

Pleistocene Speciation in North American Lichenized Fungi and the Impact of Alternative Species Circumscriptions and Rates of Molecular Evolution on Divergence Estimates

Steven D. Leavitt^{1,2,3*}, H. Thorsten Lumbsch², Soili Stenroos⁴, Larry L. St. Clair³

1 Committee on Evolutionary Biology, University of Chicago, Chicago, Illinois, United States of America, **2** Science & Education, The Field Museum, Chicago, Illinois, United States of America, **3** Department of Biology and M. L. Bean Life Science Museum, Brigham Young University, Provo, Utah, United States of America, **4** Botanical Museum, Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland

Abstract

Pleistocene climatic fluctuations influenced patterns of genetic variation and promoted speciation across a wide range of species groups. Lichens are commonly found in habitats that were directly impacted by glacial cycles; however, the role of Pleistocene climate in driving speciation in most lichen symbionts remains unclear. This uncertainty is due in part to limitations in our ability to accurately recognize independently evolving lichen-forming fungal lineages and a lack of relevant fossil calibrations. Using a coalescent-based species tree approach, we estimated divergence times for two sister clades in the genus *Xanthoparmelia* (Parmeliaceae) restricted to western North America. We assessed the influence of two different species circumscription scenarios and various locus-specific rates of molecular evolution on divergence estimates. Species circumscriptions were validated using the program BP&P. Although speciation was generally supported in both scenarios, divergence times differed between traditional species circumscriptions and those based on genetic data, with more recent estimates resulting from the former. Similarly, rates of evolution for different loci resulted in variable divergence time estimates. However, our results unambiguously indicate that diversification in the sampled *Xanthoparmelia* clades occurred during the Pleistocene. Our study highlights the potential impact of ambiguous species circumscriptions and uncertain rates of molecular evolution on estimating divergence times within a multilocus species tree framework.

Citation: Leavitt SD, Lumbsch HT, Stenroos S, Clair LLS (2013) Pleistocene Speciation in North American Lichenized Fungi and the Impact of Alternative Species Circumscriptions and Rates of Molecular Evolution on Divergence Estimates. PLoS ONE 8(12): e85240. doi:10.1371/journal.pone.0085240

Editor: Norman Johnson, University of Massachusetts, United States of America

Received: August 14, 2013; **Accepted:** November 24, 2013; **Published:** December 26, 2013

Copyright: © 2013 Leavitt et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This project was supported in part by grants from the California Lichen Society (<http://californialichens.org/>), the Brigham Young University Office of Research and Creative Activities (<http://orca.byu.edu/>), and NSF (<http://nsf.gov/>; "Hidden diversity in parmelioid lichens", DEB-0949147). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: sleavitt@fieldmuseum.org

Introduction

Improved species recognition has important implications for understanding diversity, ecological and biogeographical patterns, factors promoting diversification, and conservation [1–5]. While utilizing data from independent sources (ex. behavior, ecology, genetic, morphology, etc.) can facilitate the identification of species boundaries [2,6,7], integrating diverse data into phylogenetic and population genetic frameworks remains challenging [7]. However, modern molecular techniques allow researchers to effectively and quickly generate a large number of unlinked genetic markers from a large number of individuals (ex [8]). This information, coupled with the advent and maturation of algorithms for estimating species trees using multilocus sequence data, now provides

unprecedented insights into species delimitation and the processes of speciation [1,9,10].

The recent union of coalescent theory with phylogenetics allows gene and species trees to be estimated simultaneously [9,11–13]. Multispecies coalescent-based species tree methods facilitate an effective approach for integrating multiple genetic loci from multiple individuals per species by modeling the coalescent gene histories embedded in a shared species tree [12,14,15]; ultimately, providing a substantial improvement over single gene trees or concatenated multilocus phylogenies [11,12,14–20]. Furthermore, coalescent-based species tree methods can accurately depict relationships even in cases where incomplete lineage sorting and gene tree heterogeneity (e.g. branch length and topology) obscures phylogenetic relationships among species [13,21,22]. In addition to

improvements in estimating relationships in speciation histories, species tree methods also provide a more biologically realistic framework for temporal estimates of speciation events [23]. Recent modifications to the program BEAST [24] allow for joint estimation of divergence times and species trees [12]. Divergence dates estimated from concatenated gene trees can lead to overestimates of divergence times because gene divergence necessarily predates species divergence [23,25-27]. Furthermore, dating speciation events using priors on substitutions rates in a full probabilistic coalescent species-tree framework is especially relevant for studies investigating taxonomic groups with poor fossil records [23].

Most currently available species tree methods require that species/population assignments are made *a priori* [12,13,17,28]. However, correct assignments of samples to species may be difficult in a number of scenarios, including cryptic speciation [5,29], incongruence between morphology-based species and molecular data [30,31], or simply the fact that accurate specimen identification is a nontrivial task, particularly in complex groups with subtle or difficult to discern diagnostic characters. Inaccurate specimen identifications and equivocal, or contradictory, taxonomy may dramatically impact species tree and divergence time estimates where *a priori* species/population assignments are required. Objective assessments of alternative species delimitation hypotheses will likely result in improved perspectives on speciation histories and increased taxonomic stability [6]. Coalescent-based species delimitation methods provide a means for testing alternative hypotheses of evolutionary independence using multilocus genetic sequence data [6,32,33], and several approaches have recently been developed [13,28,32,34]. The program Bayesian Phylogenetics and Phylogeography (BP&P) delimits species using Bayesian model selection to calculate posterior probabilities of different species delimitation models using a reversible-jump Markov chain Monte Carlo (rjMCMC) [32]. This approach has been shown to outperform other coalescent-based species delimitation methods in a number of scenarios [33].

Lichens are obligate symbiotic associations consisting of a filamentous fungus and a photosynthetic partner (eukaryotic alga and/or cyanobacterium). Although the lichenized condition has been ecologically and evolutionarily successful [35-37], processes driving diversification in these symbiotic systems are poorly understood. Due to a poor fossil record and uncertainties in the interpretation of the few known lichen fossils [38-41], placement of diversification events within a temporal context remains challenging.

Assessing diversification and biogeographical patterns in lichenized fungi is further confounded by uncertainties in species boundaries. In a number of cases, traditional phenotype-based species boundaries fail to accurately characterize diversity (reviewed in 42-44). For example, in some lichen-forming genera such as *Melanelixia* and *Melanohalea* (Parmeliaceae), diagnosable phenotypic differences appear to be absent even millions of years after the initial divergence, and morphologically similar species may not be each other's closest relatives [45,46]. In contrast, other species of lichenized fungi have been shown to harbor a high

degree of intraspecific morphological and chemical variation (ex [47-49].).

The lichen-forming fungal genus *Xanthoparmelia* (Vain.) Hale (Parmeliaceae) provides an excellent study system with which to assess the impact of equivocal species delimitation scenarios and locus-specific evolutionary rates for estimating divergence times in a coalescent-based species tree framework. *Xanthoparmelia* is among the most speciose genera of lichen-forming fungi, including ca. 820 species [50]. A number of studies have provided novel insights into species relationships as well as biogeography and character evolution within the genus [48,51-57]. Although fossil evidence for *Xanthoparmelia* has not yet been found, taxon-specific (either family- or genus-level) substitution rates have been estimated for four loci – nuclear ribosomal ITS, nuclear ribosomal LSU, RPB1, and mitochondrial SSU [25,45,58]. In Parmeliaceae, rate- and fossil-calibrated phylogenies have resulted in similar estimates of divergence times for a number of deeper lineages within the family [25,45,56,58,59], supporting the use of rate-calibrated divergence estimates.

Large regions in western North America were subjected to global cooling, aridification, and major glaciation events during the Pleistocene [60-62]. These climatic shifts had a substantial impact on both vascular plant and lichen communities [25,63-67]. Additionally, Pleistocene climatic fluctuations appear to have played a major role in speciation for at least some lichen-forming fungi in western North America [64]. In contrast to Pleistocene-dominated speciation, a number of recent studies suggest that the Neogene was a major period of diversification for extant lineages in the most speciose lichen-forming fungal family, Parmeliaceae [25,45,56,58,59,68]. Based on limited sampling, species-level diversification coinciding with Pleistocene glacial cycles appears to be restricted to a limited number of clades within some parmelioid genera [25,68].

To better understand the impact of Pleistocene climate change on speciation in lichenized fungi, we assessed the timing of diversification in two *Xanthoparmelia* clades that are commonly distributed across a wide range of montane habitats, including, shrub-steppe, subalpine, and alpine communities in western North America (clades 'D' and 'E' in [48,52]). The speciation histories of these *Xanthoparmelia* clades remain unsettled due to ambiguities in species boundaries and incomplete lineage sorting [48,52]. Both clades include phenotypically polymorphic forms, representing a total of 15 traditionally circumscribed species [48,52]. In contrast to the 15 phenotype-based species recovered in these two clades, multilocus sequence data support a total of four polymorphic species-level genetic clusters. The majority of the traditionally circumscribed species recovered in clades 'D' and 'E' are also found in other distinct lineages [48]. However, formal taxonomic changes have not yet been made.

The general lineage concept [69-71] provides a framework for species delimitation using a variety of operational criteria, data sets, and analytical methods to circumscribe segments of separately evolving metapopulation lineages. In this study, we estimated speciation probabilities for both traditional, phenotype-based species and molecular-based genetic groups

using the coalescent-based species delimitation method BP&P [32]. We then compared divergence times using rate calibrated species trees, based on the two different species delimitation hypotheses, and evaluate the impact of substitution rates for different loci. The impact of equivocal species circumscriptions on divergence times has not been evaluated within a coalescent framework, and our study provides a valuable perspective on the impact of species delimitations and rates of molecular evolution for estimating divergence times within a species tree framework. In addition, our results offer insights into the speciation history of *Xanthoparmelia* species in western North America.

Materials and Methods

Sampling

We sampled a total of 180 individuals (Table S1) from two previously recognized sister clades, 'clade D' and 'clade E', which appear to be restricted to western North America [48,52]. Specimens were obtained across a variety of habitats using both opportunistic field sampling and herbarium collections. Lichens collections were made on public lands managed by the Bureau of Land Management and USDA Forest Service, and sampling permits were not required. Traditional specimen identifications were based on vegetative morphological characters and diagnostic medullary extrolites identified using thin layer chromatography (TLC; [72,73]), following established protocol [74,75]. 'Clade D' was shown to be dominated by saxicolous, isidiate morphs (*X. aff. dierythra*, *X. aff. mexicana*, *X. aff. plittii*, and *X. aff. subplittii*), but also included a limited number of non-isidiate specimens (*X. aff. lineola*) [48]. 'Clade E' included eight non-isidiate species (*X. aff. californica* [saxicolous], *X. aff. chlorochroa* [vagrant], *X. aff. coloradoënsis* [saxicolous], *X. aff. cumberlandia* [saxicolous], *X. lipochlorochroa* Hale & Elix [vagrant], *X. aff. neochlorochroa* [vagrant], *X. aff. norchlorochroa* [vagrant], *X. aff. vagans* [vagrant], and *X. wyomingica* (Gyelnik) Hale [semi-attached]), distributed across three distinct genetic clusters ('E1', 'E2', and 'E3'; [52]). *Xanthoparmelia lipochlorochroa* and *X. wyomingica* were represented by collections made from the type localities. None of the sampled species are protected by United States law.

