The Mitochondrial Genome of Nematodontous Moss

**Polytrichum commune** and Analysis of Intergenic Repeats Distribution Among Bryophyta

Denis V. Goryunov 1,*, Evgeniia A. Sotnikova 2, Svetlana V. Goryunova 3,4, Oxana I. Kuznetsova 5, Maria D. Logacheva 1, Irina A. Milyutina 1, Alina V. Fedorova 5, Vladimir E. Fedosov 6,7 and Aleksey V. Troitsky 1,*

1 Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Leninskie Gory 1, Moscow 119992, Russia; maria.loz@gmail.com (M.D.L.); irmilyutina@yandex.ru (I.A.M.)
2 National Medical Research Center for Therapy and Preventive Medicine, Petrovskiy per., 10, bld. 3, Moscow 101000, Russia; sotnikova.evgeniya@gmail.com
3 Institute of General Genetics Russian Academy of Science, GSP-1 Gubkin str. 3, Moscow 119991, Russia; orang2@yandex.ru
4 Russian Potato Research Center, 23 Lorkh Str., Kraskovo, Moscow Region 140051, Russia
5 Tsitsin Main Botanical Garden Russian Academy of Science, Botanicheskaya st. 4, Moscow 127276, Russia; oikuznets@gmail.com (O.I.K.); alina_77777@mail.ru (A.V.F.)
6 Faculty of Biology, Lomonosov Moscow State University, Leninskie Gory 1–12, Moscow 119234, Russia; fedosov_v@mail.ru
7 Botanical Garden-Institute, FEB RAS, Makovskogo Street 142, Vladivostok 690024, Russia
* Correspondence: denis.goryunov@mail.ru (D.V.G.); bobr@belozersky.msu.ru (A.V.T.)

**Abstract:** An early-branched moss *Polytrichum commune* is a widely accepted model object for ecological, environmental, physiological, and genetic studies. Its mitochondrial genome has been sequenced and annotated. The genome contains 67 genes in total and has a length equal to 114,831 bp, which exceeds the length of most known mitochondrial genomes for mosses. A phylogenetic tree based on 33 coding sequences of mitochondrial genome was constructed, and the pairwise identity of whole mitogenome sequences was estimated for 44 Bryophyta species. Based on the analysis of pairwise identity, it was shown that mitogenomes of *Tetraphis pellucida* and *Buxbaumia aphylla* sufficiently differ from those of other Bryophyta species. The first known Bryophyta mitogenome rearrangement was identified in *Pogonatum inflexum* within Polytrichopsida. Based on the intergenic repeats occurrence in 44 bryophyte mitochondrial genomes and available data on repetitive elements content in other Viridiplantae groups, it was noted for the first time that greater stability of the moss’s mitogenomes is probably associated mainly with the absence of long (>1 kb) repeats. The phenomenon of absence of the intergenic repetitive elements in the terminal clades species was discovered.

**Keywords:** Bryophyta; *Polytrichum*; mitochondrial genome; simple sequence repeats; intergenic repeats; molecular phylogeny; inversion; genome rearrangement; evolution; structural variability

---

1. **Introduction**

*Polytrichum commune* Hedw., the common hair-cap moss, is a generitype of the genus *Polytrichum*, the type genus of Polytrichaceae (Polytrichales, Polytrichopsida). Polytrichopsida includes ca. 23 genera and ca. 200 species [1–3]. The species classified in the genus *Polytrichum* are common in the boreal zone and Arctic, where they occupy different ecotopes with disturbed vegetation. For instance, *P. commune* is especially abundant in irrigated sphagnous bogs and forests, where it plays an important role replacing peat moss cover. *Polytrichum commune* is a widely accepted model object for ecological, environmental, physiological, and genetic studies. Thus, it got ca. 12,300 results, obtained from scholar.google.com.
Due to their considerable variety of forms but relatively few species, Polytrichopsida has been considered as an apparently ancient group with a position strikingly analogous to that of the conifers within the seed plant clade [2]. In terms of the peristome structure, which is traditionally used for establishing the high-level systematics of mosses, Polytrichopsida belongs to nematodontous mosses, which differ from arthrodontous mosses (Bryopsida). In nematodontous peristome, the teeth of peristome are solid and composed of bundles of dead cells. Each tooth of the arthrodontous peristome type is composed of periclinal (tangential) cell-wall remnants between two of the three concentric peristomial cell layers.

