



Article

Super-Mitobarcoding in Plant Species Identification? It Can Work! The Case of Leafy Liverworts Belonging to the Genus *Calypogeia*

Monika Ślipiko ¹, Kamil Myszczyński ², Katarzyna Buczkowska ³, Alina Bączkiewicz ³ and Jakub Sawicki ^{1,*}

¹ Department of Botany and Nature Protection, Faculty of Biology and Biotechnology, University of Warmia and Mazury in Olsztyn, 10-719 Olsztyn, Poland

² Laboratory of Translational Oncology, Intercollegiate Faculty of Biotechnology, Medicinal University of Gdansk, 80-210 Gdansk, Poland

³ Department of Biology, Institute of Experimental Biology, Adam Mickiewicz University in Poznan, 61-712 Poznan, Poland

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

Abstract: Molecular identification of species is especially important where traditional taxonomic methods fail. The genus *Calypogeia* belongs to one of the tricky taxons. The simple morphology of these species and a tendency towards environmental plasticity make them complicated in identification. The finding of the universal single-locus DNA barcode in plants seems to be ‘the Holy Grail’; therefore, researchers are increasingly looking for multiloci DNA barcodes or super-barcoding. Since the mitochondrial genome has low sequence variation in plants, species delimitation is usually based on the chloroplast genome. Unexpectedly, our research shows that super-mitobarcoding can also work! However, our outcomes showed that a single method of molecular species delimitation should be avoided. Moreover, it is recommended to interpret the results of molecular species delimitation alongside other types of evidence, such as ecology, population genetics or comparative morphology. Here, we also presented genetic data supporting the view that *C. suecica* is not a homogeneous species.

Keywords: super-barcoding; species identification; liverworts; *Calypogeia*; mitogenome



Citation: Ślipiko, M.; Myszczyński, K.; Buczkowska, K.; Bączkiewicz, A.; Sawicki, J. Super-Mitobarcoding in Plant Species Identification? It Can Work! The Case of Leafy Liverworts Belonging to the Genus *Calypogeia*. *Int. J. Mol. Sci.* **2022**, *23*, 15570.

<https://doi.org/10.3390/ijms232415570>

Academic Editor: Richard R.-C. Wang

Received: 1 November 2022

Accepted: 5 December 2022

Published: 8 December 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

It is interesting if Paul Hebert and his team, publishing their paper in 2003 [1], thought, at least for a moment, that their research would be groundbreaking and would leave a strong mark on science. The article about the molecular identification of lepidopteran species has reached almost 9000 citations today, and Hebert himself has been called the father of DNA barcoding. The proposed method for species recognition, based on short DNA sequences specific to a species, seemed to be clever and simple. It is indeed so in animals, but in plants, it occurred to be complex, and as some scientists claim, it is like searching for ‘the Holy Grail’ [2]. Whereas in animals the fragment of the ca. 600 bp-long *cox1* gene delimitates species quite efficiently, in plants, this gene fails because of its low variability [3–6]. The *cox1* gene belongs to complex IV of the respiratory chain and encodes subunit 1 of the cytochrome c oxidase. This gene is settled in the mitochondrial genome, which in plants is called peculiar by some scientists [7–9]. What amazes them in plant mitogenomes? First of all, its size. In comparison to animals, whose mitogenome is ca. 16 kb in length [7,10], plants have a broad range of mitochondrial genome size, from the smallest (66 kb) in hemiparasitic *Viscum scurruloideum* [11] to colossal, over 11 Mb in *Silene conica* (11.3 Mb) [12] and *Larix sibirica* (11.7 MB) [13]. Quite extreme differentiation in mitogenome size was found also in fungi (1.1–272.2 kb) [14,15]. However, the range of length does not exceed the kb unit. The next striking feature of plant mitogenomes is their variable structure. It is startling that despite the general acceptance of the existence of variable spatial

organization of the plant mitogenomes [16–23], still, many publications depict them in ring form without mentioning that it is just a model [9]. Most plant mitogenomes are a combination of linear DNA and some smaller in circular and branched form [10]. This unusual structural plasticity the plant mitogenomes owe to a recombination between numerous repeats scattered throughout the genome [8,24–26]. The longer the repeats, the more frequent the recombination. Repeats over 1 kb recombine the most often [11,25], and they are the main cause of structural rearrangements of plant mitogenomes [27], whereas recombination between short repeats may lead to a scramble of the gene order, even among members of the same genus [24,28,29]. The number of genes may also be variable [9,10,30], but it is not the main cause of the large size of plant mitogenomes. The primary reason for this phenomenon is an expansion of their intergenic regions by accumulating sequences from nuclear and chloroplast regions and via horizontal transfer [27,31,32]. By all appearances, plant mitogenomes seem to be very plastic. However, it turns out that their gene-coding sequences evolve slowly, and their synonymous substitution rate is around 100 lower than in animal mitogenomes [28]. Comparing this factor to the nuclear and plastid genome, plant mitochondria are 10 and 3–4 times less variable, respectively [33–35]. Plant mitochondria are dissonant: on the one hand, they are very labile in their size, structure and gene content; on the other hand, they are stable in the structure of protein-coding sequences, which is probably the result of the DNA repair mechanism, particularly of homologous recombination [27]. Nevertheless, some exceptions happen, as in Geraniaceae, Plantaginaceae, *Silene* and *Ajuga*, where nucleotide protein-coding sequences evolve faster [36–38]. The complete opposite of angiosperm mitogenomes are mitochondrial genomes of the early land plants bryophytes. They are said to be conserved in their form (circular), gene content and order [39,40]. However, it has been shown that the order of genes may be disturbed in some species [41].

Genetic stability of plant mitochondria makes species diagnosis based on these sequences problematic. Therefore, scientists, in the search for DNA barcodes, began to concentrate their attention on the plastid genome, which is more variable [27,33]. The Consortium for the Barcode of Life (CBOL) recommended a two-locus combination of *matK*+ *rbcL* as a plant barcode [42]. It suggests that species discrimination in plants is more demanding than in animals, and just a single-locus barcode is not enough to accurately identify species. Such simple barcodes are especially prone to incorrect species identification when discriminated species can share haplotypes [43,44]. Taking into account frequent hybridization and introgression processes among plants [43,45–47], a lack of a single-locus barcode is not startling. On the other hand, the two-locus barcode (*matK*+ *rbcL*) recommended by CBOL was successful only in 75% [42]. Therefore, some scientists proposed to use multiple barcodes with application sequences such as *matK*, *rbcL*, and spacers plastid *trnH-psbA* and nuclear ITS [48,49]. There were also suggestions for telling plant species apart by other chloroplast sequences, e.g., *ycf1* [50,51], *ndhF* [51], *ndhB*, *ndhH*, *trnT-trnL* [52,53], *trnL-trnF* [54,55] and their combination. Unfortunately, there are cases where even multilocus barcoding does not work. In closely related species, some researchers have proposed the application of an entire chloroplast genome as a ‘super-barcode’ [11], and others have suggested so-called ‘ultra-barcoding’, based on both the whole plastome and nuclear ribosomal DNA [56], because organellar genomes are maternally inherited.

The *Calypogeia* genus, belonging to the family Calypogeiaceae, reaches about 90 species in the world and 10 in Europe, but all of them are tricky in their diagnostics. Morphologically, they are very similar, but the lack of well-developed keys for *Calypogeia* species is problematic. A good diagnostic feature seems to be the presence of oil bodies, but they occur only in fresh plant material. Difficulties also appear when the observed oil body characteristics cannot be matched with a key [57]. Moreover, *Calypogeia* species are characterized by plasticity. Their phenotype is shaped by the environment. It can lead to masking a valid distinction between species [58]. On the other hand, among *Calypogeia* species, some of them are hidden, so-called cryptic species. Although morphologically they

look exactly alike, in fact, they form genetically different groups of species. Cryptic species can be found in *C. fissa* [59], *C. sphagnicola* [60,61] and *C. suecica* [52].

In our study, we decided to take up the issue of species identification in the *Calypogeia* genus, one of the most problematic liverwort taxa according to many taxonomists. We undertook this some time ago on the basis of super-barcoding [52]. One issue was unresolved there: the *C. suecica* case. We suggested that it could be a cryptic species and therefore collapsed correct species identification. We decided to study it again in this research.

Having access to the data and already being experienced by our previous research on liverwort mitogenomes [41,62], we also assembled 26 mitochondrial genomes of 11 species of *Calypogeia* to check if there was anything extraordinary there. Surprisingly high variability of this chondriome pushed us to ask ourselves a bold question: is there any possibility to disclose plant species using super-mitobarcoding?

2. Results

2.1. Analysis of Organelle Genome Variation

The newly assembled organellar genomes of *Calypogeia* are circular and have typical gene order like other liverworts [52,62]. Plastome length ranged from 119,628 bp in *C. arguta* to 120,170 bp in *C. muelleriana*, whereas mitogenomes reached from 159,055 bp in *C. arguta* to 163,324 bp in *C. neesiana*. Genes of the chloroplast genome were located in four regions: a large single copy (LSC), a small single copy (SSC) and two inverted repeat regions (IRs). Although the plastome of *Calypogeia* species is shorter compared to mitogenome, it has more genes: 81 protein coding, 4 rRNAs, 31 tRNAs and 6 *ycf* genes of unknown function (total 122 genes) considering only one copy of IR region. Meanwhile, the mitochondrial genome holds 70 genes: 42 protein-coding genes, 25 tRNAs and 3 rRNAs.

Analysis of two organellar genomes in *Calypogeia* revealed greater variability of the plastome. The nucleotide diversity (π) for this organellar genome ($\pi = 0.04122$) was almost twice as high as that of the mitogenome ($\pi = 0.02279$). The difference in the variation of organellar genomes is excellently presented in Figure 1, where the blue and red lines indicated the mean π value for the whole mitochondrial and chloroplast genomes, respectively. The most variable 500 bp-long region in the mitogenome of *Calypogeia* lies within the pseudogene *nad7*, and the rest of the five most mutable sliding-window regions are nested within the following spacers: *rrn5-rrn18*, *nad1-cob*, *nad2-rps12* and *nad3-nad7* pseudogene (Figure 1). On the other hand, considering full-length noncoding regions, the most variable are the following spacers: *rps8-rpl6*, *nad4L-tatC*, *rps14-rps8*, *tatC-cox2* and *atp9-rps2* (π value, Supplemental Table S1). Among the coding sequences (CDSs), the most mutable belong to the small ribosomal subunit genes *rps8* and *rps1* and to the *rpl2* gene encoding the large ribosomal subunit, as well as to the ATP synthase gene *atp4* and the reverse transcriptase gene *rtl* (π value, Supplemental Table S1).

