



RESEARCH

Open Access



Using target enrichment sequencing to study the higher-level phylogeny of the largest lichen-forming fungi family: *Parmeliaceae* (Ascomycota)

Felix Grewe^{1*} , Claudio Ametrano¹, Todd J. Widhelm¹, Steven Leavitt², Isabel Distefano¹, Wetchasart Polyiam³, David Pizarro⁴, Mats Wedin⁵, Ana Crespo⁴, Pradeep K. Divakar⁴ and H. Thorsten Lumbsch¹

Abstract

Parmeliaceae is the largest family of lichen-forming fungi with a worldwide distribution. We used a target enrichment data set and a qualitative selection method for 250 out of 350 genes to infer the phylogeny of the major clades in this family including 81 taxa, with both subfamilies and all seven major clades previously recognized in the subfamily *Parmelioideae*. The reduced genome-scale data set was analyzed using concatenated-based Bayesian inference and two different Maximum Likelihood analyses, and a coalescent-based species tree method. The resulting topology was strongly supported with the majority of nodes being fully supported in all three concatenated-based analyses. The two subfamilies and each of the seven major clades in *Parmelioideae* were strongly supported as monophyletic. In addition, most backbone relationships in the topology were recovered with high nodal support. The genus *Parmotrema* was found to be polyphyletic and consequently, it is suggested to accept the genus *Crespoa* to accommodate the species previously placed in *Parmotrema* subgen. *Crespoa*. This study demonstrates the power of reduced genome-scale data sets to resolve phylogenetic relationships with high support. Due to lower costs, target enrichment methods provide a promising avenue for phylogenetic studies including larger taxonomic/specimen sampling than whole genome data would allow.

Keywords: Next-generation sequencing, Target capture, HybPiper, Phylogenomics, Maximum likelihood, Bayesian interference, ASTRAL, *Parmotrema*, *Parmelioideae*, *Protoparmelioideae*

INTRODUCTION

Our understanding of evolutionary relationships of fungi at all phylogenetic levels has dramatically improved with the availability of genetic data from entire genomes following remarkable progress in sequencing technologies (Ametrano et al. 2019; Ebersberger et al. 2012; Robbertse et al. 2006; Spatafora et al. 2017). In addition to sequencing complete genomes, a number of more cost-efficient

methods have been developed to sample subsets of genome-scale data. These include several direct sequencing approaches, such as restriction site associated DNA sequencing (RADseq) (Andrews et al. 2016), or capture sequencing approaches using baits, such as target enrichment of specific genes (Bragg et al. 2016) or ultra-conserved elements (Faircloth et al. 2012). These methods significantly reduce costs in comparison to sequencing entire genomes and thus will enable larger taxonomic or specimen sampling in comparative studies (Jones and Good 2016). RADseq has been used to address issues of delimitation and relationships of closely related ascomycete species (Bracewell et al. 2018; Grewe

* Correspondence: fgrewe@fieldmuseum.org

¹Science & Education, The Grainger Bioinformatics Center, Negaunee Integrative Research Center, Gantz Family Collections Center, and Pritzker Laboratory for Molecular Systematics, The Field Museum, 1400 S. Lake Shore Drive, Chicago, IL, USA

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

et al. 2017; Grewe et al. 2018; Salas-Lizana and Oono 2018) and population biology (Talas and McDonald 2015). However, studies have shown that this approach is most appropriate for shallow systematics due to issues with homology at deeper evolutionary scales when genome sequences are more diverged (Harvey et al. 2016; Rubin et al. 2012).

Target enrichment sequencing particularly enhances genomic regions of interest within a heterogeneous mixture of DNA samples (i.e. metagenomes). For target enrichment sequencing, pre-designed RNA probes are added to the metagenomic DNA extracts and capture their complementary DNA sequences through hybridization. Hybridization concentrates the DNA of the targeted genomic regions and allows for selective next generation sequencing of these regions (Bragg et al. 2016; Mamanova et al. 2010). Therefore, target enrichment sequencing is a cost-effective sequencing approach compared to whole genome sequencing, which can be used for large taxonomic sampling (Dapprich et al. 2016). So far, target enrichment sequencing has not been widely used in *Ascomycota*, with notable exceptions including a study screening for pathogenicity genes (Alshuwaili et al. 2018) and another study understanding the impact of ancient hybridization in the diversification of a clade of lichenized fungi (Widhelm et al. 2019).

Parmeliaceae is the most diverse group of lichenized fungi with about 2800 currently accepted species (Jaklitsch et al. 2016) that underwent an increased diversification associated with the aridification during the Oligocene-Miocene transition (Kraichak et al. 2015). Within the family *Parmeliaceae*, two subfamilies are distinguished, *Parmelioideae* and *Protoparmelioideae* (Divakar et al. 2017), with the vast majority of species diversity occurring in *Parmelioideae* (Divakar et al. 2017). The family currently includes 69 accepted genera (Divakar et al. 2017). Previous studies suggest that the family originated during the Cretaceous and subsequently diversified after the Cretaceous-Paleogene (K-Pg) boundary (Huang et al. 2019). Speciation within genera mostly happened during the Miocene (Lumbsch 2016). The species of the family occur worldwide on all kinds of substrate and in all terrestrial ecosystems but have their centres of diversity in the tropics and temperate, winter rain areas (Crespo et al. 2010; Thell et al. 2012).

Parmelioideae includes the bulk of species in *Parmeliaceae* and consists of a number of strongly supported monophyletic clades (Crespo et al. 2010; Crespo et al. 2007; Divakar et al. 2015). Currently, this includes seven strongly supported major clades, all of which are included in this study. The clades are 1) alectoroid, 2) anzioid, 3) cetrarioid, 4) hypogymnioid, 5) parmelioid, 6) psiloparmelioid, and 7) usneoid. While previous multi-gene studies have shown the monophyly of these clades, the relationships among those remained largely unsupported. Recently, we have used 2556 genes sampled from whole genome sequences of 44 in-group taxa to study

the evolutionary relationships among major clades in the subfamily *Parmelioideae* of *Parmeliaceae* (Pizarro et al. 2018). However, this study included a limited number of species due to the high costs of sequencing entire genomes and the computational burden of analysing thousands of genes (Pizarro et al. 2018). The smaller subfamily *Protoparmelioideae* currently includes three genera: the monotypic Australian *Maronina* Hafellner & R.W. Rogers, the pantropical *Neoprotoparmelia* Garima Singh, Lumbsch & I. Schmitt (dos Santos et al. 2019; Singh et al. 2018), which includes the majority of species previously included in *Maronina* s. lat., and the temperate *Protoparmelia* M. Choisy (Poelt and Grube 1992).

The main focus of this study was to assess the phylogenetic potential of cost-effective target enrichment sequencing approach using the hyper-diverse family *Parmeliaceae* as a model system. Our specific objectives were: 1) test our recent phylogenetic hypotheses based on multi-gene and whole genome data sets, 2) address phylogenetic relationships of major clades in *Parmelioideae*, and 3) test the power of target enrichment data sets to resolve phylogenetic relationships in ascomycetes, using *Parmeliaceae* as an example. We have augmented this taxon sampling to include a total of 81 in-group taxa including samples from both *Parmelioideae* and the other subfamily of *Parmeliaceae*, *Protoparmelioideae* (Divakar et al. 2017). We compare results from a target enrichment dataset with gene extraction methods from whole genome assemblies. Ultimately, we selected the best results of all gene extraction methods to produce a robust phylogenetic tree of all taxa.

