Acta Societatis Botanicorum Poloniae

DOI: 10.5586/asbp.3573

Publication history
Received: 2017-06-28
Accepted: 2018-02-07
Published: 2018-03-28

Handling editor
Andrzej Bodył, Faculty of Biological Sciences, University of Wrocław, Poland

Authors' contributions
HK and WŁ: performed sequencing and analyzed data; HK: wrote the first draft of the paper; IJO: supervised the project, analyzed data, and prepared final version of the manuscript

Funding
This work was financially supported by grant N N303 8003 40 from the National Science Center (NCN).

Competing interests
No competing interests have been declared.

Copyright notice
© The Author(s) 2018. This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits redistribution, commercial and noncommercial, provided that the article is properly cited.

Citation
Kijak H, Łodyga W, Odrzykoski IJ. Sequence diversity of two chloroplast genes: *rps4* and tRNA^{Gly} (UCC), in the liverwort *Marchantia polymorpha*, an emerging plant model system. Acta Soc Bot Pol. 2018;87(1):3573. https://doi.org/10.5586/asbp.3573

Digital signature
This PDF has been certified using digital signature with a trusted timestamp to assure its origin and integrity. A verification trust dialog appears on the PDF document when it is opened in a compatible PDF reader. Certificate properties provide further details such as certification time and a signing reason in case any alterations made to the final content. If the certificate is missing or invalid it is recommended to verify the article on the journal website.

ORIGINAL RESEARCH PAPER

Sequence diversity of two chloroplast genes: *rps4* and tRNA^{Gly} (UCC), in the liverwort *Marchantia polymorpha*, an emerging plant model system

Hanna Kijak, Weronika Łodyga, Ireneusz J. Odrzykoski*

Faculty of Biology, Adam Mickiewicz University in Poznań, Umultowska 89, 60-614 Poznań, Poland

* Corresponding author. Email: ireko@amu.edu.pl

Abstract

The primary purpose of this study is to evaluate the sequence variation for two regions of chloroplast DNA in a collection of 27 taxonomically well-annotated clonal lines of *Marchantia polymorpha* sensu lato derived from European populations. We attempted to develop molecular markers so as to identify three taxa usually recognized as subspecies. We sequenced two regions: the *rps4* gene along with the *rps4-trnT* intergenic spacer and an intron of the tRNA^{Gly} (UCC) gene. Samples of *Marchantia paleacea* ssp. *diptera* from Japan were used for comparative purposes.

Three haplotypes (MA, MB, and MC) were identified for the species, and almost all sequence divergence between subspecies was found to occur at the level of 0.0023–0.0032 substitutions per site. The sequence divergence between *M. polymorpha* and *M. paleacea* was tenfold greater (0.0331–0.0340). We did not detect any differences between *M. paleacea* and homologous sequences from the reference chloroplast genome of *M. polymorpha* obtained from the GeneBank (NC_001319). It was confirmed that the cell suspension line A-18 used for the sequencing of the full chloroplast genome in 1986 was incorrectly taxonomically annotated.

Keywords

nucleotide polymorphism; chloroplast DNA; NC_001319.1 reference genome; *rps4* gene; tRNA^{Gly} (UCC) intron; *Marchantia polymorpha*; *Marchantia paleacea* subsp. *diptera*

Introduction

The liverwort *Marchantia polymorpha* L. has recently become one of the most important models for plant biology research and evolutionary genomics due to its relatively simple genome and unique phylogenetic position as a member of the early land plant lineage. Genomic research in *Marchantia* began in the late 1980s, when the full chloroplast, and later, mitochondrial genomes, were sequenced [1,2]. These sequences are frequently used as a reference in many comparative studies. After some delays, in comparison with other model plants (e.g., *Physcomitrella patens* among Bryophyta), the interest in *M. polymorpha* has recently been revived following the development of various molecular tools and techniques (see [3–5] for recent reviews). Furthermore, several new nuclear genome sequencing projects have recently begun and the first genome of *M. polymorpha* subsp. *ruderalis* was published this year [6].

Despite the importance of this liverwort in evolutionary studies, little is known about its genetic diversity. Since the publication of the original description, three taxa of various taxonomic ranks have been recognized within *M. polymorpha* sensu lato. After extensive revision of the genus and detailed nomenclatural studies, Bischler-Causse and Boisselier-Dubayle [7] proposed three subspecies, namely: *M. polymorpha*

subsp. *montivagans*, *M. polymorpha* subsp. *polymorpha*, and *M. polymorpha* subsp. *ruderalis* in place of the three species: *M. alpestris*, *M. aquatica*, and *M. polymorpha* [8] This new taxonomic solution – a single species comprising three subspecies – has been generally accepted (see, e.g., [9–12]). However, some authors still argue that the frequent sympatric distribution of at least two of these subspecies in Europe indicates species-level divergence [13,14].

