

Characterization of the complete chloroplast genome of *Pinus uliginosa* (Neumann) from the *Pinus mugo* complex

Konrad Celiński¹  · Hanna Kijak¹ · Jakub Barylski² · Magda Grabsztunowicz^{3,4} · Aleksandra Wojnicka-Półtorak¹ · Ewa Chudzińska¹

Received: 25 October 2016 / Accepted: 15 November 2016 / Published online: 19 November 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Taxonomic status of endangered peat-bog pine, *Pinus uliginosa* (Neumann) classified within the *Pinus mugo* complex, still remains to be elucidated. Here we present a complete chloroplast genome of *P. uliginosa*, to aid resolve its complex systematical position. The total genome size was 119,877 bp in length and contained a total of 112 genes, including 73 protein-coding genes, 35 tRNAs, and four rRNAs. The most of genes occur as a single copy. Five tRNA genes were duplicated from two to four times. Eighteen genes contain one intron, with a single gene containing two introns. No large inverted repeats were identified. The overall G+C content of *P. uliginosa* chloroplast genome is 38.5%.

Keywords *Pinus uliginosa* · Peat-bog pine · Chloroplast genome · Next-generation sequencing

Pinus uliginosa (Neumann) is one of the most controversial, mysterious and threatened taxa from an aggregate of European mountain pines—the *Pinus mugo* complex (Christensen 1987). The origin of this taxa is not accurately

resolved and taxonomical position is still unclear and thus widely discussed in the scientific literature (Hamerník and Musil 2007). Decreasing population size, lack of natural regeneration and ongoing hybridization with the other pines altogether make this taxa threatened by extinction (Wachowiak et al. 2005; Lewandowski and Dering 2006). Highly conservative genome organization, very close relations among the taxa in the *P. mugo* complex, and lack of any diagnostic markers or taxa determinants for *P. uliginosa* were all demonstrated previously using morphological and anatomical traits, polymorphism of allozymes, as well as molecular cytogenetics and flow cytometry. On the other hand, recently reported chemotaxonomic markers for *P. uliginosa* cannot be directly applied for monitoring of gene flow in natural populations or for hybrids identification (Bonikowski et al. 2015; Celiński et al. 2015). Whole-plastid-based barcodes have been shown to have a great potential in species discrimination, especially for closely related taxa where none of single barcodes or their different combinations has sufficient resolution (Li et al. 2015).

Here, we report a complete chloroplast genome sequence of the *P. uliginosa*, determined using the next-generation sequencing method. The assembled and annotated complete nucleotide sequence of *P. uliginosa* chloroplast genome has been deposited in GenBank under accession number KX833097.

Fresh pine needles were collected from a single tree of *P. uliginosa* (accession no. 06_0526_0008) grown in the Dendrological Garden, Poznań University of Life Sciences (52°25'37"N, 16°53'48"E). Intact chloroplasts fraction was isolated from needles according to Vieira et al. (2014). Plastome DNA was extracted using CTAB protocol (Doyle and Doyle 1990). The chloroplast DNA of *P. uliginosa* was fragmented using Covaris E210 (Covaris, Woburn, MA, USA). The library was prepared with a

✉ Konrad Celiński
celinski@amu.edu.pl

¹ Department of Genetics, Adam Mickiewicz University in Poznań, Umultowska 89, 61-614 Poznań, Poland

² Department of Virusology, Adam Mickiewicz University in Poznań, Umultowska 89, 61-614 Poznań, Poland

³ Department of Plant Physiology, Adam Mickiewicz University in Poznań, Umultowska 89, 61-614 Poznań, Poland

⁴ Department of Biochemistry/Molecular Plant Biology, University of Turku, Turun yliopisto, 20014 Turku, Finland

