

Whole-genome sequencing of four culturable endophytic bacteria from German hardneck garlic cloves (*Allium sativum* L.)

Girish Kumar,¹ Han Ming Gan,^{2,3} Peter Wengert,¹ Trevor Penix,¹ Anuththaman Parthasarathy,⁴ André O. Hudson,¹ Michael A. Savka¹

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT We present the whole-genome sequence of four bacterial endophytes associated with German hardneck garlic cloves (*Allium sativum* L.). Among them, *Agrobacterium fabrum* and *Pantoea agglomerans* are associated with plant protection, while *Rahnella perminowiae* and *Stenotrophomonas lactitubi* are pathogens. These data will facilitate the identification of genes to improve garlic.

KEYWORDS *Allium sativum*, German hardneck garlic cloves, endophyte genomes, disease resistant, pathogenicity

Endophytes play a vital role in plant growth, development, and stress defense without causing noticeable changes or infection (1). Garlic (*Allium sativum* L.) is a significant crop renowned for its culinary and medicinal properties and as a source of endophytic microorganisms (2). In recent studies, bacterial endophytes that are linked with garlic roots (3) and black garlic processing (4) have been isolated and identified. However, the isolation, identification, and whole-genome sequencing of bacterial endophytes from hardneck garlic cloves have not been attempted previously. In this research, we have successfully isolated and sequenced four bacterial endophytes that are associated with garlic cloves.

Mature field grown hardneck German garlic bulbs were obtained from Bischofing Farm, Rush (Longitude: -77.6374368; Latitude: 43.01495449999999), New York in August. The cloves were separated from the bulbs, peeled and subjected to surface sterilization using a 20% sodium hypochlorite (NaOCl) solution, followed by two consecutive rinses with sterile distilled water for the isolation of cultivable bacterial endophytes. The sterilized cloves were axenically chopped and then introduced into tryptic soy broth (TS) liquid medium and incubated at 28°C for 3 days. The resulting bacterial culture was 10-fold serial diluted and then plated onto TS agar media and incubated under identical conditions. Subsequently, four separate single colonies with distinct morphologies were obtained and subcultured on TS agar medium and incubated at 28°C for 48 hours to ensure strain purity.

The E.Z.N.A. bacterial DNA kit (Omega Bio-Tek, Norcross, GA, USA) was used to extract genomic DNA from 25 mg of pelleted cells. The DNA concentrations were then measured with a Qubit 3.0 fluorometer (Life Technologies, MD, USA). Subsequently, sequencing libraries were created following the Nextera XT library preparation protocols using 0.2 ng/μL of DNA and sequenced on the MiSeq sequencer (Illumina, San Diego, CA, USA) utilizing a MiSeq reagent V3 kit (2 × 300 cycles), following the manufacturer's guidelines at the Genomics Lab, Rochester Institute of Technology. Raw reads were subjected to quality control processing with the fastp version 0.23.2 (5) and then *de novo* assembled into contigs using the SPAdes genome assembler (6) (version 3.15). The NCBI Prokaryotic Genome Annotation Pipeline (7) (version 6.5, http://www.ncbi.nlm.nih.gov/genome/annotation_prok/) was used to automatically annotate the genomes upon upload. Default parameters were used for all software unless otherwise specified.

Editor Jason E. Stajich, University of California Riverside, Riverside, California, USA

Address correspondence to Michael A. Savka, massbi@rit.edu.

The authors declare no conflict of interest.

Received 12 December 2023

Accepted 28 February 2024

Published 12 March 2024

Copyright © 2024 Kumar 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/).

TABLE 1 Information on genome annotation of four bacterial endophytes associated with garlic (*Allium sativum* L.) cloves^a

| Sample | No. of raw reads | SRA accession | Assembly accession | Assigned taxonomy | Reference genome with accession number | % ANI | % cov | Assembly size (bp) | Coverage (×) | No. of contigs | N ₅₀ | %GC | No. of genes |
|----------|------------------|---------------|--------------------|-----------------------------------|--|-------|-------|--------------------|--------------|----------------|-----------------|-------|--------------|
| RIT-As-1 | 53,314,682 | SRR22245962 | GCA_0262419_95.1 | <i>Rahnella perminowiae</i> | GCA_0190497_55.1 | 99.28 | 91.23 | 5,449,955 | 10.62 | 534 | 16,306 | 51.56 | 5,376 |
| RIT-As-2 | 53,672,656 | SRR22245961 | GCA_0262420_05.1 | <i>Stenotrophomonas lactitubi</i> | GCA_0028035_15.1 | 95.54 | 85.69 | 4,527,262 | 12.22 | 245 | 28,938 | 66.06 | 4,267 |
| RIT-As-3 | 269,038,734 | SRR22245960 | GCA_0262419_55.1 | <i>Agrobacterium fabrum</i> | GCA_0000920_25.1 | 99.99 | 99.96 | 5,250,432 | 47.41 | 33 | 314,096 | 59.15 | 4,967 |
| RIT-As-4 | 246,616,136 | SRR22245959 | GCA_0262419_65.1 | <i>Pantoea agglomerans</i> | GCA_0037102_45.1 | 98.74 | 89.72 | 4,974,910 | 71.38 | 28 | 451,856 | 54.84 | 4,669 |

^a%ANI, percent average nucleotide identity; % cov, percent genome query coverage.

