

Draft Genome Sequences of a Phylogenetically Diverse Suite of *Pseudomonas syringae* Strains from Multiple Source Populations

David A. Baltrus,^a Scott Yourstone,^b Abigail Lind,^b Caroline Guilbaud,^c David C. Sands,^d Corbin D. Jones,^{b,e} Cindy E. Morris,^c Jeffrey L. Dangl^{b,e,f}

School of Plant Sciences, University of Arizona, Tucson, Arizona, USA^a; Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA^b; INRA, UR0407 Pathologie Végétale, Montfavet, France^c; Plant Sciences and Plant Pathology Department, Montana State University, Bozeman, Montana, USA^d; Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA^e; Howard Hughes Medical Institute, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA^f

Here, we report the draft genome sequences for 7 phylogenetically diverse isolates of *Pseudomonas syringae*, obtained from numerous environmental sources and geographically proximate crop species. Overall, these sequences provide a wealth of information about the differences (or lack thereof) between isolates from disease outbreaks and those from other sources.

Received 10 December 2013 Accepted 13 December 2013 Published 23 January 2014

Citation Baltrus DA, Yourstone S, Lind A, Guilbaud C, Sands DC, Jones CD, Morris CE, Dangl JL. 2014. Draft genome sequences of a phylogenetically diverse suite of *Pseudomonas syringae* strains from multiple source populations. *Genome Announc.* 2(1):e01195-13. doi:10.1128/genomeA.01195-13.

Copyright © 2014 Baltrus et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to David A. Baltrus, baltrus@email.arizona.edu.

Pseudomonas syringae is well known as a facultative phyto-pathogen of many plant species, including crops, such as tomato, bean, and wheat, as well as the laboratory model *Arabidopsis* (1). Therefore, a majority of *P. syringae*-related research has focused on identifying and understanding the virulence factors that enable this bacterium to cause disease *in planta*. More recently, increased emphasis has been placed on understanding how *P. syringae* strains survive in the environment outside plants, and greater environmental sampling has driven the collection of strains from noncrop sources and symptomless plants (2–4). Here, we describe the draft genome sequences for phylogenetically diverse isolates of *P. syringae* collected from various sources in the United States and France. The sampled locations include nonagricultural environmental sources as well as plants lacking disease symptoms. These sequences provide a wealth of data for genomic comparisons between geographically similar isolates and enable a greater understanding of the evolutionary forces that lead to the emergence of phytopathogens from environmental reservoirs.

The collection of these isolates has been previously described (3–5). Genomic DNA was prepared from populations initiated

with single colonies and purified as per Baltrus et al. (6). The sequencing of each isolate except *P. syringae* USA011 was performed on an Illumina GAII from 36-bp paired-end libraries. A genomic library from USA011 was sequenced using a portion of one Illumina HiSeq lane using 36-bp paired-end libraries. The draft genomes were assembled using SPAdes 2.5.0 without trimming for quality (7). These assemblies are part of a larger project that includes draft genome assemblies for *P. syringae* strains (and their respective accession no.) UB246 (AVEQ00000000), CC1543 (AVEJ00000000), CC94 (AVEA00000000), UB303 (AVDZ00000000), CC1416 (AVEP00000000), CC1458 (AVEN00000000), CC1544 (AVEI00000000), CC1559 (AVEG00000000), CC440 (AVEC00000000), CC457 (AVEB00000000), CC1557 (AVEH00000000), USA007 (AVDY00000000), and CC1583 (AVEF00000000), as well as *P. viridiflava* strains CC1582 (AVDW00000000) and TA043 (AVDV00000000). These assemblies are not included within this announcement due to poor assembly quality.