Representatives of Polytrichales (Polytrichopsida) and Tetraphidales (Tetraphidopsida) have a unique trait called nematodontous peristome, and therefore, they have recently been often united into a single class Polytrichopsida [1,3]. According to modern understanding, these mosses occupy a basal position in relation to the majority of mosses that bear arthrodontous peristome and are attributed to Bryopsida, while the earliest divergent lineages of mosses (Takakiopsida, Sphagnopsida, Andreaeopsida, Andreaeobryopsida and likely also Oedipodiopsida) are eperistomate. Earlier molecular research yielded conflicting results regarding relationships among the early-diverging peristomate moss lineages [4]. Thus, monotypic Oedipodiales has been sometimes included in the class Polytrichopsida [3], while Buxbaumiaceae, possessing peristome of a deviating type [5] was considered closer to Tetraphis [6] or referred to as arthrodontous Bryopsida [1,3,7,8]. According to recent phylogenetic reconstructions, Oedipodiales occupies an intermediate position between Andreaeobryophytina and Tetraphidales, while Buxbaumiaceae are classified in arthrodontous Bryopsida [7,8].

Genomic data may be crucial for large-scale phylogenetic reconstructions. In this regard, at present, the greatest hopes are pinned on the analysis of the genomes of organelles—mitochondria and plastids. The size and structure of the mitochondrial genomes is extremely variable among the Viridiplantae. Nonetheless, mitogenomes of the liverworts and mosses are very stable and conservative. Only single rearrangements in their structure have been found so far [9]. Sufficient steps toward understanding the nature of this phenomenon were performed recently [10]. However, this fundamental issue is still far from complete, and additional efforts are required to shed light on the mechanisms behind the mitogenome stability in bryophytes. In this study, we contribute to understanding the nature of the above-mentioned phenomenon.

The variety of mitochondrial genomes of bryophytes, a highly diverse group of land plants, still remains insufficiently studied in comparison with angiosperms. Large-scale trends of mitogenome evolution are of special interest; however, to address these issues properly, a denser sampling of ancient, early-diverged lineages of bryophytes is crucial. The sequencing of the mitochondrial genome of the nematodontous moss Polytrichum commune extends the already-known diversity of moss mitogenomes.

2. Materials and Methods

2.1. Sample Collection and DNA Isolation

The *P. commune* samples were collected in the summer season of 2014 in Klin Distr. of Moscow Province, Middle European Russia; the reference specimen was deposited in MW (Herbarium of Moscow University).

A Nucleospin Plant DNA Kit (Macherey Nagel, Düren, Germany) was used for total DNA extraction from whole shoots of freshly collected plants according to the manufacturer’s protocol. A yield of about 2 µg DNA was obtained according to measurements determined with a Qubit fluorometer (Invitrogen, Carlsbad, CA, USA).

2.2. Library Preparation and Sequencing

A 500 ng sample of genomic DNA was fragmented using a Covaris S220 sonicator (Covaris, Woburn, MA, USA), and a library was prepared using a TruSeq DNA sample preparation kit (Illumina, Mountain View, CA, USA). The concentration of the prepared library was measured with the Qubit fluorometer (Invitrogen, Carlsbad, CA, USA), and the
qPCR and fragment length distribution were determined with Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA). The library was diluted to 10 pM and used for cluster generation on a cBot instrument with TruSeq PE Cluster Kit v3 reagents (Illumina, Mountain View, CA, USA). Sequencing was performed on a HiSeq2000 machine (Illumina, San Diego, CA, USA) with a read length of 101 from both ends of the fragments. About 17.2 million paired-end reads were obtained.

2.3. Mitogenome Assembly and Annotation

Raw sequencing reads were preprocessed with Trimmomatic software [11]. The whole genome assembly was then accomplished using the Spades assembler [12]. A Blast database was generated from the assembled contigs, and a Blast search was performed against the *Physcomitrella patens* (Hedw.) Bruch & Schimp. mitochondrial genome sequence [13] using the standalone NCBI BLAST-2.2.29+ [14]. Among the longest hits was the *P. commune* complete mitogenome. Iterative mapping was carried out using Geneious R10 software (https://www.geneious.com) [15] to verify correctness of the assembled genome. The resulting sequence had almost 86× coverage depth. The genome boundaries were verified by PCR amplification followed by Sanger sequencing.

Genome annotation based on sequence similarity was performed using Geneious software. The mitogenome sequence of *Atrichum angustatum* (Brid.) Bruch & Schimp. (GenBank accession number KC784956) was applied as a reference as it gave a maximal percent identity and total score in a blast search on the NCBI website (https://www.ncbi.nlm.nih.gov) against a *P. commune* complete mitogenome. The annotated *P. commune* mitogenome sequence was submitted to the GenBank (accession number NC039775). A circular genome map was drawn using the CGView Server (http://stothard.afns.ualberta.ca/cgview_server) [16].

2.4. SSR Analysis

Simple sequence repeats (SSRs) were detected and located in the mitogenome of *P. commune* using GMATo v1.2 software [17].

2.5. Analysis of Repetitive Elements

The search for repeating sequences was performed using the Repeat Finder plugin in the Geneious program with the minimum repeat length of 51 bp and maximum mismatches at 15%. Each identified repeat was observed at least twice in the analyzed genome. In case all the subunits of a particular repeat were nested in those of another longer repeat, the former was excluded from the consideration. If all the repeat subunits overlapped each other or were located in the same intergenic spacer, the repeat was also filtered out.

