TECHNICAL ADVANCE

Phytopathogen Genome Announcement: Draft Genome Sequences of 62 *Pseudomonas syringae* Type and Pathotype Strains

Shalabh Thakur,¹ Bevan S. Weir,² and David S. Guttman^{1,3}

¹Department of Cell & Systems Biology, University of Toronto, Toronto, Ontario, Canada; ²Landcare Research, Auckland, New Zealand; and ³Centre for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, Ontario, Canada

Submitted 14 January 2016. Accepted 15 January 2016.

Pseudomonas syringae is a diverse species-complex that includes many important crop pathogens. Here, we report the draft genomes of 62 type and pathotype strains, which provide a genomic reference for the diversity of this species complex and will contribute to the elucidation of the genomic basis of pathogenicity and host specificity.

Genome announcement.

Pseudomonas syringae is a Gammaproteobacteria that is a causal agent of disease on a broad range of agriculturally important crop species as well as being a common isolate from habitats linked to the water cycle (Morris et al. 2008; O'Brien et al. 2011). The species complex is subdivided into approximately 64 pathovars, based on tests related to host of isolation and host-specific disease capacity, and as many as 13 distinct phylogenetic groups (phylogroups), based on multi-locus sequence typing, although the exact breadth of the species complex (and hence, number of phylogroups) is not universally agreed upon (Berge et al. 2014; Hwang et al. 2005; Sarkar and Guttman 2004; Young 2010). Here, we report draft genome sequences of 62 of the type and pathotype strains of *P. syringae*. Three additional type strains have already been sequenced and published, while an additional six were not available. Type strains are the isolate to which the scientific name of that organism is formally attached under the rules of prokaryote nomenclature. Pathotype strains are similar to type strains, with the additional requirement that a pathotype strain has the pathogenic characteristics of its pathovar (i.e., a pathogen of a particular host) (Bull and De Boer 2008).

All strains were obtained from the International Collection of Microorganisms from Plants (ICMP) culture collection. DNA was isolated using the Gentra Puregene yeast and bacteria kit (Qiagen). Purified DNA was resuspended in Tris-EDTA buffer and was quantified with the Qubit dsDNA BR assay kit (ThermoFisher Scientific). Paired-end libraries were generated with an insert size of 300 to 400 bp, using the Illumina Nextera XT DNA library prep kit, following manufacturer's protocols (Illumina). Samples were multiplexed, using the Nextera XT index kit with 96 indices. Sequencing was performed on

Illumina MiSeq, using V2 (300 cycles) chemistry at the University of Toronto Center for the Analysis of Genome Evolution and Function. The qualities of the raw reads were checked using FastQC program (Andrews 2015). Low-quality bases from paired reads were trimmed using the Trimmomatic v0.30 program (Bolger et al. 2014). The trimmed pair reads were de novo assembled into the contigs, using the CLC assembly cell v4.2 program from CLCBio. The sequence fragments <200 bp long were removed from the assembly. Additional filtering was performed to remove low-quality sequence fragments, using protocol described in IMG database version 4 for single cell data decontamination (Markowitz et al. 2014). Protein-coding genes were predicted using combined information from Glimmer, Genemark, Prodigal, and FragGeneScan (Besemer and Borodovsky 2005; Delcher et al. 2007; Hyatt et al. 2010; Rho et al. 2010). Annotations were made via BLAST search against the Uniprot/Swiss-Prot database (Altschul et al. 1997; Boutet et al. 2007).

Nucleotide sequence accession numbers.

The genome sequences have been deposited at GenBank under the accession numbers listed in Table 1.

LITERATURE CITED

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.
- Andrews, S. C. 2015. FastQC v0.11.3. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Babraham Bioinformatics, Cambridge. Published online.
- Berge, O., Monteil, C. L., Bartoli, C., Chandeysson, C., Guilbaud, C., Sands, D. C., and Morris, C. E. 2014. A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex. *PLoS One* 9:e105547.
- Besemer, J. and Borodovsky M. 2005. GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res.* 33: W451-W454.
- Bolger, A. M., Lohse, M., and Usadel, B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114-2120.
- Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M., and Bairoch, A. 2007. UniProtKB/Swiss-Prot. *Methods Mol. Biol.* 406:89-112.
- Bull, C. T., and De Boer, S. H. 2008. Demystifying the nomenclature of bacterial plant pathogens. *J. Plant Pathol.* 90:403-417.
- Delcher, A. L., Bratke, K. A., Powers, E. C., and Salzberg, S. L. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23:673-679.
- Hwang, M. S., Morgan, R. L., Sarkar, S. F., Wang, P. W., and Guttman, D. S. 2005. Phylogenetic characterization of virulence and resistance

