Draft Genome Sequence of *Pseudomonas simiae* Strain 2-36, an *In Vitro* Antagonist of *Rhizoctonia solani* and *Gaeumannomyces graminis*

Zaky Adam, Qing Chen,* Renlin Xu, Adolf E. Diange, Eden S. P. Bromfield, James Tabi Tambong

Biodiversity, Environmental Health Program, Agriculture and Agri-Food Canada, Ottawa, Ontario, Canada

* Present address: Qing Chen, Sichuan Agricultural University, Sichuan, China.

Pseudomonas simiae 2-36, isolated from a field plot under long-term mineral fertilization, exhibited strong *in vitro* antagonistic activities against *Rhizoctonia solani* and *Gaeumannomyces graminis*. We report here the draft genome sequence of *Pseudomonas simiae* 2-36, consisting of 6.4 Mbp with a 60.25% G+C content and 5,790 predicted protein-coding sequences.

*P*seudomonas simiae 2-36, isolated from a field plot under long-term mineral fertilization, exhibited strong *in vitro* antagonistic activities against *Rhizoctonia solani* and *Gaeumannomyces graminis*. The synthesis of molecules involved in antagonistic interactions and disease suppression is positively regulated by the *gacA* gene. A 500-bp *gacA* sequence derived from *P. simiae* 2-36 showed only 95.1% similarity to that of a known biocontrol agent, *Pseudomonas* sp. strain BSB1-2-37 (GenBank accession no. EF417461). Mutations in the *gacA*/*gacS* genes generally indicate phenotypic variations. We report here the draft genome sequence of *P. simiae* 2-36. The draft genome was determined by paired-end sequencing using Illumina MiSeq technology (at Génome-Québec, Montreal, Canada). A total of 7,902,862 paired-end reads, each 150 bp in length, totaling 1,185,429,300 bp, were obtained from 380-bp inserts. Quality checking using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) showed that the reads were of sufficiently good quality such that no further trimming or error correction was required. The initial *de novo* assembly using ABySS version 1.5.2 (1) produced 54 contigs contained in 46 scaffolds, from which scaffolds with length <300 bp were removed. The remaining 21 scaffolds (minimum, 966 bp; maximum, 1,186,008 bp; N_{50}, 755,011 bp; total size, 6,394,971 bp; total number of Ns, 1,715) were used for further analyses. SSPACE version 2.0 (2) was applied on the resulting scaffolds to extend and merge them into larger scaffolds based on read-pair information and short overlaps. This process reduced the number of scaffolds to 20 (minimum, 1,348 bp; maximum, 1,186,008 bp; N_{50}, 755,011 bp; total size, 6,396,512 bp; total number of Ns, 1,716). GapFiller version 1.11 (3) was used to close the gaps between the short scaffolds that are contained within the 20 large scaffolds by replacing the unknown nucleotides (Ns) with true nucleotides based on read-pair information and short overlaps. The final draft genome consists of 20 scaffolds (minimum, 1,348 bp; maximum, 1,185,661 bp; N_{50}, 754,763 bp) totaling 6,395,662 bp with 41 Ns. The G+C content of the draft genome is 60.25%, with an overall estimated coverage at 169×.

Mauve Contig Mover version 2.3.1 (4) was used to order the

draft genome of *P. simiae* strain 2-36 using *P. simiae* WCS 417 (GenBank accession no. CP007637.1) as a reference genome. Automated annotation using the RAST annotation server (5) revealed that the draft genome of *P. simiae* 2-36 contains 5,790 predicted protein-coding sequences, of which 4,489 have assigned functions, 335 have proposed functions, 955 are considered hypothetical proteins, and 11 are proteins of unknown functions. The draft genome also contains 82 predicted noncoding RNAs, including 63 tRNAs, one pseudo-tRNA, and 18 rRNAs. The number of gene copies encoding 16S rRNA, 5S rRNA, and 23S rRNA are 6, 7, and 5, respectively. Compared to RAST, Glimmer version 3.02 (6), using open reading frames (ORFs) as a training set, predicted 5,811 genes, whereas RNAmmer version 1.2 (7) predicted 13 rRNAs containing 5, 7, and 1 copies of 16S rRNA, 5S rRNA, and 23S rRNA genes, respectively. The tRNAscan-SE version 1.3.1 software (8), however, predicted 65 tRNAs genes. Sixteen gene clusters involved in the synthesis of secondary metabolites, such as bacteriocin, nonribosomal peptides, and terpenes, were identified (9).

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. JRMC00000000. The version described in this paper is the first version, JRMC01000000.

REFERENCES

1. **Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I.** 2009. ABySS: a parallel assembler for short read sequence data. Genome Res **19:** 1117–1123. http://dx.doi.org/10.1101/gr.089532.108.
2. **Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W.** 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics **27:** 578–579. http://dx.doi.org/10.1093/bioinformatics/btq683.

3. **Boetzer M, Pirovano W.** 2012. Toward almost closed genomes with Gap-Filler. Genome Biol **13**:R56. http://dx.doi.org/10.1186/gb-2012-13-6-r56.

4. **Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT.** 2009. Reordering contigs of draft genomes using the Mauve aligner. Bioinformatics **25**:2071–2073. http://dx.doi.org/10.1093/bioinformatics/btp356.

5. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Claudia Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.** 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics **9**:75. http://dx.doi.org/10.1186/1471-2164-9-75.

6. **Delcher AL, Harmon D, Kasif S, White O, Salzberg SL.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res **27**:4636–4641. http://dx.doi.org/10.1093/nar/27.23.4636.

7. **Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res **35**:3100–3108. http://dx.doi.org/10.1093/nar/gkm160.

8. **Lowe TM, Eddy SR.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res **25**:955–964. http://dx.doi.org/10.1093/nar/25.5.0955.

9. **Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T.** 2013. antiSMASH 2.0 – a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res **41**(Web Server Issue):W204–W212. http://dx.doi.org/10.1093/nar/gkt449.