, A. 2013. Polyploidy and its effect on evolutionary success: Old questions revisited with new tools. *Heredity* 110: 99–104.

Magadum, S., U. Banerjee, P. Murugan, D. Gangapur, and R. Ravikesavan. 2013. Gene duplication as a major force in evolution. *Journal of Genetics* 92: 155–161.

Mallet, J. 2005. Hybridization as an invasion of the genome. *Trends in Ecology & Evolution* 20: 229–237.

Mao, X., G. Zhu, L. Zhang, S. Zhang, and S. J. Rossiter. 2014. Differential introgression among loci across a hybrid zone of the intermediate horseshoe bat (*Rhinolophus affinis*). *BMC Evolutionary Biology* 14: 154.

Mayr, E. 1963. Animal species and evolution. Harvard University Press, Cambridge, Massachusetts, USA.

McDaniel, S. F., K. M. Neubig, A. C. Payton, R. S. Quatrano, and D. J. Cove. 2013a. Recent gene-capture on the UV sex chromosomes of the moss *Ceratodon purpureus*. *Evolution* 67: 2811–2822.

McDaniel, S. F., P. F. Perroud, A. C. Cuming, and P. Szövényi. 2016. The *Ceratodon purpureus* transcriptome ushers in the era of moss comparative genomics. *In* S. A. Rensing and P. F. Perroud [eds.], Genomes and evolution of charophytes, bryophytes, lycophytes and ferns, 141–162. Elsevier Press, Oxford, UK.

McDaniel, S. F., and A. J. Shaw. 2005. Selective sweeps and intercontinental migration in the cosmopolitan moss *Ceratodon purpureus* (Hedw.) Brid. *Molecular Ecology* 14: 1121–1132.

McDaniel, S. F., M. J. van Baren, K. S. Jones, A. C. Payton, and R. S. Quatrano. 2013b. Estimating the nucleotide diversity in *Ceratodon purpureus* (Ditrichaceae) from 218 conserved exon-primed, intron-spanning nuclear loci. *Applications in Plant Sciences* 1: 1200387.

McDaniel, S. F., M. von Stackelberg, S. Richardt, R. S. Quatrano, R. Reski, and S. A. Rensing. 2010. The speciation history of the Physcomitrium-Physcomitrella species complex. *Evolution* 64: 217–231.

McDaniel, S. F., J. H. Willis, and A. J. Shaw. 2007. A linkage map reveals a complex basis for segregation distortion in an interpopulation cross in the moss *Ceratodon purpureus*. *Genetics* 176: 2489–2500.

McDaniel, S. F., J. H. Willis, and A. J. Shaw. 2008. The genetic basis of developmental abnormalities in interpopulation hybrids of the moss *Ceratodon purpureus*. *Genetics* 179: 1425–1435.

Montelongo, T., and J. Gómez-Zurita. 2015. Nonrandom patterns of genetic admixture expose the complex historical hybrid origin of unisexual leaf beetle species in the Genus *Calligrapha*. *American Naturalist* 185: 113–134.

Müller, K. 2005. SeqState: Primer design and sequence statistics for phylogenetic DNA datasets. *Applied Bioinformatics* 4: 65–69.

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

Nadachowska-Brzyska, K., R. Burri, P. I. Olason, T. Kawakami, L. Smeds, and H. Ellegren. 2013. Demographic divergence history of Pied Flycatcher and Collared Flycatcher inferred from whole-genome re-sequencing data. *PLOS Genetics* 9: e1003942.

Neaves, W. B., and P. Baumann. 2011. Unisexual reproduction among vertebrates. *Trends in Genetics* 27: 81–88.

Nieto-Lugilde, M., O. Werner, S. F. McDaniel, and R. M. Ros. 2018. Environmental variation obscures species diversity in southern European populations of the moss genus *Ceratodon*. *Taxon* 67. In press.

Norrell, T. E., K. S. Jones, A. C. Payton, and S. F. McDaniel. 2014. Meiotic sex ratio variation in natural populations of *Ceratodon purpureus* (Ditrichaceae). *American Journal of Botany* 101: 1572–1576.