Corresponding author: D. S. Guttman; E-mail: david.guttman@utoronto.ca

Table 1. Characteristics of the sequenced *Pseudomonas syringae* type and pathotype isolates^a

Line	Species	Pathovar	Designation	Type	Accession	Genome size (bp)	No. of contigs	N50
1	<i>P. syringae</i>	<i>aceris</i>	ICMP 2802	PT	LJPM00000000	6,301,287	684	21,044
2	<i>P. syringae</i>	<i>aesculi</i>	ICMP 8947	PT	LJPN00000000	5,831,425	657	18,777
3	<i>P. syringae</i>	<i>alisalensis</i>	ICMP 15200	PT	LJPP00000000	6,102,306	400	37,474
4	<i>P. syringae</i>	<i>antirrhini</i>	ICMP 4303	PT	LJPT00000000	6,132,316	189	59,555
5	<i>P. syringae</i>	<i>apii</i>	ICMP 2814	PT	LJPR00000000	6,054,618	203	71,825
6	<i>P. syringae</i>	<i>aptata</i>	ICMP 459	PT	LJRP00000000	5,977,528	149	129,424
7	<i>P. syringae</i>	<i>atrofaciens</i>	ICMP 4394	PT	LJPO00000000	6,012,445	248	56,132
8	<i>P. syringae</i>	<i>atropurpurea</i>	ICMP 4457	PT	LJPS00000000	5,793,237	167	97,353
9	<i>P. syringae</i>	<i>berberidis</i>	ICMP 4116	PT	LJPU00000000	6,263,633	342	55,290
10	<i>P. syringae</i>	<i>broussonetiae</i>	ICMP 13650	PT	LJPV00000000	5,802,206	623	26,393
11	<i>P. syringae</i>	<i>castaneae</i>	ICMP 9419	PT	LJQD00000000	6,378,468	614	8,895
12	<i>P. syringae</i>	<i>cerasicola</i>	ICMP 17524	PT	LJQA00000000	5,526,840	970	11,460
13	<i>P. syringae</i>	<i>ciccaronei</i>	ICMP 5710	PT	LJPY00000000	5,757,646	496	39,533
14	<i>P. syringae</i>	<i>coriandricola</i>	ICMP 12471	PT	LJPZ00000000	5,846,834	426	38,916
15	<i>P. syringae</i>	<i>coryli</i>	ICMP 17001	PT	LJQC00000000	5,966,981	1,143	10,167
16	<i>P. syringae</i>	<i>cunninghamiae</i>	ICMP 11894	PT	LJQE00000000	5,782,797	389	35,340
17	<i>P. syringae</i>	<i>daphniphylli</i>	ICMP 9757	PT	LJQF00000000	5,933,542	528	29,819
18	<i>P. syringae</i>	<i>delphinii</i>	ICMP 529	PT	LJQH00000000	6,144,113	411	27,459
19	<i>P. syringae</i>	<i>dendropanacis</i>	ICMP 9150	PT	LJQG00000000	5,666,913	373	33,960
20	<i>P. syringae</i>	<i>erobotryae</i>	ICMP 4455	PT	LJQI00000000	6,161,132	440	28,807
21	<i>P. savastanoi</i>	<i>fraxini</i>	ICMP 7711	PT	LLJL00000000	6,015,697	481	38,704
22	<i>P. syringae</i>	<i>garcae</i>	ICMP 4323	PT	LJQK00000000	5,509,919	109	114,831
23	<i>P. savastanoi</i>	<i>glycinea</i>	ICMP 2189	PT	LJQL00000000	5,983,696	430	37,582
24	<i>P. syringae</i>	<i>helianthi</i>	ICMP 4531	PT	LJQM00000000	6,208,747	277	54,333
25	<i>P. syringae</i>	<i>hibisci</i>	ICMP 9623	PT	LJQN00000000	6,111,700	156	99,663
26	<i>P. syringae</i>	<i>lachrymans</i>	ICMP 3507	PT	LJQP00000000	6,121,722	512	21,694
27	<i>P. syringae</i>	<i>lapsa</i>	ICMP 3947	PT	LJQQ00000000	5,850,559	118	124,931
28	<i>P. syringae</i>	<i>maculicola</i>	ICMP 3935	PT	LJQR00000000	6,064,255	229	72,312
29	<i>P. syringae</i>	<i>mellea</i>	ICMP 5711	PT	LJQS00000000	5,895,398	260	49,083