MATERIALS AND METHODS

Taxon sampling

We included 81 representatives of lichen-forming fungal species from *Parmeliaceae* and five outgroup species in this phylogenomic study (Supplementary Table 1). Seventy-eight samples were selected to represent the seven major clades in subfamily *Parmelioideae* and three samples represented the subfamily *Protoparmelioideae* (Divakar et al. 2017; Divakar et al. 2015). Sequences from five additional species (*Arthonia rubrocincta* G. Merr. ex Lendemer & Grube, *Cladonia uncialis* (L.) Weber ex F.H. Wigg., *Lobaria pulmonaria* (L.) Hoffm., *Rhizoplaca melanophthalma* (DC.) Leuckert, and *Umbilicaria pustulata* (L.) Hoffm.) were selected as outgroups.

Target enrichment and sequencing

Baits design and target enrichment protocols were adopted from an earlier study (Widhelm et al. 2019). In short, 400 gene sequences were selected from the genome of *Lobaria pulmonaria* (available at JGI) and from a transcriptome assembly of *Evernia prunastri* [transcriptome sequence data published in (Meiser et al. 2017)].

The introns of the selected gene sequences were masked by aligning them with transcriptome assemblies of *Pseudovernia furfuracea* and *Lasallia pustulata* (Meiser et al. 2017), then all masked gene sequences were sent to Arbor Biosciences (Ann Arbor, MI, USA) for the RNA baits production.

The DNA of *Parmeliaceae* species was isolated with the ZR Fungal/Bacterial DNA MiniPrep according to the manufacturer's protocol (Zymo Research, Irvine, CA, USA). The isolated DNA was converted into libraries with the KAPA Hyper Prep Kit (KAPA Biosciences, Wilmington, MA, USA) using the Adapterama dual-indexing system to uniquely barcode all samples (Widhelm et al. 2019). A pool of all libraries was enriched for the 400 targeted genes by hybridization with the RNA baits. Enriched libraries were then paired-end sequenced with the Illumina MiSeq sequencer at the Field Museum's Pritzker Laboratory using the MiSeq 600-cycle sequencing kit version 3 (Illumina, San Diego, CA, USA).

In addition to target enrichment, we used sequences from nine *Parmeliaceae* species for which we had the whole genome sequenced: *Everniopsis trulla* (Ach.) Nyl., *Psiloparmelia denotata* Elix & T.H. Nash, *Usnea aurantiacoatra* (Jacq.) Bory, and six *Xanthoparmelia* species. For all species, DNA was isolated using the ZR Fungal/Bacterial DNA MiniPrep according to the manufacturer's protocol (Zymo Research, Irvine, CA, USA) – the same kit that was used to isolate DNA for target enrichment. Library construction and sequencing were done at the DNA services facility of the University of Illinois at Urbana-Champaign as described previously (Pizarro et al. 2018).

Gene extraction from target enrichment, whole genome sequencing data, and de novo assemblies

We used a combination of newly sequenced target enrichment data, newly sequenced whole genome data, and existing whole genome data for the target gene assembly and extraction. The Illumina MiSeq target enrichment sequencing result and Illumina whole genome sequencing results were demultiplexed in BaseSpace (Illumina) before being downloaded as raw reads to a local server. In addition, we downloaded all available whole genome sequencing data of earlier phylogenomic analyses (Abdel-Hameed et al. 2016; Greshake et al. 2016; Grewe et al. 2017; Grewe et al. 2018; Leavitt et al. 2016; McDonald et al. 2013; Meiser et al. 2017; Pizarro et al. 2018). All raw reads were trimmed with Trimmomatic v0.33 (Bolger et al. 2014), setting a quality threshold of 10 (LEADING:10 TRAILING:10) and a minimum read length of 25 bp (MINLEN:25). The surviving paired-end reads were used in HybPiper (Johnson et al. 2016) for gene assembly and extraction based on a target gene file. We generated this target gene file for HybPiper from the genome assembly of the *Parmeliaceae* species *P. furfuracea* (Meiser et al.

2017). We used BLASTn to search all gene models of the *P. furfuracea* genome with the bait sequence file as a query and identified 355 full-length genes. We extracted the complete exons of the identified *P. furfuracea* genes and used these sequences as a target gene file for all further gene extractions in this study. We translated the extracted nucleotide sequences into amino acid sequences since HybPiper did perform better using the BLASTx option, which uses an amino acid target gene file, instead of the BWA option, which uses a nucleotide target gene file. We then used the translated target gene file with the HybPiper wrapper script 'reads_first.py' on all trimmed Illumina sequence reads. HybPiper created one file folder for each species containing all assembly data and gene sequences, which we stored for further evaluation. All files that were derived from target enrichment data were tagged with "_target". Depending on the type of input sequence data, we named both methods as either "target enrichment method" or "whole genome method".

In exploratory analyses where samples prepared with the "whole genome method" recovered relatively fewer loci with the HybPiper pipeline, we subsequently assembled the whole genome sequence reads into de novo draft genomes prior to a target gene identification with Exonerate (Slater and Birney 2005) – we named this approach "de novo assembly method". De novo assemblies of the trimmed paired-end Illumina reads were constructed with SPAdes v3.5 or v3.11 (Bankevich et al. 2012). All assemblies were processed by the HybPiper script 'exonerate_hits.py' (part of the HybPiper wrapper) and the amino acid target gene file. All extracted target gene regions were transformed into the HybPiper wrapper output format with one file folder for each species and the different gene sequences separated in subfolders. The file names that resulted from the "de novo assembly method" were tagged with "_spades."

The combined output from "target enrichment," "whole genome," and "de novo assembly methods" was evaluated and visualized with the HybPiper scripts 'get_seq_lengths.py' and 'gene_recovery_heatmap.py'. The scripts provided an overview about sequencing and assembly success and allowed a direct comparison of the results of the "whole genome method" and the "de novo assembly method," which were both based on the same input data. We evaluated the sequencing success of each method by comparing the total number of recovered amino acids and the average length of recovered gene sequences. For comparative purposes, we visualized the results of the three methods in a heatmap using the script 'gene_recovery_heatmap.py'. We then used the script 'retrieve_sequences.py' to save the nucleotide and amino acid sequences of the best performing method for each species in a fasta file. We removed five fasta files of genes which contained only sequences of 50% or less of all taxa (i.e. less than 43 sequences), leaving 350 gene files for further analyses.

Selection of most informative gene sequences

We first aligned all 350 gene files with Mafft v7.450 (Kato and Standley 2013) and removed ambiguously aligned regions from each alignment using Gblocks v0.91 (Castresana 2000; Talavera and Castresana 2007) using relaxed parameter settings ($b2 = \text{half} + 1$, $b4 = 5$, $b5 = \text{half}$). We calculated the substitution rate for each gene with baseml implemented in PAML v4.9e (Yang 1997). To infer comparable substitution rates for all genes, we used the same constrained tree for each gene analysis. For the constrained tree, we concatenated all gene alignments with FASconCAT-G (Kueck and Longo 2014) and inferred the Maximum Likelihood tree with RAxML v8.2.11 (Stamatakis 2014) using the GTR + G model. We rooted the resulting tree with *Arthonia rubrocinta* and pruned it using the Python library DendroPy (Sukumaran and Holder 2010) to match the taxa of each gene alignment. The constrained tree was marked at 108 Ma for the origin of Parmeliaceae (Amo de Paz et al. 2011) prior to its use with each gene alignment for branch length and absolute substitution rate calculation in baseml (clock = 1; model = 7). Two genes showed extraordinarily high rates and hence were excluded. All other genes were sorted based on their substitution rate estimates from the slowest to the fastest evolving genes (Supplementary Table 2). We then built a concatenated alignment of the ten slowest evolving genes, then progressively increased the alignment stepwise with the next 10 faster evolving genes until 340 genes were included. We used all resulting 34 multigene alignments to calculate a Maximum Likelihood tree with RAxML v8.2.11 using the GTR + G model and the fast bootstrap option. We then calculated the average bootstrap value for each tree to identify the gene set that produced the best-supported phylogeny. The genes of this phylogeny were selected for all subsequent phylogenomic analyses. In addition, the remaining fast-evolving genes were used for separate phylogenomic analyses to identify potential conflicting phylogenetic signals.