Several morphological characters, together with habitat preference, may be used to distinguish between these three taxa [5,9]. Certain diagnostic isozyme variants, the RFLP profile of 18s rDNA, and specific RAPD markers have also been discovered [15]. However, these early molecular markers are difficult to apply for identification purposes in some samples, including herbarium specimens and some laboratory culture lines.

The primary purpose of this work was to evaluate the suitability of two commonly studied chloroplast DNA regions – the *rps4* gene and an intron of the $tRNA^{Gly}$ (UCC) gene in distinguishing between three taxa of *M. polymorpha* sensu lato. This study delivers markers for the reliable identification of various samples, including laboratory culture lines.

Material and methods

Plant material

Our analyses were of 23 carefully selected clones of *M. polymorpha* from European populations, representing all three subspecies, and a single clone (MC-57) from Japan. In addition, three samples of *M. paleacea* ssp. *diptera* from Japan were sequenced and compared against the reference sequence of the *M. polymorpha* chloroplast genome (GeneBank accession NC_001319). The selected clones were obtained from a phytotron collection maintained in Poznań since 2007. Samples were kept in 0.5-L closed containers (punched to allow ventilation and drainage) on a sterilized mineral substrate at a constant temperature of 15°C and 16-h daylight at 50 µmol photon m^{-2} s^{-1}. A list of samples together with details of the geographical location and the GenBank accession numbers are given in Tab. 1. Information on vouchers deposited in POZW can be found as part of GeneBank record.

Subspecies identification

We used certain morphological characters of the thalli (as described by Long [9]) and the electrophoretic pattern of nonspecific esterases (E.C. 3.1.1.-) for the identification of subspecies, in accordance with the work of Boisselier-Dubayle and Bischler [16]. Isozyme profiles were used in this paper as a part of the formal taxonomic description for each of the three subspecies. Enzymes were extracted by the homogenization of small apical fragments of young gametophytes in 100 µL of 0.1 M Tris-HCl buffer pH 7.5 with the addition of 10 mM $MgCl_2$, 1 mM EDTA (Na_4 salt) 0.4% Triton X-100, and 14 mM 2-mercaptoethanol. The homogenate was filtered through a strip of Miracloth tissue and transferred onto small 4 × 11 mm Whatman 3ET wicks. Samples were subjected to horizontal starch gel electrophoresis (11.5% starch; Starch Art) in a Tris-citrate/lithium boric acid buffer system in accordance with Odrzykoski and Szweykowski [17]. Zymograms displayed a similar esterase isozyme profile to the Tris-glycine system, and acrylamide gels originally used by Boisselier-Dubayle and Bischler [16]. The localization of alpha-esterases follows Method 1 of Manchenko [18].

Genomic DNA extraction

The total genomic DNA was extracted from living plants. A fragment of a single gametophyte (approximately 50 mg fresh weight) was placed in 2-mL tubes with two steel beads, and after freezing in liquid nitrogen, ground for 60 s in a tissue disruptor (Retch Ball Mill). They were then subjected to the CTAB-based procedure modified

Tab. 1 A list of studied populations with locations and sequence GeneBank accession numbers. MA, MB, MC represent three haplotypes within *Marchantia polymorpha* sensu lato corresponding to subspecies: *montivagans*, *polymorpha*, and *ruderalis*.

Taxon	Pop. No.	Locality	Province	Country	Longitude	Latitude	*rps4*	tRNA^{Gly}
MA	06	Valtavaara	Kuusamo	Finland	66°12′19″ N	29°12′43″ E	KX577604	KM007131
MA	12	Pyhatunturi	Lappland	Finland	67°00′49″ N	27°08′56″ E	KX577605	KM007132
MA	33	Ihlara Valley	Anatolia	Turkey	38°14′20″ N	34°15′51″ E	KX577606	KM007133
MA	38	Skutustadir, Myvatn Lake	Sudur-Thingeyjarsysla	Iceland	65°34′11″ N	17°02′06″ W	KX577607	KM007134
MA	39	Sejlalandfoss Waterfall	Rangárvallasýsla	Iceland	63°33′56″ N	19°59′24″ W	KX577608	KM007135
MA	42	Rilski Monastir, Rila Mts	Kjustendił	Bulgaria	42°08′00″ N	23°20′24″ E	KX577609	KM007136
MB	03	S Wornigacie	Pomorskie	Poland	53°51′35″ N	17°28′56″ E	KX577610	KM007137
MB	08	Mano Wo 1	Zachodniopomorskie	Poland	54°09′11″ N	16°17′31″ E	KX577611	KM007138
MB	25	Białowieża Forest	Podlaskie	Poland	52°37′36″ N	23°43′27″ E	KX577612	KM007139
MB	31	Trigrad George, Rhodope Mts	Smolyan	Bulgaria	41°37′00″ N	24°22′45″ E	KX577613	KM007140
MB	61	Velingrad, Rila Mts	Pazardzhik	Bulgaria	42°06′49″ N	23°54′43″ E	KX577614	KM007141
MC	04	Sarajevo	Sarajevo	Bosnia	43°84′76″ N	18°35′64″ E	KX577615	KM007154
MC	09	Mano Wo 2	Zachodniopomorskie	Poland	54°05′20″ N	16°20′69″ E	KX577616	KM007142
MC	15	Poznań	Wielkopolskie	Poland	52°25′35″ N	16°54′42″ E	KX577617	KM007143
MC	22	Tulcea	Dobruja	Romania	45°11′24″ N	28°48′00″ E	KX577618	KM007155
MC	24	Wanie Wo	Podlaskie	Poland	53°04′30″ N	22°49′30″ E	KX577619	KM007144
MC	26	Lourdes, Pyrenees Mts	Midi-Pyrenees	France	43°06′00″ N	00°03′00″ W	KX577620	KM007145
MC	29	Horodło	Lubelskie	Poland	50°53′40″ N	24°02′47″ E	KX577621	KM007156
MC	34	Kamień Pomorski	Zachodniopomorskie	Poland	53°58′11″ N	14°47′09″ E	KX577622	KM007156
MC	41	Rozenski Monastir, Pirin Mts	Blagoevgrad	Bulgaria	41°31′50″ N	23°25′35″ E	KX577623	KM007147
MC	44	Kaisaniemi Bot. Garden	Helsinki	Finland	60°10′20″ N	24°56′51″ E	KX577624	KM007148
MC	45	Kaiser-Franz-Josef-Höhe	Hohe Tauern Mts	Austria	47°04′30″ N	12°45′04″ E	KX577625	KM007157
MC	47	Marken, Waterland	Noord-Holland	Netherlands	52°27′30″ N	05°06′24″ E	KX577626	KM007149

