sciendo

Plant and Fungal Systematics 64(2): 237–247, 2019
DOI: 10.2478/pfs-2019-0020

ISSN 2544-7459 (print)
ISSN 2657-5000 (online)

Shotgun sequencing decades-old lichen specimens to resolve phylogenomic placement of type material

Steven D. Leavitt[1,2*], Rachel Keuler[1], Clayton C. Newberry[2], Roger Rosentreter[3] & Larry L. St. Clair[1,2]

Article info

Received: 23 Mar. 2019
Revision received: 31 Jul. 2019
Accepted: 1 Aug. 2019
Published: 2 Dec. 2019

Associate Editor

Damien Ertz

Abstract. Natural history collections, including name-bearing type specimens, are an important source of genetic information. These data can be critical for appropriate taxonomic revisions in cases where the phylogenetic position of name-bearing type specimens needs to be identified, including morphologically cryptic lichen-forming fungal species. Here, we use high-throughput metagenomic shotgun sequencing to generate genome-scale data from decades-old (i.e., more than 30 years old) isotype specimens representing three vagrant taxa in the lichen-forming fungal genus *Rhizoplaca*, including one species and two subspecies. We also use data from high-throughput metagenomic shotgun sequencing to infer the phylogenetic position of an enigmatic collection, originally identified as *R. haydenii*, that failed to yield genetic data via Sanger sequencing. We were able to construct a 1.64 Mb alignment from over 1200 single-copy nuclear gene regions for the *Rhizoplaca* specimens. Phylogenomic reconstructions recovered an isotype representing *Rhizoplaca haydenii* subsp. *arbuscula* within a clade comprising other specimens identified as *Rhizoplaca haydenii* subsp. *arbuscula*, while an isotype of *R. idahoensis* was recovered within a clade with substantial phylogenetic substructure comprising *Rhizoplaca haydenii* subsp. *haydenii* and other specimens. Based on these data and morphological differences, *Rhizoplaca haydenii* subsp. *arbuscula* is elevated to specific rank as *Rhizoplaca arbuscula*. For the enigmatic collection, we were able to assemble the nearly complete nrDNA cistron and over 50 Mb of the mitochondrial genome. Using these data, we identified this specimen as a morphologically deviant form representing *Xanthoparmelia* aff. *subcumberlandia*. This study highlights the power of high-throughput metagenomic shotgun sequencing in generating larger and more comprehensive genetic data from taxonomically important herbarium specimens.

Key words: fungaria, herbaria, Illumina, metagenomics, museum, natural history collections, *Rhizoplaca*, vagrant

Introduction

Natural history collections are repositories of an impressive quantity and variety of biological information (Lane 1996; Schilthuizen et al. 2015; Wen et al. 2015). Ongoing developments with DNA sequencing technologies are enabling researchers to tap into genetic data that were previously inaccessible (Cooper 1994; Lan & Lindqvist 2019). Incorporating genetic data from natural history collections representing rare/threatened species and species that are otherwise hard to access into genetic-based analyses can provide critical insights (Heckeberg et al. 2016; Hundsdoerfer et al. 2017; McGuire et al. 2018). For example, genetic data from museum specimens has been used to track temporal and spatial changes in species distributions (Lozier & Cameron 2009), investigate population genetics (Spurgin et al. 2014; Schmid et al. 2018), and evaluate conservation implications (Anco et al. 2018; Brandt et al. 2018). Additionally, accessing molecular sequence data from natural history collections may be an invaluable resource for resolving taxonomic questions (Cappellini et al. 2014), including generating genetic data from type specimens (Hawksworth 2013; Silva et al. 2017; McGuire et al. 2018). Furthermore, DNA from historical museum specimens can provide novel insights into ambiguously identified collections (Chambers & Hebert 2016).

[1] Department of Biology, Brigham Young University, 4102 Life Science Building, Provo, UT 84602, USA
[2] M. L. Bean Life Science Museum, Brigham Young University, 4102 Life Science Building, Provo, UT 84602, USA
[3] Biology Department, Boise State University, 1910 West University Drive, Boise, ID 83725, USA
[*] Corresponding author e-mail: steve_leavitt@byu.edu

Sequencing DNA from historical specimens raises a number of challenges not encountered with fresh material. In addition to natural processes that lead to DNA degradation post-mortem (Lindahl 1993), some museum procedures for pest management or storage may also adversely affect museum collections as a source of useable DNA (Cooper 1994). In many cases, museum specimens contain short, fragmented DNA molecules, limiting the utility of PCR amplification (Burrell et al. 2015; Holmes et al. 2016). Specimens collected and preserved without sequencing in mind may also contain DNA from organisms stored nearby, collected alongside the specimen, or organisms that have grown on or within museum specimens (Pääbo 1988). Furthermore, destructive sampling, including DNA extraction, may potentially damage irreplaceable specimens (Hawksworth 2013). Degraded or contaminated DNA is often not suitable for traditional PCR amplification and Sanger sequencing, limiting the ability to incorporate museum specimens into DNA barcode identification surveys or phylogenetic or population genetic analyses. However, high-throughput sequencing technologies are now routinely used to generate genome-scale data from degraded DNA commonly found in museum specimens (Staats et al. 2013; Besnard et al. 2015; McCormack et al. 2016). This type of genome-scale data based on museum specimens is playing an increasingly important role in species delimitation and phylogenomic studies (Rittmeyer & Austin 2015; Wood et al. 2018).

Natural history collections have also played a key role in phylogenetic research efforts involving lichen-forming fungi, with DNA routinely extracted from herbarium specimens (Sohrabi et al. 2010; Gueidan et al. 2015; Kistenich et al. 2019), including a 151-year-old museum specimen (Redchenko et al. 2012). However, in other cases, generating sequence data from some lichens can be problematic with the passing of even short amounts of time, on the scale of months to years, without careful, intentional preparation and storage to maximize the probability of extracting useable DNA. Even with specimens collected specifically for genetic research, additional care must be taken to preserve useable DNA, and successful extraction of high-quality DNA may be limited after only a few months. With the prevalence of cryptic species inferred from molecular systematic studies of lichen-forming fungi (Crespo & Pérez-Ortega 2009), ascertaining the relationships of name-bearing types to robustly delimited species has important taxonomic implications (Hawksworth 2013). Improving our ability to recover DNA from type specimens will minimize the need to designate interpretive types, e.g., epitypes, as well as reducing potential taxonomic redundancy (Hawksworth 2008).

Species boundaries have been relatively well studied in the *Rhizoplaca melanophthalma* species complex, including empirical analyses of multi-locus and genome-scale data (Leavitt et al. 2011a; Leavitt et al. 2013b; Leavitt et al. 2016; Grewe et al. 2017). Most recently, significant phylogeographic substructure has been discovered in a monophyletic clade comprising the vagrant *Rhizoplaca* species, *R. haydenii* subsp. *arbuscula*, *R. haydenii* subsp. *haydenii* 1979, and *R. idahoensis* (McCune & Rosentreter 2007; Keuler et al. 2019). To better understand the evolutionary relationships of vagrant *Rhizoplaca* species and the potential taxonomic implications, in this study we targeted three type specimens representing *R. haydenii* subsp. *arbuscula*, *R. idahoensis*, and *R. melanophthalma* subsp. *cerebriformis*, collected in 1985 and 1986, for high-throughput sequencing (Fig. 1). In addition, a second specimen representing *R. idahoensis* (collected in 1987) and a 30-year-old specimen identified as *R. haydenii* from a subalpine habitat outside of the known distribution of *R. haydenii*. The latter had an unusual morphology (Fig. 1C), similar to *R. haydenii* but abnormal enough to warrant sequencing the DNA to more confidently resolve its taxonomic identity. For these specimens, we used a high-throughput shotgun sequencing approach to generate genome-scale data in order to more accurately infer evolutionary relationships.

While high-throughput sequencing of PCR amplicons has been shown to effectively generate sequence data from historical lichen collections (Kistenich et al. 2019), to our knowledge this is the first attempt to use high-throughput whole-genome shotgun sequencing of historical lichen type specimens within a taxonomic context.

