




Draft Genome Sequences of Seven *Chryseobacterium* Type Strains

Andisiwe Matu,^a Adeline Lum Nde,^a Lize Oosthuizen,^a Arina Hitzeroth,^{a*} Moira Badenhorst,^a Lebogang Duba,^a Mphumzi Gidaga,^a Johan Klinck,^a Ina-Maré Kriek,^a Palesa June Lekoma,^a Lynette Nel,^a Sérgio Mendes dos Ramos,^a Jacques Rossouw,^a Nicolaas Salomane,^a Neo Segone,^a Shiela Serobe,^a Tarsisius Tiyani,^a Celia J. Hugo,^a  Jeffrey D. Newman^{a,b}

^aDepartment of Microbial, Biochemical and Food Biotechnology, University of the Free State, Bloemfontein, Free State, South Africa

^bBiology Department, Lycoming College, Williamsport, Pennsylvania, USA

ABSTRACT In an honors course on “Omics Sciences,” draft genome sequences of *Chryseobacterium elymi* KCTC 22547^T, *Chryseobacterium flavum* KCTC 12877^T, *Chryseobacterium hispanicum* KCTC 22104^T, *Chryseobacterium lathyri* KCTC 22544^T, “*Candidatus Chryseobacterium massiliae*” CCUG 51329^T, *Chryseobacterium piscium* CCUG 51923^T, and *Chryseobacterium rhizosphaerae* KCTC 22548^T were generated to facilitate phylogenomic comparisons within the genus.

Chryseobacterium species are chemoorganotrophic, Gram-negative rods with flexirubin pigments that give the colonies a yellow-orange color. The genus contains more than 100 described species from diverse habitats, including freshwater sources, soil, marine fish, and human hosts (1–6). Our laboratories have isolated many *chryseobacteria*, including novel strains, from freshwater creeks and from poultry.

To determine whether a bacterium is a member of an existing species or a novel taxon, it should be compared to previously described species at the genomic level in which an average nucleotide identity (ANI) below 95% signifies that they are different species (7). Genome sequences are now required to describe new taxa, and the Genomic Encyclopedia of Bacteria and Archaea (GEBA) project is sequencing type strain genomes for previously described species (8–11). Indeed, many *Chryseobacterium* type strain genomes have been sequenced as part of GEBA and also as part of the Genome Consortium for Active Teaching Next Generation Sequencing group (GCAT-SEEK) (12, 13). To provide a course-based undergraduate research experience (CURE) (14, 15) for the bioinformatics and -omics sciences honors course (BOCB 6834) at the University of the Free State (UFS) in Bloemfontein, South Africa, we selected seven unsequenced *Chryseobacterium* type strains from the UFS bacteria collection for analysis. Freeze-dried strains in vials were rehydrated with tryptic soy broth and streaked onto tryptic soy agar plates. After 2 days of incubation at 30°C, approximately 10 mg wet cell mass was scraped from the plates and DNA was extracted using the Zymo Quick-DNA fungal/bacterial miniprep kit according to the manufacturer’s recommendations. The purified DNA was assessed for quality by gel electrophoresis and the amount of DNA was quantified by fluorescent dye binding using a Qubit fluorometer (ThermoFisher). The source organisms’ identities were confirmed by amplification of the 16S rRNA genes from the purified DNA with universal primers 27f and 1492r (6) followed by Sanger sequencing of the amplicons with primer 27f and comparison to the type strain database on EzBioCloud (16). Nextera XT Libraries (Illumina) were prepared and 300-base paired-end sequences were generated on an Illumina MiSeq instrument at the UFS Next Generation Sequencing Unit. Reads were assembled using SPAdes v 3.10.0 with default parameters as implemented on PATRIC (17), and the assembly was assessed for quality with CheckM v 1.0.8 using the default parameters (18) on the KBase platform

Citation Matu A, Lum Nde A, Oosthuizen L, Hitzeroth A, Badenhorst M, Duba L, Gidaga M, Klinck J, Kriek I-M, Lekoma PJ, Nel L, dos Ramos SM, Rossouw J, Salomane N, Segone N, Serobe S, Tiyani T, Hugo CJ, Newman JD. 2019. Draft genome sequences of seven *Chryseobacterium* type strains. *Microbiol Resour Annot* 8:e01518-18. <https://doi.org/10.1128/MRA.01518-18>.

