

The complete chloroplast genome sequence of wild service tree *Sorbus torminalis* (L.) Crantz

Bartosz Ulaszewski¹  · Elzbieta Sandurska¹ · Ewa Sztupecka¹ · Jaroslaw Burczyk¹

Received: 20 January 2017 / Accepted: 3 February 2017 / Published online: 14 February 2017
© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract *Sorbus torminalis* is widely distributed tree species across Europe which shows interesting features from the genetic point of view and has high ecological values. Based on the whole genomic DNA we have assembled the complete 160,390 bp circular chloroplast genome of the species. The nucleotide share: 31.35% A, 18.61% C, 17.87% G, 32.12% T and 36.48% GC content is similar to those found in the *Pyrus* and *Malus* species. The genome is a standard quadripartite structure build from four subunits: large (88,029 bp) and small (19,547 bp) single copy unit and two inverted repeats (26,407 bp each). The genome contains 127 genes including: 83 protein-coding genes (77 unique), 36 tRNA genes (29 unique), and 8 rRNA genes (4 unique). Phylogenetic position of *S. torminalis* based on the whole chloroplast genome indicated that the species is phylogenetically closer to genus *Malus* than *Pyrus*.

Keywords *Sorbus torminalis* · Rosaceae · Maloideae · Chloroplast genome · Illumina sequencing · De-novo assembly

Technical note

The genus *Sorbus* consists of approximately 250 species which are spread throughout the northern hemisphere in Asia, North America, Europe and northern Africa (Phipps et al. 1990). One example is wild service tree (*Sorbus*

torminalis [L.] Crantz), which is widely distributed across western, central and southern Europe. However, the species is rare in Britain, Denmark, Germany and Poland, and therefore is being often protected at a local (nature reserves) or national scales (e.g. Poland, Bednorz 2007). Because of its specific features, such as scattered distribution of populations, insect pollination and animal seed dispersal, gametophytic self-incompatibility system, the ability for clonal propagation, and its ecological importance as a driver of biodiversity of forest ecosystems, it has been subjected to research studies often addressing the species gene conservation status (Hoebee et al. 2006; Oddou-Muratorio et al. 2004; Ludwig et al. 2013; Jankowska-Wroblewska et al. 2016). However, genomic resources of *S. torminalis*, which could further contribute to the advancement of population genetic studies of the species are currently limited. In this report we summarize our work on the *S. torminalis* chloroplast genome sequencing, assembly and annotation. The annotated sequence has been submitted to GenBank under accession number KY457242.

The individual wild service tree selected for this study was located in the Jamy Forest District (53°35'40"N, 18°50'59"E), in the north-central part of Poland, at the north-eastern distribution limits of the species. Fresh leaves were collected in late summer of 2016, stored in the dark at 20°C for 48 h to reduce the starch levels samples and allowed to dry. The total gDNA was isolated according to Wang et al. (2013) and send to the Macrogen (Macrogen Inc., Seoul, South Korea) in order to perform library preparation and sequencing on the HiSeq X Illumina system. The chloroplast genome was constructed de novo using NOVO Plasty pipeline v. 1.2.3 (Dierckxsens et al. 2016), based on 10% of the initial 242.4 M raw 150 bp pair-end reads dataset. Complete chloroplast genome sequences of *Pyrus pyrifolia* (GenBank: AP012207) (Terakami et al. 2012) and

✉ Bartosz Ulaszewski
ulaszewski@ukw.edu.pl

¹ Department of Genetics, Faculty of Natural Sciences, Institute of Experimental Biology, Kazimierz Wielki University of Bydgoszcz, Chodkiewicza 30, 85-064 Bydgoszcz, Poland

Malus prunifolia (GenBank: KU851961) (Bao et al. 2016), two closely related species from the *Roseaceae* family, were used as seed sequences in the assembly, which resulted in exactly the same chloroplast genome assemblies of *S. torminalis*. The de novo assembly generated a double-stranded

160,390 bp circular genome from 1,618,116 aligned and 372,372 assembled reads, with final 1523x coverage. The genome annotation was performed using Dual Organellar GenoMe Annotator (DOGMA) (Wyman et al. 2004) and Plann software (Huang and Cronk 2015) with manual

