Molecular Phylogenetics and Evolution 92 (2015) 217–225

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Phylogeny and species delimitations in European *Dicranum* (Dicranaceae, Bryophyta) inferred from nuclear and plastid DNA ☆

Annick S. Lang [a,b,*], Gaëlle Bocksberger [c], Michael Stech [a,b]

[a] Naturalis Biodiversity Center, P.O. Box 9517, 2300 RA Leiden, The Netherlands
[b] Leiden University, Leiden, The Netherlands
[c] Department of Primatology, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany

ARTICLE INFO

Article history:
Received 28 July 2014
Revised 24 June 2015
Accepted 29 June 2015
Available online 4 July 2015

Keywords:
Bryophyte
DNA-based taxonomy
Species boundaries
Morphological plasticity
Chloroplast markers
nrITS

ABSTRACT

DNA sequences have been widely used for taxonomy, inferring phylogenetic relationships and identifying species boundaries. Several specific methods to define species delimitations based on molecular phylogenies have appeared recently, with the generalized mixed Yule coalescent (GMYC) method being most popular. However, only few studies on land plants have been published so far and GMYC analyses of bryophytes are missing. *Dicranum* is a large genus of mosses whose (morpho-)species are partly ill-defined and frequently confused. To infer molecular species delimitations, we reconstructed phylogenetic trees based on five chloroplast markers and nuclear ribosomal ITS sequences from 27 out of 30 species occurring in Europe. We applied the species delimitation methods GMYC and Poisson tree processes (PTP) in order to compare their discriminatory power with species boundaries inferred from the molecular phylogenetic reconstructions and with the morphological species concept. Phylogenetic circumscriptions were congruent with the morphological concept for 19 species, while eight species were molecularly not well delimited, mostly forming closely related species pairs. The automated species delimitation methods achieved similar results but tended to overestimate the number of potential species and exposed several incongruences between the morphological concept and inference from molecular phylogenetic reconstructions. It is concluded that GMYC and PTP methods potentially provide a useful and objective way of delimiting bryophyte species, but studies on further bryophyte data sets are necessary to infer whether incongruences might ensue from evolutionary processes and to test the suitability of these approaches.

1. Introduction

DNA sequence data are widely used for inferring species delimitations and phylogenetic relationships. Specific methods to analyze species boundaries based on molecular phylogenetic reconstructions without prior species information, however, have been developed only recently (cf. Carstens et al., 2013 for review). Most popular is the generalized mixed Yule coalescent (GMYC) method (Fontaneto et al., 2007; Monaghan et al., 2009; Pons et al., 2006), while the Poisson tree processes (PTP) method has recently been proposed by Zhang et al. (2013) as an alternative to GMYC. Both methods estimate the point of transition between species and population, i.e. they infer species boundaries based on the differences in branching rates at species and population levels, assuming that the number of substitutions within a species is significantly lower than between species. The main difference is that GMYC requires an ultrametric tree that relies on a Bayesian tree sampling using MCMC methods to fit both Yule and coalescence models (Hudson, 1990; Yule, 1925) and finally delimit evolutionary species units (ESU; Tang et al., 2014). PTP, in contrast, uses directly the number of substitutions (instead of time) to simulate speciation and coalescent events, and ESU delimitations are based on heuristic search algorithms to estimate species boundaries with the maximum likelihood scores (Tang et al., 2014; Zhang et al., 2013). Because PTP does not require the input tree to be ultrametric, the method is much less computing-intensive than GMYC.

Generally, automated species delimitation methods are considered especially useful in organisms with unclear species boundaries, due to poor taxonomic knowledge or signals in phylogenetic reconstructions being obscured by lineage sorting

☆ This paper was edited by the Associate Editor Elizabeth Zimmer.

