



# Complete Genome Sequences of Two *Rhodococcus* sp. Strains with Large and Linear Chromosomes, Isolated from Apple Rhizosphere

Sarah Benning,<sup>a</sup> Nicolò Brugnone,<sup>a</sup> Roberto Siani,<sup>a</sup> Susanne Kublik,<sup>a</sup> Michael Schloter,<sup>a,b</sup>  Viviane Radl<sup>a</sup>

<sup>a</sup>Research Unit for Comparative Microbiome Analysis, Helmholtz Center Munich, German Research Center for Environmental Health, Neuherberg, Germany

<sup>b</sup>Chair for Soil Science, Technical University of Munich, Freising, Germany

**ABSTRACT** Members of the genus *Rhodococcus* are usually able to catalyze a number of processes, which are of great interest for ecosystem performance as well as biotechnology. Here, we report the complete genome sequences of two *Rhodococcus* strains that were isolated from rhizosphere soil from an apple orchard in northern Germany.

The genus *Rhodococcus* is known for its immense metabolic diversity, such as the production of a large variety of secondary metabolites, including antibiotics, but also for its capacity to catalyze the degradation of aromatic compounds, for example, pesticides (1, 2). Although *Rhodococcus* is known for its large genome size, with possible linear chromosomes and megaplasmids (3), many of the genes coding for enzymes that catalyze the catabolism and metabolism of organic molecules are located on several large or small, mainly linear plasmids (4).

Here, we report the complete genome sequences of *Rhodococcus* sp. strain R79 and *Rhodococcus koreensis* strain R85, which were isolated from the rhizosphere of young M26 apple rootstocks planted on former grass soil in Ellerhoop, in northern Germany (53°42'51.7"N, 9°45'12.5"E), in 2018 (experimental design as in reference 5). Complete genome sequences should give information on the pathways for the degradation of aromatic compounds, as well as the absence of pathogenicity islands. The isolates were obtained by plating serial dilutions of slurries from rhizosphere soil on actinomycete isolation agar (Sigma-Aldrich, Darmstadt, Germany) and were selected for their ability to grow in minimal medium with benzoic acid as the only carbon source. Genomic DNA for sequencing was obtained using a kit-based protocol (Genomic-tip 20/G kit; Qiagen, Hilden, Germany), after cultivating isolates for 2 days at 28°C in actinomycete broth (Sigma-Aldrich). The genomes were sequenced on a Sequel platform (Pacific Biosciences [PacBio], Menlo Park, CA). Genomic DNA was sheared according to the protocol for preparing multiplexed microbial libraries using SMRTbell Express template preparation kit v2.0 with Covaris g-TUBEs and then was concentrated using AMPure PB beads (PacBio). Libraries were prepared using the aforementioned SMRTbell Express template preparation kit v2.0 and the barcoded overhang adapter kit 8A/8B and then were loaded onto a single-molecule real-time (SMRT) cell following the instructions for diffusion loading (PacBio). No size selection was performed here because this would result in possible plasmid loss.

The demultiplexed data from sequencing were assembled using the Microbial Assembly pipeline implemented in SMRT Link v8.0.0.80529 (PacBio), which includes the circularization process. Parameters differing from the default settings were as follows: for R79, seed coverage of 40 and estimated genome size of 8.7 Mb; for R85, seed coverage of 35 and estimated genome size of 8.9 Mb. No rotation of the contigs was performed. The resulting contigs were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (6). Quality and completeness were controlled with

**Citation** Benning S, Brugnone N, Siani R, Kublik S, Schloter M, Radl V. 2021. Complete genome sequences of two *Rhodococcus* sp. strains with large and linear chromosomes, isolated from apple rhizosphere. *Microbiol Resour Announc* 10:e00159-21. <https://doi.org/10.1128/MRA.00159-21>.

**Editor** Julie C. Dunning Hotopp, University of Maryland School of Medicine

**Copyright** © 2021 Benning 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 Viviane Radl, [viviane.radl@helmholtz-muenchen.de](mailto:viviane.radl@helmholtz-muenchen.de).

**Received** 1 March 2021

**Accepted** 24 May 2021

**Published** 17 June 2021

**TABLE 1** Characteristics of genomes of strains R79 and R85

Strain and contig	Contig length (bp)	GC content (%)	Topology	GenBank accession no.
R79				
Chromosome	8,730,774	67.5	Linear	CP070619
Plasmid 1	468,864	64.9	Linear	CP070618
Plasmid 2	320,146	60.1	Linear	CP070617
Plasmid 3	167,009	64.7	Linear	CP070616
Plasmid 4	102,288	65.0	Circular	CP070615
Plasmid 5	77,479	65.7	Linear	CP070614
R85				
Chromosome	8,983,208	67.7	Linear	CP070609
Megaplasmid	1,241,328	64.8	Linear	CP070610
Plasmid 2	502,657	64.5	Linear	CP070611
Plasmid 3	126,147	65.0	Circular	CP070612
Plasmid 4	56,535	64.0	Linear	CP070613

CheckM v1.0.18 (7) as implemented in KBase (8). Taxonomy was inferred using the Type (Strain) Genome Server (TGYS) (9). Default parameters were used except where otherwise noted.