2.6. Phylogenomic and Genome-Wide Comparison Analyses

For phylogenetic reconstruction, only functional protein-coding sequences (CDS) of the whole bryophytes mitogenomes deposited in the GenBank (http://www.ncbi.nlm.nih.gov) were selected. A total of 33 CDS of genes including *cox1*, *atp4*, *rps7*, *atp1*, *atp8*, *atp6*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *rpl2*, *rpl5*, *rpl16*, *rps1*, *rps2*, *rps4*, *rps12*, *rps11*, *rps13*, *rps14*, *rps19*, *sdh3*, *sdh4*, and *tatC* were used for the phylogeny inference. The CDS of 44 mosses were aligned with the default options of MAFFT plugin of Geneious Prime 2020.2 software, and the obtained alignment was adjusted manually in BioEdit 7.2.5 [18]. The individual gene alignments were concatenated into a 24,667 bp matrix.

Phylogenetic reconstruction was performed by the maximum likelihood (ML) analyses with RAxML version 8.2.10 [19], using the GTR+GAMMA model. The number of bootstrap replications equal to 500 was set by the MRE (extended majority rules) criterion Mauve plugin with the default settings implemented in Geneious software was applied for genome-wide comparison of bryophyte mitogenomes.
3. Results
3.1. The General Mitogenome Structure of \textit{P. commune}

\textit{Polytrichum commune} mitogenome (Figure 1) consists of a single circular molecule; its length is 114,831 bp, which exceeds the length of most published bryophyte mitochondrial genomes in the GenBank database, with the exception of \textit{Sphagnum palustre} L. NC024521 (141,276 bp), \textit{Andreaea wangiana} P.C. Chen NC046759, (117,857 bp), and \textit{Atrichum angustatum} NC024520 (115,146 bp). The size of mitogenome of \textit{Pogonatum inflexum} (Lindb.) Sande Lac., which is one of three Polytrichopsida representatives involved in the analysis, is slightly smaller and equal to 113,866 bp, while the mitogenome length of all the other mosses is less than 110 Kb and varies from 109,586 bp (\textit{Funaria hygrometrica} Hedw.) to 100,342 bp (\textit{Mielichhoferia elongata} (Hoppe & Hornsch.) Nees & Hornsch.).

![Figure 1. Mitogenome map of \textit{P. commune}.](image)

The mitogenome of \textit{P. commune} contains 67 genes including genes for three rRNAs (\textit{rrn18}, \textit{rrn26}, and \textit{rrn5}), 24 tRNAs and 40 conserved mitochondrial proteins (15 ribosomal proteins, four \textit{ccm} proteins, nine nicotinamide adenine dinucleotide dehydrogenase subunits, five ATPase subunits, two succinate dehydrogenase subunits, one apocytochrome b, three cytochrome oxidase subunits, and one twin-arginine translocation complex subunit).
However, the gene set of sequenced mitogenomes differs among mosses. For instance, the rpl16 gene is absent in the GenBank annotations of Racomitrium elongatum Ehrh. ex Frissvoll, Codriophorus varius (Mitt.) Bednarek-Ochyra & Ochyra, Racomitrium lanuginosum (Hedw.) Brid., Racomitrium emersum (Müll. Hal.) A. Jaeger, Codriophorus acicularis (Hedw.) P. Beauv., Codriophorus laevigatus (A. Jaeger) Bednarek-Ochyra & Ochyra, Bucklandiella orthotrichacea (Müll. Hal.) Bednarek-Ochyra & Ochyra, and Racomitrium ericoides (Brid.) Brid., but in fact, is present in their genomes. This situation is similar to the rps4 gene in the Oxystegus tenuirostris (Hook. & Taylor) A.J.E. Sm. mitogenome and the rpl5 gene in the Tetraphis pellucida Hedw. mitogenome.

The length of genes, exones, and introns in the mitochondrial genome of *P. commune* is presented in Figure 2. The shortest genes are trnG (UCC), trnC (GCA), and trnQ (UUG), with a length of (71 bp). The first two genes trnC (GCA) and trnG (UCC) have the same length in all other 43 mitogenomes under investigation. At the same time, due to a discrepancy in annotation of the trnQ (UUG) gene its length (and orientation in Oxystegus tenuirostris) differs among Bryophyta mitogenomes: 68 bp in Pogonatum inflexum, 71 bp in Andreaea wangiana, Atrichum angustatum, Bartramia pomiformis Hedw. Buxbaumia aphylla Hedw., Funaria hygrometrica, Mielichhoferia elongata, and Sphagnum palustre, and 72 bp in all other analyzed mosses (Table S1). In all mitogenomes under consideration, the cox1 gene is the largest; the maximum length among them is 14,934 bp in Sphagnum palustre, and 9450 bp in *P. commune*. In other mosses, its length varies from 7081 bp in Tetraplodon fuegianus Besch. to 9443 bp in Pogonatum inflexum. Among exons, exon 3 of the atp9 gene (8 bp) and exon 2 of the nad2 gene (1314 bp) have boundary values in length distribution. Intron length varies from 480 bp (intron 3 of the cox1 gene) up to 2649 bp (intron 3 of the nad5 gene).

