Comparative Analysis of *Apopellia endiviifolia* Plastomes Reveals a Strikingly High Level of Differentiation between Its Terrestrial and Water Form

Jakub Sawicki [1,*], Katarzyna Krawczyk [1], Monika Ślipiko [1], Kamil Szandar [1] and Monika Szczecińska [2]

[1] Department of Botany and Nature Protection, University of Warmia and Mazury in Olsztyn, Plac Lódzki 1, 10-727 Olsztyn, Poland; katarzyna.krawczyk@uwm.edu.pl (K.K.); monika.slipiko@uwm.edu.pl (M.Ś.); kamil.szandar@uwm.edu.pl (K.S.)

[2] Department of Ecology and Environmental Protection, University of Warmia and Mazury in Olsztyn, Plac Lódzki 3, 10-727 Olsztyn, Poland; monika.szczecinska@uwm.edu.pl

* Correspondence: jakub.sawicki@uwm.edu.pl

Abstract: The simple thalloid liverwort *Apopellia endiviifolia* is a widespread Holarctic species belonging to the family Pelliaceae. European populations of this species comprise two distinct evolutionary lineages named "species A", known also as water form, and typical, mainly terrestrial forms named "species B". Newly sequenced, assembled and annotated chloroplast genomes of six European specimens belonging to the two cryptic lineages occupying different microhabitats, revealed the structure typical for liverworts and previously sequenced reference. The plastomes of *A. endiviifolia* are 120,537–120,947 bp long with a structure typical for most plants, including a pair of IR regions (each of 9,092–9,207 bp) separated by LSC (82,506–82,609 bp) and SSC (19,854–19,924 bp) regions and consist of 121 unique genes, including 81 protein-coding genes, 6 genes of unknown function (*ycf* genes), 4 ribosomal RNAs and 30 transfer RNAs. Comparative analysis of typical, terrestrial and water forms revealed 4971 molecular diagnostic characters (MDCs), which exceeds numbers found in many well recognized liverworts taxa. Moreover, beside the presence of evolutionary hotspots like *ycf*1 and *ycf*2 genes and several intergenic spacer like *ndh*B-*psb*M, *rps*4-*ndh*J and *ndh*C-*atp*E, the molecular identification of *Apopellia* cryptic species was possible by almost 98% of 500 bp long frames simulating mini barcodes. The different ecological niches can be driven by different pressures of positive selection, which was detected in nine genes including *ccs*A, *ndh*D, *ndh*F, *pet*A, *psb*B, *psb*C, *rpo*B, *ycf*1 and *ycf*2. Despite clearly genetic differences and ecological preferences, the current observation of morphological differentiation does not no allow to separate terrestrial and water forms into taxonomic species.

Keywords: *Apopellia endiviifolia*; plastome; intraspecific variation; cryptic speciation; superbarcodes; positive selection

1. Introduction

The cryptic speciation is well known to occur in all major lineages of liverworts [1–6]. In recent years integrative taxonomy approach led to description of several species of complex thalloids and leafy liverworts, previously considered as cryptic lineages distinguishable only on the basis of molecular markers [3,7].

The simple thalloid liverwort genus *Pellia* Raddi *sensu lato* is a widespread Holarctic taxon belonging to the family Pelliaceae, defined by several morphological characters. Thallus of *Pellia s.l.* is up to 1 cm wide, dichotomously branched with a poorly defined midrib. The study of cryptic speciation of genus *Pellia s.l.* has a long history, with the first report published at the beginning of the 20th century [8]. Popularization of the isoenzyme electrophoresis methods enabled characterisation of cryptic evolutionary lineages within *Pellia epiphylla* (L.) Corda. and *P. endiviifolia* (Dicks.) Dumort. species [9,10], which were

later confirmed by various markers based on direct DNA analyses [11–13]. Recently, an integrative approach based on morphology and molecular dataset enabled the split of the genus *Pellia* into *Apopellia* and *Pellia sensu stricto* [14]. The former comprises three taxonomic species, *A. apicola*, *A. megaspora* and *A. endiviifolia,* with different patterns of distribution. The European population of *Apopellia* is diverging into two evolutionary lineages, which, besides molecular characters, differ in types of microhabitats [8,9,12].

The plastid genomes are the most frequently used molecules in evolutionary studies on plants, including phylogenomics, population genetics and phylogeography [15,16]. Complete plastome sequences serve as superbarcodes enabling identification even recently divergent species [17,18] where typical single and multilocus barcodes fail. In opposition to vascular plants, infraspecific plastome diversity of liverworts are yet poorly explored.

Liverworts plastomes, as well as well-studied angiosperm plastomes, exhibit two copies of inverted repeats (IRa and IRb) separated by two single copy regions: the large single copy (LSC) and the small single copy (SSC). In contrast to vascular plants, where plastome structure variation is observed at the family or even at genus level [19,20], in liverworts the gene content and gene order in the plastome is conserved in all main liverwort lineages, including complex thalloid, simple thalloid, and leafy species [21–23], despite it being millions of years since their divergence [24].

Chloroplast genome of *Apopellia endiviifolia* was among the first sequenced liverworts plastomes [25]; however, up to date, detailed comparative study of plastomes of cryptic lineages within it was not performed. The recent advances in the next generation sequencing provided better datasets to study cryptic speciation of liverworts [26–28] and enabled detailed view into plastomes divergence between cryptic evolutionary lineages.

In this study we sequenced and analyzed six complete chloroplast genomes of two cryptic species of *Apopellia endiviifolia* with contrasting habitat preferences to determine infra and interspecific differentiation of plastomes. The potential application of chloroplast genomes as superbarcodes to identification of cryptic species within the *A. endiviifolia* complex was analyzed using different approaches allowing reliable delimitation of water and terrestrial forms.

2. Materials and Methods

DNA was isolated by the modified CTAB procedure. The liquid nitrogen-ground thalii were thoroughly mixed with 3 mL preheated CTAB isolation buffer (2% CTAB, 100 mM Tris-HCl, pH 8.0, 20 mM EDTA, 1.4 M NaCl and 2% β-mercaptoethanol) and incubated at 55 °C for 1 h. After three chloroform extractions, the DNA was precipitated and dissolved in sterile, deionized H_2O. The purity of DNA samples was assessed spectrophotometrically and it reached 90–95%. DNA quantity was estimated using the Qubit fluorometer and Qubit™ dsDNA BR Assay Kit (Invitrogen, Carsbad, NM, USA). DNA quality was checked by electrophoresis in 0.5% agarose gel stained with Euryx Simple Safe (Eurx, Gdańsk, Poland). The extracted DNA prior to long-read sequencing was carefully examined and additionally cleaned using the Genomic DNA Clean and Concentrator kit (Zymo, Irvine, CA, USA).

The preliminary delimitation based on ecological preferences and macromorphological features was confirmed using ISJ2 (intron–exon splice junction) markers, which clearly separated both cryptic lineages [11]. The PCR reaction details and primer sequences were identical to previously published [29].

The genomic libraries were constructed with Qiagen FX DNA kit (Qiagen, Hilden, Germany) and were sequenced using HiSeqX (Illumina, San Diego, CA, USA) to generate 150 bp paired-end reads at Macrogen Inc. (Seoul, Korea) with 350 bp insert in size between paired-ends. Afterwards, sequencing reads were cleaned by removing the adaptor sequences and low-quality reads with Trimmomatic v0.36 [30]. The long-read libraries were constructed using Ligation Sequencing Kit SQK-LSK109 (Oxford Nanopore Technologies, Oxford, UK) and NEBNext® Companion Module for Oxford Nanopore Technologies® Ligation Sequencing (New England Biolabs, Ipswich, MA, USA) according to manufac-

turer's protocol and sequenced using MinION MK1B portable device (ONT, Oxford, UK) and R.9.4.1 Flow Cell (ONT). The Flow Cell was prepared for sequencing with the Flow Cell Priming Kit EXP-FLP002 (ONT). Sequence reads were basecalled using high-accuracy guppy basecalling on the MinKNOW platform. The *Apopellia* chloroplast genomes were assembled using NOVOplasty [31] using previously sequenced *A. endiviifolia* chloroplast genome (NC_019628) as reference. The GeSeq 2.03 software was used for annotation [32] and genome maps were drawn using OGDRaw v1.1.1 web server [33]. The specimen details, including GenBank accession numbers of newly assembled genomes are given in Supplementary Table S1. The raw reads of all used libraries were submitted to the National Center for Biotechnology Information (NCBI, Bethesda, MD, USA) under Sequence Read Archive (SRA) number PRJNA785811.