Editor David A. Baltus, University of Arizona

Copyright © 2019 Matu 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 Jeffrey D. Newman, newman@lycoming.edu.

* Present address: Arina Hitzeroth, Institute for Microbial Biotechnology and Metagenomics, University of the Western Cape, Bellville, Western Cape, South Africa.

Received 6 November 2018

Accepted 20 November 2018

Published 3 January 2019

TABLE 1 Genome sequence information

Organism (reference)	WGS accession no.	Genome size (bp)	G+C observed/predicted (mol%)	Coverage (×)	No. of contigs	N_{50} (kbp)	No. of 16S rRNA differences ^a	CheckM completeness/contamination
<i>Chryseobacterium elymi</i> KCTC 22547 ^T (1)	QNUH01	4,591,259	37.1/36.4	34	80	249	0	100/0.245
<i>Chryseobacterium flavum</i> KCTC 12877 ^T (3)	QNUE01	4,470,829	36.7/37.2	29	40	311	25	100/0.49
<i>Chryseobacterium hispanicum</i> KCTC 22104 ^T (4)	QNUG01	3,777,625	34.6/34.3	36	148	90	4	100/0.0
<i>Chryseobacterium lathyri</i> KCTC 22544 ^T (1)	QNFY01	4,857,652	37.0/36.6	28	56	305	1	100/0.245
" <i>Candidatus Chryseobacterium massiliae</i> " CCUG 51329 ^T (5)	QNVU01	4,409,139	36.0/ND ^b	29	116	79	0	100/0.49
<i>Chryseobacterium piscium</i> CCUG 51923 ^T (2)	QNVS01	4,319,594	33.6/33.6	30	158	51	0	100/0.0
<i>Chryseobacterium rhizosphaerae</i> KCTC 22548 ^T (1)	QNUF01	5,267,751	36.3/35.9	22	98	148	1	100/0.245

^a Differences between genome-derived 16S rRNA sequences and corresponding GenBank database entries.

^b ND, not determined.

(19) and annotated by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (20). The genomes ranged in size from 3.78 to 4.86 Mbp and had G+C mol% values from 33.6 to 37.1% which were all within 0.7% of values measured by HPLC or the thermal denaturation method (Table 1). Comparison of the genome-derived 16S rRNA sequences to the original sequences in GenBank revealed that while most of the originally deposited sequences were fairly accurate, the *Chryseobacterium flavum* CW-E 2^T sequence with the GenBank accession number EF154516 had 25 differences in a small region from position 974 to 1139 relative to the genome-derived sequence and was also very different from all other *Chryseobacterium* 16S rRNA sequences. The authors of the *C. flavum* species description (3) also noted that the 16S rRNA sequence similarity to the closest relatives was below 95%. The *C. flavum* KCTC 12877^T genome-derived 16S rRNA sequence was most similar (99.1%) to that of *Chryseobacterium gleum* ATCC 35910^T, so the ANI was calculated using the Ortho-ANI Tool (21) and found to be 83.2%, confirming that *C. flavum* and *C. gleum* are indeed distinct species. The *C. flavum* KCTC 12877^T 16S rRNA sequence was deposited separately into GenBank under the accession number MK116543.

Data availability. The *Chryseobacterium* whole-genome shotgun (WGS) projects were deposited under the GenBank accession numbers QNUH00000000 (*C. elymi* KCTC 22547^T), QNUE00000000 (*C. flavum* KCTC 12877^T), QNUG00000000 (*C. hispanicum* KCTC 22104^T), QNFY00000000 (*C. lathyri* KCTC 22544^T), QNVU00000000 ("*Ca. C. massiliae*" CCUG 51329^T), QNVS00000000 (*C. piscium* CCUG 51923^T), and QNUF00000000 (*C. rhizosphaerae* KCTC 22544^T). The versions described in this paper are the first versions, QNUH01000000, QNUE01000000, QNUG01000000, QNFY01000000, QNVU01000000, QNVS01000000, and QNUF01000000, respectively. The Sequence Read Archive (SRA) accession numbers are SRX5025940, SRX5025949, SRX5025952, SRX5026294, SRX5026665, SRX5027813, and SRX5027959, respectively.