Panchy, N., M. Lehti-Shiu, and S. H. Shiu. 2016. Evolution of gene duplication in plants. *Plant Physiology* 171: 2294–2316.

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.

Pisa, S., O. Werner, A. Vanderpoorten, M. Magdy, and R. M. Ros. 2013. Elevational patterns of genetic variation in the cosmopolitan moss *Bryum argenteum* (Bryaceae). *American Journal of Botany* 100: 2000–2008.

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

Rams, S., O. Werner, and R. M. Ros. 2014. Updated checklist of the Bryophytes from the Sierra Nevada mountains (S of Spain). *Cryptogamie Bryologie* 35: 261–311.

Ramsey, J., and D. W. Schemske. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* 29: 467–501.

R Development Core Team 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Reid, N. M., S. M. Hird, J. M. Brown, T. A. Pelletier, J. D. McVay, J. D. Satler, and B. C. Carstens. 2014. Poor fit to the multispecies coalescent is widely detectable in empirical data. *Systematic Biology* 63: 322–333.

Rensing, S. A., A. K. Beike, and D. Lang. 2013. Evolutionary importance of generative polyploidy for genome evolution of haploid-dominant land plants. *In* J. Greilhuber, J. Dolezel, and J. F. Wendel [eds.], Plant genome diversity. vol 2, 295–305. Springer, Vienna, Austria.

Rieseberg, L. H., and J. H. Willis. 2007. Plant Speciation. *Science* 317: 910–914.

Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, et al. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.

Rosenberg, N. A. 2002. The probability of topological concordance of gene trees and species trees. *Theoretical Population Biology* 61: 225–247.

Såstad, S. M. 2005. Patterns and mechanisms of polyploid speciation in bryophytes. *In* F. T. Bakker, L. W. Chatrou, B. Gravendeel, and P. Pelser [eds.], Plant species-level systematics: New perspectives on pattern & process, 317–333. Gantner Verlag, Ruggell, Liechtenstein.

Seehausen, O., R. K. Butlin, I. Keller, C. E. Wagner, J. W. Boughman, P. A. Hohenlohe, C. L. Peichel, et al. 2014. Genomics and the origin of species. *Nature Reviews Genetics* 15: 176–192.

Shaner, P. J. L., T. H. Tsao, R. C. Lin, W. Liang, C. F. Yeh, X. J. Yang, F. M. Lei, et al. 2015. Climate niche differentiation between two passerines despite ongoing gene flow. *Journal of Animal Ecology* 84: 829–839.

Shaw, A. J. 1994. Systematics of *Mielichhoferia* (Bryaceae: Musci). III. Hybridization between *M. elongata* and *M. mielichhoferiana*. *American Journal of Botany* 81: 782–790.

Shaw, A. J. 1998. Genetic analysis of a hybrid zone in *Mielichhoferia* (Musci). *In* J. W. Bates, N. W. Ashton, and J. G. Duckett [eds.], Bryology for the twenty-first century, 161–174. Maney Publishing and the Brithish Bryological Society, Leeds, UK.

Shaw, A. J., O. Werner, and R. M. Ros. 2003. Intercontinental Mediterranean disjunct mosses: Morphological and molecular patterns. *American Journal of Botany* 90: 540–550.

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

Stebbins, G. L. 1974. Flowering plants: Evolution above the species level. Belknap Press of Harvard University Press, Cambridge, Massachusetts, USA.

Stöver, B. C., and K. F. Müller. 2010. TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics* 11: 7.

Szövényi, P., N. Devos, D. J. Weston, X. Yang, Z. Hock, J. A. Shaw, K. K. Shimizu, et al. 2014. Efficient purging of deleterious mutations in plants with haploid selfing. *Genome Biology and Evolution* 6: 1238–1252.

Tennessen, J. A., N. Wei, S. Straub, R. Govindarajulu, A. Liston, and T. L. Ashman. 2017. Repeated translocation of a gene cassette drives sex chromosome turnover in strawberries. *bioRxiv* 163808.

Vitte, C., and J. L. Bennetzen. 2006. Analysis of retrotransposon structural diversity uncovers properties and propensities in angiosperm genome evolution. *Proceedings of the National Academy of Sciences USA* 103: 17638–17643.