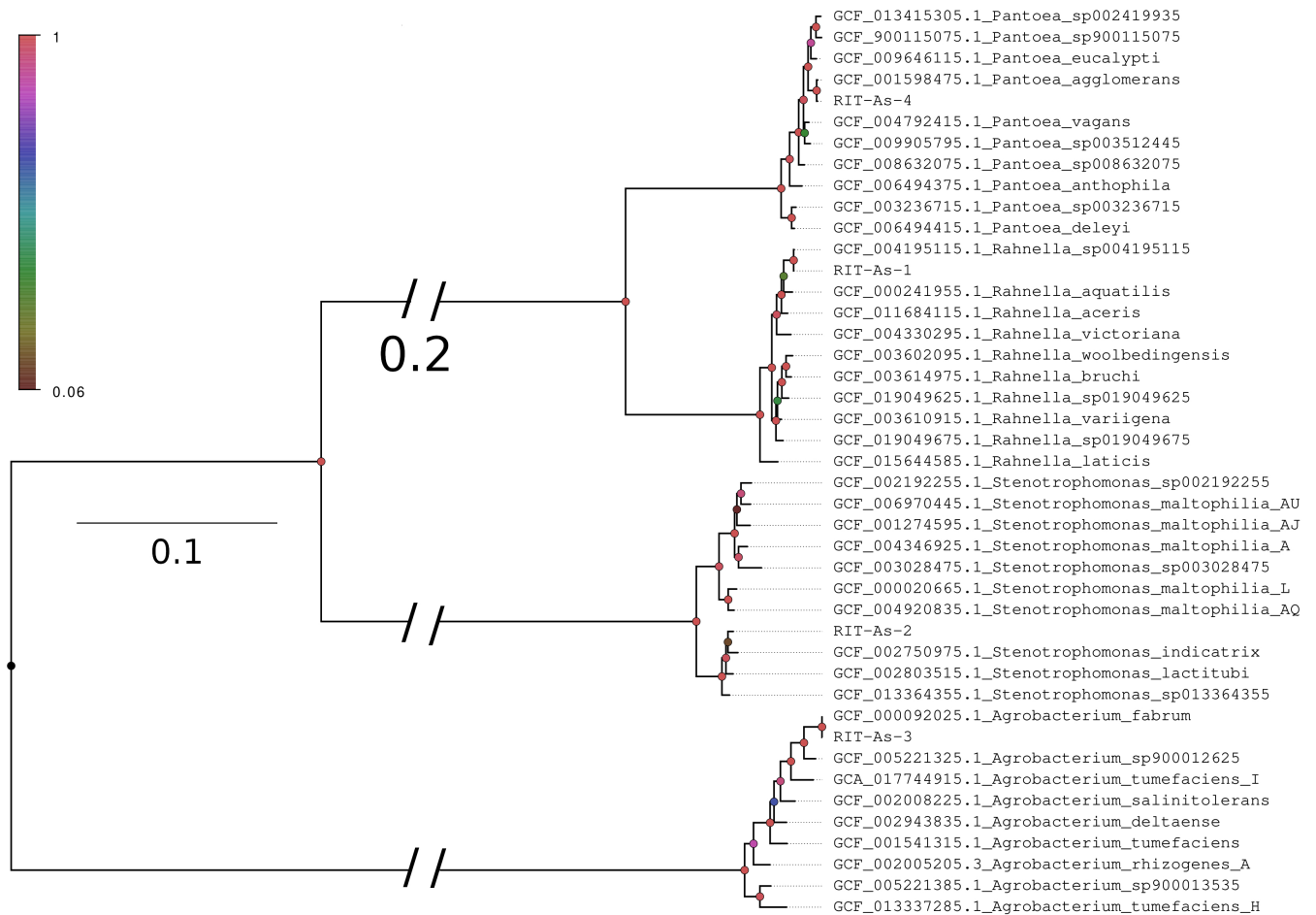


FIG 1 Maximum likelihood tree depicting the evolutionary relationships among the isolated garlic endophytes (blue tip label) and their closely related representative strains. The tree was constructed using FastTree version 2.1.10 (12) based on the concatenated alignments of highly conserved Proteobacteria proteins produced from the GToTree v1.7.07 pipeline (13). Briefly, the pipeline identified and aligned a total of 119 single-copy proteins from each genome annotation using hmmsearch v 3.3.2 (14) and muscle v 3.8.1551 (15), respectively. Nodes were colored according to their SH-like support values, and branch lengths indicate the number of substitutions per site. Branches leading to the major taxa are shortened to improve readability.

