




# Complete Genome of “*Candidatus Phytoplasma rubi*” RS, a Phytopathogenic Bacterium Associated with *Rubus* Stunt Disease

Dominik Duceck, <sup>a</sup> Christina Zübert, <sup>a</sup> Jan W. Böhm, <sup>a</sup> Gaia Carminati, <sup>b</sup> Bernd Schneider, <sup>c</sup>  Michael Kube <sup>a</sup>

<sup>a</sup>Department of Integrative Infection Biology Crops-Livestock, University of Hohenheim, Stuttgart, Germany

<sup>b</sup>Department of Agricultural, Food, Environmental, and Animal Sciences, University of Udine, Udine, Italy

<sup>c</sup>Julius Kühn-Institut, Federal Research Centre for Cultivated Plants, Institute for Plant Protection in Fruit Crops and Viticulture, Dossenheim, Germany

**ABSTRACT** The phytoplasma “*Candidatus Phytoplasma rubi*” is associated with *Rubus* stunt disease. The complete genome was determined by assembling Oxford Nanopore Technologies system-derived long reads, with short-read polishing with Illumina reads. The genome of strain RS, from Germany, is organized in one circular chromosome with a length of 762 kb.

The phytoplasma “*Candidatus Phytoplasma rubi*”, a member of the 16SrV group (elm yellows group), is known as the causative agent of *Rubus* stunt disease in raspberry and blackberry cultivars (1). This phytoplasmosis is characterized by stunting, shoot proliferation, virescence, enlarged sepals, phyllody, flower proliferation, and fruit malformations, resulting in the decline of the host and subsequent economic losses (2, 3). Due to the lack of *in vitro* cultivation for phytoplasmas, genome information is critical for understanding the pathogen-host interaction and improving diagnostics. The “*Ca. Phytoplasma rubi*” strain RS was transmitted from a German *Rubus caesius* accession to the experimental host plant *Catharanthus roseus* by the parasitic plant *Cuscuta europaea* and is maintained in *C. roseus* by periodic grafting.

A sequencing library template was generated by cetyltrimethylammonium bromide (CTAB) extraction (4) with cuttings taken from the experimental host. The DNA concentration for generation of the shotgun libraries was determined using a Qubit fluorometer (Thermo Fisher Scientific, Darmstadt, Germany). Long-read sequencing was performed by using the rapid barcoding sequencing kit and a SpotON flow cell R9 on a MinION sequencer (Oxford Nanopore Technologies [ONT], Oxford, UK), which provided 441,170 reads with MinKNOW v22.10.10 (ONT) using the default settings except for the minimal read length, which was set to 1 kb. Short, paired-read sequencing (2 × 150 nucleotides) was provided by a commercial supplier, yielding 12,612,264 pairs (NGSelect DNA library on a NovaSeq 6000 system in paired-end 150-nucleotides read mode; Eurofins Genomics Germany GmbH, Constance, Germany). Reads underwent quality trimming in CLC Genomics Workbench v21.0.5 (Qiagen, Aarhus, Denmark) with default parameters. Read selection for improving the assembly was performed by BLASTN+ v2.9.0 comparison of all ONT reads against a custom-made database comprising “*Ca. Phytoplasma*” and *C. roseus* nucleotide sequence entries in GenBank, following the exclusion of *C. roseus*-assigned reads. The remaining ONT reads ( $N_{50}$  value of 9,337 kb) were assembled in Canu v1.9 (5) with the nanopore-raw option, resulting in a circular and gapless chromosomal sequence. Short-read polishing was performed in CLC Genomics Workbench v21.0.5 (Qiagen) with default parameters. Sequence overlap was identified by BLAST analysis (6) using the contig as query and subject. The chromosome start was set to the *dnaA* gene, and overlaps were removed with Artemis v18.2.0 (7). The chromosomal sequence was automatically annotated in RAST v2.0 (8), inspected by BlastKOALA v2.2 (9), and then manually curated in Artemis v18.2.0 (7). Annotation completeness was inspected by BUSCO v5.2.2 analysis (10), using 151

**Editor** Irene L. G. Newton, Indiana University, Bloomington

**Copyright** © 2023 Duceck et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Michael Kube, michael.kube@uni-hohenheim.de.

The authors declare no conflict of interest.

**Received** 15 December 2022

**Accepted** 10 April 2023

**Published** 26 April 2023

single-copy orthologs from the *Mollicutes* class as a reference. Assembly debris was checked in RAST and against the NCBI nonredundant protein database (<https://blast.ncbi.nlm.nih.gov>). In addition, taxonomical binning of the debris sequences was performed in MEGAN v6.18.2 (11). Unless otherwise stated, the default parameters were used for all software.

The circular chromosome of “*Ca. Phytoplasma rubi*” strain RS consisted of 762,251 bp, with a GC content of 23.13%, and contained 2 rRNA operons, 32 tRNAs, 685 protein-coding genes, and 202 pseudogenes. The coding of pseudogenes is largely associated with the mobilome elements of strain RS. BUSCO analysis revealed completeness of 93%, which is weighted as normal for 16SrV group phytoplasmas. No plasmids were identified.

**Data availability.** Raw reads have been submitted to the Sequence Read Archive (SRA) under accession numbers [SRR22523853](https://www.ncbi.nlm.nih.gov/sra/SRR22523853) and [SRR22523854](https://www.ncbi.nlm.nih.gov/sra/SRR22523854). The whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under BioProject accession number [PRJNA906321](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA906321). The version described in this paper is the first version, accession number [CP114006.1](https://www.ncbi.nlm.nih.gov/nuccore/CP114006.1).

## ACKNOWLEDGMENT

We express our sincere thanks to gardener Felix Hergenbahn for his excellent technical support during the infection trials, transmitting “*Ca. Phytoplasma rubi*” to the experimental host and cultivating the plant material.

## REFERENCES

1. Malembic-Maher S, Salar P, Filippin L, Carle P, Angelini E, Foissac X. 2011. Genetic diversity of European phytoplasmas of the 16SrV taxonomic group and proposal of “*Candidatus Phytoplasma rubi*”. *Int J Syst Evol Microbiol* 61: 2129–2134. <https://doi.org/10.1099/ijs.0.025411-0>.
2. Van der Meer FA. 1987. *Rubus* stunt, p 197–203. In Converse RH (ed), *Virus diseases of small fruits*. Handbook no. 631. U.S. Department of Agriculture, Washington, DC.
3. Mäurer R, Seemüller E. 1996. Witches’ broom of *Rhamnus catharticus*: a new phytoplasma disease. *J Phytopathol* 144:221–223. <https://doi.org/10.1111/j.1439-0434.1996.tb01518.x>.
4. Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12:13–15.
5. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
6. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
7. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16:944–945. <https://doi.org/10.1093/bioinformatics/16.10.944>.
8. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>.
9. Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* 428:726–731. <https://doi.org/10.1016/j.jmb.2015.11.006>.
10. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>.
11. Huson DH, Albrecht B, Bağcı C, Bessarab I, Górka A, Jolic D, Williams RBH. 2018. MEGAN-LR: new algorithms allow accurate binning and easy interactive exploration of metagenomic long reads and contigs. *Biol Direct* 13: 6. <https://doi.org/10.1186/s13062-018-0208-7>.