DNA extraction, amplification and sequencing

DNA was extracted from all individuals in connection with previous studies [48,52], and we generated DNA sequence data for a total of nine markers. Four nuclear ribosomal loci were sampled, including the large-subunit (nrLSU, c. 844 bp), the internal transcribed spacer region (ITS, c. 548 bp), a group I intron (c. 398 bp), and the intergenic spacer region (IGS, c. 380 bp). In addition, we sampled five protein-coding loci, β -tubulin (787 bp), glyceraldehyde 3-phosphate dehydrogenase (GAPDH, 598 bp), MCM7 (541 bp), RPB1 (807 bp), and RPB2 (823 bp). Amplification of the nrLSU, IGS, ITS, group I intron, β -tubulin, and MCM7 was described previously [52]. For this study, we amplified three additional protein-coding markers, the GAPDH, RPB1, and RPB2. The GAPDH fragment was amplified using the following primers Gpd1-LM with Gpd2-LM

[76]; the RPB1 fragment was amplified using gRPB1-A [77] and fRPB1-C [78]; and primers fRPB2-5f with fRPB2-7cr were used to amplify the RPB2 fragment [79]. PCR amplifications were performed using Ready-To-Go PCR Beads (GE Healthcare, Pittsburgh, PA, USA), and cycling parameters followed a 55–50 °C touchdown reaction [80]. PCR products were visualized on 1% agarose gel and cleaned using ExoSAP-IT (USB, Cleveland, OH, USA). We sequenced complementary strands using the same primers used for amplifications, and sequencing reactions were performed using BigDye 3.1 (Applied Biosystems, Foster City, CA, USA). Products were run on an ABI 3730 automated sequencer (Applied Biosystems) at the Pritzker Laboratory for Molecular Systematics and Evolution at the Field Museum, Chicago, IL, USA.

Alignment

We assembled and edited new sequences using Sequencher 4.10 (Gene Codes Corporation, Ann Arbor, MI). Sequences were aligned using the program MAFFT 6 [81,82]. We implemented the G-INS-i alignment algorithm and '1PAM / K=2' scoring matrix, with an offset value of 0.9, and the remaining parameters were set to default values for the protein-coding (β -tubulin, GAPDH, MCM7, RPB1 and RPB2) and nrLSU markers. For the ribosomal ITS, IGS, and group I intron loci, we used the same parameters, with the exception of an offset value set to 0.1 rather than 0.9.

Population structure analyses

Previous studies have suggested that traditional morphology/chemistry-based taxonomy may misrepresent *Xanthoparmelia* diversity in western North America [48,51]. Therefore, we used multilocus Bayesian population assignment tests implemented in the programs STRUCTURE 2.3.2 [83,84] and BAPS 5 [85] to identify genetic groups as an alternative to traditional phenotype-based species. Varying approaches have been implemented to convert DNA sequence data to allelic data for Bayesian clustering (see 86). STRUCTURE is expected to perform well when there is sufficient independence across regions such that linkage disequilibrium within regions does not dominate the data (STRUCTURE manual), but can also be effective using multilocus sequence data and treating all SNPs as independent loci regardless of physical linkage within each locus (eg. 84–87). In order to assess the impact of using linked genotyped SNP data from ribosomal markers, we compared three genotyped data sets created from the aligned sequence data. The first comprised SNP data exclusively from ribosomal tandem repeat (nrLSU, ITS, IGS, group I intron); the second comprised SNPs from putatively unlinked protein-coding loci (β -tubulin, GAPDH, MCM7, RPB1, and RPB2); and the final genotype dataset comprised SNPs from all sampled loci [87–89]. STRUCTURE analyses were run for clades 'D' and 'E' separately. In the STRUCTURE analyses, we ran 10 replicate runs for each *K* value, from 1–6, using 50,000 MCMC generations for the burn-in period and an additional 50,000 generations to estimate the posterior distribution. We used an admixture model that incorporated the possibility for individuals to have mixed cluster ancestry, including an 'F model' to account for correlated allele frequencies among populations

resulting from migration or shared ancestry [84]. We used the program STRUCTURE HARVESTER [90] to process the results from the 10 replicate runs for each K value. The ΔK procedure outlined by [91] was then used to estimate the most likely number of clusters within the sample. However, we also considered other values of K that seemed biologically sensible (see STRUCTURE manual). We classified individuals with posterior probabilities < 0.70 to any cluster as “admixed”. We also used Bayesian clustering in BAPS [85] to examine the number of groups. BAPS uses a stochastic optimization procedure to identify the number of groups and is capable of handling sequence data by utilizing a linkage model [85]. In contrast to STRUCTURE, BAPS includes K as a parameter to be estimated, and the best partition of the data into K clusters is identified as the one with the highest marginal log likelihood. We selected the clustering with linked loci analysis and replicated the analysis ten times to ensure the stability of our results. Additionally, we ran BAPS with the number of clusters fixed at $K = 5$ (see STRUCTURE results below) and then replicated multiple times to ensure stability.

Phylogenetic analyses

We used the program RAxML 7.2.8 [92,93] to reconstruct maximum likelihood (ML) gene trees from each of the individual gene alignments. A search combining 200 separate ML searches was conducted, implementing the ‘GTRGAMMA’ model, and 1000 pseudoreplicates to evaluate bootstrap support for each node. Relationships were also estimated from the concatenated nine-locus data matrix using a total-evidence approach [94]. We conducted an ML analysis of the concatenated data set using locus-specific model partitions in RAxML, and all loci were treated as separate partitions. Search parameters and assessment of nodal support were performed as described above. In order to assess differences between ribosomal and protein-coding loci for tree reconstruction, we conducted separate ML analyses of the complete ribosomal dataset (ITS, nrLSU, IGS, and group I intron) and the 5-gene protein-coding dataset (β -tubulin, GAPDH, MCM7, RPB1, and RPB2). In both cases we used locus-specific model partitions, and all loci were treated as separate partitions. Search parameters and assessment of nodal support were performed as described above.

Species trees and divergence dating

Estimating a species tree using concatenated multilocus sequence data may be misleading under certain divergence scenarios [19,95,96]. In addition, divergence times are consistently overestimated in gene tree approaches, including concatenated loci, due to the fact that the time of gene divergence necessarily predates speciation events [23,26]. Therefore, we used the coalescent-based hierarchical Bayesian model *BEAST implemented in BEAST 1.7.4 to estimate species trees and divergence times [12]. *BEAST incorporates the coalescent process and the uncertainty associated with gene trees and nucleotide substitution model parameters and estimates the species tree directly from the sequence data [12]. Species assignments are required *a priori* for *BEAST analyses, and assignment errors can mislead

species tree inference and result in strong support for an incorrect species tree [96]. Therefore, we explored two different scenarios for assigning individuals to a ‘species’ group. First, we chose to explore the relationships among all potential lineages inferred using molecular sequence data. We expected strong nodal support for well-differentiated lineages in concatenated gene tree and/or coalescent-based species tree analyses, in addition to high speciation probabilities separating distinct species-level lineages [32]. We used results from the STRUCTURE and ML analyses to assigned individuals to one of seven populations. For ‘clade E’, we treated each of the genetic clusters from the $K = 5$ model as distinct species because the STRUCTURE analyses indicate a general pattern of a plateau near $K = 5$ (see Results). Individuals with inferred admixed ancestry were excluded. Although the STRUCTURE analysis did not indicate distinct populations within ‘clade D’, the ML analysis of the concatenated data matrix recovered two well-supported clades (see Results), and previous studies have shown that clustering algorithms may fail to recover genetically distinct groups represented by unbalanced sample sizes [97,98]. Therefore, individuals from ‘clade D’ were assigned to two separate groups based on the results of the ML analysis. In the second scenario, we assigned all individuals to a species group based on traditional morphological characters. We included a total of 11 species, *X. aff. chlorochroa* (54 specimens), *X. aff. coloradoensis* (34), *X. aff. cumberlandia* (43), *X. aff. lineola* (8), *X. lipochlorochroa* (3), *X. aff. mexicana* (17), *X. aff. neochlorochroa* (4), *X. aff. norchlorochroa* (2), *X. aff. plittii* (6), *X. aff. vagans* (4), and *X. wyomingica* (2). Based on previous results [48] and subsequent morphological re-examinations, specimens representing *X. neowyomingica* were combined with *X. cumberlandia*; and *X. subplittii* was merged with *X. plittii*. *Xanthoparmelia californica* and *X. dierythra* were each represented by a single specimen and therefore excluded from these analyses due to small sample sizes (see 21,99,100).

We estimated divergence using *BEAST because species tree methods provide a more biologically realistic framework for dating divergence events by directly modelling genetic divergence that pre-dates speciation [23]. In the absence of relevant fossil evidence, we assessed four different scenarios using molecular locus-specific rates of evolution for Parmeliaceae. First, we used 2.43×10^{-9} substitution/site/year (s/s/y) for the ITS marker, recently reported for the parmelioid lichen-forming genus *Melanelixia* [45]. This rate is similar to other estimates of ITS substitution rates for lichen-forming mycobionts (2.38×10^{-9} s/s/y *Oropogon*, Parmeliaceae, Lecanorales [68]; and a non-lichenized fungus (2.52×10^{-9} s/s/y, Erysiphales; [101]). In the second scenario, we used a rate estimated for Parmeliaceae (*Protoparmelia* excluded) for the nrLSU (0.70×10^{-9} s/s/y, [58]). In the third scenario we used substitution rates for both the ITS (2.38×10^{-9} s/s/y) and nrLSU (0.70×10^{-9} s/s/y) markers. In the final scenario we used the RPB1 substitution rate (1.51×10^{-9} s/s/y) estimated for *Xanthoparmelia* [58]. In all cases, substitution rates for other loci were co-estimated along the run under a uniform prior (0 – 50), relative to the specified molecular evolution rates. Locus-specific models of evolution were selected using jModeltest 2

[102], and ultrametric trees were constructed using the most similar substitution models available in BEAST. Divergence times were estimated using an uncorrelated relaxed lognormal molecular clock used for all loci [103]. We also implemented a Yule process and gamma-distributed population sizes for the species tree prior and a piecewise linear population size model with a constant root. Default values were used for the remaining priors. For each scenario, we ran two independent Markov chain Monte Carlo (MCMC) analyses for a total of 200 million generations, sampling every 2000 steps. The first 25% of sampled trees were excluded as burn-in. We assessed convergence by examining the likelihood plots through time using Tracer 1.5 [104] and compared summarized tree topologies from separate runs. The posterior probabilities of nodes were computed from the sampled trees (excluding burn-in samples) using TreeAnnotator 1.7.4 [104].

Speciation probabilities

We explored the validity of the two species delimitation scenarios. The first based on the seven groups inferred using molecular sequence data and the second treating traditionally circumscribed species as distinct species. The marginal posterior probability of speciation for each scenario was estimated using the program BP&P 2.1 [16,32]. This method accommodates the species phylogeny as well as lineage sorting due to ancestral polymorphisms. BP&P has recently been shown to outperform other coalescent-based species delimitation methods in cases of recent speciation events, with robust performance using a modest number of genetic markers [33]. We used the species trees inferred from the *BEAST analyses of the traditional species and genetic clusters as the guide trees and clades 'D' and 'E' were analyzed separately. In cases where relationships were not strongly supported in the coalescent-based species tree, we used different guide trees representing alternative topologies and assessed speciation probabilities among the topologies. The prior distributions on ancestral population size (θ) and root age (τ_0) were assigned gamma distributions of $G(1,100)$ and $G(2, 2000)$, respectively, assuming intermediate ancestral population sizes and relatively shallow divergences among species [5,64]. We used algorithm 0 and each species delimitation model was assigned equal prior probability. Each reversible-jump Monte Carlo (rjMCMC) analysis was run 500,000 generations with a burn-in of 50,000. Each analysis was run at least twice to confirm consistency between runs.

Results

Molecular data

The concatenated nine-gene data matrix consisted of 5726 aligned nucleotide positions (Table 1; TreeBase # 14929). Overall, 601 sites were variable, 273 in the ribosomal markers and 328 in the five protein-coding loci (Table 1).