Phylogenomic analyses

Evolutionary relationships were estimated from the concatenated alignment of 250 selected genes using Maximum Likelihood (ML) and Bayesian interference (BI). ML trees were estimated with the programs RAxML and IQ-TREE (Nguyen et al. 2015) using the GTR + G model for the nucleotide data set, which was partitioned by genes. For each RAxML analysis 100 bootstrap replicates and for each IQ-TREE analysis 1000 bootstrap replicates were calculated using the fast bootstrapping option implemented in RAxML and IQ-TREE, respectively. BI trees were calculated with MrBayes v3.2.6 (Huelsenbeck and Ronquist 2001; Ronquist et al. 2012) using the GTR + G model for the nucleotide data set, which was partitioned by genes. For this analysis,

two runs (four chains) with 1,500,000 iterations each were performed in parallel. Trees were sampled every 500 generations from the posterior distribution, and the first 25% of all sampled trees were discarded as the burn-in. We ensured convergence by a resulting ‘average standard deviation of split frequencies’ lower than 0.0000001 and ‘Effective Sample Size’ values higher than 200 in TRACER (Rambaut and Drummond 2007).

In addition to the concatenated-based phylogenies, we used coalescent-based methods to estimate a species tree given the individual gene trees of the 250 selected genes. All selected gene sequences were aligned with Mafft v7.450 (Kato and Standley 2013) and trimmed with Gblocks v0.91 using relaxed parameter settings ($b2 = \text{half} + 1$, $b4 = 5$, $b5 = \text{half}$). ML trees of each gene were calculated with IQ-TREE using the GTR + G model. For each analysis 1000 bootstrap replicates were calculated using the fast bootstrapping option implemented in IQ-TREE. We contracted low support branches (bootstrap < 20) of all individual trees using Newick Utilities (Junier and Zdobnov 2010) before using all trees as input for a coalescent-based species tree estimation with ASTRAL-III (Zhang et al. 2018).

In addition to the selected 250 slow-evolving genes, we concatenated all remaining fast-evolving gene sequences to estimate their phylogenetic relationship. The concatenated matrix was partitioned by genes and used for ML analyses with the programs RAxML and IQ-Tree using the GTR + G model. For each RAxML analysis 100 bootstrap replicates and for each IQ-TREE analysis 1000 bootstrap replicates were calculated using the fast bootstrapping option implemented in RAxML and IQ-TREE, respectively. All resulting phylogenetic trees were drawn with the program FigTree v1.4.2 (Rambaut 2009).

Availability of data and material

The sequences produced in this paper have been deposited in the NCBI Sequence Read Archive (SRA) with the accession numbers SRR13125477, SRR1315498, SRR13125762, SRR13167197, SRR13125985, SRR13126647, SRR13126796, SRR13126828, SRR13126859. Target enrichment gene set and multiple sequence alignments are available at FigShare DOI:<https://doi.org/10.6084/m9.figshare.13238453>.

RESULTS AND DISCUSSION

Phylogenomic data

The three different methods that were used to recover the targeted genes in Parmeliaceae executed with different success. While the “target enrichment method” used RNA baits for target capturing prior to sequencing, the “whole genome method” and “de novo assembly method” used the same whole genome sequence data either as raw data input or assembled into draft genomes to capture the targeted genes. The average length of all

recovered target genes was 82,196 (SD = 53,643) amino acids for the “whole genome method” compared to 138,069 (SD = 10,845) amino acids for the “de novo assembly method” (Supplementary Table 3). Therefore, the recovery of the target genes was comparatively improved with the “de novo assembly method,” in particular for species with low sequence coverage. The “target enrichment method” recovered a total of 123,788 (SD = 9578) amino acids, which performed better than the “whole genome method”, but the “de novo assembly method” recovered on average slightly more amino acids. The strength of each method is also visible in the heatmap showing the recovery of all genes for each species (Supplementary Figure 1). The “de novo assembly method” was more successful than the “whole genome method,” which performed weakly for many species and even completely failed for some species. The heatmap further indicated that the “target enrichment method” failed for some gene sequences (30 genes were recovered by only half of their length or less). Since many of these genes were successfully recovered with the other two methods using whole genome sequencing data, it might be a consequence of the target enrichment procedure and indicate that some baits failed to capture the Parmeliaceae gene sequences.

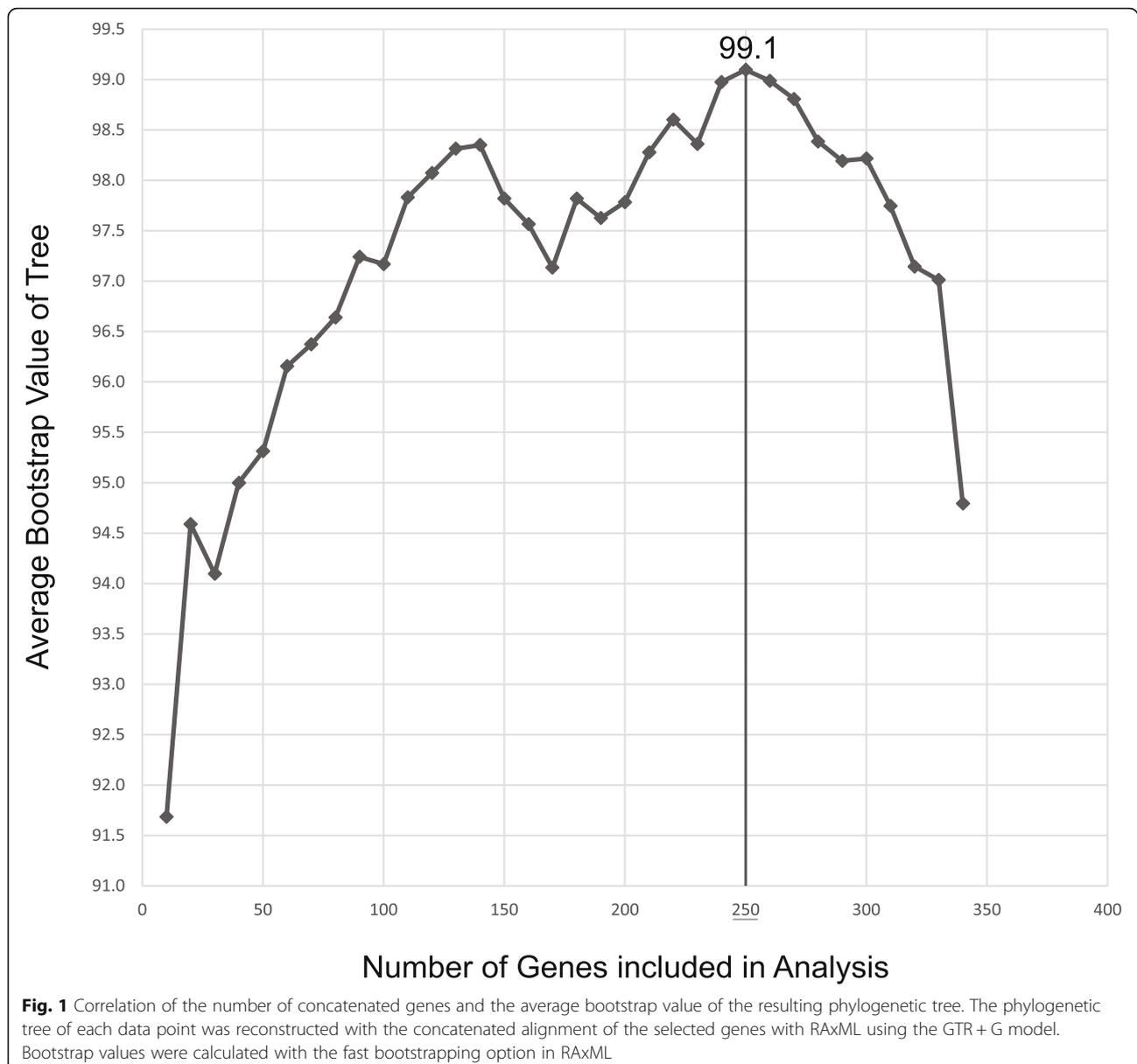
The most informative subset of genes for phylogenomic analyses was identified by the maximum number of slowest evolving genes that reconstructed the best supported phylogenetic tree. We created 34 multigene matrices that included concatenated alignments from 10 to 340 genes. All resulting phylogenetic trees had average bootstrap values above 91 (Fig. 1). Including faster evolving genes in the multigene matrix had an overall positive effect on the phylogenetic tree reconstruction until we reached 250 genes, with a maximum average bootstrap value at 99.1. Remarkably, after 250 genes, the addition of more (faster evolving) genes to the multigene matrix decreased the average bootstrap of the trees. Therefore, we selected the 250 genes that achieved the maximum average bootstrap value for all subsequent phylogenetic inferences. The final nucleotide multigene matrix had a dimension of 86 taxa and 270,160 characters and 178,588 distinct alignment patterns with 3.37% of undetermined characters or gaps.

