

Complete Genome and Methylome Sequences of Two *Salmonella enterica* spp.

Kuan Yao,^{a,b} Tim Muruvanda,^a Richard J. Roberts,^c Justin Payne,^a Marc W. Allard,^a Maria Hoffmann^a

Center for Food Safety and Nutrition, Division of Microbiology, Office of Regulatory Science, U.S. Food and Drug Administration, College Park, Maryland, USA^a; School of Systems Biology, George Mason University, Manassas, Virginia, USA^b; New England Biolabs, Inc., Ipswich, Massachusetts, USA^c

***Salmonella enterica* is responsible for major foodborne outbreaks worldwide. It can cause gastroenteritis characterized by diarrhea, vomiting, and fever. *Salmonella* infections raise public health concerns along with consequential economic impacts. In this report, we announce the first complete genome sequences of *Salmonella enterica* subsp. *enterica* serovar Choleraeuis (*S. Choleraeuis*) ATCC 10708 and *Salmonella enterica* subsp. *enterica* serovar Pullorum (*S. Pullorum*) ATCC 9120, isolated from patients with diarrhea.**

Received 19 November 2015 Accepted 27 November 2015 Published 21 January 2016

Citation Yao K, Muruvanda T, Roberts RJ, Payne J, Allard MW, Hoffmann M. 2016. Complete genome and methylome sequences of two *Salmonella enterica* spp. *Genome Announc* 4(1):e01599-15. doi:10.1128/genomeA.01599-15.

Copyright © 2016 Yao 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/4.0/).

Address correspondence to Maria Hoffmann, maria.hoffmann@fda.hhs.gov.

Salmonella enterica subsp. *enterica* serovar Choleraeuis (*S. Choleraeuis*) is a nontyphoid serotype that is highly swine-adapted and can be transmitted to humans by contaminated pork products and slaughter environments (1). According to the U.S. Centers for Disease Control and Prevention, the number of *S. Choleraeuis* infections has been gradually decreasing in the United States (2). Still, *S. Choleraeuis* infections remain prevalent in Asia. For example, *S. Choleraeuis* accounted for 5% of the reported *Salmonella* infections in northern Taiwan (3), and since 2008 between 85 and 163 cases associated with *S. Choleraeuis* infection have been reported each year in Japan (1).

S. enterica subsp. *enterica* serovar Pullorum (*S. Pullorum*) infects mainly poultry and wild birds and can lead to pullorum disease (4, 5). The strain persists in the reproductive tract of chickens, causing eggs and chicks to obtain the pathogen from vertical transmission. *S. Pullorum* is next to *S. Gallinarum* in being responsible for economic losses in poultry production. Given the importance of these serovars, we announce here the availability of complete closed reference genomes for *S. Choleraeuis* ATCC 10708 and *S. Pullorum* ATCC 9120.

Both isolates, *S. Choleraeuis* and *S. Pullorum*, were cultured in Trypticase soy broth (Becton, Dickinson, Franklin Lakes, NJ, USA) overnight at 37°C. The genomic DNA was extracted from overnight cultures using the DNeasy blood and tissue kit (Qiagen, Inc., Valencia, CA, USA). The genomes were sequenced using the

Pacific Biosciences (PacBio) RS II sequencing platform, as previously reported (6, 7). The library was prepared based on the 20-kb PacBio sample preparation protocol and sequenced using P6/C4 chemistry on four single-molecule real-time (SMRT) cells with a 240-min collection time. The continuous long read (CLR) data were *de novo* assembled using the PacBio hierarchical genome assembly process (HGAP version 3.0) with default parameters (8). The assembled sequences were annotated using the NCBI Prokaryotic Genomes Annotation Pipeline and subsequently deposited at DDBJ/EMBL/GenBank.

The *S. Choleraeuis* genome was fully closed with 182× coverage for the chromosome and 51× coverage for the plasmid sequence. The chromosome consists of 4,824,322 bp with a GC content of 52.24%, while the plasmid comprises 119,113 bp with a GC content of 48%. PHAST (9) analysis identified four “intact” prophages, including *Enterobacteria* phage Fels-2, *Enterobacteria* phage P22, prophage Gifsy-2, and one unknown phage.

The *S. Pullorum* genome was fully closed with 355× coverage for the chromosome and 178× coverage for the plasmid. The genome size was 4,694,842 bp with a GC content of 52.19% and consisted of 4,630 genes. The plasmid was found to comprise 86,650 bp. PHAST analysis identified two “intact” prophages, including *Enterobacteria* phage Fels-2 and prophage Gifsy-2.

Sequencing of these isolates also measured the kinetic variations of nucleotide incorporation rates to infer DNA methyltrans-

TABLE 1 Summary of active methylases and their recognition sequences

Strain	CAGAG	ATGCAT	CCWGG	GATC
<i>S. Choleraeuis</i>	M.Sen10708I	M.Sen10708II	M.Sen10708Dcm	M.Sen10708DamP ^a M.Sen10708ORF5545P ^b
<i>S. Pullorum</i>			M.Sen9120Dcm	M.Sen9120DamP ^a M.Sen9120ORF17650P ^b

^a Probably active.