© The Author(s) 2018 Published by Polish Botanical Society Acta Soc Bot Pol 87(1):3573

Tab. 1 Continued

Taxon	Pop. No.	Locality	Province	Country	Longitude	Latitude	*rps4*	tRNA^{Gly}
MC	57	Shinaga Wa	Gifu Pref.	Japan	36°08′49″ N	137°15′18″ E	KX577627	KM007150
MPAL	18	Ui, Ootoucho, Gojyoushi	Nara Pref.	Japan	34°40′47″ N	135°50′41″ E	KX577628	KM007151
MPAL	53	Kyoto	Kioto Pref.	Japan	35°01′28″ N	135°47′09″ E	KX577629	KM007152
MPAL	58	Tsumago	Nagano Pref.	Japan	35°34′10″ N	137°35′52″ E	KX577631	KM007158

from Doyle and Doyle [19]. Into each tube, 750 µL of extraction buffer (100 mM Tris-HCl, pH 8.0, 1.4 M NaCl, 20 mM EDTA, and 2% CTAB) were added, and after rapid vortexing, the contents incubated for 30 min at 65°C in a thermomixer (600 rpm). Subsequently, an equal volume of chloroform/isoamyl alcohol (24:1) mixture was added. Phase separation was conducted by centrifuging for 2 min at 14,000 *g*), and this step was repeated three times. The final aqueous phase was transferred to a fresh tube, and the DNA was precipitated with an equal volume of isopropanol. The DNA was again precipitated for 2 min at 4,000 *g*, and the pellet was washed twice with 1,000 µL 70% ethanol. The DNA was dissolved in 100 µL TE buffer (0.1 M Tris-HCl, 0.01 M EDTA, pH 7.5). For RNase digestion, 10 µg RNase A (10 mg/mL) was added and incubated for 30 min at 37°C. The concentration of DNA was determined spectrophotometrically (Nanodrop ND-1000) and analyzed for integrity using 0.8% agarose gel electrophoresis. The extracted DNA samples were stored at −20°C.

PCR amplification and sequencing

The PCR mixture contained the following components for 20-µL reactions: 10.8 µL Millipore H_2O, 2.0 µL 10× buffer and 0.05 units HiFi Taq DNA polymerase (Novazyme, Poland), 200 mM of each dNTP, 0.25 mM of each primer, and 2.0 ng total DNA. PCR was initiated by denaturation at 94°C for 3 min, followed by 35 cycles: 1 min denaturation at 94°C, 1 min annealing at 54°C, and 5 min elongation at 72°C. The reactions ended with 5 min elongation at 72°C (Applied Biosystems Veriti Thermal Cycler). These parameters were used for both regions. For the amplification of the *rps4* gene region, we used the primers RP1F (GCTATGTAGGCTTTTGGTC) and RP1R (CACTTGTAATGCGATGGTC), newly designed based on the reference sequence of *M. polymorpha*. Primers used for sequencing: RS2F (CTAAACGAATACGATACTGAG) and RS1R (TTTTGTAACATAAAGGAG). For the amplification and sequencing of the intron of the tRNA^{Gly} (UCC) gene, we used two primers: TG_F (CGGGTACGGGAATCGAAC) and TG_R (GCG GGT ATA GTT TAG TGC) in accordance with Szweykowska-Kulińska et al. [20]. Prior to sequencing, PCR products were purified using the Exonuclease I-Shrimp Alkaline Phosphatase enzymatic treatment. About 10 ng of PCR products were used as templates in 10-µL sequencing reactions with the BigDye Terminator v3.1 DNA Sequencing Kit (Applied Biosystems) on a 3130xl Genetic Analyzer (Applied Biosystems).

