Sequences of rDNA internal transcribed spacers from the chloroplast DNA of 26 bryophytes: properties and phylogenetic utility

Tagir H. Samigullin[a,*], Karmen M. Valiejo-Roman[a], Aleksey V. Troitsky[a], Vera K. Bobrova[a], Vladimir R. Filin[b], William Martin[c], Andrey S. Antonov[a]

[a] *A.N. Belozersky Institute of Physicochemical Biology, Moscow State University, Moscow 119899, Russia*
[b] *Biological Faculty of Moscow State University, Moscow 119899, Russia*
[c] *Institut für Genetik, Technische Universität Braunschweig, Spielmannstrasse 7, D-38023 Braunschweig, Germany*

Received 5 December 1997

Abstract We determined the sequence of the region of the chloroplast DNA inverted repeat spanning from the 3′-terminus of the 23S rRNA gene to the 5′-terminus of the tRNA^{Arg}(ACG) gene (about 700 bp) from 25 bryophytes and from the charophycean alga *Chara australis*. Phylogenetic analysis of these sequences using the neighbor-joining method suggests an early dichotomy of bryophytes and their paraphyly relative to the tracheophyte lineage. A monophyly of liverworts (Marchantiidae plus Jungermanniidae), a deep divergence of Metzgeriales among Jungermanniidae and a close affinity of the two subclasses of mosses, Sphagnidae and Andreaeidae, are evident. The branching pattern observed is consistent with the phylogenetic distribution of several prominent indels observed in the alignment.

Key words: Chloroplast; Ribosomal RNA; Internal transcribed spacer; Phylogeny; Bryophyta

1. Introduction

The early phases of the colonization of land by plants was accompanied by a rapid diversification of all main groups within a relatively short window of time from about 480 to 360 million years before the present (see [1] for a recent review). Among contemporary flora, the bryophytes (mosses, liverworts and hornworts) belong to some of the most primitive descendants of this early diversification process.

In spite of a growing accumulation of molecular data (nuclear and chloroplast-encoded rRNAs, mitochondrial *coxIII* and chloroplast *rbc*L sequences) as well as progress in the interpretation of paleobotanical data and new fossil discoveries, the precise relationships among basal lineages of land plants are still not clear [1–9].

In previous work, we have investigated the properties of non-coding sequences from the inverted repeat of chloroplast DNA (cpDNA) for the study of the evolutionary process in vascular plants. Our efforts focused on the internal transcribed spacer regions of cpDNA between the 23S rRNA, 4.5S rRNA and 5S rRNA genes [10,11]. These regions, designated cpITS2 and cpITS3, constitute about 500 bp of data from most species and can easily be amplified and sequenced from a broad spectrum of land plants with primers designed against conserved regions from the adjacent structural rRNA genes.

*Corresponding author. Fax: (7) (095) 9393181.
E-mail: tahir@bioevol.genebee.msu.su

Abbreviations: cpITS, chloroplast internal transcribed spacer; cpDNA, chloroplast DNA

In most chloroplast genomes studied to date a tRNA gene, tRNA^{Arg}(ACG), is located downstream of the 5S rRNA gene and is separated from the latter by a relatively short region of non-coding DNA about 200 nucleotides in length which is cotranscribed with the rest of the rRNA operon at least in some plants [12,13]. This region, designated here cpITS4 (chloroplast 5S rRNA-tRNA^{Arg}(ACG) spacer) represents a contiguous extension of the cpITS2 and cpITS3 regions.

Here we describe the amplification and sequencing of cpITS2, cpITS3, and cpITS4 from 25 bryophytes and the charophyte *Chara australis*, and an investigation of the phylogenetic utility of the region for addressing bryophyte evolution.

2. Materials and methods

Plant samples were either collected in the vicinity of Moscow, or taken from the Herbarium of the Department of Botany of Moscow State University. The species studied are: *Andreaea rupestris, Aneura pinguis, Atrichum undulatum, Blepharostoma trichophyllum, Buxbaumia aphylla, Calypogeia integristipula, Cephalozia bicuspidata, Ceratodon purpureus, Chiloscyphus polyanthos, Climacium dendroides, Homalia trichomanoides, Hylocomium splendens, Lophocolea heterophylla, Orthotrichum speciosum, Pellia neesiana, Plagiochila porelloides, Pleurozium schreberi, Ptilidium pulcherrimum, Racomitrium microcarpum, Rhodobryum roseum, Rhytidiadelphus triquetrus, Riccia fluitans, Schistostega pennata, Sphagnum palustre, Tetraphis pellucida. Psilotum triquetrum (syn. nudum)* was obtained from the Botanical Gardens of the Russian Academy of Sciences. *Chara australis* was obtained from the Department of Biophysics of Moscow State University where it is cultivated. The sequences are deposited in the GenBank database under the accessions numbers AF033624–AF033633, AF033635–AF033649, AF033651 and AF033652.

