

UC Santa Cruz

UC Santa Cruz Previously Published Works

Title

Complete Genome Assembly of Yersinia pseudotuberculosis IP2666pIB1

Permalink

<https://escholarship.org/uc/item/6sr691qx>

Journal

Microbiology Resource Announcements, 8(7)

ISSN

2169-8287

Authors

Zoubeidi, Adam
Schwiesow, Leah
Auerbuch, Victoria
et al.

Publication Date

2019-02-14

DOI

10.1128/mra.01592-18

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



Complete Genome Assembly of *Yersinia pseudotuberculosis* IP2666pIB1

Adam Zoubeidi,^a Leah Schwiesow,^{b*} Victoria Auerbuch,^a Hanh N. Lam^a

^aDepartment of Microbiology and Environmental Toxicology, University of California, Santa Cruz, California, USA

^bDepartment of Molecular, Cell, and Developmental Biology, University of California, Santa Cruz, California, USA

ABSTRACT *Yersinia pseudotuberculosis*, closely related to *Yersinia pestis*, is a human pathogen and model organism for studying bacterial pathogenesis. To aid in genomic analysis and understanding bacterial virulence, we sequenced and assembled the complete genome of the human pathogen *Yersinia pseudotuberculosis* IP2666pIB1.

Three species within the *Yersinia* genus are human pathogens, *Yersinia pestis*, *Yersinia pseudotuberculosis*, and *Yersinia enterocolitica*. All three pathogenic *Yersinia* species harbor a 70-kb virulence plasmid referred to as pYV, which encodes a type III secretion system critical for virulence (1–5). Three mouse virulent *Y. pseudotuberculosis* strains, YPIII, IP32953, and IP2666, are commonly used for analysis of *Y. pseudotuberculosis* pathogenicity. The IP2666pIB1 strain, which has been used as the basis for a number of studies (6–8), was generated by the Bliska lab (9) by curing the IP2666 strain of its native virulence plasmid and inserting the well-characterized pYV virulence plasmid from YPIII, called pIB1. Although many bacterial genome sequences were released, the IP2666pIB1 genome sequence was not available. Here, we present a complete sequence of the chromosome and pIB1 plasmid of *Y. pseudotuberculosis* IP2666pIB1.

Y. pseudotuberculosis IP2666pIB1 was grown in 2xYT (yeast extract-tryptone) at 26°C, shaken overnight. The culture was diluted to an optical density (OD₆₀₀) of 0.1 and grown at 26°C, and cells were pelleted when the culture reached an OD₆₀₀ of 0.8. Genomic DNA was extracted with a DNeasy blood and tissue kit (Qiagen). The samples were sent to the DNA Technologies Core at the University of California, Davis, for library preparation with the DNA sequencing kit 4.0 v2 with C4 chemistry, PacBio RS II sequencing (library preparation followed by size selection of 15 kb with Blue Pippin), and MiSeq paired-end sequencing with a 300-bp read length.

Trimmomatic version 0.36 (10) was used to trim off low-quality bases and adapter sequences from MiSeq reads. The trimmed, paired-end reads were used to assemble the genome. PacBio sequences were trimmed with Canu 1.7 (11) during assembly. The average PacBio sequence length is 20 kb. The genome sequence was assembled from two data sets from PacBio long reads (161,000 reads) and MiSeq paired-end reads (2,956,000 reads), resulting in more than 50× coverage. The two data sets were assembled together with SPAdes v3.11.1 (12), and PacBio reads were assembled with Canu 1.7 using default parameters (11). Outputs from the two programs were aligned and visualized in SeaView v4.5.2 (13). The assembled genome was manually inspected and curated in Artemis 16.0.0 (14) to a high-quality completion (15). Briefly, at positions where there were differences between the two assembled sequences aligned and visualized in SeaView, the high-quality sequence reads aligned to the genome were inspected in Artemis to reconcile the disagreements. The genome sequence was annotated with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (16, 17).

The chromosome size is 4,614,856 bp, with 47.5% GC content, 4,115 predicted

Citation Zoubeidi A, Schwiesow L, Auerbuch V, Lam HN. 2019. Complete genome assembly of *Yersinia pseudotuberculosis* IP2666pIB1. *Microbiol Resour Announc* 8:e01592-18. <https://doi.org/10.1128/MRA.01592-18>.

Editor Vincent Bruno, University of Maryland School of Medicine

Copyright © 2019 Zoubeidi 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 Hanh N. Lam, hanlam@ucsc.edu.

* Present address: Leah Schwiesow, Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA.

Received 30 November 2018

Accepted 22 January 2019

Published 14 February 2019

coding sequences, 102 ribosomal and transfer RNAs, and 182 pseudogenes. The plB1 virulence plasmid showed 44.8% GC content and 96 coding sequences. It is important to note that, unlike *Y. pestis* CO92 (GenBank accession number [NC_003143](#)) and *Y. pseudotuberculosis* IP32953 (GenBank accession number [NZ_CP009712](#)), the entire high-pathogenicity island on the *pgm* locus (18) containing yersiniabactin biosynthetic genes is absent in this strain, which is similar to *Y. pseudotuberculosis* YPIII (GenBank accession number [CP009792](#)) (19).