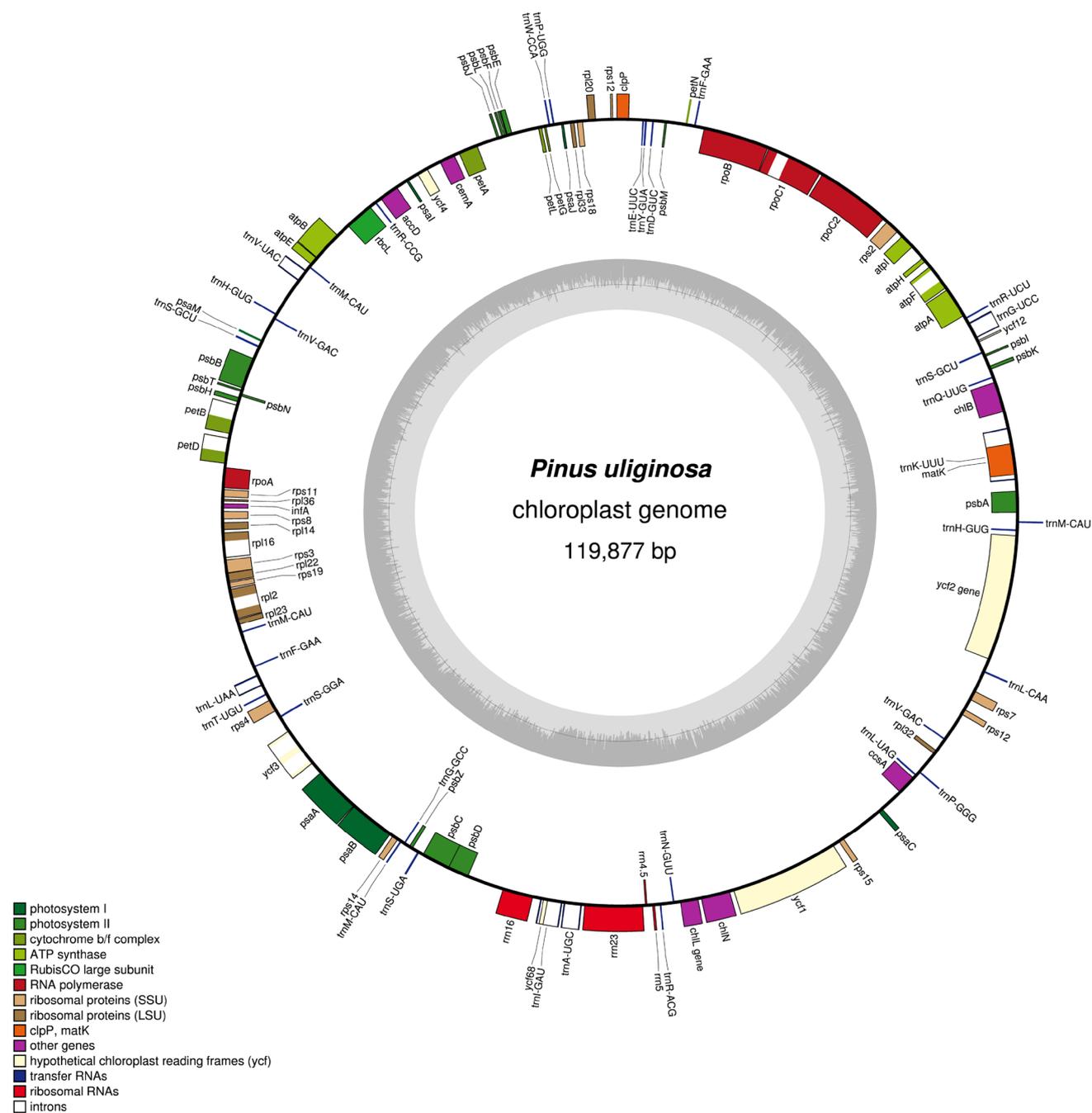


Fig. 1 Chloroplast genome map of *P. uliginosa*. Genes drawn inside the *circle* are transcribed clockwise, and those outside are counter-clockwise. Genes belonging to different functional groups are *color-*

coded. The *darker gray* in the *inner circle* corresponds to GC content, while the *lighter gray* corresponds to AT content