Sequence data for *Marchantia polymorpha* (the whole region studied) and *Psilotum nudum* (cpITS2 and cpITS3) were taken from the GenBank database (accession numbers X04465 and L41569).

Primers used to amplify and sequence cpITS2 and cpITS3 were those described earlier [10]. For the cpITS4 region, 5′-GATATTCTGGTGTCCTAGGCGTAG-3′ (cpITS4F) and 5′-CGTAGCCACGTGCTCTAATCCTC-3′ (cpITS4R) primers were used. PCR reaction mixtures contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2 mM $MgCl_2$, 200 μM of each dNTP, 0.4 μM of each primer, and 0.5 units *Taq* polymerase. The reactions were performed for 30 cycles under a regime of 50 s at 94°C, 40 s at 58°C, 1 min at 72°C.

Sequences were aligned manually using the VOSTORG package [14]. Phylogenetic trees were constructed by the neighbor-joining method with 100 bootstrap resamplings employing the TREECON package [15]. For outgroup comparison, the sequence of *Chara australis* was used.

3. Results and discussion

We determined sequences of the part of the chloroplast DNA inverted repeat region from the 3′-terminus of the 23S rRNA gene to five nucleotides upstream of the 5′-terminus of

0014-5793/98/$19.00
PII S0014-5793(97)01600-1

```
                1         11        21        31        41        51        61        71        81        91
Plagiochila     CTTGTTTGT----ACAT---GATCCGATAAAA-TA----------ATCAGATAATAATAA---CCTCTAGTCATTGTTCAA-------TTGTT--CAACG
Lophocolea      CTTGTTCGT----ACAT---GATCTGATAAAA-TA----------ATCAGATAATAATAA---CCTCTAGTCATTGTTCAA-----------------CG
Chiloscyphus    CTTGTTTGT----ACAT---GGTCTGAGAAAAACA----------ACCAGATAATAATCA---CCTCTGGTCATTGTTCAATGTTCAATTGTT--CGATG
Blepharostoma   CTTGTTTGT----ACAT---GATCCGATAAAAACA----------ATCAGACAATTATAAGAACCTCTGGTTATTGCTCAA--------TATT--CAACA
Calypogeia      CTTGTTTGT----ATAT---GGTCCGATAAAAACG----------ATCAGATAATAA---CCACCTCTGGTCATTGTTCAATGTTCAATTGTT--CGACG
Ptilidium       CTTGTTTCT----ACAT---GGTCTGATAAAAACA----------ATCAGATACTAG---CCACCTCTGTTCATTGTTCAATGTTCAATTGTT--CGACG
Cephalozia      CTTGTTCGT----AC-------TCTGAGAAAAACA---------------------------------------------------AATTGTT--TGATT
Aneura          CTTGCTCCT----ACAT---GGCCTGATAAAAGCA----------ATCTGATACTAGATA---CCTCTGTTCATTGTTCAA--------TGTT--CGGCG
Pellia          CTTATTTCC----ACAT---GGCC----AAA-GCA----------ATCAGGTACTAG------CCC-----------------------TGTT--TGGTA
Riccia          CTTGTTCCG----CTAT---GACCTGATGA----------CTTTTATCAGTTA-TAG---CCACCAACTTTCATTGTTCAA-------TTGTT--TGACA
Marchantia      CTTGTTCCG----CCAT---GACCTGATAAAAGTA----------ATCAGGTA-TAG---CCACCAACTTTCATTGTTCAA-------TTGTT--TGACA
Sphagnum        CTTGTTCCT----ATAT---GACCCGAT----------TGTTTTTATCAGGCATTAG---CCACCTATCTTCATCGTTCGG-------TTGTT--TTACA
Andreaea        CTTGTTTCT----ATAT---GACCTGAT----------TGTTTCTATCAGGC---AG---CCACCTATTTTCATTGTTCAG-------TTGTT--TGACA
Buxbaumia       CTTGTTGCT----ATAT---GACCTGAT----------TGCTTTTATCAGGC---AG---CCACTTATTTTTATCGTTCRA-------TTATT--TGACA
Atrichum        CTTGTTCCT----ATATTATGACCTGAT---------TTGTTTTTATCAGGC---AG---CCACCTATTTTCATCGTTCAG-------TTGTT--CGACA
Tetraphis       CTTGTTCCT----ATAT---AACCTGATAAAACCAATCAGGTTTTATCAGGT---AG---CTACCTATTTTCATCGTTCAG-------TTGTT--TTACA
Ceratodon       CTTGTTCCT----ATAT---GACCTGAT----------TGTTTTTATCAGGC---AA---CCATCTATTTTCATCGTTTAG-------TTGTT--TGATA
Racomitrium     CTTGTTCCT----ATAT---GACCTGAT----------TGTTTTTATCAGGC---AA---CCATCTATTTTCATCGTTTAG-------TCGTT--TAATA
Schistostega    CTTGTTCCT----ATAT---GACCTGAT----------TGTTTTTATCAGGC---AA---CTATCTATTTTCATCATTTAG-------TTATT--TGACA
Rhodobryum      CTTGTTCCT----ATAT---GACCTGAT----------TGTTTCTATCAGGT---AG---CCACCTATTTTCATCATTCAG-------TTGTT--TCACA
Orthotrichum    CTTGTTCCT----ATAT---GACCTGAT----------TGTTTTTATCAGGC---AG---CTATCTATTTTCATCGTTTAG-------TTGTT--TAAAA
Homalia         CTTGTTCCT----ATAT---GACCTGAT----------TGTTTTTATCAGGC---AG---CCACCTATTTTCATCGTTCAG-------TTGTT--TCACA
Hylocomium      CTTGTTCCT----ATAT---GACCTGAT----------TGTTTTTATCAGGC---AG---CCACCTATTTTCATCGTTTAG-------TTGTT--TCACA