Bayesian clustering

The STRUCTURE analyses revealed little evidence in support of population substructure within 'clade D'. Population

Table 1. Genetic variability of sampled markers used in this study, including the total number of samples, N (number of haplotypes); aligned length; variable and parsimony-informative (PI) sites for each sampled locus; and model of evolution selected using the program jModeltest.

Locus	N (haplotypes)	No. variable sites		Model
		Aligned bp	(PI sites)	
ITS	179 (76)	548	96 (61)	TrNef+G
nrLSU	176 (43)	844	42 (19)	TrN+I+G
IGS	164 (61)	380	68 (34)	K80+G
Intron	157 (61)	398	67 (41)	K80+G
β -tubulin	168 (42)	787	67 (33)	TrNef+G
GAPDH	155 (55)	598	56 (44)	SYM+I+G
MCM7	163 (74)	541	80 (45)	K80+I+G
RPB1	178 (35)	807	45 (30)	TrNef +I
RPB2	178 (77)	823	80 (59)	TPM1+I+G
All ribosomal loci	180 (125)	2170	273 (155)	NA
All protein-coding loci	180 (154)	3556	328 (211)	NA
All loci	180 (167)	5726	601 (366)	NA

doi: 10.1371/journal.pone.0085240.t001

clusters inferred for $K > 1$ did not correspond to morphological/chemical groups or geographic regions and tended to include a high number of individuals with inferred admixed ancestry (see Figure S1). The ΔK method indicated that a $K = 2$ model best fits the data; however, the same method does not identify the best K if $K = 1$. Furthermore, clustering using BAPS also yielded a single population cluster for 'clade D' (see Figure S1).

For 'clade E', a comparison of the three genotyped data sets – SNPs from the ribosomal tandem repeat, SNPs from the protein-coding markers, and SNPs from all sampled loci – are shown in Figure 1. Both BAPS and the ΔK method, based on results from the STRUCTURE analysis, indicated that a $K = 2$ model best fits the data (see Fig. S2). The median ML values of the Bayesian clustering analyses based on the dataset comprised of SNPs from all sampled loci with estimates of K from 1–6 are reported in Figure S2. These analyses indicate a general pattern of a plateau near $K = 5$, and we chose also to examine the phenotypic expressions and geographic distributions of population clusters within the $K = 5$ model because individual genetic clusters were generally associated with distinct phenotypic or geographic patterns. Exploratory analyses of $K > 5$ revealed a larger proportion of individuals with admixed ancestry, relative to smaller K values, and did not seem biologically sensible (data not shown). The distribution of individual population cluster membership assignments for both the $K = 2$ and $K = 5$ models are shown in Figure 1. In the $K = 5$ model, all individuals assigned to cluster 'X-1A' produced stictic acid (*X. aff. cumberlandia*). In contrast, individuals assigned to cluster 'X-1B' produced salazinic acid (*X. aff. coloradoensis*). The majority of individuals assigned membership in clusters 'X-2A', 'X-2B', and 'X-2C' also produced salazinic acid. However, a small proportion of individuals in cluster 'X-2A' produced stictic acid rather than salazinic acid, including both saxicolous (*X. aff. cumberlandia*) and vagrant morphs (*X. aff.*

vagans) (Table 2). Cluster 'X-2B' was comprised exclusively of vagrant morphs collected from southwestern Wyoming and western Montana, and included vagrant species *X. aff. chlorochroa* (salazinic acid chemosyndrome), *X. lipochlorochroa* (fatty acid chemosyndrome) and *X. aff. neochlorochroa* (norstictic acid chemosyndrome). Cluster 'X-2C' was also comprised exclusively of vagrant morphs, *X. aff. chlorochroa*, although these originated from southern Utah and a single site in Colorado, USA. Approximately 20% of the individuals assigned to cluster 'E-2' in the $K = 2$ model, were recovered with admixed ancestry in the $K = 5$ model.

ML topologies

Neither traditional, phenotype-based species nor genetic population clusters were recovered as monophyletic in any of the individual gene topologies (Figures S3 A–I). The ML phylogeny estimated from the concatenated 9-locus phylogeny is shown in Figure 2 (see Figure S4 for complete labels). The 5-gene protein-coding phylogeny was more similar to the total evidence phylogeny than the 4-gene ribosomal phylogeny (see Figures S4, S5 & S6). The total evidence phylogeny revealed two well-supported lineages within 'clade D' (Figure 2). Three non-isidiate individuals producing salazinic acid were recovered within clade 'D1-A', while clade 'D1-B' was dominated chemically by polymorphic isidiate specimens (Figure S5). Two of the five population clusters from the $K = 5$ model, 'E-1a' and 'E-2b', were not recovered as monophyletic in the 9-gene phylogeny (Figure 2).

Coalescent-based species trees and divergence time estimates

In the *BEAST analyses ESS values were > 150 for all parameters and the vast majority were well above 200. Topologies and divergence estimates were congruent across independent runs. A maximum clade credibility chronogram from the multi-locus species tree analyses based on the seven-species model inferred from the STRUCTURE and ML analyses is shown in Figure 3A. Similar to the ML analysis of the 9-locus data matrix, the *BEAST analysis supported a split between the two lineages within 'clade D'. While within 'clade E', a sister relationship was recovered for clusters 'E-1a' and 'E-1b' with weak statistical support (posterior probability [PP] = 0.68). Clusters 'E-2a', 'E-2b', and 'E-2c' were recovered within a well-supported (PP = 1.0) monophyletic lineage, corresponding to 'cluster 2' from the $K = 2$ STRUCTURE analysis. However, relationships among clusters within this clade were weakly supported (PP < 0.50).

A maximum clade credibility chronogram from the multi-locus species tree analyses based on the traditional, phenotype-based species circumscriptions is shown in Figure 3B. Relationships among the three species within 'clade D' were unresolved (PP < 0.50). In clade 'E' a well-supported lineage (PP = 0.95) was recovered and included *X. aff. chlorochroa*, *X. aff. coloradensis*, *X. lipochlorochroa*, *X. aff. neochlorochroa*, and *X. aff. norchlorochroa*. This clade generally corresponds to clusters 'E-2a', 'E-2b', and 'E-2c' from the $K = 5$ STRUCTURE analysis. However, relationships between this clade and *X. aff. cumberlandia* and *X. aff. wyomingica* were unresolved (Figure

3B). The vagrant species *X. aff. chlorochroa*, *X. lipochlorochroa*, *X. aff. neochlorochroa*, and *X. aff. norchlorochroa* were recovered as a well-supported (PP = 1.0) monophyletic clade, sister to *X. aff. vagans* (Figure 3B) with weak statistical support (PP = 0.67).

For the species tree with species-lineages inferred using sequence data, estimates of divergence times using the four different rate-specific scenarios (fixed rates for the ITS, nrLSU, RPB1, and ITS with nrLSU) are reported in Figure 3A. The substitution rates of the nine sampled loci for each of the four rate-scenarios using the genetic population clusters, estimated under a relaxed clock, are reported in Table 2. In analyses using traditional phenotype-based species, the substitution rates of the nine sampled loci estimated for the two markers resulted in the highest and lowest rates, ITS and RPB1 respectively, are reported in Table 3. In all scenarios, divergence events were estimated to have occurred during the Pleistocene (Figures 3A, B). However, estimates varied widely depending on the specific fixed substitution rate, with the fixed ITS rate resulting in the oldest estimates and the rate for the RPB1 marker yielding the most recent estimates (Figure 3A). Divergence estimates were shown to differ by between 0.6 and 1.6 million years, depending on the specified substitution rate (Figure 3A). For example, the divergence times estimated using the fixed ITS rate support the separation of clades 'D' and 'E' ca. 2.3 Ma (95% HPD = 1.7–3.0), while this split was estimated to have occurred ca. 0.7 Ma (95% HPD = 0.4–1.0) using the fixed rate for the RPB1 locus (Figure 3A). Divergence estimates indicate that in both clades 'D' and 'E' the majority of the diversification leading to extant species, including the vagrant species, occurred between 0.3 and 1.9 Ma.

For the morphology-based species tree, divergence times from the two analyses that resulted in the oldest and youngest estimates, the ITS and RPB1 rates, respectively, are shown in Figure 3B. Similar to divergence times estimated in the genetic cluster-based gene tree, divergence events were also estimated to have occurred during the Pleistocene in the morphology-based species tree. However, age estimates in the morphology-based species tree were more recent than those inferred from the genetic cluster-based gene tree analysis. In the morphology-based tree using the fixed rate for the ITS marker, the initial radiation of species within 'clade D' was estimated to have occurred ca. 0.4 Ma (95% HPD = 0.2–0.7) (Figure 3B), but ca. 1.9 Ma (95% HPD = 1.3–2.5) in the cluster-based tree (Figure 3A). Similarly, the initial radiation of 'clade E' was estimated at ca. 0.8 Ma (95% HPD = 0.5–1.4) in the morphology-based species tree (Figure 3B), but 1.5 Ma (95% HPD = 1.0–2.0) in the cluster-based tree (Figure 3A).

Speciation probabilities

Speciation probabilities based on the BP&P analyses are shown in Figures 3A, B. Within 'clade D', traditional phenotype based species (based largely on diagnostic extrolites) were not supported in the BP&P analyses (Figure 3B). In contrast, most putative species groups in 'clade E' received strong support in both the morphology-based and genetic cluster species delimitation scenarios, with the exception of the vagrant species *X. aff. chlorochroa*, *X. lipochlorochroa*, *X. aff.*