In the plastome, the five most changeable 500 bp-long regions also lie within the intergenic areas, in the large single copy (LSC) section. The most variable of them is that within the *psbA-ycf2* spacer. Interestingly, the peaks indicating the most mutable regions within genomes are much sharper in the mitogenome (Figure 1). The highest peak in the mitochondrial genome has a π value of 0.46791 (within the *nad7* pseudogene), whereas the highest peak π in the plastome is 0.22261 (within the *psbA-ycf2* spacer).

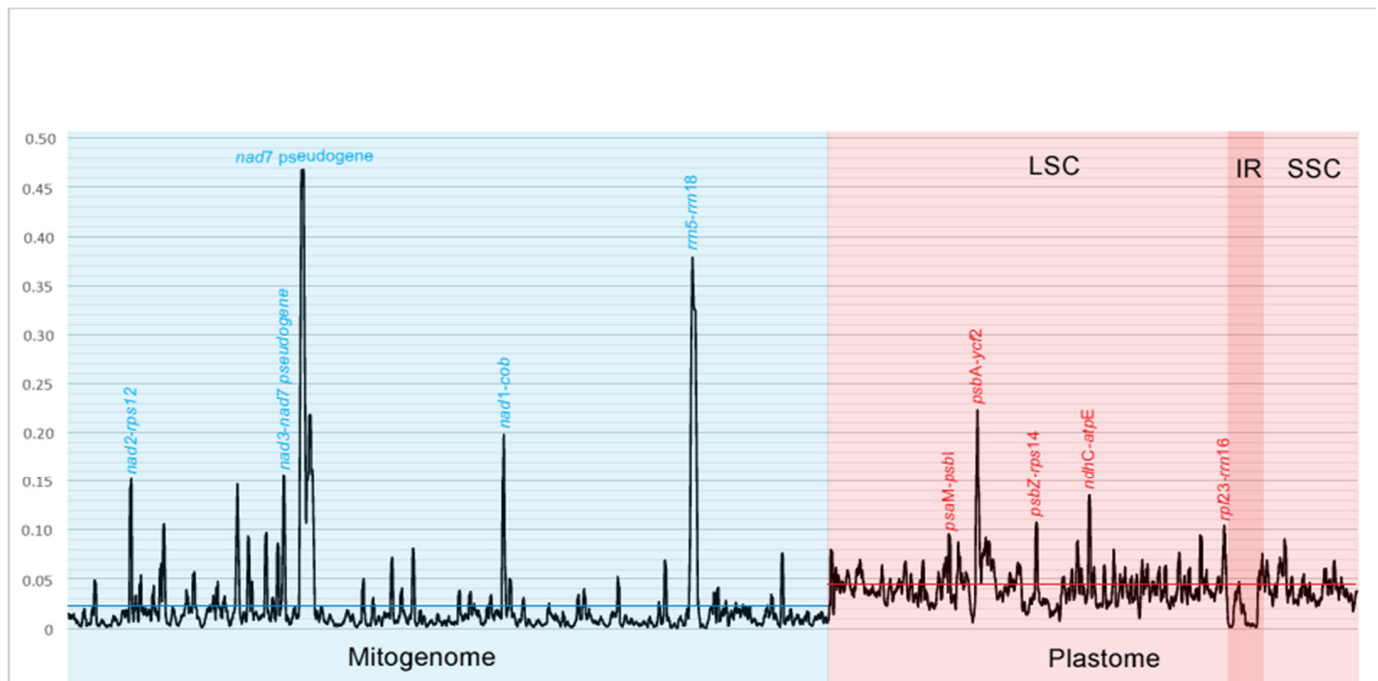


Figure 1. The five most variable regions in plastome and mitogenome of *Calypogeia* obtained by 500 bp-long sliding window. The chart shows the π values.

2.2. Phylogenetic Analysis

The phylogenetic trees based on both chloroplast and mitochondrial CDSs visibly separated all investigated species of liverworts (Figure 2). *Calypogeia muelleriana* and *C. integristipula* show the highest similarity in both organelle genome sequences among all studied species of the genus *Calypogeia*. As in previous studies [52,63], *Calypogeia arguta* was positioned as the most distant from the other *Calypogeia* species. Plastid-based phylogenetic data show relationships between *Calypogeia* species more precisely, which is supported by posterior probability values reaching one by almost all nodes. Only the relationship between *C. azurea* and *C. azorica* is supported weaker. The same doubtful phylogenetic affinity both between the two mentioned above species and between them and other *Calypogeia* species (*C. sphagnicola* and two groups of *C. suecica*) is indicated by the mitochondrial CDS-based tree. Nevertheless, all *Calypogeia* specimens were also correctly assigned to the species.

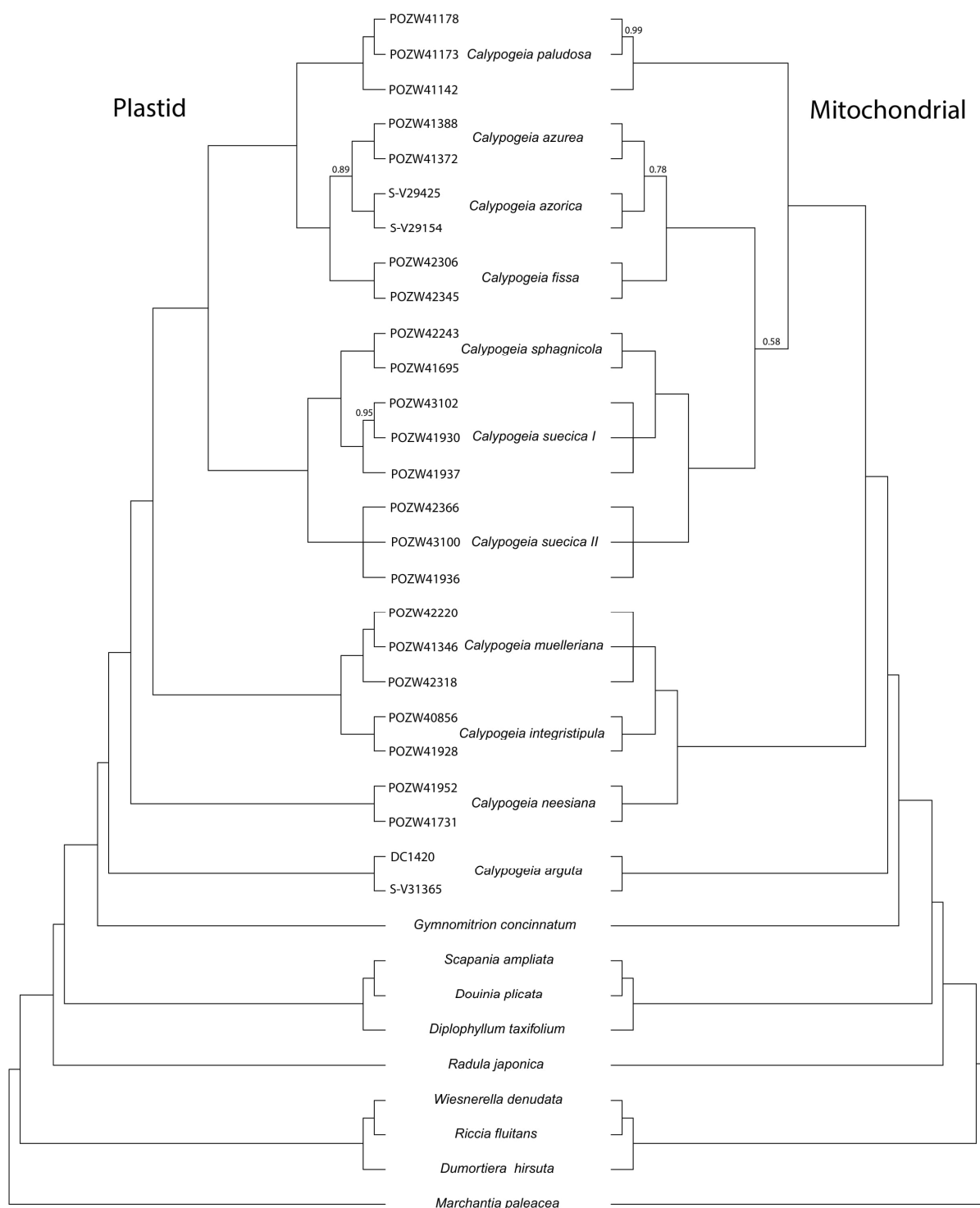


Figure 2. The phylogram based on plastid and mitochondrial CDSs derived from a Bayesian analysis. The posterior probability value lower than 1 is given at the node.

2.3. Species Delimitation

The results of tree-based analyses were not fully conclusive. The bPTP-ML analysis carried out on the chloroplast ML tree was too sensitive and recognized more species than were taken to the analysis (Table 1).

Table 1. Species delimited by bPTP-ML analysis based on chloroplast genomes.

Species Number	Species Name	Bayesian Support (BS)
1	<i>C. integristipula</i> POZW40856, <i>C. integristipula</i> POZW41928	0.435
2	<i>C. fissa</i> POZW42306, <i>C. fissa</i> POZW42345	0.365
3	<i>C. paludosa</i> POZW41142	0.655
4	<i>C. paludosa</i> POZW41173, <i>C. paludosa</i> POZW41178	0.330
5	<i>C. sphagnicola</i> POZW42243, <i>C. sphagnicola</i> POZW41695	0.417
6	<i>C. neesiana</i> POZW41952	0.409
7	<i>C. neesiana</i> POZW41731	0.409
8	<i>C. muelleriana</i> POZW42318	0.566
9	<i>C. suecica</i> I POZW41937	0.543
10	<i>C. suecica</i> I POZW41930	0.270
11	<i>C. suecica</i> I POZW43102	0.270
12	<i>C. azorica</i> S-V29425	0.362
13	<i>C. azorica</i> S-V29154	0.362
14	<i>C. azurea</i> POZW41388	0.440
15	<i>C. azurea</i> POZW41372	0.440
16	<i>C. muelleriana</i> POZW41346	0.286
17	<i>C. muelleriana</i> POZW42220	0.286
18	<i>C. suecica</i> II POZW42366	0.633
19	<i>C. suecica</i> II POZW43100, <i>C. suecica</i> II POZW41936	0.321
20	<i>C. arguta</i> DC1420	0.402
21	<i>C. arguta</i> S-V31365	0.402

The GMYC model applied for the mitochondrial coding sequences extended to other liverwort species was more precise. The tree showed only two specimens of *Calypogeia azorica* as two separated species and merged *C. muelleriana* and *C. integristipula* into one species (Figure 3).