Materials and methods

Specimen sampling

To explore the potential use of Illumina metagenomic shotgun sequencing to efficiently ascertain useable molecular sequence data from decades-old museum collections, including type collections, we targeted a total of five specimens ranging from 30 to 34 years old (Table 1). Specifically, we focused on vagrant specimens from the *Rhizoplaca melanophthalma* group, including three isotypes – *R. haydenii* subsp. *arbuscula*, *R. idahoensis*, and *R. melanophthalma* subsp. *cerebriformis*; a second historical collection representing *R. idahoensis*; and a morphologically unusual specimen originally identified as *R. haydenii* and collected in a krumholtz-dominated subalpine habitat outside of the known distribution of *R. haydenii* (Fig. 1C). Isotype collections representing *R. haydenii* subsp. *arbuscula* included specimens with somewhat deviating morphologies – specimens with more robust thalli (Fig. 1A) relative to other thalli with narrower, finely branching lobes (Fig. 1B). Narrower, finely branching lobed specimens morphologically similar to thalli from an *R. haydenii* subsp. *arbuscula* isotype collection were represented in previously generated metagenomic samples, including individuals collected near the type locality (Keuler et al. 2019). Therefore, we selected a more robust thallus from the isotype collection for metagenomic shotgun sequencing. We were unable to secure type material representing *R. haydenii* subsp. *haydenii* for DNA sequencing. Therefore, we collected *R. haydenii* subsp. *haydenii* specimens from two populations near the type locality in eastern Wyoming (Table S1) – 'Laramie Plains' (H-NYL 28325; [basionym published

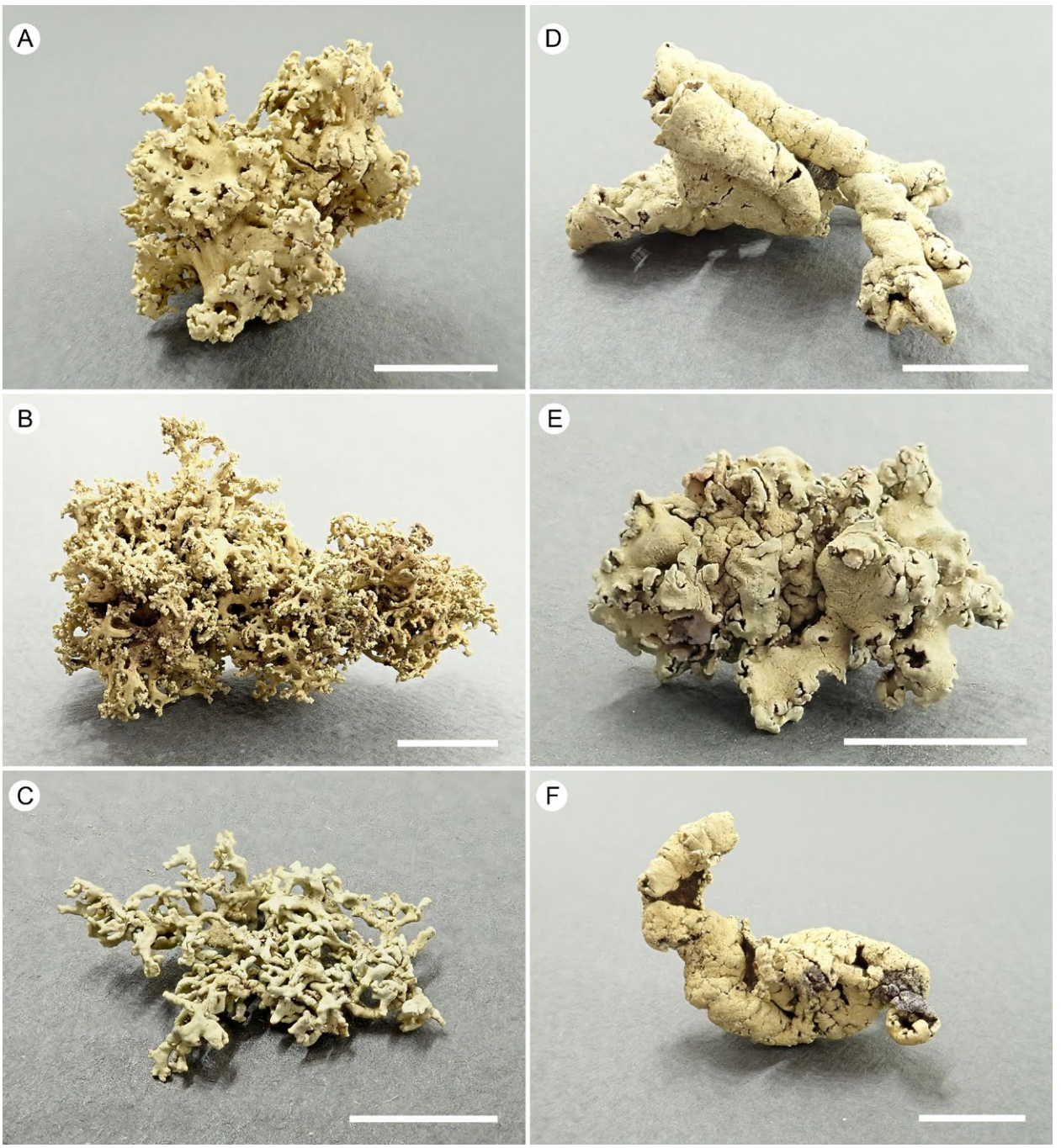

Figure 1. Vagrant *Rhizoplaca* specimens, including isotypes, and a morphologically aberrant vagrant *Xanthoparmelia* specimen. A – *Rhizoplaca haydenii* subsp. *arbuscula* (Rosentreter 3861; isotype – sequenced for this study); B – *Rhizoplaca haydenii* subsp. *arbuscula* (Rosentreter 3861; isotype – not sequenced); C – *Xanthoparmelia* aff. *subcumberlandia* (Newberry 936 – sequenced for this study), originally identified as *R. haydenii*; D – *Rhizoplaca idahoenis* (Rosentreter 3719; isotype – sequenced for this study); E – *Rhizoplaca haydenii* subsp. *haydenii* (Leavitt 18-1006 – sequenced for this study); F – *Rhizoplaca melanophthalma* subsp. *cerebriformis* (Rosentreter 3876; isotype – sequenced for this study). These specimens are housed at the Herbarium of Non-Vascular Cryptogams (BRY-C). Scales: 1 cm in all panels.

in: Tuckerman, Proc. Amer. Acad. Arts 6: 227. 1866]) for this study. Using material from the new collections, four individual specimens representing the range of morphological variation were selected for sequencing. New specimens sequenced for this study were then combined with sequences from specimens reported by Keuler et al. (2019; Table S1). All specimens were visually assessed using an Olympus SZH zoom stereo dissecting microscope, and secondary metabolites were identified using thin layer chromatography (TLC). Lichen compounds were extracted in acetone using a small, clean piece of thallus material (subsequently used for DNA isolation); the acetone extract was then separated using solvents C and G following the methods of Orange et al. (2001).

DNA extraction and sequencing

Total genomic DNA was extracted using the E.Z.N.A. Plant DNA DS Mini Kit (Omega Bio-tek) following the manufacturers' recommendations. Previous efforts to generate sequence data from five to twenty-year-old *Rhizoplaca* specimens have been successful in other studies. However, for this study, multiple attempts to generate

Table 1. Summary of herbarium specimens sequenced for this study, including number of metagenomic reads per sample, and average coverage at two loci – the multi-copy nrDNA internal transcribed spacer region (ITS) and the single-copy protein-coding region *RPB*1. The single specimen representing *Xanthoparmelia subcumberlandia* was originally identified as *R. haydenii.* All specimens are housed at the Herbarium of Non-Vascular Cryptogams (BRY-C), M.L. Bean Life Science Museum, Provo, Utah, USA.

