Complete Genome Sequence of *Clavibacter michiganensis* subsp. *insidiosus* R1-1 Using PacBio Single-Molecule Real-Time Technology

You Lu,[a,b] Deborah A. Samac,[b,c,d] Jane Glazebrook,[a,b] Carol A. Ishimaru[b,d]

Department of Plant Biology, University of Minnesota, St. Paul, Minnesota, USA[a]; Microbial and Plant Genomics Institute, University of Minnesota, St. Paul, Minnesota, USA[b]; USDA-ARS Plant Science Research Unit, St. Paul, Minnesota, USA[c]; Department of Plant Pathology, University of Minnesota, St. Paul, Minnesota, USA[d]

We report here the complete genome sequence of *Clavibacter michiganensis* subsp. *insidiosus* R1-1, isolated in Minnesota, USA. The R1-1 genome, generated by a *de novo* assembly of PacBio sequencing data, is the first complete genome sequence available for this subspecies.

Address correspondence to Carol A. Ishimaru, cishimar@umn.edu.

Clavibacter michiganensis subsp. *insidiosus* is a Gram-positive coryneform bacterium which causes bacterial wilt of alfalfa (*Medicago sativa* L.), the most widely cultivated forage legume (1). Bacterial wilt of alfalfa occurs throughout most alfalfa-growing regions worldwide. For this reason, the pathogen is considered a phytosanitary risk for the international seed movement. *C. michiganensis* subsp. *insidiosus* R1-1 was isolated from stem tissue of a *Medicago truncatula* plant inoculated with ground roots of symptomatic alfalfa plants from disease nurseries at the University of Minnesota Rosemount Research and Outreach Center in Rosemount, MN.

There are six subspecies of *C. michiganensis* identified as plant pathogens (2, 3). Complete genome sequences are currently available for a single strain of three subspecies: the tomato pathogen, *C. michiganensis* subsp. *michiganensis* (4), the potato pathogen, *C. michiganensis* subsp. *sepedonicus* (5), and the corn pathogen, *C. michiganensis* subsp. *nebraskensis* (2). Comparative genome analyses revealed extensive rearrangement between the genomes and the presence or absence of genes encoding known virulence factors (2, 5). Previous research on *C. michiganensis* subsp. *insidiosus* has focused primarily on plant resistance and diagnostic methods (6, 7). The lack of a complete genome sequence has hampered progress in understanding the molecular mechanisms of host-pathogen interactions.

An earlier *de novo* assembly of the *C. michiganensis* subsp. *insidiosus* R1-1 genome from paired-ended reads from Illumina GAIIx system sequencing resulted in 103 contigs containing ambiguous bases due to the presence of repetitive elements and a high G+C genome content. Here, we utilized PacBio single-molecule real-time (SMRT) sequencing technology (8) to generate a *de novo* assembly of the complete genome sequence of *C. michiganensis* subsp. *insidiosus* R1-1. Genomic DNA extracted from *C. michiganensis* subsp. *insidiosus* R1-1 was prepared as a 20-kb library for P6-C4 chemistry, followed by BluePippin size selection at 15 kb. The PacBio RSII sequencing system generated 64,530 reads, with a mean read length of 10,432 bp from one SMRT cell. The initial assembly was conducted using Hierarchical Genome Assembly

Process 3 (HGAP3) (9) in PacBio SMRT portal version 2.2.0, and four contigs were assembled, with a mean coverage of 163.44×. This assembly was corrected with the Quiver consensus algorithm to obtain a high-accuracy genome assembly (9). Further improvement of the quality of the genome sequence was performed with Pilon (10) using data generated by Illumina GAIIx platform sequencing. This process corrected 43 indel errors in the PacBio assembly. The resulting genome assembly contains four circular contigs, corresponding to one chromosome (3,207,520 bp; 72.96% G+C content) and three circular plasmids, pCI1 (47,690 bp; 67.78% G+C content), pCI2 (49,401 bp; 67.58% G+C content), and pCI3 (103,451 bp; 66.16% G+C content).