Sequence analysis

The CodonCode Aligner v. 5.0.1 (CodonCode Co., USA) was used for the visual inspection and editing of the chromatograms from the forward and reverse sequencing primers, sequence alignment, and detection of polymorphic sites. Regions of the chloroplast DNA corresponding to the chloroplast DNA reference sequences for *M. polymorpha* (NC-01319) were used as outgroups. The evolutionary distances were computed in MEGA6 [21] using the number of base differences method and the Tamura-3-parameter model of nucleotide substitution. The detailed phylogenetic analysis was not performed, but in order to visualize the results, a maximum likelihood (ML) tree was constructed for a matrix that combines both the investigated regions.

Results

Electrophoretic phenotypes of esterase isozymes

All selected samples were classified into three groups (MA, MB, MC) based on the electrophoretic phenotypes of alpha-esterase. Of the 27 clones, Variant 1, which is characteristic of *M. polymorpha* subsp. *ruderalis* (group MC), was detected in 13 clones. Variant 2, specific for *M. polymorpha* subsp. *polymorpha* (group MB),

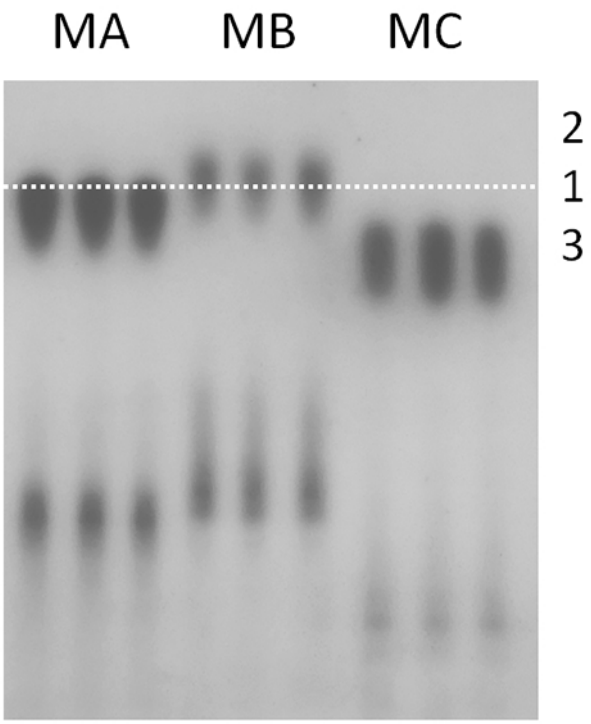

Fig. 1 Electrophoretic phenotypes of alpha-esterases from nine samples of three taxa (MA, MB, MC) of *Marchantia polymorpha* sensu lato. Positions 1–3 from the right (Variant 1) positions 4–6 (Variant 2), and positions 7–9 (Variant 3). Variant numbering is based on that of Bischler-Causse and Boisselier-Dubayle [7].

was present in five clones, and Variant 3, specific for *M. polymorpha* subsp. *montivagans* (group MA), in six clones (Fig. 1).

Sequence characteristics

In the first region (*rps4*), the obtained sequences included a small portion of intergenic spacer between tRNA^{Ser} and the *rps4* gene, the whole sequence of *rps4* and *rps4*-tRNA^{Thr} intergenic spacer. For further comparison, sequences were trimmed to include a complete sequence of the *rps4* gene (609 bp long) and intergenic spacer *rps4-trnT* (228 bp) only and aligned to the corresponding region of the full chloroplast genome of *M. polymorpha* (NC_001319, positions 49,425–50,261).

Within the *rps4* gene, two polymorphic sites were identified in *M. polymorpha* complex: (***i***) T to C at position 98 commencing at the beginning of the gene in group MB only, and (***ii***) C to T at position 492 in group MC. Within *rps4*-tRNA^{Thr} spacer, SNP polymorphism was detected at two sites: at position 55 (C to T) in a single sample (MC_22), and at position 167 (G to T) in five of 13 samples of group MC only. The entire 836-bp regions were highly differentiated between *M. polymorpha* and *M. paleacea*, with 16 substitutions and four single nucleotide indels (Tab. 2).

The investigated fragment of an intron of the tRNA^{Gly} (UGG) gene was slightly shorter in *M. polymorpha* than in *M. paleacea*. Three diagnostic substitutions were detected in *M. polymorpha*, characteristic for each group: A–G substitution at position 120 present in MC only, and at position 341 present in MB only. Unique for group MA is a C–T substitution at position 525. In a comparison between *Marchantia polymorpha* and *M. paleacea*, differences were detected at 23 positions within 593 bp of investigated sequence (Tab. 3). In the case of both investigated regions, all samples of *M. paleacea* had an identical sequence to that of the reference sample.