Quite similar results were revealed by ASAP (distance-based method) for mitogenome data (Figure 4A). The lowest asap-score at the level of 4.5 had three partitions. One of them showed more species than was expected, because it treated specimens of *C. azorica* as two species. On the other hand, another partition was less sensitive and pointed out two species fusions. The third partition merged only *C. muelleriana* with *C. integristipula* together, but the dendrogram supporting the bar chart indicated that this fusion is dubious. The orange dot on the node between specimens of *C. muelleriana* and *C. integristipula* suggests that such a merger is rather unlikely, and these individuals correspond to two different species. The correctness of the above division is supported by the ASAP analysis based on chloroplast genomes (Figure 4B). The node between *C. muelleriana* and *C. integristipula* is marked with a black dot, which denotes an even lower probability than for mitogenomic data, that these specimens belong to one group. Two partitions with the lowest asap-score values (1.5 and 3.0) generated for chloroplast genomes perfectly distinguished 11 species of *Calypogeia*.

The species delimitation analysis, based on genetic distance conducted in Spider, revealed the existence of a barcoding gap for all 11 tested species of *Calypogeia* (Supplemental Figure S1). All results were statistically significant. The largest barcoding gaps both for plastome and mitogenome were identified for *C. arguta* and for *C. neesiana*. For these species, the number of molecular diagnostic characters (MDCs) was also the largest: 9720 for *C. arguta* and 894 for *C. neesiana* (Supplemental Table S3).

A screening of the mitochondrial genome fragment by fragment with a length of 500 bp revealed some hotspots in genetic variation and indicated nucleotide sequences laying claim to DNA barcodes. It followed that many such places in the mitogenome can be found, which illustrate many peaks on the Spider charts (Figure 5A,E).

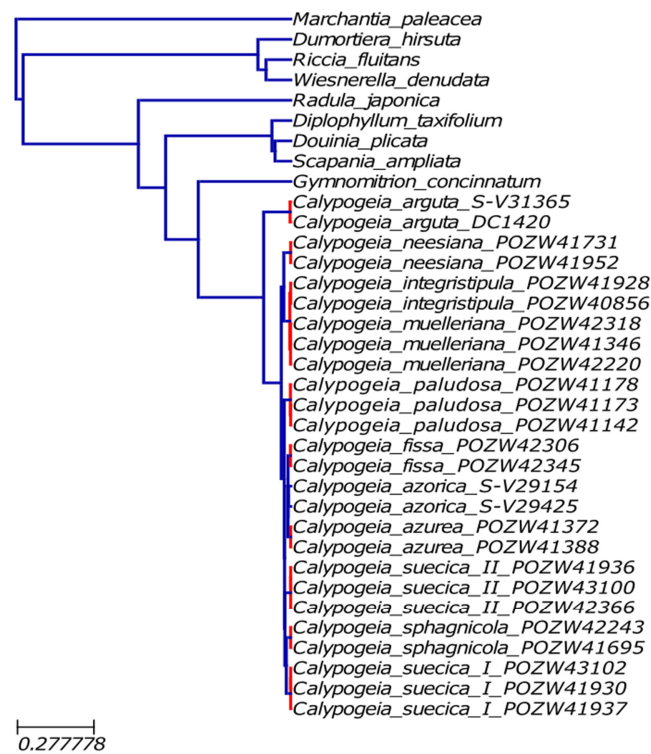
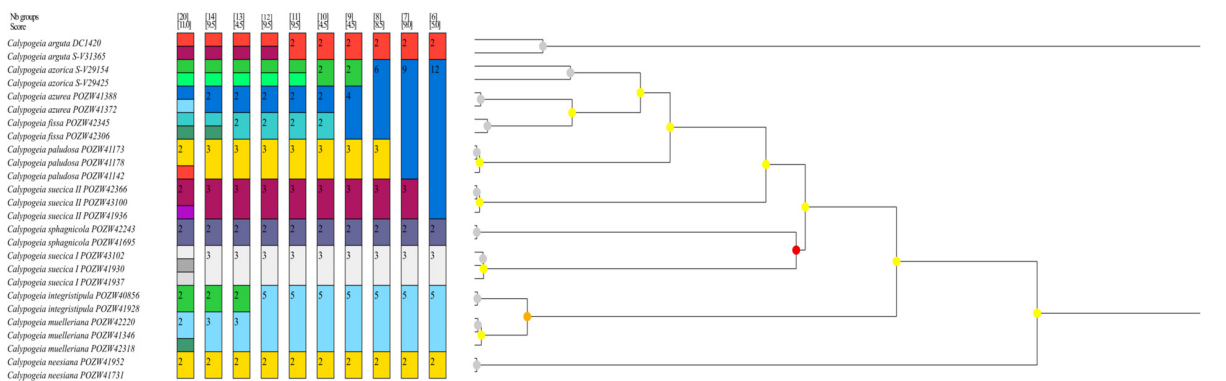


Figure 3. Species delimitation by GMYC analysis. Red lines indicate the same species group.

A. mitochondrial genome



B. chloroplast genome

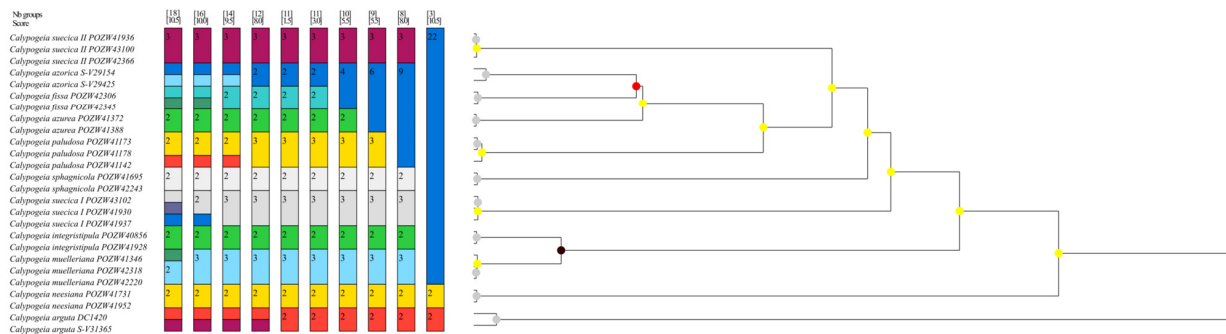


Figure 4. Species delimitation by ASAP analysis for mitochondrial (A) and chloroplast (B) genomes. The colorful fields indicate the groups of species. Every field contains the number of individuals. Above the colorful bars, the coefficient asap-score (the lower value) and number of species (the upper value) recognized for whole dataset are presented. The cladogram supports the species delimitation presented on the bar charts.

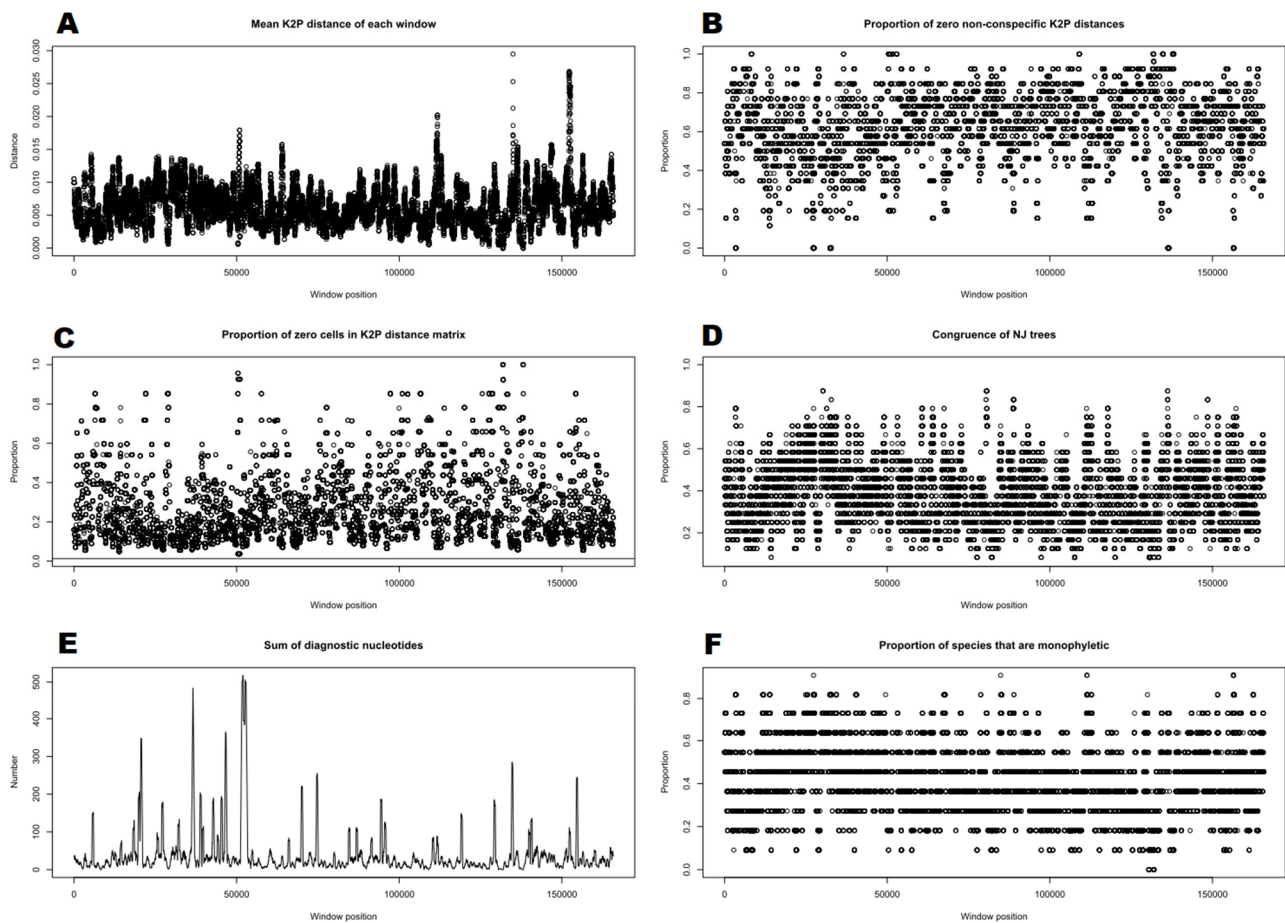


Figure 5. Results of several analyses across the mitochondrial sequences of *Calypogeia* sp. using the sliding-window method. (A) The plot of the mean Kimura 2-parameter distance matrix for each 500 bp-long window. (B) The proportion of zero non-conspecific distances. (C) The proportion of zero cells in the distance matrix. The unbroken horizontal line crossing the *y*-axis at 0 is the proportion of zero cells in the distance matrix created from the full dataset. (D) The proportion of clades that are identical between the windows and the full dataset. (E) The sum of diagnostic nucleotide positions for all species. (F) The proportion of species that are monophyletic.