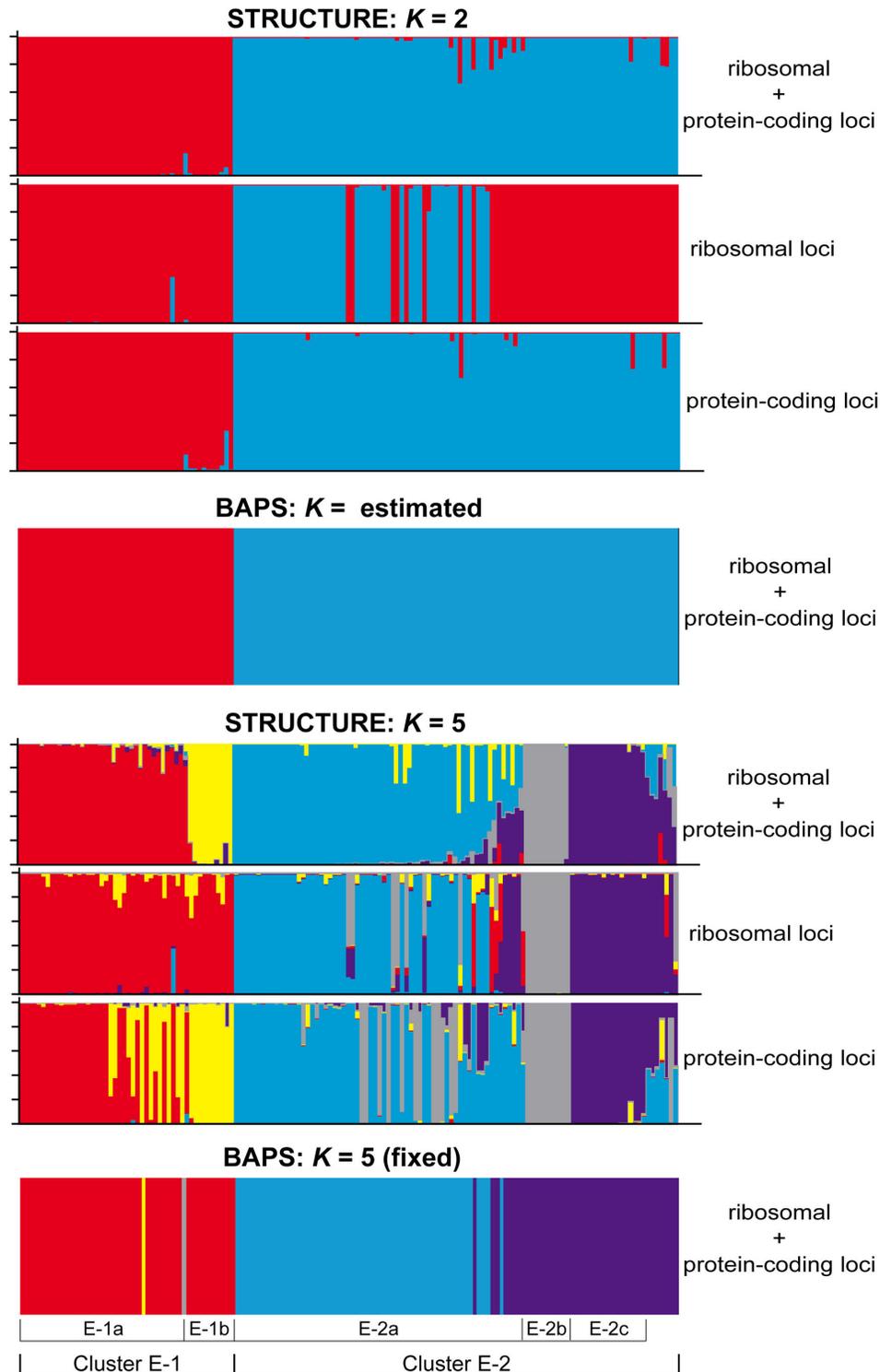


Figure 1. Population assignments to genetic clusters in *Xanthoparmelia* 'clade E'. Population membership was inferred using the Program BAPS and STRUCTURE using SNP data from nine sampled loci. (nrLSU, IGS, ITS, group I intron, β -tubulin, GAPDH, MCM7, RPB1, and RPB2), ribosomal loci only (nrLSU, IGS, ITS, group I intron) and protein-coding loci only (β -tubulin, GAPDH, MCM7, RPB1, and RPB2). STRUCUTRE analyses include estimates for $K = 2$ and $K = 5$ models; and the BAPS analyses included the estimated number of clusters ($K = 2$) and the number of clusters fixed at $K = 5$. Each accession is shown by a thin vertical line that is partitioned into colored segments representing the proportion of each individual's genome assigned to a genetic cluster.

doi: 10.1371/journal.pone.0085240.g001

Table 2. Estimated substitution rates for western North American *Xanthoparmelia* clades 'D' and 'E' from the *BEAST analysis based on population clusters inferred from SNP data from a nine-locus data set.

	nrLSU	ITS	RPB1	ITS/LSU
IGS	3.38 (HPD = 1.92 - 5.02)	1.95 (HPD = 1.27 - 2.68)	7.14 (HPD = 4.06 - 10.54)	2.10 (HPD = 1.43 - 2.83)
Intron	3.93 (HPD = 2.18 - 5.86)	2.28 (HPD = 1.45 - 3.16)	8.29 (HPD = 4.73 - 12.39)	2.44 (HPD = 1.61 - 3.31)
nrLSU	0.70 (fixed)	0.57 (HPD = 0.34 - 0.83)	2.07 (HPD = 1.09 - 3.18)	0.69 (fixed)
ITS	4.86 (HPD = 2.82 - 7.13)	2.42 (fixed = 2.43)	10.27 (HPD = 6.00 - 15.10)	2.43 (fixed)
β -tubulin	1.66 (HPD = 0.94 - 2.48)	0.96 (HPD = 0.62 - 1.33)	3.50 (HPD = 1.99 - 5.18)	1.03 (HPD = 0.69 - 1.40)
GAPDH	2.54 (HPD = 1.38 - 3.86)	1.48 (HPD = 0.92 - 2.13)	5.43 (HPD = 2.97 - 8.26)	1.60 (HPD = 1.02 - 2.24)
MCM7	4.61 (HPD = 2.55 - 7.02)	2.72 (HPD = 1.69 - 3.85)	9.87 (HPD = 5.54 - 14.73)	2.81 (HPD = 1.84 - 3.88)
RPB1	1.00 (HPD = 0.54 - 1.53)	0.58 (HPD = 0.35 - 0.83)	1.56 (fixed = 1.51)	0.62 (HPD = 0.39 - 0.87)
RPB2	2.45 (HPD = 1.38 - 3.65)	1.42 (HPD = 0.90 - 1.96)	5.17 (HPD = 2.97 - 7.68)	1.52 (HPD = 1.02 - 2.06)

Rates were estimated under an uncorrelated lognormal relaxed molecular clock using fixed substitution rates for four different scenarios (fixed rates shown in parentheses): nrLSU, ITS, RPB1, and ITS/nrLSU markers. Units: substitution/site/ 10^9 years.

doi: 10.1371/journal.pone.0085240.t002

neochlorochroa, and *X. aff. norchlorochroa* which were collapsed as a single species (Figure 3). Comparing alternative guide trees representing alternative relationships at weakly supported nodes resulted in similar levels of support at those nodes (results not shown).

Discussion

In this study, we compared two alternative, though somewhat overlapping, species delimitation scenarios in two clades in the lichenized fungal genus *Xanthoparmelia* from western North America. In the first scenario we used Bayesian clustering and phylogenetic inference to identify putative lineages based on molecular sequence data; and the second scenario was based on traditional morphology/chemistry-based species circumscriptions (Table 4). We found that neither genetic clustering nor Bayesian species delimitation supported traditional phenotype-based species in a clade dominated by isidiate morphs ('clade D'; Figures 1 and 3B). In contrast, the Bayesian species delimitation method implemented in this study generally provided strong support for both the phenotype- and allelic-based species circumscriptions in 'clade E'. However, divergence estimates differed between the two approaches for delimiting species, demonstrating that differences in phenotype- and molecular-based species circumscriptions may have important implications for estimating divergence times within a coalescent-based framework (Figure

3). Further confounding our ability to assess the timing of species diversification in *Xanthoparmelia*, we found that rates of molecular evolution for different loci, estimated from closely related taxa, resulted in substantial differences in divergence times (Figure 3A,B). In spite of these challenges, our results indicate that the Pleistocene was an important period of diversification for *Xanthoparmelia* in western North America, and this study provides additional insight into species boundaries for this group.

Equivocal species boundaries and the impact on divergence estimates

Despite increasing amounts of data and improved perspectives on species boundaries, both species discovery and specimen identification remains challenging [1,105]. In general, lichens display relatively few taxonomically useful characters, and studies have repeatedly shown that our current interpretation of morphological and chemical characters often fails to accurately characterize species diversity [42]. Our study of *Xanthoparmelia* lineages in western North America empirically demonstrated that equivocal taxonomy could potentially have a significant impact on species tree and divergence time estimates where *a priori* species assignments are required.

We assessed the validity of both phenotype- and genetic-based species circumscriptions in two *Xanthoparmelia* clades by calculating speciation probabilities using the Bayesian species delimitation program BP&P [32]. Although coalescent-based species delimitation methods are likely a step towards improved objectivity and comparability [6], our empirical study of species boundaries in *Xanthoparmelia* indicates that selection between competing hypotheses of species boundaries remains challenging. Below we explore several reasons which may explain the overall support of species in the differing species delimitation scenarios assessed here, particularly in 'clade E'.

First, there was some overlap between the phenotype-based species circumscriptions and the genetic clusters (Table 4). For example, the vast majority of specimens identified as *X. aff. cumberlandia* were recovered within genetic cluster 'E-1a', although a limited number of specimens identified as *X. aff. cumberlandia* were also assigned membership in cluster 'E-2a'. Similarly, a large number of vagrant morphs were recovered in two distinct genetic clusters, 'E-2b' and 'E-2c' (see Figures 1 & S4), supporting the inference of genetic differentiation of vagrants species from attached saxicolous morphs in the morphology-based scenario. Ideally, both genetic and phenotypical data should support similar hypotheses of species boundaries [7]; however, in practice different data sets and operational criteria may give conflicting or ambiguous results [106-109]. Our empirical study of *Xanthoparmelia* species supports the potential for strong speciation probabilities in differing species delimitation scenarios in cases where species hypotheses based on phenotypical characters and genetic data overlap to some degree.

Second, in the phenotype-based species delimitation scenario *X. aff. norchlorochroa* and *X. wyomingica* were represented by only two individuals. Previous studies suggest



Figure 2. Maximum likelihood topology of *Xanthoparmelia* clades 'D' and 'E'. Phylogeny estimated from ribosomal (nrLSU, IGS, ITS, and a group I intron) and protein-coding (β -tubulin, GAPDH, MCM7, RPB1, and RPB2) markers representing *Xanthoparmelia* clades 'D' and 'E'. Thickened branches indicate ML bootstrap values $\geq 70\%$, and colored branches correspond to population genetic clusters from the $K=5$ model in the STRUCTURE analyses (see Figure 1).

doi: 10.1371/journal.pone.0085240.g002

that this sampling density (both in number of individuals and loci) may be adequate to accurately infer speciation probabilities, even in cases of recent divergence [33,100]. However, *X. wyomingica* sensu lato (s.l.) often occurs in intermediate morphs, integrating between nearly completely

vagrant forms – similar to *X. chlorochroa* s.l. – to loosely attached forms – similar to *X. coloradoensis* s.l. [74,110]. Our study included only two characteristic specimens collected from a single population near the type locality in the Bighorn Mountains, Wyoming, USA. Including additional *X. wyomingica*

Table 4. Provisional taxonomy for montane *Xanthoparmelia* clades 'D' and 'E' in western North America.

Provisional taxon	Bayesian clustering (see Phylogeny (see		Traditional species	General comments/notes
	Figure 1)	Figure 2)		
' <i>X. aff. lineola</i> '	'D'	'Clade D-1'	<i>X. aff. lineola</i>	Non-isidiate forms expressing the salazinic acid chemosyndrome, known from three specimens collected in montane habitats in Utah (Duchesne & Washington Counties), USA.
' <i>X. isidiomontana</i> '*	'D'	'Clade D-2'	<i>X. aff. dierythra</i> , <i>X. aff. lineola</i> , <i>X. aff. mexicana</i> , and <i>X. aff. plitti</i>)	Lineage dominated by the chemically polymorphic (norstictic, salazinic, and stictic acid chemosyndromes), isidiate forms in continental habitats in the Intermountain Western USA, east to the western Dakotas, USA.
' <i>X. subcumberlandia</i> '	'E-1a'	'Cluster E-1a'	<i>X. aff. cumberlandia</i>	Non-isidiate forms expressing the stictic acid chemosyndrome, from continental habitats in the Intermountain Western USA.
' <i>X. utahensis</i> '*	'E-1b'	'Clade E-1b'	<i>X. aff. coloradoensis</i>	Non-isidiate forms expressing the salazinic acid chemosyndrome, from continental habitats in the Intermountain Western USA.
' <i>X. wyomingica</i> '	'E-2a', 'E-2b', 'E-2c'	'Clade E-2'	<i>X. aff. chlorochroa</i> , <i>X. aff. coloradoensis</i> , <i>X. aff. cumberlandia</i> , <i>X. aff. lipochlorochroa</i> , <i>X. aff. neochlorochroa</i> , <i>X. aff. norchlorochroa</i> , <i>X. aff. vagans</i> , and <i>X. aff. wyomingica</i> ,	Morphologically variable forms, ranging from adnate saxicolous forms to semi-attached and completely vagrant terricolous forms. Saxicolous forms generally express the salazinic acid chemosyndrome, but vagrant forms are chemically polymorphic, producing salazinic, stictic, norstictic, and fatty acid chemosyndromes.

* new putative species

doi: 10.1371/journal.pone.0085240.t004

Table 3. Estimated substitution rates for western North American *Xanthoparmelia* clades 'D' and 'E' from the *BEAST analysis using traditional, phenotype-based species circumscriptions.