Phylogenetic relationships

Phylogenies inferred from the concatenated target enrichment dataset under Maximum Likelihood and Bayesian inference recovered identical topologies and hence only the IQ-TREE tree is shown here (Fig. 2). The coalescent-based ASTRAL-III tree conflicted in one node altering the placement of the *Coelopogon/Menegazzia* clade (Fig. 2, Supplementary Figure 2). All nodes in the Bayesian analysis received a posterior probability of 1.0, whereas five nodes in the IQ-TREE analysis and

nine nodes in the RaxML analysis received bootstrap support of less than 100% (Fig. 2, Supplementary Figure 2). However, with the exception of the low supported sister-group relationships of *Arctoparmelia* and *Pseudevernia* (86% in the IQ-TREE and 75% in the RAXML analyses), of a clade consisting of the genera *Notoparmelia* + *Parmelia* and a clade including the genera *Bulborrhizina*, *Bulbothrix*, *Parmelinella*, and *Parmelina* (89% in the RAXML tree), and of *Nephromopsis chlorophylla* and *N. cucullata* (67% in the RAXML analysis), the bootstrap support for those was above 95% and is considered here to be strong. The monophyly of the two subfamilies Protoparmelioideae and Parmelioideae was strongly supported. The phylogenetic tree of all remaining fast-evolving genes resulted in a similar topology as the tree of the 250 slow-evolving genes, with the exception of a strongly supported cluster of *Omphalora arizonica*, *Everniopsis trulla*, *Psiloparmelia denotata*, *Oropogon secalonicus*, and *Platismatia glauca* (Supplementary Figure 2). Therefore, some of the fast-evolving genes may reflect a different evolutionary history of these taxa than the 250 selected genes. In addition, many relationships in the phylogenetic inference of the fast-evolving genes were either unresolved or less supported indicated by poor bootstrap support. The different evolutionary history and the lack of phylogenetic signal of some fast-evolving genes may have caused the decrease of the average bootstrap support when more fast-evolving genes were included to the 250-gene phylogeny (Fig. 1).

The overall phylogeny of the 250 selected genes was similar to topologies inferred from a multi-gene data set with a larger taxon sampling (274 ingroup taxa) (Divakar et al. 2015) and a genomic data set with a smaller taxon sampling (44 ingroup taxa) (Pizarro et al. 2018). Here we focus on describing the phylogenetic relationships of major clades and differences to either of these previous studies. The phylogenetic position of the *Coelopogon/Menegazzia* clade was unresolved in previous studies, being either sister to the usneoid clade without support (Divakar et al. 2015), sister to the alectorioid clade, or sister to all remaining *Parmelioideae* – again without support (Pizarro et al. 2018). In our concatenated-based analyses, the predominantly southern Hemisphere *Coelopogon/Menegazzia* clade was an early divergent lineage within the subfamily *Parmelioideae*, as it formed a strongly supported sister-group relationship with all remaining *Parmelioideae* (Fig. 2, Supplementary Figure 2). However, in the coalescent-based analysis, the *Coelopogon/Menegazzia* clade was sister to the alectorioid clade (Fig. 2, Supplementary Figure 2). An inconsistent placement of *Menegazzia* depending on the use of concatenated or coalescent-based analyses was also documented in an earlier study on *Parmeliaceae* (Pizarro et al. 2018). In our study, we added *Coelopogon* as a sister taxon



to *Menegazzia* to resolve the placement of *Coelopogon*/*Menegazzia*; however, the phylogenetic position of the two taxa remained conflicting between the two analyses. Concatenation-based analyses have been shown to overestimate phylogenetic relationships in large datasets (Edwards et al. 2007). Furthermore, coalescent-based analyses can be inconsistent under gene flow and incomplete lineage sorting (Solis-Lemus and Ane 2016; Solis-Lemus et al. 2016) and when individual gene trees are erroneous due to evolutionary rate heterogeneity (Koch et al. 2018). In this study, we particularly selected slow-evolving genes, which reduces the evolutionary rate heterogeneity of the used genes. However, the position of *Coelopogon*/*Menegazzia*

remained in conflict whether coalescent-based or concatenated-based analyses were used (Fig. 2).

The alectorioid clade was resolved as monophyletic and was sister to the remaining *Parmelioideae*, which agreed with the placement in the ML tree of the genomic analysis of Pizarro et al. (2018). In contrast to Krog's hypothesis (Krog 1982), *Letharia* and *Lethariella* did not form a sister-group relationship in our study (Fig. 2). The genus *Letharia* was only included in Divakar et al. (2015) but its phylogenetic relationships remained unresolved. In our study, the two *Letharia* species formed a strongly supported sister-group to the usneoid clade, which was also strongly supported as