NEBNext DNA Library Prep Kit (New England Biolabs Inc., Ipswich, MA, USA) using dual-indexing TruSeq HT (Illumina Inc., San Diego, CA, USA). Paired-end (PE, 2×250nt) sequencing of a library was performed using Illumina MiSeq (MiSeq Reagent kit v2) at the Genomed sequencing facility (Warsaw, Poland) following manufacturer’s protocols (Illumina, Inc., San Diego, CA, USA). The

automatic primary analysis and the de-multiplexing of the raw reads were performed on MiSeq with the use of MiSeq Reporter (MSR) v2.4 (BaseSpace). In total, 1,355,084 paired-end reads (2×250 bp) were obtained. The obtained reads were trimmed using Cutadapt 1.4 and assembled with CLC Genomics Workbench 7.0.4 in reference based assembly mode (with D17510.1.1—*Pinus thunbergii* chloroplast

Table 1 List of genes present in *P. uliginosa* chloroplast genome

Group of genes	Name of genes
Photosystem I (8)	<i>psaA, B, C, I, J, M, ycf3^b, 4</i>
Photosystem II (16)	<i>psbA, B, C, D, E, F, H, I, J, K, L, M, N, T, Z, ycf12</i>
Cytochrome b/f complex (6)	<i>petA, B^a, D^a, G, L, N</i>
ATP synthase (6)	<i>atpA, B, E, F^a, H, I</i>
Large subunit of Rubisco (1)	<i>rbcL</i>
Chlorophyll biosynthesis (3)	<i>chlB, L, N</i>
Ribosomal proteins large subunit (9)	<i>rpl2^a, 14, 16^a, 20, 22, 23, 32, 33, 36</i>
Ribosomal proteins small subunit (11)	<i>rps2, 3, 4, 7, 8, 11, 12, 14, 15, 18, 19</i>
Subunits of RNA polymerase (4)	<i>rpoA, B, C1^a, C2</i>
Conserved open reading frames (2)	<i>ycf2, 68</i>
Other genes (7)	<i>infA, matK, accD, cemA, clpP, ccsA, ycf1</i>
Ribosomal RNAs (4)	<i>rrn16, 23, 4.5, 5</i>
Transfer RNAs (35)	<i>trnA-TGC^a, trnD-GUC, trnE-UUC, trnF-GAA(x2), trnG-GCC, trnG-UCC^a, trnH-GUG(x2), trnI-GAU^a, trnK-UUU^a, trnL-CAA, trnL-UAA^a, trnL-UAG, trnM-CAU(x4), trnN-GUU, trnP-GGG, trnP-UGG, trnQ-UUG, trnR-ACG, trnR-CCG, trnR-UCU, trnS-GCU(x2), trnS-GGA, trnS-UGA, trnT-UGU, trnV-GAC(x2), trnV-UAC^a, trnW-CCA, trnY-GUA</i>

The numbers in parentheses represents the number of genes

^aGenes containing a single intron

^bGenes containing two introns

DNA as a reference). As a result, 27,419 paired-end reads with the average length 239,22 bp were assembled into 45 contigs with mean coverage of 54,40×. Gaps and ambiguous sites were PCR-amplified and re-sequenced using Sanger method, using 3130x Genetic Analyzer (Applied Biosystems). Protein-coding genes and genes for non-coding RNAs were annotated with DOGMA (Dual Organellar GenoMe Annotator; Wyman et al. 2004) and cpGAVAS (Chloroplast Genome Annotation, Visualization, Analysis, and GenBank Submission Server; Liu et al. 2012). Annotated sequence was aligned to reference genome using MAFFT (Katoh et al. 2002) and all features were manually inspected to cross-validate the predictions. A circular map of the chloroplast genome was drawn with the OGDRAW software (Lohse et al. 2013). The overall GC content was analyzed using Geneious 9.1.2 (Kearse et al. 2012). REPuter (Kurtz et al. 2001) was used to visualize the large (>1000 bp) duplicated sequences in *P. uliginosa* cp genome by forward versus reverse complement (palindromic) alignment.