	ITS	RPB1
IGS	2.11 (HPD = 1.39 - 2.94)	7.30 (HPD = 4.19 - 10.66)
Intron	2.27 (HPD = 1.48 - 3.17)	7.85 (HPD = 4.53 - 11.62)
nrLSU	0.61 (HPD = 0.36 - 0.88)	2.08 (HPD = 1.10 - 3.16)
ITS	2.39 (fixed 2.43)	9.51 (HPD = 5.60 - 13.75)
BT	1.04 (HPD = 0.66 - 1.44)	3.56 (HPD = 2.03 - 5.25)
GAPDH	1.58 (HPD = 0.97 - 2.27)	5.29 (HPD = 2.96 - 8.03)
MCM7	2.86 (HPD = 1.77 - 4.04)	9.61 (HPD = 5.50 - 14.16)
RPB1	0.60 (HPD = 0.36 - 0.86)	1.55 (fixed = 1.51)
RPB2	1.45 (HPD = 0.93 - 2.02)	4.96 (HPD = 2.89 - 7.37)

Rates were estimated under an uncorrelated lognormal relaxed molecular clock using fixed substitution rates for the two analyses that resulted in the oldest and youngest estimates, the ITS and RPB1 rates, respectively (fixed rates shown in parentheses). Units: substitution/site/10⁹ years.

doi: 10.1371/journal.pone.0085240.t003

s.l. specimens from other populations, including those with deviating morphologies, could potentially have a substantial impact on our inference of species boundaries.

A third reason that may have contributed to the general support for competing species delimitation hypotheses is the exclusion of individuals with evidence of admixed ancestry. Although some degree of gene flow between incipient species may be a common phenomenon [13,111], coalescent-based approaches do not account for this source of gene-tree discordance [21,96,112]. In order to mitigate potentially

confounding effects of gene flow on species tree inference, sampling schemes tend to avoid including individuals from phylogeographic lineage boundaries or individuals with evidence of admixed ancestry [48,96]. While eliminating individuals affected by introgression may improve accuracy of species tree reconstructions, the removal of individuals showing evidence of an admixed genome likely has a significant impact on speciation probabilities by inflating the perceived amount of genetic divergence among populations.

The exclusion of specimens with admixed ancestries in BP&P is largely a practical issue when individual population assignments are equivocal or if there is evidence of hybridization between two distinct populations (ex. hybrid boundary, intermediate morphologies, etc.), thereby violating the assumption of complete cessation of gene flow following species divergence [32,100]. As expected, the accuracy of BP&P has been shown to decrease with increasing levels of interspecific introgression, although the program still performs relatively well using a modest number of loci (4-10) and individuals per species, despite low levels of gene flow [21,100]. BP&P has been shown to be indecisive at intermediate levels of migration and tends to lump populations into a single species at higher migration rates (> 10 migrants per generation) [100].

The exclusion of individuals with ambiguous population membership in groups where population boundaries are not clear likely has a substantial impact on estimating speciation probabilities. In our study of montane *Xanthoparmelia* populations in western North America, ca. 20% of the individuals within 'cluster E-2' were inferred with admixed ancestry in the STRUCTURE analyses (see Figure 1) and were excluded from the BP&P analysis. The admixed individuals included characteristic forms of *X. aff. chlorochroa*, *X. aff.*

coloradoënsis, and *X. wyomingica*, and none of these specimens displayed any phenotypical evidence of hybridization. The specimens with inferred admixed ancestry may simply be an artifact of our limited molecular sampling or other methodological limitations, rather than actually representing truly admixed individuals. Ultimately, the treatment of individuals with inferred admixed ancestry must be carefully considered on a case-by-case basis.

Potential violations of assumptions may also limit the practical application of speciation probabilities estimated in BP&P [5,32,100]. BP&P requires the input of a user-specified guide tree and misspecification of the guide tree can result in strong support for models containing more species by artificially increasing divergence among populations [5]. In our study of *Xanthoparmelia*, we compared different guide trees representing alternative topologies in cases where relationships were not strongly supported in the coalescent-based species tree and found similar speciation probabilities in all cases (data not shown). Therefore, we assert that the high speciation probabilities observed in this study are likely not the result of a mis-specified guide tree. Prior distributions for ancestral population sizes (θ) and the root age (τ_0) may also have a significant impact on species delimitation results [5,32]. In many cases, including *Xanthoparmelia*, little is known of ancestral population sizes and root age, requires a largely subjective specification for θ and τ_0 . In general, a combination of large θ and small τ_0 provide a more conservative estimate supporting fewer species, although species delimitation results appear to be robust with changes to τ_0 [5,32]. For species delimitation in *Xanthoparmelia* we assumed intermediate ancestral population sizes and relatively shallow divergences among species (see Methods), but speciation probabilities were similar under a more conservative combination of priors, 0.1 and 0.001 for θ and τ_0 , respectively (data not shown).

Given general support for both species delimitation scenarios, we assessed divergence estimates for both hypotheses within a coalescent-based multispecies species tree approach. Our results revealed some differences in estimates of divergence times. Divergence time estimates of the initial split between clades 'E' and 'D' were consistent for both morphology- and genetic-based species trees, depending on the specified rate calibration (Figures 3A, B). However, differences in divergence time estimates between morphology- and molecular-based species trees become apparent in clades 'D' and 'E', with non-overlapping HPD intervals for the initial radiation of 'clade D' (Figures 3A, B). Estimates for the initial radiation of 'clade E' differed by up to 0.7 million years, 1.5 Ma in the population cluster-based species tree and 0.8 Ma in the phenotype-based species tree, although HPD intervals overlapped between 1.0 - 1.3 Ma (Figures 3A, B). The potential for discrepancy among divergence estimates from ambiguous species delimitation hypotheses has important implications for inferring specific factors driving diversification in western North American lichens, particularly in more recent diversification scenarios such as we see in western North American *Xanthoparmelia* species.

In this study, we used a number of genetic markers that have commonly been implemented to assess relationships among

closely related species, and intraspecific evolutionary processes [67,80,113-115]. However, morphology and genetic-based *Xanthoparmelia* lineages showed strong evidence of incomplete lineage sorting across all sampled loci. Qualitatively, the RPB1 gene tree (Figure S3) was most similar to the 9-gene total evidence phylogeny and the 5-gene protein-coding phylogeny (Figures S4, S6). This finding is consistent with other studies highlighting the utility of the RPB1 for accurately inferring relationships at multiple scales [78,116,117]. In contrast, relationships in the four-gene ribosomal phylogeny generally lacked support and did not correspond to either morphology- or genetic-based lineages (Figure S5). Overall, our results highlight the pressing need to develop additional markers in order to resolve species relationships in montane *Xanthoparmelia* in western North America with greater confidence.

The impact of substitution rates on divergence estimates

Our results show that using mutation rates estimated for different loci may result in more substantial differences in divergence estimates than those inferred using alternative species delimitation scenarios in the sampled *Xanthoparmelia* lineages (Figure 3). In the absence of relevant fossil evidence for *Xanthoparmelia*, we assessed divergence times using Parmeliaceae-specific mutation rates from three loci, ITS, nrLSU, and RPB1 [45,58]. Divergence estimates using substitution rates for the ITS, nrLSU, and the combination of ITS/nrLSU loci resulted in at least partially overlapping highest posterior density (HPD) regions, but differed significantly from estimates using the substitution rate for the RPB1 marker (Figure 3). Previous studies have shown relatively consistent substitution rates (ca. 1.5 substitution/site/year $\times 10^{-9}$) for a number of protein coding loci, including MCM7, RPB1, RPB2, across multiple genera in Parmeliaceae [25,45]. However, estimated substitution rates for protein-coding loci used in this study varied widely, with the RPB1 marker showing markedly slower mutation rates, when compared to the MCM7 and RPB2 markers (Tables 2 & 3). Amo de Paz et al. [58] also showed that substitution rates for the RPB1 marker varied widely among major lineages in Parmeliaceae. In contrast, estimates of ITS substitution rates for a range of lichen-forming mycobionts and a non-lichenized fungus were relatively similar [25,45,68,101]. Additional research will be required to better understand differences in evolutionary rates across loci, lineages, and different temporal scales in order to identify loci that may provide more reliable rate-calibrated divergence estimates [118-120]. However, in the absence of fossil calibration points, an exploration of all reasonable substitution rates from different loci will be important in showing a more complete range of potential divergence times.

Pleistocene-driven speciation in *Xanthoparmelia* in western North America

Our study indicates that Pleistocene glacial cycles played a major role in *Xanthoparmelia* diversification in montane habitats in western North America. Although alternative species hypotheses and rates of molecular evolution for different loci

resulted in varying and occasionally non-overlapping divergence times (Figure 3), diversification events are consistently estimated to have occurred during the Pleistocene in the two *Xanthoparmelia* clades investigated here (clades 'D' and 'E'; see 48. Recent studies have also revealed previously unrecognized species-level divergence coinciding with Pleistocene climatic fluctuations in the *Rhizoplaca melanophthalma* species-complex (Lecanoraceae) [64] and two species complexes in the genus *Melanohalea* (Parmeliaceae), *M. elegantula* s.l. and *M. subolivacea* s.l. [25]. Interestingly, *Rhizoplaca melanophthalma* s.l. and *Melanohalea subolivacea* s.l. often co-occur to varying degrees with the sampled *Xanthoparmelia* lineages across a broad range of habitats in western North America, ranging from pinyon-juniper woodlands into subalpine and even alpine habitats. The shared temporal diversification pattern in these lichen-forming fungal groups suggests that speciation may have been driven by similar factors during the Pleistocene glaciations in western North America.

Subtle differences in microhabitats have been shown to have a strong impact on the occurrence and distribution of some lichens [121-124], and ecological processes may be central to the formation of new species through ecology-based divergent selection [125,126]. However, there is little evidence of ecological divergence among the closely related *Xanthoparmelia* lineages found in both clades 'D' and 'E'. Most species-level lineages were found across a broad range of habitats and were commonly found occurring in sympatry at local scales [127]. Similarly, distinct lineages within the *Rhizoplaca melanophthalma* and *Melanohalea subolivacea* species complexes showed no apparent preference for distinct environmental conditions [25,64,88]. Ultimately, empirically assessing environmental parameters and modeling ecological niches may provide important insights into the specific processes driving speciation in these groups.

The biological roles of the different medullary compounds remain uncertain [128-130]. Medullary extrolites have traditionally played a major role in *Xanthoparmelia* taxonomy [74]. However, our results indicate that differences in secondary chemistry in isidiate morphs in 'clade D' and vagrant morphs in 'clade E' do not correspond with distinct evolutionary lineages (Table 4). In contrast, saxicolous forms in 'clade E' producing stictic acid are supported by molecular data as a lineage distinct from morphologically similar forms producing salazinic acid (Table 4). Differences in morphology and chemistry across substrata and habitats have been documented for *Xanthoparmelia* populations in the Sonoran Desert [131], and additional research may provide novel insights into the role of distinct extrolites on the fitness of *Xanthoparmelia* species.

Although species boundaries in the sampled *Xanthoparmelia* clades remain somewhat equivocal, these data provide valuable insights for improving our perspective concerning the species boundaries of montane *Xanthoparmelia* in western North America. A number of recent studies have demonstrated that some morphological and chemical characters traditionally used to separate species likely represent intraspecific polymorphisms rather than taxonomically diagnostic traits

[47,49,132,133]. Our population clustering analysis supports a broader phenotypic circumscription for the *Xanthoparmelia* lineages sampled here. Based on the results from the Bayesian clustering and phylogenetic analyses and general patterns in morphological and/or chemical characters, we propose an alternative hypothesis of species boundaries within the sample clades and recommend a number of provisional names (Table 4). A complete taxonomic revision of species within clades 'D' and 'E' will be treated in detail in a companion paper (Leavitt et al. *in prep*).