Pleurozium      CTTGTTCCT----ATAT---GACCTGAT----------TGTTTTTATCAGGC---AG---CCACCTATTTTCATCGTTTAG-------TTGTT--TCACA
Rhytidiadelphus CTTGTTTCT----ATAT---GACCTGAT----------TGTTTTTATCAGGC---AG---CCACCTATTTTCATCGTTTAG-------TTGTT--TCACA
Climacium       CTTGTTCCT----ATAT---GACCTGAT----------TGTTTTTATCAGGC---AG---CCACCTATTTTCATCGTTTAG-------TTGTT--TCACA
Psilotum        CTTGTTCCTTCCTACAG---GAACTGA-GAAT-----TTGGTT-----AGGT---AG---CC-------------GTTCGA-------TTGCT--TGACA
Chara           CTT-TTCCT----A--------CCT-------------------AATCA---AAT-------------------CG---AA--------TGAAAATGA-A

                101       111       121       131       141       151       161       171       181       191
Plagiochila     ACATT-AACAT---------ACCCTG-----CCTCCTTCT-----------------------------------AGAA---GGAGGCAGGGTAT-AAGAG
Lophocolea      ACATT-AACAT---------GCTCTGCCCCCTCCCCTTCT---------------------------AAGGGG--------AGGGGGCAGGGCAT-AAGAG
Chiloscyphus    ATATT-AACCT---------GCCCTG------CTCCTTCTT-----------------------------------GAA------GCAGGGCAT-AAGAG
Blepharostoma   ACATT-AACCT---------ACCCTGCCCC-CCTCTTTCT------------------------------------------GGGGGCAGGGTAT-AAGA-
Calypogeia      ACATT-AACCT---------GCCCTGCCC-----CT---------------------------------------AGAA--AGGGGGCGGGGTAT-AAGAG
Ptilidium       ACATC-AACCTAACAGTTTAGCCCTGCCC-------------------------------------------ACGAGAA----GGGCAGGGCAT-GAGAG
Cephalozia      ATGTT-AACCT-------NNGCCCTG------CTCT--------------------------------------AAGAA-----------ACATCAANNG
Aneura          ATACCGAA-TT-ACGGTT-AGCCCCG-------TCCTTCTT-----------------------------------------GTGGGCGGGGCAT-AGGAG
Pellia          AC--------T--------------------------------------------------------------CGAGAA-GGCTGGGCAGG---------
Riccia          ACATA-AACCTAACAACTTGACCCTG------CTCTTATTTT------------------------------------------GGGCAGGGTTTC-AAAG
Marchantia      ACATA-AACCTAACAACTTTACCCTG------CTCTTATTTT------------------------------------------GGGCAGGGTTTC-AAAG
Sphagnum        ACATC-AACCTAATAG------------------------------------------------------------------------------------
Andreaea        ACATC-AACTTAACAG------------------------------------------------------------------------------------
Buxbaumia       ACATC-AACCTAACAA------------------------------------------------------------------------------------
Atrichum        ACATC-AACTTAAAAG------------------------------------------------------------------------------------
Tetraphis       ACATC-AACTTAACAG------------------------------------------------------------------------------------
Ceratodon       ATACA-AACTTAACAG------------------------------------------------------------------------------------
Racomitrium     ATAGA-AACTTAACAG------------------------------------------------------------------------------------
Schistostega    ATACA-AACTTAACAG------------------------------------------------------------------------------------
Rhodobryum      ACATC-AACCTAACAG------------------------------------------------------------------------------------
Orthotrichum    ACATC-AACCTAACAG------------------------------------------------------------------------------------
Homalia         ACATC-AACCTAAAAG------------------------------------------------------------------------------------
Hylocomium      ACATC-AACCTAACAG------------------------------------------------------------------------------------
Pleurozium      ACATC-AACCTAACAG------------------------------------------------------------------------------------
Rhytidiadelphus ACATC-AACCTAACAG------------------------------------------------------------------------------------
Climacium       ACATC-AACCTAACAG------------------------------------------------------------------------------------
Psilotum        ATATC-AACCTAAGAGCTTCGCCCTGCC---TCTCTGTCCATC------------------------AAGGGATGGGA-TGGAAGGGCAGG---TC-GGAG
Chara           ATT-----CC---GTGCT-AGTGTTCCT---ACTCTTTCTTACCTGGCCTCGTCGTTATCCTTTTT------------------AGCAGAA--TCAAAAG
```