Voglmayr, H. 2000. Nuclear DNA amounts in mosses (Musci). *Annals of Botany* 85: 531–546.

Vrijenhoek, R. C. 1978. Coexistence of clones in a heterogeneous environment. *Science* 199: 549–552.

Weeks, A. R., and A. A. Hoffmann. 2008. Frequency-dependent selection maintains clonal diversity in an asexual organism. *Proceedings of the National Academy of Sciences USA* 105: 17872–17877.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159.

Wright, S. I., S. Kalisz, and T. Slotte. 2013. Evolutionary consequences of self-fertilization in plants. *Proceedings of the Royal Society B* 280: 20130133.

Yousefi, N., K. Hassel, K. I. Flatberg, P. Kemppainen, E. Trucchi, A. J. Shaw, M. O. Kyrkjeeide, et al. 2017. Divergent evolution and niche differentiation within the common peatmoss *Sphagnum magellanicum*. *American Journal of Botany* 104: 1060–1072.

APPENDIX 1. Voucher information for the studied specimens. For each sequenced sample, information is given as follows: herbarium code (acronyms follow Index Herbariorum); geographic origin, collection date (year-month-day), sex if known (F = female, M = male), presence of sporophyte if appropriate (indicated by asterisk), and GenBank accession numbers for the six loci studied in the order *hp23.9*, *PPR*, *rpL23A*, *TBP*, *TRc1b3.05*, *trn*L. Sequences obtained by cloning are indicated by their GenBank accession number (in parentheses).