Supporting Information

Figure S1. Population assignments to genetic clusters in *Xanthoparmelia* 'clade D'. Population membership was inferred using the Program BAPS and STRUCTURE using SNP data from nine sampled loci (nrLSU, IGS, ITS, group I intron, β -tubulin, GAPDH, MCM7, RPB1, and RPB2). STRUCTURE analyses include estimates for $K = 2$ and $K = 4$ models; and the BAPS analyses included the estimated number of clusters ($K = 1$). Each accession is shown by a thin vertical line that is partitioned into colored segments representing the proportion of each individual's genome assigned to a genetic cluster. (PDF)

Figure S2. Evaluation of the STRUCTURE results from the *Xanthoparmelia* 'clade E'. SNP data from nine sampled loci (nrLSU, IGS, ITS, group I intron, β -tubulin, GAPDH, MCM7, RPB1, and RPB2), ribosomal loci only (nrLSU, IGS, ITS, group I intron) and protein-coding loci only (β -tubulin, GAPDH, MCM7, RPB1, and RPB2) were included. (A) Likelihood values estimated from ten independent runs from each K from 1 - 6. (B) Results of the ΔK method for inferring the number of distinct population clusters. (PDF)

Figure S3. Maximum likelihood gene trees for *Xanthoparmelia* clades 'D' and 'E'. Topologies were estimated from each of the nine sampled loci (nrLSU, IGS, ITS, group I intron, β -tubulin, GAPDH, MCM7, RPB1, and RPB2), with bootstrap support indicated at nodes. Colors in 'clade E' correspond to genetic clusters inferred from the $K = 5$ model; colors in 'clade D' correspond to two well-supported clades recovered in the concatenated multilocus phylogeny (see Figure S4). Tip labels include the individual STRUCTURE assignment for the $K = 2$ ($K = 1$ for 'clade D'), $K = 5$ ($K = 4$ for 'clade D'), morphological identification, and specimen number, in that order. (PDF)

Figure S4. Total evidence ML tree estimated from the concatenated nine loci data matrix. Loci included: nrLSU, IGS, ITS, group I intron, β -tubulin, GAPDH, MCM7, RPB1, and RPB2; and bootstrap support indicated at nodes. Colors in 'clade E' correspond to genetic clusters inferred from the $K = 5$ model; colors in 'clade D' correspond to two well-supported clades recovered in this total evidence phylogeny.

(PDF)

Figure S5. ML tree estimated from the concatenated ribosomal loci. Loci include: nrLSU, IGS, ITS, and group I intron; and bootstrap support indicated at nodes. Colors in 'clade E' correspond to genetic clusters inferred from the $K = 5$ model; colors in 'clade D' correspond to two well-supported clades recovered in the concatenated multilocus phylogeny (see Figure S4).

(PDF)

Figure S6. ML tree estimated from the concatenated protein-coding loci. Loci include: β -tubulin, GAPDH, MCM7, RPB1, and RPB2; and bootstrap support indicated at nodes. Colors in 'clade E' correspond to genetic clusters inferred from the $K = 5$ model; colors in 'clade D' correspond to two well-supported clades recovered in the concatenated multilocus phylogeny (see Figure S4).

(PDF)

Table S1. Collection information for all *Xanthoparmelia* specimens included in the present study : ID, individual code; morphological/chemical species identification; inferred population cluster; Brigham Young

References

- Camargo A, Sites JW (2013) Species Delimitation: A Decade After the Renaissance. In: I Pavlinov. The Species Problem - Ongoing Issues: InTech.
- Sites J, Marshall J (2003) Delimiting species: a renaissance issue in systematic biology. *Trends Ecol Evol* 18: 462–470. doi:10.1016/S0169-5347(03)00184-8.
- Wiens JJ (2007) Species delimitation: new approaches for discovering diversity. *Syst Biol* 56: 875–878. doi:10.1080/10635150701748506. PubMed: 18027280.
- Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R et al. (2007) Cryptic species as a window on diversity and conservation. *Trends Ecol Evol* 22: 148–155. doi:10.1016/j.tree.2006.11.004. PubMed: 17129636.
- Leaché AD, Fujita MK (2010) Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *Proc Biol Sci* 277: 3071–3077. doi:10.1098/rspb.2010.0662. PubMed: 20519219.
- Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C (2012) Coalescent-based species delimitation in an integrative taxonomy. *Trends Ecol Evol* 9: 480–488.
- Yeates DK, Seago A, Nelson L, Cameron SL, Joseph LEO et al. (2011) Integrative taxonomy, or iterative taxonomy? *Syst Entomol* 36: 209–217. doi:10.1111/j.1365-3113.2010.00558.x.
- Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S et al. (2013) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Mol Ecol* 22: 787–798. doi:10.1111/mec.12023. PubMed: 23057853.
- Edwards SV (2009) Is a new and general theory of molecular systematics emerging? *Evolution* 63: 1–19. doi:10.1111/j.1558-5646.2008.00549.x. PubMed: 19146594.
- Knowles LL (2009) Estimating species trees: methods of phylogenetic analysis when there is incongruence across genes. *Syst Biol* 58: 463–467. doi:10.1093/sysbio/syp061. PubMed: 20525600.
- Liu L, Yu L, Kubatko L, Pearl DK, Edwards SV (2009) Coalescent methods for estimating phylogenetic trees. *Mol Phylogenet Evol* 53: 320–328. doi:10.1016/j.ympev.2009.05.033. PubMed: 19501178.
- Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Mol Biol Evol* 27: 570–580. doi:10.1093/molbev/msp274. PubMed: 19906793.
- Knowles LL, Carstens BC (2007) Delimiting species without monophyletic gene trees. *Syst Biol* 56: 887–895. doi:10.1080/10635150701701091. PubMed: 18027282.
- Edwards SV, Liu L, Pearl DK (2007) High-resolution species trees without concatenation. *Proc Natl Acad Sci U S A* 104: 5936–5941. doi:10.1073/pnas.0607004104. PubMed: 17392434.
- Liu L, Pearl DK (2007) Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Syst Biol* 56: 504–514. doi:10.1080/10635150701429982. PubMed: 17562474.
- Rannala B, Yang Z (2003) Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* 164: 1645–1656. PubMed: 12930768.
- Liu L (2008) BEST: Bayesian estimation of species trees under the coalescent model. *Bioinformatics* 24: 2542–2543. doi:10.1093/bioinformatics/btn484. PubMed: 18799483.
- Kubatko LS, Carstens BC, Knowles LL (2009) STEM: species tree estimation using maximum likelihood for gene trees under coalescence. *Bioinformatics* 25: 971–973. doi:10.1093/bioinformatics/btp079. PubMed: 19211573.
- Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol Evol* 24: 332–340. doi:10.1016/j.tree.2009.01.009. PubMed: 19307040.
- Weisrock DW, Smith SD, Chan LM, Biebow K, Kappeler PM et al. (2012) Concatenation and concordance in the reconstruction of mouse lemur phylogeny: an empirical demonstration of the effect of allele sampling in phylogenetics. *Mol Biol Evol* 29: 1615–1630. doi:10.1093/molbev/mss008. PubMed: 22319174.
- Camargo A, Avila LJ, Morando M, Sites JW (2012) Accuracy and precision of species trees: effects of locus, individual, and base-pair sampling on inference of species trees in lizards of the *Liolaemus darwini* group (Squamata, Liolaemidae). *Syst Biol* 61: 272–288. doi:10.1093/sysbio/syr105. PubMed: 22076301.
- Weisrock DW, Rasoloarison RM, Fiorentino I, Ralison JM, Goodman SM et al. (2010) Delimiting species without nuclear monophyly in Madagascar's mouse lemurs. *PLOS ONE* 5: e9883. doi:10.1371/journal.pone.0009883. PubMed: 20360988.
- McCormack JE, Heled J, Delaney KS, Peterson AT, Knowles LL (2011) Calibrating divergence times on species trees versus gene trees: Implications for speciation history of *Aphelocoma* jays. *Evolution* 65: 184–202. doi:10.1111/j.1558-5646.2010.01097.x. PubMed: 20681982.
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7: 214. doi:10.1186/1471-2148-7-214. PubMed: 17996036.

University Herbarium of Non-vascular Cryptogams (BRY) voucher accession number; major acid, diagnostic secondary chemistry; Location; Lat., latitude; Lon., longitude; Ele., altitude in meters a.s.l.; collector(s), and GenBank accession numbers. (XLSX)

Acknowledgements

We thank Trevor Goward, Matthew Nelsen, Roger Rosentreter, Jack Sites, and Leigh Johnson for invaluable conceptual help and insightful comments. We are also indebted to various colleagues for providing material and field assistance, notably C. Björk, A. DeBolt, M. DeVito, M. felix, R. Fuller, T. Hardle, J. Hollinger, the Howell families, the Leavitt family, B. McCune, M. Robinson, and R. Rosentreter. We gratefully acknowledge anonymous reviewers for their constructive feedback.

Author Contributions

Conceived and designed the experiments: SL HTL LS SS. Performed the experiments: SL. Analyzed the data: SL. Contributed reagents/materials/analysis tools: SL HTL LS. Wrote the manuscript: SL HTL SS LS.