3. Discussion

3.1. Variation of Organelle Genomes

At present, whole-genome sequencing has become one of the fundamental ways of obtaining scientific information. Although more and more mitochondrial genomes are sequenced among plants, still, little is directly written about their genetic variation at the species level [12]. Many studies rather touch on the issue of structural plasticity between closely related species [9,64,65]. Research taking up the subject of intraspecies variation is scarce [12], which raises the problem with an assessment of the degree of variability obtained in one's own studies. Most papers generally state that a low mutation rate is typical for plant mitogenomes [27], especially for protein-coding regions [65]. It is even more difficult to collect information about the genetic variation of mitogenomes in bryophytes. Most studies report that mitochondrial genomes of bryophytes are invariant, but it refers to their stability at composition and architecture levels [39,40,66]. Among liverworts, the level of genetic variation has been determined only for *Calypogeia* plastomes [52] and for organelle genomes of cryptic species of the *Aneura pinguis* complex [67]. Plastid genomes of bryophytes seem to evolve faster and be slightly more variable than mitogenomes, but they require verification in the future. Currently, it is confirmed only by one study regarding *Aneura pinguis* [67]. However, there are reports stating that the substitution rate of the

plastome was three times higher than that of mitogenome in angiosperms and twice as high as the substitution rate of mitogenome in gymnosperms [65]. The greater nucleotide diversity in various regions of plant mitogenome was noticed by van de Paer et al. in their study on the olive family [68]. Our results are in line with the above observations, because the chloroplast genome ($\pi = 0.04122$) of the genus *Calypogeia* is twice as variable as the mitogenome ($\pi = 0.02279$). Similar disparity values of organelle genome variation (based on π value) were observed in the Oleaceae family. However, it is worth noting that the value of disparateness between mitogenome and plastome in *Calypogeia* is at the species level and in olives at the subspecies (disparity of size 1.5–2 times) and subtribe level (disparity of size 3–5 times) [68]. This may also support the hypothesis that bryophytes evolve more slowly than angiosperms [69], hence a similar level of π -value discrepancy between organelle genomes is shaped at different taxonomic levels as a potential for evolution.

A huge disparity in the nucleotide diversity (π) between organelle genomes was found in the cryptic species complex of *Aneura pinguis* [67]. The nucleotide diversity for the mitochondrial genome in this species was 0.00233 and was almost 22 times lower than in the plastome. In comparison, most π values in the mitochondrial genome for some subgroups of *Oryza sativa* [70] were lower than 0.005, but in the complex of *Olea europaea* (olive) were about 0.031 and about 0.173 in the subtribe Oleinae [68]. So, the π value of 0.02279 for the mitogenome of the genus *Calypogeia* seems to be pretty high. This huge variability of the mitochondrial genome of *Calypogeia* is also noticeable in the number of SNPs (4781) and indels (1068) (Supplemental Table S1). The mitogenome of *Aneura pinguis* counted 953 SNPs and 1940 indels [67], whereas in genus *Orthotrichum* s.l., 427 SNPs and 6 indels were found [71]. The species *Scapania ampliata*, belonging like the *Calypogeia* species to leafy liverworts, was detected to have 823 SNPs and 2242 indels in the mitochondrial genome [72]. Meanwhile, species belonging to the complex thalloids showed a much lower level of genetic variation in this organelle genome. There were only 7 SNPs and no indels in *Marchantia polymorpha* subsp. *ruderalis* [73], 12 SNPs and 24 indels in *Dumortiera hirsuta* [74], 18 SNPs and 19 indels in *Riccia fluitans* [75], and 14 SNPs and 7 indels in *Monosolenium tenerum* [40]. Our findings for genus *Calypogeia* and results of other researchers for complex thalloids are in line to some extent with reports that Marchantiopsida (complex thalloid liverworts) show a lower evolution rate than Jungermanniopsida (leafy liverworts) [69,76].

It is not a mystery that noncoding regions in plants evolve rapidly, whereas protein-coding regions evolve leisurely [65,68]. It means that a mutation within the non-coding sequence is more likely. Mitochondrial non-coding sequences of *Calypogeia* have over 6 times more SNPs and almost 50 times more indels than the coding sequences. The average π value for protein-coding regions here was 1.5 times lower than for noncoding regions. In our previous study regarding the variation of the *Calypogeia* plastid genome, the π value for coding and noncoding sequences differed twice with an advantage for the noncoding sequences [52]. The sliding-window analysis indicated exactly the most variable regions in the noncoding sequences (Figure 1). The highest peak of variation in the mitogenome of *Calypogeia* was for the *nad7* pseudogene. A nonfunctional form of a gene can behave like a noncoding sequence and gather more mutations. The other five most variable mitochondrial regions were located within spacers. The same was true of the plastid genome of *Calypogeia* (Figure 1). The most variable region was the spacer *psbA-ycf2*, as we have already reported in our previous study [52].

The general π value obtained in our former research for the *Calypogeia* plastome ($\pi = 0.035076$) [52] was slightly lower than in the current study ($\pi = 0.04122$), but it was assuredly due to the extension of the number of specimens analyzed in the current study. The variation of the chloroplast genome of the genus *Calypogeia* is outstanding, but the variation of the seemingly stable liverwort mitogenome is striking.

3.2. Phylogenetic Relationships

The phylogenetic tree presents *C. arguta* as the most distant from other *Calypogeia* species. This distinctiveness of *C. arguta* was noticed by Bakalin et al., who suggested lumping *C. arguta* and related species into the new genus *Asperifolia* [77]. *Calypogeia muelleriana* and *C. integristipula* show the highest similarity in the mitochondrial genome sequence among all studied species of the genus *Calypogeia* (Figure 2). A close relationship between these species has also been revealed by studies based on the selected cpDNA loci [63] and the whole chloroplast genomes [52]. The latter is also visible in the current research (Figure 2). *Calypogeia muelleriana* and *C. integristipula* are considered distinct species, and they have diagnostic morphological features that allow them to be distinguished [78–80]. The species also differ in terms of isozyme and molecular markers [81,82]. Thus, the high similarity in the sequence of mitogenomes revealed in these studies proves the common origin of these two species. These species most likely inherited the organelle genomes from a common ancestor.

In our previous research on species identification based on whole plastomes, we suggested that two specimens of *C. suecica* could belong to two different cryptic species [52]. This study confirmed this hypothesis. We have now divided the species of *C. suecica* into two separate species labeled as Group 1 and 2 and extended the size of each group to three individuals. Both plastid and mitochondrial genomic data indicated these two species groups as genetically distinct (Figure 2). Surprisingly, *C. suecica* Group 1 is more closely related to *C. sphagnicola* than to *C. suecica* Group 2. Our morphological observations point to some differences between the two groups of *C. suecica* [83].

3.3. Species Delimitation

The organelle genomes of *Calypogeia* among the rather “evolutionarily lazy” genomes of liverworts were revealed to be unusually variable, which makes them a great foundation for the genetic identification of species. While the DNA barcodes for plants are rather sought among chloroplast sequences [42], the use of mitochondrial sequences for species identification is quite unconventional because of a generally reported lower nucleotide variation of plant mitogenomes [33–35]. As determined by CBOL, plastid single-locus DNA barcodes do not cope with plant species delimitation [42]. What is more, even two- or multi-locus DNA barcoding fails when plant species are closely related and the sequence variation is insufficient [42,84,85] or taxa are just taxonomically difficult to separate [46,47,86,87]. If in such situations scientists recommend using complete chloroplast genomes as super-barcodes to identify plant species [12], then applying a similar method to much less variable mitogenomes seems reasonable.

As far as we know, it is the first application of mitogenomes as super-barcodes in plants. We called this method ‘super-mitobarcoding’. Thus far, mitochondrial sequences in plants, in contrast to animals [1], have been useless in identifying species because of their nucleotide stability. The *Calypogeia* case with a high variability level of mitogenomic sequences provides a strong foundation for the idea of super-mitobarcoding. We tested the effectiveness of mitogenomes in *Calypogeia* species using different methods. This approach was recommended by some researchers, especially when closely related species are to be distinguished [88]. Some of the applied methods were not completely successful. The best results were obtained in the Spider analysis based on barcoding gaps. All *Calypogeia* species had a statistically significant barcoding gap and were correctly separated (Supplemental Figure S1), no matter if it was a mitochondrial or plastid genome. Positive separation results with this method were also obtained in our previous study regarding species delimitation of *Calypogeia* based on classic super-barcoding [52]. The graphs generated in the Spider analysis showed the ‘dynamics’ of the *Calypogeia* mitogenome, where many peaks and dots rise and fall but are densely spaced (Figure 5). For comparison, the plastome was more stable, as our analyses from the year 2020 showed [52]. In the mitochondrial genome can be found nucleotide positions, where the mean K2P distances among species are the greatest (Figure 5A), and the diagnostic nucleotides accumulated in

the most number (Figure 5E) or clades are identical (Figure 5D). In the *Calypogeia* plastome, genetic distances between species were at a higher level (most under 0.1) [52] than in mitogenomes (most under 0.02; Figure 5A), but the number of diagnostic nucleotides in various positions of the organelle genome was definitely greater in the mitogenome. On the one hand, this may be due to a more evolutionarily stable mitochondrial genome [65,67]; on the other hand, it may be due to the ‘dynamics’ of the *Calypogeia* mitogenome and high π value of some mitogenomic sites (Figure 1).