ACKNOWLEDGMENTS

Sequencing was funded by the UFS Department of Microbial, Biochemical and Food Biotechnology and a Lycoming College Professional Development Grant. J.D.N. was supported by a Fulbright Grant from the U.S. State Department.

REFERENCES

1. Cho SH, Lee KS, Shin DS, Han JH, Park KS, Lee CH, Park KH, Kim SB. 2010. Four new species of *Chryseobacterium* from the rhizosphere of coastal sand dune plants, *Chryseobacterium elymi* sp. nov., *Chryseobacterium hagamense* sp. nov., *Chryseobacterium lathyri* sp. nov. and *Chryseobacterium rhizosphaerae* sp. nov. Syst Appl Microbiol 33:122–127. <https://doi.org/10.1016/j.syapm.2009.12.004>.
2. De Beer H, Hugo CJ, Jooste PJ, Vancanneyt M, Coenye T, Vandamme P. 2006. *Chryseobacterium piscium* sp. nov., isolated from fish of the South

- Atlantic Ocean off South Africa. *Int J Syst Evol Microbiol* 56:1317–1322. <https://doi.org/10.1099/ijs.0.64014-0>.
3. Zhou Y, Dong J, Wang X, Huang X, Zhang KY, Zhang YQ, Guo YF, Lai R, Li WJ. 2007. *Chryseobacterium flavum* sp. nov., isolated from polluted soil. *Int J Syst Evol Microbiol* 57:1765–1769. <https://doi.org/10.1099/ijs.0.65046-0>.
 4. Gallego V, García MT, Ventosa A. 2006. *Chryseobacterium hispanicum* sp. nov., isolated from the drinking water distribution system of Sevilla, Spain. *Int J Syst Evol Microbiol* 56:1589–1592. <https://doi.org/10.1099/ijs.0.64264-0>.
 5. Greub G, La Scola B, Raoult D. 2004. Amoebae-resisting bacteria isolated from human nasal swabs by amoebal coculture. *Emerg Infect Dis* 10:470–477. <https://doi.org/10.3201/eid1003.020792>.
 6. Strahan BL, Failor KC, Batties AM, Hayes PS, Cicconi KM, Mason CT, Newman JD. 2011. *Chryseobacterium piperi* sp. nov., isolated from a freshwater creek. *Int J Syst Evol Microbiol* 61:2162–2166. <https://doi.org/10.1099/ijs.0.027805-0>.
 7. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 57:81–91. <https://doi.org/10.1099/ijs.0.64483-0>.
 8. Whitman WB, Woyke T, Klenk HP, Zhou Y, Lilburn TG, Beck BJ, De Vos P, Vandamme P, Eisen JA, Garrity G, Hugenholtz P, Kyrpides NC. 2015. Genomic encyclopedia of bacterial and archaeal type strains, phase III: the genomes of soil and plant-associated and newly described type strains. *Stand Genomic Sci* 10:26. <https://doi.org/10.1186/s40793-015-0017-x>.
 9. Kyrpides NC, Hugenholtz P, Eisen JA, Woyke T, Göker M, Parker CT, Amann R, Beck BJ, Chain PSG, Chun J, Colwell RR, Danchin A, Dawyndt P, Dedeurwaerdere T, DeLong EF, Detter JC, De Vos P, Donohue TJ, Dong XZ, Ehrlich DS, Fraser C, Gibbs R, Gilbert J, Gilna P, Glöckner FO, Jansson JK, Keasling JD, Knight R, Labeda D, Lapidus A, Lee JS, Li WJ, Ma J, Markowitz V, Moore ERB, Morrison M, Meyer F, Nelson KE, Ohkuma M, Ouzounis CA, Pace N, Parkhill J, Qin N, Rossello-Mora R, Sikorski J, Smith D, Sogin M, Stevens R, Stingsl U, Suzuki KI, Taylor D, Tiedje JM, Tindall B, Wagner M, Weinstock G, Weissenbach J, White O, Wang J, Zhang L, Zhou YG, Field D, Whitman WB, Garrity GM, Klenk HP. 2014. Genomic encyclopedia of bacteria and archaea: sequencing a myriad of type strains. *PLoS Biol* 12:e1001920. <https://doi.org/10.1371/journal.pbio.1001920>.
