




Complete Genome Sequence of a Wild-Type Isolate of *Caulobacter vibrioides* Strain CB2

 Derrick C. Scott,^a LaTia Scott,^a Kiesha Wilson,^b Keshawn Ross,^a Damyen Ingram,^a Tajah Lewter,^a Jasmine Herring,^a LaTrice Linton,^a David Duncan,^a Anthea Aikins,^a Bert Ely^b

^aDepartment of Biological Sciences, Delaware State University, Dover, Delaware, USA

^bDepartment of Biological Sciences, University of South Carolina, Columbia, South Carolina, USA

ABSTRACT The complete genome of *Caulobacter vibrioides* strain CB2 consists of a 4,123,726-bp chromosome, a GC content of 67.2%, and 3,896 coding DNA sequences. It has no rearrangements but numerous indels relative to the reference NA1000 genome. This will allow us to study the impact of horizontal gene transfer on caulobacter genomes.

Caulobacters are Gram-negative bacteria which produce stalks at one end of the cell. The bacterium differentiates into two cell types and divides asymmetrically at each cell cycle. *Caulobacter crescentus* strain CB15 has a well-developed system of genetics (1, 2) and was the first caulobacter to be fully sequenced (3). It was developed into a single-celled model system to study cellular differentiation, asymmetric division, and cell cycle progression (4, 5). Studies of this obligatory differentiation during the cell cycle made caulobacter the dominant prokaryotic model system for studying the mechanisms of cell cycle control and cellular differentiation. Today, however, caulobacters have attracted interest because it was discovered that even closely related strains of caulobacters exhibit extremely high levels of genome rearrangements (6, 7). To discover the molecular mechanism behind this “genome scrambling,” more genomes are needed for genomic comparisons. Here, we report the full sequence of the genomic material of an additional *Caulobacter vibrioides* wild-type strain, CB2, sampled from tap water in California (8).

The *Caulobacter vibrioides* strain CB2 was grown in peptone yeast extract for 48 h at 30°C as previously described (9). Total genomic DNA was extracted with a Qiagen DNeasy tissue kit following the manufacturer’s protocol. Previous studies have shown that high-GC-content genomes can be challenging to sequence with short-read technology (10). As such, the total genome was sequenced with a PacBio RS II sequencer at the Delaware Bioinformatics Institute. The DNA library was prepared for sequencing with the PacBio blunt-end ligation protocol. The subsequent raw reads were assembled with the Hierarchical Genome Assembly Process (HGAP3) (11) in SMRT Portal with the default *de novo* parameters. The computational requirements needed for the analysis were leveraged through Amazon Machine Image (AMI) EC2 with the smrtanalysis-2.3.0-ami-20fb4848 image. The number of raw assembly reads was 45,003 with a N_{50} read length of 23,435 bp. Total coverage was 125×. All pipelines began with adapter removal and subread filtering with a final contig polish done by Quiver, the last step in the HGAP3 assembly process. The sequence was annotated with the Rapid Annotation using Subsystem Technology (RAST) server (<http://rast.nmpdr.org>) and the NCBI Prokaryotic Genome Annotation Pipeline (12–14) and then visualized in Artemis (release 16.0.0) (15). The CB2 genome aligns well with the circular *Caulobacter* NA1000 genome. The ends of the contig we obtained extensively duplicated each other, indicating circularity. The complete genome of *Caulobacter vibrioides* CB2 is 4,123,726 bp, with a

Received 17 September 2018 **Accepted** 4 October 2018 **Published** 1 November 2018

Citation Scott DC, Scott L, Wilson K, Ross K, Ingram D, Lewter T, Herring J, Linton L, Duncan D, Aikins A, Ely B. 2018. Complete genome sequence of a wild-type isolate of *Caulobacter vibrioides* strain CB2. *Microbiol Resour Announc* 7:e01215-18. <https://doi.org/10.1128/MRA.01215-18>.

Editor J. Cameron Thrash, Louisiana State University

Copyright © 2018 Scott 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 Derrick C. Scott, dcscott@desu.edu.

GC content of 67.2%. There are 3,896 predicted coding sequences (CDSs), 52 tRNAs, and 2 rRNA operons.

