AMERICAN SOCIETY FOR MICROBIOLOGY

Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequences of 57 *Salmonella enterica* Strains from Selected U.S. Swine Feed Mills

Sara Lomonaco,[a] Gabriela Magossi,[b] Maria Sanchez Leon,[a] Daniela Miller,[a] George J. Kastanis,[a] Sandra M. Tallent,[a] Marc W. Allard,[a] Eric W. Brown,[a] Valentina Trinetta[b]

[a]Center for Food Safety and Applied Nutrition, U.S. Food and Drug Administration, College Park, Maryland, USA

[b]Food Science Institute, Kansas State University, Manhattan, Kansas, USA

ABSTRACT The number of *Salmonella* infection cases linked to pork products has increased. Pathogen presence in the feed mill environment is one of the many potential transmission routes into the food production chain. Here, we describe the draft genome sequences of 57 *Salmonella enterica* isolates from selected U.S. swine feed mills.

Salmonella enterica, one of the most important pathogens in the food industry, is the leading cause of bacterial foodborne disease in the United States (1). While the impact on human health alone is enormous, there are also substantial economic repercussions of *Salmonella* outbreaks across the food industry. Therefore, it is important to have molecular typing data for well-characterized isolates from various sources. Whole-genome sequencing (WGS) has been increasingly used to characterize food and clinical *Salmonella* isolates (2).

Here, we describe the draft genome sequences of 57 *S. enterica* isolates collected during 2016 and 2017 from 11 U.S. swine feed mills (from 8 U.S. states) producing mash feed and/or pelleted feed (3). Within each mill, 12 sites were sampled, as follows: floor surfaces (n = 5), equipment (n = 4), brooms (n = 1), worker shoes (n = 1), and finished feed (n = 1). Samples were analyzed following the USDA Food Safety and Inspection Service (USDA-FSIS) guidelines (4), and identification was confirmed by PCR (5). Overnight tryptone soy broth cultures of positive isolates were subjected to DNA extraction using the DNeasy blood and tissue kit (Qiagen, Hilden, Germany). Paired-end DNA libraries were prepared with the Nextera XT DNA library preparation kit, and WGS was carried out on either a MiSeq or NextSeq sequencer, using a 500-cycle MiSeq reagent V2 kit or a 300-cycle NextSeq 500/550 high-output V2 kit, respectively (Illumina, San Diego, CA). Unless otherwise noted, default parameters were used in all analyses. *De novo* genome assemblies were obtained with the Shovill pipeline version 0.9 (https://github.com/tseemann/shovill), available in the GalaxyTrakr pipeline (https://www.galaxytrakr.org; see reference 6). The "trim reads" option was selected, and 500 bp was set as the minimum contig length. ResFinder and SPIFinder from the Center for Genomic Epidemiology were used to identify antibiotic resistance (AMR) genes and *Salmonella* pathogenicity islands (SPIs) in the draft genomes, respectively (https://cge.cbs.dtu.dk/services/). Draft genomes were annotated using the NCBI's Prokaryotic Genome Annotation Pipeline (7). *Salmonella* serotypes were predicted from draft genomes with SeqSero 1.0 (http://www.denglab.info/SeqSero; see reference 8). Draft genomes ranged between 4,484,198 and 5,116,995 Mb in size, with 52% average GC content. The number of contigs for each isolate ranged from 29 to 86. The isolates belonged to 15 different serotypes, namely, *Salmonella enterica* serotypes Agona (n = 14), Mbandaka (n = 13), Senftenberg (n = 7), Schwarzengrund (n = 6), Rissen (n = 3), and Hartford and Typhimurium (n = 2 each). Nine serotypes were observed in only one isolate each,

Received 28 August 2018 **Accepted** 28 September 2018 **Published** 1 November 2018

Citation Lomonaco S, Magossi G, Sanchez Leon M, Miller D, Kastanis GJ, Tallent SM, Allard MW, Brown EW, Trinetta V. 2018. Draft genome sequences of 57 *Salmonella enterica* strains from selected U.S. swine feed mills. Microbiol Resour Announc 7:e01191-18. https://doi.org/10.1128/MRA.01191-18.

Editor David A. Baltrus, University of Arizona

Address correspondence to Sara Lomonaco, sara.lomonaco@fda.hhs.gov.

TABLE 1 Metadata for the 57 *Salmonella enterica* strains isolated from different feed mills in the United States between 2016 and 2017

