This is the submitted version of a paper published in *Organisms Diversity & Evolution*.

Citation for the original published paper (version of record):

https://doi.org/10.1007/s13127-016-0320-4

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

http://urn.kb.se/resolve?urn=urn:nbn:se:nrm:diva-2554
Using multi-locus sequence data for addressing species boundaries in commonly accepted lichen-forming fungal species (*Diploschistes*, Graphidaceae)

Xin Zhao¹,², Samantha Fernández-Brime³, Mats Wedin³, Marissa Locke⁴, Steven D. Leavitt⁵, H. Thorsten Lumbsch⁵

¹ College of Life Sciences, Liaocheng University, Liaocheng 252059, P. R. China
² College of Life Sciences, Shandong Normal University, Jinan 250014, P. R. China
³ Department of Botany, Swedish Museum of Natural History, P.O. Box 50007, 104 05 Stockholm, Sweden
⁴ Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, 63110, USA
⁵ Science & Education, The Field Museum, 1400 S. Lake Shore Drive, Chicago, IL 60605, USA
Abstract

Accurate species delimitations are of great importance for effectively characterizing biological diversity. Our criteria for delimiting species have changed dramatically over the last decades with the increasing availability of molecular data and improvement of analytical methods to evaluate these data. Whereas reciprocal monophyly is often seen as an indicator for identifying distinct lineages, recently diverged species often fail to form monophyletic groups. At the same time, cryptic species have repeatedly been detected in numerous organismal groups. In this study we addressed the species delimitation in the crustose lichen-forming fungal genus *Diploschistes* using multilocus sequence data from specimens representing 16 currently accepted species. Our results indicate the presence of previously undetected, cryptic species-level lineages in the subgenus *Limborina*. In the subgenus *Limborina*, samples from different continents currently classified under the same species were shown to be only distantly related. At the same time, in two clades within subgen. *Diploschistes* characterized by short branches, none of the currently accepted species formed monophyletic groups. In spite of the lack of monophyly in phylogenetic reconstructions, a multispecies coalescent method provided support for eight of the nine accepted species in subgen. *Diploschistes* as distinct species. We propose to reduce *D. neutrophilus* to synonymy with *D. diacapsis* and point out that additional sampling will be necessary before accepting additional species in subgen. *Limborina*.

**Key words:** *Diploschistes*, classification, *Limborina*, molecular phylogeny, BPP, multispecies coalescent method
Introduction

In a broad sense, species delimitation is the process of identifying how individuals and populations fit into natural, species-level clusters, which are not simply constructs of classification (Carstens et al. 2013). Species are generally considered to represent a fundamental unit in biology and provide valuable context for organizing, evaluating, and communicating important biological concepts and principles (Coyne and Orr 2004; Mayr 1963). Therefore, accurate species circumscriptions are integral to interpreting biological patterns and processes across a wide range of sub-disciplines in biology.

Over the last decade, multi-locus sequence data have increasingly been used for assessing species (Camargo and Sites 2013), and the ongoing development of empirical approaches facilitates more objective species delimitation using molecular data (Leavitt et al. 2015b; Leavitt et al. 2016). DNA sequence data have revealed previously unrecognized species-level lineages hidden within nominal taxa in all organismal groups studied to date. Furthermore, many traditional taxonomic concepts conflict to various degrees with species-level lineages circumscribed using molecular sequence data. The bulk of species delimitation research highlights the fact that finding and applying the appropriate character sets and analytical tools remains one of the greatest challenges with empirical species delimitation (Lumbsch and Leavitt 2011).

Similar to other biological groups, molecular sequence data have been central to improving hypotheses of species boundaries in lichen-forming fungi (Crespo and Lumbsch 2010). Traditionally, differences in morphological, chemical, and ecological features have been the predominant source of diagnostic taxonomic characters for circumscribing lichen-forming fungal species (Printzen 2010). However, lichenized fungi generally display few taxonomically useful characters, and varying levels of intraspecific variation among different species groups may confound accurate taxonomic
circumscriptions. Therefore, molecular genetic data play an increasingly prominent role in delimiting fungal species and understanding evolutionary relationships in lichenized fungi.