Taxon	Voucher	Locality	Collection date	# of filtered Illumina reads	ITS mean coverage	*RBP*1 mean coverage
R. haydenii subsp. *arbuscula* (usnic acid chemotype)	R. Rosentreter 3861 (isotype)	USA, Idaho, Lemhi Co.	08 June 1986	31,069,276 (2×150)	402.1	29.1
R. idahoensis (usnic acid chemotype)	R. Rosentreter 3719 (isotype)	USA, Idaho, Lemhi Co.	27 August 1985	28,832,106 (2×150)	134.1	40.4
R. idahoensis (usnic acid chemotype)	A. DeBolt 750	USA, Idaho, Lemhi Co.	26 June 1987	26,370,812 (2×150)	171.9	31.4
R. melanophthalma subsp. *cerebriformis* (usnic and psoromic acids chemotype)	R. Rosentreter 3876 (isotype)	USA, Idaho, Lemhi Co.	08 June 1986	28,558,924 (2×150)	662.5	39.7
Xanthoparmelia aff. *subcumberlandia* (usnic, norstictic, and stictic acids chemotype)	C. Newberry 936	USA, Utah, Duchesne Co.	26 July 1989	8,234,622 (2×150)	24.1	<0.5

genetic data via Sanger sequencing for the *R. haydenii*-like specimens collected in 1989 from subalpine habitat in the Uinta Mountains of northeastern Utah were unsuccessful. DNA extractions from these specimens revealed low quantities of highly fragmented DNA. From this total genomic DNA, high-throughput sequencing libraries were prepared using the standard Illumina whole-genome sequencing (WGS) library preparation process with Adaptive Focused Acoustics for shearing (Covaris) followed by an AMPure cleanup step. This DNA was then processed using the NEBNext® Ultra™ II End Repair/dA-Tailing Module and the NEBNext® Ultra™ II Ligation Module (New England Biolabs), using standard Illumina index primers. Libraries were run on the Illumina HiSeq with 2×125 paired-end sequencing at the DNA Sequencing Center located on the campus of Brigham Young University, Provo, Utah, USA.

Read filtering and phylogenomic datasets

All paired-end (PE) reads were filtered using TRIMMOMATIC v0.33 (Bolger et al. 2014) before assembly to remove low-quality reads and/or included contamination from Illumina adaptors using the following parameters: ILLUMINACLIP; LEADING:3; TRAILING:3; SLIDINGWINDOW:4:15; and MINLEN:36. Retained, high-quality reads were assembled de novo using SPAdes 3.12 (Bankevich et al. 2012), with the parameters '-meta -k 35,55,85'. Contigs containing the nuclear ribosomal operon (nrDNA) were identified using a custom BLAST search. ITS sequences from the contig containing nrDNA were then searched against NCBI's GenBank database using BLAST (Wheeler et al. 2006) to infer the samples' identities. To assess coverage, PE reads were initially mapped back to sequences representing: (i) the multi-copy nrDNA contig and (ii) single copy protein-code gene regions, including the *MCM*7, *RPB*1, and *RPB*2 (Lofgren et al. 2019). Read mapping was implemented using the Geneious Prime Read Mapper (Kearse et al. 2012), using the 'medium-low sensitivity/fast' settings and iterated 5 times. Reads that were successfully mapped to the nrDNA contig were subsequently assembled *de novo* using the native Geneious Primer assembler at 'medium-low sensitivity' (Kearse et al. 2012).

For all specimens with inferred identity in the *R. haydenii*/*idahoensis* group (Keuler et al. 2019), we generated a nuclear phylogenomic alignment using REALPHY v1.12 (Bertels et al. 2014), which has empirically been shown to construct genome-scale datasets (Leavitt et al. 2016; Zeng et al. 2018). New data were combined with metagenomic reads from previous studies (Leavitt et al. 2016) and mapped to BUSCO gene regions > 1 kb from a pure fungal culture of *R. melanophthalma* (see Keuler et al. 2019) in REALPHY v1.12, implementing Bowtie 2.1.0 (Langmead & Salzberg 2012) for read mapping with the following parameters: –readLength 75 –perBaseCov 5 –gapThreshold 0.2 with the remaining parameters set to default values. With the –gapThreshold parameter set to 0.2, each site had no more than 20% missing data.

For the specimen from the Uinta Mountains originally identified as *R. haydenii* (Newberry 936 [BRY-C]; Fig. 1C), we were only able to recover nuclear ribosomal DNA (nrDNA) sequences and no single-copy nuclear markers (see Results). BLAST searches using markers within the nrDNA, e.g., the internal transcribed spacer region (ITS1, 5.8S, and ITS2), a fragment of the large subunit (nrLSU) and a fragment of the intergenic spacer region (IGS), suggested a close relationship to a number of *Xanthoparmelia* species known from western North America. ITS, nrLSU, and IGS sequences from the Newberry 936 (BRY-C) specimen were extracted from the nrDNA contig and combined with a previously published alignment for western North American *Xanthoparmelia* species (TreeBase #14929; https://treebase.org), comprising four nrDNA loci – ITS, nrLSU, IGS and a group I intron, and five protein-coding loci – β-tubulin, *GAPDH*, *MCM*7, *RPB*1, and *RPB*2. New nrDNA sequences (ITS, nrLSU, and IGS) were added to the TreeBase alignment using the –addfragments feature in the online MAFFT v7 server (Rozewicki et al. 2017) followed by implementation of the –multipair strategy for a more accurate alignment, using the '--multipair' strategy and 200PAM /

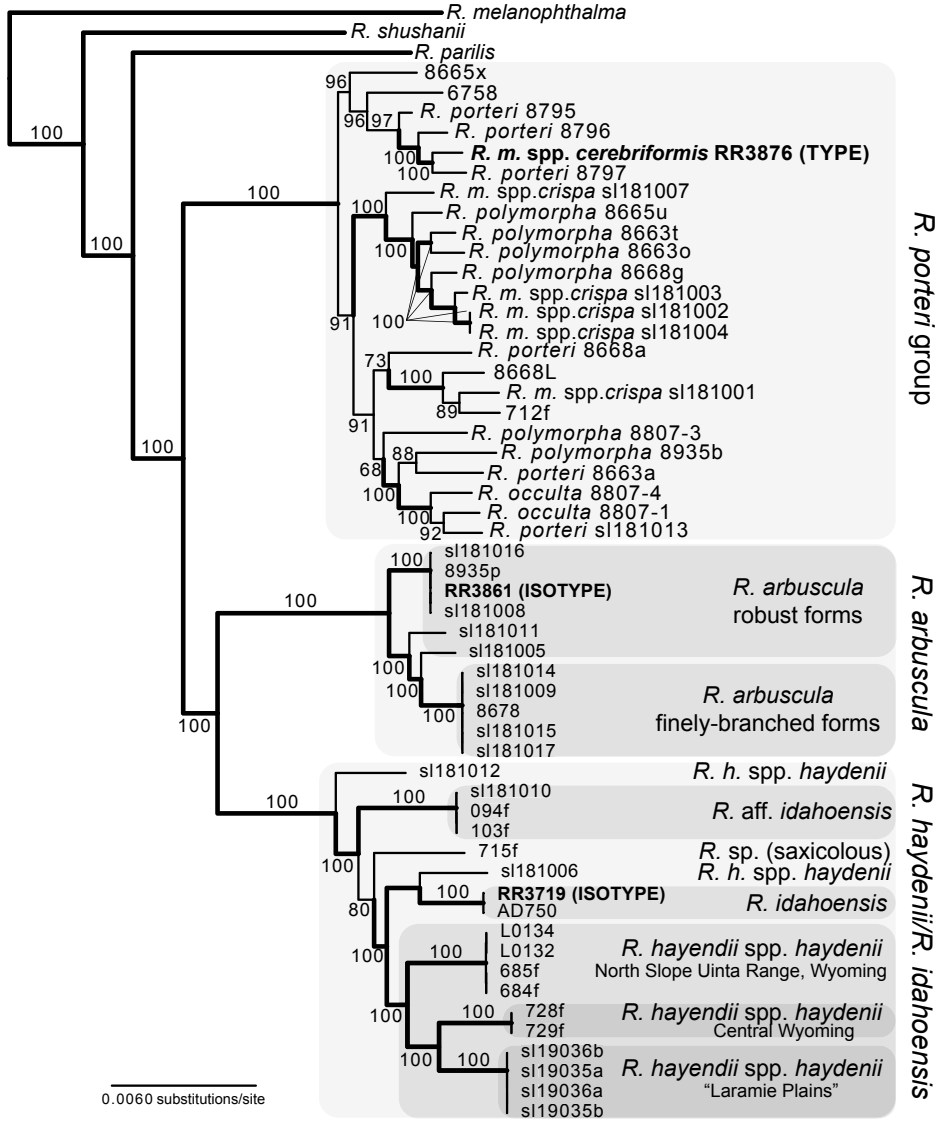

Figure 2. Phylogeny of the *Rhizoplaca melanophthalma* species group inferred from a 1.64 Mb BUSCO alignment. Isotype specimens are indicated in bold text. Bootstrap support values are indicated at nodes, with thickened branches highlighting bootstrap support values = 100%. Distinct forms in the '*arbuscula*' and '*haydenii*/*idahoensis*' clades are indicated.