The complete chloroplast genomes of *Apopellia endiviifolia* from this study and previously sequenced plastome from GenBank were compared using mVISTA 1.4 [34] with the LAGAN 2.0 alignment program [35]. NC_019628 accession (typical form) used as a reference.

Comparative analysis of chloroplast genomes was carried out in the Spider 1.1-5 program [36] based on inter- and intraspecific distances that were calculated using the Kimura 2-parameter model (K2P) of nucleotide substitution. Barcoding analyzes of entire *Apopellia* chloroplast genomes and their 500 bp-long fragments generated by sliding window were made in Spider. The sliding window nucleotide diversity analysis was performed using TASSEL 5.0 software [37]. Ten most variable regions were extracted, realigned using MUSCLE v5 [38] and analyzed using TASSEL 5.0.

The discrete molecular diagnostic characters (MDCs) for each species were calculated according to the Jörger and Schrödl [39] approach using FASTACHAR 0.2.3 software [40].

The Poisson Tree Processes (PTP) method was applied to delimitate species boundaries [41]. The analysis was performed using rooted ML (Maximum likelihood) and BI (Bayesian inference) trees, the MCMC algorithm was run for 1,000,000 generations, with 100 thinning and 0.2 burn-in.

As the backbone of phylogenetic analysis, we used plastid CDS (coding DNA sequence) alignment from previous study [23]. Existing dataset, which includes a set of plastid genes, was expanded by adding CDS from six newly assembled plastomes of *Apopellia*. To keep a congruence with a previous study, we applied Maximum Likelihood and Bayesian inference methods using RAxML 8.0 [42] and MrBayes 3.2 [43] with parameters used by Yu et al. [23].

Mixed Effects Model of Evolution (MEME) analysis was used to detect individual sites evolving under positive selection in a proportion of branches [44]. The BUSTED algorithm [45] was used to identify episodic positive selection in cryptic lineages. To test branch-specific positive selection, we used the aBSREL algorithm [46] which uses a model selection procedure for each branch to determine the proportion of sites per branch that have undergone positive selection. All the three tests were performed using Datamonkey server [47] with default settings.

3. Results and Discussion

3.1. Characteristics of Newly Sequenced Apopellia Chloroplast Genomes

Liverwort plastid genomes are known for their conserved structure and results obtained in this study confirm this observation. The *Apopellia* plastome was among the first sequenced liverwort plastomes [25] and the newly sequenced *Apopellia* specimens have identical gene order and content. The plastomes of *A. endiviifolia* are 120,537–120,947 bp long with a structure typical for most plants, including a pair of IR regions (each of 9,092–9,207 bp) separated by LSC (82,506–82,609 bp) and SSC (19,854–19,924 bp) regions (Figure 1). As in the most known liverworts plastome sequences [23], plastomes of *Apopellia* consist of 121 unique genes, including 81 protein-coding genes, 6 genes of unknown function (*ycf* genes), 4 ribosomal RNAs and 30 transfer RNAs. The most significant change in the water form of *Apopellia* is a length reduction of *cys*T and *ycf* 66 genes by 37 and

6 amino acids, respectively, due to the presence of termination codons. The total plastome size of a typical *A. endiviifolia* form ranged from 120,537 to 120,546 bp, while water forms fall within 120,944–120,946 bp range.
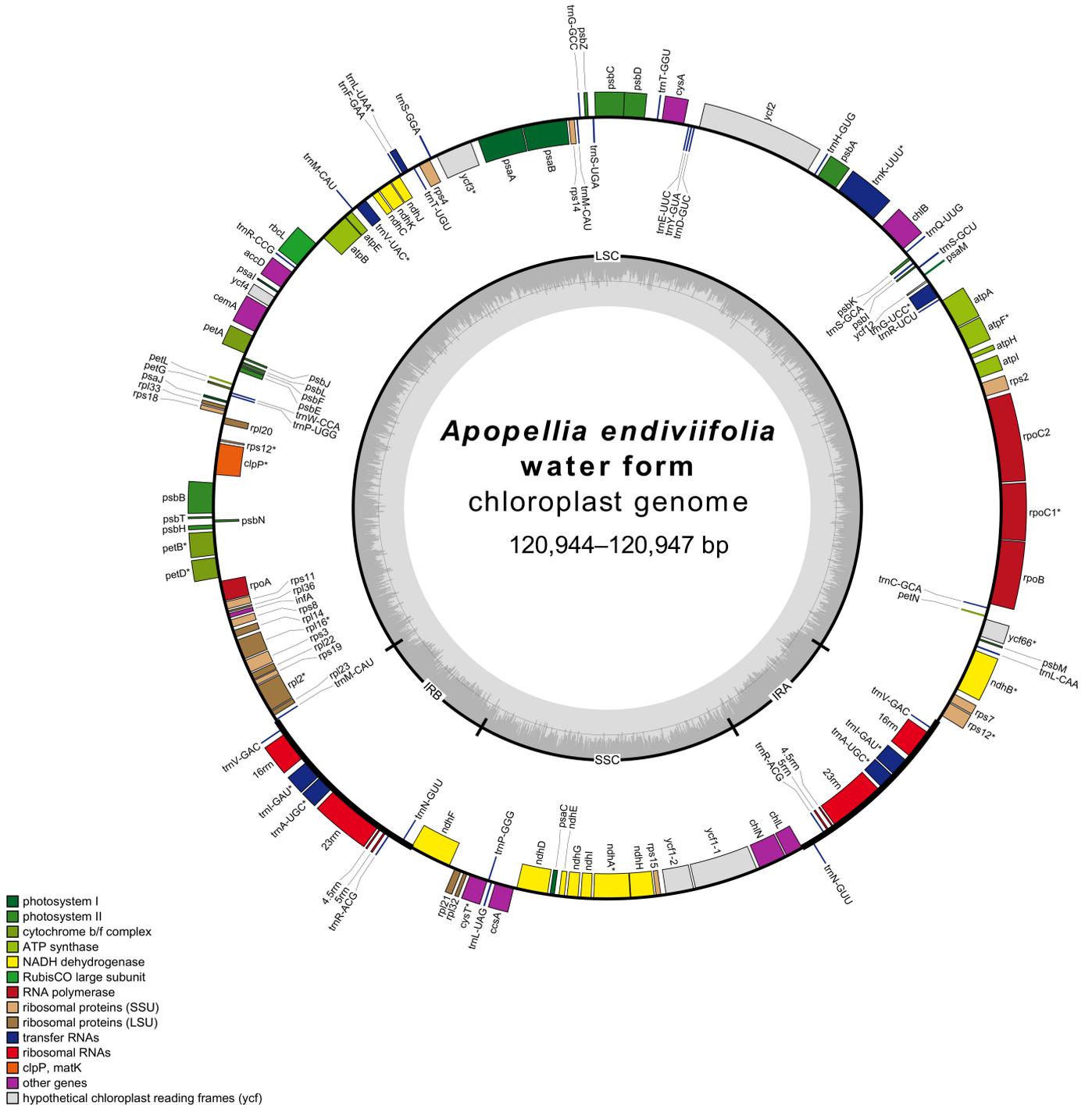

Figure 1. Gene map of the chloroplast genome of *Apopellia endiviifolia* water form. Genes inside and outside the outer circle are transcribed in counterclockwise and clockwise direction, respectively. The genes are color-coded based on their function. The inner circle visualizes the GC content. * indicate genes containing introns.

