Phylogeny of the European species of the genus *Pellia* (Hepaticae; Metzgeriales) based on the molecular data from nuclear *tRNA_{Leu}^{CAA}* intergenic sequences

Pawel Fiedorow, Ireneusz Odyzykoski, Jerzy Szweykowski, Zofia Szweykowska-Kulinska

1. Introduction

Genetic and phylogenetic relationships in *Pellia* species have been studied using different methods. The genus *Pellia* (Hepaticae, Metzgeriales) is represented in Poland by two diploid, dioecious species (*P. epiphylla*, *P. neesiana*) and one tetraploid, monoecious species (*P. borealis*). Unfortunately these species are morphologically similar and their proper identification causes difficulties (especially when they are sterile as is frequently the case). To solve this problem isozyme markers were developed (Zielinski, 1987). Multilocus genotypes provided good diagnostic characters to distinguish easily and reliably all *Pellia* species. Additionally, isoenzyme studies revealed two different multienzyme genotypes in *Pellia endiviifolia*. The high genetic distance suggests that these genotypes are actually two different species. Because they are impossible to distinguish morphologically, they could not be recognized by formal taxonomy; thus we treat them as two cryptic species: *P. endiviifolia*-species A and species B. Similar results were also obtained earlier in *P. epiphylla* where also two different multienzyme genotypes were found (Zielinski, 1987; Szweykowski and Odrzykoski, 1990). This taxon represents a group of two morphologically indistinguishable and allopatric (at least in Poland), cryptic species: *P. epiphylla*-species N and *P. epiphylla*-species S.
Species-N grows in the northern part of Poland, while species -S is found mostly in its southern part (Szweykowski et al., 1995). Comparison of these two taxa with diploid *P. borealis* produced very interesting results. Three independent isozyme markers suggested an allopolyploid origin of *P. borealis* as a result of hybridization between the two cryptic species: *P. epiphylla-* species N and *P. epiphylla-* species S (Odrzykoski et al., 1996). This hypothesis had to be examined using different, independent methodology, particularly in the light of some previous studies that suggested an autoploidy of *P. borealis* (Zielinski, 1987).

To solve this controversy we have decided to study all Polish *Pellia* species at the DNA level. In order to have well-defined plant material for our research, an in vitro collection of all Polish species was established (Fiedorow and Szweykovska-Kulinska, 1998a). Because of the putative very close relationship between: *P. epiphylla-* species N and species S and *P. endiviifolia-* species A and species B we used DNA sequences that evolve rapidly. Based on the concerted evolution phenomenon, we choose sequences of tandemly repeated gene families. Concerted evolution is a process that produces and maintains homogeneity within related gene families within a species in, for example rDNA, histones, tRNA or 5S rRNA genes (Dover, 1982; Amstutz et al., 1985; Hillis and Dixon, 1991; Drouin and Moniz de Sa, 1995). Thus these sequences should show high homogeneity within one species and should differ strongly even between closely related species. Positive results concerning identification of closely related species were obtained using ITS (internal transcribed spacer) 1 and 2 sequences (Da Rocha and Bertrand, 1995). Additionally we developed a new type molecular marker that can differentiate closely related species (in peat-mosses and in liverworts), tRNA genes arranged in tandem arrays. Transfer RNA genes can be arranged in the nuclear genome in two different ways: they can be either scattered throughout the genome or they can form clusters containing one or several types of tRNA genes (Beier et al., 1991; Szweykovska-Kulinska and Badzmierowska, 1992). We assumed that tRNA gene spacers of tandem repeated tRNA genes can be used as phylogenetic marker analogous to ITS sequences (Fiedorow and Szweykovska-Kulinska, 1998b). To analyze *Pellia* species we analyzed intergenic spacers that separate nuclear tRNA*Leu* genes organized in tandem repeats.

2. Materials and methods

2.1. Liverworts

Plants for this study were collected from twelve Polish populations (see Table 1). Species were identified using multienzyme genotyping (starch gel electrophoresis) (Szweykowski et al., 1981; Zielinski, 1987; Odrzykoski et al., 1996). We established in vitro culture lines from two samples of each species of *Pellia* and additionally from *Conocephalum conicum* – species J (Marchantiales) (Odrzykoski and Szweykowski, 1991) used as an outgroup in phylogenetic reconstruction.