![Figure 2](image_url)

**Figure 2.** Distribution of gene, exon, and intron lengths in *P. commune* mitochondrial genome.

Most of the compared mitochondrial genomes do not differ from the one presented herein in the number of exons. Among the analyzed genomes, the cox1 gene contains the largest number of exons: 7 exons in Sphagnum palustre and 5 in all others. Fifty genes of *P. commune* consist of a single exon.

When comparing the structures of the mitogenomes of *P. commune* and other Polytrichopsida species, an inversion of about 4700 bp was revealed in the Pogonatum inflexum (Figure 3). The atp6, rps7, and rps12 genes, as well as the trnE, trnR, and trnG tRNA genes, are located within the inverted genome fragment of *P. inflexum*. It is interesting to note that two forward repeat subunits A7(TA)T8 were identified in the regions flanking the inversion. When comparing 44 bryophyte mitogenomes in the Mauve program, this rearrangement was not discovered in other mosses.
that two forward repeat subunits A7(TA)T8 were identified in the regions flanking the inversion. When comparing 44 bryophyte mito genomes in the Mauve program, this rear-arrangement was not discovered in other mosses.

3.2. Microsatellite Content

In total, 89 microsatellite loci (SSR) were found in the mitogenome of *P. commune* (Table S2) under the following identification criteria: minimal number of repeating units ≥10 for mononucleotides, ≥5 for dinucleotides, ≥4 for trinucleotides, and ≥3 for tetra-, penta-, and hexa-nucleotides. Of these, 13 are located in coding sequences, and 76 in noncoding sequences. The SSR length distribution is shown in Figure 4.

Most microsatellites refer to mono- and di-nucleotides classes (33 and 35 loci, respectively). Penta- and hexa-nucleotides are the least frequent SSR groups in the genome (one locus each). Among all the SSRs, 84.44% are composed only of A/T bases. The total length of the SSR loci is 1021 bp, which comprises approximately 0.89% of the genome length.

3.3. Phylogenetic Analysis

The identity between all the analyzed complete genome sequences was calculated (Figure 5). The mitogenome of *P. commune* is most similar to those of other Polytrichopsida representatives (*Atrichum angustatum* (97.9%) and *Pogonatum inflexum* (93.2%)), and it was
the least similar to *Sphagnum palustre* (65.3%) and *Buxbaumia aphylla* (67.8%). Overall, the least identity was between *Buxbaumia aphylla* and *Sphagnum palustre* (54.3%).

| Group I          | P. commune | A. angustatum | P. inflexum | A. wangi | T. bellica | C. aciculata | C. laevigatus | R. elongatum | R. enoides | C. varius | R. lanuginosum | R. emerum | B. orthotrichaceae |
|------------------|------------|---------------|-------------|----------|------------|--------------|--------------|--------------|-------------|-----------|----------------|-----------|-------------------|
| Group II         |            |               |             |          |            |              |              |              |             |           |                 |          |                   |
| Group III        | O. callistomum | O. diaphanum | O. macrocephalum | N. obtusifolia | N. gymnostoma | O. bicolor | O. speciosum | U. phyllanthera | U. crispa | U. hutchinsiae | S. mirum | S. bunaense | O. stellatum | Z. viridissimus |
| Group IV         | E. rivulare | S. uncinata | C. americanum | H. impolens | A. attenuatus | A. rugelli | P. cygniisetum | B. aphylla | S. palustre |

**Figure 5.** Heatmap of bryophyte whole mitogenome pairwise identity. The color key on the figure shows the percentage of pairwise identity between mitogenomes. Groups formed by species with a high pairwise identity are indicated.

A phylogenetic tree constructed from 33 mitogenome CDS is presented in Figure 6. *Polytrichum commune* is found among other nematodontous mosses, occupying a basal position to two other Polytrichaceae species. The overall topology of the phylogenetic tree corresponds to the topology of a coalescent tree based on 272 protein-coding genes of the nucleus and organelles [7] and reconstructions from 39 mitochondrial protein genes of 157 bryophytes [8].
3.4. Analysis of Repeating Elements

It was suggested based on the analysis of 14 mitogenomes that the complete lack of intergenic repeat sequences, considered to be essential for intragenomic recombinations, likely accounts for the evolutionary stability of moss mitogenomes [10]. Study of the presence of repeats, especially those localized in intergenic spacers, is extremely important for understanding the forces that maintain the highly stable mitogenome structure in mosses. We carried out a search for such repetitive elements in 44 moss mitogenomes using the same search criteria as Liu et al. [10]: nonadjacent repeats longer than 50 bp and having at least 85% of pairwise identity between repeat subunits. It was believed that repeats with such parameters are most efficiently involved in the intragenomic recombination processes [10].