Caulobacter vibrioides CB2 has no inversions, but it has numerous insertions and deletions compared to the closely related NA1000 strain. The absence of rearrangements indicates that this well-documented phenomenon (6, 7), occurs relatively rarely in these two compared genomes. Further studies will be able to shed light on the horizontal gene transfer that happens much more frequently among caulobacters.

Data availability. The complete genome sequence of *Caulobacter vibrioides* CB2 has been deposited in GenBank under the accession number [CP023313](https://doi.org/10.1093/nar/gkx1068). The raw reads are also available under SRA accession number [SRX4603198](https://doi.org/10.1093/nar/gkx1068).

ACKNOWLEDGMENT

This research was supported by an Institutional Development Award (IDeA) from the National Institutes of Health's National Institute of General Medical Sciences (NIGMS) under grant number P20GM103446.

REFERENCES

- Ely B. 1991. Genetics of *Caulobacter crescentus*. *Methods Enzymol* 204: 372–384. [https://doi.org/10.1016/0076-6879\(91\)04019-K](https://doi.org/10.1016/0076-6879(91)04019-K).
- Shapiro L. 1976. Differentiation in the *Caulobacter* cell cycle. *Annu Rev Microbiol* 30:377–407. <https://doi.org/10.1146/annurev.mi.30.100176.002113>.
- Nierman WC, Feldblyum TV, Laub MT, Paulsen IT, Nelson KE, Eisen J, Heidelberg JF, Alley MRK, Ohta N, Maddock JR, Potocka I, Nelson WC, Newton A, Stephens C, Phadke ND, Ely B, DeBoy RT, Dodson RJ, Durkin AS, Gwinn ML, Haft DH, Kolonay JF, Smit J, Craven MB, Khouri H, Shetty J, Berry K, Utterback T, Tran K, Wolf A, Vamathevan J, Ermolaeva M, White O, Salzberg SL, Venter JC, Shapiro L, Fraser CM. 2001. Complete genome sequence of *Caulobacter crescentus*. *Proc Natl Acad Sci U S A* 98: 4136–4141. <https://doi.org/10.1073/pnas.061029298>.
- Brun YV, Janakiraman R. 2000. The dimorphic life cycle of *Caulobacter* and stalked bacteria, p 297–317. In Brun YV, Shimkets LJ (ed), *Prokaryotic development*. ASM Press, Washington, DC.
- Laub MT, McAdams HH, Feldblyum T, Fraser CM, Shapiro L. 2000. Global analysis of the genetic network controlling bacterial cell cycle. *Science* 290:2144–2148. <https://doi.org/10.1126/science.290.5499.2144>.
- Scott D, Ely B. 2016. Conservation of the essential genome among *Caulobacter* and *Brevudimonas* species. *Curr Microbiol* 72:503–510. <https://doi.org/10.1007/s00284-015-0964-x>.
- Ash K, Brown T, Watford T, Scott LE, Stephens C, Ely B. 2014. A comparison of the *Caulobacter* NA1000 and K31 genomes reveals extensive genome rearrangements and differences in metabolic potential. *Open Biol* 4:140128. <https://doi.org/10.1098/rsob.140128>.
- Poindexter JS. 1964. Biological properties and classification of the *Caulobacter* group. *Bacteriol Rev* 28:231–295.
- Johnson RC, Ely B. 1977. Isolation of spontaneously derived mutants of *Caulobacter crescentus*. *Genetics* 86:25–32.
- Scott D, Ely B. 2015. Comparison of genome sequencing technology and assembly methods for the analysis of a GC-rich bacterial genome. *Curr Microbiol* 70:338–344. <https://doi.org/10.1007/s00284-014-0721-6>.
- Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
- Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O. 2008. Toward an online repository of standard operating procedures (SOPs) for (meta)genomic annotation. *OMICS* 12:137–141. <https://doi.org/10.1089/omi.2008.0017>.
- Tatusova T, DiCuccio M, Badretdin A, Chetverin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- Haft DH, DiCuccio M, Badretdin A, Brover V, Chetverin 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>.
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16:944–945. <https://doi.org/10.1093/bioinformatics/16.10.944>.