25. Leavitt SD, Esslinger TL, Divakar PK, Lumbsch HT (2012) Miocene and Pliocene dominated diversification of the lichen-forming fungal genus *Melanohalea* (Parmeliaceae, Ascomycota) and Pleistocene population expansions. *BMC Evol Biol* 12: 176. doi:10.1186/1471-2148-12-176. PubMed: 22963132.
26. Edwards SV, Beerli P (2000) Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* 54: 1839–1854. Available online at: doi:10.1554/0014-3820(2000)054[1839:PGDPDA]2.0.CO;2. PubMed: 11209764
27. Brandley MC, Wang Y, Guo X, Montes de Oca AN, Feria-Ortiz M et al. (2011) Accommodating heterogeneous rates of evolution in molecular divergence dating methods: an example using intercontinental dispersal of *Plestiodon* (Eumeces) lizards. *Syst Biol* 60: 3–15. doi:10.1093/sysbio/syr036. PubMed: 20952756.
28. Ence DD, Carstens BC (2011) SpedeSTEM: a rapid and accurate method for species delimitation. *Mol Ecol Resour* 11: 473–480. doi:10.1111/j.1755-0998.2010.02947.x. PubMed: 21481205.
29. Niemiller ML, Near TJ, Fitzpatrick BM (2012) Delimiting species using multilocus data: diagnosing cryptic diversity in the southern cavefish, *Typhlichthys subterraneus* (Teleostei: Amblyopsidae). *Evolution* 66: 846–866. doi:10.1111/j.1558-5646.2011.01480.x. PubMed: 22380444.
30. Ross KG, Gotzek D, Ascunce MS, Shoemaker DD (2010) Species delimitation: a case study in a problematic ant taxon. *Syst Biol* 59: 162–184. doi:10.1093/sysbio/syp089. PubMed: 20525628.
31. Rivas Plata E, Lumbsch HT (2011) Parallel evolution and phenotypic divergence in lichenized fungi: A case study in the lichen-forming fungal family Graphidaceae (Ascomycota: Lecanoromycetes: Ostropales). *Mol Phylogenet Evol* 61: 45–63. doi:10.1016/j.ympev.2011.04.025. PubMed: 21605691.
32. Yang Z, Rannala B (2010) Bayesian species delimitation using multilocus sequence data. *Proc Natl Acad Sci U S A* 107: 9264–9269. doi:10.1073/pnas.0913022107. PubMed: 20439743.
33. Camargo A, Morando M, Avila LJ, Sites JW (2012) Species delimitation with ABC and other coalescent-based methods: A test of accuracy with simulations and an empirical example with lizards of the *Liolaemus darwini* complex (Squamata: Liolaemidae). *Evolution* 66: 2834–2849. doi:10.1111/j.1558-5646.2012.01640.x. PubMed: 22946806.
34. O'Meara BC (2010) New heuristic methods for joint species delimitation and species tree inference. *Syst Biol* 59: 59–73. doi:10.1093/sysbio/syp077. PubMed: 20525620.
35. Honegger R (1991) Functional aspects of the lichen symbiosis. *Annu Rev Plant Physiol Plant Mol Biol* 42: 553–578. doi:10.1146/annurev.pp.42.060191.003005.
36. Gargas A, DePriest PT, Grube M, Tehler A (1995) Multiple origins of lichen symbioses in fungi suggested by SSU rDNA phylogeny. *Science* 268: 1492–1495. doi:10.1126/science.7770775. PubMed: 7770775.
37. Lutzoni F, Pagel M, Reeb V (2001) Major fungal lineages are derived from lichen symbiotic ancestors. *Nature* 411: 937–940. doi:10.1038/35082053. PubMed: 11418855.
38. Taylor JW, Berbee ML (2006) Dating divergences in the Fungal Tree of Life: review and new analyses. *Mycologia* 98: 838–849. doi:10.3852/mycologia.98.6.838. PubMed: 17486961.
39. Lücking R, Huhndorf S, Pfister DH, Plata ER, Lumbsch HT (2009) Fungi evolved right on track. *Mycologia* 101: 810–822. doi:10.3852/09-016. PubMed: 19927746.
40. Berbee ML, Taylor JW (2010) Dating the molecular clock in fungi – how close are we? *Fungal. Biol Rev* 24: 1–16.
41. Prieto M, Wedin M (2013) Dating the diversification of the major lineages of Ascomycota (Fungi). *PLOS ONE* 8: e65576. doi:10.1371/journal.pone.0065576. PubMed: 23799026.
42. Lumbsch HT, Leavitt SD (2011) Goodbye morphology? A paradigm shift in the delimitation of species in lichenized fungi. *Fungal Divers* 50: 59–72. doi:10.1007/s13225-011-0123-z.
43. Crespo A, Lumbsch HT (2010) Cryptic species in lichen-forming fungi. *IMA Fungus* 1: 167–170. doi:10.5598/imafungus.2010.01.02.09. PubMed: 22679576.
44. Crespo A, Pérez-Ortega S (2009) Cryptic species and species pairs in lichens: A discussion on the relationship between molecular phylogenies and morphological characters. *Anales del Jardín Botánico de Madrid* 66: 71–81. doi:10.3989/ajbm.2225.
45. Leavitt SD, Esslinger TL, Divakar PK, Lumbsch HT (2012) Miocene divergence, phenotypically cryptic lineages, and contrasting distribution patterns in common lichen-forming fungi (Ascomycota: Parmeliaceae). *Biol J Linn Soc* 1007: 920–937.
46. Leavitt SD, Esslinger TL, Spribille T, Divakar PK, Lumbsch HT (2013) Multilocus phylogeny of the lichen-forming fungal genus *Melanohalea* (Parmeliaceae, Ascomycota): Insights on diversity, distributions, and a comparison of species tree and concatenated topologies. *Mol Phylogenet Evol* 66: 138–152. doi:10.1016/j.ympev.2012.09.013. PubMed: 23017822.
47. Pérez-Ortega S, Fernández-Mendoza F, Raggio J, Vivas M, Ascaso C et al. (2012) Extreme phenotypic variation in *Cetraria aculeata* (lichenized Ascomycota): adaptation or incidental modification? *Ann Bot* 109: 1133–1148. doi:10.1093/aob/mcs042. PubMed: 22451601.
48. Leavitt SD, Johnson LA, Goward T, St Clair LL (2011) Species delimitation in taxonomically difficult lichen-forming fungi: An example from morphologically and chemically diverse *Xanthoparmelia* (Parmeliaceae) in North America. *Mol Phylogenet Evol* 60: 317–332. doi:10.1016/j.ympev.2011.05.012. PubMed: 21627994.
49. Pino-Bodas R, Martín M, Burgaz A (2012) *Cladonia subburgida* and *C. iberica* (Cladoniaceae) form a single, morphologically and chemically polymorphic species. *Mycological Progress* 11: 269–278. doi:10.1007/s11557-011-0746-1.
50. Thell A, Crespo A, Divakar PK, Kärnefelt I, Leavitt SD et al. (2012) A review of the lichen family Parmeliaceae – history, phylogeny and current taxonomy. *Nordic J Bot* 30: 641–664. doi:10.1111/j.1756-1051.2012.00008.x.
51. Leavitt SD, Nelsen MP, St Clair LL (2012) Treading in murky waters: making sense of diversity in *Xanthoparmelia* (Parmeliaceae, Ascomycota) in the Western United States. *Bull California Lichen. Soc* 19: 58–70.
52. Leavitt SD, Johnson L, St Clair LL (2011) Species delimitation and evolution in morphologically and chemically diverse communities of the lichen-forming genus *Xanthoparmelia* (Parmeliaceae, Ascomycota) in western North America. *Am J Bot* 98: 175–188. doi:10.3732/ajb.1000230. PubMed: 21613107.
53. Hodgkinson BP, Lendemer JC (2011) Molecular analyses reveal semi-cryptic species in *Xanthoparmelia tasmanica*. *Bibliotheca Lichenologica* 106: 108–119.
54. Blanco O, Crespo A, Elix JA, Hawksworth DL, Lumbsch HT (2004) A molecular phylogeny and a new classification of parmelioid lichens containing *Xanthoparmelia*-type lichenan (Ascomycota: Lecanorales). *Taxon* 53: 959–975. doi:10.2307/4135563.
55. Amo de Paz G, Lumbsch HT, Cubas P, Elix JA, Crespo A (2010) The genus *Karooia* (Parmeliaceae, Ascomycota) includes unrelated clades nested within *Xanthoparmelia*. *Australian Syst Bot* 23: 173–184. doi:10.1071/SB09055.
56. Amo de Paz G, Cubas P, Crespo A, Elix JA, Lumbsch HT (2012) Transoceanic dispersal and subsequent diversification on separate continents shaped diversity of the *Xanthoparmelia pulla* group (Ascomycota). *PLOS ONE* 7: e39683. doi:10.1371/journal.pone.0039683. PubMed: 22745810.
57. Thell A, Elix JA, Søchting U (2009) *Xanthoparmelia lineola* s. l. Australia and North America. *Bibliotheca Lichenologica* 99: 393–404.
58. Amo de Paz G, Cubas P, Divakar PK, Lumbsch HT, Crespo A (2011) Origin and diversification of major clades in parmelioid lichens (Parmeliaceae, Ascomycota) during the Paleogene inferred by Bayesian analysis. *PLOS ONE* 6: e28161. doi:10.1371/journal.pone.0028161. PubMed: 22174775.
59. Divakar PK, Del-Prado R, Lumbsch HT, Wedin M, Esslinger TL et al. (2012) Diversification of the newly recognized lichen-forming fungal lineage *Montanelia* (Parmeliaceae, Ascomycota) and its relation to key geological and climatic events. *Am J Bot* 99: 2014–2026. doi:10.3732/ajb.1200258. PubMed: 23204485.
60. Van Devender TR, Spaulding WG (1979) Development of vegetation and climate in the Southwestern United States. *Science* 204: 701–710. doi:10.1126/science.204.4394.701. PubMed: 17840214.
61. Graham A (1999) Late Cretaceous and Cenozoic History of North American Vegetation. New York: Oxford University Press.
62. Osborn G, Bevis K (2001) Glaciation in the Great Basin of the Western United States. *Quaternary Sci Rev* 20: 1377–1410. doi:10.1016/S0277-3791(01)00002-6.
63. Coats LL, Cole KL, Mead JI (2008) 50,000 years of vegetation and climate history on the Colorado Plateau, Utah and Arizona, USA. *Quatern Res* 70: 322–338. doi:10.1016/j.yqres.2008.04.006.
64. Leavitt SD, Fernández-Mendoza F, Pérez-Ortega S, Sohrabi M, Divakar PK et al. (2013) Local representation of global diversity in a cosmopolitan lichen-forming fungal species complex (*Rhizoplaca*, Ascomycota). *J Biogeogr* 40: 1792–1806. doi:10.1111/jbi.12118.
65. Thompson RS, Anderson KH (2000) Biomes of western North America at 18,000, 6000 and 0 14C yr bp reconstructed from pollen and packrat midden data. *J Biogeogr* 27: 555–584. doi:10.1046/j.1365-2699.2000.00427.x.
66. Loera I, Sosa V, Ickert-Bond SM (2012) Diversification in North American arid lands: Niche conservatism, divergence and expansion of habitat explain speciation in the genus *Ephedra*. *Mol Phylogenet Evol* 65: 437–450. doi:10.1016/j.ympev.2012.06.025. PubMed: 22776548.

