

# Draft Genome Sequences of Five *Yersinia pseudotuberculosis* ST19 Isolates and One Isolate Variant

Mikhail E. Platonov,<sup>a</sup> Yann Blouin,<sup>b,c</sup> Vera V. Evseeva,<sup>a</sup> Maxim V. Afanas'ev,<sup>d</sup> Christine Pourcel,<sup>b,c</sup> Sergey V. Balakhonov,<sup>d</sup> Gilles Vergnaud,<sup>b,c,e</sup> Andrey P. Anisimov<sup>a</sup>

State Research Center for Applied Microbiology and Biotechnology, Obolensk, Russia<sup>a</sup>; Université Paris-Sud, Institut de Génétique et Microbiologie, Orsay, France<sup>b</sup>; CNRS, Centre National de la Recherche Scientifique, Orsay, France<sup>c</sup>; Irkutsk Antiplague Research Institute of Siberia and the Far East, Irkutsk, Russia<sup>d</sup>; DGA/MRIS—Mission pour la Recherche et l'Innovation Scientifique, Bagneux, France<sup>e</sup>

**We report the first draft genome sequences of five *Yersinia pseudotuberculosis* isolates of sequence type (ST) 19 and of a variant from one of the five isolates. The total length of assemblies ranged from 4,226,485 bp to 4,274,148 bp, including between 3,808 and 3,843 predicted coding sequences.**

Received 22 February 2013 Accepted 13 March 2013 Published 11 April 2013

**Citation** Platonov ME, Blouin Y, Evseeva V, Afanas'ev MV, Pourcel C, Balakhonov SV, Vergnaud G, Anisimov AP. 2013. Draft genome sequences of five *Yersinia pseudotuberculosis* ST19 isolates and one isolate variant. *Genome Announc*. 1(2):e00122-13. doi:10.1128/genomeA.00122-13.

**Copyright** © 2013 Platonov et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Gilles Vergnaud, gilles.vergnaud@u-psud.fr, or Andrey P. Anisimov, anisimov@obolensk.org.

It has been proposed that *Yersinia pestis*, the etiologic agent of plague, emerged from a *Yersinia pseudotuberculosis* O:1b ancestor (1, 2) somewhere in Eurasia (3–5). Compared to the information available about the genetically and phenotypically monomorphic *Y. pestis*, which has been the focus of detailed investigations in terms of genetic diversity, population genetics, and microevolution (3–7), relatively little is known about *Y. pseudotuberculosis* (8). This pair, *Y. pseudotuberculosis* and *Y. pestis*, represent a very promising model for the investigation of microevolution and speciation in bacteria.

To gain insights into the evolutionary origin of *Y. pestis*, we have analyzed by multilocus variable-number tandem-repeat (VNTR) analysis (MLVA) a collection of 143 *Y. pseudotuberculosis* isolates. The isolates were sampled from 1940 to 2008 from different sources in Siberia, the Far East, central Asia, and the Caucasus and northwest regions of the Russian Federation. The MLVA assay used 14 loci (ms01, ms04, ms05, ms24, ms25, ms27, ms31, ms35, ms38, ms40, ms41, ms56, ms68, and ms74) (9). The five *Y. pseudotuberculosis* isolates that clustered closest to *Y. pestis* were further analyzed by multilocus sequence typing and were shown to belong to the globally spread *Y. pseudotuberculosis* clone of sequence type 19 (ST19) O:3 (8, 10). They were isolated from small wild rodents, and interestingly, three of the selected strains were misidentified as *Y. pestis* at the moment of isolation.

Whole-genome sequencing was performed using Illumina GA IIx (Illumina, Inc.) by generating paired-end libraries with an insert size of 300 bp, according to the manufacturer's instructions. The read length was 75 bp, and the number of reads that passed Illumina quality filters varied from 7.7 to 16 million, corresponding to 760 Mb to 1.6 Gb of high-quality data. For each genome, six million reads were *de novo* assembled using Velvet with a *k* value of 29, and the coding sequences (CDSs) were predicted using BioNumerics v6.6 (Applied Maths, Belgium). A comparison with the complete genomes of *Y. pestis* CO92 and *Y. pseudotuberculosis* strains IP31758, IP32953, and YPIII for function annotations was

also performed using BioNumerics v6.6. Finally, we obtained 221 to 262 contigs for each genome. The mean total length of contigs for each strain was 4,261,235 bp, comprising, on average, 3,831 predicted CDSs. This corresponds to 90% (total length) and 93% (predicted CDSs) of the values for the complete genomes of *Y. pseudotuberculosis* strains IP31758, IP32953, and YPIII (accession no. NC\_009708.1, NC\_006155.1, and NC\_010465.1, respectively). The average G+C content of the draft genomes is 47.53%.