Sequence divergence between *M. polymorpha*, *M. paleacea*, and the reference sample

In order to estimate the sequence divergence between the investigated samples, we calculated the number of base substitutions per site. We used a total of 1,297 positions, excluding indels. The number of substitutions between the subspecies of *M. polymorpha* varied from 0.0023 to 0.0032, the greatest being between *M. polymorpha* subsp. *polymorpha* (MB) and two other subspecies. The divergence between *M. polymorpha* sensu lato and *M. paleacea* was tenfold greater, between 0.0331 and 0.0340 (Tab. 4). The maximum likelihood tree for the combination of all investigated sequences and corresponding sequences from the reference sample (NC_001319) showed three clades with a bootstrap of greater than 60%, which corresponded to infraspecific groups of *M. polymorpha* (Fig. 2). All samples of *Marchantia paleacea* were placed together with homologous sequences from the complete cpDNA genome of *M. polymorpha* (NC_001319).

Discussion

Two species of *Marchantia* (*M. polymorpha* L. and *M. paleacea* Bertol.), have played a significant role in our understanding of developmental biology and evolutionary

Tab. 2 Polymorphic sites detected within the *rps4* gene and *rps4*-tRNA^{Thr} spacer of investigated samples of *Marchantia polymorpha* sensu lato (MA, MB, MC) and *M. paleacea* compared against the reference sequence for "*M. polymorpha*" NC_001319. The numbers in the first row correspond to nucleotide position commencing from the beginning of the gene and spacer.

	***rps4* gene**										***rps4*-tRNA^{Thr} spacer**													
	83	**98**	**195**	**228**	**249**	**273**	**443**	**486**	**492**	**522**	**7**	**13**	**15**	**19**	**20**	**24**	**25**	**30**	**36**	**55**	**86**	**149**	**152**	**167**
MA_06	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MA_12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MA_33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MA_38	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MA_39	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MA_42	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MB_03	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MB_08	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MB_25	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MB_31	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MB_61	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MC_04	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MC_09	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MC_15	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	-	.	.	T
MC_22	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	-	.	.	T
MC_24	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MC_26	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MC_29	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MC_34	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MC_41	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	-	.	.	T
MC_44	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
MC_45	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	-	.	.	K
MC_47	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	-	.	.	T
MC_57	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.
M. polymorpha	C	T	G	C	A	C	C	G	C	G	T	A	A	C	T	G	A	T	A	C	-	T	C	G
MP_18	T	T	A	A	G	T	T	A	C	A	-	T	-	A	-	A	C	A	T	C	T	G	A	G
MP_53	T	T	A	A	G	T	T	A	C	A	-	T	-	A	-	A	C	A	T	C	T	G	A	G
MP_54	T	T	A	A	G	T	T	A	C	A	-	T	-	A	-	A	C	A	T	C	T	G	A	G
MP_58	T	T	A	A	G	T	T	A	C	A	-	T	-	A	-	A	C	A	T	C	T	G	A	G
M. paleacea	T	T	A	A	G	T	T	A	C	A	-	T	-	A	-	A	C	A	T	C	T	G	A	G
NC_001319	T	T	A	A	G	T	T	A	C	A	-	T	-	A	-	A	C	A	T	C	T	G	A	G

© The Author(s) 2018 Published by Polish Botanical Society

Tab. 3 Polymorphic sites detected within the tRNA^{Gly} (UCC) intron between investigated samples of *Marchantia polymorpha* sensu lato and *M. paleacea* compared against the reference sequence for "*M. polymorpha*" NC_001319. The numbers in the first row correspond to nucleotide position commencing from the beginning of the gene.

	120	170	171	212	254	261	303	324	339	341	370	430	438	459	482	483	488	489	492	493	494	495	496	502	503	504	506	521	525	526	527	553
MA-06	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	T	.	.	.
MA-12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	T	.	.	.
MA-33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	T	.	.	.
MA-38	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	T	.	.	.
MA-39	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	T	.	.	.
MA-42	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	T	.	.	.
MB-03	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MB-08	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MB-25	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MB-31	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MB-61	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MC-04	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MC-09	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MC-15	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MC-22	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MC-24	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MC-26	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MC-29	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MC-34	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MC-41	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MC-44	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MC-45	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MC-47	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
MC-57	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	-	-	.	.	.	.	.	.
M. polymorpha	A	T	T	C	T	T	T	T	C	A	G	A	T	A	G	T	A	A	A	A	T	-	-	-	-	-	T	T	C	A	C	C
MP-18	A	A	A	A	A	G	C	A	A	A	A	G	C	-	C	G	G	T	T	T	C	A	A	A	C	A	G	C	C	C	G	G
MP-53	A	A	A	A	A	G	C	A	A	A	A	G	C	-	C	G	G	T	T	T	C	A	A	A	C	A	G	C	C	C	G	G
MP-56	A	A	A	A	A	G	C	A	A	A	A	G	C	-	C	G	G	T	T	T	C	A	A	A	C	A	G	C	C	C	G	G
MP-58	A	A	A	A	A	G	C	A	A	A	A	G	C	-	C	G	G	T	T	T	C	A	A	A	C	A	G	C	C	C	G	G
M. paleacea	A	A	A	A	A	G	C	A	A	A	A	G	C	-	C	G	G	T	T	T	C	A	A	A	C	A	G	C	C	C	G	G
NC_001319	A	A	A	A	A	G	C	A	A	A	A	G	C	-	C	G	G	T	T	T	C	A	A	A	C	A	G	C	C	C	G	G