Fig. 1. Alignement of the cpITS3 region from 26 bryophytes, *Psilotum* and *Chara*. For species names, see Section 2.

the $tRNA^{Arg}$(ACG) gene from 25 bryophytes and the alga *Chara australis*. The region studied contains chloroplast DNA spacers cpITS2, cpITS3 and cpITS4, as well as the coding sequences for the 4.5S and 5S rRNAs. As in our earlier survey of non-coding regions from vascular plants [10,11], we observed here that the evolution of cpDNA spacers in bryophytes was accompanied by gross changes in their length, visible as numerous indels in the sequence alignment. Among bryophytes, the length of cpITS2, cpITS3, and cpITS4 varies from 62 to 110, 130 to 230, and 101 to 238 bp, respectively. The cpITS regions of *Chara australis* are considerably shorter and comprise only 52, 132, and 69 bp for cpITS2, cpITS3 and cpITS4, respectively.

The alignment of the regions sequenced contains 22 3′-terminal nucleotides of the 23S rRNA gene, cpITS2, the 4.5S rRNA gene, cpITS3, the 5S rRNA gene and cpITS4 without five nucleotides adjacent to the $tRNA^{Arg}$(ACG) gene. The full length of the alignment is 1131 positions. The alignment for the cpITS3 region for taxa studied shown in Fig. 1 exemplifies the difficulty of unambiguously aligning portions of these non-coding sequences across higher bryophyte taxa.