* Corresponding author at: Naturalis Biodiversity Center, P.O. Box 9517, 2300 RA Leiden, The Netherlands.
E-mail address: annick.lang55@gmail.com (A.S. Lang).

http://dx.doi.org/10.1016/j.ympev.2015.06.019
1055-7903/

or introgression (O'Meara, 2010 and references therein). Most GMYC studies so far focused on different animal groups (e.g. Poulakakis et al., 2012; Zaldívar-Riverón et al., 2010) and very few examples of analyses of other organisms such as algae (e.g. Leliaert et al., 2009), fungi (e.g. Parnmen et al., 2012) and land plants (e.g. Hernández-León et al., 2013) have been published. GMYC analyses of bryophyte datasets are still missing. Bryophytes are an important component of terrestrial ecosystems and count up to 18,000 known species (Goffinet and Shaw, 2009). The limited number of morphological characters available, high morphological plasticity, and often broad geographical distributions pose serious problems on species delimitations and taxonomy in many bryophyte lineages. Therefore, species delimitation methods such as GMYC and PTP could potentially make an important contribution to delimit bryophyte species and evaluate the significance of morphological characters for species identification, but their performance on bryophyte datasets remains to be tested. While molecular data can facilitate the circumscription of (closely related) bryophyte species, (e.g. Dong et al., 2012; Hedenäs and Eldenäs, 2007; Heinrichs et al., 2009; Stech et al., 2013), multiple DNA markers are often required in order to obtain supported species delimitations due to low levels of genetic variability (Hollingsworth et al., 2009, 2011; Lang et al., 2014a), which may pose a problem on the accuracy of species delimitation methods.

Species circumscription and identification in the Holarctic moss genus *Dicranum* (Dicranaceae, Bryophyta) has been notoriously difficult. The genus counts more than 90 species (www.tropicos.org; Frey and Stech, 2009), many of which are broadly distributed and display a great range of morphological plasticity, with only few habitat-specific species (Hedenäs and Bisang, 2004). Moreover, *Dicranum* and related genera display little molecular variation (Cox et al., 2010; La Farge et al., 2002; Stech, 1999; Stech et al., 2012). Thus, assessing species delimitations in *Dicranum* is challenging both at the morphological and molecular level. Our recent studies on the *Dicranum scoparium* and *D. acutifolium* species complexes (Lang and Stech, 2014; Lang et al., 2014b) as well as on boreal-arctic *Dicranum* species (Lang et al., 2014a) showed that in several cases conclusive species delimitations could only be obtained from combined analyses of several chloroplast markers and nuclear ribosomal ITS sequences.

The present study aims to elucidate species boundaries within *Dicranum* on a broad geographic scale, including 27 of the 29 *Dicranum* species occurring in Europe (Hedenäs and Bisang, 2004) plus *D. septentrionale* Tubanova and Ignatova, a newly recorded species in Scandinavia (Lang et al., 2014b). Molecular phylogenetic reconstructions based on five chloroplast markers (*trnH*$_{GUG}$-*psbA*, *rps4-trnT*$_{UGU}$ and *trnL*$_{UAA}$–*trnF*$_{GAA}$ intergenic spacers, *rps19-rpl2*, *rpoB*) plus the nrITS1-5.8S-ITS2 region will be used to test, to our knowledge for the first time in bryophytes, the congruence of two automated species delineation approaches, the general mixed Yule-coalescent (GMYC) and Poisson tree processes (PTP) methods.

