The Macaronesian endemic moss *Andoa berthelotiana* (Myuriaceae, Bryophyta):

Phylogenetic relationships and cryptic speciation

*Soraia Martins*\(^1,2,^*\), *Manuela Sim-Sim*\(^1,2\) & *Michael Stech*\(^3,4\)

\(^1\) Universidade de Lisboa, Faculdade de Ciências de Lisboa, DBV, 3cÉ3c - Centre for Ecology, Evolution and Environmental Changes, C2, Campo Grande, 1749-016 Lisboa, Portugal

\(^2\) Museu Nacional de História Natural e da Ciência, Universidade de Lisboa, Rua da Escola Politécnica, 58, 1250-102 Lisboa, Portugal

\(^3\) Naturalis Biodiversity Center, P.O. Box 9517, 2300 RA Leiden, The Netherlands

\(^4\) Leiden University, Leiden, The Netherlands

* Corresponding author: ssmartins@fc.ul.pt

With 5 figures and 1 table

Short title: Phylogenetic inference in *Andoa berthelotiana*
Abstract: The Macaronesian endemic pleurocarpous moss species *Andoa berthelotiana* occurs in the Madeira, Azores and Canary Islands archipelagos. It has a checkered taxonomic history with placements in different genera or, as monospecific genus *Andoa*, in different families, most frequently in the Hypnaceae. Earlier molecular phylogenetic analyses indicated a close relationship of *Andoa* with *Ctenidium*, *Hyocomium* and *Myurium* in a clade corresponding to Myuriaceae, which was supported in the present study based on a larger sampling and markers from all three genomes (chloroplast *trnL-trnF*, nuclear ribosomal ITS and mitochondrial *nad5*). Haplotype networks of the *nad5* and ITS sequences as well as AFLP fingerprinting data revealed the existence of two intraspecific lineages in *Andoa berthelotiana*, one occurring in Madeira and the Canary Islands and the other in the Azores. The *trnL-trnF* haplotype network contradicts this geographic pattern, however, the position of the two mutations differentiating the *trnL-trnF* haplotypes suggests that they may not be phylogenetically informative. A detailed morphological analysis of plants from the Azores and Madeira indicates the existence of distinguishing characters that correspond with the two molecular lineages. However, the considerable morphological overlap observed would hamper the identification of specimens if both lineages were formally described as separate taxa. We therefore suggest treating the two lineages of *Andoa* as semi-cryptic species.

Key words: AFLP; *Andoa*; haplotypes; Macaronesia; molecular phylogeny; *nad5*; nrITS; *trnL-trnF*
Introduction

Macaronesia is traditionally considered a distinct biogeographic region including five archipelagos, Azores, Madeira, Selvagens, Canary Islands, and Cape Verde (Vanderpoorten et al. 2007, Fernández-Palacios et al. 2010). Its circumscription has been related with the idea that its flora constitutes a relic from the subtropical Tertiary flora (Vanderpoorten et al. 2007). More specifically, the laurel forest (Laurissilva) present in northern Macaronesia is considered to contain a relic flora of the late Tertiary, which survived the glaciations that affected the Mediterranean region during that period (Capelo et al. 2005), in contrast to the Macaronesian area that was hardly affected (Sérgio 1984).

The Macaronesian flora contains a number of endemic bryophyte genera and species that were considered paleoendemics, i.e. taxa with a wider distribution range in the Tertiary that disappeared from the continents during the Ice Age and survived in Macaronesia, particularly species of the laurel forests (Frahm 2004, 2006, Frahm & Häusler 2006). Arguments in favour of the paleoendemics hypothesis were the amount of time required for endemic genera to develop as well as fossils found in Baltic and Saxon amber, which were ascribed to species presently considered extinct in Europe (Echinodium spp. and Andoa berthelotiana Ochyra; Frahm 2004, 2008). Furthermore, Sérgio (1984) argued that the location of the Macaronesian archipelagos between Europe, Africa and America offers a privileged situation for the persistence of relict species of either Gondwanan or Laurasian origin.