As a result, the presence of intergenic repeats was revealed in 7 out of 44 analyzed species, which is about 16% (Table 1). Intergenic repeats were found in *Sphagnum*, *Tetraphis*, *Polytrichum* commune, and *Funariaidae* (*Physcomitrella patens* and *Funaria hygrometrica*). All members of the Dicranales, Grimmiales, Hypnales, and Orthotrichales completely lack intergenic repeats. The largest number of repeats in intergenic spacers was found in *Sphagnum*. 

---

Figure 6. Maximum likelihood (ML) phylogenetic tree of bryophytes mitogenome sequence with bootstrap support values indicated. Revised species names are in green. *Species with repeats in intergenic spacers. For Bryophytina species, higher rank taxa are indicated.
Table 1. Intergenic repetitive elements in mitogenomes of mosses.

| GenBank Species Name          | GenBank Accession Number | Number of Repeats in Intergenic Spacers | Genome Size, bp | Taxonomical Group      | Mean Genome Size in Taxonomical Group, bp |
|------------------------------|--------------------------|----------------------------------------|-----------------|------------------------|------------------------------------------|
| Sphagnum palustre            | NC_024521                | 1                                      | 141,276         | Sphagnopsida           | 141,276                                  |
| Andreaea wangiana            | NC_046759                | 0                                      | 117,857         | Andreaeopsida          | 117,857                                  |
| Pogonatum inflexum           | MK131350                 | 2                                      | 113,866         |                        |                                          |
| Polytrichum commune          | NC_039775                | 1                                      | 114,831         | Polytrichopsida        | 114,614.3                                |
| Atrichum angustatum          | NC_024520                | 1                                      | 115,146         |                        |                                          |
| Tetraphis pellucida          | NC_024290                | 1                                      | 107,730         | Tetraphidopsida        | 107,730                                  |
| Buxbaumia aphylla            | NC_024518                | 0                                      | 100,725         | Buxbaumiidae          | 100,725                                  |
| Physcomitrella patens        | NC_007945                | 1                                      | 105,340         |                        |                                          |
| Funaria hygrometrica         | NC_024523                | 1                                      | 109,586         | Funariidae            | 107,463                                  |
| Oxystegus tenuirostris       | NC_028040                | 0                                      | 105,001         |                        |                                          |
| Chorisodontium aciphyllum    | NC_050308                | 0                                      | 105,766         | Dicranales             | 105,703.3                                |
| Syntrichia filaris           | NC_027515                | 0                                      | 106,343         |                        |                                          |
| Codriophorus varius          | NC_026891                | 0                                      | 106,358         |                        |                                          |
| Racemitrium ericoides        | NC_026540                | 0                                      | 106,727         |                        |                                          |
| Racomitrium elongatum        | NC_026890                | 0                                      | 106,746         |                        |                                          |
| Racomitrium lanuginosum      | NC_029452                | 0                                      | 106,795         |                        |                                          |
| Codriophorus laevisanus      | NC_025931                | 0                                      | 106,809         |                        |                                          |
| Codriophorus accicularis     | NC_026784                | 0                                      | 106,818         |                        |                                          |
| Racomitrium emersum          | NC_026975                | 0                                      | 107,186         |                        |                                          |
| Bucklandiella orthotrichacea | NC_026974                | 0                                      | 107,215         |                        |                                          |
| Tetraplodon fiegiannus       | NC_028191                | 0                                      | 104,741         | Splachnales            | 104,741                                  |
| Bartramia pomiformis         | NC_024519                | 0                                      | 106,198         | Bartramiales           | 106,198                                  |
| Mielichhoferia elongata      | NC_036945                | 0                                      | 100,342         | Bryales                | 100,342                                  |
| Zygodon viridisinus          | NC_031846                | 0                                      | 103,098         |                        |                                          |
| Orthotrichium stellatum      | NC_024522                | 0                                      | 104,131         | Orthotrichales         | 104,479.9                                |
| Stoneobryum mirum            | NC_034907                | 0                                      | 104,345         |                        |                                          |
Table 1. Cont.