^b Probably inactive, as it is encoded on a prophage (*Enterobacteria* phage Fels-2, NC_010463).

ferase activities (10). The SMRT data of the methylomes were analyzed and are summarized in Table 1. They are also deposited in REBASE (11). DNA methyltransferase recognition motifs, which were detected by SMRT sequencing, and the genes encoding the various motifs can be found for *S. Choleraesuis* at <http://rebase.neb.com/cgi-bin/pacbioget?16669> and for *S. Pullorum* at <http://rebase.neb.com/cgi-bin/pacbioget?10067>.

Nucleotide sequence accession numbers. The complete genome and plasmid sequences of *S. Choleraesuis* are available in GenBank under the accession numbers CP012344 and CP012345. The complete chromosome and plasmid sequences of *S. Pullorum* were deposited under the accession numbers CP012347 and CP012348.

ACKNOWLEDGMENTS

This project was supported by the U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, Office of Regulatory Science and by the Small Business Innovation Research Program (NIGMS) of the National Institutes of Health under award number R44GM105125 to R.J.R.

R.J.R. works for New England Biolabs, a company that sells research reagents, including restriction enzymes and DNA methylases to the scientific community.

FUNDING INFORMATION

This project was supported by the U.S. FDA, Center for Food Safety and Applied Nutrition, Office of Regulatory Science, and by the Small Business Innovation Research Program (NIGMS) of the National Institutes of Health under award number R44GM105125 to R.J.R.

REFERENCES

1. Matayoshi M, Kitano T, Sasaki T, Nakamura M. 2015. Resistance phenotypes and genotypes among multiple-antimicrobial-resistant *Salmonella enterica* subspecies *enterica* serovar Choleraesuis strains isolated between 2008 and 2012 from slaughter pigs in Okinawa Prefecture, Japan. *J Vet Med Sci* 77:705–710. <http://dx.doi.org/10.1292/jvms.14-0683>.
2. Su L, Wu T, Chiu C. 2014. Decline of *Salmonella enterica* serotype choleraesuis infections, Taiwan. *Emerg Infect Dis* 20:715–716. <http://dx.doi.org/10.3201/eid2004.130240>.
3. Chiu C-H, Su L-H, Chu C. 2004. *Salmonella enterica* serotype choleraesuis: epidemiology, pathogenesis, clinical disease, and treatment. *Clin Microbiol Rev* 17:311–322. <http://dx.doi.org/10.1128/CMR.17.2.311-322.2004>.
4. Liu G-R, Rahn A, Liu W-Q, Sanderson KE, Johnston RN, Liu S-L. 2002. The evolving genome of *Salmonella enterica* serovar pullorum. *J Bacteriol* 184:2626–2633. <http://dx.doi.org/10.1128/JB.184.10.2626-2633.2002>.
5. Wigley P, Berchieri A, Jr, Page KL, Smith AL, Barrow PA. 2001. *Salmonella enterica* serovar Pullorum persists in splenic macrophages and in the reproductive tract during persistent, disease-free carriage in chickens. *Infect Immun* 69:7873–7879. <http://dx.doi.org/10.1128/IAI.69.12.7873-7879.2001>.
6. Hoffmann M, Muruvanda T, Allard MW, Korch J, Roberts RJ, Timme R, Payne J, McDermott PF, Evans P, Meng J, Brown EW, Zhao S. 2013. Complete genome sequence of a multidrug-resistant *Salmonella enterica* serovar Typhimurium var. 5– strain isolated from chicken breast. *Genome Announc* 1(6):e01068-13. <http://dx.doi.org/10.1128/genomeA.01068-13>.
7. Hoffmann M, Muruvanda T, Pirone C, Korch J, Timme R, Payne J, Evans P, Meng J, Brown EW, Allard MW. 2014. First fully closed genome sequence of *Salmonella enterica* subsp. *enterica* serovar Cubana associated with a foodborne outbreak. *Genome Announc* 2(5):e01112-14. <http://dx.doi.org/10.1128/genomeA.01112-14>.
8. Chin C, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <http://dx.doi.org/10.1038/nmeth.2474>.
9. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. *Nucleic Acids Res* 39:W347–W352. <http://dx.doi.org/10.1093/nar/gkr485>.
10. Korlach J, Turner SW. 2012. Going beyond five bases in DNA sequencing. *Curr Opin Struct Biol* 22:251–261. <http://dx.doi.org/10.1016/j.sbi.2012.04.002>.
11. Roberts RJ, Vincze T, Posfai J, Macelis D. 2015. REBASE-A database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Research* 43:D298–D299. <http://dx.doi.org/10.1093/nar/gku1046>.