More recently, molecular data did indeed reveal connections of the Macaronesian bryoflora to the Neotropics, Paleotropics, North America, and Europe (e.g., Feldberg et al. 2004, Sim-Sim et al. 2005, Stech et al. 2006, Vanderpoorten & Long 2006, Aigoin et al. 2009a, Stech et al. 2010a,

In addition, intraspecific genetic variation in Macaronesian bryophytes may indicate processes of (cryptic) speciation as well as different colonization routes to, and patterns of diversification between, the Macaronesian archipelagos. An integrative analysis of the pleurocarpous moss genus Isothecium in Macaronesia, for example, led to the description of a new, probably recently evolved species (Draper et al. 2015). Haplotype diversity in the acrocarpous moss species Grimmia montana Bruch & Schimp. suggested strong transatlantic gene flow for Madeira but a close relationship of Canary Islands populations to south-western Europe or southern Africa (Vanderpoorten et al. 2008). In liverworts, higher haplotype diversity within Madeira and Canary Islands compared to the Azores (Laenen et al. 2011, Lopes et al. 2015) was observed. To infer general patterns, however, studies of further taxa from different main lineages of bryophytes are necessary that may add to the growing evidence of intraspecific molecular differentiation within Macaronesia.

Andoa berthelotiana, the single species of the Macaronesian endemic genus Andoa, was first described as Hypnum berthelotianum Mont., and subsequently placed in various genera belonging to different families, such as Ctenidium (Hylocomiaceae or Ctenidiaceae), Hylocomium (Hylocomiaceae), Eurhynchium (Brachytheciaceae), as well as Gollania and Hyocomium (Hypnaceae). It was first considered a separate genus by Ando (1973), as Allorgea
berhelotiana (Mont.) Ando, which was replaced by its current name by Ochyra (1982).

Although mostly considered as belonging to the Hypnaceae (e.g. Ochyra 1982, Frey & Stech 2009), Hedenäs (1992) considered Andoa as part of the Ctenidiaceae.

Molecular data resolved a sistergroup relationship of Andoa with Myurium hochstetteri (Schimp.) Kindb., another Macaronesian endemic (Myuriaceae in Frey & Stech 2009), which together formed a well-supported clade with Ctenidium molluscum (Hedw.) Mitt. and Hyocomium armoricum (Brid.) Wijk & Margad. (Cox et al. 2010, Huttunen et al. 2012, Kučera et al. 2019). However, only single samples of Andoa and Myurium (composite sample of Myurium based on sequences of different specimen vouchers in Huttunen et al. 2012) were included in these phylogenies, and no study has so far investigated the molecular circumscription of Andoa and its delimitation from Myurium in more detail.

Andoa berthelotiana is currently present in the Madeira, Azores and Canary Islands archipelagos (Gabriel et al. 2010, Brugués & González-Mancebo 2014, González-Mancebo et al. 2012, Sim-Sim et al. 2014, González-Mancebo et al. 2019). The frequency of Andoa diminishes from Azores to the Canary Islands, which represents the meridional limit of the species (González-Mancebo et al. 2019). Andoa is common in the forests of northern and central Madeira Island but is also present in other areas of the island (Hedenäs 1992). Currently it is considered Vulnerable in Madeira (Sim-Sim et al. 2014). In the Azores it exists in all nine islands (Frahm 2005, Gabriel et al. 2010) not being currently under threat. In the Canary Islands Andoa is red-listed and considered Endangered (EN) (González-Mancebo et al. 2012, Brugués and González-Mancebo 2014), since it is known from only few populations on three islands: Tenerife and Fuerteventura, where it is very rare, and La Palma, where it is rare (González-Mancebo et al. 2008, 2019). In the
IUCN red list *Andoa berthelotiana* is currently assessed as Vulnerable (González-Mancebo et al. 2019).

Based on an integrative approach including analyses of DNA sequences (chloroplast *trnL-trnF*, nuclear ribosomal ITS, mitochondrial *nad5*), AFLP fingerprinting, and morphology, the present study aims at (i) revisiting the phylogenetic relationships of *Andoa berthelotiana* and its closest relatives, and (ii) assessing patterns of intraspecific variation and possible (cryptic) speciation within *A. berthelotiana*.