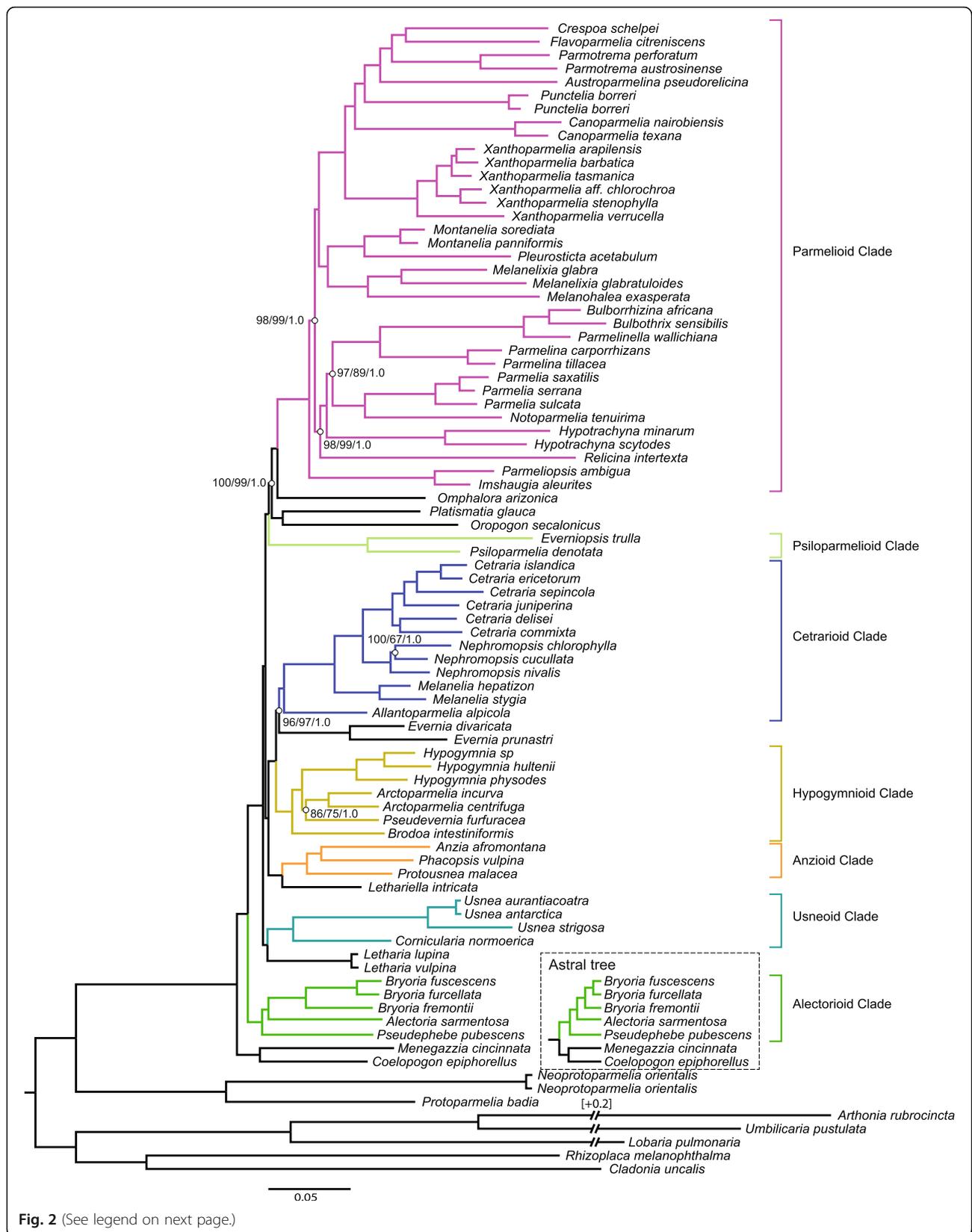


Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 Phylogenetic relationships among major lineages of Parmeliaceae, represented by 81 specimens. The tree shown was generated by Maximum Likelihood inference using IQ-TREE of a data set containing the 250 most phylogenetic informative genes of the target enrichment gene set. Additional trees generated by Maximum Likelihood using RAxML and Bayesian inference using MrBayes resulted with the same topology. All nodes of the three trees received 100% bootstrap support (BS) with IQ-TREE and RAxML or 1.0 posterior probability (PP) with MrBayes unless highlighted by an open circle and the respective support values (IQ-TREE BS/RAxML BS/MrBayes PP). Major clades of Parmeliaceae are highlighted by different colors. The dotted box highlights the conflict between the concatenated-based trees (IQ-TREE, RAxML, MrBayes) and coalescent-based tree (ASTRAL-III). The unit of branch length is substitutions per site

monophyletic. Our results are in accordance with the phenotypic similarities of *Letharia* with usneoid lichens (Crespo et al. 2007; Krog 1982). The anzioid clade formed a strongly supported sister-group with *Lethariella intricata* (Moris) Krog, the type species of the genus. Previously, *Lethariella* clustered with *Letharia* (Divakar et al. 2015) but in that study, two other species of the genus, *L. cashmeriana* Krog and *L. togashii* (Asahina) Krog were included and which could indicate that the genus is not monophyletic. The anzioid clade+*Lethariella* formed a strongly supported sister-group relationship to a clade including the hypogymnioid and cetrarioid clades and the genus *Evernia*. The relationships of *Evernia* and the cetrarioid clade was also strongly supported in a previous study (Pizarro et al. 2018) but lacked support in another (Divakar et al. 2015). Also, the strongly supported relationships of the hypogymnioid clade with the cetrarioid clade+*Evernia* has been found previously (Pizarro et al. 2018). The psiloparmelioid clade that was not included in Pizarro et al. (2018) had an unsupported placement in Divakar et al. (2015). In our study it formed a strongly supported sister-group to a clade that includes the parmelioid clade (including the early diverging clade consisting of *Imshaugia* and *Parmeliopsis*), the monotypic genus *Omphalora*, and a clade including *Oropogon* and *Platismatia*. This agrees with the tree topology in Pizarro et al. (2018) but it lacked support in that analysis. All genera for which more than one species was included formed strongly supported, monophyletic groups with the exception of the genus *Parmotrema*. *Parmotrema schelpei* (Hale) D. Hawksw., which was placed in the subgenus *Crespoa* by some authors (Hawksworth 2011; Kirika et al. 2016) does not cluster with the two other species of *Parmotrema* included in this study. This supports the segregation of the clade at the generic level and the acceptance of the genus *Crespoa* (Lendemer and Hodkinson 2012). In previous studies the relationship of *Pleurosticta* was unresolved (Crespo et al. 2010; Divakar et al. 2012) and it was grouped with species corresponding to the genus *Montanelia* (Crespo et al. 2010; Divakar et al. 2012; Leavitt et al. 2015) In our study, *Pleurosticta* formed an independent lineage that was the strongly supported as sister-group to *Montanelia*.

Utility of target capture data sets to resolve phylogenetic relationships

The *Parmeliaceae* phylogenies include 35 taxa that were sequenced by target enrichment, together with 46 other ingroup taxa for these whole genomes were sequenced. A comparison of target enrichment to whole-genome sequencing or multi-gene sequencing of earlier *Parmeliaceae* studies revealed the limits of both whole-genome and multi-gene methods. Multi-gene phylogenies are affordable and allow the inclusion of many taxa, but they can be limited on the amount of phylogenetic information (Crespo et al. 2007; Divakar et al. 2017; Divakar et al. 2015). In comparison, whole genomes provide thousands of gene sequences for phylogenetic analyses, but the high sequencing cost and computational burden can limit the number of taxa (Pizarro et al. 2018). Target enrichment sequencing overcomes the limits of both whole-genome and multi-gene sequencing methods, since it allows for the affordable sequencing of hundreds of genes of multiple taxa, all can be pooled together in one sequencing run. This method specifically enriches targeted gene sequences prior to next-generation sequencing, which reduces sequencing costs and has the additional benefit of sorting symbiotic metagenomes, such as whole lichen DNA isolates. The enrichment of genes also works well when only low amounts of DNA are available and therefore qualifies as a method for rare species and older herbarium specimens, for which traditional molecular methods would fail. As a reduced genome representation method, target enrichment usually recovers fewer genes than whole genome sequencing, however a well-chosen gene set might improve phylogenies and outperform large genomic datasets, as shown by the selection of the most phylogenetically informative 250 genes out of 350 genes to reconstruct a remarkably well supported phylogeny for *Parmeliaceae* (Fig. 1). In the only other phylogenomic study using target enrichment data in ascomycetes so far, patterns consistent with ancient hybridization could be detected (Widhelm et al. 2019). This demonstrated that, in addition to reconstructing phylogenies, these data sets are also powerful in identifying historical processes shaping diversity of organisms.