The gene order and -content seem to be stable in liverworts, but some gene losses are scattered along different evolutionary lineages including loss of *cys*A and *cys*T genes in *Ptilidium pulcherrimum* and *Cheilolejeunea xanthocarpa* [21,23] and additional four genes

(*ndh*F, *rpl*21, *rps*32 and *ccs*A) in *Cololejeunea lanciloba* [23]. Single genes are also missing in *Aneura pinguis* (*ycf*66), *Haplomitrium blumei* (*psa*M) and *Schistochila macrodonta* (*ndh*F) [23,26]. Heterotrophic *Aneura mirabilis* has a reduced plastome due to the loss of six genes (*ndh*A-E-G-K-I and *ycf*66) or pseudogenization (19 genes) of genes involved in the process of photosynthesis [48]. Two liverworts species increase the number of plastid genes by transferring them to IRs: *Haplomitrium blumei* transfered *ndh*F gene from SSC [23], while *Conocephalum salebrosum* transfered *rps*7 and *rps*12 from LSC [27].

3.2. Barcoding and Molecular Identification of Apopellia Cryptic Species

Application of organellar genomes as superbarcodes enabled identification of both forms of *A. endiviifolia*. Using the barcoding gap method, molecular delimitation was possible using 2375 out of 2427 generated 500-bp long frames. The mean number of diagnostic nucleotides was 20.7 per 500 bp, which is 4.41% of nucleotides. Number of diagnostic nucleotides per frame ranged from 0 to 51 in IR regions and from 1 to 144 in hotspots of single copy regions (Figure 2). Comparative analysis using the mVista program revealed that most of the variation among *A. endiviifolia* cryptic species was found in intergenic regions of LSC and SSC parts of the plastome (Supplementary Figure S1). Several intergenic spacers like *ndh*B-*psb*M, *rps*4-*ndh*J and *ndh*C-*atp*E could be identified as hotspots of interspecific variation (Figure 3, Supplementarty Table S3). Among coding regions, *ycf*1 and *ycf*2 were revealed as the most variable which is congruent with data obtained from other liverworts genera like *Aneura* [26] and *Calypogeia* [49]. *Ycf*1 and *ycf*2 are also considered as promising barcodes in vascular plants [50,51].
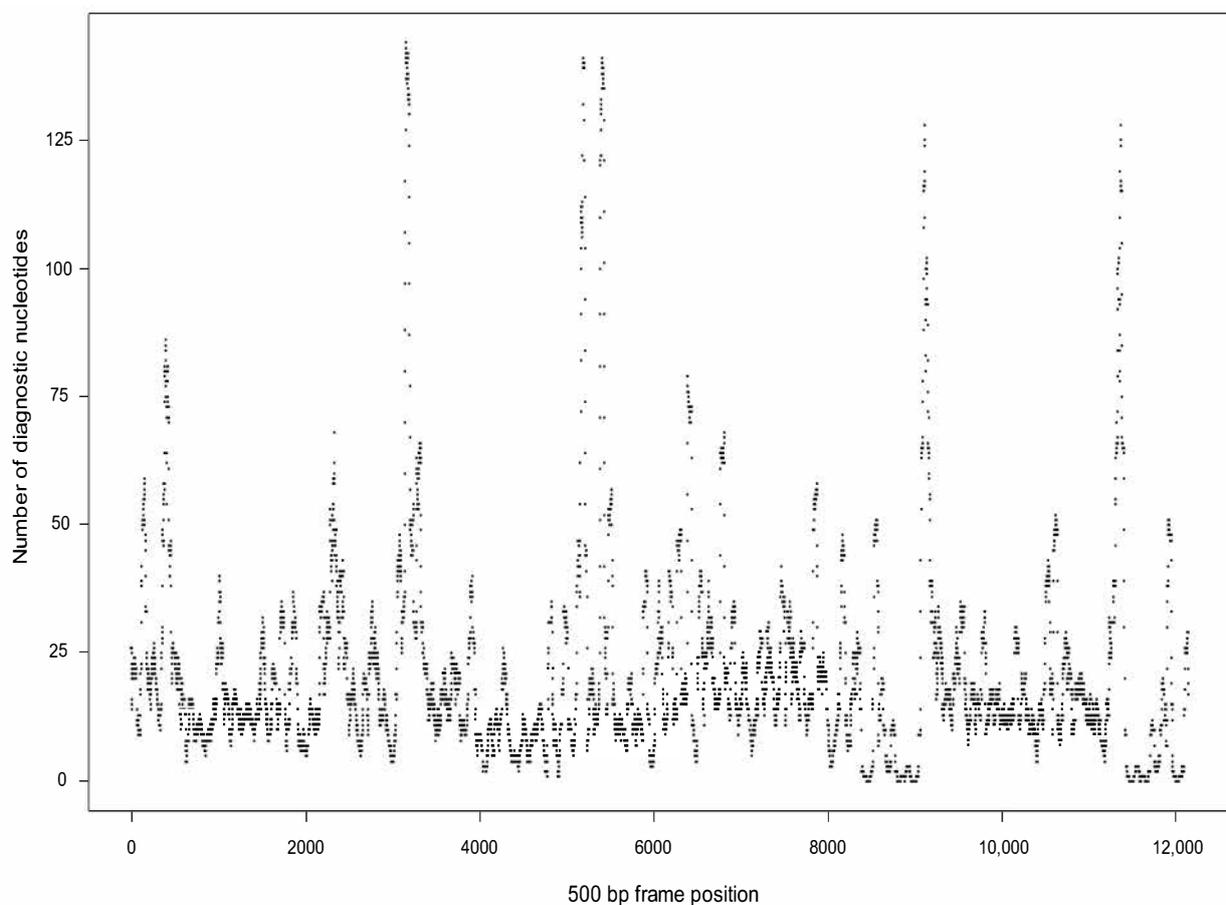
Figure 2. The sum of diagnostic nucleotides position in 500bp frames for terrestrial and water forms of *Apopellia endiviifolia*.
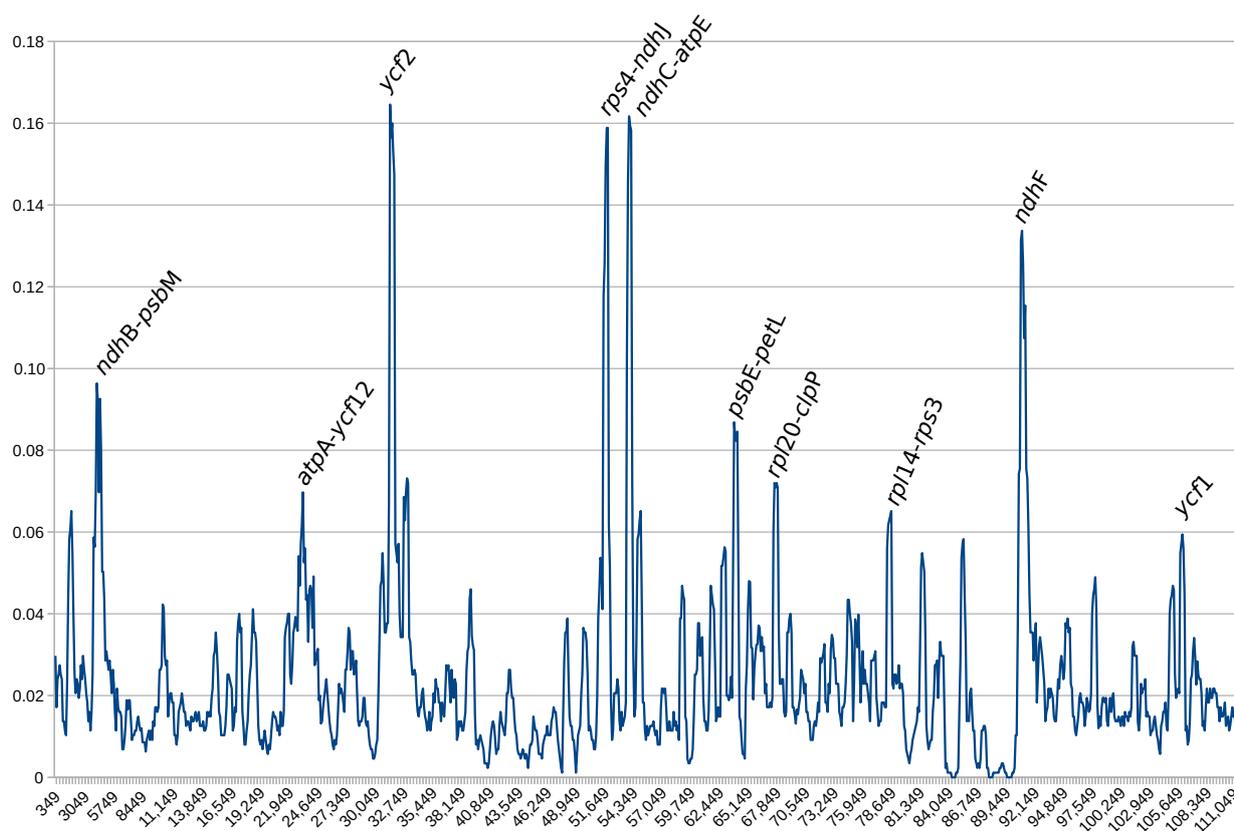
Figure 3. Nucleotide diversity (Pi) by sliding window analysis in the aligned whole plastomes (excluding second IR) of seven *Apopellia* specimens. Window length 500 bp, step size 100 bp.