Annotation was performed with the NCBI Prokaryotic Genome Annotation Pipeline (http://www.ncbi.nlm.nih.gov/genome/annotation_prok/). In total, 3,012 protein-coding genes, two rRNA operons (16S, 23S, 5S), 46 tRNA genes, and 27 copies of the insertion element IS*1122* were identified in the R1-1 genome.

The availability of the complete genome sequence of *C. michiganensis* subsp. *insidiosus* R1-1 will be valuable for further comparative genomic analyses and the development of molecular diagnostic tools.

Nucleotide sequence accession numbers. The *C. michiganensis* subsp. *insidiosus* R1-1 genome sequence was deposited in GenBank under the accession numbers CP011043 to CP011046. The versions described here are the first versions.

REFERENCES

1. **Metzler MC, Laine MJ, De Boer SH.** 1997. The status of molecular biological research on the plant pathogenic genus *Clavibacter*. FEMS Microbiol Lett **150**:1–8. http://dx.doi.org/10.1111/j.1574-6968.1997.tb10342.x.

2. **Eichenlaub R, Gartemann KH.** 2011. The *Clavibacter michiganensis* subspecies: molecular investigation of Gram-positive bacterial plant pathogens. Annu Rev Phytopathol **49:**445–464. http://dx.doi.org/10.1146/annurev-phyto-072910-095258.

3. **González AJ, Trapiello E.** 2014. *Clavibacter michiganensis* subsp. *phaseoli*, a new subspecies pathogen in bean. Int J Syst Evol Microbiol **64:**1752–1755. http://dx.doi.org/10.1099/ijs.0.058099-0.

4. **Gartemann KH, Abt B, Bekel T, Burger A, Engemann J, Flügel M, Gaigalat L, Goesmann A, Gräfen I, Kalinowski J, Kaup O, Kirchner O, Krause L, Linke B, McHardy A, Meyer F, Pohle S, Rückert C, Schneiker S, Zellermann EM, Puhler A, Eichenlaub R, Kaiser O, Bartels D.** 2008. The genome sequence of the tomato-pathogenic actinomycete *Clavibacter michiganensis* subsp. *michiganensis* NCPPB382 reveals a large island involved in pathogenicity. J Bacteriol **190:**2138–2149. http://dx.doi.org/10.1128/JB.01595-07.

5. **Bentley SD, Corton C, Brown SE, Barron A, Clark L, Doggett J, Harris B, Ormond D, Quail MA, May G, Francis D, Knudson D, Parkhill J, Ishimaru CA.** 2008. Genome of the actinomycete plant pathogen *Clavibacter michiganensis* subsp. *sepedonicus* suggests recent niche adaptation. J Bacteriol **190:**2150–2160. http://dx.doi.org/10.1128/JB.01598-07.

6. **Samac DA, Nix RJ, Oleson AE.** 1998. Transmission frequency of *Clavibacter michiganensis* subsp. *insidiosus* to alfalfa seed and identification of the bacterium by PCR. Plant Dis **82:**1362–1367. http://dx.doi.org/10.1094/PDIS.1998.82.12.1362.

7. **Marefat A, Ophel-Keller K, McKay A.** 2007. A real-time PCR assay for detection of *Clavibacter michiganensis* subsp. *insidiosus* in Lucerne. Austral Plant Pathol **36:**262–269. http://dx.doi.org/10.1071/AP07018.

8. **Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S.** 2009. Real-time DNA sequencing from single polymerase molecules. Science **323:**133–138. http://dx.doi.org/10.1126/science.1162986.

9. **Chin CS, 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. http://dx.doi.org/10.1038/nmeth.2474.

10. **Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM.** 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One **9:**e112963. http://dx.doi.org/10.1371/journal.pone.0112963.