The target enrichment method can reach its limits when the baits sequences differ too much from the target

sequences, hence different baits sets are required for diverse taxon samplings. To overcome this limitation, a universal bait set of 353 conserved flowering plant genes was developed and is publicly available for phylogenomic analyses of flowering plants (Brewer et al. 2019; Johnson et al. 2019). A similar universal bait set could be optimized for fungal groups and take over as the next generation of phylogenetic marker genes. Hence, target enrichment sequencing provides a powerful novel avenue with great potential for the future of fungal phylogenomic research.

CONCLUSION

For phylogenomic analyses, target enrichment sequencing represents an effective and inexpensive alternative for sequencing hundreds of genes when compared to other methods such as whole genome sequencing. We used target enrichment sequencing to generate data for the phylogenetic reconstruction of the largest family of lichen-forming fungi: *Parmeliaceae*. All sequenced genes were filtered for the 250 most informative genes. These genes were implemented in coalescent-based and concatenated-based phylogenetic applications, which estimated highly-supported phylogenetic trees. The trees of both methods differed in the placement of *Menegazzia*, potentially due to misleading effects of incomplete lineage sorting or evolutionary rate heterogeneity. While future phylogenomic methods might help to address these issues in phylogenetic reconstructions of large datasets, this study highlights the advantages that target enrichment has for fungal research. The inexpensive and straightforward target enrichment sequencing approach will open new possibilities for fungal researchers to incorporate the use of large genomic data sets into their research.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s43008-020-00051-x>.

Additional file 1: Supplementary Figure 1. Heatmap summarizing the results of target gene recovery. Each field represents the percentage of recovered amino acids of every gene for a respective taxon. All taxa were sorted by the methods that was used for the gene recovery.

Additional file 2: Supplementary Figure 2. Phylogenetic relationships among major lineages of *Parmeliaceae*. Four trees were generated by Maximum Likelihood inference using a) IQ-TREE, b) RAXML, c) Bayesian interference using MrBayes, or d) a coalescent-based species tree calculation of a data set containing the 250 most phylogenetic informative genes of the target enrichment gene set. In addition, two trees were generated by Maximum Likelihood inference using e) IQ-Tree and f) RAXML of data containing the 89 fastest evolving genes. Numbers at tree branches represents IQ-TREE bootstrap, RAXML bootstrap, or MrBayes posterior probability values, respectively. The unit of branch lengths of the IQ-TREE, RAXML, and MrBayes trees is substitutions per site. Branch lengths of the tree generated by ASTRAL-III are in coalescent units.

Additional file 3: Supplementary Table 1. Collection details of taxa used in this study.

Additional file 4: Supplementary Table 2. Absolute substitution rates calculated by baseml using the GTR model; the unit of rates is substitutions per site per million years.

Additional file 5: Supplementary Table 3. Results of target gene recovery using the "target enrichment method", the "whole genome method", or the "de novo assembly method".

Acknowledgements

We thank our colleagues Mikhail Andreev, Carlos Boluda, John A. Elix, Bernard Goffinet, Trevor Goward, Paul Kirika, Elisa Lagostina, Christian Printzen, Joana Vilagrara, and Janet Voight for providing samples for this study and Sarah Cowles for comments on the manuscript.

Adherence to national and international regulations

Not applicable.

Authors' contributions

FG, PKD, and HTL designed the research project; TJW and ID performed the molecular lab work; FG, CA, and HTL analysed the data; SL, WP, DP, MW, AC, and PKD collected samples; FG, CA, and HTL wrote the paper. All authors read and approved the final manuscript.

Funding

The project was financially supported by the Spanish Ministerio de Ciencia e Innovación (PID2019-105312GB-I00), the Santander-Universidad Complutense de Madrid (PR87/19-22637), the Swedish Research Council VR (VR2016-03589), and The Grainger Foundation through The Grainger Bioinformatics Center.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Science & Education, The Grainger Bioinformatics Center, Negaunee Integrative Research Center, Gantz Family Collections Center, and Pritzker Laboratory for Molecular Systematics, The Field Museum, 1400 S. Lake Shore Drive, Chicago, IL, USA. ²Department of Biology and M. L. Bean Life Science Museum, Brigham Young University, Provo, UT, USA. ³Lichen Research Unit, Biology Department, Faculty of Science, Ramkhamhaeng University, Ramkhamhaeng 24 Road, Bangkok 10240, Thailand. ⁴Departamento de Farmacología, Farmacognosia y Botánica, Facultad de Farmacia, Universidad Complutense de Madrid, 28040 Madrid, Spain. ⁵Department of Botany, Swedish Museum of Natural History, PO Box 50007, SE-104 05 Stockholm, Sweden.

Received: 24 August 2020 Accepted: 29 November 2020

Published online: 14 December 2020

References

- Abdel-Hameed M, Bertrand RL, Piercey-Normore MD, Sorensen JL (2016) Putative identification of the usnic acid biosynthetic gene cluster by de novo whole-genome sequencing of a lichen-forming fungus. *Fungal Biology* 120:306–316
- Alshuwaili F, Zaccaron ML, Sharma S, Bluhm BH (2018) A genetic screen for pathogenicity genes in the fungus *Diaporthe longicolla* causing *Phomopsis* seed decay of soybean. *Phytopathology* 108(12):S1–S1
- Ametrano CG, Grewe F, Crous PW, Goodwin SB, Liang C, Selbmann L, Lumbsch HT, Leavitt SD, Muggia L (2019) Genome-scale data resolve ancestral rock-inhabiting lifestyle in Dothideomycetes (Ascomycota). *IMA Fungus* 10:19. <https://doi.org/10.1186/s43008-019-0018-2>
- 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(12):e2816. <https://doi.org/10.1371/journal.pone.0028161>
- Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA (2016) Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* 17(2):81–92. <https://doi.org/10.1038/nrg.2015.28>