2.2. In vitro culture

Apical parts of liverworts thalli were sterilized in 0.5% sodium hypochlorite (NaOCl) for 15 min. Sterilized sections of thallus were put on a sterile solid medium containing per one liter: NH$_4$NO$_3$ – 0.12 g, KH$_2$PO$_4$ – 0.7 g, MgSO$_4$ $\times$ 7H$_2$O – 0.246 g, CaCl$_2$ – 0.02 g, FeCl$_3$ – 0.03 g, sucrose – 10 g, agar – 10 g. The agar medium was adjusted to pH 5.6 (Lukavsky et al., 1991; Fiedorow and Szweykovska-Kulinska, 1998a).

2.3. Total genomic DNA isolation

DNA was extracted according to the modified procedure described by Junghans and Metzlaff (Junghans and Metzlaff, 1990; Fiedorow and Szweykovska-Kulinska, 1997). To extract larger amounts of pure DNA, DNeasy Plant Mini Kits and DNeasy Plant Maxi Kits (Qiagen) were used according to the manufacturer’s protocol.

2.4. Polymerase chain reactions (PCR) - amplification of tRNA*Leu* intergenic sequences

The PCR reaction mixture contained: 7.5 ng of total DNA per 10 µl reaction, 0.5 µM of each primer, 10 mM Tris–HCl (pH 8.8), 1.5 mM MgCl$_2$, 50 mM KCl, 0.1% (v/v) Triton X100, 1 mM spermidine (Fiedorow and Szweykovska-Kulinska, 1997), 200 µM of each dNTP, and 0.2 unit of Prime Zyme DNA polymerase (Finnzymes OY).

Primers L1 and L2 were used for PCR amplification of intergenic sequences of tRNA*Leu* genes. Their nucleotide sequences were as follows:

L1 5’ GGTTCAATCCACTTCTCT 3’
L2 5’ GAACTTGAGTCGGGCG 3’

Primers L1 spans the nucleotide sequence identical to the nuclear tRNA*Leu* gene from *Pellia neesiana* (AF056029) between nucleotides 61–78. Primer L2 is complementary to nucleotides 26–41 of the tRNA*Leu* gene from *P. neesiana* (AF056029) (explained in Fiedorow and Szweykovska-Kulinska, 1998b).

PCR was initiated by denaturation at 95°C for 5 min, followed by 30 cycles of: 1 min denaturation at 94°C, 1 min. hybridization at 56°C and 1 min elongation at 72°C. The reaction was terminated by a 5 min elongation cycle at 72°C. Products were analyzed in a 10% polyacrylamide gel.

2.5. Cloning and sequencing of tRNA*Leu* intergenic spacers

Cloning of the PCR products was carried out using the pGEM-T Easy Vector System from Promega, according to the manufacturer’s protocol.

DNA sequencing was carried out using the fmol Sequencing System from Promega, according to the manufacturer’s protocol.
2.6. Restriction analysis of PCR amplified tRNA\textsubscript{Leu} spacers

PCR amplification products were treated with the following restriction enzymes:
- MvaI 3h incubation in 37\degree C (5 units of enzyme per 100 ng of PCR product),
- MaeI 3h incubation in 45\degree C (5 units of enzyme per 100 ng of PCR product).

Both enzymes were produced by Boehringer Mannheim and restriction analysis was carried out according to the manufacturer’s protocol, with buffers provided with the enzymes. Restriction fragments were analyzed in a 10% polyacrylamide gel.

2.7. Computer analyses of DNA sequences data

Sequences data were evaluated with ClustalX program (Thompson et al., 1997) to perform sequences alignment. Further analyses were performed using PAUP v. 3.0s (Swofford, 1991) program.

3. Results

3.1. Identification of proper material for Pellia phylogenetic studies

Total genomic DNA was isolated from the following Pellia species: P. neesiana, P. epiphylla- species N, P. epiphylla- species S, P. borealis, P. endiviifolia- species A and P. endiviifolia- species B. To perform phylogenetic analysis using an outgroup, total DNA from Conocephalum conicum- species J (Marchantiales) was also extracted. DNA from all species was used as templates for PCR reactions to
amplify specific DNA sequences. To evaluate phylogenetic differences and to identify Pellia species, we used spacer regions which separate nuclear tRNA\textsubscript{CAA} genes. We were not able to generate homologous PCR products from the closely related Aneura pinguis. The absence of PCR product from tRNA\textsubscript{CAA} spacer in Aneura pinguis, can be explained in one of two ways: the spacer region is too long for PCR amplification, or less likely there are no tandemly repeated tRNA\textsubscript{CAA} genes in the Aneura pinguis genome. To restrict our analysis to the interspecific differences only and to avoid any interpopulation differences, for each of the species tested we collected material from several geographically distant Polish populations (see Table 1). On average, ten separate populations for each of the species were analyzed using PCR technique. All species used in our studies have been previously identified using electrophoretic separation of isozymes.