| GenBank Species Name       | GenBank Accession Number | Number of Repeats in Intergenic Spacers | Genome Size, bp | Taxonomical Group | Mean Genome Size in Taxonomical Group, bp |
|----------------------------|--------------------------|----------------------------------------|-----------------|-------------------|------------------------------------------|
| *Stoneobryum bunyaense*    | NC_031392                | 0                                      | 104,352         |                   |                                          |
| *Nyholmiella obtusifolia*  | NC_031767                | 0                                      | 104,603         |                   |                                          |
| *Ulota hutchinsiae*        | NC_024517                | 0                                      | 104,608         |                   |                                          |
| *Orthotrichum macrocephalum* | NC_029355              | 0                                      | 104,624         |                   |                                          |
| *Ulota crispa*             | NC_031393                | 0                                      | 104,636         |                   |                                          |
| *Nyholmiella gymnostoma*   | NC_031391                | 0                                      | 104,638         |                   |                                          |
| *Ulota phyllantha*         | NC_031394                | 0                                      | 104,671         |                   |                                          |
| *Orthotrichum bicolor*     | NC_031389                | 0                                      | 104,736         |                   |                                          |
| *Orthotrichum diaphanum*   | NC_029356                | 0                                      | 104,744         |                   |                                          |
| *Orthotrichum speciosum*   | NC_026121                | 0                                      | 104,747         |                   |                                          |
| *Orthotrichum callistomum* | NC_031390                | 0                                      | 104,785         |                   |                                          |
| *Ptychomnion cygnisetum*   | NC_024514                | 0                                      | 104,480         | Ptychomniales      | 104,480                                  |
| *Hypnum imponens*          | NC_024516                | 0                                      | 103,830         |                   |                                          |
| *Anomodon rugelii*         | NC_016121                | 0                                      | 104,239         |                   |                                          |
| *Anomodon attenuatus*      | NC_021931                | 0                                      | 104,252         |                   |                                          |
| *Brachythecium rivulare*   | NC_016121                | 0                                      | 104,460         |                   |                                          |
| *Sanionia uncinata*        | NC_027974                | 0                                      | 104,497         |                   |                                          |
| *Climacium americanum*     | NC_024515                | 0                                      | 105,048         |                   |                                          |

Thus, our data on a wider sample confirm the conclusion about the small repeat numbers in the intergenic spacers of the mosses mitogenomes.

4. Discussion

Due to the high level of conservation of bryophyte mitogenomes, it is possible to construct a reliable multiple alignment not only for CDS but for complete genome sequences. When analyzing pairwise identity of whole mitogenome sequences on a heatmap, four groups are well distinguished (Figure 5). The first one consists of *P. commune*, *A. angustatum*, and *P. inflexum* from Polytrichopsida (intragroup identity varies from 92.85% to 97.90%). The second group is composed of Grimmiales species (intragroup identity varies from 97.62% to 99.56%). The third group is formed by members of the Orthotrichales order (intragroup identity ranges from 96.22% to 99.98%), and the fourth group includes representatives of Hypnales (intragroup identity ranges from 95.66% to 99.87%). All groups are presented as maximally supported clades on the phylogenetic tree (Figure 6). All Bryopsida representatives are combined in a large group with a minimal pairwise identity of 79.19%. Thus, clustering of bryophytes taxa on pairwise identity of the whole genomes corresponds
to phylogenetic reconstruction only from coding sequences but with the two exceptions explained in detail below.

It is worth noting that along with the basal eperistomate mosses *Sphagnum palustre* and *Andreaea wangiana*, there are two other taxa with surprisingly lower identity to other Bryophyta: *Tetraphis pellucida* and *Buxbaumia aphylla*. The pairwise identity between these and all the other mosses was not higher than 68.81% for *B. aphylla* and 74.33% for *T. pellucida*, respectively. Remarkably, *T. pellucida* possesses the same nematodontous peristome type as other Polytrichopsida species but differs sufficiently from them in the value of the pairwise identity of entire genomes. This finding questions the treatment of these four-toothed mosses within Polytrichopsida [1,3]. Moreover, the joining of *Tetraphis pellucida* with other Polytrichopsida species is not highly supported on the phylogenetic tree. The taxonomical placement of the *Buxbaumia* genus, which has a peculiar peristome structure [5], is not currently well defined and significantly differs according to various estimations [1,3,6–8]. Unfortunately, it was not possible to conclude whether *Tetraphis* or arthrodontous Bryopsida species are significantly closer to *Buxbaumia* based on the performed whole mitogenomes identity analysis. However, *B. aphylla* forms maximally supported clade together with another Bryopsida representatives on the phylogenetic tree.

The application of organellar CDS for phylogeny reconstruction is widely used for plants. However, the introduction in the analysis of complete genome sequences may contribute to the relatedness assessment. Moreover, it is important to involve the noncoding sequences for phylogeny reconstructions, as those are not constrained by selection.

The low structural variability of bryophyte mitochondrial genomes was broadly noted recently. Despite the inversion in the *Pogonatum inflexum* mitogenome detected in the present study, the performed analysis shows a high evolutionary conservatism of the mitogenome structure in mosses. By contrast, it is known that in flowering plants, mitochondrial genomes undergo extensive and high frequency recombination that results in rearrangements and rapid evolution of their general structure. In this case, recombinations occur most often between long (several kb in size) repeats [20,21]. Most angiosperms have several pairs of such repeats in their mitogenomes. For example, in *Silene conica* L., a species with an extremely large mitogenome, the number of large (>1 kb) repeats is 1121, and in the *Silene* species, the relative frequency of recombinant genomes increases with repeat size [22]. At the same time, recombination between the intermediate size repeats (ISRs, usually ranging from 50 to 600 bp in size) is generally infrequent [20].