K = 2 scoring matrix. For the misidentified *Xanthoparmelia* specimen (Newberry 936 [BRY-C]; Fig. 1C), contigs from the SPAdes assembly containing portions of the mitochondrial genome were then identified using a customized BLAST search (Wheeler et al. 2006).

Phylogenetic analysis

Evolutionary relationships were inferred using a supermatrix approach, which has been shown to accurately infer relationships across a wide range of scenarios (Tonini et al. 2015). We inferred relationships of the four vagrant *Rhizoplaca* samples, including three isotypes, with other members of the *R. melanophthalma* species complex using the REALPHY phylogenomic dataset analyzed under a maximum likelihood (ML) criterion as implemented in IQ-TREE v1.6.7 (Nguyen et al. 2015), with 1,000 ultrafast bootstrap replicates (Hoang et al. 2018), after which the best-fitting substitution model for the entire REALPHY matrix was selected using ModelFinder (Kalyaanamoorthy et al. 2017) based on 1,000 ultra-fast bootstrap replicates (Hoang et al. 2018). The concatenated nine-locus *Xanthoparmelia* dataset was analyzed in IQ-TREE v1.6.7, as described above, with the exception that each locus was treated as a separate partition, with no additional partitioning within loci (e.g., codon position, etc.). In all cases, nodal support was mapped onto the best-scoring ML topologies.

Results

TLC indicated that the specimens representing *R. haydenii* subsp. *arbuscula*, *R. haydenii* subsp. *haydenii* and *R. idahoensis* (a total of two specimens) contained usnic acid, while *R. melanophthalma* subsp. *cerebriformis* contained usnic and psoromic acids (Table 1). The morphologically unusual specimen originally identified as *R. haydenii* (Newberry 936 [BRY-C]) contained usnic, norstictic, and stictic acids.

Illumina metagenomic shotgun sequencing generated between 8.2 and 31.1 million PE reads for each specimen

(Table 1; Table S1), and newly generated reads, including those from isotype collections, are available in the NCBI Short Read Archive under project PRJNA576709. From each metagenomic assembly we were able to identify a single contig containing the complete nrDNA small subunit (nrSSU), ITS, and the nrLSU; the contig also contained the majority of the IGS region. Exploratory BLAST searches and phylogenetic analyses showed that the morphologically unusual specimen originally identified as *R. haydenii* (Newberry 936 [BRY-C]) clearly falls within the genus *Xanthoparmelia*, while the other three specimens belonged to the *R. melanophthlama* complex. Coverage of the single copy *RPB*1 marker ranged from 29× to 40× for the *Rhizoplaca* species; and < 0.5× for the *Xanthoparmelia* specimen (Table 1); coverage of the multi-copy ITS ranged from 24× (Newberry 936 [BRY-C]) to 663× (Rosentreter 3876 [BRY-C]).

The BUSCO alignment of the *R. melanophthalma* species complex reconstructed in REALPHY comprised a total of 1,650,966 aligned nucleotide position characters (File S1). Phylogenomic inference using this alignment in IQ-TREE recovered a generally well-supported (e.g., bootstrap values >95%) topology consistent with previous reconstructions (Fig. 2).

Both *R. idahoensis* specimens sequenced for this study, including an isotype (Rosentreter 3719 [BRY-C]), were recovered within the '*R. haydenii*/*R. idahoensis*' clade as a distinct, well-supported clade within specimens representing the polyphyletic *R. haydenii*. Another lineage containing two specimens morphologically similar to *R. idahoensis* (Leavitt 094f & Leavitt 103f [BRY-C]) was also recovered within the '*R. haydenii*/*R. idahoensis*' clade but was clearly distinct from the type specimen (Fig. 2). The four *R. haydenii* subsp. *haydenii* specimens collected from locations near the type locality – Laramie, WY – were nearly genetically identical and recovered within the '*R. haydenii*/*R. idahoensis*' clade; and all *R. haydenii* subsp. *haydenii* specimens from Wyoming were recovered within a single monophyletic clade, with phylogenetic substructure corresponding to distinct geographic regions. The *R. haydenii* subsp. *arbuscula* isotype (Rosentreter 3861 [BRY-C]), was recovered within a second major clade, distinct from the '*R. haydenii*/*R. idahoensis*' clade, the '*R. arbuscula*' group (Fig. 2), comprised of other *R. haydenii* subsp. *arbuscula* specimens collected from the region. All of the *R. haydenii* subsp. *haydenii* specimens with narrower, finely branching lobes (see Fig. 1B) included in this study (all from Lemhi Valley, ID, USA) were phylogenetically distinct from the more robust forms. The final isotype specimen, representing *R. melanophthalma* subsp. *cerebriforms* (Rosentreter 3876 [BRY-C]) was recovered within the '*R. porteri*' clade, and was closely related to three specimens representing *R. porteri* (Fig. 2).

The *Xanthoparmelia* phylogeny (Fig. 3) inferred from a nine-locus dataset (File S2), including the unusual *Xanthoparmelia* specimen from the Uinta Mountains of northeastern Utah (Fig. 1C), was congruent with previous reconstructions, supporting two distinct clades provisionally named clades 'D' and 'E' (Fig. 3). The clade containing the unusual *Xanthoparmelia* specimen was recovered within a weakly supported clade comprised of other stictic acid-containing *Xanthoparmelia* species, which has previously been supported as a distinct genetic cluster, cluster 'E1-a' (Leavitt et al. 2013c). Many members of this clade correspond morphologically with *X. subcumberlandia* Elix & T. H. Nash. Because we were unable to generate data from single-copy nuclear genomic regions, we provide the nearly complete nrDNA operon and fragments of the mitochondrial genome assembly as supplementary files (File S3).

Taxonomy

Rhizoplaca arbuscula Rosentreter, St. Clair & Leavitt, comb. et stat. nov. (Fig. 1A, B).

MycoBank MB 832846

Basionym: *Rhizoplaca haydenii* subsp. *arbuscula* Rosentreter, Monogr. N. Am. Lichenol. 1: 91. 2007.

Diagnosis: similar to *Rhizoplaca haydenii* but with more narrowly divided lobes and evolutionarily distinct from the lineage that includes *R. haydenii*.

Type: USA, Idaho: Lemhi Co., 3 km NW of Leadore, 44.695° –113.386°, 1800 m, 8 June 1986. Rosentreter 3861 (Holotype – OSC; Isotypes – BRY-C, ASU, CANL, COLO, GZU, US).

Description. Free on soil; thallus globose with abundant marginal whitish nodules; lobes narrowly divided, flat to round in cross section; medulla lacking both placodiolic and psoromic acids.

Etymology. Named after the tree-like resemblance of the narrowly divided lobes.

Discussion

Here we report the first, to our knowledge, successful use of high-throughput sequencing specifically targeting lichen type specimens to resolve the evolutionary relationships of the mycobiont. Based on the genome-scale data generated in this study and morphological differences, we propose to elevate *R. haydenii* subsp. *arbuscula* to species, formally named here as *R. arbuscula.* We also show that data from high-throughput metagenomic sequencing can be used to identify ambiguously identified historical lichen collections. Below we discuss the implications for vagrant members of the *Rhizoplaca melanophthalma* species complex (Leavitt et al. 2011a), as well as lichen research in general.