3.2. tRNA\textsubscript{Leu} intergenic spacers as the genetic markers

Using PCR primers based on tRNA\textsubscript{Leu} gene sequence from Pellia neesiana, DNA spacers between tRNA\textsubscript{Leu} genes were amplified from the following DNA templates: P. neesiana (Fig. 1, lines 1,2), P. epiphylla- species S (Fig. 1, lines 3,4), P. borealis (Fig. 1, lines 5,6), P. epiphylla- species N (Fig. 1, lines 7,8), P. endiviifolia- species A (Fig. 1, lines 9,10), P. endiviifolia- species B (Fig. 1, lines 11,12) and Conocephalum conicum- species J. All PCR products showed significant size difference. Each of the two PCR products generated from P. borealis template (products (1) and (2)) had the same size as single PCR products amplified separately from P. epiphylla- species-S and P. epiphylla- species N templates, respectively. The same results were obtained when the length of ITS 1 PCR fragments of P. epiphylla- species N, P. epiphylla- species S and P. borealis was compared (data not shown).

As mentioned above, single amplification products of intergenic sequences between tRNA\textsubscript{Leu} genes differ in length in the case of the two cryptic Pellia species: P. epiphylla- species N and P. epiphylla- species S. Two products identical in length to respective P. epiphylla- species S and P. epiphylla- species N single products can be found in P. borealis. Therefore we tested the hypothesis that tRNA intergenic sequences from the two cryptic species are also identical in sequence to amplification products obtained from the polyploid P. borealis template. PCR products from all three taxa were analyzed using the Mael and Mval enzymes. We found that enzyme Mael cuts the PCR product generated from P. epiphylla- species N template (Fig. 2, lines 5,6) and the correspondingly sized PCR product (2) from P. borealis template (Fig. 2, lines 3,4). On the contrary, PCR fragments generated from P. epiphylla- species S DNA (Fig. 2, lines 1,2) and the homologous product (1) obtained from P. borealis DNA (Fig. 2, lines 3,4) were not the substrates for Mael I enzyme. Similar results were obtained using the Mval enzyme, which digests the PCR product amplified from P. epiphylla- species S (Fig. 2, lines 7,8) and correspondingly sized PCR product (1) from P. borealis (Fig. 2, lines 9,10). However, the PCR product from P. epiphylla- species N (Fig. 2, lines 11,12) and the homologous product [2] obtained form P. borealis template (Fig. 2, lines 9,10) were not digested. This result provided strong evidence that the polyploid P. borealis is an allopolyploid formed after hybridization of P. epiphylla- species N, and P. epiphylla- species S.
3.4. Sequence analysis of tRNA<sub>Leu</sub> gene spacer

Amplified intergenic tRNA<sub>Leu</sub> spacers from all studied species were cloned into pGEM-T Easy and sequenced. Since the PCR primers do not hybridize to the 5′ and 3′ ends of the tRNA<sub>Leu</sub> genes, the PCR products contain fragments of two tRNA<sub>Leu</sub> genes sequences at their 5′ and 3′ ends. Transfer RNA<sub>Leu</sub> gene partial sequences from all tested products were identical to respective sequences from tRNA<sub>Leu</sub> genes from Pellia neesiana (Fiedorow and Szweykowska-Kulinśka, 1998b) and Phaseolus vulgaris (Stienberg et al., 1993). The length of obtained sequences of intergenic tRNA<sub>Leu</sub> spacers were as follows: P. neesiana –102 bp; P. epiphylla- species S –104 bp; P. borealis-two products (1) (104 bp) and (2) (101 bp); P. epiphylla-species N –101 bp; P. endiviifolia species A –95 bp; P. endiviifolia- species B –78 bp and C. conicum- species J –355 bp. The intergenic tRNA<sub>Leu</sub> spacers of P. epiphylla-species N and P. borealis [2] are identical in length and have identical nucleotide sequence. In the case of P. epiphylla- species S, respective DNA spacer from P. borealis (1) is the same size and has the same nucleotide sequence except for one nucleotide (in P. borealis, cytidine replaces thymidine at position 10 of the DNA intergenic spacer).