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Son P, Pribliski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19(5):455–477. <https://doi.org/10.1089/cmb.2012.0021>
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bracewell RR, Vanderpool D, Good JM, Six DL (2018) Cascading speciation among mutualists and antagonists in a tree - beetle - fungi interaction. *Proceedings of the Royal Society B-Biological Sciences* 285(1881). <https://doi.org/10.1098/rspb.2018.0694>
- Bragg JG, Potter S, Bi K, Moritz C (2016) Exon capture phylogenomics: efficacy across scales of divergence. *Molecular Ecology Resources* 16(5):1059–1068. <https://doi.org/10.1111/1755-0998.12449>
- Brewer GE, Clarkson JJ, Maurin O, Zuntini AR, Barber V, Bellot S, Biggs N, Cowan RS, Davies NMJ, Dodsworth S, Edwards SL, Eiserhardt WL, Epitawalage N, Frisby S, Grall A, Kersey PJ, Pokorny L, Leitch IJ, Forest F, Baker WJ (2019) Factors affecting targeted sequencing of 353 nuclear genes from herbarium specimens spanning the diversity of angiosperms. *Frontiers in Plant Science* 10:1102. <https://doi.org/10.3389/fpls.2019.01102>
- Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* 17(4):540–552
- Crespo A, Kauff F, Divakar PK, Amo G, Arguello A, Blanco O, Cubas P, del Prado R, Elix JA, Esslinger TL, Ferencova Z, Hawksworth DL, Lutzoni F, Millanes AM, Molina MC, Perez-Ortega S, Wedin M, Ahti T, Bungartz F, Calvelo S, Aptroot A, Barreno E, Candan M, Cole M, Ertz D, Goffinet B, Lindblom L, Lücking R, Mattsson JE, Messuti M, Miadlikowska J, Piercey-Normore M, Rico V, Sipman HJM, Schmitt I, Spribille T, Thell A, Thor G, Lumbsch HT (2010) Phylogenetic generic classification of parmelioid lichens (Parmeliaceae, Ascomycota) based on molecular, morphological and chemical evidence. *Taxon* 59:1735–1753
- Crespo A, Lumbsch HT, Mattsson JE, Blanco O, Divakar PK, Articus K, Wiklund E, Bawingan PA, Wedin M (2007) Testing morphology-based hypotheses of phylogenetic relationships in Parmeliaceae (Ascomycota) using three ribosomal markers and the nuclear RPB1 gene. *Molecular Phylogenetics and Evolution* 44(2):812–824
- Dapprich J, Ferriola D, Mackiewicz K, Clark PM, Rappaport E, D'Arcy M, Sasson A, Gai XW, Schug J, Kaestner KH, Monos D (2016) The next generation of target capture technologies - large DNA fragment enrichment and sequencing determines regional genomic variation of high complexity. *BMC Genomics* 17:486. <https://doi.org/10.1186/s12864-016-2836-6>
- Divakar PK, Crespo A, Kraichak E, Leavitt SD, Singh G, Schmitt I, Lumbsch HT (2017) Using a temporal phylogenetic method to harmonize family- and genus-level classification in the largest clade of lichen-forming fungi. *Fungal Diversity* 84:101–117
- Divakar PK, Crespo A, Wedin M, Leavitt SD, Hawksworth DL, Myllys L, McCune B, Randlane T, Bjerke JW, Ohmura Y, Schmitt I, Boluda CG, Alors D, Roca-Valiente B, Del-Prado R, Ruihal C, Buaruang K, Núñez-Zapata J, Amo de Paz G, Rico VJ, Molina MC, Elix JA, Esslinger TL, Tronstad IKK, Lindgren H, Ertz D, Gueidan C, Saag L, Mark K, Singh G, Dal Grande F, Parmmen S, Beck A, Benatti MN, Blanchon D, Candan M, Clerc P, Goward T, Grube M, Hodkinson BP, Hur J-S, Kantvilas G, Kirika PM, Lendemer J, Mattsson J-E, Messuti M, Miadlikowska J, Nelsen M, Ohlson JI, Pérez-Ortega S, Saag A, Sipman HJM, Sohrabi M, Thell A, Thor G, Truong C, Yahr R, Upreti DK, Cubas P, Lumbsch HT (2015) Evolution of complex symbiotic relationships in a morphologically derived family of lichen-forming fungi. *New Phytologist* 208(4):1217–1226. <https://doi.org/10.1111/nph.13553>
- Divakar PK, Del-Prado R, Lumbsch HT, Wedin M, Esslinger TL, Leavitt SD, Crespo A (2012) Diversification of the newly recognized lichen forming fungal lineage *Montanelia* (Parmeliaceae, Ascomycota) and its relation to key geological and climatic events. *American Journal of Botany* 99:2014–2026
- dos Santos LA, Aptroot A, Lücking R, da Silva Caceres ME (2019) High diversification in the *Neoprotomarmelia multifera* complex (Ascomycota, Parmeliaceae) in Northeast Brazil revealed by DNA barcoding and phenotypical characters. *Bryologist* 122(4):539–552. <https://doi.org/10.1639/0007-2745-122.4.539>
- Ebersberger I, Simoes RM, Kucyzok A, Gube M, Kothe E, Voigt K, von Haeseler A (2012) A consistent phylogenetic backbone for the fungi. *Molecular Biology and Evolution* 29(5):1319–1334. <https://doi.org/10.1093/molbev/msr285>
- Edwards SV, Liu L, Pearl DK (2007) High-resolution species trees without concatenation. *Proceedings of the National Academy of Sciences, USA* 104:5936–5941
- Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC (2012) Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology* 61(5):717–726. <https://doi.org/10.1093/sysbio/sys004>
- Greshake B, Zehr S, Dal Grande F, Meiser A, Schmitt I, Ebersberger I (2016) Potential and pitfalls of eukaryotic metagenome skimming: a test case for lichens. *Molecular Ecology Resources* 16(2):511–523. <https://doi.org/10.1111/1755-0998.12463>
- Grewe F, Huang JP, Leavitt SD, Lumbsch HT (2017) Reference-based RADseq resolves robust relationships among closely related species of lichen-forming fungi using metagenomic DNA. *Scientific Reports* 7:9884
- Grewe F, Lagostina E, Wu H, Printzen C, Lumbsch HT (2018) Population genomic analyses of RAD sequences resolves the phylogenetic relationship of the lichen-forming fungal species *Usnea antarctica* and *Usnea aurantiacoatra*. *Myckeys* 43:91–113. <https://doi.org/10.3897/myckeys.43.29093>
- Harvey MG, Smith BT, Glenn TC, Faircloth BC, Brumfield RT (2016) Sequence capture versus restriction site associated DNA sequencing for shallow systematics. *Systematic Biology* 65(5):910–924. <https://doi.org/10.1093/sysbio/syw036>
- Hawksworth DL (2011) *Parmotrema* subgen. *Crespoa* subgen. nov. for the *Canoparmelia crozalsiana* clade. *Lichenologist* 43(4):647–648
- Huang J-P, Kraichak E, Leavitt SD, Nelsen MP, Lumbsch HT (2019) Accelerated diversifications in three diverse families of morphologically complex lichen-forming fungi link to major historical events. *Scientific Reports* 9:8518. <https://doi.org/10.1038/s41598-019-44881-1>
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17(8):754–755
- Jaklitsch WM, Baral HO, Lücking R, Lumbsch HT (2016) Ascomycota. In: Frey W (ed) *Syllabus of plant families - Adolf Engler's syllabus der Pflanzenfamilien*, vol 1/2, 13th edn. Gebr. Borntraeger Verlagsbuchhandlung, Stuttgart, pp 1–150
- Johnson MG, Gardner EM, Liu Y, Medina R, Goffinet B, Shaw AJ, Zerega NJC, Wickett NJ (2016) HYBIPIPER: extracting coding sequences and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences* 4(7). <https://doi.org/10.3732/app.1600016>
- Johnson MG, Pokorny L, Dodsworth S, Botigue LR, Cowan RS, Devault A, Eiserhardt WL, Epitawalage N, Forest F, Kim JT, Leebens-Mack JH, Leitch IJ, Maurin O, Soltis DE, Soltis PS, Wong GKS, Baker WJ, Wickett NJ (2019) A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic Biology* 68(4):594–606. <https://doi.org/10.1093/sysbio/syy086>
- Jones MR, Good JM (2016) Targeted capture in evolutionary and ecological genomics. *Molecular Ecology* 25(1):185–202. <https://doi.org/10.1111/mec.13304>
- Junier T, Zdobnov EM (2010) The Newick utilities: high-throughput phylogenetic tree processing in the Unix shell. *Bioinformatics* 26(13):1669–1670. <https://doi.org/10.1093/bioinformatics/btq243>
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30(4):772–780
- Kirika P, Divakar PK, Crespo A, Leavitt SD, Mugambi GK, Gatheri GW, Lumbsch HT (2016) Polyphyly of the genus *Canoparmelia* - another example of incongruence between phenotype-based classification and molecular phylogeny within lichenized Ascomycota (Parmeliaceae). *Phytotaxa* 289:36–48
- Koch NM, Coppard SE, Lessios HA, Briggs DEG, Mooi R, Rouse GW (2018) A phylogenomic resolution of the sea urchin tree of life. *BMC Evolutionary Biology* 18:189. <https://doi.org/10.1186/s12862-018-1300-4>
- Kraichak E, Divakar PK, Crespo A, Leavitt SD, Nelsen MP, Lücking R, Lumbsch HT (2015) A tale of two hyper-diversities: diversification dynamics of the two largest families of lichenized fungi. *Scientific Reports* 5:e10028
- Krog H (1982) Evolutionary trends in foliose and fruticose lichens of the Parmeliaceae. *Journal of the Hattori Botanical Laboratory* 52:303–311
- Kueck P, Longo GC (2014) FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Frontiers in Zoology* 11:81. <https://doi.org/10.1186/s12983-014-0081-x>
- Leavitt SD, Divakar PK, Ohmura Y, Wang L-S, Esslinger TL, Lumbsch HT (2015) Who's getting around? Assessing species diversity and phylogeography in the widely distributed lichen-forming fungal genus *Montanelia* (Parmeliaceae, Ascomycota). *Molecular Phylogenetics and Evolution* 90:85–96. <https://doi.org/10.1016/j.ympev.2015.04.029>