30	<i>P. syringae</i>	<i>mori</i>	ICMP 4331	PT	LJQU00000000	5,818,523	619	20,579
31	<i>P. syringae</i>	<i>myricae</i>	ICMP 7118	PT	LJQV00000000	5,953,908	403	37,332
32	<i>P. savastanoi</i>	<i>nerii</i>	ICMP 16943	PT	LJQW00000000	5,715,739	341	38,255
33	<i>P. syringae</i>	<i>oryzae</i>	ICMP 9088	PT	LJQX00000000	5,502,925	86	172,792
34	<i>P. syringae</i>	<i>papulans</i>	ICMP 4048	PT	LJRB00000000	6,097,618	298	65,788
35	<i>P. savastanoi</i>	<i>phaseolicola</i>	ICMP 2740	PT	LJQZ00000000	5,967,551	287	52,236
36	<i>P. syringae</i>	<i>philadelphii</i>	ICMP 8903	PT	LJQY00000000	6,150,448	479	55,294
37	<i>P. syringae</i>	<i>photiniae</i>	ICMP 7840	PT	LJQO00000000	5,954,075	624	23,771
38	<i>P. syringae</i>	<i>porri</i>	ICMP 8961	PT	LJRA00000000	5,828,252	320	63,648
39	<i>P. syringae</i>	<i>primulae</i>	ICMP 3956	PT	LJRC00000000	5,895,099	310	36,169
40	<i>P. savastanoi</i>	<i>retacarpa</i>	ICMP 16945	PT	LJRD00000000	5,673,161	510	27,300
41	<i>P. syringae</i>	<i>rhapiolepidis</i>	ICMP 9756	PT	LJRE00000000	5,483,229	446	28,865
42	<i>P. syringae</i>	<i>ribicola</i>	ICMP 3882	PT	LJRF00000000	5,893,023	219	52,114
43	<i>P. savastanoi</i>	<i>savastanoi</i>	ICMP 4352	T/PT	LJRJ00000000	6,023,516	321	55,596
44	<i>P. syringae</i>	<i>sesame</i>	ICMP 763	PT	LJRG00000000	6,028,879	462	44,661
45	<i>P. syringae</i>	<i>solidagae</i>	ICMP 16925	PT	LJRH00000000	5,980,624	382	37,149
46	<i>P. syringae</i>	<i>spinaceae</i>	ICMP 16929	PT	LJRI00000000	6,039,219	1,657	5,991
47	<i>P. syringae</i>	<i>syringae</i>	ICMP 3023	PT	LJRK00000000	6,073,368	179	76,745
48	<i>P. syringae</i>	<i>tabaci</i>	ICMP 2835	PT	LJRL00000000	6,239,582	170	106,290
49	<i>P. syringae</i>	<i>tagetis</i>	ICMP 4091	PT	LJRM00000000	6,042,642	208	69,652
50	<i>P. syringae</i>	<i>theae</i>	ICMP 3923	PT	LJRU00000000	6,233,091	369	38,965
51	<i>P. syringae</i>	<i>tomato</i>	ICMP 2844	PT	LJRN00000000	6,261,696	344	45,004
52	<i>P. tremiae</i>	<i>tremiae</i>	ICMP 9151	T/PT	LJRO00000000	5,930,024	510	133,195
53	<i>P. syringae</i>	<i>ulmi</i>	ICMP 3962	PT	LJRW00000000	6,244,367	455	40,538
54	<i>P. syringae</i>	<i>viburni</i>	ICMP 3963	PT	LJRR00000000	6,405,699	469	35,691
55	<i>P. syringae</i>	<i>zizaniae</i>	ICMP 8921	PT	LJRT00000000	5,729,547	152	257,833
56	<i>P. amygdali</i>		ICMP 3918	T	LJPQ00000000	5,934,986	222	54,621
57	<i>P. cannabina</i>	<i>cannabina</i>	ICMP 2823	T/PT	LJPX00000000	5,916,843	702	26,269
58	<i>P. caricapapayae</i>		ICMP 2855	T	LJPW00000000	6,257,942	193	78,662
59	<i>P. congelans</i> ^b		ICMP 19117	T	LJQB00000000	5,630,811	86	121,405
60	<i>P. ficuserectae</i>		ICMP 7848	T	LJQJ00000000	5,655,233	624	17,896
61	<i>P. meliae</i>		ICMP 6289	T	LJQT00000000	5,098,132	504	25,414
62	<i>P. viridiflava</i>		ICMP 2848	T	LJRS00000000	5,902,365	254	45,282