Sample identifier	BioSample no.	Run no.	Accession no.	No. of contigs	Total length (bp)	N_{50} (bp)	GC content (%)	Genome coverage (×)
CFSAN071932	SAMN08057838	SRR6959189	QUTN00000000	52	5,012,146	260,770	51.8	71
CFSAN071933	SAMN08057839	SRR6959212	QUTO00000000	49	5,010,824	235,309	51.8	91
CFSAN071934	SAMN08057843	SRR6957504	QUTP00000000	49	5,010,127	260,945	51.8	89
CFSAN071935	SAMN08057815	SRR6957525	QUTQ00000000	86	5,116,995	271,078	52.1	135
CFSAN071936	SAMN08057823	SRR6959053	QUTR00000000	56	4,974,144	235,297	51.8	123
CFSAN071937	SAMN08057820	SRR6959057	QUTS00000000	39	4,780,443	292,436	52.1	148
CFSAN071938	SAMN08057858	SRR6959072	QUTT00000000	43	4,782,578	268,283	52.1	133
CFSAN071940	SAMN08057866	SRR6959115	QUTU00000000	41	4,484,295	381,546	52.2	112
CFSAN071941	SAMN08057869	SRR6958537	QUTV00000000	39	4,484,198	381,546	52.2	139
CFSAN071943	SAMN08057863	SRR7545387	QUTW00000000	36	4,943,381	341,411	51.8	135
CFSAN071944	SAMN08057865	SRR6957683	QUTX00000000	56	4,797,481	381,076	52.2	139
CFSAN071945	SAMN08057879	SRR6959319	QUTY00000000	61	4,689,856	236,673	52.1	127
CFSAN071946	SAMN08057880	SRR6959315	QUTZ00000000	45	4,973,164	389,983	52	139
CFSAN071947	SAMN08057876	SRR6959018	QUUA00000000	39	4,974,888	390,021	52	154
CFSAN071948	SAMN08057877	SRR6959052	QUUB00000000	43	4,672,093	235,297	52.2	74
CFSAN071949	SAMN08057888	SRR6958605	QUUC00000000	40	4,974,485	412,899	52	140
CFSAN071950	SAMN08057887	SRR6958972	QUUD00000000	50	4,971,687	317,512	52	127
CFSAN071951	SAMN08057883	SRR6957691	QUUE00000000	39	4,978,782	412,745	52	51
CFSAN071952	SAMN08057882	SRR6959190	QUUF00000000	42	4,830,818	401,049	52.2	186
CFSAN071953	SAMN08057886	SRR6959322	QUUG00000000	53	4,689,837	278,573	52.1	164
CFSAN071954	SAMN08057885	SRR6959202	QUUH00000000	29	4,780,916	408,171	52	164
CFSAN071955	SAMN08057836	SRR6959274	QUUI00000000	33	4,707,452	389,545	52.1	125
CFSAN071956	SAMN08057825	SRR6958149	QUUJ00000000	58	4,696,591	245,410	52.1	160
CFSAN071957	SAMN08057821	SRR6958145	QUUK00000000	60	4,696,272	259,971	52.1	181
CFSAN071958	SAMN08057819	SRR6958473	QUUL00000000	57	4,690,445	259,968	52.1	182
CFSAN071959	SAMN08057816	SRR6958470	QUUM00000000	59	4,690,960	235,297	52.1	187
CFSAN071960	SAMN08057868	SRR6958115	QUUN00000000	45	4,813,521	322,240	52.2	187
CFSAN071961	SAMN08057881	SRR6956165	QUUO00000000	39	4,781,312	440,059	52.1	177
CFSAN071964	SAMN08057864	SRR6957524	QUUP00000000	38	4,721,388	363,871	52.1	127
CFSAN071965	SAMN08057889	SRR6957526	QUUQ00000000	74	4,978,222	214,013	52.1	133
CFSAN071966	SAMN08057809	SRR6957442	QUUR00000000	78	4,980,304	214,013	52.1	134
CFSAN071967	SAMN08057806	SRR6956728	QUUS00000000	67	4,936,335	146,406	52.1	117
CFSAN071968	SAMN08057804	SRR6956712	QUUT00000000	35	4,708,769	437,088	52.3	148
CFSAN071969	SAMN08057824	SRR6956669	QUUU00000000	46	4,780,193	267,610	52.1	144
CFSAN071970	SAMN08057822	SRR6651925	QUUV00000000	32	4,829,698	447,189	52.1	175
CFSAN071971	SAMN08057833	SRR6653453	QUUW00000000	34	4,827,673	447,721	52.1	182
CFSAN071972	SAMN08057832	SRR6653448	QUUX00000000	41	4,760,816	280,947	52	156
CFSAN071973	SAMN08057831	SRR6651923	QUUY00000000	55	5,010,453	350,213	51.8	143
CFSAN071974	SAMN08057842	SRR6653464	QUUZ00000000	54	4,994,979	307,967	51.8	146
CFSAN071975	SAMN08057848	SRR6653465	QUVA00000000	51	4,994,206	307,967	51.8	125
CFSAN071976	SAMN08057856	SRR6654229	QUVB00000000	48	4,994,668	360,164	51.8	194
CFSAN071977	SAMN08057857	SRR6654195	QUVC00000000	47	4,995,733	329,065	51.8	212
CFSAN071978	SAMN08057853	SRR6654125	QUVD00000000	55	5,008,977	350,306	51.8	189
CFSAN071979	SAMN08057849	SRR6654123	QUVE00000000	50	4,996,561	307,967	51.8	181
CFSAN071980	SAMN08057850	SRR6651152	QUVF00000000	50	5,007,268	307,967	51.8	231
CFSAN071981	SAMN08057847	SRR6651155	QUVG00000000	56	5,009,733	307,967	51.8	208
CFSAN071982	SAMN08057855	SRR6651156	QUVH00000000	49	4,994,374	360,164	51.8	228
CFSAN071983	SAMN08057854	SRR6650949	QUVI00000000	50	4,995,973	307,793	51.8	227
CFSAN071984	SAMN08057851	SRR6650951	QUVJ00000000	54	5,008,558	350,213	51.8	143
CFSAN071985	SAMN08057871	SRR6651157	QUVK00000000	52	5,010,076	328,176	51.8	164
CFSAN071986	SAMN08057870	SRR6650950	QUVL00000000	55	5,008,380	307,967	51.8	129
CFSAN071990	SAMN08057875	SRR6958062	QUVM00000000	41	4,745,986	454,049	52.2	110
CFSAN071991	SAMN08057874	SRR6957337	QUVN00000000	51	4,835,035	246,700	52.1	48
CFSAN071992	SAMN08057873	SRR7771352	QUVO00000000	47	4,843,802	262,144	52.1	88
CFSAN071993	SAMN08057805	SRR6958582	QUVP00000000	37	4,794,792	281,696	52	74
CFSAN071994	SAMN08057807	SRR6957350	QUVQ00000000	39	4,794,849	276,964	52	128
CFSAN074091	SAMN09695778	SRR7545388	QUZX00000000	41	4,974,987	412,331	52	322