In the analyzed Bryophyta mitochondrial genomes (Table 1), the abundance of the intergenic repeats is wider than previously indicated [10], but the length of the repeats does not exceed 257 bp. Mitochondrial genomes of liverworts are also characterized by evolutionary stability, yet contain a considerable number of repeats in intergenic spacers. In *Gymnomitrion concinnatum* (Lightf.) Corda, the only liverwort species in which structural rearrangements of the mitogenome were revealed, ten pairs of intergenic repeats were identified, with sizes varying from 107 to 566 bp. The two longest repeats, 566 bp (97.7% sequence identity) and 435 bp length (95.9% sequence identity), were located on junctions between conservative blocks. In each of the six analyzed liverwort species except *Aneura pinguis* (L.) Dumort., two pairs of repeated sequences exceeding 400 bp and 90% of identity were found [9]. Thus, evolutionarily stable liverwort genomes contain only intermediate size repeats but lack long repeats. In chlorophytes and charophytes, which are the groups of algae closest to land plants, the length of repeats in the mitochondrial genome also does not exceed five hundred base pairs [23].

Thus, long repeats exceeding 1 kb were observed only in vascular plants. Hence, it can be concluded that long repeats do have a key influence on mitogenome stability in Viridiplantae, and the greater stability of the mitochondrial genomes of mosses apparently is associated mainly with the absence of long repeats, although recombination events between shorter repetitive sequences can occur with a low frequency. Therefore, the inversion observed in the polytrichopsid moss *Pogonatum inflexum* that is flanked by a short
17 bp repeat demonstrates that even those shorter than 50 bp facilitate rearrangements in mitogenomes.

Thus, the extremely low structural variability and lack of the long repetitive elements in moss mitogenomes is an ancient trait. This statement is in good agreement with the conclusion of Emily Wynn and Alan Christensen, that the first vascular common ancestor of lycophytes, ferns, gymnosperms, and angiosperms acquired new mechanisms of mitochondrial genome replication and repair that led to an expansion of repeats and increasing repeat size and mitochondrial genome length [23].

It is worth noting that distribution of intergenic repeats along the phylogenetic tree of mosses is not random (Figure 6). The repeats are present in mitogenomes of most basal species but absent in the species of terminal clades. A mitogenome size reduction during the evolution of mosses has recently been noted in the analysis of the *Andreaea wangiana* mitogenome [24]. The above-mentioned trends to miniaturization of the mitochondrial genome and loss of intergenic repeats may indicate a strong selection toward further stabilization of the genome structure in mosses.

The revealed inversion within *Pogonatum inflexum* mitogenome is the only currently known rearrangement in bryophytes, and the sequencing of new moss mitogenomes is crucial for an adequate assessment of the level of conservatism in its structure as well as the deep insight into the evolutionary history of bryophytes.

**Supplementary Materials:** The following material is available online at https://www.mdpi.com/1424-2818/13/2/54/s1. Table S1: Length and orientation of *trnQ* (UUG) genes in Bryophyta mitogenomes according their annotations in GenBank. Table S2: Microsatellite loci (SSR) in the mitogenome of *Polytrichum commune*.

**Author Contributions:** Conceptualization, D.V.G., S.V.G., and A.V.T.; methodology, M.D.L. and S.V.G.; validation, V.E.F.; formal analysis, D.V.G., S.V.G., and A.V.T.; investigation, O.I.K., M.D.L., I.A.M., D.V.G., and A.V.E.; resources, A.V.T.; data curation, E.A.S. and M.D.L.; writing—original draft preparation, D.V.G., S.V.G., E.A.S., V.E.F., and A.V.T.; visualization, E.A.S., D.V.G., and A.V.T.; project administration, D.V.G. and A.V.T., funding acquisition, A.V.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** The work was supported by the Russian Foundation for Basic Research: 18-04-00574.

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available in the supplementary material. Sequence of this study is available from GenBank (accession number NC039775). Other publicly available sequences from GenBank (https://www.ncbi.nlm.nih.gov/nuccore) were also analyzed.

**Acknowledgments:** Computational resources of the Makarich HPC cluster were provided by the Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, Russia. We are grateful to Paul Stothard for assistance in genome map preparation, Michael Ignatov for the valuable discussion, and Arthur Zalevsky for help with the software installation.