(continued on next page)

^a Type strains that have been previously sequenced: *P. syringae* pv. *avellanae* ICMP 9746 (ATDK00000000), *P. syringae* pv. *actinidiae* ICMP 9617 (AFTH01000000), *P. syringae* pv. *coronafaciens* ICMP 3113 (JSED01000000); Type strains that have not been sequenced: *P. syringae* pv. *avii* ICMP14479, *P. syringae* pv. *persicae* ICMP5846, *P. syringae* pv. *striaefaciens* ICMP3961, *P. syringae* pv. *morsprunorum* ICMP18416, *P. syringae* pv. *pisi* ICMP2452, *P. syringae* pv. *passiflorae* ICMP 129. N50 is the weighted median of the size of the genome assembly, such that 50% of the entire assembly is contained in contigs equal to or larger than this value; CDS = coding sequences.

^b This strain is not a confirmed plant pathogen.

Table 1. (continued from previous page)

Line	Genome coverage	No. of CDS	Phylogroup	Country of isolation	Host common name	Host scientific name
1	18.76	5,772	2	U.S.A.	Maple	<i>Acer</i> sp.
2	28.82	5,701	3	India	Horse chestnut	<i>Aesculus indica</i> Colebr.
3	54.89	5,259	5	U.S.A.	Broccoli raab	<i>Brassica rapa</i> subsp. <i>rapa</i>
4	48.95	5,584	1	U.K.	Snapdragon	<i>Antirrhinum majus</i>
5	72.17	5,579	1	U.S.A.	Celery	<i>Apium graveolens</i> var. <i>dulce</i> (Mill.) Pers.
6	39.99	5,318	2	U.S.A.	Sugar beet	<i>Beta vulgaris</i>
7	32.69	5,278	2	New Zealand	Wheat	<i>Triticum aestivum</i>
8	32.36	5,154	4	Japan	Italian ryegrass	<i>Lolium multiflorum</i> Lam.
9	167.11	5,785	1	New Zealand	Barberry	<i>Berberis</i> sp.
10	54.39	5,582	3	Japan	Paper mulberry	<i>Broussonetia kazinoki</i> Sieb. × <i>Broussonetia papyrifera</i> Vent.
11	65.81	6,389	3	Japan	Japanese chestnut	<i>Castanea crenata</i> Sieb. & Zucc.
12	17.45	5,395	3	Japan	Cherry	<i>Prunus</i> × <i>yedoensis</i>
13	87.57	5,528	3	Italy	Carob	<i>Ceratonia siliqua</i>
14	49.93	5,257	5	Germany	Cilantro	<i>Coriandrum sativum</i>
15	15.01	5,554	2	Italy	European hazelnut	<i>Corylus avellana</i>
16	38.75	5,463	3	China	Chinese fir	<i>Cunninghamia lanceolata</i>
17	112.73	5,550	3	Japan	Himeyuzuriha	<i>Daphniphyllum teijsmanni</i> Zoll.
18	50.72	5,386	1	New Zealand	Delphinium	<i>Delphinium</i> sp.
19	51.39	5,326	3	Japan	Kakuremino	<i>Dendropanax trifidus</i>
20	49.85	6,069	3	U.S.A.	Loquat	<i>Eriobotrya japonica</i>
21	67.98	5,606	3	Netherlands	European ash	<i>Fraxinus excelsior</i>
22	39.56	4,875	4	Brazil	Coffee	<i>Coffea arabica</i>
23	27.1	5,741	3	New Zealand	Soybean cv. disoy	<i>Glycine max</i>
24	41.45	5,585	6	Mexico	Sunflower	<i>Helianthus annuus</i>
25	102.65	5,608	3	U.S.A.	Chinese hibiscus	<i>Hibiscus rosa-sinensis</i>
26	34.74	6,152	3	Hungary	Cucumber	<i>Cucumis sativus</i>
27	42.19	5,156	2		Corn hybrid	<i>Zea</i> sp.
28	99.1	5,588	1	New Zealand	Cauliflower	<i>Brassica oleracea</i> var. <i>Botrytis</i>
29	20.31	5,445	3	Japan	Tobacco	<i>Nicotiana tabacum</i>
30	41.5	5,837	3	Hungary	White mulberry	<i>Morus alba</i>
31	50.44	5,599	3	Japan	Chinese bayberry	<i>Myrica rubra</i>
32	45.19	5,394	3	Spain	Oleander	<i>Nerium oleander</i>
33	99.77	4,825	4	Japan	Rice cv. akihikari	<i>Oryza sativa</i>
34	31.27	5,636	2	Canada	Apple	<i>Malus</i> × <i>domestica</i> Borkh.
35	32.17	5,720	3	Canada	Bean	<i>Phaseolus vulgaris</i>
36	153.36	5,751	1	U.K.	Mock orange cv. aureus	<i>Philadelphus coronarius</i>
37	39.53	5,724	3	Japan	Japanese photinia	<i>Photinia glabra</i>
38	33.99	5,195	4	France	Leek	<i>Allium porrum</i>
39	70.03	4,982	7	U.S.A.	Polyanthus	<i>Primula</i> sp.
40	45.29	5,412	3	Spain	Broom bush	<i>Retama sphaerocarpa</i>
41	70.93	5,280	3	Japan	Sharinbai	<i>Rhaphiolepis umbellata</i>
42	80.2	4,953	7	U.S.A.	Golden currant	<i>Ribes aureum</i>
43	66.92	5,728	3	Yugoslavia	Olive	<i>Olea europaea</i>
44	33.11	5,677	3	Yugoslavia	Sesame	<i>Sesamum indicum</i>
45	32.43	5,412	2	Japan	Tall golden rod	<i>Solidago altissima</i>
46	16.85	6,599	1	Japan	Spinach	<i>Spinacia oleracea</i>
47	72.61	5,350	2	U.K.	Lilac	<i>Syringa vulgaris</i>
48	92.91	5,707	3	Hungary	Tobacco	<i>Nicotiana tabacum</i>
49	55.37	5,383	6	Zimbabwe	African marigold	<i>Tagetes erecta</i>
50	96.66	5,423	1	Japan	Tea	<i>Camellia sinensis</i>
51	62.33	5,626	1	U.K.	Tomato	<i>Lycopersicon esculentum</i>
52	76.53	5,545	4	Japan	Urajiroenoki	<i>Trema orientalis</i> Blume
53	73.41	5,839	3	Yugoslavia	Elm	<i>Ulmus</i> sp.
54	59.32	5,866	1	U.S.A.	Viburnum	<i>Viburnum</i> sp.
55	278.06	5,146	4	U.S.A.	Wild rice	<i>Zizania aquatica</i>
56	50.99	5,398	3	Greece	Almond	<i>Prunus dulcis</i>
57	62.72	5,463	5	Hungary	Indian hemp	<i>Cannabis sativa</i>
58	76.76	5,638	6	Brazil	Pawpaw, papaya	<i>Carica papaya</i>
59	91.26	4,906	2	Germany	Grass phyllosphere	NA
60	19.02	5,400	3	Japan	Inubiwa	<i>Ficus erecta</i>
61	34.26	5,155	3	Japan	Chinaberry	<i>Melia azedarach</i>
62	66.78	4,986	7	Switzerland	Bean	<i>Phaseolus vulgaris</i>