The complete chloroplast (cp) genome of *P. uliginosa* is 119,877 bp in length (Fig. 1). The overall G+C content of the *P. uliginosa* chloroplast genome is 38.5%, which is similar to the other reported for *Pinaceae* cp genomes (Duan et al. 2016; Fang et al. 2016). The cp genome of *P. uliginosa* contains a total of 112 predicted functional genes including 73 protein-coding genes, 4

ribosomal RNA (rRNA) and 35 transfer RNA (tRNA) genes (Table 1). Five tRNA genes were duplicated from two (*trnF-GAA*, *trnH-GUG*, *trnS-GCU* and *trnV-GAC*) to four times (*trnM-CAU*). Twelve genes (*petB*, *petD*, *atpF*, *rpl2*, *rpl16*, *rpoC1*) and six tRNA genes contain one intron, while one gene (*ycf3*) contains two introns. There were no large inverted repeats (IRs) in the *P. uliginosa* plastome.

Here reported chloroplast genome of *P. uliginosa* provides useful genomic resources for further studies on determination of taxa-diagnostic markers in the *P. mugo* complex, analysis of hybridization and introgression processes in this complex, characterization of genetic diversity of *P. uliginosa* and more effective protection of its threatened gene pools.

Acknowledgements This work was financially supported by the National Science Centre in Poland under Project No. NN304060339. The founders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Funding The funding was provided by Narodowe Centrum Nauki (Grant No. NN304 060339).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a

link to the Creative Commons license, and indicate if changes were made.

References

- Bonikowski R, Celiński K, Wojnicka-Pótorak A, Maliński T (2015) Composition of essential oils isolated from the needles of *P. uncinata* Ramond ex. Mirb and *P. uliginosa* Neumann ex Wimmer grown in Poland. *Nat Prod Commun* 10:371–373
- Celiński K, Bonikowski R, Wojnicka-Pótorak A, Chudzińska E, Maliński T (2015) Volatiles as chemosystematic markers for distinguishing closely related species within the *Pinus mugo* complex. *Chem Biodivers* 12:1208–1213
- Christensen KI (1987) Taxonomic revision of the *Pinus mugo* complex and *P. rhaetica* (*P. mugo* x *sylvestris*) (Pinaceae). *Nord J Bot* 7:383–408
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13–15
- Duan RY, Yang LM, Lv T, Wu GL, Huang MY (2016) The complete chloroplast genome sequence of *Pinus dabeshanensis*. *Conservation Genet Resour*. doi:10.1007/s12686-016-0567-2
- Fang MF, Wang YJ, Zu YM, Dong WL, Wang RN, Deng TT, Li ZH (2016) The complete chloroplast genome of the Taiwan red pine *Pinus taiwanensis* (Pinaceae). *Mitochondr DNA Part A* 27:2732–2733
- Hamerník J, Musil I (2007) The *Pinus mugo* complex—its structuring and general overview of the used nomenclature. *J For Sci* 53:253–266
- Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059–3066
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Mentjies P, Drummond A (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649
- Kurtz S, Choudhuri J, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R (2001) REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res* 29:4633–4642
- Lewandowski A, Dering M (2006) Crossability between *Pinus uliginosa* and its putative parental species *Pinus sylvestris* and *Pinus mugo*. *Silvae Genet* 55:52–54
- Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S (2015) Plant DNA barcoding: from gene to genome. *Biol Rev* 90:157–166
- Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X (2012) CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. *BMC Genomics* 13:715
- Lohse M, Drechsel O, Kahlau S, Bock R (2013) Organellar Genome DRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res* 41:W575–W581
- Vieira LN, Faoro H, Fraga HPF, Rogalski M, Souza EM, Pedrosa FO, Nodari RO, Guerra MP (2014). An improved protocol for intact chloroplasts and cpDNA isolation in conifers. *PLoS One* 9 e84792. doi:10.1371/journal.pone.0084792
- Wachowiak W, Celiński K, Prus-Głowacki W (2005) Evidence of natural reciprocal hybridisation between *Pinus uliginosa* and *P. sylvestris* in the sympatric population of the species. *Flora* 200:563–568
- Wyman SK, Jansen RK, Boore JL (2004) Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* 20:3252–3255