The biggest interspecific variation within IRs was found in the fragment flanking SSC region.

Whole plastome analysis (excluding one copy of IR) revealed 4971 molecular diagnostic characters (MDCs) enabling molecular delimitation of typical and water forms, which comprised 4.4% of the plastome. The presence of significant numbers of MDCs is crucial for splitting and description of new species [52]. Low variation at the infraspecific level resulted in only six potential MDCs as defined in CAOS [53] and QUIDDICH [54]. The number of MDCs that differentiate cryptic species of *A. endiviifolia* is much higher than identified for good taxonomic species of leafy liverwort genus *Calypogeia* where chloroplast MDCs ranged from 159 to 1369 [49]. The lower number of MDCs was also identified for species of complex thalloid liverwort genera *Marchantia* and *Conocephalum* [27]. The liverwort model species *Marchantia polymorpha* and *M. palaceae* revealed 4076 MDCs, while recently described *Conocephalum salebrosum* [3] and *C. conicum* were characterized by 5878 MDCs.

The plastid superbarcoding approach is getting more attention as an effective tool for molecular delimitation of recently evolved species [17,19,55–57]. With the genomics resources presented in this study, the researchers can identify cryptic species within *Apopellia endiviifolia* using any primers targeting LSC or SSC region, without need ordering special primes for this purpose.

Application of superbarcodes in bryophytes is poorly explored. Recent studies apply plastid superbarcodes for species identification of genus *Calypogeia* [49] and *Conocephalum* [27] and prove their usefulness in effective distinguishing even problematic, allopolyploid taxa.

Genetic distinctiveness of water and typical forms was also confirmed by PTP analysis. The fundamental assumption of PTP analysis is that the number of substitutions between species is significantly higher than the number of substitutions within species. The results

obtained using this approach clearly support both forms as separated species with support varying from 0.509 to 0.522 for typical and water forms respectively. In comparison, species support in complex thalloid liverworts clade was often twice lower in the case of *Riccia* (0.188), *Dumortiera* (0.207), *Reboulia* (0.211) or *Conocephalum* (0.229) and similar to those found in the leafy liverworts like *Nowellia* (0.588), *Delavayell*a (0.570) or *Herbertus* (0.535).

Obtained results based on complete plastome sequences are congruent with previous studies. The identification of cryptic species within *Apopellia endiviifolia* beginning in the 1980s using isoenzyme markers [10]. The isoenzymes of peroxidases, esterases and mhd enable molecular delimitation of *A. endiviifolia* cryptic lineages. The presence of cryptic speciation was also confirmed using various genome scanning DNA markers, like RAPD and ISJ [11,12]. These studies, based on a genetic similarity approach, revealed values between 0.2 and 0.5, which characterize separate biological species [29,58]. Later studies using *rps*4, *rpl*16 and *trn*L-*trn*F [14] confirmed monophyly of these *Apopelia* lineages.

3.3. Ecological, Geographical and Morphological Differentiation

The low complexity of simple thalloid liverworts reduces the potential number of morphological diagnostic characters, which are essential for taxonomic description of novel species, including separation of cryptic species lineages. Thalli of typical form usually show characteristic dichotomically branched outgrows on the thallus apex, which are absent in water forms [59]. However, these characteristics outgrows can appear only periodically, and do not form in in vitro conditions [11]. Morphological studies on *A. endiviifolia* revealed significant differences between land and water forms in quantitative treats like length of basal and marginal cells, even after six months of in vitro culture [11], but those quantitative characters are not preferred as main diagnostics features. Described as *Apopellia megaspora* in the early 1980s, one of the North American *A. endiviifolia* cryptic species beside quantitative differences of spore diameters differs from the remaining *Apopellia* species by lacking thallus proliferation in autumn [60]. The better-studied cryptic species of another simple thalloid liverwort, *Aneura pinguis*, despite the presence of various lineage-specific DNA markers [4,26,61] and many biochemical compounds [62], still lacks significant morpho-anatomical differentiation. Moreover, the cryptic lineages *Apopellia endiviifolia* and some of *Aneura pinguis* differ in the type of preferred microhabitat, which could be defined not only by general description [10,14,61], but also by specific parameters, like pH and concentration of Ca, Mg, K, and Na [63]. The earlier studies suggest that water form (named species A) is known mainly from mountains, with very restricted occurrence in the lowlands, which correspond to preferred microhabitats distribution [59]. The sites with clear and fast running shallow water, which this lineage depends on, are more common in the mountains. The geographic distribution of the typical form (named species B) is concentrated in lowlands, with rare occurrences in mountains. Typical form is independent of running water and prefers denuded soil and rock detritus [59].

3.4. Apopellia Chloroplast Genomes Variation at Infraspecific Level

The comparative analysis of plastomes revealed very low variation at cryptic species level (Supplementary Table S2). Three sequenced individuals of water form differ only by two substitutions and three indels found in intergenic regions (Supplementary Table S2). Among four accessions of typical form 21 substitutions and eight indels were found, including 15 in CDSs (*rpo*A, *pet*A, *psb*J, *psb*B, *ndh*F, *ccs*A and *ycf*1). Most of them (10) were concentrated in the *ccs*A-*ndh*D region of SSC and had an impact on the length of both genes. Earlier studies on genus *Apopellia* revealed infraspecific variation in *rps*4, *rpl*16 and *trn*L-*trn*F plastid regions; however, both cryptic species (as genetic type EU and genetic type A1, for typical and water form respectively) were analyzed as one [14]. Considering water and typical forms as biological species [10] the potential of using plastome sequences for infraspecific studies like population genetics and phylogeography is very limited. In the studied dataset only the *ccs*A-*ndh*D region revealed variation big enough to identify every

specimen of the typical form; however, this region was identical in all the three individuals of water form (Supplementary Figure S1).