The *Rhizoplaca melanophthalma* species complex (Leavitt et al. 2011a) has received considerable attention, including the description of a number of new species supported largely by genetic data (Leavitt et al. 2013a; Leavitt et al. 2016; Grewe et al. 2017). Most recently, sampling of vagrant forms in this complex has revealed striking phylogeographic structure in the '*R. haydenii*' clade including *R. haydenii* subsp. *arbuscula*, *R. haydenii* subsp. *haydenii*, and *R. idahoensis* (Keuler et al. 2019). It is now apparent that extensive, distribution-wide sampling

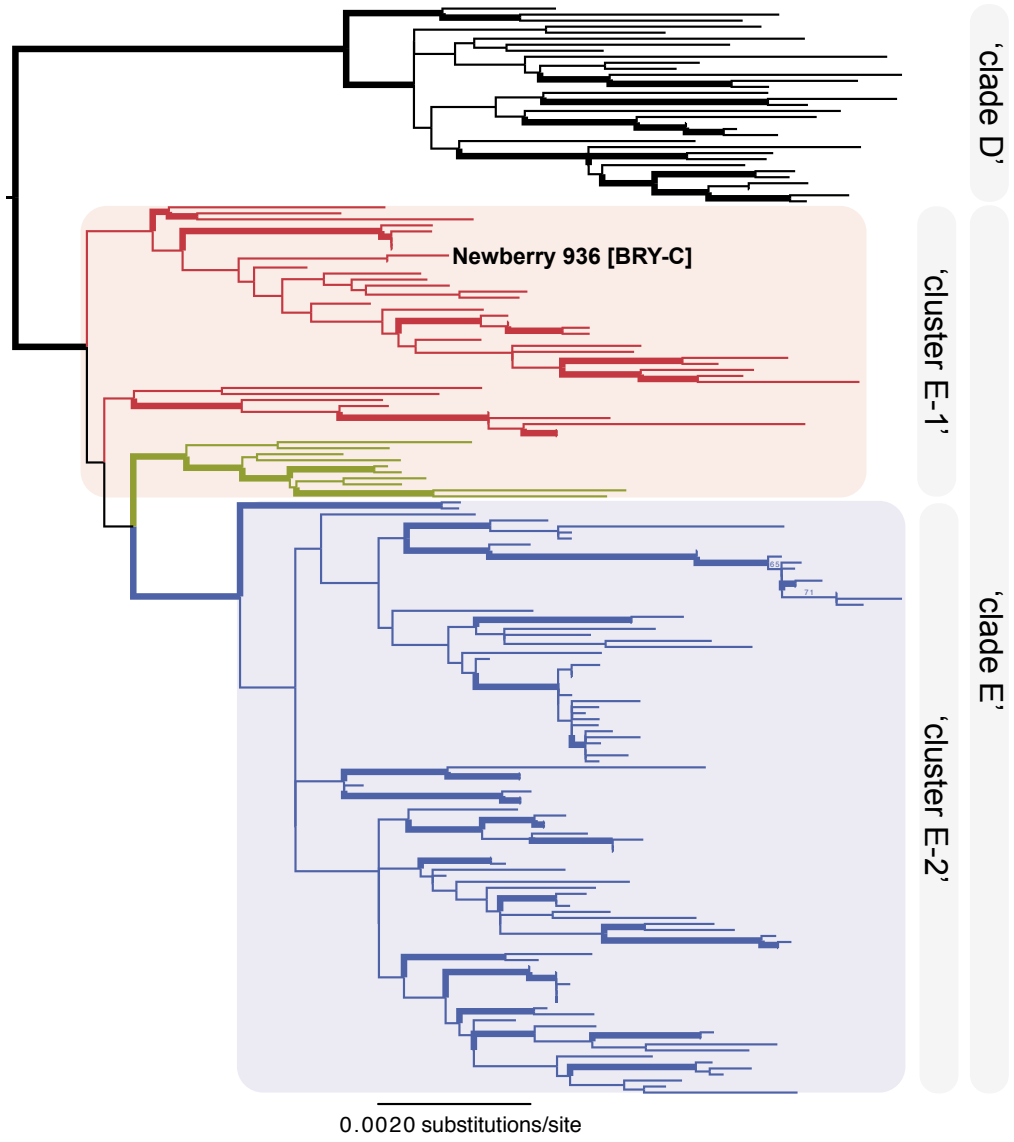

Figure 3. Phylogeny of *Xanthoparmelia* clades 'D' and 'E' estimated from concatenated ribosomal (nrLSU, IGS, ITS, and a group I intron) and protein-coding (*β-tubulin*, GAPDH, *MCM7*, *RPB1*, and *RPB2*) genetic loci. Thickened branches indicate ML bootstrap values ≥ 70%, and all other branches were supported by < 70% bootstrap values. Colored branches correspond to population genetic clusters. The morphologically deviant specimen, originally identified as *R. haydenii* (Newberry 936 [BRY-C]), was recovered within the genetic group 'cluster E-1', highlighted in red.

of members of this clade will be essential for inferring the evolutionary history of this group in order to accurately make appropriate taxonomic revisions and better understand phylogeographic patterns and processes of speciation.

Within the broad '*R. haydenii*' clade, two well-supported clades were recovered as clearly distinct – the '*R. arbuscula*' and '*R. haydenii*/*R. idahoensis*' lineages (Fig. 2). Based on morphological differences between specimens in the '*R. arbuscula*' clade and specimen recovered in the '*R. haydenii*/*R. idahoensis*' clade and corroborating support from phylogenomic data, we propose that *R. haydenii* subsp. *arbuscula* be formally described as a distinct species – *R. arbuscula*. At a finer scale, within the '*R. arbuscula*' clade, two distinct forms were observed, one with a shrubbier thallus and the others with more narrowly divided and elongated lobes, both of which are represented in isotype collections (Fig. 1A, B). However, both forms are closely related and distinct from *R. haydenii* subsp. *haydenii* (Fig. 2). The range of variation in *R. arbuscula* specimens is distinct from the variation in *R. haydenii* subsp. *haydenii* specimens, which have much broader lobes and are generally found with bluish/dark green pigments.

In this study, the relationships of isotypes representing *R. haydenii* subsp. *arbuscula* (elevated here to *R. arbuscula*) and *R. idahoensis* were recovered with strong support within the *Rhizoplaca melanophthalma* species complex (Fig. 2). Specimens representing *R. haydenii* subsp. *haydenii* were recovered in multiple distinct clades. *Rhizoplaca haydenii* subsp. *haydenii* was recovered as polyphyletic, with *R. idahoensis*, *R.* cf. *idahoensis* and an interesting saxicolous specimen from western Montana (Leavitt 715f [BRY-C]), nested within the broader '*R. haydenii*/*R. idahoensis*' clade (Fig. 2). The phylogeographic structure in the '*R. haydenii*/*R. idahoensis*' clade is striking, and additional studies will be required to more accurately determine population structure,

phylogeographic patterns, and species boundaries in this group. Different forms of *Rhizoplaca haydenii* s.lat. grow intermixed in populations across Idaho, Wyoming, and Montana. More extensive sampling of these populations will be required in order to accurately determine the taxonomic importance of these variable forms, e.g., the role of environmental modification or local adaptations vs. conserved diagnostic traits separating evolutionarily independent lineages. Current sampling of *R. arbuscula* is limited to central Idaho, although this taxon occurs in additional populations in Idaho, Montana, Utah, and Wyoming. Whether *R. arbuscula* has phylogeographic structure similar to the '*R. haydenii*/*R. idahoensis*' clade remains unknown.

The isotype representing *R. melanophthalma* subsp. *cerebriformis* (Fig. 1F) was recovered with the '*R. porteri* group' (Leavitt et al. 2016), comprising three formally described species – *R. occulta*, *R. polymorpha* and *R. porteri* – and two formally described subspecies, *R. melanophthalma* subsp. *cerebriformis* and *R. melanophthalma* subsp. *crispa*. While data from the standard DNA barcoding marker for fungi, the ITS region and reduced-representation genome-scale data support the segregation of the three species-level taxa as distinct lineages (Leavitt et al. 2013a; Grewe et al. 2017), ongoing research suggests a more dynamic evolutionary history in this group (Keuler et al. 2019). Pending a robust species-delimitation study of the members in this clade, we propose that all specimens recovered therein should be treated as the '*R. porterii* complex', with less emphasis on the currently recognized species names.

In this study, we demonstrate the utility of high-throughput sequencing for resolving the position of ambiguous herbarium collections. In this case, a specimen tentatively identified as *R. haydenii* (Fig. 1C) was inferred to actually belong to *Xanthoparmelia* based on rDNA and mitochondrial DNA data generated from Illumina short metagenomic reads (Fig. 3). While previous attempts to generate sequence data for this specimen via polymerase chain reaction (PCR) and Sanger sequencing consistently failed, our high-throughput sequencing approach generated a nearly complete rDNA operon (18S, ITS1, 5.8S. ITS, 28S, and intergenic spacer regions) and over 50 kb of the mitochondrial genome (File S3). These data document an incredibly deviant form of *Xanthoparmelia* (Fig. 4). Specifically, upper and lower cortices were largely indistinguishable, lobes demonstrated atypical shapes, and rhizines were absent. Another *Xanthoparmelia* species from western North America, *X. idahoensis* Hale, also has unusual contorted lobes (Fig. 4C). However, representatives of *X. idahoensis* are much larger than the deviant *Xanthoparmelia* specimen sequenced here, occur in relatively barren 'badlands', in contrast to subalpine habitats where the deviant *Xanthoparmelia* specimen was collected, and belong to a distinct lineage of *Xanthoparmelia* (Hale 1990; Leavitt et al. 2011b). Other closely related *Xanthoparmelia* specimens occurring as vagrant/erratic forms and also collected in the Uinta Mountains had more typical features of other *Xanthoparmelia* species (Fig. 4B).