**Materials and Methods**

**Taxon sampling:** A total of 64 *Andoa berthelotiana* specimens collected between 1994 and 2010 were analyzed, covering all Macaronesian islands where the genus currently occurs except Corvo and Graciosa in the Azores archipelago. Of these, 13 specimens (Azores 7, Madeira 5, Canary Is. 1) were included in the phylogenetic, 51 (Azores trnL-trnF/ITS/nad5 27/11/6, Madeira 21/9/6, Canary Is.1/2/1) in the haplotype, 64 (Azores 34, Madeira 27, Canary Is. 3) in the AFLP, and 29 (Azores 19, Madeira 10) in the morphological analysis, respectively. The material from the Canary Islands was not enough for a morphological evaluation. For phylogenetic analysis, one newly sequenced *Myurium hochstetteri* specimen and 11 specimens from GenBank, viz. two *Andoa*, two *Myurium*, two *Ctenidium molluscum*, one *Hyocomium armoricum* as well as two Brachytheciaceae and two Meteoriaceae specimens as outgroup representatives (following Huttunen et al. 2012) were included in the data set. Voucher information and GenBank accession numbers for newly generated sequences are given in Appendix 1, accession numbers taken from Genbank in Appendix 2, and voucher information of samples used for DNA fingerprinting in Appendix 3.
DNA extraction, PCR and sequencing: The plants were cleaned manually with distilled water under a dissecting microscope and ultrasonic treatment, and dried with silical gel before DNA extraction. Genomic DNA was extracted using the protocol described in Quandt et al. (2007) and the DNeasy Plant Mini Kit (Qiagen). The chloroplast trnL-trnF, nuclear ribosomal ITS, and mitochondrial nad5 regions were amplified with primers and C and F (Taberlet et al. 1991), 18F and 25R (Stech & Frahm 1999), as well as nad5_4F and nad5_3R (Shaw et al. 2003), respectively. The amplification reactions (final volume 25 µl) included 1µl of diluted 1:10 stock DNA as template, 250µM of each dNTP, 2.5mM of buffer including MgCl₂, and 2 units of Taq DNA polymerase (HotStarTaq, Qiagen or BIOTAQ™, Bioline). The amplification conditions were: 2 min at 94°C, 35 cycles with 1 min at 94°C, 1 min at 55°C and 1 min at 72°C plus a final 5 min extension step at 72°C for trnL-trnF, 5 min at 94°C, 35 cycles with 45 s at 94°C, 45 s at 50°C and 1 min at 72°C plus a final 4 min extension step at 72°C for ITS, and 90 s at 96°C, 35 cycles with 45 s at 96°C, 1 min at 55°C and 1 min at 72°C plus a final 7 min extension step at 72°C for nad5. PCR products were checked by agarose gel electrophoresis, and afterwards cleaned and sequenced at Macrogen (www.macrogen.com) using the amplification primers.

Phylogenetic and haplotype analysis: The DNA sequences were manually aligned using PhyDE® v0.995 (Müller et al. 2006) according to the criteria laid out in Kelchner (2000) and Quandt & Stech (2005).

Separate phylogenetic reconstructions of each marker did not reveal incongruence in terms of well-supported clades. Accordingly, the datasets were combined. Phylogenetic reconstructions under maximum parsimony (MP) and maximum likelihood (ML) were performed with PAUP
Heuristic searches under parsimony were implemented by means of random sequence addition with 1000 replicates and employing the default settings otherwise. Gaps were treated as missing data or coded as informative by simple indel coding (SIC) (Simmons & Ochoterena, 2000) as implemented in SeqState (Müller 2004). Heuristic bootstrap searches under parsimony were performed with 1000 replicates and 10 random addition cycles per bootstrap replicate with the same options in effect. The best fit evolutionary models, inferred using MrModeltest 2.3 (Nylander 2004) under the Akaike information criterion (AIC) were GTR+I for \textit{trnL-trnF}, \textit{nad5} and the combined dataset, and HKY+G for ITS. Maximum likelihood analyses were executed in PAUP 4.0b10 under the model GTR+I with the settings inferred for the combined dataset: Lset Base=(0.2572 0.2253 0.2410), Nst=6, Rmat=(1.0300 2.2914 0.2561 1.0817 3.9282), Rates=equal, Pinvar=0.8303. Bayesian posterior probabilities were calculated using MrBayes v.3.1.2 (Huelsenbeck & Ronquist 2001). The settings Nst=6 as well as rates=propinv for \textit{trnL-trnF} and \textit{nad5}, and Nst=2 rates=gamma for ITS, were employed. In addition, a Bayesian analysis including indels coded by SIC was performed, with sequence and indel data treated as separated and unlinked partitions. The a priori probabilities used were the ones mentioned in the default settings of the program. Two runs with two chains (10 million generations each) were run simultaneously. Trees were sampled each 1,000th generation. Fifty-percent majority-rule consensus tree and posterior probabilities of clades were calculated by combining the runs and using the trees sampled after the chains converged (average deviation of split frequencies <0.01) and discarding the first 25% of the trees as “burn-in”.