Despite recent advances in sequencing and assembly liverwort plastomes [22,23,26,49,64,65] the information about intraspecific variation of these molecules is still limited and only few species, excluding cryptic species of *Aneura pinguis* [26] and *Nowellia curvifolia* [28], have two or more complete chloroplast genome sequences available. However, based on the results of this and previous studies, the liverworts plastome variation at intraspecific level seem to be rather low [27,49,66] in comparison to vascular plants [19,67,68]. The specimens of model species *M. polymorpha* ssp. *ruderalis* from Korea and Japan differed by only four SNPs [66] and the *Riccia fluitans* from Korea and Poland differed by 59 SNPs [27]. The variation of plastomes seems to be inconsistent even at the generic level as in *Conocephalum*, where samples of *C. salebrosum* were differentiated by a single SNP, while in *C. conicum* number of detected SNPs was 32 in samples from Central Europe [27].

Since these four species are complex thalloid liverworts, which are considered a slowly evolving group [24], the infraspecific variation of simple thalloid and leafy was expected to be higher. However, the comparative analyses of European *Calypogeia* species revealed a low level of intraspecific variation, limited to few SNPs [49]. Analysis of the complete chloroplast genomes of leafy *Nowellia curvifolia* from Europe revealed only one SNP, but the European and North American specimens differed by 35 SNPs [28]. Considering currently available complete genome sequences of liverworts, low interspecific differentiation of complex thalloid plastomes does not reflect infraspecific variation, which seems to be comparable to simple thalloid and leafy liverworts plastomes.

3.5. Genes Undergoing Positive Selection

Applied tests revealed positive selection in nine genes (*ccs*A, *ndh*D, *ndh*F, *pet*A, *psb*B, *psb*C, *rpo*B, *ycf*1 and *ycf*2). For most of these genes only a single site under episodic diversifying selection was revealed, except for *rpo*B where two sites were identified. Among these genes, branch-site analysis (aBSREL) confirmed selection only in *ccs*A and *ycf*2. The former was the only gene under selection according to the gene-wide test (BUSTED). The positive selection at genus level is poorly explored in liverworts. The *ccs*A, *ndh*F, *pet*A and *ycf*1 genes revealed signatures of positive selection in *Aneura pinguis* [26]. More evidence of positive selection at genus level could be found in wider studied vascular plants. Both *ycf* genes were detected under positive selection in genus *Fagopyrum* [69], *Holcoglossum* [70] and *Salix* [71], while *ycf*1 in *Caragana* [51]. The *ccs*A gene, which was pointed out as undergoing positive selection by all three tests, plays an important role in cytochrome biogenesis [72]. Earlier studies revealed its potential role in adaptation to different habitats in the genus *Cardamine* [73].

3.6. Phylogeny

Phylogenetic position of Pellideae is congruent with previous studies based on plastid datasets [23,27]. Both forms of *A. endiviifolia* form monophyletic, well supported clade sister to *Pallavicinia longispina* (Figure 4, Supplementary Figure S2). This clade contains two, well supported (1.00 PP and 100% BS) subclades for each of the forms. Additionally, inside the clade ML analysis grouped typical forms specimens from Beskydy Mts and Germany were distinguished with a statistically supported clade (96% BS), which is congruent with the variation pattern revealed within this group (Supplementary Table S2). However, Bayesian Inference analysis resolved this clade as not statistically supported (0.77 PP). ML analysis also clustered together typical forms of *A. endiviifolia* from Tatra Mts and Pomeranian (Supplementary Figure S2) with moderate support (86% BS), but BI does not distinguish this clade (Figure 4).
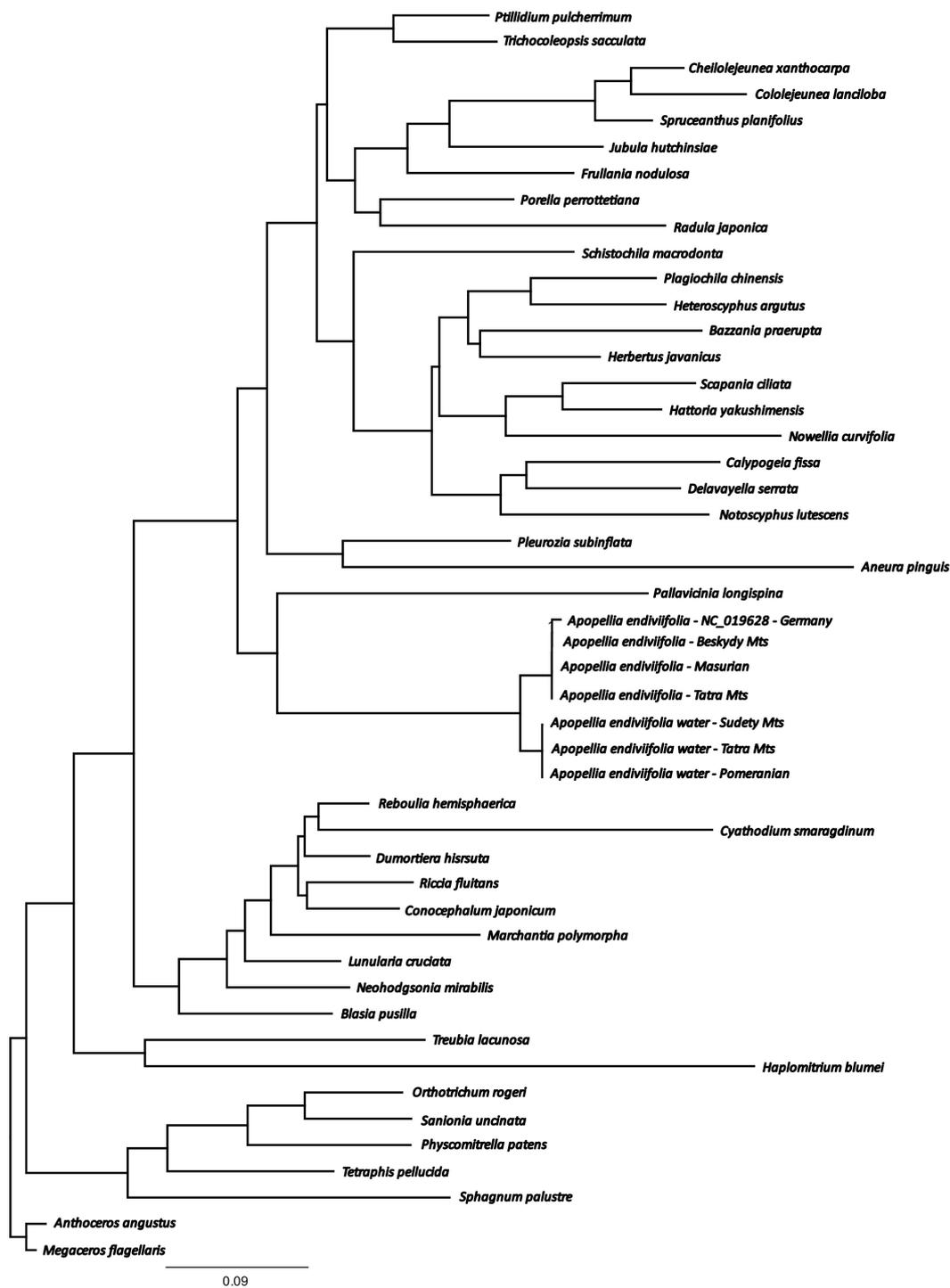
Figure 4. Phylogenetic relationships of liverworts based on protein-coding sequences obtained using Bayesian inference. All resolved clades have maximal posterior probability value.