© The Author(s) 2018 Published by Polish Botanical Society

Tab. 4 Estimates of evolutionary divergence (number of base substitutions per site and standard error above diagonal) between three subspecies of *M. polymorpha* sensu lato (MA, MB, and MC), *M. paleacea* subsp. *diptera* (MP), and the reference sequence (NC_001319). Calculations based on a combined matrix of the three investigated regions with 1,297 positions in the final dataset.

		MA	MB	MC	MP
M. p. montivagans	MA		0.0012	0.0012	0.0047
M. p. polymorpha	MB	0.0023		0.0014	0.0047
M. p. ruderalis	MC	0.0024	0.0032		0.0046
M. paleacea	MP	0.0334	0.0342	0.0342	

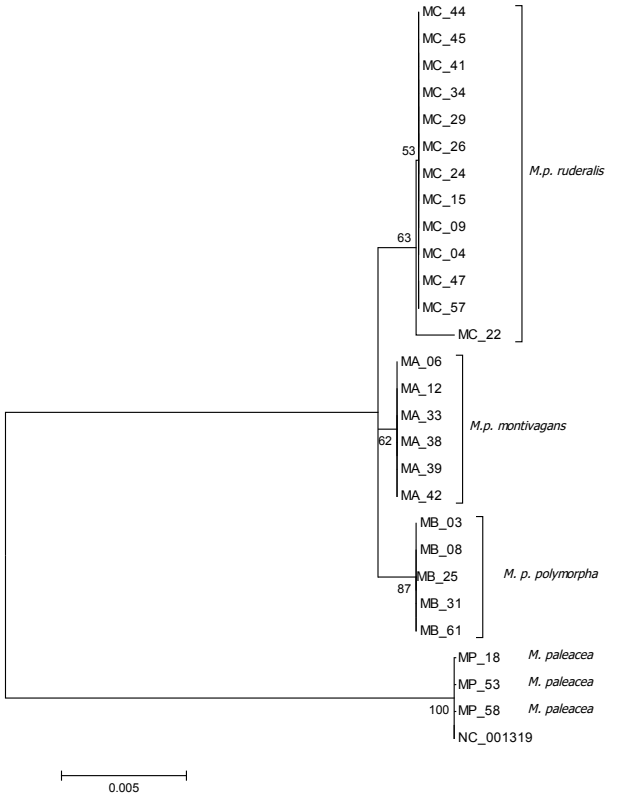

Fig. 2 Maximum likelihood phylogenetic tree based on the combined set of all investigated cpDNA regions (a total of 1,315 positions in the final dataset) and the Tamura-3-parameter model of nucleotide substitutions (best BIC score in MEGA model test). The tree is drawn to scale, with branch lengths measured based on the number of substitutions per site. Bootstrap values (1,000 replicates) are shown next to the branches.

studies [4,5,22]. The first species was formally described in Europe, although it has an almost cosmopolitan geographic range. The original description distinguishes between three unnamed infraspecific taxa (α, β, γ), with later varieties being named: *alpestris communis*, *aquatica*, and *domestica communis* [3]. Experiments involving the crossing of these varieties under greenhouse conditions revealed partial reproductive isolation, one of the reasons for their recognition as separate species by Burgeff [8]. This solution was accepted in some liverwort floras [13], but today, these species (*M. alpestris*, *M. aquatica*, and *M. polymorpha*) are mainly considered to be infraspecific taxa (subspecies) following taxonomic revision of the genus by Bischler-Causse [23] and subsequent lectotypification [7]. This taxonomic rank was adopted in a recent checklist of hornworts and liverworts [12]. Some additional arguments to support species rank of these taxa have been recently summarized by Bowman et al. [24].

The known morphological markers for distinguishing between these taxa include some features of the thallus (width, color, branching pattern, presence or absence of a black median line on the thallus surface) and type (entire or toothed) of appendage margin [9]. Owing to environmental plasticity, morphological markers alone can seldom be used for the identification of some specimens present in scientific collections (e.g., in vitro or cell suspension cultures). Additional markers include electrophoretic variants of some isozymes (nonspecific esterases), RAPD profile, and RFLP of 18s rDNA [15]. Recently, several organellar and nuclear sequences have also become available in public sequence databases and have frequently been used in phylogenetic studies of the major taxonomic groups (e.g., [25–27]).