Three prevalent types of insertion events can be distinguished in these non-coding regions from the chloroplast DNA internal repeat: (i) tandem duplications of 2–4 bp stretches which arise apparently due to replication slippage mechanisms [16,17], (ii) elongation of preexisting homopolymeric stretches, and (iii) insertions of oligonucleotide stretches with no recognizable similarity to adjoining sequences. Nucleotide substitutions superimposed upon these insertion-deletion processes result in a complex pattern of ITS sequence

```
                 201       211       221       231       241       251       261       271       281       291
Plagiochila      G---------TTCCTTCTTCC----GGGAC------T---AAT---------------------------------------------------------
Lophocolea       G--------TTTGTTTCTTCT------GACACTTATT---AAT------------------------------------------TC-AG----------
Chiloscyphus     G---------CTCCTTCTTTC----GGGACACTTATT-------------CGG--------------------------------ACTAG----------
Blepharostoma    GAA--------TCCTTCTTTC----GGGACAC--ATTATTAATATTTATTCGG-----------------------------------------------
Calypogeia       G-------------TTCTTTC----GGGACACTTTTT-------------------------------------------CAA--ACTAG----------
Ptilidium        G--G----------TCATTTC----GGGACACTTATC-------------CGGA--------------------------CAATTACTAG----------
Cephalozia       GAAG---------CTTAC-------GGGAGCTTCTT--------------TGG--------------------------------ACTAA----------
Aneura           G---------TTCCT------GGAGAGGGTACTTCC--------------CGC--------------------------------ACCAG----------
Pellia           ----------CTCC-------GGAAGG-ACACTGNNN--CAAT-------TGCT----------------------------ATTACTAG----------
Riccia           G--------CTTTTTTGCCT-GGAAGGGACACTTCT---------------AGTGCCCTNTTCCAGAATG-AAAGCCTCACAATTACTCG----------
Marchantia       GGG------TTTTTTT-CCT-GGAAGGGACACTTCT---------------AGTGCCCT-TTCCAGAATG-AAAGACTCACAATTACTTG----------
Sphagnum         ------------CTCTGCCT--GAAGGTATACCTAT---------------GGTGTA-C-TTCCGGAATG-AGAGCCTCACAATTACTAG----------
Andreaea         ------------CTCTGCCT-GGAAGGTACACC-ATA---AATA-------GGTGTACC-TTCCAGAATT-AAAGCCTGACAATAATTAG----------
Buxbaumia        ----------CTCTCTGCCT-GGAAGGTACATCTATA---TAT--------GATGTACC-TTCCAGAATT-AGAGCCTCATAATTACTAG--CCAA----
Atrichum         ----------CTCTCTACCT-GGAAGATACATC-ATA---------------GTGTATC-TTCCAGAATT-AGAGCCTCACAATCACTA-TA--------
Tetraphis        ----------CTCTCTGCCT-GGAAGGTACATC-ATA---TATA-------GGTGTATC-TTCCAGAATT-AGAGGCTCACAACAACTAG----------
Ceratodon        ----------CTCTCTACCT-GGAAGGTACATC-ATA---AATA-------GATGTATC-TTCCAGAATT-AGAGCCTCATAATCATTAGCACTAA----
Racomitrium      ----------CTCTCTGCCT-GGAATATATATCTATT---TAT--------GATGTATC-TTCCAGAATT-AGAGCCTCA------TTAGCACTAA----
Schistostega     ----------CTCTAATTCT-GGAAGGTACATC-ATA---AATA-------GATGTACC-TTCCAGAATT-AGAGCCTCACAATCACTAGCACTAG----
Rhodobryum       ----------TTCTAATTCT-GGAAGATACATCTA------------------TGTATC-TTCCAGAATT-AGAGCCTCACAATCATTAGCACTAG----
Orthotrichum     ----------TTCTCTGCCT-GGAAGATACATC-ACA---AATA-------GATGTATC-TTCCAGAATT-AGAGCCTGA------CTAGCATTAG----
Homalia          ----------TTCTCTGCCT-GGAAGATACATC-ATA---AATA-------GATTTATC-TTCCAGAATT-AGAGCCTCACAATAACTAGCATTAG----
Hylocomium       ----------TTCTCTGCCT-GGAAGATACATC-ATA---AATA-------GATTTATC-TTCCAGAATT-AGAGCCTCACAATAACTAGCACTAG----
Pleurozium       ----------TTCTCTGCCT-GGAAGATACATC-ATA---AATA-------GATTTATC-TTCCAGAATT-AGAGCCTCACAATAACTAGCACTAG----
Rhytidiadelphus  ----------TTCTCTGCCT-GGAAGATACATC-ATA---AATA-------GATTTATC-TTCCAGAATT-AGAGCCTCACAATAACTAGCACTAGCACT
Climacium        ----------TTCTCTGCCT-GGAAGATACATC-ATA---AATA-------GATTTATC-TTCCAGAATT-AGAGCCTCACAATAGCTAGCACTAG----
Psilotum         GAAGG----TTCTCTGTCT-GGAAGGTGCACC-A----------------GAGG--CC-TTCCAGAATCGGAAGCCTCACAAGAACTAG----------
Chara            -AAGAAAAA-------------------------------------------------------AGAATA-AGAGG------------------------

                 301       311       321       331       341       351       361       371
Plagiochila      -------------TTTTATG---TTTTTT-CCC--ATG-----------------CCTTTTTCCGTTCATGGGTTGA
Lophocolea       -------------ACTTATG---TTTCTCTCTCT-CT------------------CTTTCTTCCATTTATGGATTGA
Chiloscyphus     ------------TTCTTATG---TTTTTT-CCC--ATG-----------------CTTTTTTCCGTTTATGGGTTGA
Blepharostoma    -------------ACTTAT----TTTT---CC-----G--TTTATTTATGGGTT----------GTTTATGGGTTGA
Calypogeia       ------------TTTTTATG---TTTTTCTCCC--ATG-----------------CCTCTTTCCGTTCATGGGTTGA
Ptilidium        ------------TTTTTATG---TTTT-CTCCC--ATG-----------------CCTCTCTCCGTTTATGGGTTGA
Cephalozia       --------------TTTATA---TTCCTTTCCC--CCG--TTT---------------------GTT---GTGTTGA
Aneura           -----------------------TTTTT--CC-----G--TTTCTTTATCCACGACTCTTCCT-ATCTATGGGTTGA
Pellia           -----------------------TTT----CT---ATG-----------------CTCTCCCT--TTCCCGGGTTGA
Riccia           -----TTTTTGTCTANTATTATACCTTT--CT---ATG-----------------CTCTTC----------GGCTGA
Marchantia       -------GTTTTTTTTTATTATACTTTT--CT---TTG---------------------------TTCATGGGTTGA
Sphagnum         --CCA------------ATATAGCTTTT--CCCCTATGCTTTCGTTTATGGG---TTCTTT---GTTTATGGGTTGA
Andreaea         --CCA------------ATA-TGCTTTT--CCCTTATG-----------------CTCTCTTC-GTTGATGGGTTGA
Buxbaumia        --TCG------------ATA-TACTTTT--CTCTTATG-----------------TTCTCCTC-ATTCATGGGTTGA
Atrichum         ---CA------------ATA-TGCTTTTT-CC-TTATG-----------------CTTTCTTC-ATTCATGGGTTGA
Tetraphis        --CCG------------ATA-TGCTTTTT-CCCTTATG-----------------CTCTCTTA-GTTCATGGGTTGA
Ceratodon        --CCG------------ATATTACTTTTT-CT-TTCTG-----------------TTTTTTTT-GTTTATGGATTGA
Racomitrium      --CCG------------ATATTACTTTTT-CC-TTATG-----------------TTTTTTTT-GTTTATGGATTGA
Schistostega     --CCG------------ATA-TACTTTTT-CC-TTATG-----------------TTTTTTTT-GTTTATGGATTGA
Rhodobryum       --CCG------------ATACTACTTTTT-CC-TTATG-----------------TTTTTC----TTTATGGGTTGA
Orthotrichum     --CCG------------ATACTACTTTTT-CC-TTATG-----------------TTTTT-----TTTATGGGTTGA
Homalia          --CCG------------ATACTACTTTTT-CC-TTATG-----------------TTTTT-----TTTATGGGTTGA
Hylocomium       --TCG------------ATACTACTTTTT-CC-TTATA-----------------TTTTTTTT--TTTATGGGTTGA
Pleurozium       --TCG------------ATACTACTTTTT-CC-TTATA-----------------TTTTT-----TTTATGGGTTGA
Rhytidiadelphus  AGTCG------------ATACTACTTTTT-CC-TTATA-----------------TTTTT-----TTTATGGGTTGA
Climacium        --TCG------------ATACTACTTTTT-CC-TTATA-----------------TTTT------TTTATGGGTTGA
Psilotum         --CCA------------AT----CTTTT---CTTTCTGC----------------TTCTTTCTCGTCCATAGGTTGA
Chara            ------------------------------------------------------------------CCATCTGTTGA
```