It is not surprising that most key positions for species delimitation are within the most variable mitochondrial regions. The greatest genetic distances between species are placed within 500 bp-long fragments of the *nad7* pseudogene (position ca. 50,000; Figures 5A and 1) and the spacers *nad1-cob* (position ca. 90,000–95,000; Figures 5A and 1), *rrn5-rrn18* (position ca. 132,000–134,000; Figures 5A and 1), but also within the most variable protein-coding genes such as *rps1* (position ca. 65,000; Supplemental Table S1) and *rpl2* (position ca. 55,000; Supplemental Table S1). Nevertheless, some regions such as *atp1-cox1* (position ca. 110,000; Figure 5A) or *ccmB-ccmC* (position ca. 153,000; Figure 5A), which are not among the most mutable, generate quite a high peak in the genetic distance chart. On the other hand, some genetic distances between species, generated by some of the top variable fragments such as spacers *nad2-rps12* (position ca. 12,000–20,000; Figure 5A) or the *nad3-nad7* pseudogene (position ca. 40,000–48,000; Figure 5A) are lost among others.

Similar observations can be made for the number of diagnostic nucleotides (Figure 5E). Most of them are located within the top variable regions: spacers (*nad2-rps12*, *nad3-nad7* pseudogene, *rrn5-rrn18*; Figure 1), the pseudogene *nad7* (Figure 1) and the gene *rpl2* (Supplemental Table S1). On the other hand, quite a lot is located within the spacers *sdh3-nad3* (position ca. 35,000; Figure 5E) and *ccmFC-rpl10* (position ca. 160,000; Figure 5E), which were ranked only 15th and 25th, respectively, among the most variable noncoding sequences (Supplemental Table S1). The above observations are congruent with our earlier report showing that sometimes good DNA barcodes are not among the top mutable regions [52]. What is more, a species identification based on the whole chloroplast genome can help with finding in *Calypogeia* single-locus DNA barcodes, which can distinguish species pretty well (e.g., *ndhB*, *ndhH*, *trnT-trnL*) [52]. Nevertheless, in the present research, single-locus DNA barcodes were not tested.

The first group of methods for species delimitation we used was based on the application of genetic distances. The methods using a comparison of intra- and interspecific genetic distances are debatable, because they do not appeal to the clear species concept [89]. However, they are widely applied [52,90,91] according to the first suggestion of the barcoding forerunner Hebert [1]. This type of method is also proven to have pretty good results [52,90,92,93]. The assemble species by automatic partitioning (ASAP), based partly on the barcoding gap, gives a possibility of a choice, of which partition is most relevant, prompting the correct answer by the asap-score. Both the Spider analysis (Supplemental Figure S1) and 1 of 10 ASAP partitions perfectly separated all *Calypogeia* species (Figure 4). Nevertheless, the solution of the ASAP for the plastome was more supported by the lower coefficient asap-score and by the ultrametric clustering tree.

The second group of methods was based on a phylogenetic approach, where all individuals belonging to the same species have a common evolutionary history [94]. Although GMYC and PTP were initially developed for single loci, they are increasingly used for multilocus datasets [95]. GMYC follows a speciation model [96], based on the mitochondrial genome, and almost showed accurate species division of *Calypogeia* (Figure 3). However, the GMYC tree presented both false-negative (*C. muelleriana* and *C. integristipula* as one species) and false-positive (*C. azorica* as two species) results. As discussed above, *C. muelleriana* and *C. integristipula* probably have organelles inherited from a common ancestor. Hence the genetic similarity. However, the oversplit of *C. azorica* into two distinct species is not entirely clear. Because the mitogenome sequences of two representatives of *C. azorica* are very similar to each other (99.9% pairwise identity), which is also visible in the phylogram (Figure 2), it must be a result of GMYC shortcomings. Luo et al. [95] reported

that GMYC is vulnerable to the number of loci and individuals, giving false positives. The mitogenomes do not suffer from the number of loci, but in the current research, *C. azorica* was represented only by two specimens (Supplemental Table S2). The second deficiency of GMYC is sensitivity to the presence of gene flow [95]. In our dataset, it could have happened between *C. muelleriana* and *C. integristipula*. Hence, the false-negative result for these two species. Considering that some researchers [96,97] reported general oversplitting of GMYC, the present results for *Calypogeia* species are quite passable.

The second approach used in the group of phylogenetic-based methods was PTP, but it was applied for plastid genomic data in accordance with a reference [98]. It performed the bPTP-ML version, because it is recommended when interspecies distances are predicted to be small [98], as it is at the low taxonomic level or among closely related species, like in the *Calypogeia* case. bPTP-ML was not very helpful in the present study, because it was too sensitive and indicated too many species (Table 1). Similar to GMYC, the sample size and gene flow can affect the power of the bPTP-ML resolution [95]. Moreover, shallow speciation events may increase the number of false negatives. This could explain the merging of *C. muelleriana* and *C. integristipula* into one species.

The issue of cryptic speciation of *C. suecica*, which we already noted in our previous research from 2020, was resolved. The current analyses, both based on mitogenomic and chloroplast data, definitely divided all specimens of *C. suecica* into two distinct groups: two cryptic species. The present results confirmed the usefulness of every type of DNA barcoding in discovering new species and detecting cryptic speciation [1,49,89]. However, our outcomes supported the scientific suggestion that a single method of molecular species delimitation should be avoided [94], because each method has its own shortcomings, which can be compensated for by the strength of another one. Moreover, it is recommended to interpret the results of molecular species delimitation alongside other types of evidence, such as ecology, population genetics or comparative morphology [99].

4. Materials and Methods

4.1. Plant Material

Calypogeia species used in this study came from the following herbaria: Herbarium of Adam Mickiewicz University (POZW), Herbarium D.A. Callaghan and Herbarium Schäfer-Verwimp (Supplemental Table S2). The 26 specimens represent 11 taxa of *Calypogeia*: *C. integristipula*, *C. fissa*, *C. sphagnicola*, *C. paludosa*, *C. muelleriana*, *C. azurea*, *C. arguta*, *C. azorica* and *C. neesiana*. *Calypogeia suecica* was divided into two groups marked with numbers 1 and 2 (Group 1, Group 2) as a result of our previous study, where *C. suecica* turned out to be genetically heterogeneous [52]. The total genomic DNA from specimens of each European *Calypogeia* species was extracted using the ZR Plant/Seed DNA MiniPrep™ kit (Zymo Research, Irvine, CA, USA). The species *C. paludosa* and *C. muelleriana* and two groups of *C. suecica* were represented by three specimens, and the rest of the species were represented by two. DNA quantity was estimated using a Qubit fluorometer and the Qubit™ dsDNA BR Assay Kit (Invitrogen, Carlsbad, NM, USA).

4.2. Organellar Genome Assembly

The genomic library was constructed with the TruSeq Nano DNA kit (Illumina, San Diego, CA, USA) and was sequenced using HiSeqX (Illumina) to generate 150 bp paired-end reads at Macrogen Inc. (Seoul, Korea) with a 350 bp insert size between paired ends. In view of the low amount of available plant material of *C. azorica*, the genomic library had to be constructed using an alternative kit (Qiagen FX library kit), which allows the use of a lower concentration of input DNA. After sequencing, reads were cleaned by removing the adaptor sequences and low-quality reads with Trimmomatic v0.36 [100]. The filtered reads were de novo assembled using the NOVOPlasty 4.3.1 tool [101] and previously published organellar genomes of *Calypogeia* [52,62] as references. Circularized genomes were verified and annotated using Geneious R8 software [102]. Annotated sequences of

Calypogeia chloroplast and mitochondrial genomes were submitted to GenBank with the accession numbers specified in Supplemental Table S2.

4.3. Analysis of Organelle Genome Variation

Twenty-six mitochondrial and the same number of plastid genomes of *Calypogeia* species were aligned using the MAFFT plugin [103] within Geneious R8. The second copy of IR was excluded from analyses. The genetic variation and barcoding analyses were conducted using the Spider package [104]. Nucleotide diversity (π) was calculated for each 500 bp-long fragment generated by sliding-window analysis (fragments were generated every 50 bp). The π value, number of SNPs and indels were also calculated for each coding and noncoding region using a custom Python script.

4.4. Phylogenetic Analyses

Phylogenetic analyses were performed using mitogenomes and plastomes of 35 species: 26 of the *Calypogeia* genus and 9 other known liverworts. MAFFT software [103] was used to create an alignment of 35 CDS of 42 genes and MrBayes 3.2.6 [105] software was applied for Bayesian analysis. The MCMC algorithm was run for 5,000,000 generations (sampling every 500) with four incrementally heated chains (starting from random trees). Tracer 1.7.1 software [106] was used to determine the number of generations needed to reach stationarity, which occurred at approximately 300,000 generations. Therefore, the first 600 trees were discarded as burn-in, and the remaining trees were used to develop a Bayesian consensus tree. The same methodical approach was applied to plastomes of the 35 above-mentioned liverwort species. In our latest research, we discovered genetic heterogeneity within *C. suecica* [52]. Species delimitation analyses suggested that two specimens of *C. suecica* belonged to two different species. In the current research, *C. suecica* was treated as a species complex and divided into two groups. We added two more specimens to each group in comparison to the previous study. *Marchantia paleacea* was used as an outgroup in both obtained phylogenetic trees.