67. Fernández-Mendoza F, Printzen C (2013) Pleistocene expansion of the bipolar lichen *Cetraria aculeata* into the Southern hemisphere. *Mol Ecol* 22: 1961–1983. doi:10.1111/mec.12210. PubMed: 23402222.
68. Leavitt SD, Esslinger TL, Lumbsch HT (2012) Neogene-dominated diversification in neotropical montane lichens: dating divergence events in the lichen-forming fungal genus *Oropogon* (Parmeliaceae). *Am J Bot* 99: 1764–1777. doi:10.3732/ajb.1200146. PubMed: 23092994.
69. de Queiroz K (2007) Species concepts and species delimitation. *Syst Biol* 56: 879–886. doi:10.1080/10635150701701083. PubMed: 18027281.
70. de Queiroz K (1998) The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. Oxford: Oxford University Press.
71. de Queiroz K (1999) The general lineage concept of species and the defining properties of the species category. *Species, New Interdisciplinary Essays*: 49–89.
72. Orange A, James PW, White FJ (2001) Microchemical methods for the identification of lichens. London: British Lichen Society.
73. Culberson CF (1972) Improved conditions and new data for identification of lichen products by standardized thin-layer chromatographic method. *J Chromatogr A* 72: 113–125. doi:10.1016/0021-9673(72)80013-X.
74. Hale ME (1990) A synopsis of the lichen genus *Xanthoparmelia* (Vainio) Hale (Ascomycotina, Parmeliaceae). Washington D.C.: Smithsonian Institution Press.
75. Nash III, Elix JA (2004) *Xanthoparmelia*. In: III Nash BD, Ryan P, Diederich C, Gries F, BUNGARTZ. Lichen Flora of the Greater Sonoran Desert Region. Tempe: Lichens Unlimited.
76. Myllys L, Stenroos S, Thell A (2002) New genes for phylogenetic studies of lichenized fungi: Glyceraldehyde-3-Phosphate Dehydrogenase and Beta-Tubulin genes. *Lichenologist* 34: 237–246. doi:10.1006/lich.2002.0390.
77. Stiller JW, Hall BD (1997) The origin of red algae: Implications for plastid evolution. *Proc Natl Acad Sci of the USA* 94: 4520–4525. doi:10.1073/pnas.94.9.4520. PubMed: 9114022.
78. Matheny PB, Liu YJ, Ammirati JF, Hall BD (2002) Using RPB1 sequences to improve phylogenetic inference among mushrooms (Inocybe, Agaricales). *Am J Bot* 89: 688–698. doi:10.3732/ajb.89.4.688. PubMed: 21665669.
79. Liu YJ, Whelen S, Hall BD (1999) Phylogenetic relationships among ascomycetes: evidence from an RNA polymerase II subunit. *Mol Biol Evol* 16: 1799–1808. doi:10.1093/oxfordjournals.molbev.a026092. PubMed: 10605121.
80. Lindblom L, Ekman S (2006) Genetic variation and population differentiation in the lichen-forming ascomycete *Xanthoria parietina* on the island Storfosna, central Norway. *Mol Ecol* 15: 1545–1559. doi:10.1111/j.1365-294X.2006.02880.x. PubMed: 16629810.
81. Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* 9: 286–298. doi:10.1093/bib/bbn013. PubMed: 18372315.
82. Katoh K, Kuma K-i, Toh H, Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* 33: 511–518. doi:10.1093/nar/gki198. PubMed: 15661851.
83. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959. PubMed: 10835412.
84. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587. PubMed: 12930761.
85. Corander J, Marttinen P, Sirén J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9: 539. doi:10.1186/1471-2105-9-539. PubMed: 19087322.
86. O'Neill EM, Schwartz R, Bullock CT, Williams JS, Shaffer HB et al. (2013) Parallel tagged amplicon sequencing reveals major lineages and phylogenetic structure in the North American tiger salamander (*Ambystoma tigrinum*) species complex. *Mol Ecol* 22: 111–129. doi:10.1111/mec.12049. PubMed: 23062080.
87. Falush D, Wirth T, Linz B, Pritchard JK, Stephens M et al. (2003) Traces of human migrations in *Helicobacter pylori* populations. *Science* 299: 1582–1585. doi:10.1126/science.1080857. PubMed: 12624269.
88. Leavitt SD, Fankhauser JD, Leavitt DH, Porter LD, Johnson LA et al. (2011) Complex patterns of speciation in cosmopolitan “rock posy” lichens – Discovering and delimiting cryptic fungal species in the lichen-forming *Rhizoplaca melanophthalma* species-complex (Lecanoraceae, Ascomycota). *Mol Phylogenet Evol* 59: 587–602. doi:10.1016/j.ympev.2011.03.020. PubMed: 21443956.
89. Conrad DF, Jakobsson M, Coop G, Wen X, Wall JD et al. (2006) A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nat Genet* 38: 1251–1260. doi:10.1038/ng1911. PubMed: 17057719.
90. Earl DA, von Holdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet Res* 4: 359–361.
91. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14: 2611–2620. doi:10.1111/j.1365-294X.2005.02553.x. PubMed: 15969739.
92. Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web Servers. *Syst Biol* 57: 758–771. doi:10.1080/10635150802429642. PubMed: 18853362.
93. Stamatakis A (2006) RAxML-VI-HP: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690. doi:10.1093/bioinformatics/btl446. PubMed: 16928733.
94. Kluge AG (1989) A concern for evidence and a phylogenetic hypothesis for relationships among *Epicrates* (Boidae, Serpentes). *Syst Zoology* 38: 7–25. doi:10.2307/2992432.
95. Degnan JH, Rosenberg NA (2006) Discordance of species trees with their most likely gene trees. *PLoS Genet* 2: e68. doi:10.1371/journal.pgen.0020068. PubMed: 16733550.
96. Leaché AD (2009) Species tree discordance traces to phylogeographic clade boundaries in North American fence lizards (*Sceloporus*). *Syst Biol* 58: 547–559. doi:10.1093/sysbio/syp057. PubMed: 20525608.
97. Kalinowski ST (2011) The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. *Heredity* (Edinb) 106: 625–632. doi:10.1038/hdy.2010.95. PubMed: 20683484.
98. Nelson MF, Anderson NO (2013) How many marker loci are necessary? Analysis of dominant marker data sets using two popular population genetic algorithms. *Ecol Evol* 3: 3455–3470. PubMed: 24223282.
99. Lanier HC, Knowles LL (2012) Is recombination a problem for species-tree analyses? *Syst Biol* 61: 691–701. doi:10.1093/sysbio/syr128. PubMed: 22215721.
100. Zhang C, Zhang D-X, Zhu T, Yang Z (2011) Evaluation of a Bayesian coalescent method of species delimitation. *Syst Biol* 60: 747–761. doi:10.1093/sysbio/syr071. PubMed: 21876212.
101. Takamatsu S, Matsuda S (2004) Estimation of molecular clocks for ITS and 28S rDNA in Erysiphales. *Mycoscience* 45: 340–344. doi:10.1007/S10267-004-0187-7.
102. Posada D (2008) jModelTest: Phylogenetic Model Averaging. *Mol Biol Evol* 25: 1253–1256. doi:10.1093/molbev/msn083. PubMed: 18397919.
103. Drummond AJ, Ho SY, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biol* 4: e88. doi:10.1371/journal.pbio.0040088. PubMed: 16683862.
104. Rambaut A, Drummond AJ (2009) TreeAnnotator. version 1.6.1. Available: <http://beast.bio.ed.ac.uk/TreeAnnotator>.
105. Collins RA, Cruickshank RH (2013) The seven deadly sins of DNA barcoding. *Mol Ecol Resour* 13: 969–975. PubMed: 23280099.
106. Ruiz-Sanchez E, Sosa V (2010) Delimiting species boundaries within the Neotropical bamboo *Oatea* (Poaceae: Bambusoideae) using molecular, morphological and ecological data. *Mol Phylogenet Evol* 54: 344–356. doi:10.1016/j.ympev.2009.10.035. PubMed: 19897047.
107. Dayrat B (2005) Towards integrative taxonomy. *Biol J Linn Soc* 85: 407–415. doi:10.1111/j.1095-8312.2005.00503.x.
108. Hey J, Waples RS, Arnold ML, Butlin RK, Harrison RG (2003) Understanding and confronting species uncertainty in biology and conservation. *Trends Ecol Evol* 18: 597–603. doi:10.1016/j.tree.2003.08.014.
109. Roe AD, Sperling FAH (2007) Population structure and species boundary delimitation of cryptic *Dioryctria* moths: an integrative approach. *Mol Ecol* 16: 3617–3633. doi:10.1111/j.1365-294X.2007.03412.x. PubMed: 17845435.
110. Rosentreter R (1993) Vagrant lichens in North America. *Bryologist* 96: 333–338. doi:10.2307/3243861.
111. Pinho C, Harris DJ, Ferrand N (2008) Non-equilibrium estimates of gene flow inferred from nuclear genealogies suggest that Iberian and North African wall lizards (*Podarcis* spp.) are an assemblage of incipient species. *BMC Evol Biol* 8: 63. doi:10.1186/1471-2148-8-63. PubMed: 18302778.
112. Eckert AJ, Carstens BC (2008) Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. *Mol Phylogenet Evol* 49: 832–842. doi:10.1016/j.ympev.2008.09.008. PubMed: 18845264.
113. Printzen C, Ekman S, Tønsgaard T (2003) Phylogeography of *Cavernularia hultenii*: Evidence of slow genetic drift in a widely disjunct

- lichen. *Mol Ecol* 12: 1473–1486. doi:10.1046/j.1365-294X.2003.01812.x. PubMed: 12755876.
114. Fernández-Mendoza F, Domaschke S, García MA, Jordan P, Martín MP et al. (2011) Population structure of mycobionts and photobionts of the widespread lichen *Cetraria aculeata*. *Mol Ecol* 20: 1208–1232. doi: 10.1111/j.1365-294X.2010.04993.x. PubMed: 21324011.
 115. Wirtz N, Printzen C, Lumbsch HT (2012) Using haplotype networks, estimation of gene flow and phenotypic characters to understand species delimitation in fungi of a predominantly Antarctic *Usnea* group (Ascomycota, Parmeliaceae). *Org Divers Evol* 12: 17–37. doi:10.1007/s13127-011-0066-y.
 116. Crespo A, Lumbsch HT, Mattsson J-E, Blanco O, Divakar PK et al. (2007) Testing morphology-based hypotheses of phylogenetic relationships in Parmeliaceae (Ascomycota) using three ribosomal markers and the nuclear RPB1 gene. *Mol Phylogenet Evol* 44: 812–824. doi:10.1016/j.ympev.2006.11.029. PubMed: 17276700.
 117. Truong C, Divakar PK, Yahr R, Crespo A, Clerc P (2013) Testing the use of ITS rDNA and protein-coding genes in the generic and species delimitation of the lichen genus *Usnea* (Parmeliaceae, Ascomycota). *Mol Phylogenet Evol* 68: 357–372. doi:10.1016/j.ympev.2013.04.005. PubMed: 23603312.
 118. Lutzoni F, Pagel M (1997) Accelerated evolution as a consequence of transitions to mutualism. *Proc Natl Acad Sci* 94: 11422–11427.
 119. Lumbsch HT, Hipp AL, Divakar PK, Blanco O, Crespo A (2008) Accelerated evolutionary rates in tropical and oceanic parmelioid lichens (Ascomycota). *BMC Evol Biol* 8: e257. doi: 10.1186/1471-2148-8-257. PubMed: 18808710.
 120. Wang H, Guo S, Huang M, Thorsten LH, Wei J (2010) Ascomycota has a faster evolutionary rate and higher species diversity than Basidiomycota. *Sci China Life Sci* 53: 1163–1169. doi:10.1007/s11427-010-4063-8. PubMed: 20953937.
 121. Palmqvist K, Sundberg B (2000) Light use efficiency of dry matter gain in five macro-lichens: relative impact of microclimate conditions and species-specific traits. *Plant Cell Environ* 23: 1–14. doi:10.1046/j.1365-3040.2000.00529.x.
 122. Renhorn KE, Esseen P-A, Palmqvist K, Sundberg B (1996) Growth and vitality of epiphytic lichens. *Oecologia* 109: 1–9.
 123. Ranius T, Johansson P, Berg N, Niklasson M (2008) The influence of tree age and microhabitat quality on the occurrence of crustose lichens associated with old oaks. *J Veg Sci* 19: 653–662. doi: 10.3170/2008-8-18433.
 124. Rivas Plata E, Lücking R, Lumbsch H (2008) When family matters: an analysis of Thelotremataceae (Lichenized Ascomycota: Ostropales) as bioindicators of ecological continuity in tropical forests. *Biodiversity Conserv* 17: 1319–1351. doi:10.1007/s10531-007-9289-9.
 125. Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters* 8: 336–352. doi:10.1111/j.1461-0248.2004.00715.x.
 126. Funk DJ (1998) Isolating a role for natural selection in speciation: host adaptation and sexual isolation in *Neochlamisus bebbianae* leaf beetles. *Evolution* 52: 1744–1759. doi:10.2307/2411347.
 127. Leavitt SD, St Clair LL (2008) Lichens of Boulder Mountain Plateau, Wayne, County, Utah, USA. *Evansia* 26: 85–89.
 128. Pöykkö H, Hyvärinen M, Bačkor M (2005) Removal of lichen secondary metabolites affects food choice and survival of lichenivorous moth larvae. *Ecology* 86: 2623–2632. doi:10.1890/04-1632.
 129. Huneck S (1999) The significance of lichens and their metabolites. *Naturwissenschaften* 86: 559–570. doi:10.1007/s001140050676. PubMed: 10643590.
 130. Asplund J, Solhaug KA, Gauslaa Y (2009) Fungal depsidones - an inducible or constitutive defence against herbivores in the lichen *Lobaria Pulmonaria*? *Basic Appl Ecol* 10: 273–278. doi:10.1016/j.baae.2008.04.003.
 131. Nash TH, Zavada M (1977) Population studies among Sonoran Desert species of *Parmelia* subg. *Xanthoparmelia* (Parmeliaceae). *Am J Bot* 64: 664–669. doi:10.2307/2441718.
 132. Pino-Bodas R, Burgaz A, Martín M, Lumbsch HT (2011) Phenotypical plasticity and homoplasy complicate species delimitation in the *Cladonia gracilis* group (Cladoniaceae, Ascomycota). *Org Divers Evol* 11: 343–355. doi:10.1007/s13127-011-0062-2.
 133. Velmala S, Myllys L, Halonen P, Goward T, Ahti T (2009) Molecular data show that *Bryoria fremontii* and *B. tortuosa* (Parmeliaceae) are conspecific. *Lichenologist* 41: 231–242. doi:10.1017/S0024282909008573.