Draft Genome Sequence of *Duganella* sp. Strain DN04, Isolated from Cultivated Soil

Rachel Raths,ᵃ Vincent Peta,ᵃ Heike Bückingᵃ

ᵃBiology and Microbiology Department, South Dakota State University, Brookings, South Dakota, USA

ABSTRACT Here, we sequenced *Duganella* sp. strain DN04, a novel species within the genus *Duganella* that was isolated from a maize field in North Carolina. The assembled draft genome size is 6,562,230 bp, with a total of 6,039 protein coding sequences and 3,889 functionally assigned genes, including genes putatively involved in the colonization of plants.

The genus *Duganella* within the family *Oxalobacteraceae* was identified in 1997 (1) and consists of Gram-negative, motile, aerobic, mesophilic bacteria that are mainly found in soils. Most *Duganella* strains produce the bis-indole pigment violacein, which has antifungal and antibacterial properties (2, 3). Other *Duganella* strains show proteolytic and lipolytic activities (4).

Duganella sp. strain DN04 was isolated from a maize field in North Carolina on 9 June 2016 (geographical coordinates, 36.1034, −78.4114; soil pH, 5.8). Soil was added to phosphate-buffered saline, and dilutions were plated on Reasoner's 2A (R2A) agar plates and incubated at 30°C for 1 to 2 days, followed by 20°C for 1 to 2 days. Repeated colony transfers and streaks were performed on R2A agar to acquire pure colonies. Genomic DNA was extracted from a freshly grown R2A culture using the AllPrep bacterial DNA/RNA/protein kit (Qiagen, Inc., Germantown, MD) following the kit protocol. A genomic library was prepared using a Nextera kit (Illumina, San Diego, CA), size selected to an average fragment length of 475 bp, and sequenced using Illumina NextSeq paired-end v2 chemistry on v2.5 flow cells at 150 bp per read. A target coverage of 20× was used, and the genome was assembled using SPAdes v3.13.0 (5). Default parameters were used for all software unless otherwise specified. The genome was screened for possible contamination by a BLAST search of annotated coding regions that represented at least 60% of the contigs against a diverse set of genomes. If less than 10% of the hits were to proteins from other species in the same family, the contamination was considered to be low.

A total of 11,788,390 reads were obtained, with a total read length of 1,738,895,421 bp and an average read length of 148 bp. The genome length was 6,562,230 bp. In total, we found 281 contigs with an N_{50} value of 40,161 bp (range, 671 to 141,855 bp) and an L_{50} value of 49. Assembly quality assessment using Benchmarking Universal Single-Copy Orthologs (BUSCO) (6) revealed a measured completeness (40 single-copy BUSCOs) of 100%. Following genome annotation by PATRIC v3.5.28 (7), we identified a total of 6,039 protein coding sequences with 2,150 hypothetical genes and 3,889 genes with functional assignments. Strain DN04 has 48 tRNA genes, 4 rRNA genes, 43 antibiotic resistance genes, and a GC content of 64.4%. Based on 16S rRNA, *Duganella sacchari* strain Sac-22 is the closest related species (98.82% identity).

Galaxy and RAST v2.0 were used to annotate and identify specific genes (8, 9). We discovered several genes that are putatively involved in plant growth promotion, including genes needed to catalyze the decomposition of hydrogen peroxide (*katE*) and genes involved in the production of urease (*ureA* to *ureG*), biofilms (*bdcA* and *wspC*)

(10), and biotin (*bioA* to *bioD* and *bioF*) (11), several genes of the *pst* operon, and a two-component signal transduction system involved in phosphate uptake (*phoR*, *phoB*, *phoD*, *pstS*, *pstC*, *pstA*, *pstB*, *phoU*, and *ppk*) (12). We also identified 60 putative virulence genes, including 45 antibiotic and toxic compound resistance genes and 15 invasion and intracellular resistance genes.

Data availability. The complete genome sequence described here has been deposited in NCBI/GenBank under BioProject number PRJNA529278, BioSample number SAMN11263562, accession number NZ_SPVG00000000, and SRA number SRX6098754.

ACKNOWLEDGMENTS

REFERENCES

1. Hiraishi A, Shin YK, Sugiyama J. 1997. Proposal to reclassify *Zoogloea ramigera* IAM 12670 (P. R. Dugan 115) as *Duganella zoogloeoides* gen. nov., sp. nov. Int J Syst Bacteriol 47:1249–1252. https://doi.org/10.1099/00207713-47-4-1249.

2. Haack FS, Poehlein A, Kröger C, Voigt CA, Piepenbring M, Bode HB, Daniel R, Schäfer W, Streit WR. 2016. Molecular keys to the *Janthinobacterium* and *Duganella* spp. interaction with the plant pathogen *Fusarium graminearum*. Front Microbiol 7:1668. https://doi.org/10.3389/fmicb.2016.01668.

3. Choi SY, Kim S, Lyuck S, Kim SB, Mitchell RJ. 2015. High-level production of violacein by the newly isolated *Duganella violaceinigra* str. Sci Rep 5:15598. https://doi.org/10.1038/srep15598.

4. Aranda S, Montes-Borrego M, Landa BB. 2011. Purple-pigmented violacein-producing *Duganella* spp. inhabit the rhizosphere of wild and cultivated olives in southern Spain. Microb Ecol 62:446–459. https://doi.org/10.1007/s00248-011-9840-9.

5. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski 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.

6. Simão 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.

7. Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, Conrad N, Dietrich EM, Disz T, Gabbard JL, Gerdes S, Henry CS, Kenyon RW, Machi D, Mao C, Nordberg EK, Olsen GJ, Murphy-Olson DE, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Vonstein V, Warren A, Xia FF, Yoo H, Stevens RL. 2017. Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. Nucleic Acids Res 45:D535–D542. https://doi.org/10.1093/nar/gkw1017.

8. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Grüning BA, Guerler A, Hillman-Jackson J, Hiltemann S, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res 46:W537–W544. https://doi.org/10.1093/nar/gky379.

9. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, de Crécy-Lagard V, Diaz N, Disz T, Edwards R, Fonstein M, Frank ED, Gerdes S, Glass EM, Goesmann A, Hanson A, Iwata-Reuyl D, Jensen R, Jamshidi N, Krause L, Kubal M, Larsen N, Linke B, McHardy AC, Meyer F, Neuweger H, Olsen G, Olson R, Osterman A, Portnoy V, Pusch GD, Rodionov DA, Rückert C, Steiner J, Stevens R, Thiele I, Vassieva O, Ye Y, Zagnitko O, Vonstein V. 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res 33:5691–5702. https://doi.org/10.1093/nar/gki866.

10. Barraud N, Kjelleberg S, Rice SA. 2015. Dispersal from microbial biofilms. Microbiol Spectrum 3:MB-0015-2014. https://doi.org/10.1128/microbiolspec.MB-0015-2014.

11. Streit WR, Joseph CM, Phillips DA. 1996. Biotin and water-soluble vitamins are key growth factors for alfalfa root colonization by *Rhizobium meliloti* 1021. Mol Plant Microbe Interact 9:330–338. https://doi.org/10.1094/MPMI-9-0330.

12. Asaf S, Khan AL, Khan MA, Al-Harrasi A, Lee I-J. 2018. Complete genome sequencing and analysis of endophytic *Sphingomonas* sp. LK11 and its potential in plant growth. 3 Biotech 8:389. https://doi.org/10.1007/s13205-018-1403-z.