The lichen-forming genus *Diploschistes* Norman (Lecanoromycetes: Ostropales: Graphidaceae) (Lumbsch and Huhndorf 2010) currently includes about 30 crustose species (Jaklitsch et al. 2016), which grow on rocks, soil or over mosses and other lichens. Traditionally, *Diploschistes* has been characterized by having a carbonized proper excipulum with lateral paraphyses, *Trebuixia* Puymaly as its photobiont, and by the absence of a columella (Lumbsch 1989). Morphologically, *Diploschistes* was regarded as consisting of three main groups: the *D. actinostomus* group with perithecioid ascomata, the *D. scrucopus* group with urceolate ascomata, and the *D. ocellatus* group with lecanoroid ascomata. These informal groups were supported as monophyletic clades in phylogenetic studies using phenotypical (Lumbsch and Tehler 1998) or molecular data (Fernández-Brime et al. 2013; Martín et al. 2003).

The monophyly of the genus *Diploschistes* has been discussed in the literature. Morphologically two species were unique in the genus: *D. bisporus* (Bagl.) J.Steiner and *D. ocellatus* (Vill.) Norman, which both lack lateral paraphyses (Lumbsch 1989). Morphological studies already revealed that the former species is distinct in numerous characters and consequently it was separated at the generic level as *Ingvariella bispora* (Bagl.) Guderley & Lumbsch (Guderley et al. 1997). Recent studies showed that this genus is not only unrelated to other *Diploschistes* but belongs to different family in Ostropales (Fernández-Brime et al. 2011). *Diploschistes ocellatus* was also found to be distinct from other *Diploschistes* species with some studies showing the genus being non-monophyletic (Parnmen et al. 2013; Rivas Plata et al. 2013; Kraichak et al. 2014). Consequently, *D. ocellatus* has been segregated as *Xalocoa ocellata* (Kraichak et al. 2014).
Currently, *Diploschistes* includes species placed in the *actinostomus* group, which is now recognized as subgenus *Limborina* Fdez.-Brime, Gaya & Llimona, and the *scruposus* group or subgenus *Diploschistes* (Fernández-Brime et al. 2013).

While recent molecular studies yielded in a better understanding of the circumscription and phylogenetic placement of *Diploschistes* and the phylogeny of major clades within the genus, the delimitation of species is still unresolved. Previous studies (Fernández-Brime et al. 2013; Martín et al. 2003) indicated that the circumscription of some species needs re-examination and the species delimitation based on phenotypic characters differ between authors (Clauzade and Roux 1985; Llimona1974; Lumbsch 1989, 1988; Pant and Upreti 1993). Hence we assembled a data set consisting of six loci (two nuclear ribosomal, one mitochondrial ribosomal, and three nuclear protein-coding genes) to address the species delimitations within *Diploschistes* and test morphology-based hypotheses.

**Methods and materials**

**Taxon sampling**

Our sampling of *Diploschistes* species included a total of 93 specimens representing 16 currently recognized species (Table 1). Species of the subgenus *Diploschistes* included: *D. cinereocaesius*, *D. diacapsis*, *D. gypsaceus*, *D. interpediens*, *D. muscorum*, *D. neutrophilus*, *D. rampoddensis*, *D. scruposus*, and *D. thunbergianus*. Taxa included in this study that belong to subgenus *Limborina* included: *D. actinostomus*, *D. caesioplumbeus*, *D. candidissimus*, *D. diploschistoides*, *D. elixii*, *D. euganeus*, and *D. sticticus*. Based on previous studies, the tree was rooted with subgenus *Limborina* (Martín et al. 2003; Fernández-Brime et al. 2013). We attempted to sample specimens across the range of each species’ distribution, and overall specimens from Africa, Australia, Central America, Europe, North America and South America were selected.
Additional specimens were selected from GenBank to improve our taxonomic sampling.