Ingroup

Mediterranean mountain areas

MUB 43730: Spanish southeastern mountains, 2011-11-13, F, KP825628, KP826017, KP826181, KP826402, KP826531, KY229001. MUB 49304: Spanish Sierra Nevada (hereafter Sierra Nevada), 2012-07-20, F, KP825703, KP826091, KP826265, KP826473, KP826601, MG050779. MUB 49306: Sierra Nevada, 2012-07-20, F, KP825701, KP826089, KP826263, KP826471, KP826599, KY229023. MUB 49318: Sierra Nevada, 2012-07-20, KP825698, KP826086, KP826260, KP826468, KP826596, –. MUB 49319: Sierra Nevada, 2012-07-20, KP825697, KP826085, KP826259, KP826467, KP826595, MG050780. MUB 49323: Sierra Nevada, 2012-07-20, F, KP825696, KP826084, KP826258, KP826466, KP826594, KY229040. MUB 49326: Sierra Nevada, 2012-07-20, F, KP825693, KP826081, KP826255, KP826463, KP826591, MG050781. MUB 49327: Sierra Nevada, 2012-07-20, F, KP825692, KP826080, KP826254, KP826462, KP826590, –. MUB 49329: Sierra Nevada, 2012-07-20, F, KP825690, KP826078, KP826252, KP826460, KP826588, KY229024. MUB 49331: Sierra Nevada, 2012-07-20, F, KP825688, KP826076, KP826250, KP826459, KP826586, –. MUB 49339: Sierra Nevada, 2012-07-20, F, (MG050789, MG050790, MG050791, MG050792, MG050793, MG050794, MG050795, MG050796, MG050797, MG050798, MG050799), (KP826073, MG050748, MG050749, MG050750, MG050751, MG050752), KP826248, (KP826456, MG050761, MG050762, MG050763, MG050764, MG050765), KP826583, KY229035. MUB 49341: Sierra Nevada, 2012-07-20, F, KP825683, KP826071, KP826246, KP826454, KP826581, MG050782. MUB 49342: Sierra Nevada, 2012-07-20, F, KP825682, KP826070, KP826245, KP826453, KP826580, –. MUB 49351: Sierra Nevada, 2012-07-20, F, KP825681, KP826069, KP826244, KP826452, KP826579, –. MUB 49353: Sierra Nevada, 2012-07-20, F, KP825679, KP826067, KP826242, KP826450, –, –. MUB 49356: Sierra Nevada, 2012-07-20, KP825677, KP826065, KP826239, KP826448, KP826577, KY229030. MUB 49357: Sierra Nevada, 2012-07-20, F, KP825676, KP826064, KP826241, KP826447, KP826576, –. MUB 49366: Sierra Nevada, 2012-07-21, F, KP825670, KP826058, KP826238, KP826442, KP826570, KY229011. MUB 49370: Sierra Nevada, 2012-07-21, KP825674, KP826062, KP826234, KP826446, KP826574, KY229015. MUB 49373: Sierra Nevada, 2012-07-21, F, KP825671, KP826059, KP826233, KP826443, KP826571, –. MUB 49382: Sierra Nevada, 2012-07-21, F, KP825669, KP826057, KP826180, KP826441, KP826569, –. MUB 49387: Sierra Nevada, 2012-07-21, F, KP825666, KP826054, KP826230, KP826438, KP826565, –. MUB 49399: Sierra Nevada, 2012-07-21, F, KP825663, KP826051, KP826224, KP826435, KP826563, KY229033. MUB 49403: Sierra Nevada, 2012-07-21, F, KP825660, KP826048, KP826182, KP826432, KP826560, –. MUB 49408: Sierra Nevada, 2012-07-21, F, KP825657, KP826045, KP826222, –, KP826557, KY229005. MUB 49410: Sierra Nevada, 2012-07-21, F, KP825655, KP826043, KP826220, KP826428, KP826555, –. MUB 49411: Sierra Nevada, 2012-07-21, F, KP825654, KP826042, KP826219, KP826427, KP826554, MG050783. MUB 49412: Sierra Nevada, 2012-07-21, F, KP825653, KP826041, KP826218, KP826426, KP826553, –. MUB 49413: Sierra Nevada, 2012-07-21, F, KP825652, KP826040, KP826217, KP826425, KP826552, KY229008. MUB 49424: Sierra Nevada, 2012-07-21, F, KP825651, KP826039, KP826216, KP826424, KP826551, –. MUB 49426: Sierra Nevada, 2012-07-21, F, KP825649, KP826037, KP826214, KP826422, KP826549, –. MUB 49427: Sierra Nevada, 2012-07-21, F, KP825648, KP826036, KP826213, KP826421, KP826548, –. MUB 49442: Sierra Nevada, 2012-07-21, F, KP825643, KP826031, KP826208, KP826417, KP826544, –. MUB 49443: Sierra Nevada, 2012-07-21, F, KP825642, KP826030, KP826207, KP826416, KP826543, –. MUB 49444: Sierra Nevada, 2012-07-21, F, KP825641, KP826029, KP826206, KP826415, KP826542, MG050784. MUB 49445: Sierra Nevada, 2012-07-21, KP825640, KP826028, KP826209, KP826414, KP826541, –. MUB 49451: Sierra Nevada, 2013-07-12, F, (KP825639, MG050800, MG050801, MG050802, MG050803, MG050804, MG050805, MG050806, MG050807, MG050808), (KP826027, MG050753), (KP826204, MG050869, MG050870), KP826413, KP826540, KY229045. MUB 49461: Sierra Nevada, 2013-07-12, F, KP825638, KP826026, KP826203, KP826412, KP826539, KY229052. MUB 49471: Sierra Nevada, 2013-07-12, M, KP825706, KP826094, KP826201, KP826476, KP826604, KY229043. MUB 49473: Sierra Nevada, 