Haplotype analyses were carried out with DNAsp5 (Librado & Rozas 2009) for each marker separately, using 13 \textit{nad5}, 22 ITS and 49 \textit{trnL-trnF} sequences. In addition, haplotype diversity (Hd) and nucleotide diversity (\Pi) were calculated.
**AFLP analysis:** AFLP fingerprinting followed Vos et al. (1995). DNA samples, each with 300 to 500 ng of DNA, were submitted to restriction-digestion by EcoRI/MseI endonucleases (5U each) at 37 °C for 16 hours. After ligation to their respective adapters during another incubation period of 16 hours at 37 °C, the samples were diluted 1:10 in ultrapure water. The first PCR amplification was carried out by using pre-selective primers which are complementary to the adapters with addition of one 3’ nucleotide (Eco-A and Mse-C primers). PCR products were diluted 1:10 for selective amplification, with the FAM-labelled fixed selective primer Eco-AAC together with selective primers Mse-CAA or Mse-CTG, respectively. Products of the selective amplifications were analyzed on a MegaBACE 1000 sequencer and fragments were scored using the MegaBACE Genetic Profiler Software Suite v2.2 (GE Healthcare Life Sciences Europe, Eindhoven, The Netherlands). The dataset obtained for each selective primer combination was transformed into a binary matrix (0 - fragment absent /1 – fragment present). Separate maximum parsimony analysis of the binary matrices of both selective primer combinations were performed with PAUP 4.0b10, heuristic searches under parsimony were implemented by means of random sequence addition with 1000 replicates and employing the default settings otherwise. The result did not show incongruence with respect to well supported branches, indicating that the matrices could be combined.

The combined AFLP data were analyzed by AMOVA (Analysis of Molecular Variance) and PCoA (Principal coordinates analysis), using GenAlex 6.5 (Peakall & Smouse 2012). In PCoA, the data analyzed is a similarity or distance matrix for a set of objects. The similarity or distance matrix is assembled from the data variables (Gower 1966), here the scored AFLP fragments. The AMOVA was carried out to assess the amount of variation within the populations and to
calculate pairwise PhiPT values (Fst analogue for binary data in GenAlex; Bolibok-Bragoszewska et al. 2014) to estimate genetic differentiation. The number of permutations used to estimate the probability values was 999 (Yamasaki & Ideta 2013). Following Yaacov et al. (2012), we considered PhiPT values 0.15–0.25 representing large, 0.05–0.15 moderate, and <0.05 little genetic variation with p-value = 0.001. The PCoA was used to investigate graphically the genetic relationships within *Andoa berthelothiana* among the different archipelagos.

**Morphological analysis:** Morphological characters scored from *Andoa berthelothiana* samples from Madeira and the Azores included plant size, leaf shape, leaf margins type, leaf size, length and width of alar, basal and median leaf cells, seta length and operculum length and type, length and width of capsule, exostome teeth and exothecial cells, and spore size. For the capsule the color and shape were also registered. Average and median values were calculated for the quantitative traits. Per specimen 5 leaves were measured, in each leaf 5 alar, basal and median leaf cells were measured. Due to the lower number of sporophytes available in proper conservation state, the number of setae and capsules measured by specimen, when present, varied from 1 to 5. Per specimen also 5 exostome teeth, 5 exothecium cells and 5 spores were measured, all randomly selected.