**Molecular data**

Sample preparation, DNA isolation, PCR and direct sequencing were performed as described previously (Fernández-Brime et al. 2013; Leavitt et al. 2012). Molecular data were generated for six loci: the internal transcribed spacer (ITS), nuclear large subunit (nucLSU), mitochondrial small subunit (mtSSU), minichromosome maintenance complex component 7 (MCM7), the largest subunit of the RNA polymerase II gene (RPB1), and the second largest subunit of RNA polymerase II gene (RPB2). Primers and PCR cycling parameters used for amplifying the six loci are listed in Table 2.

**Sequence alignments**

New sequences were assembled and edited using the program Sequencher v4.10 (Gene Code Corporation, Ann Arbor, MI) and were subjected to BLAST searches for a first verification of their identities. Sequences of each locus were aligned using the program MAFFT v7 (Katoh et al. 2009). For ITS sequences, we used the L-ING-i alignment algorithm with the remaining parameters set to default values. For nucLSU, G-ING-i algorithm and “leave gappy regions” were selected. Then we used E-ING-i algorithm for mtSSU and RPB1, and G-ING-i algorithm for MCM7 and RPB2, with the remaining parameters set to default values. The alignments were adjusted manually to exclude missing data and concatenated. Ambiguous positions of the ITS and mtSSU alignments were removed using the Gblocks web server (Castresana 2000), implementing all the options for a less stringent selection.

**Phylogenetic analysis**

Exploratory phylogenetic analyses of individual loci revealed a general pattern of poorly resolved topologies. Therefore, the six single-locus alignments were concatenated in
Geneious v6.1.2 (Biomatters Ltd., Auckland, NZ) for subsequent phylogenetic analyses.

Only specimens that were represented by at least two of six targeted loci were included in the concatenated data matrix (Table 1). A maximum likelihood (ML) analysis was carried out on the multilocus matrix using the locus-specific model partitions (ITS, nucLSU, mtSSU, MCM7, RPB1 and RPB2) in RAxML v8.1.24 (Stamatakis 2006). A search combining 200 separate ML searches was conducted, implementing a GTRGAMMA model, and 1000 pseudoreplicates to evaluate bootstrap support for each node. In addition to the ML analysis, a Bayesian analysis with MrBayes v3.2.3 (Ronquist et al. 2012) was also used for phylogenetic inference from our multilocus dataset. The most appropriate nucleotide substitution model for each of the six loci was selected using the Akaike information criterion in jModelTest v2.1.7 (Darriba et al. 2012). The Bayesian analysis was run for 10,000,000 generations with four independent chains and sampling every 1000th tree. All model parameters were unlinked. Two independent Bayesian runs were conducted to ensure that stationarity was reached and the runs converged at the same log-likelihood level (Nylander et al. 2008). After discarding the burn-in, the remaining 7500 trees of each run were pooled to calculate a 50% majority rule consensus tree. Clades that received bootstrap support ≥ 70% under ML and posterior probabilities ≥ 0.95 were considered significant. Phylogenetic trees were visualized using FigTree v1.4.2 (Rambaut 2009).