The comparative genomic analysis among strains B-6796 (C-2), B-6862 (797), B-6863 (1384), B-6864 (8421-S), B-6865 (2430-1), and B-6866 (2430-2) and other available *Y. pseudotuberculosis* and *Y. pestis* isolates (i) will provide information on the speciation and evolution of *Y. pestis* and (ii) might explain the means and mechanisms of the global spread of *Y. pseudotuberculosis* clone ST19 O:3. A detailed report of a full comparative genomic analysis will be included in a future publication.

**Nucleotide sequence accession numbers.** The draft genome sequences for the strains B-6796, B-6862, B-6863, B-6864, B-6865, and B-6866 have been included in the European Nucleotide Archive at EMBL-EBI (<http://www.ebi.ac.uk/ena/data/view/>) under accession no. CAQT01000001 to CAQT01000262, CAQU01000001 to CAQU01000247, CAQV01000001 to CAQV01000221, CAQW01000001 to CAQW01000237, CAQX01000001 to CAQX01000240, and CAQY01000001 to CAQY01000243, respectively.

## ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 10-04-91057) and the Centre National de la Recherche Scientifique (CNRS), France (PICS project no. 5516).

## REFERENCES

1. Achtman M, Zurth K, Morelli G, Torrea G, Guiyoule A, Carniel E. 1999. *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* 96:14043–14048.
2. Skurnik M, Peippo A, Ervelä E. 2000. Characterization of the O-antigen

- gene clusters of *Yersinia pseudotuberculosis* and the cryptic O-antigen gene cluster of *Yersinia pestis* shows that the plague bacillus is most closely related to and has evolved from *Y. pseudotuberculosis* serotype O:1b. *Mol. Microbiol.* 37:316–330.
3. Cui Y, Li Y, Gorgé O, Platonov ME, Yan Y, Guo Z, Pourcel C, Dentovskaya SV, Balakhonov SV, Wang X, Song Y, Anisimov AP, Vergnaud G, Yang R. 2008. Insight into microevolution of *Yersinia pestis* by clustered regularly interspaced short palindromic repeats. *PLoS One* 3:e2652. <http://dx.doi.org/10.1371/journal.pone.0002652>.
  4. Li Y, Cui Y, Hauck Y, Platonov ME, Dai E, Song Y, Guo Z, Pourcel C, Dentovskaya SV, Anisimov AP, Yang R, Vergnaud G. 2009. Genotyping and phylogenetic analysis of *Yersinia pestis* by MLVA: insights into the worldwide expansion of Central Asia plague foci. *PLoS One* 4:e6000. <http://dx.doi.org/10.1371/journal.pone.0006000>.
  5. Morelli G, Song Y, Mazzoni CJ, Eppinger M, Roumagnac P, Wagner DM, Feldkamp M, Kusecek B, Vogler AJ, Li Y, Cui Y, Thomson NR, Jombart T, Leblois R, Lichtner P, Rahalison L, Petersen JM, Balloux F, Keim P, Wirth T, Ravel J, Yang R, Carniel E, Achtman M. 2010. *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. *Nat. Genet.* 42:1140–1143.
  6. Riehm JM, Vergnaud G, Kiefer D, Damdindorj T, Dashdavaa O, Khurelsukh T, Zöller L, Wölfel R, Le Flèche P, Scholz HC. 2012. *Yersinia pestis* lineages in Mongolia. *PLoS One* 7:e30624. <http://dx.doi.org/10.1371/journal.pone.0030624>.
  7. Cui Y, Yu C, Yan Y, Li D, Li Y, Jombart T, Weinert LA, Wang Z, Guo Z, Xu L, Zhang Y, Zheng H, Qin N, Xiao X, Wu M, Wang X, Zhou D, Qi Z, Du Z, Wu H, Yang X, Cao H, Wang H, Wang J, Yao S, Rakin A, Li Y, Falush D, Balloux F, Achtman M, Song Y, Wang J, Yang R. 2013. Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*. *Proc. Natl. Acad. Sci. U. S. A.* 110:577–582.
  8. Laukkanen-Ninios R, Didelot X, Jolley KA, Morelli G, Sangal V, Kristo P, Brehony C, Imori PF, Fukushima H, Siitonen A, Tseneva G, Voskresenskaya E, Falcao JP, Korkeala H, Maiden MC, Mazzoni C, Carniel E, Skurnik M, Achtman M. 2011. Population structure of the *Yersinia pseudotuberculosis* complex according to multilocus sequence typing. *Environ. Microbiol.* 13:3114–3127.
  9. Le Flèche P, Hauck Y, Onteniente L, Prieur A, Denoeud F, Ramisse V, Sylvestre P, Benson G, Ramisse F, Vergnaud G. 2001. A tandem repeats database for bacterial genomes: application to the genotyping of *Yersinia pestis* and *Bacillus anthracis*. *BMC Microbiol.* 1:2. <http://dx.doi.org/10.1186/1471-2180-1-2>.
  10. Bogdanovich T, Carniel E, Fukushima H, Skurnik M. 2003. Use of O-antigen gene cluster-specific PCRs for the identification and O-genotyping of *Yersinia pseudotuberculosis* and *Yersinia pestis*. *J. Clin. Microbiol.* 41:5103–5112.