**Results**

**Phylogeny:** The combined *trnL-trnF*, ITS and *nad5* alignment comprised 2576 positions, of which 95 were parsimony-informative. With indels coded by SIC included, the combined dataset comprised 2729 characters, of which 174 were parsimony-informative. Phylogenetic reconstructions (Fig. 1) resolved *Andoa berthelothiana* as monophyletic, but with statistical
support only in the maximum parsimony analysis with indels included (bootstrap support [BS] 90%). Andoa appeared as sister to a well-supported clade of Myurium hochstetteri (MP BS 98%–100%, ML BS 99%, Bayesian inference posterior probability [PP] 1). This sistergroup relationships was supported by MP BS 99–100%, ML BS 90%, PP 0.99–1. The clade of both Ctenidium molluscum accessions received maximum support and was sister to Hyocomium armoricum without support. The clade of all four ingroup genera was maximally supported in all analyses. Within Andoa berthelotiana, two clades were resolved, one from Madeira and the Canary Islands (MP BS with indels 85%, PP 0.97–1), and another one from the Azores without support.

**Haplotype networks:** The haplotype networks resulting from the ITS (431 positions, excluding sites with gaps/missing data), nad5 (909), and trnL-trnF (378) alignments are shown in Fig. 2. Numbers of haplotypes were 4 for ITS (Fig. 2A), 2 for nad5 (Fig. 2B), and 3 for trnL-trnF (Fig. 2C). Haplotype diversity (Hd) and nucleotide diversity (Π) were 0.636 / 0.00183 for ITS, 0.538 / 0.00059 for nad5, and 0.583 / 0.00175 for trnL-trnF. All samples from Madeira and Canary Islands shared one ITS and nad5 haplotype, whereas the samples from Azores were divided into three haplotypes (ITS) or formed a single haplotype (nad5), respectively. The most frequent Azorean ITS haplotype comprised one sample each from the islands of Faial, São Jorge, Santa Maria, and Terceira as well as three samples from Pico, the second most frequent haplotype comprised one sample each from Flores and São Jorge, and the least frequent haplotype was composed of a single sample from Pico. In the trnL-F haplotype network, one haplotype was shared by samples from all three archipelagos (Madeira 12; Azores 11: two each from São Jorge and Pico, three from Faial, one each from Flores, Terceira and São Miguel; Canary Islands one),
another by samples from Azores (13: three each from São Jorge, Terceira and Pico, two from Flores and one from Santa Maria) and Madeira (9), and the third haplotype included exclusively samples from the Azores (two from Terceira, one each from Flores and Pico).

**AFLP analysis:** The AFLP fingerprinting resulted in 549 discernible bands for both selective primer combinations concatenated. The PCoA analysis of this dataset (Fig. 3) showed a clear differentiation between the *Andoa berthelotiana* specimens from Madeira plus Canary Islands archipelagos versus those from the Azores. Within the Azores, the samples belonging to Faial (1158/1159/1160/1541/1542/1543) are closer together, in the upper left quarter of the coordinate space. Five out of six samples from Terceira stand close, in the lower left quarter, forming two groups (1116, 1215 and 1219 very close together; 1127 very close to 1220). Pairwise PhiPT values were 0.057 between Madeira and Canary Islands, 0.208 between Madeira and Azores, and 0.194 between Canary Islands and Azores, with p-value=0.001.

**Morphological observations:** The studied morphological traits are summarized and compared between Madeiran and Azorean samples in Table 1. Gametophyte characters (alar, basal, median cells and leaf size) and sporophyte characters (capsule, exothecial cells and exostome tooth, spore) for which a length/width ratio was calculated, as well as seta length, are represented graphically by boxplots in Fig. 4 and Fig. 5, respectively. Plants from Madeira and the Azores tend to be different in several characters. In Madeira plants seem to be bigger and the leaves longer and less wide when comparing to the Azores. Alar cells are also apparently longer in Madeira while the basal cells and seta are seemingly longer in the Azores. In addition, the exothecial cells length/width ratio is higher in the Azores and the exostome tooth length/width
ratio is higher in Madeira archipelago.

Discussion

The present study confirms the results of earlier phylogenetic analyses (Cox et al. 2010, Huttunen et al. 2012, Kučera et al. 2019) that _Andoa berthelotiana_ is closest to _Myurium hochstetteri_, and that both species form a well-supported lineage of pleurocarpous mosses together with _Ctenidium molluscum_ and _Hyocomium armoricum_. This lineage corresponds to the family Myuriaceae, a small family that has so far been considered to comprise 17 species in four genera, _Eumyurium, Myurium, Oedicladium_, and _Palisadula_ (Frey & Stech 2009), which is extended by a part of the genera formerly placed in the subfamily Ctenidioideae of the Hypnaceae (_Andoa, Ctenidium, Hyocomium_; Nishimura et al. 1984, Kučera et al. 2019). The exact circumscription of Myuriaceae has still to be determined based on molecular data, since Arikawa et al. (2008) resolved _Oedicladium_ as separate from _Myurium_ and _Ctenidium_ based on _rbcL_ sequences. That study did, however, provide support for the monophyly of _Ctenidium_, whereas other studies have only included one species, either _Ctenidium malacodes_ Mitt. (Buck et al. 2000) or _C. molluscum_ (Cox et al. 2010, Huttunen et al. 2012, Kučera et al. 2019).