4.5. Species Delimitation

Species delimitation was performed using two different approaches: (1) based on the phylogenetic tree and (2) based on a genetic distance. In the first approach, an input is a tree, and the method assumes that the phylogenetic tree accurately reflects a species differentiation [95]. The generalized mixed Yule-coalescent model (GMYC) [96] was applied to the mitochondrial dataset to delineate species. As an input, the Bayesian tree (obtained earlier in phylogenetic analysis) was used, but previously, the tree was ultraparameterized according to GMYC requirements. Ultrametrization of the tree was performed in the PATHd8 program [107], whereas species identification based on chloroplast genome data was carried out with the use of Poisson tree processes with Bayesian implementation (bPTP) [98]. In comparison to GMYC analysis, bPTP is run on a phylogenetic tree without ultrametrization. In such a case, the length of branches represents the number of substitutions, while an ultrametric tree in the GMYC model presents the stable rate of mutations across all lineages and the vertical branches mirror the time which elapsed between the tree root and the internal nodes. bPTP analysis was carried out on the ML phylogenetic tree (bPTP-ML) obtained using RAxML [108], but in this case, the tree was based on whole plastomes and calculated only for the *Calypogeia* species. bPTP analysis was made with the following settings: unrooted tree; MCMC generations: 500,000; thinning: 100; seed: 123. The bPTP analysis was performed on the server, but GMYC was called directly from the Linux command line (<https://github.com/zhangjiajie/pGMYC>; accessed on 11 September 2021).

Species delimitation analyses based on genetic distance were carried out in Spider [104], and the assemble species by automatic partitioning (ASAP) [94] was applied. The ASAP generates 10 partitions computing for each partition the asap-score, which is the average of two components: probability of panmixia (p -value) and barcoding gap width.

The lower the score, the better the partition for the dataset. As input data, alignments of plastomes and mitogenomes of *Calypogeia* species were separately used in each analysis.

5. Conclusions

Sometimes, something that looks predictable can surprise. With genomic sequencing data, the stability of the mitochondrial genome of *Calypogeia* was tested. The high nucleotide variability of the mitogenome of the European species of this genus astonished and pushed to check their usefulness as super-barcodes. As it turned out, it might work! The *Calypogeia* species were correctly identified. However, the precision of the species delimitation depended on the applied method and should be interpreted in the light of the general knowledge about the investigated taxon.

The results of the research raise an important issue. Lack of data or scant data on nucleotide variation of genomes among and within bryophyte species may make scientists blind to some lines of research. The science world seems to keep looking at bryophytes as “evolutionary sphinxes of the past” whose organelle genomes evolve slowly. So, can these “sphinxes” surprise with anything else? Fortunately, they can hide many curiosities, but to discover them, you have to think outside the box and not hesitate to check seeming facts.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms232415570/s1>, Figure S1. Line plot of the barcode gap for the *Calypogeia* species. A. Data based on mitogenome. B. Data based on the plastome. For each individual in the dataset, the lines represent the furthest intraspecific distance (bottom of line value), and the closest interspecific distance (top of line value); Table S1 SNPs, indel variation and π values within mitochondrial coding (CDS) and noncoding regions of *Calypogeia* species.; Table S2. Species used in this study, sampling data and sequencing results.; Table S3. Number of MDCs for every *Calypogeia* species based on mitochondrial genome.

Author Contributions: M.Ś., Writing—original draft preparation, Conceptualization, Methodology, Formal analysis, Visualization; K.M., Formal analysis, Methodology; K.B. and A.B., Resources, Writing—original draft preparation; J.S., Conceptualization, Formal analysis, Funding acquisition, Visualization, Supervision. All authors have read and agreed to the published version of the manuscript.

Funding: The research was financially supported by the National Science Center, 30-312 Kraków, Poland. Thanks to Grant No. 2015/19/B/NZ8/03970, we were able to pay for sequencing of the organelle genomes of *Calypogeia*. This was commissioned to a commercial company. The design of the study, collection of plant material, analyses and interpretation of data and writing the paper were made independently only by the authors of the manuscript without financial support.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The newly sequenced organellar genomes can be found in GenBank under the accession numbers given in Supplementary Table S2. Other supplementary data are also available in the Supplementary Materials.

Acknowledgments: We are grateful to the Tatra National Park and the Bieszczady National Park for the possibility of collecting samples. Thanks also to D.A. Callaghan and Alfons Schäfer-Verwimp for sharing the samples.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hebert, P.D.N.; Cywinska, A.; Ball, S.L.; deWaard, J.R. Biological identification through DNA barcodes. *Proc. Royal Soc. B.* **2003**, *270*, 313–321. [[CrossRef](#)] [[PubMed](#)]
2. Rubinoff, D.; Cameron, S.; Kipling, W. Are plant DNA barcodes a search for the Holy Grail? *Trends Ecol. Evol.* **2006**, *21*, 1–2. [[CrossRef](#)] [[PubMed](#)]
3. Cho, Y.; Qiu, Y.-L.; Kuhlman, P.; Palmer, J.D. Explosive invasion of plant mitochondria by a group I intron. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14244–14249. [[CrossRef](#)] [[PubMed](#)]

4. Cho, Y.; Mower, J.P.; Qiu, Y.-L.; Palmer, J.D. Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 17741–17746. [[CrossRef](#)]
5. Fazekas, A.J.; Burgess, K.S.; Kesanakurti, P.R.; Graham, S.W.; Newmaster, S.G.; Husband, B.C.; Percy, D.M.; Hajibabaei, M.; Barrett, S.C.H. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE* **2008**, *3*, e2802. [[CrossRef](#)]
6. Hollingsworth, P.M.; Graham, S.W.; Damon, D.P. Choosing and Using a Plant DNA Barcode. *PLoS ONE* **2011**, *6*, e19254. [[CrossRef](#)]
7. Galtier, N. The intriguing evolutionary dynamics of plant mitochondrial DNA. *BMC Biol.* **2011**, *9*, 61. [[CrossRef](#)]
8. Gualberto, J.M.; Mileshina, D.; Wallet, C.; Niazi, A.K.; Weber-Lotfi, F.; Dietrich, A. The plant mitochondrial genome: Dynamics and maintenance. *Biochimie* **2013**, *100*, 107–120. [[CrossRef](#)]
9. Kozik, A.; Rowan, B.A.; Lavelle, D.; Berke, L.; Schranz, M.E.; Michelmore, R.W.; Christensen, A.C. The alternative reality of plant mitochondrial DNA: One ring does not rule them all. *PLoS Genet.* **2019**, *15*, e1008373. [[CrossRef](#)]
10. Morley, S.A.; Brent, L.N. Plant mitochondrial DNA. *Front. Biosci.* **2017**, *22*, 1023–1032. [[CrossRef](#)]
11. Skippington, E.; Barkman, T.J.; Rice, D.W.; Palmer, J.D. Miniaturized mitogenome of the parasitic plant *Viscum scurruloideum* is extremely divergent and dynamic and has lost all nad genes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E3515–E3524. [[CrossRef](#)] [[PubMed](#)]
12. Sloan, D.B.; Alverson, A.J.; Chuckalovcak, J.P.; Wu, M.; McCauley, D.E.; Palmer, J.D.; Taylor, D.R. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol.* **2012**, *10*, e1001241. [[CrossRef](#)] [[PubMed](#)]
13. Putintseva, Y.A.; Bondar, E.I.; Simonov, E.P.; Sharov, V.V.; Oreshkova, N.V.; Kuzmin, D.A.; Konstantinov, Y.M.; Shmakov, V.N.; Belkov, V.I.; Sadovsky, M.G.; et al. Siberian larch (*Larix sibirica* Ledeb.) mitochondrial genome assembled using both short and long nucleotide sequence reads is currently the largest known mitogenome. *BMC Genom.* **2020**, *21*, 654. [[CrossRef](#)]
14. Forget, L.; Ustinova, J.; Wang, Z.; Huss, V.A.R.; Franz Lang, B. *Hyaloraphidium curvatum*: A linear mitochondrial genome, tRNA editing, and an evolutionary link to lower fungi. *Mol. Biol. Evol.* **2002**, *19*, 310–319. [[CrossRef](#)]
15. Liu, W.; Cai, Y.; Zhang, Q.; Chen, L.; Shu, F.; Ma, X.; Bian, Y. The mitochondrial genome of *Morchella importuna* (272.2 kb) is the largest among fungi and contains numerous introns, mitochondrial non-conserved open reading frames and repetitive sequences. *Int. J. Biol. Macromol.* **2019**, *143*, 373–381. [[CrossRef](#)] [[PubMed](#)]
16. Oldenburg, D.J.; Bendich, A.J. Size and structure of replicating mitochondrial DNA in cultured tobacco cells. *Plant Cell* **1996**, *8*, 447–461. [[CrossRef](#)]
17. Backert, S.; Börner, T. Phage T4-like intermediates of DNA replication and recombination in the mitochondria of the higher plant *Chenopodium album* (L.). *Curr. Genet.* **2000**, *37*, 304–314. [[CrossRef](#)] [[PubMed](#)]
18. Bendich, A.J. Structural analysis of mitochondrial DNA molecules from fungi and plants using moving pictures and pulsed-field gel electrophoresis. *J. Mol. Biol.* **1996**, *255*, 564–588. [[CrossRef](#)]
19. Oldenburg, D.J.; Bendich, A.J. The structure of mitochondrial DNA from the liverwort, *Marchantia polymorpha*. *J. Mol. Biol.* **1998**, *276*, 745–758. [[CrossRef](#)]
20. Manchekar, M.; Scissum-Gunn, K.; Song, D.; Khazi, F.; McLean, S.L.; Nielsen, B.L. DNA recombination activity in soybean mitochondria. *J. Mol. Biol.* **2006**, *356*, 288–299. [[CrossRef](#)]
21. Mower, J.P.; Case, A.L.; Floro, E.R.; Willis, J.H. Evidence against equimolarity of large repeat arrangements and a predominant master circle structure of the mitochondrial genome from a monkeyflower (*Mimulus guttatus*) lineage with cryptic CMS. *Genome Biol. Evol.* **2012**, *4*, 670–686. [[CrossRef](#)] [[PubMed](#)]
22. Cheng, N.; Lo, Y.; Ansari, M.I.; Ho, K.; Jeng, S.; Lin, N.; Dai, H. Correlation between mtDNA complexity and mtDNA replication mode in developing cotyledon mitochondria during mung bean seed germination. *New Phytol.* **2017**, *213*, 751–763. [[CrossRef](#)] [[PubMed](#)]
23. Palmer, J.D. Intraspecific variation and multicircularity in *Brassica* mitochondrial DNAs. *Genetics* **1988**, *118*, 341–351. [[CrossRef](#)] [[PubMed](#)]
24. Palmer, J.D.; Adams, K.L.; Cho, Y.; Parkinson, C.L.; Qiu, Y.-L.; Song, K. Dynamic evolution of plant mitochondrial genomes: Mobile genes and introns and highly variable mutation rates. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6960–6966. [[CrossRef](#)] [[PubMed](#)]
25. Alverson, A.J.; Zhuo, S.; Rice, D.W.; Sloan, D.B.; Palmer, J.D. The mitochondrial genome of the legume *Vigna radiata* and the analysis of recombination across short mitochondrial repeats. *PLoS ONE* **2011**, *6*, e16404. [[CrossRef](#)] [[PubMed](#)]
26. Dong, S.; Zhao, C.; Chen, F.; Liu, Y.; Zhang, S.; Wu, H.; Zhang, L.; Liu, Y. The complete mitochondrial genome of the early flowering plant *Nymphaea colorata* is highly repetitive with low recombination. *BMC Genom.* **2018**, *19*, 614. [[CrossRef](#)]
27. Chevigny, N.; Schatz-Daas, D.; Lotfi, F.; Gualberto, J.M. DNA Repair and the stability of the plant mitochondrial genome. *Int. J. Mol. Sci.* **2020**, *21*, 328. [[CrossRef](#)]
28. Palmer, J.D.; Herbon, L.A. Plant mitochondrial DNA evolved rapidly in structure, but slowly in sequence. *J. Mol. Evol.* **1988**, *28*, 87–97. [[CrossRef](#)]
29. Fauron, C.M.R.; Moore, B.; Casper, M. Maize as a model of higher plant mitochondrial genome plasticity. *Plant Sci.* **1995**, *112*, 11–32. [[CrossRef](#)]
30. Dong, S.; Chen, L.; Liu, Y.; Wang, Y.; Zhang, S.; Yang, L.; Lang, X. The draft mitochondrial genome of *Magnolia biondii* and mitochondrial phylogenomics of angiosperms. *PLoS ONE* **2020**, *15*, e0231020. [[CrossRef](#)]