Salmonella enterica serotypes Bareilly, Braenderup, Cubana, Javiana, Kiambu, Poona, Soerenga, Worthington, and the potential monophasic variant of *S. enterica* serotype Typhimurium (I 4,[5],12:i:−). AMR genes were detected in 30 out of 57 strains, including those encoding tetracycline, phenicol, aminoglycoside, and beta-lactam resistance. Up to 12 SPIs were detected in these genomes, including SPI-1, SPI-2 and SPI-4, which encode predicted type I and type III secretion systems (9). The described draft genome sequences will be useful in comparative genomic analyses of *S. enterica* from the pork production chain in terms of phylogenetic insights into their evolution and support of further epidemiological investigations of outbreaks.

Data availability. This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under accession numbers QUTN00000000 to QUVQ00000000 and QUZX00000000 (Table 1). The first versions (QUTN01000000 to QUVQ01000000 and QUZX01000000) are described here.

ACKNOWLEDGMENTS

This work was supported by the Center for Food Safety and Applied Nutrition (CFSAN) at the U.S. Food and Drug Administration (FDA). This project was also supported in part by an appointment to the Research Participation Program at the CFSAN, U.S. FDA, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the FDA.

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the U.S. FDA. The use of trade names and commercial sources is for identification purposes only and does not imply endorsement.

REFERENCES

1. Centers for Disease Control and Prevention. 2011. National enteric disease surveillance: *Salmonella* surveillance overview. Centers for Disease Control and Prevention, Atlanta, GA. https://www.cdc.gov/nationalsurveillance/salmonella-surveillance.html.
2. Allard MW, Strain E, Melka D, Bunning K, Musser SM, Brown EW, Timme R. 2016. Practical value of food pathogen traceability through building a whole-genome sequencing network and database. J Clin Microbiol 54:1975–1983. https://doi.org/10.1128/JCM.00081-16.
3. Magossi G, Cernicchiaro N, Dritz SS, Houser T, Woodworth JC, Jones CK, Trinetta V. 2018. Preliminary evaluation of *Salmonella* presence in selected United States feed mills. Microbiol Open https://doi.org/10.1002/mbo3.711.
4. US Department of Agriculture Food Safety and Inspection Service. 2014. Isolation and identification of *Salmonella* from meat, poultry, pasteurized egg, and catfish products and carcass and environmental sponges, 4.09. *In* Laboratory guidebook. US Department of Agriculture Food Safety and Inspection Service, Washington, DC.
5. Bai J, Trinetta V, Shi X, Noll LW, Magossi G, Zheng W, Porter EP, Cernicchiaro N, Renter DG, Nagaraja TG. 2018. A multiplex real-time PCR assay, based on *invA* and *pagC* genes, for the detection and quantification of *Salmonella enterica* from cattle lymph nodes. J Microbiol Methods 148:110–116. https://doi.org/10.1016/j.mimet.2018.03.019.
6. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Grüning BA, Guerler A, Hillman-Jackson J, Hiltemann S, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res 46:W537–W544. https://doi.org/10.1093/nar/gky379.
7. 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.
8. Zhang S, Yin Y, Jones MB, Zhang Z, Deatherage Kaiser BL, Dinsmore BA, Fitzgerald C, Fields PI, Deng X. 2015. *Salmonella* serotype determination utilizing high-throughput genome sequencing data. J Clin Microbiol 53:1685–1692. https://doi.org/10.1128/JCM.00323-15.
9. Amavisit P, Lightfoot D, Browning GF, Markham PF. 2003. Variation between pathogenic serovars within *Salmonella* pathogenicity islands. J Bacteriol 185:3624–3635. https://doi.org/10.1128/JB.185.12.3624-3635.2003.