Nucleotide sequence accession numbers. The nucleotide sequence accession numbers for Genbank are found in [Table 1](#). These sequences and associated metadata can be publicly found at

TABLE 1 Characteristics of the seven *P. syringae* isolates

Isolate	NCBI accession no. ^a	JGI taxon ID	Genome size (bp)	No. of contigs	N ₅₀	Substrate of isolation	Country of isolation	Reference
<i>Pseudomonas syringae</i> CC1417	AVEO00000000	2506783030	5,648,464	210	50,166	Epilithon	USA	3
<i>Pseudomonas syringae</i> CC1466	AVEM00000000	2506783019	5,591,749	294	60,479	<i>Dodecatheon pulchellum</i>	USA	3
<i>Pseudomonas syringae</i> CC1513	AVEL00000000	2506783020	5,725,032	164	85,568	<i>Hutchinsia alpina</i>	France	3
<i>Pseudomonas syringae</i> CC1524	AVEK00000000	2506783021	5,828,366	264	44,224	Stream water	France	4
<i>Pseudomonas syringae</i> CC1629	AVEE00000000	2524614586	5,932,928	261	75,208	Oats	USA	4
<i>Pseudomonas syringae</i> CC1630	AVED00000000	2524614587	6,056,064	283	52,218	Sainfoin	USA	4
<i>Pseudomonas syringae</i> USA011	AVDX00000000	2509276053	6,429,288	198	91,862	Stream water	USA	4

^a Projects have been deposited at GenBank under these accession numbers, but the version described in this paper is version XXXX02000000.

the Joint Genome Institute (JGI) portal (<http://img.jgi.doe.gov/>) with the JGI taxon IDs listed in [Table 1](#).

ACKNOWLEDGMENTS

Funding was provided by National Institutes of Health (NIH) grant no. GM066025 and the Department of Energy Basic Energy Sciences grant no. ER20187 to J. L. Dangl and by NIH National Research Service Ruth Kirchstein Award no. GM082279-01 and startup funds from the University of Arizona to D. A. Baltrus.

J. L. Dangl is an Investigator of the Howard Hughes Medical Institute.

We thank P. Miecskowski, H. Kelkar, and J. Walsh of the UNC-CH HTS core and Amy Chen, Susannah Tringe, and Konstantinos Liolios of the Lawrence Berkeley National Laboratory for their technical assistance.

REFERENCES

- O'Brien HE, Thakur S, Guttman DS. 2011. Evolution of plant pathogenesis in *Pseudomonas syringae*: a genomics perspective. *Annu. Rev. Phytopathol.* **49**: 269–289. <http://dx.doi.org/10.1146/annurev-phyto-072910-095242>.
- Demba Diallo M, Monteil CL, Vinatzer BA, Clarke CR, Glaux C, Guilbaud C, Desbiez CEC, Morris CE. 2012. *Pseudomonas syringae* naturally lacking the canonical type III secretion system are ubiquitous in nonagricultural habitats, are phylogenetically diverse and can be pathogenic. *ISME J.* **6**:1325–1335. <http://dx.doi.org/10.1038/ismej.2011.202>.
- Morris CE, Sands DC, Vinatzer BA, Glaux C, Guilbaud C, Buffière A, Yan S, Dominguez H, Thompson BM. 2008. The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *ISME J.* **2**:321–334. <http://dx.doi.org/10.1038/ismej.2007.113>.
- Morris CE, Sands DC, Vanneste JL, Montarry J, Oakley B, Guilbaud C, Glaux C. 2010. Inferring the evolutionary history of the plant pathogen *Pseudomonas syringae* from its biogeography in headwaters of rivers in North America, Europe, and New Zealand. *mBio* **1**(3):e00107-10. <http://dx.doi.org/10.1128/mBio.00107-10>.
- Morris CE, Glaux C, Latour X, Gardan L, Samson R, Pitrat M. 2000. The relationship of host range, physiology, and genotype to virulence on cantaloupe in *Pseudomonas syringae* from cantaloupe blight epidemics in France. *Phytopathology* **90**:636–646. <http://dx.doi.org/10.1094/PHYTO.2000.90.6.636>.
- Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, Cherkis K, Roach J, Grant SR, Jones CD, Dangl JL. 2011. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathog.* **7**:e1002132. <http://dx.doi.org/10.1371/journal.ppat.1002132>.
- 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. <http://dx.doi.org/10.1089/cmb.2012.0021>.