4. Conclusions

Application of plastid superbarcodes for molecular delimitation of *Apopelia* cryptic species clearly separates two evolutionary lineages. The number of detected MDCs revealed genetic divergence at a level which in many cases exceeds values known from well-established taxonomic liverworts species. However, the base to comparison is rather low, since most complete plastome sequences of liverworts concern only single representatives of the species. Besides detected hotspots of variation, almost every 500 bp-long

frame enabled identification of water and terrestrial forms. Recently published studies significantly expand our knowledge of liverworts chloroplast genomes, but their infra-specific variation remained still unexplored and was limited to five genera. Two cryptic species of *Apopellia* are confirmed to have low intraspecific plastome variation, which was especially low in the water form. Performed analysis revealed positive selection in nine out of 87 protein coding genes, but all the three tests were congruent only in the case of *ccs*A, which seems to play an important role in adaptation to different habitats also in other species.

References

1. Therrien, J.P.; Crandall-Stotler, B.J.; Stotler, R.E. Morphological and Genetic Variation in *Porella platyphylla* and *P. Platyphylloidea* and Their Systematic Implications. *Bryologist* **1998**, *101*, 1–19. [CrossRef]
2. Heinrichs, J.; Hentschel, J.; Bombosch, A.; Fiebig, A.; Reise, J.; Edelmann, M.; Kreier, H.-P.; Schäfer-Verwimp, A.; Caspari, S.; Schmidt, A.R.; et al. One Species or at Least Eight? Delimitation and Distribution of *Frullania tamarisci* (L.) Dumort. s. l. (Jungermanniopsida, Porellales) Inferred from Nuclear and Chloroplast DNA Markers. *Mol. Phylogenet. Evol.* **2010**, *56*, 1105–1114. [CrossRef] [PubMed]
3. Szweykowski, J.; Buczkowska, K.; Odrzykoski, I.J. *Conocephalum salebrosum* (Marchantiopsida, Conocephalaceae)—A New Holarctic Liverwort Species. *Plant Syst. Evol.* **2005**, *253*, 133–158. [CrossRef]
4. Bączkiewicz, A.; Sawicki, J.; Buczkowska, K.; Polok, K.; Zieliński, R. Application of Different DNA Markers in Studies on Cryptic Species of *Aneura pinguis* (Jungermanniopsida, Metzgeriales). *Cryptogam. Bryol.* **2008**, *29*, 3–21.
5. Wachowiak, W.; Bączkiewicz, A.; Chudzińska, E.; Buczkowska, K. Cryptic Speciation in Liverworts—A Case Study in the *Aneura pinguis* Complex. *Bot. J. Linn. Soc.* **2007**, *155*, 273–282. [CrossRef]
6. Bakalin, V.A.; Vilnet, A.A.; Choi, S.S.; Nguyen, V.S. *Blepharostoma trichophyllum* SL (Marchantiophyta): The Complex of Sibling Species and Hybrids. *Plants* **2020**, *9*, 1423. [CrossRef]
7. Buczkowska, K.; Bakalin, V.; Bączkiewicz, A.; Aguero, B.; Gonera, P.; Ślipiko, M.; Szczecińska, M.; Sawicki, J. Does *Calypogeia azurea* (Calypogeiaceae, Marchantiophyta) Occur Outside Europe? Molecular and Morphological Evidence. *PLoS ONE* **2018**, *13*, e0204561. [CrossRef]
8. Lorbeer, G. Die Zytologie Der Lebermoose Mit Besonderer Berücksichtigung Allgemeiner Chromosomenfragen: Tl 1. *Ja. Bot.* **1934**, *80*, 567–818.
9. Krzakowa, M. Evolution and Speciation in Pellia, with Special Reference to the *Pellia megaspora-endiviifolia* Complex (Metzgeriales), IV. Isozyme Investigations. *J. Bryol.* **1981**, *11*, 447–450. [CrossRef]
10. Zieliński, R. Interpretation of Electrophoretic Patterns in Population Genetics of Bryophytes: VI. Genetic Variation and Evolution of the Liverwort Genus *Pellia* with Special Reference to Central European Territory. *Lindbergia* **1987**, *12*, 87–96.
11. Polok, K.; Sawicki, J.; Kubiak, K.; Szczecinska, M.; Korzekwa, K.; Szandar, K.; Zielinski, R. Evolutionary Divergence within *Pellia endiviifolia* [Dicks.] Dum. from Poland. *Ser. Biol. Uniw. Adama Mickiewicza Pozn.* **2005**, *72*, 241–252.

12. Pacak, A.; Fiedorow, P.; Dabert, J.; Szweykowska-Kulińska, Z. RAPD Technique for Taxonomic Studies of *Pellia epiphylla*-complex (Hepaticae, Metzgeriales). *Genetica* **1998**, *104*, 179–187. [CrossRef]

13. Szweykowska-Kulińska, Z.; Pacak, A.; Jankowiak, K. New DNA Markers for Discrimination between Closely-Related Species and for the Reconstruction of Historical Events; an Example Using Liverworts. *Cell. Mol. Biol. Lett.* **2002**, *7*, 403–416.

14. Schütz, N.; Quandt, D.; Nebel, M. The Position of the Genus Apopellia Stat. Nov. within the Pelliales (Marchantiophytina: Jungermanniopsida). *Taxon* **2016**, *65*, 221–234. [CrossRef]

15. Abdullah; Mehmood, F.; Rahim, A.; Heidari, P.; Ahmed, I.; Poczai, P. Comparative Plastome Analysis of *Blumea*, with Implications for Genome Evolution and Phylogeny of Asteroideae. *Ecol. Evol.* **2021**, *11*, 7810–7826. [CrossRef]

16. Huang, B.; Ruess, H.; Liang, Q.; Colleoni, C.; Spooner, D.M. Analyses of 202 Plastid Genomes Elucidate the Phylogeny of Solanum Section Petota. *Sci. Rep.* **2019**, *9*, 4454. [CrossRef]

17. Krawczyk, K.; Nobis, M.; Myszczyński, K.; Klichowska, E.; Sawicki, J. Plastid Super-Barcodes as a Tool for Species Discrimination in Feather Grasses (Poaceae: Stipa). *Sci. Rep.* **2018**, *8*, 1924. [CrossRef]

18. Zhang, W.; Sun, Y.; Liu, J.; Xu, C.; Zou, X.; Chen, X.; Liu, Y.; Wu, P.; Yang, X.; Zhou, S. DNA Barcoding of Oryza: Conventional, Specific, and Super Barcodes. *Plant Mol. Biol.* **2021**, *105*, 215–228. [CrossRef]

19. Szczecińska, M.; Sawicki, J. Genomic Resources of Three Pulsatilla Species Reveal Evolutionary Hotspots, Species-Specific Sites and Variable Plastid Structure in the Family Ranunculaceae. *Int. J. Mol. Sci.* **2015**, *16*, 22258–22279. [CrossRef]

20. Nevill, P.G.; Howell, K.A.; Cross, A.T.; Williams, A.V.; Zhong, X.; Tonti-Filippini, J.; Boykin, L.M.; Dixon, K.W.; Small, I. Plastome-Wide Rearrangements and Gene Losses in Carnivorous Droseraceae. *Genome Biol. Evol.* **2019**, *11*, 472–485. [CrossRef]

21. Forrest, L.L.; Wickett, N.J.; Cox, C.J.; Goffinet, B. Deep Sequencing of *Ptilidium* (Ptilidiaceae) Suggests Evolutionary Stasis in Liverwort Plastid Genome Structure. *Plant Ecol. Evol.* **2011**, *144*, 29–43. [CrossRef]

22. Myszczyński, K.; Górski, P.; Ślipiko, M.; Sawicki, J. Sequencing of Organellar Genomes of *Gymnomitrion concinnatum* (Jungermanniales) Revealed the First Exception in the Structure and Gene Order of Evolutionary Stable Liverworts Mitogenomes. *BMC Plant Biol.* **2018**, *18*, 321. [CrossRef]

