

UC Irvine

UC Irvine Previously Published Works

Title

Draft Genome Sequence of Salmonella enterica subsp. enterica Serovar Lille Strain CRJJGF_000101 (Phylum Gammaproteobacteria)

Permalink

<https://escholarship.org/uc/item/93q2k5js>

Journal

Microbiology Resource Announcements, 4(4)

ISSN

2576-098X

Authors

Gupta, Sushim K
McMillan, Elizabeth A
Jackson, Charlene R
et al.

Publication Date

2016-08-25

DOI

10.1128/genomea.00603-16

Peer reviewed

Draft Genome Sequence of *Salmonella enterica* subsp. *enterica* Serovar Lille Strain CRJJGF_000101 (Phylum *Gammaproteobacteria*)

Sushim K. Gupta,^a Elizabeth A. McMillan,^{a,b} Charlene R. Jackson,^a Prerak T. Desai,^c Steffen Porwollik,^c Michael McClelland,^c Lari M. Hiott,^a Shaheen B. Humayoun,^a Jonathan G. Frye^a

Bacterial Epidemiology and Antimicrobial Resistance Research Unit, U.S. National Poultry Research Center, U.S. Department of Agriculture, Agricultural Research Service, Athens, Georgia, USA^a; Department of Microbiology, University of Georgia, Athens, Georgia, USA^b; Department of Microbiology and Molecular Genetics, University of California Irvine, Irvine, California, USA^c

Here, we report a 4.98 Mbp draft genome sequence of *Salmonella enterica* subsp. *enterica* serovar Lille strain CRJJGF_000101, isolated from ground beef in 2007.

Received 6 May 2016 Accepted 23 May 2016 Published 14 July 2016

Citation Gupta SK, McMillan EA, Jackson CR, Desai PT, Porwollik S, McClelland M, Hiott LM, Humayoun SB, Frye JG. 2016. Draft genome sequence of *Salmonella enterica* subsp. *enterica* serovar Lille strain CRJJGF_000101 (phylum *Gammaproteobacteria*). *Genome Announc* 4(4):e00603-16. doi:10.1128/genomeA.00603-16.

Copyright © 2016 Gupta 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 Jonathan G. Frye, Jonathan.Frye@ars.usda.gov.

Salmonellosis is a major public-health concern worldwide. Consumption of raw or improperly prepared or contaminated foods, mostly of animal origin, is considered the major route of human salmonellosis (1). There are more than 2,500 serovars of *Salmonella*, exhibiting more than 46 different O groups and 114 different H groups; all of them are considered a potential threat to human health (2). The *Salmonella* strain CRJJGF_000101 was isolated from ground beef in 2007 using standard microbiology techniques and serotyped using SMART (3). The isolate was serogrouped using serogroup-specific antisera (Difco Laboratories, Detroit, MI) and the serovar was determined at the National Veterinary Services Laboratories, APHIS, USDA (Ames, IA). This bacterium belonged to antigenic group O:7(C1), along with *Salmonella* serovar Rumford, and contained somatic O antigen 7 and only phase1 flagellar H antigens z38 (7:z38:-) (4). Using pulsed-field gel electrophoresis (PFGE) as described by PulseNet (5), the isolate was assigned PFGE pattern LPPX01.0020. MICs ($\mu\text{g/ml}$) were determined by broth microdilution using the Sensititre semi-automated antimicrobial susceptibility system (TREK Diagnostics Systems, Thermo Fisher Scientific, Inc., Oakwood Village, OH). Results were interpreted according to Clinical and Laboratory Standards Institute (CLSI) guidelines (6).

The genomic DNA was isolated using a GenElute bacterial genomic DNA kit (Sigma-Aldrich, St. Louis, MO) and the DNA library was constructed using a Nextera-XT DNA preparation kit and paired-end sequencing was performed on an Illumina HiSeq2500 (Illumina Inc., San Diego, CA) using a 500-cycle MiSeq reagent kit. About 4,304,692 reads with quality score >30 were assembled using Velvet assembler (7), which resulted in 196 contigs with minimum contig length ≥ 200 bp. The total assembly size was 4.98 Mbp, with N_{50} values of 70.7 kbp, and G+C content of 51.98%. The contigs were ordered with MAUVE using *Salmonella* LT2 as a reference (8), and prodigal (9) and ARAGORN (10) were used to predict coding sequences (CDS) and tRNAs. A total of 4,764 coding sequences (≥ 50 amino acids) and 48 tRNAs were predicted within the genome. Prophages, signal peptides, and re-

sistance genes were predicted using PFAST (11), signalp (12), and ARG-ANNOT (13), respectively. We identified signal peptides in 462 genes, two clustered regularly interspaced short palindromic repeat (CRISPR) (14) loci, and 5-intact/6-incomplete phages in the genome. We detected an *aac6-Iaa* resistance gene, which remains cryptic and a tetracycline resistance gene *tetC*, which was confirmed phenotypically. The detected arsenic resistance genes (*pstBACS*) (15) correlated with the phenotypes. The strain's MIC was 411 and 58 $\mu\text{g/ml}$ compared to a wild type median MIC, of 51 and 15 $\mu\text{g/ml}$ for arsenate and arsenite, respectively. The pathogenesis analysis of *Salmonella* serovar Lille in the mouse model categorized the strain in the low invasive group (16). However, *Salmonella* serovar Lille has been reported during salmonellosis outbreaks (17). Host specific invasion may be the reason for low invasiveness in mice. The information generated from the genome may enhance our understanding of the variations in the virulence genes for host specific infection.