In this work, we have attempted to find diagnostic mutations for the infraspecific taxa of *M. polymorpha* sensu lato within two regions of chloroplast DNA. Our sample of plants was taxonomically annotated carefully using the original description, including the esterase isozyme profile [7,17] and selected to include both a broad geographical range and major types of environments (tundra, mountains, anthropogenic). The selected genes (*rps4* and $tRNA^{Gly}$) are amongst the most frequently used in phylogenetic studies of bryophytes (e.g., [28–30]).

Three haplotypes were recognized (MA, MB, and MC) with only minor variations in MC, and these appear to be diagnostic for the three subspecies. In both regions, single mutations allow for the correct identification of infraspecific taxa within the investigated sample of plants. Much greater differences were detected between *M. polymorpha* sensu lato and samples of *M. paleacea* ssp. *diptera* from Japan. This second species belongs to subgenus *Chlamidium* (Corda) Bischl., the largest in genus *Marchantia*, with 21 accepted species [12]. Comparisons of both investigated genes against corresponding regions of the reference chloroplast DNA genome obtained from the NCBI GeneBank (NC_001319.1) place this sequence with *M. paleacea*, not with *M. polymorpha* subspecies studied here.

This result confirms our hypothesis concerning the incorrect taxonomic annotation of a laboratory cell line (A18) used in the complete cpDNA sequencing project (see Kijak et al. [31]). This conclusion was recently validated by Villarreal et al. [27] in an extensive phylogenetic study of the class Marchantiopsida (79 species and 11 loci from

both organellar genomes). It is probable that these studies were taken into account in preparing a new description of *Marchantia* chloroplast DNA reference genomes. The same sequence has two GenBank accession numbers (NC_001319 *M. polymorpha* and a new: X04465.1 *M. paleacea*).

It appears that the taxonomic rank proposed by Bischler-Cause and Boisselier-Dubayle [7] for the three taxa occurring within *M. polymorpha* is inadequate for explaining the degree of morphological, ecological, and molecular differentiation present, since the divergence time between the three subspecies of *M. polymorpha* was estimated by Villareal et al. [27] to be about 5 Ma (2–11 Ma). The results of our study, by comparison, would appear to support the concept of three separate species [13,14]. Further studies into genetic differentiation, however, are necessary.

In this paper, we present the results of screening of cpDNA sequences for diagnostic mutations which can be used to identify these taxa. No other sequence regions have previously been tested on such a large scale, including short sequences proposed as "DNA barcodes" for land plants [32]. Our results indicate that one of the investigated regions ($tRNA^{Gly}$ intron) could be used as such a marker since it contains a diagnostic mutation for each taxon and can be amplified and sequenced using a single pair of primers. The diagnostic value of this region requires verification based on a larger sample size spanning the whole geographical range of this scientifically important species.

Acknowledgments

We thank Professor Robert Wyatt for helpful comments on the early version of this paper. Dr Makasuni Kimura, Dr Iwona Melosik, Izolda Blicharska-Odrzykoska, and Samuel J. Odrzykoski assisted in the collection of the samples used in this study.

References

1. Ohyama K, Fuzukawa H, Kohchi T, Shirai H, Sano T, Sano S, et al. Chloroplast gene organization deduced from complete sequence from liverwort *Marchantia polymorpha* chloroplast DNA. Nature. 1986;322:572–574. https://doi.org/10.1038/322572a0
2. Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, et al. Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive form of plant mitochondrial genome. J Mol Biol. 1992;223:1–7. https://doi.org/10.1016/0022-2836(92)90708-R
3. Bowman J. A brief history of *Marchantia* from Greece to genomics. Plant Cell Physiol. 2015;57(2):210–219. https://doi.org/10.1093/pcp/pcv044
4. Bowman J, Araki T, Kohchi T. *Marchantia*: past, present and future. Plant Cell Physiol. 2016;57:205–209. https://doi.org/10.1093/pcp/pcw023
5. Shimamura M. *Marchantia polymorpha*; taxonomy, phylogeny and morphology of a model system. Plant Cell Physiol. 2015;57:230–256. https://doi.org/10.1093/pcp/pcv192
6. Bowman JL, Kohchi T, Yamato KT, Jenkins J, Shu S, Ishizaki K, et al. Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. Cell. 2017;171:287–304. https://doi.org/10.1016/j.cell.2017.09.030
7. Bischler-Causse H, Boisselier-Dubayle MC. Lectotypification of *Marchantia polymorpha* L. J Bryol. 1991;16:361–365. https://doi.org/10.1179/jbr.1991.16.3.361
8. Burgeff H. Genetische Studien an *Marchantia*. Jena: Fisher; 1943.
9. Long DG. *Marchantia polymorpha* in Britain and Ireland. Bulletin of the British Bryological Society. 1995;66:29–37.
10. Paton JA. The liverwort flora of the British Isles. Colchester: Harley Books; 1999.
11. Grolle R, Long DG. An annotated check-list of the Hepaticae and Anthocerotae of Europe and Macaronesia. J Bryol. 2000;22:103–140. https://doi.org/10.1179/jbr.2000.22.2.103
12. Söderström L, Hagborg A, von Konrat M, Bartholomew-Began S, Bell D, Briscoe L, et al. World checklist of hornworts and liverworts. PhytoKeys. 2016;59:1–828. https://doi.org/10.3897/phytokeys.59.6261