The amended Myuriaceae remain predominantly Eurasian (including Macaronesia), except _Ctenidium_, which is present on four continents, the Americas (Parra et al. 2002, Staples et al. 2004, Costa et al. 2011, Shevock et al. 2019), Asia (Redfearn Jr. 1990, O’Shea 2002), Europe (Hodgetts et al. 2020) including Macaronesia (Azores, Gabriel et al. 2010; Canary Islands – La Palma, Losada-Lima et al. 2004, absent in Madeira, Hodgetts & Lockhart 2020) and Oceania (Tan 2000, Higuchi & Nishimura 2006). _Ctenidium molluscum_ is the single representative of the
genus in Europe and Macaronesia. *Myurium* is currently known from Macaronesia (Azores, Madeira and Canary Islands) plus Scotland and was present at a single location in Ireland, where it is currently considered extinct (Sim-Sim et al. 2019). *Hyocomium armoricum* occurs in Europe (Hodgetts et al. 2020) including Macaronesia (Azores, Frahm 2005; Madeira, Sérgio et al. 2006, Sjögren 2006), and extends to northern Africa (Ros et al. 2000). Morphologically, *Andoa, Myurium* and *Ctenidium* have in common the dioicous sexual condition and the associated (facultative) presence of dwarf males, which may shorten effective fertilization distances in dioicous mosses (Hedenäs & Bisang 2011). *Andoa* has facultative dwarf males, while *Myurium* has obligate dwarf males (Hedenäs & Bisang 2011). Several *Ctenidium* species present facultative dwarf males, such as *Ctenidium molluscum* and *Ctenidium malacodes* Mitt., where others present apparently obligate ones, such as *Ctenidium capillifolium* (Mitt.) Broth. (Hedenäs & Bisang 2011). *Hyocomium* is dioicous as well, but dwarf males have not been observed (Hedenäs & Bisang 2015). Other characters shared by the four genera in different combinations include julaceous stems (*Andoa, Myurium*), prorate lamina cells (*Andoa, Ctenidium, Hyocomium*), weakly differentiated (*Andoa, Ctenidium, Hyocomium*) or few differentiated (*Myurium*) alar cells, and a whitish-yellow (*Andoa*) or whitish-hyaline (*Myurium*) exostome (Nishimura 1985, Hedenäs 1992, Frey & Stech 2009). The current study indicates the presence of two molecular lineages within *Andoa berthelotiana*, one occurring in Madeira and the Canary Islands, and the other in the Azores. Although sequence variation in the DNA markers is low and the two groups receive no or only low statistical support in the phylogenetic trees, the genetic differentiation (PhiPT) between the Azores and the other two archipelagos is clearly higher than between Madeira and Canary Islands. Furthermore, the AFLP fingerprinting data corroborates the phylogenetic and ITS/nad5
haplotype analyses, resolving the same two geographically separated lineages. The closer molecular relationship between Madeira and the Canary Island populations is paralleled by the greater similarity between the bryophyte floras of the western Canary Islands (namely La Palma and Tenerife) and Madeira, especially as far as mosses are concerned (González-Mancebo et al. 2008). There are several species shared by Madeira and the Canary Islands, which are close relatives to species present in the Azores, such as in the genera Echinodium and Leucodon (Stech et al. 2008, González-Mancebo et al. 2009, Stech et al. 2011).

The trnL-trnF haplotype network seems to contradict the other data concerning the division Azores vs. Madeira/Canaries in Andoa berthelotiana. However, the trnL-trnF haplotypes are based only two positions with A/C substitutions. The first one is located in the loop of the hairpin P9.1 of the trnL intron secondary structure (Quandt & Stech 2005), and probably has occurred several times independently. The second one is located in a variable area at the end of the trnL-trnF spacer, and probably has also occurred several times independently. Consequently, these mutations might actually not be phylogenetically informative.

