Molecular Phylogenetics and Evolution xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Exploring the Udoteaceae diversity (Bryopsidales, Chlorophyta) in the Caribbean region based on molecular and morphological data

Laura Lagourgue[a,b,*], Nicolas Puillandre[c], Claude E. Payri[b]

[a] Sorbonne Universités, UPMC Univ Paris 06, IFD, 4 Place Jussieu, 75252 Paris Cedex 05, France
[b] UMR ENTROPIE (IRD, UR, CNRS), Institut de Recherche pour le Développement, B.P. A5 Nouméa Cedex, Nouvelle-Calédonie, 98848, France.
[c] Institut Systématique Evolution Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, 57 rue Cuvier, CP 26, Paris 75005, France

ARTICLE INFO

Keywords:
Bryopsidales
Udoteaceae
Udotea
Species delimitation
Multilocus phylogeny
Morphological clustering

ABSTRACT

The Udoteaceae family (Bryopsidales, Chlorophyta) is known to be highly diverse morphologically in the Caribbean region, but only few studies have studied its genetic diversity. Using an integrative taxonomic approach, this study aimed at (1) exploring the Udoteaceae species diversity using a combination of five DNA-based species delimitation methods and morpho-anatomical data for confirmation; (2) estimating the discriminatory power of traditional diagnostic characters using a morphology-based clustering method and statistical analyses focused on the genus *Udotea*; and (3) reconstructing the phylogeny of the family based on a multilocus analysis (*tuf*A, *rbc*L, 18S rDNA). Our results revealed strong congruence between species hypotheses across delimitation methods and markers. Morpho-anatomical characters proved essential to validate these hypotheses, to assign species names and to unveil new species. Morphological analyses led to relevant results for accurately discriminating *Udotea* morphospecies. Siphon features and cortication were key characters to define supra-specific groups and to revise the taxonomy of the genus *Udotea*. Phylogenetic analyses confirmed the polyphyly of *Udotea, Rhipocephalus* and *Penicillus*, which led us to propose a revised definition of *Udotea sensu stricto* based on both genetic and morphological data. Finally, our study emphasizes the importance of combining genetic and morphological data for the taxonomic revision of the Udoteaceae, but stresses the need of including more taxa from other geographical regions to better resolve taxonomic issues.

1. Introduction

Udoteaceae is a family of green macroalgae belonging to the order Bryopsidales (Chlorophyta). It comprises both calcified and non-calcified species, and has a significant ecological function through its contribution to primary production and carbonate fluxes. The Udoteaceae are particularly diverse morphologically, the most among the bryopsidalean families. Their internal structure is siphonous and characterized by a unique giant multinucleate cell that ranges from simple siphonous filaments to complex multiaxial structures. They include tufts of uncalcified and free filaments (*e.g.*, *Chlorodesmis*), brush-like calcified thalli (*e.g.*, *Penicillus*) or calcareous compact fan-shaped blades (*e.g.*, *Udotea*). Although they occur in the Mediterranean Sea, Udoteacean species are mostly tropical and particularly abundant and diverse in the Caribbean zone. Several authors have investigated the North Atlantic marine flora over the last decades (see Wynne (2017) for a review), and several species of *Udotea* were morphologically described by Littler and Littler (1990) and recorded by Collado-Vides et al. (2009) in their revision of the genus diversity from Cuba and Mexico. According to these works, only four out of the fourteen currently accepted Udoteaceae genera are present in the tropical western Atlantic, with one of them endemic to this region (*Rhipocephalus*). All genera included, twenty-four species have been listed, of which seventeen are only found in the tropical western Atlantic and ten are restricted to the Caribbean region. However, most of these species are based on morphological descriptions and genetic data is still unavailable or fragmentary, a common trend in the Udoteaceae. Indeed, since J. Agardh (1887) described the family, the taxonomy of the Udoteaceae has been established predominantly on morphological and anatomical traits (Ernst, 1904; Farghaly, 1980; Gepp and Gepp, 1911; Littler and Littler, 1990). The most significant taxonomical work on Udoteaceae is that of Gepp and Gepp (1911) based on material sampled during the Siboga expedition to the East Indian archipelago. Life cycles and sexual reproduction patterns were also the subject of a few studies but have not been reviewed since the work of Meinesz (1980) and Vroom et al. (1998), the latter performed cladistics analyses on morphological, anatomical and reproductive characters. The first phylogenetic analysis has been realized using nuclear-encoded ribosomal DNA (nrDNA) (Kooistra, 2002). This study