All four endophyte species were phylogenetically placed into their respective genera clades in the constructed maximum likelihood tree, corroborating with their initial taxonomic assignment based on average nucleotide identity (ANI; Fig. 1). Of the four genomes sequenced (Table 1), two bacterial endophyte species (*Agrobacterium fabrum* and *Pantoea agglomerans*) confer protection to plants against abiotic stress (8, 9). On the other hand, endophytes *Rahnella perminowiae* and *Stenotrophomonas lactitubi* have been found to be pathogens (10, 11). The whole-genome sequencing data presented in this study can assist in identifying genes associated with endophyte-mediated resistance and pathogenicity in garlic, thereby contributing to the progress of garlic cultivation.

ACKNOWLEDGMENTS

We are grateful for the ongoing support from the Thomas H. Gosnell School of Life Sciences (GSoLS) and the College of Science (COS) at the Rochester Institute of Technology (RIT).

AUTHOR AFFILIATIONS

¹Thomas H. Gosnell School of Life Sciences, Rochester Institute of Technology, Rochester, New York, USA

²Department of Biological Sciences, Sunway University, Petaling Jaya, Malaysia

³Patriot Biotech Sdn. Bhd., Subang Jaya, Malaysia

⁴School of Chemistry and Biosciences, University of Bradford, West Yorkshire, United Kingdom

AUTHOR ORCID*s*

André O. Hudson  <https://orcid.org/0000-0001-5690-4322>

Michael A. Savka  <http://orcid.org/0000-0003-1328-2978>

AUTHOR CONTRIBUTIONS

Girish Kumar, Data curation, Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review and editing | Han Ming Gan, Conceptualization, Formal analysis, Visualization, Writing – original draft, Writing – review and editing | Peter Wengert, Data curation, Methodology, Writing – original draft, Writing – review and editing | Trevor Penix, Data curation, Methodology, Writing – original draft | Anutthaman Parthasarathy, Data curation, Methodology, Writing – original draft | André O. Hudson, Data curation, Methodology, Writing – original draft, Writing – review and editing | Michael A. Savka, Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Resources, Supervision, Visualization, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

The whole genome sequences reported in this article were deposited in GenBank, and accession numbers for each SRA and whole-genome assembly are available in Table 1.

REFERENCES

- Pacifico D, Squartini A, Crucitti D, Barizza E, Lo Schiavo F, Muresu R, Carimi F, Zottini M. 2019. The role of the endophytic microbiome in the grapevine response to environmental triggers. *Front Plant Sci* 10:1256. <https://doi.org/10.3389/fpls.2019.01256>
- Lanzotti V, Scala F, Bonanomi G. 2014. Compounds from *Allium* species with cytotoxic and antimicrobial activity. *Phytochem Rev* 13:769–791. <https://doi.org/10.1007/s11101-014-9366-0>
- Costa Júnior PSP, Cardoso FP, Martins AD, Teixeira Buttrós VH, Pasqual M, Dias DR, Schwan RF, Dória J. 2020. Endophytic bacteria of garlic roots promote growth of micropropagated meristems. *Microbiol Res* 241:126585. <https://doi.org/10.1016/j.micres.2020.126585>
- Qiu Z, Lu X, Li N, Zhang M, Qiao X. 2018. Characterization of garlic endophytes isolated from the black garlic processing. *Microbiologyopen* 7:e00547. <https://doi.org/10.1002/mbo3.547>
- Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. Spades: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>
- Zafar-ul-Hye M, Danish S, Abbas M, Ahmad M, Munir TM. 2019. ACC deaminase producing PGPR *Bacillus amyloliquefaciens* and *Agrobacterium fabrum* along with biochar improve wheat productivity under drought stress. *Agronomy* 9:343. <https://doi.org/10.3390/agronomy9070343>
- Dunleavy JM. 1989. *Curtobacterium plantarum* sp. nov. is ubiquitous in plant leaves and is seed transmitted in soybean and corn. *Int J Syst Evol Microbiol* 39:240–249. <https://doi.org/10.1099/00207713-39-3-240>
- Brady C, Asselin JA, Beer S, Brurberg MB, Crampton B, Venter S, Arnold D, Denman S. 2022. *Rahnella perminowiae* sp. nov., *Rahnella bonaserana* sp. nov., *Rahnella rivi* sp. nov. and *Rahnella ecdela* sp. nov., isolated from diverse environmental sources, and emended description of the genus *Rahnella*. *Int J Syst Evol Microbiol* 72:005190. <https://doi.org/10.1099/ijsem.0.005190>
- Bian D-R, Xue H, Piao C-G, Li Y. 2020. *Stenotrophomonas cyclobalanopsis* sp. nov., isolated from the leaf spot disease of *Cyclobalanopsis patelliformis*. *Anton Leeuw* 113:1447–1454. <https://doi.org/10.1007/s10482-020-01453-y>
- Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>
- Lee MD. 2019. GToTree: a user-friendly workflow for phylogenomics. *Bioinformatics* 35:4162–4164. <https://doi.org/10.1093/bioinformatics/btz188>
- Johnson LS, Eddy SR, Portugaly E. 2010. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* 11:1–8. <https://doi.org/10.1186/1471-2105-11-431>
- Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:1–9. <https://doi.org/10.1186/1471-2105-5-113>