Results of the nucleotide sequence alignment are presented in Fig. 3. Respective sequence amplified from the Conocephalum conicum nuclear genome (AF225700) that was used during this analysis is omitted from the figure. Because of significant length differences of this sequence its presentation in the alignment could cause lower comprehensibility of the figure. Wagner parsimony analysis of the tRNA<sub>Leu</sub> gene spacer sequences, performed using PAUP v.

P. epiphylla-S ATGCCACTTGGTCGCTCCTACTTGGTCGCGGCGACAT---
P. borealis[1] ATGCCACTTGGTCGCTCCTACTTGGTCGCGGCGACAT---
P. borealis[2] ACATACCTTGGTCGCTCCTACTTGGTCGCGGCGACAT---
P. epiphylla-N AGCATACCTTGGTCGCTCCTACTTGGTCGCGGCGACAT---
P. neesiana ATTTAAACTATT---T---T---T---T---T---T---
P. endiviifolia-A TTTTAAACATT---T---T---T---T---T---T---
P. endiviifolia-A TTTTAAACATT---T---T---T---T---T---T---
P. epiphylla-S TTGGAGCTTTTTTCTCTGATTCTGGTTTTGGAATACTACAT---
P. borealis[1] TTGGAGCTTTTTTCTCTGATTCTGGTTTTGGAATACTACAT---
P. borealis[2] TTGGAGCTTTTTTCTCTGATTCTGGTTTTGGAATACTACAT---
P. epiphylla-N TTGGAGCTTTTTTCTCTGATTCTGGTTTTGGAATACTACAT---
P. neesiana TTGGAGCTTTTTTCTCTGATTCTGGTTTTGGAATACTACAT---
P. endiviifolia-B --AGCTGATTTTT---CTAATAATTATATAATTACGCAA---
P. endiviifolia-A AAGCTGATAAAGAGACTACCAAG---GATAATAAATACAT---
P. epiphylla-S CACTATCACATTCTCCAGGTTT---T---T---T---T---T---T---
P. epiphylla-N CACTATCACATTCTCCAGGTTT---T---T---T---T---T---T---
P. neesiana TATTGTGAGCACCTACCAAC---T---T---T---T---T---T---
P. endiviifolia-B TAATTGTGACACCTACCAAC---T---T---T---T---T---T---
P. endiviifolia-A TACTCTAGAACATGCTAA---T---T---T---T---T---T---

P. borealis-S P. epiphylla-N P. neesiana P. endiviifolia-A


Fig. 3. Alignment of Pellia species intergenic tRNA<sub>Leu</sub> sequences using ClustalX program. Sequences derived from polyploid P. borealis have sequences nearly identical to P. epiphylla-N and P. epiphylla-S. The only nucleotide difference between P. epiphylla-S and P. borealis is marked with bold letters.

Fig. 4. Pellia phylogeny inferred from tRNA<sub>Leu</sub> gene spacer sequences. Cladogram of bootstrap analysis using PAUP v. 3.0s. Conocephalum conicum was used as an outgroup. The one shown is the most parsimonious tree, 173 steps, consistency index CI = 0.832, retention index RI = 0.532, rescaled consistency index RC = 0.443. The CI value exceeds the expected 0.776 for the given taxa number (Sanderson and Donoghue, 1989). Bootstrap values (percentage) of 100 branch-and-bound replications are shown at the internal nodes.

3.0 s program (Swofford, 1991), yielded one most parsimonious tree (consistency index 0.832, retention index 0.532, rescaled consistency index 0.443), clustering together spacer sequences from two cryptic species of the P. epiphylla complex with P. neesiana spacer sequence immediately below (Fig. 4). Sequences from two cryptic species of the P. endiviifolia complex are more distantly related and the taxa are paraphyletic. Trees of the same topology were obtained using distance analysis (Kimura 2-parameter model) performed with Neighbor-Joining and Fitch-Margoliash methods (data not shown).

The intergenic regions of some studied liverworts contain a CAA triplet and TATA sequence that are found in many plant nuclear rRNA gene 5′-flanking sequences (Choisne et al., 1997). These sequences are part of the transcription initiation signals for plant RNA polymerase III. Both CAA and TATA sequences were found in spacer regions of P. neesiana and P. endiviifolia- species B. Only CAA appears in the spacer sequence of P. epiphylla- species S and P. borealis (1). The TATA box is found in the tRNA<sub>Leu</sub> spacer of P. endiviifolia- species A. There is no CAA motif in P. epiphylla- species N and P. endiviifolia- species A tRNA<sub>Leu</sub> DNA spacers, and there is no TATA box in the case of P. neesiana. Trees of the same topology were obtained using distance analysis (Kimura 2-parameter model) performed with Neighbor-Joining and Fitch-Margoliash methods (data not shown).