13. Schuster RM. The Hepaticae and Anthocerotae of North America. Vol. 6. Chicago, IL: Field Museum of Natural History; 1992.
14. Damsholt K. Illustrated flora of Nordic liverworts and hornworts. 2nd ed. Lund: Nordic Bryological Society; 2009.
15. Boisselier-Dubayle MC, Jubier MF, Lejeune B, Bischler H. Genetic variability in the three subspecies of *Marchantia polymorpha* (Hepaticae): isozymes, RFLP and RAPD markers. Taxon. 1995;44:363–376. https://doi.org/10.2307/1223406
16. Boisselier-Dubayle MC, Bischler H. Electrophoretic studies in *Marchantia polymorpha* L. Journal of the Hattori Botanical Laboratory. 1989;67:297–311.
17. Odrzykoski IJ, Szweykowski J. Genetic differentiation without concordant morphological divergence in the thallose liverwort *Conocephalum conicum*. Plant Syst Evol. 1991;178:135–152.
18. Manchenko GP. Handbook of detection of enzymes on electrophoretic gels. Boca Raton, FL: CRC Press; 2003. https://doi.org/10.1021/ja0335420
19. Doyle JJ, Doyle JL. A rapid DNA isolation procedure from small quantities of fresh leaf tissues. Phytochemical Bulletin. 1987;19:11–15.
20. Szweykowska-Kulińska Z, Pacak A, Jankowiak K. New DNA markers for discrimination between closely related species and for the reconstruction of historical events: an example using liverworts. Cell Mol Biol Lett. 2002;7:403–416.
21. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013;30:2725–2729. https://doi.org/10.1093/molbev/mst197
22. Alam A, Pandey S. *Marchantia polymorpha* L.: an emerging model plant system to study contemporary plant biology – a review. Plant Science Today. 2016;3:88–99. https://doi.org/10.14719/pst.2016.3.2.221
23. Bischler-Causse H. *Marchantia* L. The European and African taxa. Berlin: J. Cramer; 1993. (Bryophytorum Bibliotheca; vol 45).
24. Bowman JL, Araki T, Arteaga-Vazquez MA, Berger F, Dolan L, Haseloff J, et al. The naming of names: guidelines for gene nomenclature in *Marchantia*. Plant Cell Physiol. 2016;57(2):257–261. https://doi.org/10.1093/pcp/pcv193
25. Boisselier-Dubayle MC, Lambourdiere J, Bischler H. Molecular phylogenies support multiple morphological reductions in the liverwort subclass Marchantiidae (Bryophyta). Mol Phylogenet Evol. 2002;24:66–77. https://doi.org/10.1016/S1055-7903(02)00201-4
26. Forrest LL, Davis EC, Long DG, Crandall-Stotler BJ, Clark A, Hollingsworth ML. Unravelling the evolutionary history of the liverworts (Marchantiopyta): multiple taxa, genomes and analyses. Bryologist. 2006;109(3):303–334. https://doi.org/10.1639/0007-2745(2006)109[303:UTEHOT]2.0.CO;2
27. Villarreal JC, Crandall-Stotler BJ, Hart ML, Long DG, Forrest LL. Divergence times and the evolution of morphological complexity in an early land plant lineage (Marchantiopsida) with a slow molecular rate. New Phytol. 2015;209(4):1734–1746. https://doi.org/10.1111/nph.13716
28. Goffinet B, Cox CJ, Shaw J, Hedderson TAJ. The Bryophyta (mosses): systematic and evolutionary inferences from an *rps4* gene (cpDNA) phylogeny. Ann Bot. 2001;87:191–208. https://doi.org/10.1006/anbo.2000.1318
29. Wang J, Zhu RL. Lejeuneaceae (Marchantiophyta) share an indel in chloroplast *rps4* gene. J Bryol. 2008;30:174–177. https://doi.org/10.1179/174328208X300642
30. Stech M, Quandt D. 20,000 species and five key markers: the status of molecular bryophyte phylogenetics. Phytotaxa. 2010;9:196–228. https://doi.org/10.11646/phytotaxa.9.1.11
31. Kijak H, Kisiel K, Miwa H, Cieślewicz A, Odrzykoski IJ. What has been actually sequenced? Organellar genome sequences of *Marchantia polymorpha* from GenBank belong to *M. paleacea* subsp. *diptera* not to *M. polymorpha* sensu lato. In: Book of abstracts: International Association of Bryologist Conference; 2013 Oct 15–19; London, England. London: Natural History Museum; 2013. p. 35–36.
32. CBOL Plant Working Group. A DNA barcode for land plants. Proc Natl Acad Sci USA. 2009;106(31):12794–12797. https://doi.org/10.1073/pnas.0905845106