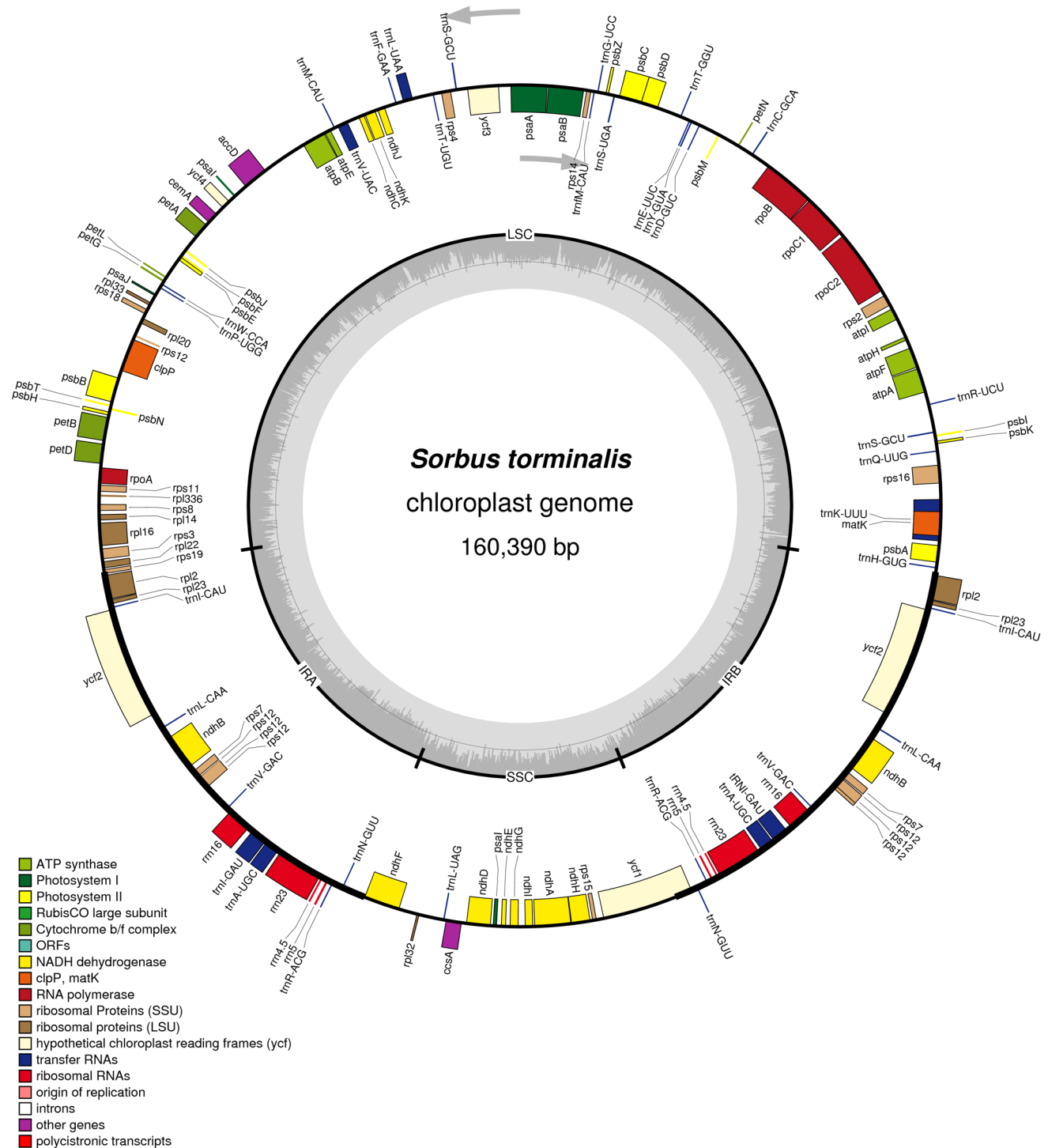
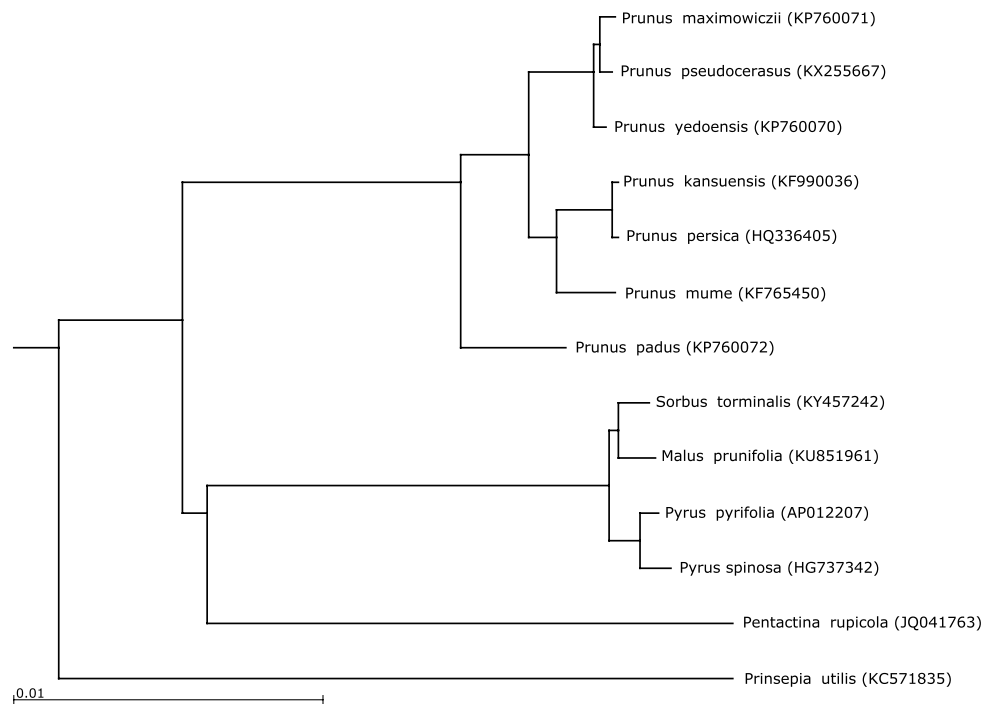