4. Discussion

Our results confirmed phylogenetic and taxonomic hypothesis concerning species delimitation in the liverwort
genus *Pellia*, based on results obtained with isozyme and RAPD (random amplified polymorphic DNA) methods (Odrzykoski et al., 1996; Pacak et al., 1998). Analysis of rRNA*Leu* gene spacer DNA shows large differences between nucleotide sequences of particular pairs of cryptic species *P. epiphylla*-species N, *P. epiphylla*-species S and *P. endiviifolia*-species A, *P. endiviifolia*-species B (Fig. 1).

As the result of rRNA*Leu* gene spacer analysis, one most parsimonious tree for species of the genus *Pellia* was found (Fig. 4). According to these data, species of the *Pellia epiphylla* complex are grouped together and show the closest phylogenetic relation with *P. neesiana*. *P. endiviifolia*-species A and *P. endiviifolia*-species B are classified as two different, more distant species.

Restriction analysis of *Pellia epiphylla* complex gave strong support for the hypothesis that *P. borealis* is an allopolyploid formed by diploidization of a hybrid of the two cryptic species *P. epiphylla*-species N and *P. epiphylla*-species S. Restriction analysis and sequencing of rRNA*Leu* gene intergenic spacers confirmed that *P. epiphylla*-species N and *P. epiphylla*-species S are the most probable donors of nuclear DNA for polyploid *P. borealis*. The allopolyploid character of *P. borealis*, with *P. epiphylla*-species N and *P. epiphylla*-species S as the source of nuclear DNA, is the most probable explanation for the appearance of two different rRNA*Leu* gene spacers in the genome of *P. borealis* with virtually the same sequence as the single sequences found in *P. epiphylla*-species N and *P. epiphylla*-species S. Our results are in full agreement with the conclusion of Odrzykoski et al. (1996), and later studies performed with RAPD markers (Pacak et al., 1998) about the allopolyploid character of *P. borealis* and sources of its nuclear DNA.

Transfer RNA*Leu* genes analysis supports the conclusion that at least in the case of some *Pellia* species, rRNA*Leu* genes arranged in tandems can be transcriptionally active. The spacer sequences that separate the rRNA*Leu* genes are AT rich. Long T-tracts found in this region may function as putative transcription termination signals for polymerase III (Bogenhagen and Brown, 1981). The presence of CAA triplets and TATA box motifs in *P. neesiana* and *P. endiviifolia*-species B spacer sequences also supports the possibility that these genes are transcriptionally active (Choisine et al., 1997). It is difficult to elucidate whether the lack of the TATA motif and CAA triplet imply no transcription activity, or lower transcription activity of rRNA*Leu* genes arranged in tandems in *P. epiphylla*-species N.

Our results confirm the idea that tandemly repeated rRNA genes can be used as taxonomic and phylogenetic markers. All of the genetic markers used on *Pellia* (isozymatic, RAPD and molecular markers: ITS 1 and tandemly repeated rRNA*Leu* genes) supports the hypothesis that *P. epiphylla*-species N and *P. epiphylla*-species S as well as *P. endiviifolia*-species A and *P. endiviifolia*-species B, represent sibling species. Further more *P. borealis* is an allopolyploid species that originated from two cryptic species: *P. epiphylla*-species N and *P. epiphylla*-species S.

**Acknowledgements**

This work was supported by research grant from the State Committee for Scientific Research (KBN 0417/P04/99/16).

**References**


Fiedorow, P., Szewykowska-Kulińska, Z., 1998a. In vitro collection of *Pellia epiphylla*-complex gave two different, more distant species. Further more *P. borealis* is an allopolyploid species that originated from two cryptic species: *P. epiphylla*-species N and *P. epiphylla*-species S. Our results confirm the idea that tandemly repeated rRNA genes can be used as taxonomic and phylogenetic markers. All of the genetic markers used on *Pellia* (isozymatic, RAPD and molecular markers: ITS 1 and tandemly repeated rRNA*Leu* genes) supports the hypothesis that *P. epiphylla*-species N and *P. epiphylla*-species S as well as *P. endiviifolia*-species A and *P. endiviifolia*-species B, represent sibling species. Further more *P. borealis* is an allopolyploid species that originated from two cryptic species: *P. epiphylla*-species N and *P. epiphylla*-species S.


