Whole-Genome Sequences of *Corynebacterium macginleyi* CCUG 32361^T and Clinical Isolates NML 080212 and NML 120205

Anne-Marie Bernier,[a] Kathryn Bernard[b,c]

[a]Department of Biology, Université de Saint-Boniface, Winnipeg, Manitoba, Canada
[b]National Microbiology Laboratory (NML)–CSCHAH Site, Public Health Agency of Canada, Winnipeg, Manitoba, Canada
[c]Department of Medical Microbiology, University of Manitoba, Winnipeg, Manitoba, Canada

ABSTRACT   Draft genome sequences of *Corynebacterium macginleyi* CCUG 32361^T and clinical isolates NML 080212 and NML 120205 were assembled and studied. Genome sizes ranged from 2.35 Mb to 2.42 Mb, with G+C contents ranging from 57.1% to 57.2%.

The lipophilic species *Corynebacterium macginleyi*, an ocular pathogen for some coryneforms provisionally called CDC group G-1, was described by Riegel et al. in 1995 (1). The first bilateral eye infection caused by this agent in Canada (NML 080212) was reported in 2010 (2). Upon review, the NCBI genome database was found to lack whole-genome sequence (WGS) data for this species. Therefore, in this study, *C. macginleyi* CCUG 32361^T, previously acquired from the Culture Collection of the University of Gothenburg (CCUG; NML 080212), and a second clinical isolate recovered from an ocular infection in Vancouver, British Columbia (NML 120205) were characterized by WGS.

Bacteria were subcultured after storage at −80°C in Microbank vials (PRO-LAB) from NML stocks or from lyophilized CCUG 32361 and passed twice on Colombia blood agar (CBA) plates for 24 h at 35°C in 5% CO_2. A loopful of plate culture was grown aerobically in Trypticase soy broth (BD) for 18 h at 35°C, and DNA was extracted from broth cultures using the DNA minikit (Qiagen). Paired-end whole-genome shotgun libraries were constructed using the Nextera XT library preparation kit, and samples were run separately for sequencing on the MiSeq 600-cycle kit (version 3) on a MiSeq sequencer (Illumina, Munich, Germany). Read quality was assessed with FastQC version 0.11.8 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and short paired-end reads were then merged with Fast Length Adjustment of SHort reads (FLASH, version 1.2.9) (3). Reads were then assembled using SPAdes (version 3.9.0) (4) with default settings. Sequencing and assembly metrics for the three strains are presented in Table 1. The assemblies produced draft genomes composed of 2,349,818 to 2,419,073 bp, with an average G+C content of 57.1% and an average coverage of 100-fold. Contigs from assembled shotgun data were annotated for genes and other features using Prokka (version 1.13) (5), which revealed an average of 2,321 coding DNA sequences (CDS). One clustered regularly interspaced short palindromic repeat (CRISPR) sequence with 11 spacers was found only in the genome of strain NML 080212 and was otherwise absent from the other two strains. Prophages were not found in any of the three genomes using Phaster (6).

Average nucleotide identity values using BLAST+ (ANIb) as calculated using JSpeciesWS (7) were used to compare the NML strains to CCUG 32361^T and to each other. All ANIb scores were greater than 98.74%. Similarly, the *in silico* DNA-DNA

TABLE 1 *Corynebacterium macginleyi* sequencing, assembly, and annotation metrics

Strain	No. of reads	No. of bases	No. of contigs >1 kb	Genome size (bp)	Coverage (×)	Contig N_{50} (bp)	G+C content (%)	No. of CDS (Prokka)
CCUG 32361^T	391,644	116,274,866	56	2,419,073	96	128,157	57.1	2,366
NML 080212	462,220	134,561,054	77	2,417,567	95	57,952	57.09	2,337
NML 120205	477,946	136,596,203	40	2,349,818	109	163,397	57.21	2,260

hybridization (*is*DDH) values as calculated using formula 2 of the Genome-to-Genome Distance Calculator (8) were greater than 90.2%. Isolates lacked fatty acid synthase genes, which were previously observed for lipophilic *Corynebacterium* species (9), but genes associated with mycolate synthesis were detected (10).

Data availability. Draft genome sequences of *Corynebacterium macginleyi* CCUG 32361^T, NML 080212, and NML 120205 were deposited at DDBJ/ENA/GenBank under accession numbers REGE00000000, REGD00000000, and REGC00000000, respectively. The versions described in this paper are REGE01000000, REGD01000000, and REGC01000000, respectively. Sequence Read Archive accession numbers are SRX4938091, SRX4938092, and SRX4938093.

REFERENCES

1. Riegel P, Ruimy R, de Briel D, Prévost G, Jehl F, Christen R, Monteil H. 1995. Genomic diversity and phylogenetic relationships among lipid-requiring diphtheroids from humans and characterization of *Corynebacterium macginleyi* sp. nov. Int J Syst Bacteriol 45:128–133. https://doi.org/10.1099/00207713-45-1-128.

2. Alsuwaidi AR, Wiebe D, Burdz T, Ng B, Reimer A, Singh C, Bernard K. 2010. *Corynebacterium macginleyi* conjunctivitis in Canada. J Clin Microbiol 48:3788–3790. https://doi.org/10.1128/JCM.01289-10.

3. Magoc T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27:2957–2963. https://doi.org/10.1093/bioinformatics/btr507.

4. 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.

5. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

6. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–W21. https://doi.org/10.1093/nar/gkw387.

7. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: a Web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32:929–931. https://doi.org/10.1093/bioinformatics/btv681.

8. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60. https://doi.org/10.1186/1471-2105-14-60.

9. Tauch A, Schneider J, Szczepanowski R, Tilker A, Viehoever P, Gartemann KH, Arnold W, Blom J, Brinkrolf K, Brune I, Gotker S, Weisshaar B, Goesmann A, Droge M, Puhler A. 2008. Ultrafast pyrosequencing of *Corynebacterium kroppenstedtii* DSM44385 revealed insights into the physiology of a lipophilic corynebacterium that lacks mycolic acids. J Biotechnol 136:22–30. https://doi.org/10.1016/j.jbiotec.2008.03.004.

10. Tippelt A, Mollmann S, Albersmeier A, Jaenicke S, Ruckert C, Tauch A. 2014. Mycolic acid biosynthesis genes in the genome sequence of *Corynebacterium atypicum* DSM 44849. Genome Announc 2:e00845-14. https://doi.org/10.1128/genomeA.00845-14.