31. Mower, J.P.; Jain, K.; Hepburn, N.J. The role of horizontal transfer in shaping the plant mitochondrial genome. In *Advances in Botanical Research, Volume 63: Mitochondrial Genome Evolution*; Maréchal Drouard, L., Ed.; Elsevier: Amsterdam, The Netherlands, 2012; pp. 41–69.
32. Szandar, K.; Krawczyk, K.; Myszczyński, K.; Ślipiko, M.; Sawicki, J.; Szczecińska, M. Breaking the limits-multichromosomal structure of an early eudicot *Pulsatilla patens* mitogenome reveals extensive RNA-editing, longest repeats and chloroplast derived regions among sequenced land plant mitogenomes. *BMC Plant Biol.* **2022**, *22*, 109. [[CrossRef](#)]
33. Wolfe, K.H.; Li, W.H.; Sharp, P.M. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 9054–9058. [[CrossRef](#)]
34. Drouin, G.; Daoud, H.; Xia, J. Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Mol. Phylogenet. Evol.* **2008**, *49*, 827–831. [[CrossRef](#)] [[PubMed](#)]
35. Richardson, A.O.; Rice, D.W.; Young, G.J.; Alverson, A.J.; Palmer, J.D. The “fossilized” mitochondrial genome of *Liriodendron tulipifera*: Ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. *BMC Biol.* **2013**, *11*, 29. [[CrossRef](#)] [[PubMed](#)]
36. Bakker, F.T.; Breman, F.; Merckx, V. DNA sequence evolution in fast evolving mitochondrial DNA *nad1* exons in Geraniaceae and Plantaginaceae. *Taxon* **2006**, *55*, 887–896. [[CrossRef](#)]
37. Sloan, D.B.; Oxelman, B.; Rautenberg, A.; Taylor, D.R. Phylogenetic analysis of mitochondrial substitution rate variation in the angiosperm tribe Sileneae. *BMC Evol. Biol.* **2009**, *9*, 260. [[CrossRef](#)]
38. Zhu, A.D.; Guo, W.H.; Jain, K.; Mower, J.P. Unprecedented heterogeneity in the synonymous substitution rate within a plant genome. *Mol. Biol. Evol.* **2014**, *31*, 1228–1236. [[CrossRef](#)]
39. Liu, Y.; Medina, R.; Goffinet, B. 350 My of mitochondrial genome stasis in mosses, an early land plant lineage. *Mol. Biol. Evol.* **2014**, *31*, 2586–2591. [[CrossRef](#)]
40. Dong, S.; Zhao, C.; Zhang, S.; Zhang, L.; Wu, H.; Liu, H.; Zhu, R.; Jia, Y.; Goffinet, B.; Liu, Y. Mitochondrial genomes of the early land plant lineage liverworts (Marchantiophyta): Conserved genome structure, and ongoing low frequency recombination. *BMC Genom.* **2019**, *20*, 953. [[CrossRef](#)]
41. Myszczyński, K.; Górski, P.; Ślipiko, M.; Sawicki, J. Sequencing of organellar genomes of *Gymnomitrium concinnatum* (Jungermanniales) revealed the first exception in the structure and gene order of evolutionary stable liverworts mitogenomes. *BMC Plant Biol.* **2018**, *18*, 321. [[CrossRef](#)]
42. CBOL Plant Working Group. A DNA barcode for land plants. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12794–12799. [[CrossRef](#)] [[PubMed](#)]
43. Hollingsworth, P.M.; Li, D.-Z.; van der Bank, M.; Twyford, A.D. Telling plant species apart with DNA: From barcodes to genomes. *Phil. Trans. R. Soc. B* **2016**, *371*, 20150338. [[CrossRef](#)] [[PubMed](#)]
44. Liu, J.; Jiang, J.; Song, S.; Tornabene, L.; Chabarría, R.; Naylor, G.J.P.; Li, C. Multilocus DNA barcoding—Species identification with multilocus Data. *Sci. Rep.* **2017**, *7*, 16601. [[CrossRef](#)]
45. Naciri, Y.; Caetano, S.; Salamin, N. Plant DNA barcodes and the influence of gene flow. *Mol. Ecol. Resour.* **2012**, *12*, 575–580. [[CrossRef](#)] [[PubMed](#)]
46. Percy, D.M.; Argus, G.W.; Cronk, Q.C.; Fazekas, A.J.; Kesanakurti, P.R.; Burgess, K.S.; Husband, B.C.; Newmaster, S.G.; Barrett, S.C.; Graham, S.W. Understanding the spectacular failure of DNA barcoding in willows (*Salix*): Does this result from a trans-specific selective sweep? *Mol. Ecol.* **2014**, *23*, 4737–4756. [[CrossRef](#)]
47. Twyford, A.D. Testing evolutionary hypotheses for DNA barcoding failure in willows. *Mol. Ecol.* **2014**, *23*, 4674–4676. [[CrossRef](#)]
48. Kress, W.J.; Wurdack, K.J.; Zimmer, E.A.; Weigt, L.A.; Janzen, D.H. Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 8369–8374. [[CrossRef](#)]
49. Li, D.Z.; Gao, L.M.; Li, H.T.; Wang, H.; Ge, X.J.; Liu, J.Q.; Chen, Z.D.; Zhou, S.L.; Chen, S.L.; China Plant BOL Group. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 19641–19646. [[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. Amar, M.H. *ycf1-ndhF* genes, the most promising plastid genomic barcode, sheds light on phylogeny at low taxonomic levels in *Prunus persica*. *J. Genet. Eng. Biotechnol.* **2020**, *18*, 42. [[CrossRef](#)]
52. 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](#)] [[PubMed](#)]
53. 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](#)] [[PubMed](#)]
54. Bączkiewicz, A.; Szczecińska, M.; Sawicki, J.; Stebel, A.; Buczkowska, K. DNA barcoding, ecology and geography of cryptic species of *Aneura pinguis* and their relationship with *Aneura maxima* and *Aneura mirabilis* (Metzgeriales, Marchantiophyta). *PLoS ONE* **2017**, *12*, e0188837. [[CrossRef](#)] [[PubMed](#)]
55. Liu, Y.; Yan, H.F.; Cao, T.; Ge, X.J. Evaluation of 10 plant barcodes in Bryophyta (mosses). *J. Syst. Evol.* **2010**, *48*, 36–46. [[CrossRef](#)]
56. Ji, Y.; Liu, C.; Yang, J.; Jin, L.; Yang, Z.; Yang, J.B. Ultra-barcoding discovers a cryptic species in *Paris yunnanensis* (Melanthiaceae), a medicinally important plant. *Front. Plant Sci.* **2020**, *22*, 411. [[CrossRef](#)] [[PubMed](#)]