- phenotypes of *Pseudomonas syringae*. Appl. Environ. Microbiol. 71: 5182-5191.
- Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. 2010. Prodigal: Prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11: 119.
- Markowitz, V. M., Chen, I. M., Palaniappan, K., Chu, K., Szeto, E., Pillay, M., Ratner, A., Huang, J., Woyke, T., Huntemann, M., Anderson, I., Billis, K., Varghese, N., Mavromatis, K., Pati, A., Ivanova, N. N., and Kyrpides, N. C. 2014. IMG 4 version of the integrated microbial genomes comparative analysis system. Nucleic Acids Res. 42:D560-D567.
- Morris, C. E., Sands, D. C., Vinatzer, B. A., Glaux, C., Guilbaud, C., Buffière, A., Yan, S., Dominguez, H., and Thompson, B. M. 2008. The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. ISME J. 2:321-334.
- O'Brien, H. E., Thakur, S., and Guttman, D. S. 2011. Evolution of plant pathogenesis in *Pseudomonas syringae*: A genomics perspective. Annu. Rev. Phytopathol. 49:269-289.
- Rho, M., Tang, H., and Ye, Y. 2010. FragGeneScan: Predicting genes in short and error-prone reads. Nucleic Acids Res. 38:e191.
- Sarkar, S. F., and Guttman, D. S. 2004. Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. Appl. Environ. Microbiol. 70:1999-2012.
- Young, J. M. 2010. Taxonomy of *Pseudomonas syringae*. J. Plant Pathol. 92:S5-S14.

AUTHOR-RECOMMENDED INTERNET RESOURCE

International Collection of Microorganisms from Plants (ICMP) culture collection: www.landcareresearch.co.nz/resources/collections/icmp