When working with historical specimens where PCR and Sanger sequencing may fail to yield data, high-throughput sequencing provides an effective alternative for generating crucial molecular data (Green et al. 2006; Rowe et al. 2011). High-coverage metagenomic data from type specimens is also more versatile, in terms of scientific utility, as the data can be integrated into a variety of genetic datasets, ranging from single- and multi-locus datasets (Leavitt et al. 2018) to reduced-representation phylogenomic datasets (Grewe et al. 2017) to other phylogenomic and genomic applications, including investigations into other associated symbiotic partners (Paul et al. 2018).

While guidelines for collecting botanical samples intended for genomic work have been developed (Funk et al. 2017), to our knowledge established best practices for collecting lichen samples for genomic work have not yet been developed. Extraction of DNA from members of the *Rhizoplaca melanophthalma* group depends to a large degree on specimen processing and curation. Specimens collected dry, processed right after collection and housed in herbaria with low humidity may yield high-quality DNA years after they were collected (personal observation). However, in other cases we have been unable to extract DNA from relatively recent collections (1 to 5 years) of *Rhizoplaca* specimens. A variety of factors likely influence DNA quantity and quality in lichen collections. Relative to herbaria predominantly housing vascular plants, lichen collections tend to not have as much DNA degradation since most lichen herbaria are not treated with chemicals (Hall 1988). Lichens from arid regions are often dry when collected or are more easily dried without the use of artificial drying methods; and based on our anecdotal observations, the integrity of the DNA tends to be preserved at a higher level even in older collections. In contrast, extracting high-quality DNA from lichens occurring in humid areas is generally more problematic. In many cases, lichen specimens from humid regions must be dried using artificial heating methods which often degrade the DNA. Furthermore, our experience suggests that lichen collections stored in plastic containers/bags, even for short periods of time, are less likely to yield high-quality DNA. Similarly, moistening and pressing the specimens after collection, especially if a heated drier is used, may damage DNA from lichens intended for genomic work.

Successfully generated DNA from name-bearing types will be particularly valuable for resolving taxonomic issues in complexes comprised of morphologically cryptic lineages (Hawksworth 2013). With the increased recognition of morphologically cryptic species of lichen-forming fungi (Crespo & Pérez-Ortega 2009), attempts to recover genetic data from previously designated types where ample material exists will minimize the risk of erroneously designating epitypes (Hawksworth 2013). However, we note that in many cases non-destructive microscopic studies are sufficient to unequivocally place type specimens in relationship to recently collected specimens. Similarly, shotgun sequencing approaches for generating genome-scale data may not be necessary in all cases. Standard PCR amplification and Sanger sequencing may

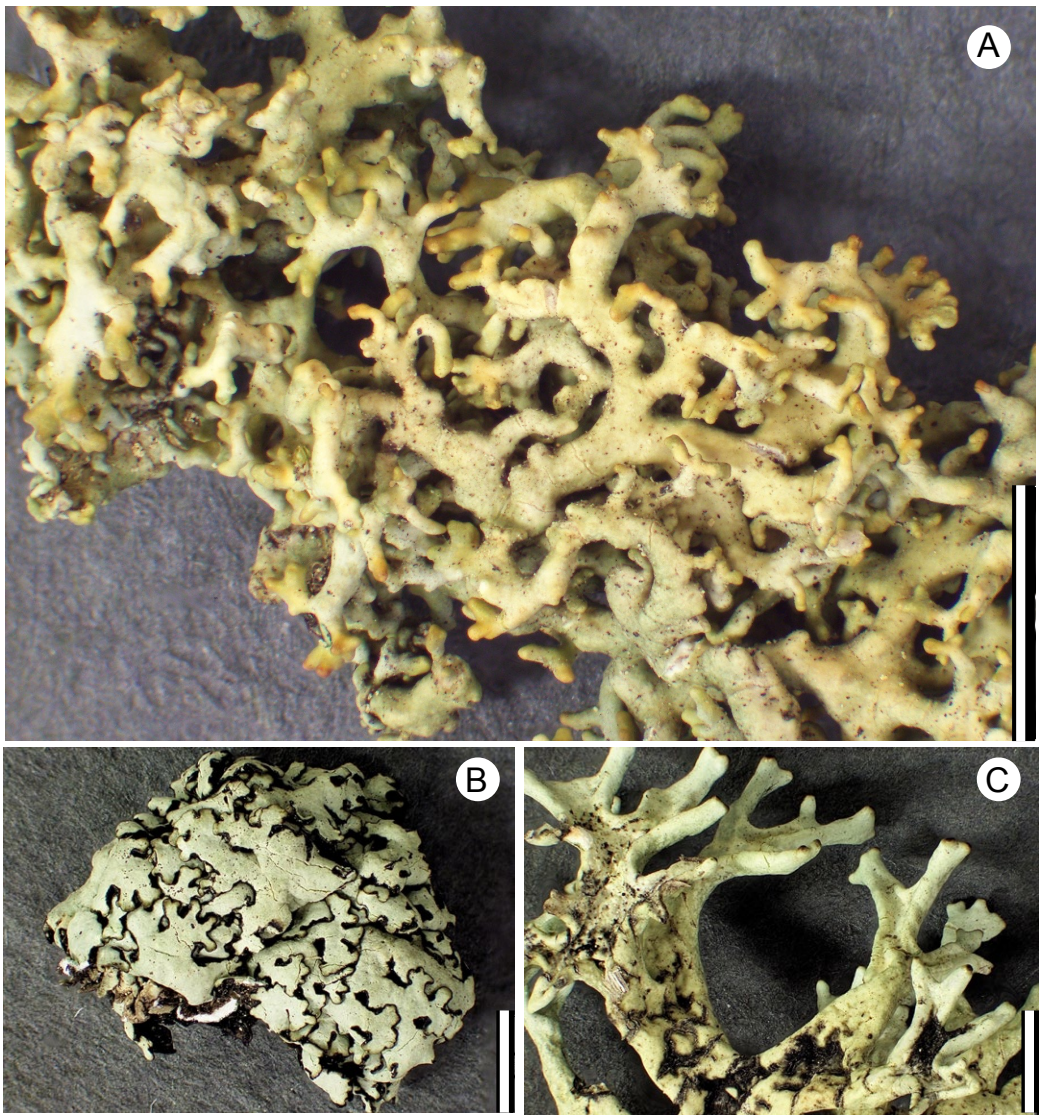

Figure 4. Close-up images of *Xanthoparmelia* specimens. A – *Xanthoparmelia* aff. *subcumberlandia* (Newberry 936 [BRY-C]), originally identified as *R. haydenii* – no rhizines, no typical lobe margins, no differentiated upper or lower cortex; B – *Xanthoparmelia* aff. *subcumberlandia* (Leavitt 121f [BRY-C]) with characteristic features of *Xanthoparmelia*; and C – *Xanthoparmelia idahoensis* (Leavitt 787f), a taxon distantly related to *X. subcumberlandia* but with contorted and twisted lobes somewhat similar to those in panel 'A'. Specimens in panels A and B are closely related and belong to clade 'E' – see Figure 3. Scales: A = 1 mm; B, C = 2 mm.

be sufficient to accurately assign name-bearing types; and high-throughput sequencing of PCR amplicons has also been shown to effectively generate sequence data from historical lichen collections (Kistenich et al. 2019). As sequencing technologies continue to change and the cost of generating genome-scale data decreases, researchers must carefully evaluate the most appropriate data for effectively addressing specific questions. In conclusion, our study provides additional evidence of the value of historical lichen collections as a source of genetic material (Bruns et al. 1990; Brock et al. 2009), highlighting the role of high-throughput sequencing technologies in generating larger and more comprehensive genetic data from taxonomically important herbarium specimens.