2013-07-12, F, (KP825637, MG050809, MG050810, MG050811, MG050812, MG050813, MG050814, MG050815, MG050816, MG050817, MG050818, MG050819), KP826025, (MG050871, MG050872, MG050873, MG050874, MG050875, MG050876), (MG050766, MG050767, MG050768, MG050769, MG050770), KP826538, KY229041. MUB 49480: Sierra Nevada, 2013-07-12, F, (KP825636, MG050820, MG050821, MG050822, MG050823, MG050824, MG050825, MG050826), KP826024, KP826199, KP826410, KP826537, KY229046. MUB 49485: Sierra Nevada, 2013-07-12, F, (KP825635, MG050827, MG050828, MG050829, MG050830, MG050831, MG050832, MG050833), (KP826023, MG050754, MG050755, MG050756, MG050757, MG050758), (MG050877, MG050878, MG050879, MG050880, MG050881, MG050882), (KP826409, MG050771, MG050772, MG050773, MG050774, MG050775, MG050776), KP826536, KY229032. MUB 49492: Sierra Nevada, 2013-07-13, F, (MG050834, MG050835, MG050836, MG050837, MG050838, MG050839, MG050840), KP826022, KP826198, KP826408, –, KY229037. MUB 49501: Sierra Nevada, 2013-07-13, F, KP825633, –, KP826197, KP826407, KP826535, KY229042. MUB 49504: Sierra Nevada, 2013-07-13, F, (KP825632, MG050841, MG050842, MG050843, MG050844, MG050845), KP826021, KP826196, KP826406, (MG050867, MG050868), KY229047. MUB 49505: Sierra Nevada, 2013-07-13, F, KP825631, KP826020, KP826195, KP826405, KP826534, KY229031. MUB 49518: Sierra Nevada, 2013-07-13, F, (KP825630, MG050846, MG050847, MG050848, MG050849, MG050850, MG050851, MG050852, MG050853, MG050854), (KP826019, MG050759), KP826194, KP826404, KP826533, KY229038. MUB 49528: Sierra Nevada, 2013-07-13, F, (KP825629, MG050855, MG050856, MG050857, MG050858, MG050859, MG050860), (KP826018, MG050760), KP826193, (MG050777, MG050778), KP826532, KY229027. MUB 49538: Spanish central mountain ranges, 2011-10-27, F, KP825762, KP826150, KP826192, KP826528, KP826659, KY229021. MUB 49540: Spanish central mountain ranges, 2011-10-27, F, KP825760, KP826148, KP826191, KP826526, KP826657, –. MUB 49541: Spanish central mountain ranges, 2011-10-27, F, KP825759, KP826147, KP826190, KP826525, KP826656, MG050785. MUB 49542: Spanish central mountain ranges, 2011-10-27, F, KP825758, KP826146, KP826188, KP826524, KP826655, –. MUB 49545: Spanish central mountain ranges, 2011-10-27, KP825755, KP826143, KP826186, KP826521, KP826652, KY229029. MUB 49550: Spanish central mountain ranges, 2011-10-27, F*, KP825750, KP826138, KP826179, KP826516, KP826647, MG050786. MUB 49552: Spanish central mountain ranges, 2011-10-27, F*, KP825748, KP826136, KP826177, KP826514, KP826645, MG050787. MUB 49553: Spanish central mountain ranges, 2011-10-27, F*, KP825747, KP826135, KP826176, KP826513, KP826644, –. MUB 49554: Spanish central mountain ranges, 2011-10-27, F*, KP825746, KP826134, KP826175, KP826512, KP826643, KY229017. MUB 49555: Spanish central mountain ranges, 2011-10-27, KP825745, KP826133, KP826174, KP826511, KP826642, –. MUB 49557: Spanish central mountain ranges, 2011-10-27, F, KP825743, KP826131, KP826173, KP826509, KP826640, MG050788. MUB 49558: Spanish central mountain ranges, 2011-10-28, F, KP825742, KP826130, KP826172, KP826508, KP826639, –. MUB 49560: Spanish central mountain ranges, 2011-10-28, F*, KP825740, KP826128, KP826170, KP826506, KP826637, KY229013. MUB 49562: Spanish central mountain ranges, 2011-10-28, KP825738, KP826126, –, KP826504, KP826635, –. MUB 49564: Spanish central mountain ranges, 2011-10-28, F*, KP825736, KP826124, KP826168, KP826502, KP826633, –. MUB 49566: Spanish central mountain ranges, 2011-10-28, F, KP825734, KP826122, KP826167, KP826500, KP826631, KY229044. MUB 49567: Spanish central mountain ranges, 2011-10-28, F*, KP825733, KP826121, KP826166, KP826499, KP826630, KY229003. MUB 49568: Spanish central mountain ranges, 2011-10-29, F, KP825732, –, KP826165, –, KP826629, KY229048. MUB 49569: Spanish central mountain ranges, 2011-10-29, F*, KP825731, KP826119, KP826164, KP826497, KP826628, KY229009. MUB 49570: Sicilian Mount Etna, 2013-09-07, F, KP825714, KP826107, –, KP826478, KP826606, KY229016. MUB 49593: Sicilian Mount Etna, 2013-09-08, F, KP825715, KP826106, KP826163, KP826479, KP826607, KY229034. MUB 49600: Spanish southeastern mountains, 2013-11-15, F*, KP825722, KP826104, KP826159, KP826486, KP826613, KY229022. MUB 49602: Spanish southeastern mountains, 2013-11-15, F, KP825723, KP826105, KP826160, KP826487, KP826614, KY229050.