23. Yu, Y.; Yang, J.-B.; Ma, W.-Z.; Pressel, S.; Liu, H.-M.; Wu, Y.-H.; Schneider, H. Chloroplast Phylogenomics of Liverworts: A Reappraisal of the Backbone Phylogeny of Liverworts with Emphasis on Ptilidiales. *Cladistics* **2020**, *36*, 184–193. [CrossRef]

24. Villarreal, A.J.C.; Crandall-Stotler, B.J.; Hart, M.L.; Long, D.G.; Forrest, L.L. Divergence Times and the Evolution of Morphological Complexity in an Early Land Plant Lineage (Marchantiopsida) with a Slow Molecular Rate. *New Phytol.* **2016**, *209*, 1734–1746. [CrossRef] [PubMed]

25. Grosche, C.; Funk, H.T.; Maier, U.G.; Zauner, S. The Chloroplast Genome of *Pellia endiviifolia*: Gene Content, RNA-Editing Pattern, and the Origin of Chloroplast Editing. *Genome Biol. Evol.* **2012**, *4*, 1349–1357. [CrossRef] [PubMed]

26. Myszczyński, K.; Bączkiewicz, A.; Buczkowska, K.; Ślipiko, M.; Szczecińska, M.; Sawicki, J. The Extraordinary Variation of the Organellar Genomes of the *Aneura pinguis* Revealed Advanced Cryptic Speciation of the Early Land Plants. *Sci. Rep.* **2017**, *7*, 9804. [CrossRef] [PubMed]

27. Sawicki, J.; Bączkiewicz, A.; Buczkowska, K.; Górski, P.; Krawczyk, K.; Mizia, P.; Myszczyński, K.; Ślipiko, M.; Szczecińska, M. The Increase of Simple Sequence Repeats during Diversification of Marchantiidae, An Early Land Plant Lineage, Leads to the First Known Expansion of Inverted Repeats in the Evolutionarily-Stable Structure of Liverwort Plastomes. *Genes* **2020**, *11*, 299. [CrossRef] [PubMed]

28. Sawicki, J.; Krawczyk, K.; Ślipiko, M.; Szczecińska, M. Sequencing of Organellar Genomes of *Nowellia curvifolia* (Cephaloziaceae Jungermanniales) Revealed the Smallest Plastome with Complete Gene Set and High Intraspecific Variation Suggesting Cryptic Speciation. *Diversity* **2021**, *13*, 81. [CrossRef]

29. Sawicki, J.; Szczecinska, M. A Comparison of PCR-Based Markers for the Molecular Identification of Sphagnum Species of the Section Acutifolia. *Acta Soc. Bot. Pol.* **2011**, *80*, 185–192. [CrossRef]

30. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef]

31. Dierckxsens, N.; Mardulyn, P.; Smits, G. NOVOPlasty: De Novo Assembly of Organelle Genomes from Whole Genome Data. *Nucleic Acids Res.* **2017**, *45*, e18. [CrossRef] [PubMed]

32. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq –Versatile and Accurate Annotation of Organelle Genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11. [CrossRef]

33. Greiner, S.; Lehwark, P.; Bock, R. OrganellarGenomeDRAW (OGDRAW) Version 1.3.1: Expanded Toolkit for the Graphical Visualization of Organellar Genomes. *Nucleic Acids Res.* **2019**, *47*, W59–W64. [CrossRef] [PubMed]

34. Brudno, M.; Poliakov, A.; Minovitsky, S.; Ratnere, I.; Dubchak, I. Multiple Whole Genome Alignments and Novel Biomedical Applications at the VISTA Portal. *Nucleic Acids Res.* **2007**, *35*, W669–W674. [CrossRef] [PubMed]

35. Brudno, M.; Do, C.B.; Cooper, G.M.; Kim, M.F.; Davydov, E.; Green, E.D.; Sidow, A.; Batzoglou, S. NISC Comparative Sequencing Program LAGAN and Multi-LAGAN: Efficient Tools for Large-Scale Multiple Alignment of Genomic DNA. *Genome Res.* **2003**, *13*, 721–731. [CrossRef] [PubMed]

36. Brown, S.D.; Collins, R.A.; Boyer, S.; Lefort, M.; Malubres-Olarte, J.; Vink, C.J.; Cruickshank, R.H. Spider: An R Package for the Analysis of Species Identity and Evolution, with Particular Reference to DNA Barcoding. *Mol. Ecol. Resour.* **2012**, *12*, 562–565. [CrossRef] [PubMed]

37. Bradbury, P.J.; Zhang, Z.; Kroon, D.E.; Casstevens, T.M.; Ramdoss, Y.; Buckler, E.S. TASSEL: Software for Association Mapping of Complex Traits in Diverse Samples. *Bioinformatics* **2007**, *23*, 2633–2635. [CrossRef]

38. Edgar, R.C. MUSCLE: A Multiple Sequence Alignment Method with Reduced Time and Space Complexity. *BMC Bioinform.* **2004**, *5*, 113. [CrossRef]

39. Jörger, K.M.; Schrödl, M. How to Describe a Cryptic Species? Practical Challenges of Molecular Taxonomy. *Front. Zool.* **2013**, *10*, 1–27. [CrossRef]

40. Merckelbach, L.M.; Borges, L.M. Make Every Species Count: Fastachar Software for Rapid Determination of Molecular Diagnostic Characters to Describe Species. *Mol. Ecol. Resour.* **2020**, *20*, 1761–1768. [CrossRef]

41. Kapli, P.; Lutteropp, S.; Zhang, J.; Kobert, K.; Pavlidis, P.; Stamatakis, A.; Flouri, T. Multi-Rate Poisson Tree Processes for Single-Locus Species Delimitation under Maximum Likelihood and Markov Chain Monte Carlo. *Bioinformatics* **2017**, *33*, 1630–1638. [CrossRef]

42. Stamatakis, A. RAxML Version 8: A Tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [CrossRef]

43. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Syst. Biol.* **2012**, *61*, 539–542. [CrossRef]

44. Murrell, B.; Wertheim, J.O.; Moola, S.; Weighill, T.; Scheffler, K.; Kosakovsky Pond, S.L. Detecting Individual Sites Subject to Episodic Diversifying Selection. *PLoS Genet.* **2012**, *8*, e1002764. [CrossRef]

45. Murrell, B.; Weaver, S.; Smith, M.D.; Wertheim, J.O.; Murrell, S.; Aylward, A.; Eren, K.; Pollner, T.; Martin, D.P.; Smith, D.M. Gene-Wide Identification of Episodic Selection. *Mol. Biol. Evol.* **2015**, *32*, 1365–1371. [CrossRef]

46. Smith, M.D.; Wertheim, J.O.; Weaver, S.; Murrell, B.; Scheffler, K.; Kosakovsky Pond, S.L. Less Is More: An Adaptive Branch-Site Random Effects Model for Efficient Detection of Episodic Diversifying Selection. *Mol. Biol. Evol.* **2015**, *32*, 1342–1353. [CrossRef]

47. Weaver, S.; Shank, S.D.; Spielman, S.J.; Li, M.; Muse, S.V.; Kosakovsky Pond, S.L. Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes. *Mol. Biol. Evol.* **2018**, *35*, 773–777. [CrossRef]

48. Wickett, N.J.; Zhang, Y.; Hansen, S.K.; Roper, J.M.; Kuehl, J.V.; Plock, S.A.; Wolf, P.G.; Depamphilis, C.W.; Boore, J.L.; Goffinet, B. Functional Gene Losses Occur with Minimal Size Reduction in the Plastid Genome of the Parasitic Liverwort Aneura Mirabilis. *Mol. Biol. Evol.* **2008**, *25*, 393–401. [CrossRef]