Nucleotide sequence accession number. The genome sequence of *Salmonella enterica* subsp. *enterica* serovar Lille strain CRJJGF_000101 has been deposited in GenBank under the accession number **JQWF00000000**. This paper describes the first version of the genome.

ACKNOWLEDGMENTS

We thank Calvin Williams for all the technical support.

J.G.F. and C.R.J. were supported by USDA projects 6040-32000-006-00 and 6040-32000-009-00, and a grant from the Foundation for Meat and Poultry Research and Education. M.M. was supported in part by NIH grants R01AI052237, AI039557 AI052237, AI073971, AI075093, AI077645 AI083646, USDA grants 2009-03579 and 2011-67017-30127, the Binational Agricultural Research and Development Fund, and a grant from the Center for Produce Safety.

The mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

FUNDING INFORMATION

This work, including the efforts of Charlene R. Jackson and Jonathan Gray Frye, was funded by Foundation for Meat and Poultry Research and Education. This work, including the efforts of Michael McClelland, was funded by HHS | National Institutes of Health (NIH) (R01AI052237, AI039557, AI052237, AI073971, AI075093, AI077645, and AI083646). This work, including the efforts of Charlene R. Jackson and Jonathan Gray Frye, was funded by U.S. Department of Agriculture (USDA) (6040-32000-006-00 and 6040-32000-009-00). This work, including the efforts of Michael McClelland, was funded by U.S. Department of Agriculture (USDA) (2009-03579 and 2011-67017-30127). This work, including the efforts of Michael McClelland, was funded by United States - Israel Binational Agricultural Research and Development Fund (BARD). This work, including the efforts of Michael McClelland, was funded by Center for Produce Safety (CPS).

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

1. D'Aoust JY. 1991. Pathogenicity of foodborne *Salmonella*. *Int J Food Microbiol* 12:17–40. [http://dx.doi.org/10.1016/0168-1605\(91\)90045-Q](http://dx.doi.org/10.1016/0168-1605(91)90045-Q).
2. Malorny B, Hauser E, Dieckmann R. 2011. New approaches in subspecies-level *Salmonella* classification, p. 1–23. In Porwollik S (ed). *Salmonella from genome to function*. Caister Academic Press, Norfolk, United Kingdom.
3. Leader BT, Frye JG, Hu J, Fedorka-Cray PJ, Boyle DS. 2009. High-throughput molecular determination of *Salmonella enterica* serovars by use of multiplex PCR and capillary electrophoresis analysis. *J Clin Microbiol* 47:1290–1299. <http://dx.doi.org/10.1128/JCM.02095-08>.
4. Grimont P, Weill F. 2007. Antigenic formulae of the *Salmonella* serovars. 9th ed. WHO Collaborating Centre for Reference and Research on *Salmonella*, Geneva, Switzerland. https://www.pasteur.fr/sites/www.pasteur.fr/files/wklm_en.pdf.
5. Gerner-Smidt P, Hise K, Kincaid J, Hunter S, Rolando S, Hyytiä-Trees E, Ribot EM, Swaminathan B, Pulsenet Taskforce. 2006. PulseNet USA: a five-year update. *Foodborne Pathog Dis* 3:9–19. <http://dx.doi.org/10.1089/fpd.2006.3.9>.
6. Clinical and Laboratory Standards Institute. 2015. Performance standards for antimicrobial susceptibility testing: 25th informational supplement (m100–S25). Clinical and Laboratory Standards Institute, Wayne, PA.
7. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <http://dx.doi.org/10.1101/gr.074492.107>.
8. Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. 2009. Reordering contigs of draft genomes using the Mauve aligner. *Bioinformatics* 25:2071–2073. <http://dx.doi.org/10.1093/bioinformatics/btp356>.
9. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <http://dx.doi.org/10.1186/1471-2105-11-119>.
10. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 32:11–16. <http://dx.doi.org/10.1093/nar/gkh152>.
11. 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>.
12. Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8:785–786. <http://dx.doi.org/10.1038/nmeth.1701>.
13. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, Rolain JM. 2014. Arg-Annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother* 58:212–220. <http://dx.doi.org/10.1128/AAC.01310-13>.
14. Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 35:W52–W57. <http://dx.doi.org/10.1093/nar/gkm360>.
15. Pal C, Bengtsson-Palme J, Rensing C, Kristiansson E, Larsson DG. 2014. BacMet: antibacterial biocide and metal resistance genes database. *Nucleic Acids Res* 42:D737–D743. <http://dx.doi.org/10.1093/nar/gkt1252>.
16. McWhorter AR, Davos D, Chousalkar KK. 2015. Pathogenicity of *Salmonella* strains isolated from egg shells and the layer farm environment in Australia. *Appl Environ Microbiol* 81:405–414. <http://dx.doi.org/10.1128/AEM.02931-14>.
17. Lehmacher A, Bockemühl J, Aleksic S. 1995. Nationwide outbreak of human salmonellosis in Germany due to contaminated paprika and paprika-powdered potato chips. *Epidemiol Infect* 115:501–511. <http://dx.doi.org/10.1017/S0950268800058660>.