Fig. 1 Visualization of the *Sorbus torminalis* chloroplast genome, with marked genes and transcription directions on the *outer circle*. The *inner circle* highlights contents of G+C (dark grey) and A+T (light grey) and show locations of the LSC, SSC, IRA and IRB regions

Fig. 2 Phylogenetic relationship within the subfamily *Maloideae*, calculated based on available complete chloroplast genome sequences using NJ method



correction using the AP012207 sequence as a reference. Genome was visualized with OrganellarGenomeDRAW (OGDRAW) (Lohse et al. 2007).

The nucleotide share: 31.35% A, 18.61% C, 17.87% G, 32.12% T and 36.48% GC content is similar to those found in *P. pyrifolia* (Terakami et al. 2012) and *M. prunifolia* (Bao et al. 2016). The cpDNA structure is quadripartite and is build by large (LSC, 88,029 bp) and small (SCC, 19,547 bp) singlecopy regions separated by two inverted repeat regions (IRA and IRB, 26,407 bp each) (Fig. 1). The genome contains 127 genes including: protein-coding genes (PCG), transfer RNA (tRNA) and ribosomal RNA (rRNA). PCG group includes 83 genes (77 unique genes), in which 6 of them are in double copies (*ndhB*, *rpl2*, *rpl23*, *rps7*, *rps12*, *ycf2*). Notably, 13 PCGs contain at least one intron (*atpF*, *ndhA*, *ndhB*, *petB*, *petD*, *clpP*, *rpoC1*, *rpl2*, *rpl16*, *rps16*, *rps12*, *ycf3*). The tRNA group includes 36 genes in which 7 of them are in double copies (*trnA-UGC*, *trnI-CAU*, *trnI-GAU*, *trnL-CAA*, *trnN-GUU*, *trnR-ACG*, *trnV-GAC*), and 6 of them contain at least one intron (*trnA-UGC*, *trnG-UCC*, *trnI-GAU*, *trnK-UUU*, *trnL-UAA*, *trnV-UAC*). In the rRNA group all genes are in double copies (*rrn4.5*, *rrn5*, *rrn16*, *rrn23*) without introns.

In order to determine the phylogenetic position of *S. torminalis* based on the whole chloroplast genome sequence we have selected 12 available complete genomes from the *Maloideae* subfamily (Fig. 2). The sequences were aligned using MAFFT software (Katoh and Standley 2013) online server (<http://mafft.cbrc.jp/alignment/server/>) and a neighborhood-joining (NJ) tree was created

using Archeopteryx (Han and Zmasek 2009) integrated with MAFFT server. The resulting phylogenetic tree suggests that *S. torminalis* is more closely related to the genus *Malus* than *Pyrus*, which confirms earlier findings on the phylogeny within *Rosaceae* family (Campbell et al. 2007).