Species delimitation analysis

For a subgroup of species in the *D. scruposus* group, the multispecies coalescent model implemented in the program BPP v3.2 (Yang and Rannala 2010; Rannala and Yang 2013; Yang and Rannala 2014) was used to infer support for the separation of the sampled *Diploschistes* species. The *D. scruposus* group has recently been shown to have a recent diversification history (Rivas Plata 2011; Kraichak et al. 2015), and recently diverged species may not be recovered as monophyletic due to incomplete lineage sorting (Knowles and Carstens 2007). Given the recent diversification history for the *D.*
scruposus group, lack of resolution and short branches in phylogenetic reconstructions for this group (see Results), and support from phenotypic and ecological evidence (Lumbsch and Tehler 1998), it may be reasonable to assume that traditionally circumscribed species in the D. scruposus group represent distinct evolutionary lineages, in spite of their lack of monophyly in phylogenetic reconstructions. Therefore, for the BPP analyses, which accounts for incomplete lineage sorting within a multispecies coalescent framework, specimens within the D. scruposus group were assigned to candidate species based on phenotype-based identifications. BPP incorporates coalescent theory and phylogenetic uncertainty into parameter estimation; and the posterior distribution for species delimitation models is sampled using a reversible-jump Markov Chain Monte Carlo (rjMCMC) chain. We used the unguided species delimitation analysis ‘A11’ (Yang 2015), which explores different species delimitation models and different species phylogenies, with fixed specimen assignments to populations. Specimens were assigned to nine currently accepted species: D. cinereocaesius, D. diacapsis, D. gypsaecus, D. interpediens, D. muscorum, D. neutrophilus, D. rampoddensis, D. scruposus, and D. thunbergianus. Using analysis ‘A11’, the algorithm attempts to merge populations into one species, and uses the nearest neighbor interchange (NNI) or subtree pruning and regrafting (SPR) algorithms to change the species tree topology (Yang and Rannala 2014). Analysis ‘A00’ (Yang 2015), a within-model inference, was used to generate the posterior distribution of the parameters theta ($\theta$) and tau ($\tau$) under the multispecies coalescent model (MSC) model to infer a reasonable combination of priors given the data (Rannala 2015). Based on the results from the ‘A00’ analyses, the gamma prior G for $\theta$ was set to $\sim$ G(1, 85), and the gamma prior G for $\tau$ was set to $\sim$ G(1, 200). Under the unguided species delimitation model, ‘A11’, we used two different search algorithms (algorithm 0 or 1), equal probabilities for the labeled histories, to assign probabilities to the models, rates were allowed to vary among loci (locus rate=1), and the analyses were set for automatic fine-tune adjustments. The rjMCMC analysis was run for 100,000 generations, sampling every 2 generations discarding the first 10% as burn-in. The analysis was run twice to confirm consistency between runs.
Results

Molecular data
For this study, 217 new sequences were generated (Table 1). The multilocus matrix we used in this study was deposited in TreeBase (ID# pending). The concatenated, six-locus matrix consisted of 93 individuals and 5074 aligned nucleotide position characters (Table 3). A summary of alignment information for the multilocus dataset was also provided in Table 3.

Phylogenetic analysis
Phylogenies derived from the ML and B/MCMC analyses were generally concordant. Minor differences in the arrangement of some terminals occurred, but relationships at deeper nodes and in well-supported clades were identical. We chose to present the ML topology, with nodal support values from both ML bootstrap analysis and posterior probabilities from the Bayesian inference (Fig. 1).

The two main groups of *Diploschistes* were recovered in our phylogenetic trees with strong support (both are BS=98, PP=1.0; Fig. 1). In our ML tree, subgen. *Diploschistes* was shown to include three poorly supported subclades, which were labeled as Clade 1, Clade 2 and Clade 3, respectively. Clade 1 contained five species – *D. diacapsis* (three specimens), *D. gypsaceus* (five specimens), *D. interpediens* (12 specimens), *D. muscorum* (two specimens), and *D. scruposus* (26 specimens), and two samples which could not be identified with certainty, with seven supported internodes. Clade 2 contained three species – *D. diacapsis* (five specimens), *D. muscorum* (seven specimens), and *D. neutrophilus* (seven specimens), with five supported internodes. Clade 3 contained three species – *D. cinereocaesius* (four specimens), *D. rampoddensis* (three specimens) and *D. thunbergianus* (two specimens), with six supported internodes. Within subgen. *Limborina*, most relationships were also unresolved, and only two internodes were supported. In this
clade, only *D. diploschistoides* (BS=75, PP=1.0) was recovered as monophyletic with strong support, whereas all other species were either not monophyletic or their monophyly was not strongly supported.