The number of raw polymerase reads for the SMRT cell was 388,809 reads. The numbers of polymerase reads and  $N_{50}$  values from the subread statistics were 44,862 reads and 7,477 bp, respectively, for R79 and 69,894 reads and 6,877 bp, respectively, for strain R85. For both strains, the completeness of the genomes was 99.7%. Strain R79 had an overall genome size of 9,866,560 bp (GC content of 67.28%), with one linear chromosome, four linear plasmids, and one circular plasmid (Table 1). In total, 9,028 genes were detected for the genome of R79. For strain R85, we obtained an even larger genome size of 10,909,875 bp (GC content of 67.18%), with one linear chromosome, one linear megaplasmid, two linear plasmids, and one circular plasmid (Table 1). Here, 9,955 genes in total were detected. From these genes, we conclude that both strains are potentially able to metabolize aromatic compounds and to produce nonribosomal peptides and polyketides. For both strains, we observed more than 6,000 predicted protein features related to secondary metabolism, which again shows the potential of *Rhodococcus* strains for bioremediation purposes. They partly have the potential to support plant growth. Thus, the two *Rhodococcus* strains might be used in the future as biostimulants for plants.

**Data availability.** The full genome sequences and the raw reads for R79 and R85 have been deposited in the NCBI database under BioProject number [PRJNA700828](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA700828). The genomes can be found in GenBank with the accession numbers [CP070614](https://www.ncbi.nlm.nih.gov/nuccore/CP070614) to [CP070619](https://www.ncbi.nlm.nih.gov/nuccore/CP070619) for R79 and [CP070609](https://www.ncbi.nlm.nih.gov/nuccore/CP070609) to [CP070613](https://www.ncbi.nlm.nih.gov/nuccore/CP070613) for R85. The raw reads have been deposited in the Sequence Read Archive (SRA) with the accession numbers [SRX10094454](https://www.ncbi.nlm.nih.gov/sra/SRX10094454) for R79 and [SRX10094455](https://www.ncbi.nlm.nih.gov/sra/SRX10094455) for R85.

## ACKNOWLEDGMENTS

We thank the German Federal Ministry of Education and Research (BMBF) for funding through the funding initiative BonaRes (grant 031B0512).

Warm thanks also go to the colleagues from the ORDIAmur project for providing soil from their experimental field site.

## REFERENCES

- Bell KS, Philp JC, Aw DW, Christofi N. 1998. The genus *Rhodococcus*. *J Appl Microbiol* 85:195–210. <https://doi.org/10.1046/j.1365-2672.1998.00525.x>.
- Larkin MJ, Kulakov LA, Allen CCR. 2006. Biodegradation by members of the genus *Rhodococcus*: biochemistry, physiology, and genetic adaptation. *Adv Appl Microbiol* 59:1–29. [https://doi.org/10.1016/S0065-2164\(06\)59001-X](https://doi.org/10.1016/S0065-2164(06)59001-X).
- Larkin MJ, Kulakov LA, Allen CCR. 2010. Genomes and plasmids in *Rhodococcus*. *Microbiol Monogr* 16:73–90. [https://doi.org/10.1007/978-3-642-12937-7\\_3](https://doi.org/10.1007/978-3-642-12937-7_3).
- Gürtler V, Seviour RJ. 2010. Systematics of members of the genus *Rhodococcus* (Zopf 1891) emend Goodfellow et al. 1998. *Microbiol Monogr* 16:1–28. [https://doi.org/10.1007/978-3-642-12937-7\\_1](https://doi.org/10.1007/978-3-642-12937-7_1).

5. Mahnkopp F, Simon M, Lehndorff E, Pätzold S, Wrede A, Winkelmann T. 2018. Induction and diagnosis of apple replant disease (ARD): a matter of heterogeneous soil properties? *Sci Horticul* 241:167–177. <https://doi.org/10.1016/j.scienta.2018.06.076>.
6. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the Prokaryotic Genome Annotation Pipeline reach with protein family model curation. *Nucleic Acids Res* 49:D1020–D1028. <https://doi.org/10.1093/nar/gkaa1105>.
7. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
8. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia JM, Chia JM, Colasanti R, Conrad N, Davis JJ, Davison BH, Dejongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Meyer F, Mills M, Novichkov PS, Oh T, Olsen GJ, Olson R, Parrello B, Pasternak S, Pearson E, Poon SS, Price GA, Ramakrishnan S, Ranjan P, Ronald PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S, Yu D. 2018. KBase: the United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol* 36:566–569. <https://doi.org/10.1038/nbt.4163>.
9. Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 10:2182. <https://doi.org/10.1038/s41467-019-10210-3>.