Cases of cryptic speciation in bryophytes have increasingly been reported based on molecular data (reviewed in Renner et al. 2020). However, as Renner et al. (2020, p. 41) pointed out, “fully two-thirds of all studies on cryptic bryophyte species rested their claims of morphological cryptis on previous taxonomic investigations, without revision of morphology to confirm cryptic species status”. Considering the frequently subtle morphological differences between bryophyte species, an integrative approach is needed to infer to which degree morphological and molecular divergence are actually decoupled in cryptic species. Struck et al. (2018) proposed a conceptual framework that would recognize cryptic species based on low levels of morphological variation relative to their level of genetic differentiation and divergence times, and proposed an
interdisciplinary approach to test for significant differences in rates of morphologically
discrepancy between cryptic and non-cryptic species. Re-analyzing morphological variation in
comparison with molecular variation indeed revealed, or clarified, morphological discontinuities
among (possibly cryptic) Macaronesian bryophytes already, resulting in their formal description
at the appropriate taxonomic level (e.g. Hedenås et al. 2014, Draper et al. 2015). According to
the present morphological observations, the two molecular lineages of *Andoa berthelotiana* show
variation in several characters that coincides with the resolved molecular lineages. However, the
considerable morphological overlap between both molecular groups would hamper the
identification of specimens if both lineages were formally described as separate taxa. As a
preliminary conclusion we therefore suggest to treat the two lineages of *Andoa berthelotiana* as
semi-cryptic species. The presence of fossils identified as *A. cf. berthelotiana* from the Middle
Miocene (Frahm 2004, 2007) might suggest a Continental European origin and subsequent
diversification in Macaronesia, where different ecological constraints might have led to the
genetic differentiation between plants from the Azores and the other two archipelagos. Further
study, including morphological analyses of Canary Islands populations as well as extended
sampling and molecular dating analyses of the Myuriaceae is necessary to elucidate the evolution
of the *Andoa* lineages. From a conservation point of view, the present results already provide
important information, emphasizing the need of protecting the habitats where *Andoa* occurs, in
particular in Madeira and the Canary Islands, where the species is under threat, to conserve its
phylogenetic diversity.

**Acknowledgements**
Sincere thanks are due to Foundation of Science and Technology for the PhD scholarship SFRH/BD/46274/2008, Madeira Botanical Garden former Director Luisa Maria Gouveia for permission to collect material and help with the field work and herbarium material observation, Drª. Susana Fontinha for help with logistics and field work in Madeira, Carlos Lobo for help with logistics and field work, Madeira Herbarium for allowing observation of herbarium material, Direção Regional dos Açores and Eng. Carlos Pimentel for permission to collect material and help with the field work, Prof. Rosalina Gabriel for providing advice and help with field work in the Azores, Eduarda Goulart, Eng. José Costa and co-worker forest guard for helping with logistics and field work in the Azores, Drª. Cecília Sérgio for providing information and advice regarding Andoa, Dr. César Garcia for providing help with field work, Profª. Juana González-Mancebo for providing samples from the Canary Islands, Dr. G. M. Dirkse for providing samples from S. Miguel Island, the Collection department of Naturalis Biodiversity Center for support with collection work, and the technicians of the laboratories of Naturalis Biodiversity Center and Institute of Biology, Leiden University, support of the molecular lab work.

References


Nylander, J.A.A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University. – Available from: https://github.com/nylander/MrModeltest2


Table 1. Main morphological traits of *Andoa berthelotiana* in Madeira and the Azores based on the studied material.