2. Material and methods

2.1. Sampling

A total of 202 *Dicranum* specimens were sampled (Appendix 1), representing 27 species of the 29 European species recognized by Hedenäs and Bisang (2004) and including the new European species record of *D. septentrionale* (Lang et al., 2014b): six *Dicranum acutifolium* (Lindb. and Arnell) C.E.O. Jensen, nine *D. angustum* Lindb., six *D. bonjeanii* De Not., five *D. brevifolium* (Lindb.) Lindb., three *D. canariense* Hampe ex Müll.Hal., five *D. crassifolium* Sérgio, Ochyra and Séneca, one *D. dispersum* Engelmark, one *D. drummondii* Müll.Hal., four *D. elongatum* Schleich. ex Schwägr., three *D. flagellare* Hedw., 11 *D. flexicaule* Brid., four *D. fragilifolium* Lindb., six *D. fuscescens* Turner, two *D. groenlandicum* Brid., 11 *D. laevidens* R.S. Williams, three *D. leioneuron* Kindb., eight *D. majus* Turner, four *D. montanum* Hedw., four *D. polysetum* Sw., 65 *D. scoparium* Hedw., two *D. scottianum* Turner ex R. Scott, nine *D. septentrionale*, 15 *D. spadiceum* J.E. Zetterst., three *D. spurium* Hedw., four *D. tauricum* Sapjegin, four *D. undulatum* Schrad. ex Brid. and four *D. viride* (Sull. and Lesq.) Lindb. specimens. Fourty specimens were newly sequenced for all six markers employed here, except four specimens which ITS sequences have been generated by Tubanova et al. (2010), Ignatova and Fedosov (2008). The other 162 specimens were sequenced for previous studies (Lang and Stech, 2014; Lang et al., 2014a, 2014b; Stech, 1999; Stech et al., 2006). We chose as outgroup four specimens of *Holomitrium*, sister genus of the *Dicranum* s.l. clade (La Farge et al., 2002; Stech et al., 2006).

2.2. DNA extraction, amplification and sequencing

The greenest parts of single gametophyte stems were selected for DNA extraction and cleaned manually with demineralised water under a binocular. Total DNA extraction was carried out using the NucleoSpin® Plant II Kit (Macherey–Nagel, Düren, Germany). Six markers employed to delimit closely related *Dicranum* species in Lang and Stech (2014), Lang et al. (2014a, 2014b) were amplified and sequenced, i.e. five chloroplast regions (partial *rpoB* gene, *trnH*$_{GUG}$-*psbA*, *rps19-rpl2*, *rps4-trnT*$_{UGU}$ and *trnL*$_{UAA}$–*trnF*$_{GAA}$ intergenic spacer) and the nuclear ribosomal nrITS1-5.8S-ITS2 region. PCR amplifications were performed as described in Lang and Stech (2014). All PCR products were purified and sequenced at Macrogen Inc. (www.macrogen.com). GenBank accession numbers of all sequences are listed in Appendix 1.

2.3. Alignment and phylogenetic reconstruction

Sequences were aligned in Geneious v5.3.6 (Biomatters, 2010) using 65% similarity matrix costs, and manually adjusted. Short hairpin-associated inversions in the *trnH-psbA* spacer, which can flip at the population level and may significantly reduce phylogenetic structure if undetected (Borsch and Quandt, 2009; Quandt and Stech, 2004; Whitlock et al., 2010), were positionally separated in the alignment and the corresponding indels were excluded.

Phylogenetic inferences were based on maximum likelihood (ML) and Bayesian inference (BI) analyses. Gaps were coded as informative by a simple indel coding strategy (SIC) (Simmons and Ochoterena, 2000) implemented in SeqState (Müller, 2004). To check for incongruence, phylogenetic reconstructions based on chloroplast and nuclear sequences were visually compared. In addition, an incongruence length difference test (ILD, Farris et al., 1994) as implemented in PAUP* 4.0b10 (Swofford, 2002) was performed with 100 replicates. As both visual inspections and the ILD test indicated that the plastid and nuclear tree topologies were congruent ($p = 0.06$), the two datasets were combined.

Three nucleotide partitions were used in ML and BI, namely the non-coding chloroplast markers (*rps4-trnT*, *trnL–trnF*, *trnH-psbA*, *rps19-rpl2*), the coding chloroplast region *rpoB* and the nuclear ribosomal ITS region. Partitions were unlinked in both ML and BI inferences. ML analyses were carried out with RAxML v.7.2.6 (Stamatakis, 2006) employing the graphical user interface raxmlGUI v.0.93 (Silvestro and Michalak, 2012) with the default GTR model of nucleotide substitution and Γ rate heterogeneity for all partitions. Bootstrap analyses under ML were done using the thorough bootstrap heuristics algorithm with 20 runs and 1000 replicates. BI analyses were run on the CIPRES science gateway (Miller et al., 2010). Bayesian posterior probabilities were