49. Ślipiko, M.; Myszczyński, K.; Buczkowska, K.; Bączkiewicz, A.; Szczecińska, M.; Sawicki, J. Molecular Delimitation of European Leafy Liverworts of the Genus Calypogeia Based on Plastid Super-Barcodes. *BMC Plant Biol.* **2020**, *20*, 243. [CrossRef]

50. Dong, W.; Xu, C.; Li, C.; Sun, J.; Zuo, Y.; Shi, S.; Cheng, T.; Guo, J.; Zhou, S. Ycf1, the Most Promising Plastid DNA Barcode of Land Plants. *Sci. Rep.* **2015**, *5*, 8348. [CrossRef]

51. Jiang, M.; Chen, H.; He, S.; Wang, L.; Chen, A.J.; Liu, C. Sequencing, Characterization, and Comparative Analyses of the Plastome of *Caragana rosea* car. *rosea*. *Int. J. Mol. Sci.* **2018**, *19*, 1419. [CrossRef] [PubMed]

52. DeSalle, R.; Egan, M.G.; Siddall, M. The Unholy Trinity: Taxonomy, Species Delimitation and DNA Barcoding. *Philos. Trans. R. Soc. B Biol. Sci.* **2005**, *360*, 1905–1916. [CrossRef] [PubMed]

53. Sarkar, I.N.; Planet, P.J.; Desalle, R. CAOS Software for Use in Character-based DNA Barcoding. *Mol. Ecol. Resour.* **2008**, *8*, 1256–1259. [CrossRef] [PubMed]

54. Kühn, A.L.; Haase, M. QUIDDICH: QUick IDentification of DIagnostic CHaracters. *J. Zool. Syst. Evol. Res.* **2020**, *58*, 22–26. [CrossRef]

55. Wu, L.; Wu, M.; Cui, N.; Xiang, L.; Li, Y.; Li, X.; Chen, S. Plant Super-Barcode: A Case Study on Genome-Based Identification for Closely Related Species of Fritillaria. *Chin. Med.* **2021**, *16*, 52. [CrossRef]

56. Liu, Z.-F.; Ma, H.; Ci, X.-Q.; Li, L.; Song, Y.; Liu, B.; Li, H.-W.; Wang, S.-L.; Qu, X.-J.; Hu, J.-L.; et al. Can Plastid Genome Sequencing Be Used for Species Identification in Lauraceae? *Bot. J. Linn. Soc.* **2021**, *197*, 1–14. [CrossRef]

57. Chen, Q.; Wu, X.; Zhang, D. Comparison of the Abilities of Universal, Super, and Specific DNA Barcodes to Discriminate among the Original Species of *Fritillariae cirrhosae bulbus* and Its Adulterants. *PLoS ONE* **2020**, *15*, e0229181. [CrossRef]

58. Szczecińska, M.; Sawicki, J.; Polok, K.; Hołdyński, C.; Zieliński, R. Comparison of Three Polygonatum Species from Poland Based on DNA Markers. *Ann. Bot. Fennici* **2006**, *43*, 379–388.

59. Szweykowski, J. *An Annotated Checklist of Polish Liverworts and Hornworts*; Polish Academy of Sciences, W Szafer Institute of Botany: Kraków, Poland, 2006; ISBN 83-89648-33-4.

60. Schuster, R.M. Evolution and Speciation in Pellia, with Special Reference to the *Pellia megaspora-endiviifolia* Complex (Metzgeriales), I. Taxonomy and Distribution. *J. Bryol.* **1981**, *11*, 411–431. [CrossRef]

61. Bączkiewicz, A.; Szczecińska, M.; Sawicki, J.; Stebel, A.; Buczkowska, K. DNA Barcoding, Ecology and Geography of the Cryptic Species of *Aneura pinguis* and Their Relationships with *Aneura maxima* and *Aneura mirabilis* (Metzgeriales, Marchantiophyta). *PLoS ONE* **2017**, *12*, e0188837. [CrossRef]

62. Wawrzyniak, R.; Wasiak, W.; Bączkiewicz, A.; Buczkowska, K. Volatile Compounds in Cryptic Species of the *Aneura pinguis* Complex and *Aneura maxima* (Marchantiophyta, Metzgeriidae). *Phytochemistry* **2014**, *105*, 115–122. [CrossRef]

63. Diatta, J.; Bączkiewicz, A.; Drapikowska, M.; Rodkiewicz, P.; Sawicki, J.; Szczecińska, M.; Buczkowska, K. Geochemical Alkalinity and Acidity as Preferential Site-Specific for Three Liverwort Lineages of *Aneura pinguis* Cryptic Species A. *Sci. Rep.* **2021**, *11*, 4334.

64. Choi, S.S.; Kwon, W.; Park, J. The Complete Chloroplast Genome of *Scapania ampliata* Steph., 1897 (Scapaniaceae, Jungermanniales). *Mitochondrial DNA Part B* **2020**, *5*, 2890–2892. [CrossRef]

65. Kwon, W.; Kim, Y.; Park, J. The Complete Chloroplast Genome Sequence of *Dumortiera hirsuta* (Sw.) Nees (Marchantiophyta, Dumortieraceae). *Mitochondrial DNA Part B* **2019**, *4*, 318–319. [CrossRef]

66. Kwon, W.; Kim, Y.; Park, J. The Complete Chloroplast Genome of Korean *Marchantia polymorpha* subsp. *ruderalis* Bischl. & Boisselier: Low Genetic Diversity between Korea and Japan. . *Mitochondrial DNA Part B* **2019**, *4*, 959–960.

67. Cho, M.-S.; Kim, J.H.; Yamada, T.; Maki, M.; Kim, S.-C. Plastome Characterization and Comparative Analyses of Wild Crabapples (*Malus baccata* and *M. Toringo*): Insights into Infraspecific Plastome Variation and Phylogenetic Relationships. *Tree Genet. Genomes* **2021**, *17*, 41. [CrossRef]

68. Park, S.; An, B.; Park, S. Reconfiguration of the Plastid Genome in Lamprocapnos Spectabilis: IR Boundary Shifting, Inversion, and Intraspecific Variation. *Sci. Rep.* **2018**, *8*, 13568. [CrossRef]

69. Wang, X.; Zhou, T.; Bai, G.; Zhao, Y. Complete Chloroplast Genome Sequence of *Fagopyrum dibotrys*: Genome Features, Comparative Analysis and Phylogenetic Relationships. *Sci. Rep.* **2018**, *8*, 12379. [CrossRef]

70. Li, Z.-H.; Ma, X.; Wang, D.-Y.; Li, Y.-X.; Wang, C.-W.; Jin, X.-H. Evolution of Plastid Genomes of *Holcoglossum* (Orchidaceae) with Recent Radiation. *BMC Evol. Biol.* **2019**, *19*, 63. [CrossRef]

71. Zhou, J.; Jiao, Z.; Guo, J.; Wang, B.S.; Zheng, J. Complete Chloroplast Genome Sequencing of Five Salix Species and Its Application in the Phylogeny and Taxonomy of the Genus. *Mitochondrial DNA Part B* **2021**, *6*, 2348–2352. [CrossRef]

72. Xie, Z.; Merchant, S. The Plastid-Encoded CcsA Gene Is Required for Heme Attachment to Chloroplast c-Type Cytochromes (∗). *J. Biol. Chem.* **1996**, *271*, 4632–4639. [CrossRef] [PubMed]

73. Hu, S.; Sablok, G.; Wang, B.; Qu, D.; Barbaro, E.; Viola, R.; Li, M.; Varotto, C. Plastome Organization and Evolution of Chloroplast Genes in Cardamine Species Adapted to Contrasting Habitats. *BMC Genom.* **2015**, *16*, 306. [CrossRef] [PubMed]