Continued

APPENDIX 1. Continued.

Other mountainous systems

CBFS 6159: Alps, 1997-08-14, KP825712, KP826100, –, KX503294, –, –. CBFS 6162: Alps, 1997-08-14, F, KP825711, KP826099, KP826154, KP826483, KP826611, KY229028. CBFS 13557: Alps, 2009-07-23, F, KP825708, KP826096, KP826151, –, KP826608. MUB 49604: Alps, 2012-09-14, F*, KP825627, KP826016, KP826162, KP826401, KP826530, KY229053. MUB 49606: Alps, 2012-09-14, F*, KP825727, KP826115, KP826161, KP826493, KP826624, –. MUB 49613: Alps, 2012-09-16, F*, KP825726, KP826114, –, KP826492, KP826623, KY229051. MUB 49617: Alps, 2012-09-16, F*, KP825725, KP826113, –, KP826491, KP826622, KY229002. MUB 49619: Alps, 2012-09-17, M, KP825724, KP826112, –, KP826490, KP826621, KY229000. MUB 49624: Pyrenees, 2012-08-31, F*, KP825730, KP826118, –, KP826496, KP826627, KY229007. MUB 49629: Pyrenees, 2012-08-31, F*, KP825729, KP826117, KP826158, KP826495, KP826626, KY229055. MUB 49650: Pyrenees, 2012-09-02, F*, KP825728, KP826116, KP826157, KP826494, KP826625, KY229004.

Lowlands

BOL 46302: South Africa, 2002-10-25, F*, KP825717, KP826109, –, KX503295, KP826618, KY229010. BOL 46303: South Africa, 2002-11-16, F*, KP825716, KP826108, –, –, KP826617, –. MUB 49652: Germany, 2011-08-11, F*, KP825718, KP826110, KP826156, KP826488, KP826619, KY229039. MUB 49653: Germany, 2011-02-09, F*, KP825719, KP826111, –, KP826489, KP826620, KY229020. MUB 49654: Czech Republic, 2014-04-11, F*, KX503276, –, KX503286, KX503291, KX503306, KY229012. MUB 49655: Czech Republic, 2014-04-11, F*, KX503275, –, KX503288, KX503290, KX503305, KY228999. MUB 49659: Czech Republic, 2014-04-12, F*, KX503274, –, KX503287, KX503289, KX503304, KY229006. MUB 52185: United Kingdom, 2014-12-10, KX503277, KX503282, KX503284, KX503292, KX503307, KY229049. MUB 52186: United Kingdom, 2014-04-03, (MG050861, MG050862, MG050863, MG050864, MG050865, MG050866), KX503283, KX503285, KX503293, KX503308, KY229054. S B201182: Sweden, 1985-07-16, F*, KP825721, KP826103, –, KX503296, KP826616, KY229018. S B201183: Sweden, 1985-07-21, F*, KP825720, KP826102, –, KP826485, KP826615, KY229014.

Outgroup: *Cheilothela chloropus*

MUB 52416: Sierra Nevada, 2011-04-28, KX503273, KX503281, –, KX503299, KX503303, KY229025. MUB 52417: Sierra Nevada, 2011-04-28, –, KX503280, –, KX503298, KX503302, –. MUB 52418: Sierra Nevada, 2011-04-28, –, KX503279, –, KX503297, KX503301, KY229026. MUB 52419: Sierra Nevada, 2011-04-28, –, KX503278, –, –, KX503300, –.