<table>
<thead>
<tr>
<th>Character</th>
<th><em>Andoa berthelotiana</em> – Madeira (n=10)</th>
<th><em>Andoa berthelotiana</em> - Azores (n=19)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant size</td>
<td>Small to medium-sized, stems up to 10 cm long</td>
<td>Small to medium-sized, stems up to 9 cm long</td>
</tr>
<tr>
<td>Leaf shape</td>
<td>From cordate base ovate-triangular; some long triangular, then upper part subulate, some acuminate (1/2 to 1/3 of leaf length)</td>
<td>From cordate base, (long-)ovate; some more triangular or long triangular, then subulate, some acute or acuminate (1/2 to 1/3 of leaf length)</td>
</tr>
<tr>
<td>Leaf margins</td>
<td>Denticulate, teeth increasing in size towards subula, upper part toothed</td>
<td>Denticulate, teeth increasing in size towards subula, upper part slightly toothed or toothed</td>
</tr>
<tr>
<td>Leaf size</td>
<td>To 0.274 mm long and 0.80 mm wide</td>
<td>To 0.284 mm long and 0.108 mm wide</td>
</tr>
<tr>
<td>Basal cells</td>
<td>19.6–61.4 μm x 3.3–11 μm</td>
<td>15.2–72.4 μm x 4.4–9.9 μm</td>
</tr>
<tr>
<td>Alar cells</td>
<td>6.6–28.4 μm x 5.5–13.2 μm</td>
<td>5.5–28.4 μm x 4.4–13.2 μm</td>
</tr>
<tr>
<td>Median cells</td>
<td>41.6–74.6 μm x 3.3–7.7 μm</td>
<td>37.2–94.2 μm x 2.2–5.5 μm</td>
</tr>
<tr>
<td>Seta length</td>
<td>2–4.3 cm long</td>
<td>2–3.7 cm long</td>
</tr>
<tr>
<td>Capsule</td>
<td>Reddish orange, horizontal and cylindrical</td>
<td>Reddish orange, horizontal and cylindrical</td>
</tr>
<tr>
<td>Component</td>
<td>Size 1</td>
<td>Size 2</td>
</tr>
<tr>
<td>--------------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Spores</td>
<td>12.1–18.5 µm x 11–17.4 µm</td>
<td>14.2–20.7 µm x 11–19.6 µm</td>
</tr>
<tr>
<td>Exothelial cells</td>
<td>22.5–45 µm x 13.5–36 µm</td>
<td>13.5–40 µm x 13.5–27 µm</td>
</tr>
<tr>
<td>Operculum</td>
<td>Rostellate (0.4 mm)</td>
<td>Rostellate (0.6–0.7 mm)</td>
</tr>
<tr>
<td>Exostome teeth</td>
<td>350–546 µm x 63–112 µm</td>
<td>378–570 µm x 77–112 µm</td>
</tr>
</tbody>
</table>
Fig. 1. Bayesian phylogenetic analysis of *Andoa berthelotiana* and its potential closest relatives within Hypnales, based on the combined plastid *trnL-trnF*, mitochondrial *nad5* intron sequence and data nuclear ribosomal ITS, including indels coded by simple indel coding according to Simmons and Ochoterena (2000). Four species of the Hypnales grade (Meteoriaceae: *Papillaria crocea* (Hampe) A. Jaeger and *Meteorium polytrichum* Dozy & Molk.; Brachytheciaceae: *Pseudoscleropodium purum* (Hedw.) Fleisch. and *Brachythecium rivulare* Schimp.) were used as outgroup representatives. Clade support indicated on the branches include: maximum parsimony bootstrap values 70% without and with simple indel coding / maximum likelihood bootstrap support 70% / Bayesian posterior probabilities >95 without and with simple indel.
coding

**Fig. 2.** Haplotype networks resulting from alignments of A ITS (20 samples), B *nad5* (13) and C *trnL-trnF* (47) sequences of *Andoa berthelotiana*. Black circles indicate haplotypes from Madeira + Canary Islands, white from Azores, and grey from Azores + Madeira + Canary Islands (left circle in C) or Azores + Madeira (central circle in C), respectively. Circle size corresponds to number of sequences sharing the same haplotype. Hypothetical intermediate haplotypes are indicated by small dots.
**Fig. 3.** PCoA analysis of *Andoa berthelotiana* samples from Madeira (grey diamonds), the Canary Islands (black squares) and the Azores (white triangles), for the first two axes based on the AFLP distance matrices.

**Fig. 4.** Boxplots of length/width ratios of gametophyte characters: a) alar cells, b) basal cells, c) median cells, d) leaf size. In soft grey the values for Azores and in dark grey for Madeira.
**Fig. 5.** Boxplots of length/width ratios of sporophyte characters: a) capsules, b) exothecial cells, c) exostome tooth, d) spore as well as e) seta length. In soft grey the values for Azores and in dark grey for Madeira.