Branch lengths between clades in subgen. *Limborina* and clade 3 differed considerably from those in clades 1 and 2. Branch lengths and support for clades was generally low in clades 1 and 2. Further, species in clades 1 and 2 did not form monophyletic groups. However, their monophyly could not be rejected using alternative topology tests (data not shown). Hence we employed multispecies coalescent species delimitation using BPP to evaluate separation of currently accepted species in subgenus *Diploschistes*. An 8-species delimitation scenario had the highest probability, followed by a 7-species scenario (Table 4). All other species delimitation models had probabilities < 0.05. Currently accepted species in subgenus *Diploschistes* received the highest supported, with the exception of *D. diacapsis* and *D. neutrophilus*, which were collapsed into a single species with high probability (Table 5).

**Discussion**

We used a six-locus dataset including three ribosomal (ITS, nuLSU, mtSSU) and three protein-coding markers (MCM7, RPB1, RPB2) of 93 specimens representing 16 currently accepted species to test the species delimitation in the genus *Diploschistes*. Our results indicate both the presence of previously undetected, cryptic species and difficulties in separating species using molecular markers. Species in this genus have largely been separated based on ascomatal characters, such as apthecial morphology, exciple thickness, number of ascospores per ascus, ascospore-size, -form and -amyloidity, secondary metabolites, and thallus morphology (Lumbsch 1989; Lumbsch and Elix 1989; Rivas Plata et al. 2010). Whereas species delimitation based on these phenotypical characters have largely been in agreement among authors, variability of number of ascospores, thallus
morphology, and ecology of species of the subgenus *Diploschistes* have differed somewhat among authors (Clauzade and Roux 1985; Llimona 1974; Lumbsch 1989, 1988; Pant and Upreti 1993).

Among species of the subgenus *Limborina*, samples from different continents currently classified under the same species were often only distantly related. This includes *D. actinostomus* with the samples from North America and Africa being separated, *D. euganeus* with samples from Australia and Europe not clustering together, and *D. sticticus* with samples from Australia and Africa not forming a monophyletic group. Only two species in this subgenus that included more than one sample formed monophyletic groups, these are *D. caesioplumbeus* (both collections from Spain) and *D. diploschistoides* (three samples from Australia). These results suggest that phenotypically similar specimens occurring on different continents in fact represent distinct lineages. Our taxon sampling is insufficient to address species delimitation in subgen. *Limborina* but these results demonstrate that additional studies are necessary to better understand species delimitation in this subgenus. However, the presence of distinct lineages within nominal species on different continents has repeatedly been shown in other groups of lichen-forming fungi (Amo de Paz et al. 2012; Arguello et al. 2007; Divakar et al. 2010; Hodkinson and Lendemer 2011; Thell et al. 2009; Otálora et al. 2010; Parnmen et al. 2012; Alors et al. 2016; Leavitt et al. 2015a; Zhao et al. 2015).

In subgen. *Diploschistes* three major clades were found. In clade 3, all species were monophyletic and strongly supported. In contrast, none of the currently accepted species were monophyletic in clades 1 and 2, and the relationships within those clades were inferred with short branches and mostly poorly supported. Monophyly is not a prerequisite of a species, and thus, a lack thereof is not necessarily evidence that these lineages are conspecific (Leavitt et al. 2016). For example, in North American species of
the genus *Xanthoparmelia* recent diversification during Pliocene and Pleistocene was estimated and independent species-level lineages were not supported by concordant evolutionary histories across multiple, independent loci (Leavitt et al. 2011a; Leavitt et al. 2011b; Leavitt et al. 2013).

In spite of our attempt to reconstruct phylogenetic relationships for the *D. scrupsosus* group using multilocus sequence data, species boundaries and relationships within this group remained unresolved. Therefore, we used the program BPP to delimit species boundaries in this group within a statistical framework modeled under the multispecies coalescent. This analysis strongly supported eight of the nine accepted species in subgen. *Diploschistes* as separate species. Previously it has been shown that the genus *Diploschistes*, and especially subgen. *Diploschistes*, has diversified recently (Rivas Plata 2011; Kraichak et al. 2015). We hypothesize that the difficulties in separating species in clades 1 and 2 of subgen. *Diploschistes* are due to recent diversification. The multispecies coalescent analysis, which accounts for incomplete lineage sorting, supports that most of them are in fact are distinct species. The only exception is *D. neutrophilus*, which was supported to belong to *D. diacapsis*. Consequently, we propose to reduce *D. neutrophilus* to synonymy with *D. diacapsis*.