Acknowledgements

We dedicate this publication to our friend and colleague, Dr. James Lawrey, on the occasion of his 70th birthday. Dr. Lawrey's academic career and exemplary professional interactions have been inspiring. We acknowledge fruitful collaboration with the Snake River Plain herbarium at Boise State University. This research was supported by the M. L. Bean Life Science Museum, Brigham Young University, Provo, Utah, USA. We thank Ed Wilcox, DNA Sequencing Center, Brigham Young University, Provo, Utah, USA, for help with sequencing.

Supplementary electronic material

Table S1. Summary of specimens included in the study, including taxon, clade, DNA code, locality, GPS coordinates, number of filtered reads, and voucher ID. Download file

File S1. The BUSCO alignment of the *R. melanophthalma* species complex reconstructed in REALPHY comprised a total of 1,650,966 aligned nucleotide position characters. Download file

File S2. Concatenated alignment of ribosomal (nrLSU, IGS, ITS, and a group I intron) and protein-coding (*β-tubulin*, GAPDH, *MCM7*, *RPB1*, and *RPB2*) sequence data from *Xanthoparmalia* clades 'D' and 'E'. Download file

File S3. *Xanthoparmelia* nrDNA cistron; and three mitochondrial contigs from the *de novo* assembly of short reads. Download file

References

Anco, C., Kolokotronis, S. O., Henschel, P., Cunningham, S. W., Amato, G. & Hekkala, E. 2018. Historical mitochondrial diversity in African leopards (*Panthera pardus*) revealed by archival museum specimens. *Mitochondrial DNA Part A* 29: 455–473.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A. & Pevzner, P. A. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19: 455–477.

Bertels, F., Silander, O. K., Pachkov, M., Rainey, P. B. & van Nimwegen, E. 2014. Automated reconstruction of whole-genome phylogenies from short-sequence reads. *Molecular Biology and Evolution* 31: 1077–1088.

Besnard, G., Bertrand, J. A., Delahaie, B., Bourgeois, Y. X., Lhuillier, E. & Thébaud, C. 2015. Valuing museum specimens: high-throughput DNA sequencing on historical collections of New Guinea crowned pigeons (*Goura*). *Biological Journal of the Linnean Society* 117: 71–82.

Bolger, A. M., Lohse, M. & Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.

Brandt, J. R., de Groot, P. J. V., Witt, K. E., Engelbrektsson, P. K., Helgen, K. M., Malhi, R. S., Ryder, O. A. & Roca, A. L. 2018. Genetic structure and diversity among historic and modern populations of the Sumatran rhinoceros (*Dicerorhinus sumatrensis*). *Journal of Heredity* 109: 553–565.

Brock, P. M., Döring, H. & Bidartondo, M. I. 2009. How to know unknown fungi: the role of a herbarium. *New Phytologist* 181: 719–724.

Bruns, T. D., Fogel, R. & Taylor, J. W. 1990. Amplification and sequencing of DNA from fungal herbarium specimens. *Mycologia* 82: 175–184.

Burrell, A. S., Disotell, T. R. & Bergey, C. M. 2015. The use of museum specimens with high-throughput DNA sequencers. *Journal of Human Evolution* 79: 35–44.

Cappellini, E., Gentry, A., Palkopoulou, E., Ishida, Y., Cram, D., Roos, A. M., Watson, M., Johansson, U. S., Fernholm, B., Agnelli, P., Barbagli, F., Littlewood, D. T. J., Kelstrup, C. D., Olsen, J. V., Lister, A. M., Roca, A. L., Dalen, L. & Gilbert, M. T. P. 2014. Resolution of the type material of the Asian elephant, *Elephas maximus* Linnaeus, 1758 (Proboscidea, Elephantidae). *Zoological Journal of the Linnean Society* 170: 222–232.

Chambers, E. A. & Hebert, P. D. N., 2016. Assessing DNA barcodes for species identification in North American reptiles and amphibians in natural history collections. *PlosONE* 11: e0154363.

Cooper, A. 1994. DNA from Museum Specimens. In: Herrmann, B. & Hummel, S. (eds), *Ancient DNA: Recovery and Analysis of Genetic Material from Paleontological*, *Archaeological*, *Museum*, *Medical*, *and Forensic Specimens*, pp. 149–165. Springer New York, New York, NY.

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 Jardin Botanico de Madrid* 66: 71–81.

Funk, V. A., Gostel, M., Devine, A., Kelloff, C. L., Wurdack, K., Tuccinardi, C., Radosavljevic, A., Peters, M. & Coddington, J. 2017. Guidelines for collecting vouchers and tissues intended for genomic work (Smithsonian Institution): Botany Best Practices. *Biodiversity Data Journal* 5: e11625.

Green, R. E., Krause, J., Ptak, S. E., Briggs, A. W., Ronan, M. T., Simons, J. F., Du, L., Egholm, M., Rothberg, J. M., Paunovic, M. & Pääbo, S. 2006. Analysis of one million base pairs of Neanderthal DNA. *Nature* 444: 330.

Grewe, F., Huang, J.-P., Leavitt, S. D. & Lumbsch, H. T. 2017. Reference-based RADseq resolves robust relationships among closely related species of lichen-forming fungi using metagenomic DNA. *Scientific Reports* 7(1): 9884.

Gueidan, C., Aptroot, A., Cáceres, M. E.d.S. & Binh, N. Q. 2015. Molecular phylogeny of the tropical lichen family Pyrenulaceae: contribution from dried herbarium specimens and FTA card samples. *Mycological Progress* 15: 7.

Hale, M. E. 1990. *A synopsis of the lichen genus Xanthoparmelia* (*Vainio*) *Hale* (*Ascomycotina*, *Parmeliaceae*). Smithsonian Institution Press, Washington D.C.

Hall, A. V. 1988. Pest control in herbaria. *Taxon* 37: 885–907.

Hawksworth, D. L. 2008. The need for a more effective biological nomenclature for the 21st century. *Botanical Journal of the Linnean Society* 109: 543–567.

Hawksworth, D. L. 2013. The oldest sequenced fungal specimen. *The Lichenologist* 45: 131–132.

Heckeberg, N. S., Erpenbeck, D., Worheide, G. & Rossner, G. E. 2016. Systematic relationships of five newly sequenced cervid species. *PeerJ* 4: e2307.

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. 2018. UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution* 35: 518–522.

Holmes, M. W., Hammond, T. T., Wogan, G. O., Walsh, R. E., LaBarbera, K., Wommack, E. A., Martins, F. M., Crawford, J. C., Mack, K. L., Bloch, L. M. & Nachman, M. W. 2016. Natural history collections as windows on evolutionary processes. *Molecular Ecology* 25: 864–881.

Hundsdoerfer, A. K., Packert, M., Kehlmaier, C., Strutzenberger, P. & Kitching, I. J. 2017. Museum archives revisited: Central Asiatic hawkmoths reveal exceptionally high late Pliocene species diversification (Lepidoptera, Sphingidae). *Zoologica Scripta* 46: 552–570.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., von Haeseler, A. & Jermiin, L. S. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* 14: 587.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. & Drummond, A. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649.

Keuler, R. A., Garretson, A., Saunders, T., Erickson, R., St. Andre, N., Grewe, F., Smith, H., Lumbsch, H., St Clair, L., Leavitt, S. 2019. Genome-scale data reveal the potential role of hybrid speciation in lichen-forming fungi. Botany 2019 – Sky Islands & Desert Seas, Tucson, Arizona, July 27–31. Botanical Society of America. Abstract ID: 190.

Kistenich, S., Halvorsen, R., Schrøder-Nielsen, A., Thorbek, L., Timdal, E. & Bendiksby, M. 2019. DNA sequencing historical lichen specimens. *Frontiers in Ecology and Evolution* 7: 5.

Lan, T. & Lindqvist, C. 2019. *Technical advances and challenges in genome-scale analysis of ancient DNA*. In: Lindqvist, C. & Rajora, O. P. (eds), Paleogenomics: Genome-Scale Analysis of Ancient DNA, pp. 3–29. Springer International Publishing, Cham.

Lane, M. A. 1996. Roles of natural history collections. *Annals of the Missouri Botanical Garden* 83: 536–545.

Langmead, B. & Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9: 357–359.

Leavitt, S. D., Fankhauser, J. D., Leavitt, D. H., Porter, L. D., Johnson, L. A. & St. Clair, L. L. 2011a. 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). *Molecular Phylogenetics and Evolution* 59: 587–602.