- Leavitt SD, Grewe F, Widhelm T, Muggia L, Wray B, Lumbsch HT (2016) Resolving evolutionary relationships in lichen-forming fungi using diverse phylogenomic datasets and analytical approaches. *Scientific Reports* 6:22262
- Lendemer JC, Hodkinson BP (2012) Recognition of the *Parmelia crozalsiana* group as the genus *Crespoa*. *North American Fungi* 7(2):1–5
- Lumbsch HT (2016) Lichen-forming Fungi, diversification of. In: Kliman RM (ed) *Encyclopedia of evolutionary biology*, vol 2. Academic Press, Oxford, pp 305–311
- Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J, Turner DJ (2010) Target-enrichment strategies for next-generation sequencing. *Nature Methods* 7(2):111–118. <https://doi.org/10.1038/nmeth.1419>
- McDonald TR, Mueller O, Dietrich FS, Lutzoni F (2013) High-throughput genome sequencing of lichenizing fungi to assess gene loss in the ammonium transporter/ammonia permease gene family. *BMC Genomics* 14:225. <https://doi.org/10.1186/1471-2164-14-225>
- Meiser A, Otte J, Schmitt I, Grande FD (2017) Sequencing genomes from mixed DNA samples - evaluating the metagenome skimming approach in lichenized fungi. *Scientific Reports* 7:14881
- Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32(1):268–274. <https://doi.org/10.1093/molbev/msu300>
- Pizarro D, Divakar PK, Grewe F, Leavitt SD, Huang J-P, Dal Grande F, Schmitt I, Wedin M, Crespo A, Lumbsch HT (2018) Phylogenomic analysis of 2556 single-copy protein-coding genes resolves most evolutionary relationships for the major clades in the most diverse group of lichen-forming fungi. *Fungal Diversity* 92(1):31–41. <https://doi.org/10.1007/s13225-018-0407-7>
- Poelt J, Grube M (1992) Lichen flora of the Himalayas. 5. The genus *Protoparmelia* Choisy. *Nova Hedwigia* 55(3–4):381–395
- Rambaut A (2009) FigTree 1.2.2. <http://tree.bio.ed.ac.uk/software/figtree/>
- Rambaut A, Drummond AJ (2007) Tracer v1.4, Available from <http://beast.bio.ed.ac.uk/Tracer>
- Robbertse B, Reeves JB, Schoch CL, Spatafora JW (2006) A phylogenomic analysis of the Ascomycota. *Fungal Genetics and Biology* 43(10):715–725. <https://doi.org/10.1016/j.fgb.2006.05.001>
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61:539–542
- Rubin BER, Ree RH, Moreau CS (2012) Inferring phylogenies from RAD sequence data. *PLoS One* 7(4):12. <https://doi.org/10.1371/journal.pone.0033394>
- Salas-Lizana R, Oono R (2018) Double-digest RADseq using standard Illumina indexes improve deep and shallow phylogenetic resolution of *Lophodermium*, a widespread fungal endophyte of pine needles. *Ecology and Evolution* 8(13):6638–6651. <https://doi.org/10.1002/ece3.4147>
- Singh G, Aptroot A, Rico V, Otte J, Divakar PK, Crespo A, Caceres MED, Lumbsch HT, Schmitt I (2018) *Neoprotoparmelia* gen. nov. and *Maronina* (Lecanorales, Protoparmelioideae): species description and generic delimitation using DNA barcodes and phenotypical characters. *MycKeys* 44:19–50
- Slater GS, Birney E (2005) Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6:81. <https://doi.org/10.1186/1471-2105-6-31>
- Solis-Lemus C, Ane C (2016) Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. *PLoS Genetics* 12(3). <https://doi.org/10.1371/journal.pgen.1005896>
- Solis-Lemus C, Yang MY, Ane C (2016) Inconsistency of species tree methods under gene flow. *Systematic Biology* 65(5):843–851. <https://doi.org/10.1093/sysbio/syw030>
- Spatafora JW, Aime MC, Grigoriev IV, Martin F, Stajich JE, Blackwell M (2017) The fungal tree of life: from molecular systematics to genome-scale phylogenies. *Microbiology Spectrum* 5(5). <https://doi.org/10.1128/microbiolspec.FUNK-0053-2016>
- Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313
- Sukumaran J, Holder MT (2010) DendroPy: a Python library for phylogenetic computing. *Bioinformatics* 26(12):1569–1571. <https://doi.org/10.1093/bioinformatics/btq228>
- Talas F, McDonald BA (2015) Genome-wide analysis of *Fusarium graminearum* field populations reveals hotspots of recombination. *BMC Genomics* 16:996. <https://doi.org/10.1186/s12864-015-2166-0>
- Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* 56(4):564–577
- Thell A, Crespo A, Divakar PK, Kärnefelt I, Leavitt SD, Lumbsch HT, Seaward MRD (2012) A review of the lichen family Parmeliaceae - history, phylogeny and current taxonomy. *Nordic Journal of Botany* 30(6):641–664. <https://doi.org/10.1111/j.1756-1051.2012.00008.x>
- Widhelm TJ, Grewe F, Huang J-P, Mercado-Diaz JA, Goffinet B, Luecking R, Moncada B, Mason-Gamer R, Lumbsch HT (2019) Multiple historical processes obscure phylogenetic relationships in a taxonomically difficult group (Lobariaceae, Ascomycota). *Scientific Reports* 9:8968. <https://doi.org/10.1038/s41598-019-45455-x>
- Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applied Biosciences* 13(5):555–556
- Zhang C, Rabiee M, Sayyari E, Mirarab S (2018) ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19(6):153. <https://doi.org/10.1186/s12859-018-2129-y>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