Fig. 1 (*continued*).

evolution which renders the unambiguous alignment of portions of the non-coding regions difficult. The complete alignment of the region from the 3′-end of the 23S gene to the 5′-end of the $tRNA^{Arg}$(ACG) gene for all species investigated is available from the authors on request.

Fig. 2 shows a phylogenetic tree, constructed by the neighbor-joining method using a distance matrix based on Kimura's two parameter evolution model [18]. For distance matrix estimation gaps were counted as single nucleotide changes regardless of their length as in [15].

Despite the possible ambiguities in the alignment mentioned above, the tree topology is quite stable with respect to different versions of alignment; the only difference concerns the relative positions of Andreaea and Sphagnum – they do not form a single cluster as in Fig. 2 but branch off one after another, with Andreaea forming a sister lineage to Bryidae, and Sphagnum occupying the most basal position among the Bryopsida. According to some traditional views, Sphagnidae and Andreaeidae represent two subclasses of mosses with no close connection either to the third subclass, Bryidae, or to each other [19]. The overall tree topology, when Kimura's distance with gaps not taken into account was employed, or when *p*-distances with or without gaps were calculated, did not change. These options may result in uniting *Atrichum* and *Buxbaumia* into a single cluster or in adding *Cephalozia* to the *Pellia* plus *Aneura* clade, however, with a low bootstrap support (22–37%). In these cases the mean bootstrap value was somewhat lower.

The main divisions of the tree are in agreement with the separation of the Bryophyta into three main groups accepted