Leavitt, S. D., Johnson, L. A., Goward, T. & St. Clair, L. L. 2011b. Species delimitation in taxonomically difficult lichen-forming fungi: An example from morphologically and chemically diverse *Xanthoparmelia* (Parmeliaceae) in North America. *Molecular Phylogenetics and Evolution* 60: 317–332.

Leavitt, S. D., Fernández-Mendoza, F., Pérez-Ortega, S., Sohrabi, M., Divakar, P. K., Lumbsch, H. T. & St. Clair, L. L. 2013a. DNA barcode identification of lichen-forming fungal species in the *Rhizoplaca melanophthalma* species-complex (Lecanorales, Lecanoraceae), including five new species. *MycoKeys* 7: 1–22.

Leavitt, S. D., Fernández-Mendoza, F., Pérez-Ortega, S., Sohrabi, M., Divakar, P. K., Vondrák, J., Thorsten Lumbsch, H. & St. Clair, L. L. 2013b. Local representation of global diversity in a cosmopolitan lichen-forming fungal species complex (*Rhizoplaca*, Ascomycota). *Journal of Biogeography* 40: 1792–1806.

Leavitt, S. D., Lumbsch, H. T., Stenroos, S. & St. Clair, L. L. 2013c. 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: e85240.

Leavitt, S. D., Grewe, F., Widhelm, T., Muggia, L., Wray, B. & Lumbsch, H. T. 2016. Resolving evolutionary relationships in lichen-forming fungi using diverse phylogenomic datasets and analytical approaches. *Scientific Reports* 6: 22262

Leavitt, S. D., Kirika, P. M., Amo De Paz, G., Huang, J.-P., Hur, J.-S., Elix, J. A., Grewe, F., Divakar, P. K. & Lumbsch, H. T. 2018. Assessing phylogeny and historical biogeography of the largest genus of lichen-forming fungi, *Xanthoparmelia* (Parmeliaceae, Ascomycota). *The Lichenologist* 50: 299–312.

Lindahl, T. 1993. Instability and decay of the primary structure of DNA. *Nature* 362: 709–715.

Lofgren, L. A., Uehling, J. K., Branco, S., Bruns, T. D., Martin, F. & Kennedy, P. G. 2019. Genome-based estimates of fungal rDNA copy number variation across phylogenetic scales and ecological lifestyles. *Molecular Ecology*.

Lozier, J. D. & Cameron, S. A. 2009. Comparative genetic analyses of historical and contemporary collections highlight contrasting demographic histories for the bumble bees *Bombus pensylvanicus* and *B. impatiens* in Illinois. *Molecular Ecology* 18: 1875–1886.

McCormack, J. E., Tsai, W. L. & Faircloth, B. C. 2016. Sequence capture of ultraconserved elements from bird museum specimens. *Molecular Ecology Resources* 16: 1189–1203.

McCune, B. & Rosentreter, R. 2007. Biotic Soil Crust Lichens of the Columbia Basin. *Monographs in North American Lichenology* 1: 1–105.

McGuire, J. A., Cotoras, D. D., O'Connell, B., Lawalata, S. Z. S., Wang-Claypool, C. Y., Stubbs, A., Huang, X., Wogan, G. O. U., Hykin, S. M., Reilly, S. B., Bi, K., Riyanto, A., Arida, E., Smith, L. L., Milne, H., Streicher, J. W. & Iskandar, D. T. 2018. Squeezing water from a stone: high-throughput sequencing from a 145-year old holotype resolves (barely) a cryptic species problem in flying lizards. *PeerJ* 6: e4470.

Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. 2014. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32: 268–274.

Orange, A, James, P. W. & White, F. J. 2001. *Microchemical methods for the identification of lichens*. Twayne Publishers.

Pääbo, S. 1988. Ancient DNA: Extraction, characterization, molecular cloning, and enzymatic amplification. *Proceedings of the Nattional Academy of Sciences* 86: 1939–1943.

Paul, F., Otte, J., Schmitt, I. & Dal Grande, F. 2018. Comparing Sanger sequencing and high-throughput metabarcoding for inferring photobiont diversity in lichens. *Scientific Reports* 8: 8624.

Redchenko, O., Vondrák, J. & Košnar, J. 2012. The oldest sequenced fungal herbarium sample. *The Lichenologist* 44: 715–718.

Rittmeyer, E. N. & Austin, C. C. 2015. Combined next-generation sequencing and morphology reveal fine-scale speciation in Crocodile Skinks (Squamata: Scincidae: *Tribolonotus*). *Molecular Ecology* 24: 466–483.

Rowe, K. C., Singhal, S., Macmanes, M. D., Ayroles, J. F., Morelli, T. L., Rubidge, E.M, Bi, K. E. & Moritz, C. C. 2011. Museum genomics: low-cost and high-accuracy genetic data from historical specimens. *Molecular Ecology Resources* 11: 1082–1092.

Rozewicki, J., Yamada, K. D. & Katoh, K. 2017. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*: bbx108

Schilthuizen, M., Vairappan, C. S., Slade, E. M., Mann, D. J. & Miller, J. A. 2015. Specimens as primary data: museums and 'open science'. *Trends in Ecology and Evolution* 30: 237–238.

Schmid, S., Neuenschwander, S., Pitteloud, C., Heckel, G., Pajkovic, M., Arlettaz, R. & Alvarez, N. 2018. Spatial and temporal genetic dynamics of the grasshopper *Oedaleus decorus* revealed by museum genomics. *Ecology and Evolution* 8: 1480–1495.

Silva, P. C., Malabarba, M. C. & Malabarba, L. R. 2017. Using ancient DNA to unravel taxonomic puzzles: the identity of *Deuterodon pedri* (Ostariophysi: Characidae). *Neotropical Ichthyology* 15: e160141.

Sohrabi, M., Myllys, L. & Stenroos, S. 2010. Successful DNA sequencing of a 75 year-old herbarium specimen of *Aspicilia aschabadensis* (J. Steiner) Mereschk. *The Lichenologist* 42: 626–628.

Spurgin, L. G., Wright, D. J., van der Velde, M., Collar, N. J., Komdeur, J., Burke, T. & Richardson, D. S. 2014. Museum DNA reveals the demographic history of the endangered Seychelles warbler. *Evolution Applications* 7: 1134–1143.

Staats, M., Erkens, R. H. J., van de Vossenberg, B., Wieringa, J. J., Kraaijeveld, K., Stielow, B., Geml, J., Richardson, J. E. & Bakker, F. T. 2013. Genomic treasure troves: Complete genome sequencing of herbarium and insect museum specimens. *PlosONE* 8: e69189.

Tonini, J., Moore, A., Stern, D., Shcheglovitova, M. & Orti, G. 2015. Concatenation and species tree methods exhibit statistically indistinguishable accuracy under a range of simulated conditions. *PLoS Currents* 7.

Wen, J., Ickert-Bond, S. M., Appelhans, M. S., Dorr, L. J. & Funk, V. A. 2015. Collections-based systematics: Opportunities and outlook for 2050. J*ournal of Systematics and Evolution* 53: 477–488.

Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., DiCuccio, M., Edgar, R., Federhen, S., Geer, L. Y., Kapustin, Y., Khovayko, O., Landsman, D., Lipman, D. J., Madden, T. L., Maglott, D. R., Ostell, J., Miller, V., Pruitt, K. D., Schuler, G. D., Sequeira, E., Sherry, S. T., Sirotkin, K., Souvorov, A., Starchenko, G., Tatusov, R. L., Tatusova, T. A., Wagner, L. & Yaschenko, E. 2006. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*: gkl1031.

Wood, H. M., González, V. L., Lloyd, M., Coddington, J. & Scharff, N. 2018. Next-generation museum genomics: Phylogenetic relationships among palpimanoid spiders using sequence capture techniques (Araneae: Palpimanoidea). *Molecular Phylogenetics and Evolution* 127: 907–918.

Zeng, Q., Cui, Z., Wang, J., Childs, K. L., Sundin, G. W., Cooley, D. R., Yang, C. H., Garofalo, E., Eaton, A., Huntley, R. B., Yuan, X. & Schultes, N. P. 2018. Comparative genomics of Spiraeoideae-infecting *Erwinia amylovora* strains provides novel insight to genetic diversity and identifies the genetic basis of a low-virulence strain. *Molecular Plant Pathology* 19: 1652–1666.