



# Draft Genome Sequence of the Plant Growth-Promoting Rhizobacterium *Pseudomonas fluorescens* Strain CREA-C16 Isolated from Pea (*Pisum sativum* L.) Rhizosphere

 Nunzio D'Agostino,<sup>a</sup> Roberto Sorrentino,<sup>a</sup> Riccardo Scotti,<sup>a</sup> Melania Salzano,<sup>a</sup> Vincenzo Aurilia,<sup>b</sup> Massimo Zaccardelli<sup>a</sup>

CREA-ORT, Consiglio per la Ricerca in Agricoltura e l'Analisi dell'Economia Agraria, Centro di Ricerca per l'Orticoltura, Pontecagnano Faiano, Salerno, Italy<sup>a</sup>; National Research Council of Italy, Institute for Mediterranean Agriculture and Forest Systems, Ercolano, Naples, Italy<sup>b</sup>

**ABSTRACT** Herein, we report the draft genome sequence of *Pseudomonas fluorescens* strain CREA-C16, a plant growth-promoting rhizobacterium that was isolated from the rhizosphere of *Pisum sativum* L. plants. The genome sequence is ~6 Mb in size, with a G+C content of 60.1%, and includes 4,457 candidate protein-encoding genes.

*Pseudomonas fluorescens* strain CREA-C16 was originally isolated from the rhizosphere of *Pisum sativum* L. plants cultivated in the Sele Valley (Campania, Italy) during 2007 and examined for its plant growth-promoting properties. The ability of this strain to promote root elongation and growth of tomato (*Solanum lycopersicum* L.) plants (72% increase in leaf dry weight compared to noninoculated plants), to release antifungal as well as phytohormone-like substances, and to produce fluorescent pigments makes it an efficient plant growth-promoting rhizobacterium (PGPR).

Furthermore, tomato plant roots are subjected to major transcriptome reprogramming after the early interaction with *P. fluorescens* strain CREA-C16. Genes encoding calmodulin, calmodulin-like proteins, and ethylene-responsive element binding factors are involved in plant root response to this PGPR.

Genomic DNA was extracted from a pellet of cells using the cetyltrimethylammonium bromide (CTAB) method, according to the protocol described by Ausbel et al. (1).

A draft genome sequence of *Pseudomonas fluorescens* strain CREA-C16 was generated from a paired-end library constructed with the Nextera library preparation kit. The sequencing ( $2 \times 101$  bp) was performed on a HiSeq 2500 device and generated 1.150 million reads (estimated insert size, ~1,100 bp). After filtering out low-quality reads, trimming of 10 bases from the 3' end, and adapter removal, preprocessed reads were re-paired using the fastqCombinedPairedEnd.py script.

High-quality reads were *de novo* assembled with Velvet (2) into 132 contigs (k-mer size, 41) with minimum length of 200 bp, ranging from 204 bp to 415,383 bp and a  $N_{50}$  length of 110,307 bp.

Those contigs were ordered by Ragout version 0.2 (3) using *Pseudomonas fluorescens* PF-0 (GenBank accession number CP000094.2) as the reference genome.

The assembly was further improved by running SOAP GapCloser version 1.12 (4), based on high-quality Illumina paired- and single-end reads. The result was a draft genome sequence of 5,758,958 bp, including gaps, with a G+C content of 60.1%. Per-base sequence coverage was calculated by mapping high-quality reads to the

**Received** 14 November 2016 **Accepted** 22 November 2016 **Published** 26 January 2017

**Citation** D'Agostino N, Sorrentino R, Scotti R, Salzano M, Aurilia V, Zaccardelli M. 2017. Draft genome sequence of the plant growth-promoting rhizobacterium *Pseudomonas fluorescens* strain CREA-C16 isolated from pea (*Pisum sativum* L.) rhizosphere. Genome Announc 5:e01456-16. <https://doi.org/10.1128/genomeA.01456-16>.

**Copyright** © 2017 D'Agostino 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 Nunzio D'Agostino, [nunzio.dagostino@crea.gov.it](mailto:nunzio.dagostino@crea.gov.it), or Massimo Zaccardelli, [massimo.zaccardelli@crea.gov.it](mailto:massimo.zaccardelli@crea.gov.it).

newly assembled genome using bowtie2 version 2.1.0 (5) and then running the coverage tool from the BEDtools package (6). The mean and median coverage depths were 23.24× and 26×, respectively.

Reference-based gene identification was performed running the prokaryotic gene finder tool EasyGene (7) using *Pseudomonas aeruginosa* PAO1 as reference, the PRODIGAL gene-finding algorithm (8), and the FgenesB (Softberry) using the “archae bacterial generic” as the closest organism. The final annotation includes only those genes predicted by at least two of three prediction tools. The annotation process resulted in 4,457 candidate protein-encoding genes.

Functional annotation was performed by BLAST search (e-value, < 10<sup>-5</sup>) of the predicted protein sequences against UniProtKB (9). Gene model descriptions were refined using InterProScan (10) for the identification of conserved protein domains and GO terms retrieval, and KAAS (11) was used for KEGG Orthology (KO) assignments and KEGG pathway mapping.

The annotation includes genes encoding enzymes involved in the indole-3-acetic acid (IAA) production from tryptophan, as well as a gene encoding the enzyme 1-aminocyclopropane-1-carboxylate (ACC) deaminase, a key factor in facilitating plant growth promotion.

**Accession number(s).** This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the GenBank accession no. [CP017951](https://doi.org/10.1101/074492). The version described in this paper is the first version.

## ACKNOWLEDGMENTS

This work was carried out in the frame of the “GenoPom-pro—Integrating post-genomic platforms to enhance the tomato production chain” project (PON02\_00395\_3082360), supported by the PON R&C 2007-2013 grant funded by the Italian Ministry of Education, University and Research in cooperation with the European Regional Development Fund (ERDF).

## REFERENCES

1. Ausbel FM, Brent R, Kingston RE, Moore DD, Smith JA, Seidman JC, Struhl K. 1987. Preparation of genomic DNA from bacteria. *Curr Protoc Mol Biol* Chapter 2:Unit 2.
2. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
3. Kolmogorov M, Raney B, Paten B, Pham S. 2014. Ragout—a reference-assisted assembly tool for bacterial genomes. *Bioinformatics* 30: i302–i309. <https://doi.org/10.1093/bioinformatics/btu280>.
4. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam TW, Wang J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *GigaScience* 1:18. <https://doi.org/10.1186/2047-217X-1-18>.
5. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
6. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842. <https://doi.org/10.1093/bioinformatics/btq033>.
7. Larsen TS, Krogh A. 2003. EasyGene—a prokaryotic gene finder that ranks ORFs by statistical significance. *BMC Bioinformatics* 4:21. <https://doi.org/10.1186/1471-2105-4-21>.
8. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
9. Pundir S, Magrane M, Martin MJ, O'Donovan C, UniProt Consortium. 2015. Searching and navigating UniProt databases. *Curr Protoc Bioinformatics* 50:1.27.1–1.27.10. <https://doi.org/10.1002/0471250953.bi0127550>.
10. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30: 1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>.
11. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 35:W182–W185. <https://doi.org/10.1093/nar/gkm321>.