* Corresponding author at: UMR ENTROPIE (IRD, UR, CNRS), Institut de Recherche pour le Développement, B.P. A5 Nouméa Cedex, Nouvelle-Calédonie, 98848, France
E-mail address: laura.lagourgue@ird.fr (L. Lagourgue).

https://doi.org/10.1016/j.ympev.2018.06.023
Received 26 December 2017; Received in revised form 3 June 2018; Accepted 11 June 2018

emphasized the polyphyly of several genera among the Udoteaceae (*Udotea*, *Penicillus*, *Chlorodesmis*), which was later confirmed by Lam and Zechman (2006), Curtis et al. (2008) and Coppejans et al. (2011) based on the chloroplast *rbc*L marker, and by Verbruggen et al. (2009a, 2009b) based on multilocus analyses. All the results, together with the presence of cryptic diversity and/or morphological variability, illustrate how challenging the study of Udoteaceae can be and emphasize the need for a taxonomic revision.

Sequence-based species delimitation approaches are efficient tools for discriminating species and could help to better describe the Udoteaceae diversity and to resolve taxonomic ambiguities. A number of methods have been developed to detect discontinuities in DNA sequence variation associated with species boundaries. They have been applied at various taxonomic levels for the Phaeophyceae (Montecinos et al., 2017; Silberfeld et al., 2013; Vieira et al., 2014), the Rhodophyta (*e.g.*, Jesus et al., 2016; Guillemin et al., 2016; Pardo et al., 2014) and the Chlorophyta (Leliaert et al., 2009; Sauvage et al., 2016; Zou et al., 2016). In this study, we have chosen five exploratory species delimitation methods: the Automatic Barcode Gap Discovery (ABGD, Puillandre et al., 2012a), the General Mixed Yule Coalescent (GMYC) (Pons et al., 2006) and its Bayesian implementation, bGMYC (Reid and Carstens, 2012), the Poisson tree process model (PTP, Zhang et al., 2013), and the Multi-rate PTP (mPTP, Kapli et al., 2017). The association of these methods is interesting for the different assumptions they rely on. On the one hand, ABGD is a distance-based method which aims at identifying a barcode gap that delimits the intraspecific from the interspecific distances in the distribution of genetic distances. On the other hand, the (b)GMYC and (m)PTP methods use information from a phylogenetic tree and focus on branching and mutation rates to identify a shift between speciation (between species) and coalescence events (within species). While (b)GMYC rely on branching rates through time and thus require an ultrametric time-calibrated tree, (m)PTP relies on the number of substitutions along branches, and, in the case of mPTP, accounts for divergent intraspecific variation. The second group of methods (tree-based methods) assumes the monophyly of the delimited species, which is not the case of the distance-based methods such as ABGD. To converge towards robust species hypotheses, several authors strongly recommend using multilocus datasets and assessing the congruence between several methods (Carstens et al., 2013; Carstens and Knowles, 2007; Dupuis et al., 2012; Leliaert et al., 2014; Puillandre et al., 2012b; Rannala, 2015). The analysis of non-genetic data is also recommended by many authors to discuss DNA-based species hypotheses (Carstens et al., 2013; Carstens and Knowles, 2007; Fujita et al., 2012; Talavera et al., 2013; Wiens, 2007).

Following these recommendations, we based our study on molecular data, using several genetic markers to propose Primary Species Hypotheses (PSHs) and infer their phylogenetic relationship, combined with morpho-anatomical data to propose Secondary Species Hypothesis (SSHs), following an integrative taxonomy strategy. Our objectives were to: (i) explore the diversity of the Udoteaceae in the Caribbean region, (ii) clarify species boundaries and (iii) produce a more accurate phylogeny of the family. To reach these objectives, we (i) applied the five species delimitation methods described above using two chloroplast markers; (ii) validated the resulting DNA-based species hypotheses using cladistic analyses of morpho-anatomical traits, and (iii) reconstructed the phylogeny of the family based on the analysis of the chloroplast *tufA* and *rbc*L, and the nuclear 18S rDNA markers.