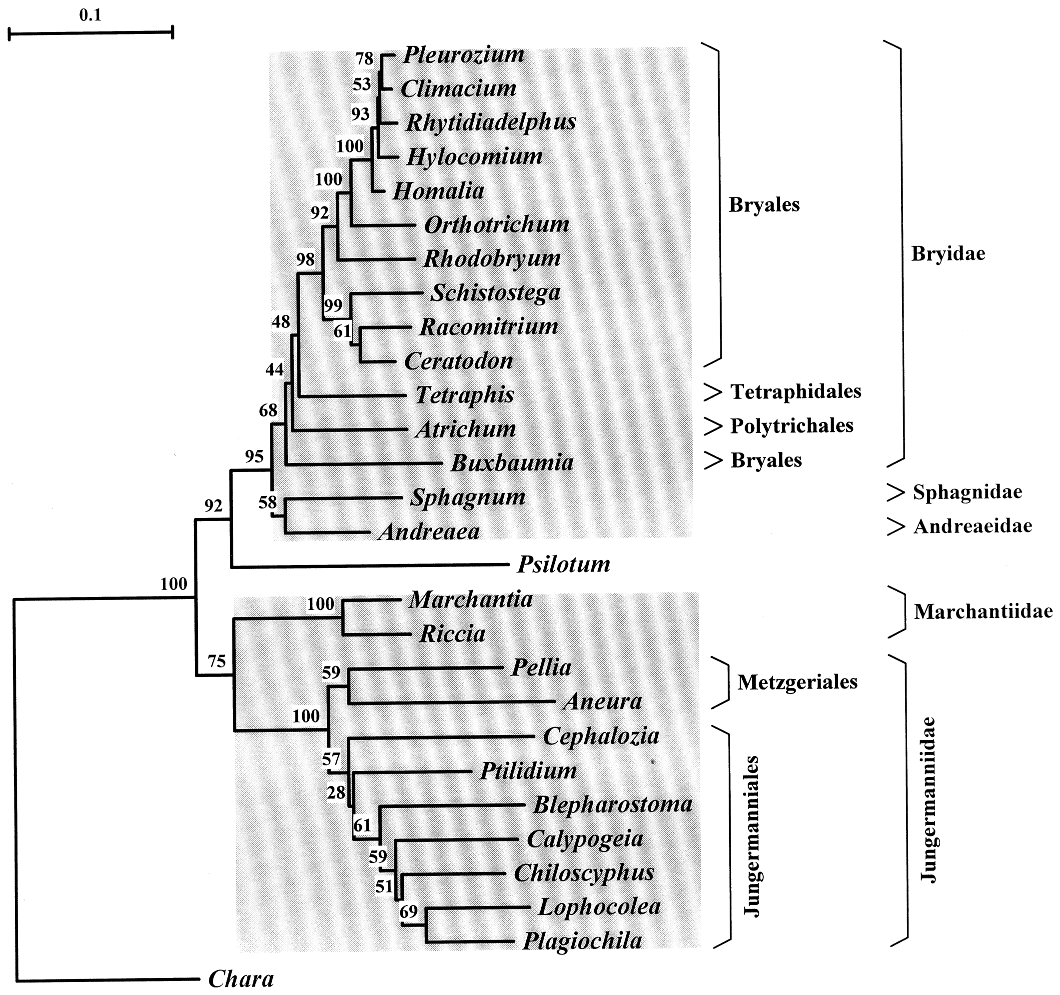

Fig. 2. Bryophyte phylogeny inferred from cpITS1–cpITS4 region sequences. Neighbor-joining tree based on Kimura's distances with gaps taken into account is presented. The scale bar indicates Kimura's distance. Bootstrap values from 100 resamplings are given at the nodes of the tree. For species names, see Section 2. Bryopsida and Jungermanniopsida classes are shadowed.

by most modern bryologists [20] and referred to here as classes, i.e. Bryopsida (Sphagnidae, Andreaeidae and Bryidae), Jungermanniopsida (Marchantiidae and Jungermanniidae), and Anthocerotopsida (not presented in the current analysis). Within Jungermanniopsida, two clusters supported in all bootstrap samples are observed. Again, this dichotomy correlates well with the separation between liverworts with reduced sporogonium (Marchantiidae) and well-developed sporogonium (Jungermanniidae). The observed monophyly of Jungermanniopsida is in agreement with traditional systematic views [1,20], but does not coincide completely with molecular data on *rbc*L [6] and 18S rRNA [9]. According to [9], the most basal dichotomy of bryophytes occurred between Marchantiidae and all other groups of bryophytes.

Within Jungermanniidae the first dichotomy occurs between Metzgeriales and Jungermanniales. Again, it is in agreement with a hypothesis according to which the Metzgeriales is the oldest group among liverworts [21].

There exist uncertainties and/or controversies about the results of phylogenetic analyses based on different macromolecules with respect to Bryophyta as a monophyletic taxon or a paraphyletic group relative to vascular plants with mosses or liverworts being the sister group of tracheophytes [4–9,22]. Our chloroplast ITS data posit that Bryophyta is an artificial paraphyletic group because Jungermanniopsida (Marchantiidae and Jungermanniidae) form one clade, and Bryopsida together with vascular plants (presented in our analysis by Psilotum) comprise another.

The existence of these main clades of bryophytes is strongly supported by analysis of indels within these groups. The most prominent of them are (positions are numbered from the beginning of the whole alignment): Jungermanniidae – ITS2 97–135 deletion, ITS3 504–528 deletion, ITS4 1001–1006 deletion; Jungermanniales – ITS2 40–44 deletion, ITS3 471–474 insertion, ITS4 860–876 deletion; Jungermanniopsida – ITS3 indel pattern in positions 366–459, ITS4 1007–1035 deletion; Bryopsida – ITS4 insertion pattern in the region 780–790, 1036–1069 deletion.

The predicted secondary structures of bryophyte ITS sequences (data not shown) correspond only roughly to secondary structure models of cpITS2–4 previously proposed for vascular plants [13,23,24]. This indicates a high degree of structural flexibility of these regions with regard to pre-rRNA processing. A more detailed analysis of secondary structure evolution as well as relationships between bryophyte orders based on a broader set of species will be presented elsewhere.

Notably, in *Chara australis* chloroplast DNA there is a nucleotide sequence highly homologous to 4.5S rRNA and separated from a 3′-end part of 23S rDNA by insertion of an alien sequence. The presence of 4.5S rRNA, which originated from the 3′-terminal region of 23S rRNA due to insertion of the ITS2 sequence [25], has been shown for all land plants studied [26]. The occurrence of 4.5S rDNA and the corresponding transcript was discovered initially in another species of *Chara* (V.V. Goremykin, personal communication). It is absent in the chloroplast genome of the unicellular green alga *Chlorella* [27]. Earlier we failed to find the 4.5S rRNA among low molecular weight RNAs from a multicellular green alga *Cladophora* sp. as well [28].

The presence of cpITS2 in *Chara* is consistent with the view of the origin of land plants from a charophycean ancestor supported by molecular and morphological analyses [5,22,29]. In future studies, it will be of interest to more precisely circumscribe the phylogenetic distribution of the 4.5S rRNA, and examine its utility for resolving filiation processes during the earliest stages of land plant evolution.

