

Draft Genome Sequences of 21 *Salmonella enterica* Serovar Enteritidis Strains

Ruth E. Timme,^a Marc W. Allard,^a Yan Luo,^b Errol Strain,^b James Pettengill,^a Charles Wang,^a Cong Li,^a Christine E. Keys,^a Jie Zheng,^a Robert Stones,^c Mark R. Wilson,^d Steven M. Musser,^a and Eric W. Brown^a

Office of Regulatory Science^a and Office of Food Defense, Communications, and Emergency Response,^b Center for Food Safety & Applied Nutrition, U.S. Food & Drug Administration, College Park, Maryland, USA; Food & Environment Research Agency, Sand Hutton, York, United Kingdom^c; and Forensic Science Program, Western Carolina University, Cullowhee, North Carolina, USA^d

***Salmonella enterica* subsp. *enterica* serovar Enteritidis is a common food-borne pathogen, often associated with shell eggs and poultry. Here, we report draft genomes of 21 *S. Enteritidis* strains associated with or related to the U.S.-wide 2010 shell egg recall. Eleven of these genomes were from environmental isolates associated with the egg outbreak, and 10 were reference isolates from previous years, unrelated to the outbreak. The whole-genome sequence data for these 21 human pathogen strains are being released in conjunction with the newly formed 100K Genome Project.**

Salmonella enterica is one of the primary causes of food-borne illness in the United States, leading to more deaths than any other food-related pathogen (4). This bacterial species is extremely diverse, comprising over 2,500 serovars (1), one of which, *Salmonella enterica* subsp. *enterica* serovar Enteritidis, has had an international increase in infection rates over the past 20 years (3). *S. Enteritidis*, most commonly associated with eggs and poultry, caused the largest shell egg recall in U.S. history (2010). Because of its highly clonal nature during outbreaks, traditional pulsed-field gel electrophoresis (PFGE) and phage typing have not been useful subtyping tools for this serovar (5); however, whole-genome sequencing overcomes this barrier by providing the discriminatory power needed for differentiating highly clonal strains (2).

Currently there is only one complete *S. Enteritidis* genome

available in GenBank (*S. Enteritidis* strain P125109) and one draft genome (*S. Enteritidis* strain LA5). We announce the availability of 21 new high-quality draft *Salmonella enterica* subsp. *enterica* serovar Enteritidis genomes, 11 associated with the shell egg outbreak of 2010 and 10 related reference strains. The PFGE patterns determined were JEGX01.0004 for strains 622731-39, 639016-6,

Received 20 July 2012 Accepted 22 August 2012

Address correspondence to Ruth E. Timme, ruth.timme@fda.hhs.gov.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JB.01289-12

TABLE 1 DDBJ/EMBL/GenBank accession numbers, average sequence coverage, and contig numbers for *S. enterica* subsp. *enterica* serovar Enteritidis strains^a

| S. Enteritidis strain | Accession no(s). | | WGS | Coverage (fold) | No. of contigs |
|-----------------------|------------------|----------------------|------------------------------|-----------------|----------------|
| | BioProject | SRA | | | |
| 622731-39 | 52615 | SRR518786 | ALEI00000000 | 22 | 60 |
| 639016-6 | 52617 | SRR518813 | ALEJ00000000 | 23 | 62 |
| 640631 | 52619 | SRR518800 | ALEK00000000 | 20 | 51 |
| 77-0424 | 53259 | SRR518840 | ALEL00000000 | 21 | 51 |
| 607307-6 | 53263 | SRR518859 | ALEM00000000 | 23 | 61 |
| 485549-17 | 59531 | SRR518788 | ALEN00000000 | 20 | 50 |
| 596866-22 | 59533 | SRR518816, SRR518817 | ALEO00000000 | 27 | 51 |
| 596866-70 | 59535 | SRR518755 | ALEP00000000 | 22 | 47 |
| 629164-26 | 59537 | SRR518756 | ALEQ00000000 | 20 | 41 |
| 629164-37 | 59539 | SRR518757 | ALER00000000 | 22 | 56 |
| 639672-46 | 59541 | SRR518770 | ALES00000000 | 22 | 61 |
| 639672-50 | 59543 | SRR518818 | ALET00000000 | 16 | 50 |
| 77-1427 | 60069 | SRR518841 | ALEU00000000 | 18 | 54 |
| 77-2659 | 60071 | SRR518843 | ALEV00000000 | 17 | 54 |
| 78-1757 | 60073 | SRR518811 | ALEW00000000 | 38 | 49 |
| 22510-1 | 60075 | SRR518784 | ALEX00000000 | 19 | 49 |
| 8b-1 | 60511 | SRR518767 | ALEY00000000 | 20 | 53 |
| 648905 5-18 | 62825 | SRR518823 | ALEZ00000000 | 19 | 61 |
| 648901 6-18 | 62829 | SRR518763 | ALFA00000000 | 19 | 63 |
| 50-3079 | 73685 | SRR518824 | ALFB00000000 | 17 | 51 |
| 58-6482 | 77695 | SRR518825 | ALFC00000000 | 18 | 55 |