Data availability. The *Yersinia pseudotuberculosis* IP2666plB1 project has been deposited in the National Center of Biotechnology Information (NCBI) under the accession numbers [CP032566](#) and [CP032567](#) (BioProject number [PRJNA475632](#)). The raw sequencing reads have also been submitted to the Sequence Read Archive (SRA) under accession numbers [SRR8061175](#), [SRR8061176](#), and [SRR8061177](#).

ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award number R01AI119082 (to V.A.).

We thank Jim Bliska for his gift of *Yersinia pseudotuberculosis* IP2666plB1 for our study.

REFERENCES

- Portnoy DA, Wolf-Watz H, Bolin I, Beeder AB, Falkow S. 1984. Characterization of common virulence plasmids in *Yersinia* species and their role in the expression of outer membrane proteins. *Infect Immun* 43: 108–114.
- Ben-Gurion R, Shafferman A. 1981. Essential virulence determinants of different *Yersinia* species are carried on a common plasmid. *Plasmid* 5:183–187. [https://doi.org/10.1016/0147-619X\(81\)90019-6](https://doi.org/10.1016/0147-619X(81)90019-6).
- Gemski P, Lazere JR, Casey T. 1980. Plasmid associated with pathogenicity and calcium dependency of *Yersinia enterocolitica*. *Infect Immun* 27:682–685.
- Ferber DM, Brubaker RR. 1981. Plasmids in *Yersinia pestis*. *Infect Immun* 31:839–841.
- Gemski P, Lazere JR, Casey T, Wohlhieter JA. 1980. Presence of a virulence-associated plasmid in *Yersinia pseudotuberculosis*. *Infect Immun* 28:1044–1047.
- Garrity-Ryan LK, Kim OK, Balada-Llasat J-M, Bartlett VJ, Verma AK, Fisher ML, Castillo C, Songsungthong W, Tanaka SK, Levy SB, Mecsas J, Aleksun MN. 2010. Small molecule inhibitors of LcrF, a *Yersinia pseudotuberculosis* transcription factor, attenuate virulence and limit infection in a murine pneumonia model. *Infect Immun* 78:4683–4690. <https://doi.org/10.1128/IAI.01305-09>.
- Fisher ML, Castillo C, Mecsas J. 2007. Intranasal inoculation of mice with *Yersinia pseudotuberculosis* causes a lethal lung infection that is dependent on *Yersinia* outer proteins and PhoP. *Infect Immun* 75:429–442. <https://doi.org/10.1128/IAI.01287-06>.
- Auerbuch V, Isberg RR. 2007. Growth of *Yersinia pseudotuberculosis* in mice occurs independently of Toll-like receptor 2 expression and induction of interleukin-10. *Infect Immun* 75:3561–3570. <https://doi.org/10.1128/IAI.01497-06>.
- Bliska JB, Guan KL, Dixon JE, Falkow S. 1991. Tyrosine phosphate hydrolysis of host proteins by an essential *Yersinia* virulence determinant. *Proc Natl Acad Sci USA* 88:1187–1191. <https://doi.org/10.1073/pnas.88.4.1187>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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. <https://doi.org/10.1089/cmb.2012.0021>.
- Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 27:221–224. <https://doi.org/10.1093/molbev/msp259>.
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16:944–945. <https://doi.org/10.1093/bioinformatics/16.10.944>.
- Chain PSG, Grafham DV, Fulton RS, FitzGerald MG, Hostetler J, Muzny D, Ali J, Birren B, Bruce DC, Buhay C, Cole JR, Ding Y, Dugan S, Field D, Garrity GM, Gibbs R, Graves T, Han CS, Harrison SH, Highlander S, Hugenholtz P, Khouri HM, Kodira CD, Kolker E, Kyrpides NC, Lang D, Lapidus A, Malfatti SA, Markowitz V, Metha T, Nelson KE, Parkhill J, Pitluck S, Qin X, Read TD, Schmutz J, Sozhamannan S, Sterk P, Strausberg RL, Sutton G, Thomson NR, Tiedje JM, Weinstock G, Wollam A, Genomic Standards Consortium Human Microbiome Project Jumpstart Consortium, Detter JC. 2009. Genome project standards in a new era of sequencing. *Science* 326:236–237. <https://doi.org/10.1126/science.1180614>.
- Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res* 46:D851–D860. <https://doi.org/10.1093/nar/gkx1068>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44: 6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- Buchrieser C, Rusniok C, Frangeul L, Couve E, Billault A, Kunst F, Carniel E, Glaser P. 1999. The 102-kilobase *pgm* locus of *Yersinia pestis*: sequence analysis and comparison of selected regions among different *Yersinia pestis* and *Yersinia pseudotuberculosis* strains. *Infect Immun* 67:4851–4861.
- Rakin A, Schneider L, Podladchikova O. 2012. Hunger for iron: the alternative siderophore iron scavenging systems in highly virulent *Yersinia*. *Front Cell Infect Microbiol* 2:151. <https://doi.org/10.3389/fcimb.2012.00151>.