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

**References**

1. Goffinet, B.; Buck, W.R.; Shaw, A.J. Morphology and classification of the Bryophyta. In *Bryophyte Biology*, 2nd ed.; Goffinet, B., Shaw, J., Eds.; Cambridge University Press: Cambridge, UK, 2009; pp. 55–138, ISBN 978-0-521-87225-6.
2. Bell, N.E.; Hyvönen, J. Phylogeny of the moss class Polytrichopsida (BRYOPHYTA): Generic-level structure and incongruent gene trees. *Mol. Phylogenetics Evol.* **2010**, 55, 381–398. [CrossRef] [PubMed]
3. Goffinet, B.; Buck, W.R. Classification of the Bryophyta. Available online: http://bryology.uconn.edu/classification/ (accessed on 11 November 2020).
4. Cox, C.J.; Goffinet, B.; Wickett, N.J.; Boles, S.B.; Shaw, A.J. Moss diversity: A molecular phylogenetic analysis of genera. *Phytotaxa* **2010**, 9, 175–195. [CrossRef]
5. Ignatov, M.S.; Spirina, U.N.; Kolesnikova, M.A.; Volosnova, L.F.; Polevova, S.V.; Ignatova, E.A. Buxbaumia: A moss peristome without a peristomial formula. *Arctoa* **2018**, 27, 172–202. [CrossRef]
6. Newton, A.E.; Cox, C.J.; Duckett, J.G.; Wheeler, J.A.; Goffinet, B.; Hedderon, T.A.J.; Mishler, B.D. Evolution of the major moss lineages: Phylogenetic analyses based on multiple gene sequences and morphology. *Bryologist* 2000, 103, 187–211. [CrossRef]

7. Liu, Y.; Johnson, M.G.; Cox, C.J.; Medina, R.; Devos, N.; Vanderpoorten, A.; Hedenäs, L.; Bell, N.E.; Shevock, J.R.; Aguero, B.; et al. Resolution of the ordinal phylogeny of mosses using targeted exons from organellar and nuclear genomes. *Nat. Commun.* 2019, 10, 1485. [CrossRef]

8. Bell, D.; Lin, Q.; Gerelle, W.K.; Joya, S.; Chang, Y.; Taylor, Z.N.; Rothfels, C.J.; Larsson, A.; Villarreal, J.C.; Li, F.-W.; et al. Organelomic data sets confirm a cryptic consensus on (unrooted) land-plant relationships and provide new insights into bryophyte molecular evolution. *Am. J. Bot.* 2019, 107, 1–25. [CrossRef]

9. Myszczyński, K.; Górski, P.; Slipiko, M.; Sawicki, J. Sequencing of organellar genomes of Gymnomitron concinnatum (Jungermanniales) revealed the first exception in the structure and gene order of evolutionary stable liverworts mitogenomes. *BMC Plant Biol.* 2018, 18, 321. [CrossRef]

10. 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]

11. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 2014, 30, 2114–2120. [CrossRef]

12. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 2012, 19, 455–477. [CrossRef]

13. Terasawa, K.; Odahara, M.; Kabeya, Y.; Kikugawa, T.; Sekine, Y.; Sato, N. The mitochondrial genome of the moss *Physcomitrella patens* sheds new light on mitochondrial evolution in land plants. *Mol. Biol. Evol.* 2007, 24, 699–709. [CrossRef] [PubMed]

14. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* 1990, 215, 403–410. [CrossRef]

15. Kearse, 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] [PubMed]

16. Grant, J.R.; Stothard, P. The CGView server: A comparative genomics tool for circular genomes. *Nucl. Acids Res.* 2008, 36, W181–W184. [CrossRef] [PubMed]

17. Wang, X.; Lu, P.; Luo, Z. GMATo: A novel tool for the identification and analysis of microsatellites in large genomes. *Bioinformation* 2013, 9, 541–544. [CrossRef]

18. Hall, T.A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* 1999, 41, 95–98. [CrossRef]

19. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014, 30, 1312–1313. [CrossRef]

20. Gualberto, J.M.; Mileshina, D.; Wallet, C.; Niazi, A.K.; Weber-Lotfi, F.; Dietrich, A. The plant mitochondrial genome: Dynamics and maintenance. *Biochimie* 2014, 100, 107–120. [CrossRef]

21. Gualberto, J.M.; Newton, K.J. Plant mitochondrial genomes: Dynamics and mechanisms of mutation. *Annu. Rev. Plant Biol.* 2017, 68, 225–252. [CrossRef]

22. 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]

23. Wynn, E.L.; Christensen, A.C. Repeats of unusual size in plant mitochondrial genomes: Identification, incidence and evolution. *G3 Genes Genomes Genet.* 2019, 9, 549–559. [CrossRef] [PubMed]

24. Huang, W.-Z.; Ma, W.-Z.; Schneider, H.; Yu, Y.; Wu, Y.-H. Mitochondrial genome from *Andreaea wangiana* reveals structural conservatism and a trend of size reduction in mosses. *Bryologist* 2019, 122, 597–606. [CrossRef]