Acknowledgements This work was supported by Ministry of Science and Higher Education in Poland (statutory funds).

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

- Bao L, Li K, Liu Z, Han M, Zhang D (2016) Characterization of the complete chloroplast genome of the Chinese crabapple *Malus prunifolia* (Rosales: Rosaceae: Maloideae). *Cons Genet Resour* 8:227–229. doi:10.1007/s12686-016-0540-0
- Bednorz L (2007) Conservation of genetic resources of *Sorbus torminalis* in Poland. *Dendrobiology* 58:3–7
- Campbell CS, Evans RC, Morgan DR, Dickinson TA, Arsenault MP (2007) Phylogeny of subtribe Pyrinae (formerly the Maloideae, Rosaceae): limited resolution of a complex evolutionary history. *Plant Syst Evol* 266:119–145. doi:10.1007/s00606-007-0545-y
- Dierckxsens N, Mardulyn P, Smits G (2016) NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res* doi:10.1093/nar/gkw955

- Han MV, Zmasek CM (2009) phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinformatics* 10:356. doi:[10.1186/1471-2105-10-356](https://doi.org/10.1186/1471-2105-10-356)
- Hoebbe SE, Menn C, Rotach P, Finkeldey R, Holderegger R (2006) Spatial genetic structure of *Sorbus torminalis*: the extent of clonal reproduction in natural stands of a rare tree species with a scattered distribution. *For Ecol Manag* 226 :1–8. doi:[10.1016/j.foreco.2005.12.024](https://doi.org/10.1016/j.foreco.2005.12.024)
- Huang DI, Cronk QC (2015) Plann: a command-line application for annotating plastome sequences. *Appl Plant Sci* 3:1500026. doi:[10.3732/apps.1500026](https://doi.org/10.3732/apps.1500026)
- Jankowska-Wroblewska S, Meyza K, Sztupecka E, Kubera L, Burczyk J (2016) Clonal structure and high genetic diversity at peripheral populations of *Sorbus torminalis*(L.) Crantz. *iForest* 9:892–900. doi:[10.3832/ifer1885-009](https://doi.org/10.3832/ifer1885-009)
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. doi: [10.1093/molbev/mst010](https://doi.org/10.1093/molbev/mst010)
- Lohse M, Drechsel O, Bock R (2007) OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr Genet* 52:267–274. doi: [10.1007/s00294-007-0161-y](https://doi.org/10.1007/s00294-007-0161-y)
- Ludwig S, Robertson A, Rich TC, Djordjević M, Cerović R, Houston L, Harris SA, Hiscock SJ (2013) Breeding systems, hybridization and continuing evolution in Avon Gorge *Sorbus*. *Ann Bot* 111:563–575. doi:[10.1093/aob/mct013](https://doi.org/10.1093/aob/mct013)
- Oddou-Muratorio S, Demesure-Musch B, Pelissier R, Gouyon PH (2004) Impacts of gene flow and logging history on the local genetic structure of a scattered tree species, *Sorbus torminalis* L. Crantz. *Mol Ecol* 13:3689–3702. doi:[10.1111/j.1365-294X.2004.02373.x](https://doi.org/10.1111/j.1365-294X.2004.02373.x)
- Phipps J, Robertson K, Rohrer J, Smith P (1990) The genus *Sorbus*. *Can J Bot* 68:2209–2269
- Terakami S, Matsumura Y, Kurita K, Kanamori H, Katayose Y, Yamamoto T, Katayama H (2012) Complete sequence of the chloroplast genome from pear (*Pyrus pyrifolia*): genome structure and comparative analysis. *Tree Genet Genome* 8 :841–854. doi: [10.1007/s11295-012-0469-8](https://doi.org/10.1007/s11295-012-0469-8)
- Wang N, Thomson M, Bodles WJ, Crawford RM, Hunt HV, Featherstone AW, Pellicer J, Buggs RJ (2013) Genome sequence of dwarf birch (*Betula nana*) and cross-species RAD markers. *Mol Ecol* 22:3098–3111. doi: [10.1111/mec.12131](https://doi.org/10.1111/mec.12131)
- Wyman SK, Jansen RK, Boore JL (2004) Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* 20:3252–3255. doi: [10.1093/bioinformatics/bth352](https://doi.org/10.1093/bioinformatics/bth352)