^a SRA, NCBI short-read archive; WGS, NCBI whole-genome shotgun assembly database.

640631, 607307-6, 485549-17, 596866-22, 596866-70, 629164-37, 639672-46, 639672-50, 648905 5-18, and 648901 6-18 and JEGX01.0034 for strain 629164-26.

DNA was isolated from a pure culture of each strain using a Qiagen DNeasy blood and tissue kit (Qiagen Inc., Valencia, CA). Genome sequencing was performed using 454 Titanium sequencing technology (Roche, Branford, CT), achieving 15 to 20× average genome coverage. *De novo* assemblies were created for each genome using the 454 Life Sciences Newbler software package, v.2.6 (Roche), and annotated with the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (<http://www.ncbi.nlm.nih.gov/genomes/static/Pipeline.html>). Contig numbers ranged from 41 to 63 (Table 1). An in-depth, comparative genomic analysis of these data will be provided in a future publication.

This large data release contributes toward the efforts of the 100K Genome Project consortium. The U.S. Food and Drug Administration (FDA), Agilent, and University of California, Davis, along with many other federal and private partners, will sequence 100,000 pathogen genomes over the next 5 years (<http://100kgenome.vetmed.ucdavis.edu>). The product of this enormous effort will be a public molecular epidemiology reference database useful for designing pathogen detection assays, providing evolutionary context for emerging global outbreaks, and many other applications yet to be realized. The public database will be

housed at the National Center for Biotechnology Information (NCBI) in Bethesda, MD.

Nucleotide sequence accession numbers. The draft genome sequences of the 21 *Salmonella* Enteritidis strains are available in DDBJ/EMBL/GenBank under the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

We thank the NCBI rapid annotation pipeline team and our FDA partners in the Center for Veterinary Medicine, Division of Field Sciences, and regional field laboratories for providing isolates, namely, Shaohua Zhao, Rebecca Dreisch, Peggy Carter, Norma Duran, and Palmer Orlandi.

This work was supported by the Center for Food Safety and Applied Nutrition at the U.S. Food and Drug Administration.

REFERENCES

1. Grimont PAD, Weill F-X. 2007. Antigenic formulae of the *Salmonella* serovars, 9th ed. WHO Collaborating Centre for Reference and Research on *Salmonella*, Paris, France.
2. Lienau EK, et al. 2011. Identification of a salmonellosis outbreak by means of molecular sequencing. *N. Engl. J. Med.* 364:981–982.
3. Rodrigue DC, Tauxe RV, Rowe B. 1990. International increase in *Salmonella* enteritidis: a new pandemic? *Epidemiol. Infect.* 105:21–27.
4. Scallan E, et al. 2011. Foodborne illness acquired in the United States—major pathogens. *Emerg. Infect. Dis.* 17:7–15.
5. Zheng J, Keys CE, Zhao S, Meng J, Brown EW. 2007. Enhanced subtyping scheme for *Salmonella* Enteritidis. *Emerg. Infect. Dis.* 13:1932.