 10. Mukherjee S, Seshadri R, Varghese NJ, Eloe-Fadrosh EA, Meier-Kolthoff JP, Göker M, Coates RC, Hadjithomas M, Pavlopoulos GA, Paez-Espino D, Yoshikuni Y, Visel A, Whitman WB, Garrity GM, Eisen JA, Hugenholtz P, Pati A, Ivanova NN, Woyke T, Klenk HP, Kyrpides NC. 2017. 1,003 reference genomes of bacterial and archaeal isolates expand coverage of the tree of life. *Nat Biotechnol* 35:676–683. <https://doi.org/10.1038/nbt.3886>.
 11. Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN, Kunin V, Goodwin L, Wu M, Tindall BJ, Hooper SD, Pati A, Lykidis A, Spring S, Anderson IJ, D'haeseleer P, Zemla A, Singer M, Lapidus A, Nolan M, Copeland A, Han C, Chen F, Cheng J-F, Lucas S, Kerfeld C, Lang E, Gronow S, Chain P, Bruce D, Rubin EM, Kyrpides NC, Klenk H-P, Eisen JA. 2009. A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature* 462:1056–1060. <https://doi.org/10.1038/nature08656>.
 12. Buonaccorsi VP, Boyle MD, Grove D, Praul C, Sakk E, Stuart A, Tobin T, Hosler J, Carney SL, Engle MJ, Overton BE, Newman JD, Pizzorno M, Powell JR, Trun N. 2011. GCAT-SEEKquence: genome consortium for active teaching of undergraduates through increased faculty access to next-generation sequencing data. *CBE Life Sci Educ* 10:342–345. <https://doi.org/10.1187/cbe.11-08-0065>.
 13. Buonaccorsi V, Peterson M, Lamendella G, Newman J, Trun N, Tobin T, Aguilar A, Hunt A, Praul C, Grove D, Roney J, Roberts W. 2014. Vision and change through the genome consortium for active teaching using next-generation sequencing (GCAT-SEEK). *CBE Life Sci Educ* 13:1–2. <https://doi.org/10.1187/cbe.13-10-0195>.
 14. Kloser MJ, Brownell SE, Chiariello NR, Fukami T. 2011. Integrating teaching and research in undergraduate biology laboratory education. *PLoS Biol* 9:9–11. <https://doi.org/10.1371/journal.pbio.1001174>.
 15. Drew JC, Triplett EW. 2008. Whole genome sequencing in the undergraduate classroom: outcomes and lessons from a pilot course. *J Microbiol Biol Educ* 9:3–11. <https://doi.org/10.1128/jmbe.v9.89>.
 16. Yoon S, Ha S, Kwon S, Lim J, Kim Y, Seo H, Chun J. 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 67:1613–1617. <https://doi.org/10.1099/ijssem.0.001755>.
 17. Wattam AR, Brettin T, Davis JJ, Gerdes S, Kenyon R, Machi D, Mao C, Olson R, Overbeek R, Pusch GD, Shukla MP, Stevens R, Vonstein V, Warren A, Xia F, Yoo H. 2018. Assembly, annotation, and comparative genomics in PATRIC, the all bacterial bioinformatics resource center. In: Setubal J, Stoye J, Stadler P. (eds), *Comparative Genomics. Methods in Molecular Biology*, 1704:79–101. Humana Press, New York, NY.
 18. 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>.
 19. 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>.
 20. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res* 46:D851–D860. <https://doi.org/10.1093/nar/gkx1068>.
 21. Lee I, Kim YO, Park SC, Chun J. 2016. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 66:1100–1103. <https://doi.org/10.1099/ijssem.0.000760>.