57. Bakalin, V.A.; Klimova, K.G.; Nguyen, V.S. A review of *Calypogeia* (Marchantiophyta) in the eastern Sino-Himalaya and Meta-Himalaya based mostly on types. *PhytoKeys* **2020**, *153*, 111–154. [[CrossRef](#)]
58. Warnstorff, C. Die europäische Artgruppen der Gattung *Calypogeia* Raddi (1820). *Bryolog. Zeitschr.* **1917**, *1*, 97–114.
59. Buczkowska, K. Genetic Differentiation of *Calypogeia fissa* Raddi (Hepaticae, Jungermanniales) in Poland. *Plant Syst. Evol.* **2004**, *247*, 187–201. [[CrossRef](#)]
60. Buczkowska, K.; Baczkiewicz, A.; Sawicki, J.; Szczecińska, M. Genetic Differentiation within *Calypogeia sphagnicola* (Jungermanniales, Hepaticae) in Poland. *Nowellia Bryol.* **2009**, 34–42.
61. Buczkowska, K.; Sawicki, J.; Szczecińska, M.; Klama, H.; Baczkiewicz, A. Allopolyploid speciation of *Calypogeia sphagnicola* (Jungermanniopsida, Calypogeiaceae) based on isozyme and DNA markers. *Plant Syst. Evol.* **2012**, *298*, 549–560. [[CrossRef](#)]
62. Ślipiko, M.; Myszczynski, K.; Buczkowska-Chmielewska, K.; Baczkiewicz, A.; Szczecińska, M.; Sawicki, J. Comparative analysis of four *Calypogeia* species revealed unexpected change in evolutionarily-stable liverwort mitogenomes. *Genes* **2017**, *8*, 395. [[CrossRef](#)] [[PubMed](#)]
63. Buczkowska, K.; Bakalin, V.; Baczkiewicz, 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](#)] [[PubMed](#)]
64. Asaf, S.; Khan, A.L.; Khan, A.R.; Waqas, M.; Kang, S.-M.; Khan, M.A.; Shahzad, R.; Seo, C.-W.; Shin, J.-H.; Lee, I.-J. Mitochondrial genome analysis of wild rice (*Oryza minuta*) and its comparison with other related species. *PLoS ONE* **2016**, *11*, e0152937. [[CrossRef](#)] [[PubMed](#)]
65. Kan, S.; Liao, X.; Wu, Z. The roles of mutation and selection acting on mitochondrial genomes inferred from intraspecific variation in seed plants. *Genes* **2022**, *13*, 1036. [[CrossRef](#)]
66. Dong, S.; Liu, Y. The mitochondrial genomes of bryophytes. *Bry. Div. Evo.* **2021**, *43*, 112–126. [[CrossRef](#)]
67. Myszczynski, K.; Baczkiewicz, 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](#)]
68. Van de Paer, C.; Bouchez, O.; Besnard, G. Prospects on the evolutionary mitogenomics of plants: A case study on the olive family (Oleaceae). *Mol. Ecol. Res.* **2017**, *18*, 407–423. [[CrossRef](#)]
69. Linde, A.-M.; Eklund, D.M.; Cronberg, N.; Bowman, J.L.; Lagercrantz, U. Rates and patterns of molecular evolution in bryophyte genomes, with focus on complex thalloid liverworts, Marchantiopsida. *Mol. Phylogenet. Evol.* **2021**, *165*, 107295. [[CrossRef](#)]
70. Tong, W.; He, Q.; Park, Y.J. Genetic variation architecture of mitochondrial genome reveals the differentiation in Korean landrace and weedy rice. *Sci. Rep.* **2017**, *7*, 43327. [[CrossRef](#)]
71. Sawicki, J.; Plášek, V.; Ochyra, R.; Szczecińska, M.; Ślipiko, M.; Myszczynski, K.; Kulik, T. Mitogenomic analyses support the recent division of the genus *Orthotrichum* (Orthotrichaceae, Bryophyta). *Sci. Rep.* **2017**, *7*, 4408. [[CrossRef](#)]
72. Choi, S.S.; Min, J.; Kwon, W.; Park, J. The complete mitochondrial genome of *Scapania ampliata* Steph., 1897 (Scapaniaceae, Jungermanniales). *Mitochondrial DNA B Resour.* **2021**, *6*, 686–688. [[CrossRef](#)] [[PubMed](#)]
73. Kwon, W.; Kim, Y.; Park, J. The complete mitochondrial genome of Korean *Marchantia polymorpha* subsp. *ruderalis* Bischl. & Boisselier: Inverted repeats on mitochondrial genome between Korean and Japanese isolates. *Mitochondrial DNA B* **2019**, *4*, 769–770. [[CrossRef](#)]
74. Kwon, W.; Kim, Y.; Park, J. The complete mitochondrial genome of *Dumortiera hirsuta* (Sw.) Nees (Dumortieraceae, Marchantiophyta). *Mitochondrial DNA B* **2019**, *4*, 1586–1587. [[CrossRef](#)]
75. Min, J.; Kwon, W.; Xi, H.; Park, J. The complete mitochondrial genome of *Riccia fluitans* L. (Ricciaceae, Marchantiophyta): Investigation of intraspecific variations on mitochondrial genomes of *R. fluitans*. *Mitochondrial DNA B* **2020**, *5*, 1220–1222. [[CrossRef](#)] [[PubMed](#)]
76. 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](#)]
77. Bakalin, V.A.; Maltseva, Y.D.; Müller, F.; Klimova, K.G.; Nguyen, V.S.; Choi, S.S.; Troitsky, A.V. *Calypogeia* (Calypogeiaceae, Marchantiophyta) in Pacific Asia: Updates from molecular revision with particular attention to the genus in North Indochina. *Plants* **2022**, *11*, 983. [[CrossRef](#)] [[PubMed](#)]
78. Schuster, R.M. *The Hepaticae and Anthocerotae of North America East of the Hundredth Meridian*; Columbia University Press: New York, NY, USA; London, UK, 1969; Volume 2, pp. 98–215.
79. Paton, J.A. *The Liverwort Flora of the British Isles*; Harley Books: Colchester, UK, 1999; pp. 77–92.
80. Buczkowska, K. The genus *Calypogeia* Raddi (Hepaticae Jungermanniales) in Poland: biometrical analysis of morphological and anatomical variation. *Nova Hedwig.* **2004**, *78*, 121–146. [[CrossRef](#)]
81. Buczkowska, K.; Odrzykoski, I.J.; Chudzińska, E. Delimitation of some European species of *Calypogeia* Raddi (Jungermanniales, Hepaticae) based on cytological characters and multienzyme phenotype. *Nova Hedwig.* **2004**, *78*, 147–163. [[CrossRef](#)]
82. Buczkowska, K.; Dabert, M. The development of species-specific SCAR markers for delimitation of *Calypogeia* species. *J. Bryol.* **2011**, *33*, 291–299. [[CrossRef](#)]
83. Buczkowska, K. *Characteristics of Calypogeia Suecica Groups, Unpublished Data*; Department of Biology, Institute of Experimental Biology, Adam Mickiewicz University in Poznań: Poznań, Poland, 2022.

84. Li, X.; Yang, Y.; Henry, R.J.; Rossetto, M.; Wang, Y.; Chen, S. Plant DNA barcoding: From gene to genome. *Biol. Rev. Camb. Phil. Soc.* **2015**, *90*, 157–166. [[CrossRef](#)]
85. Zhang, D.; Mo, X.; Xiang, J.; Zhou, N. Molecular identification of original plants of *Fritillariae cirrhosae* bulbs, a traditional chinese medicine (TCM) using plant DNA barcoding. *Afr. J. Tradit. Complement. Altern. Med.* **2016**, *13*, 74–82. [[CrossRef](#)] [[PubMed](#)]
86. Wu, F.-H.; Chan, M.-T.; Liao, D.-C.; Hsu, C.-T.; Lee, Y.-W.; Daniell, H.; Duvall, M.R.; Lin, C.-S. Complete chloroplast genome of *Oncidium* Gower Ramsey and evaluation of molecular markers for identification and breeding in *Oncidiinae*. *BMC Plant Biol.* **2010**, *10*, 68. [[CrossRef](#)] [[PubMed](#)]
87. Carbonell-Caballero, J.; Alonso, R.; Ibañez, V.; Terol, J.; Talon, M.; Dopazo, J. A phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus *Citrus*. *Mol. Biol. Evol.* **2015**, *32*, 2015–2035. [[CrossRef](#)] [[PubMed](#)]
88. Jamdade, R.; Mosa, K.A.; El-Keblawy, A.; Al Shaer, K.; Al Harthi, E.; Al Sallani, M.; Al Jasmi, M.; Gairola, S.; Shabana, H.; Mahmoud, T. DNA barcodes for accurate identification of selected medicinal plants (Caryophyllales): Toward barcoding flowering plants of the United Arab Emirates. *Diversity* **2022**, *14*, 262. [[CrossRef](#)]
89. Rubinoff, D.; Cameron, S.; Will, K. A genomic perspective on the shortcomings of mitochondrial DNA for "barcoding" identification. *J. Hered.* **2006**, *97*, 581–594. [[CrossRef](#)]
90. Krawczyk, K.; Szczecińska, M.; Sawicki, J. Evaluation of 11 single-locus and seven multilocus DNA barcodes in *Lamium* L. (Lamiaceae). *Mol. Ecol. Res.* **2014**, *14*, 272–285. [[CrossRef](#)]
91. Mallo, D.; De Oliveira, M.L.; Posada, D. SimPhy: Phylogenomic simulation of gene, locus, and species trees. *Syst. Biol.* **2016**, *65*, 334–344. [[CrossRef](#)]
92. 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* bulb and its adulterants. *PLoS ONE* **2020**, *15*, e0229181. [[CrossRef](#)]
93. Chen, Q.; Hu, H.; Zhang, D. DNA Barcoding and Phylogenomic Analysis of the genus *Fritillaria* in China based on complete chloroplast genomes. *Front. Plant Sci.* **2022**, *13*, 764255. [[CrossRef](#)]
94. Puillandre, N.; Brouillet, S.; Achaz, G. ASAP: Assemble species by automatic partitioning. *Mol. Ecol. Resour.* **2021**, *21*, 609–620. [[CrossRef](#)] [[PubMed](#)]
95. Luo, A.; Ling, C.; Ho, S.Y.W.; Zhu, C.D. Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Syst. Biol.* **2018**, *67*, 830–846. [[CrossRef](#)] [[PubMed](#)]
96. Fujisawa, T.; Barraclough, T.G. Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: A revised method and evaluation on simulated data sets. *Syst. Biol.* **2013**, *62*, 707–724. [[CrossRef](#)] [[PubMed](#)]
97. Kekkonen, M.; Hebert, P.D. DNA barcode-based delineation of putative species: Efficient start for taxonomic workflows. *Mol. Ecol. Resour.* **2014**, *14*, 706–715. [[CrossRef](#)] [[PubMed](#)]
98. Zhang, J.; Kapli, P.; Pavlidis, P.; Stamatakis, A. A general species delimitation method with applications to phylogenetic placements. *Bioinform.* **2013**, *29*, 2869–2876. [[CrossRef](#)]
99. Sukumaran, J.; Knowles, L.L. Multispecies coalescent delimits structure, not species. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 1607–1612. [[CrossRef](#)]
100. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
101. Dierckxens, N.; Mardulyn, P.; Smits, G. NOVOPlasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* **2016**, *45*, e18. [[CrossRef](#)]
102. Kears, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [[CrossRef](#)]
103. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
104. Brown, S.D.; Collins, R.A.; Boyer, S.; Lefort, M.C.; Malumbres-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](#)]
105. Huelsenbeck, J.P.; Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **2001**, *17*, 754–755. [[CrossRef](#)] [[PubMed](#)]
106. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [[CrossRef](#)] [[PubMed](#)]
107. Britton, T.; Anderson, C.L.; Jacquet, D.; Lundqvist, S.; Bremer, K. Estimating divergence times in large phylogenetic trees. *Syst. Biol.* **2007**, *56*, 741–752. [[CrossRef](#)] [[PubMed](#)]
108. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)] [[PubMed](#)]