Acknowledgements: This work was supported by Grants 436 RUS 17/78/96 (S) and Ma 1426/1-3 from the Deutsche Forschungsgemeinschaft to W.M. and by Grants N 96-04-0012G and 96-04-50932 from the Russian Foundation of Basic Research to A.A.

References

[1] Kendrick, P. and Crane, P.R. (1997) Nature 389, 33–39.
[2] Mishler, B.D. and Churchill, S.P. (1985) Cladistics 1, 305–328.
[3] Mishler, B.D. (1986) J. Bryol. 14, 71–81.
[4] Hiesel, R., von Haeseler, A. and Brennicke, A. (1994) Proc. Natl. Acad. Sci. USA 91, 634–638.
[5] Mishler, B.D., Lewis, L.A., Buchheim, M.A., Renzaglia, K.S., Garbary, D.J., Delwiche, C.F., Zechman, F.W., Kantz, T.S. and Chapman, R.L. (1994) Ann. Missouri Bot. Gard. 81, 451–483.
[6] Lewis, L.A., Mishler, B.D. and Vilgalys, R. (1997) Mol. Phyl. Evol. 7, 377–393.
[7] Malek, O., Lattig, K., Hiesel, R., Brennicke, A. and Knoop, V. (1996) EMBO J. 15, 1403–1411.
[8] Manhart, J.R. (1994) Mol. Phyl. Evol. 3, 114–127.
[9] Bopp, M. and Capesius, I. (1996) Bot. Acta 109, 368–372.
[10] Goremykin, V., Bobrova, V., Pahnke, J., Troitsky, A., Antonov, A. and Martin, W. (1996) Mol. Biol. Evol. 13, 383–396.
[11] Pahnke, J., Goremykin, V., Bobrova, V., Troitsky, A., Antonov, A. and Martin, W. (1996) in: Pteridology in Perspective (Camus, J.M., Gibby, M., Johns, R.J., Eds.), pp. 217–230, Royal Botanical Gardens, Kew.
[12] Dormann-Przybyl, D., Strimatter, G. and Kossel, H. (1986) Plant Mol. Biol. 7, 419–431.
[13] Leal-Klevezas, D.S. (1991) Thesis, The University of Guelph, Ontario.
[14] Zharkikh, A.A., Rzhetsky, A.Yu., Morozov, P.S., Sitnikova, T.L., Krushkal, J.S. and Matushkin, Yu.G. (1990) in: Modelling and Computer Methods in Molecular Biology and Genetics, pp. 217–218, Novosibirsk.
[15] Van de Peer, Y. and De Wachter, R. (1994) Comput. Applic. Biosci. 10, 569–570.
[16] Wolfson, R., Higgins, K.G. and Sears, B.B. (1991) Mol. Biol. Evol. 8, 709–720.
[17] Vogler, A.P., Welsh, A. and Hancock, J.M. (1997) Mol. Biol. Evol. 14, 16–19.
[18] Kimura, M. (1980) J. Mol. Evol. 16, 111–120.
[19] Vitt, D.H. (1984) in: New Manual of Bryology (Schuster R.M., Ed.), Vol. 2, pp. 697–759, Hattory Bot. Lab., Nichinan, Japan.
[20] New Manual of Bryology (Schuster R.M., Ed.) (1983, 1984), Vol. 1–2, Hattory Bot. Lab., Nichinan, Japan.
[21] Renzaglia, K.S., Brown, R.C., Lemmon, B.E., Duckett, J.G. and Ligrone, R. (1994) Can. J. Bot. 72, 65–72.
[22] Kranz, H.D., Mik, D., Sieglel, M.-L., Capesius, I., Sensen, C.W. and Huss, V.A.R. (1995) J. Mol. Evol. 41, 74–84.
[23] Strittmatter, G. and Koessel, H. (1984) Nucleic Acids Res. 12, 7633–7647.
[24] Nazar, R.N., McDougall, J.M. and VanRyk, D.I. (1987) Nucleic Acids Res. 15, 7593–7603.
[25] Takaiwa, F. and Sugiura, M. (1982) Eur. J. Biochem. 124, 13–19.
[26] Troitsky, A.V., Bobrova, V.K., Ponomarev, A.G. and Antonov, A.S. (1984) FEBS Lett. 105-109.
[27] Wakasugi, T., Nagai, T., Kapoor, M., Sugita, M., Ito, M., Ito, S., Tsudzuki, J., Nakashima, K., Tsudzuki, T., Suzuki, Y., Hamada, A., Ohta, T., Inamura, A., Yoshinaga, K. and Sugiura, M. (1997) Proc. Natl. Acad. Sci. USA 94, 5967–5972.
[28] Bobrova, V.K. (1985) Ph.D. Thesis, Biological Faculty, The Moscow State University.
[29] Graham, L.E., Delwiche, C.F. and Mishler, B.D. (1991) Adv. Bryol. 4, 213–244.