2. Material and methods

2.1. Sampling and study region

A total of 205 specimens of Udoteaceae from the Caribbean region were included in the study. Most of the samples (127) were collected in the Lesser Antilles from 10 islands, including 102 specimens from the scientific campaign PACOTILLES in 2015 (DOI, https://doi.org//10.17600/15005200) onboard the vessel Antea, and 25 specimens collected in 2014 in Guadeloupe (Supplementary Figure A.1). All our samples were collected by SCUBA from a total of 20 sites between the surface and 40 m deep. Sites were georeferenced and habitats described using bathymetric and substrata descriptors. When possible, specimens were photographed on site prior to collection in order to document their *in-situ* habits. Samples were sorted, labeled and photographed while fresh. Subsamples were preserved in both 95% ethanol and silica gel for later DNA analyses. Specimens were then pressed-dried as herbarium vouchers and are currently housed at NOU in New Caledonia. The other 78 specimens were collected and sequenced by collaborators of the National Museum of Natural History and the School of Biosciences of the University of Melbourne, which enabled us to extend our geographical coverage to most subregions of the Caribbean zone. The latter specimens are housed at PC in Paris and GENT respectively (herbarium abbreviations follow Thiers (2016), continuously updated).

2.2. DNA extraction, amplification and sequencing

From the 127 Udoteaceae specimens collected in the Lesser Antilles, 114 samples were successfully extracted using the DNeasy Plant Mini Kit (Qiagen Inc, Valencia, CA, USA) following a modified and optimized protocol based on the manufacturer's instructions (available on request). Sequences for two chloroplast and one nuclear genes were produced: *tuf*A (approximately 800 base pairs –bp), *rbc*L (approx. 1400 bp) and 18S rDNA (approx. 1300 bp) using previously published primers (Händeler et al., 2010; Kooistra, 2002; Lam and Zechman, 2006; Verbruggen et al., 2009a) (Supplementary material, Table A.1). The *rbc*L and 18S rDNA markers were sequenced in two fragments. PCR reactions were conducted in a final volume of 25 µL including of 12.5 µL of AmpliTaq Gold 360 Master Mix (Applied Biosystems), 0.75 µL of dimethylsulfoxyde (DMSO), 1 µL of bovine serum albumin (BSA), 1 µL of each primer (10 µM), 2.5 µL of DNA, and 6.25 µL of ultra-pure water (PCR programs available on request). The Sanger sequencing reaction was carried out using 20 µL of PCR product by Genoscreen (Lille, FRANCE). Sequences were edited in Geneious version 7.1.9 (http://www.geneious.com, Kearse et al., 2012) and aligned for each marker separately using the MUSCLE algorithm available in the software. In addition to the sequences produced during this study, Udoteaceae sequences obtained from collaborators (National Museum of Natural History: 42 *tuf*A sequences, and the School of Biosciences, University of Melbourne: 33 *tuf*A and 6 *rbc*L sequences) and GenBank (16 *rbc*L and 15 18S rDNA sequences) were added to the datasets. The detailed list of specimens used and the accession numbers are recorded in Supplementary material (Table A.2).

2.3. Trees inference for species delimitation

The species delimitation methods were used for the analysis of the two chloroplast markers, *tuf*A and *rbc*L, known for their discriminatory power at the species level in green macroalgae (Leliaert et al., 2014; Saunders and Kucera, 2010; Verbruggen et al., 2009a). The 18S rDNA was not used for species delimitation due to the low intra and interspecific variability observed in preliminary analyses (results available upon request). Both Bayesian ultrametric and Maximum Likelihood (ML) trees were constructed to produce the two types of input required for tree-based delimitation methods. Identical sequences (numerous for the *tuf*A dataset) were removed using the Collapsetypes v4.6 perl script (Chesters, 2013) to keep only one sequence for each haplotype, in order