The coalescent-based BPP program accounts for ancestral polymorphisms and incomplete lineage sorting. However, other factors, such as occasional gene flow, hybridization, and recombination, are other evolutionary factors potentially influencing species delimitation inferences. While BPP performs quite robustly under a range of scenarios, speciation probabilities decrease with increasing levels of gene flow (Camargo et al. 2012; Zhang et al. 2011).

BPP also requires *a priori* assignment of individuals to candidate species and the impact
of incorrectly assigned specimens remains unclear. While traditionally accepted nominal species have been well-studied morphology (Lumbsch 1989; Lumbsch and Elix 1989), our multilocus phylogenetic reconstructions failed to provide strong support either for or against the traditionally, phenotype-based species in the *D. scruposus* group. Therefore, we based our specimen assignments to candidate species for the BPP analysis on phenotype-based identifications. Arguably, as the use of genome-wide molecular data becomes more commonplace in lichen research, species boundaries and evolutionary relationships in lineages with recent diversification histories, including the *D. scruposus* group, will better understood. In the meantime, we propose that the 8-species model inferred for the *D. scruposus* group (Table 4) represents a useful working hypothesis of species boundaries for this group.

**Taxonomic conclusions**

*Diploschistes diacapsis* (Ach.) Lumbsch


In a recent revision of *Diploschistes* (Fernández-Brime et al. 2013), the phylogeny supported two separate clades corresponding to the morphological concepts of *D. diacapsis* and *D. diacapsis* subsp. *neutrophilus*. The authors also noticed morphological and ecological differences: *D. diacapsis* had thicker and convex thalli (up to 2mm), which were detached from the substrate, and grew in gypsiferous or highly calcareous soils from inland continental areas, while *D. diacapsis* subsp. *neutrophilus* had thin flat thalli completely attached to the substratum, and grew in decarbonized soils in coastal
areas. Based on these results, Fernández-Brime et al. (2013) raised *D. dicapsis* ssp, *neutrophilus* to species level. Our present study, however, includes a larger number of samples and loci, and our phylogeny (Fig. 1) clearly shows that specimens with *D. diacapsis* and *D. neutrophilus* morphologies do not form distinct clades. Furthermore, several samples identified as *D. neutrophilus* were collected in the Tabernas desert (Table 1), a typical inland semi-arid *D. diacapsis* locality. In the light of these results based on more data, and finding the thallus thickness and shape a much more inconsistent character than previously believed, we formally synonymize *D. neutrophilus* with *D. diacapsis*.

**Acknowledgements**

This study was financially supported by the National Science Foundation (DEB 0516116 to The Field Museum; PI, H. T. Lumbsch; Co-PI, R. Lücking; and DEB-1025861 to The Field Museum; PI, T. Lumbsch; Co-PI, R. Lücking) and Swedish Research Council grant (VR 621-2012-3990 to the Swedish Museum of Natural History; PI M. Wedin).
References


Parnmen, S., Cáceres, M. E. S., Lücking, R., & Lumbsch, H. T. (2013). *Myriochapsa* and *Nitidochapsa*, two new genera in Graphidaceae (Ascomycota:
Ostropales) for chroodiscoid species in the *Ocellularia* clade. *Bryologist, 116*, 127-133.


---

**Fig. 1** Maximum likelihood (ML) phylogenetic relationships of *Diploschistes* taxa inferred from a combined 6-locus analysis. Values at each node indicate nonparametric bootstrap support (BS)/posterior probability (PP), branches in bold received maximum likelihood bootstrap support values equal or above 70 and posterior probabilities equal or above 0.95.