

DATA NOTE

Open Access



217 closed *Salmonella* reference genomes using PacBio sequencing

Yan Luo^{1*}, Jae Hee Jang², Maria Balkey¹ and Maria Hoffmann¹

Abstract

Objectives Whole Genome Sequencing (WGS) is widely used in food safety for the detection, investigation, and control of foodborne bacterial pathogens. However, the WGS data in most public databases, such as the National Center for Biotechnology Information (NCBI), primarily consist of Illumina short reads which lack some important information for repetitive regions, structural variations, and mobile genetic elements, and the genomic location of certain important genes like antimicrobial resistance genes (AMR) and virulence genes. To address this limitation, we have contributed 217 closed circular *Salmonella enterica* genomes that were generated using PacBio sequencing to the NCBI Pathogen Detection (PD) database and GenBank. This dataset provides a higher level of accuracy to genome representations in the database.

Data description High-quality complete reference genomes generated from PacBio long reads can provide essential details that are not available in draft genomes from short reads. A complete reference genome allows for more accurate data analysis and researchers to establish connections between genome variations and known genes, regulatory elements, and other genomic features. The addition of 217 complete genomes from 78 different *Salmonella* serovars, each representing either a distinct SNP cluster within the NCBI PD database or a unique strain, significantly enriches the diversity of the reference genome database.

Keywords Pacbio sequencing, Closed genomes, Diverse *Salmonella* serovars

Objective

In 2012, the U.S. Food and Drug Administration's Center for Food Safety and Applied Nutrition (FDA-CFSAN) launched the GenomeTrakr network [1], the first distributed network of laboratories that utilize WGS for foodborne pathogen identification. GenomeTrakr data is submitted to the NCBI PD database to assist in foodborne pathogen surveillance and

outbreak detection. To date, WGS data of over 1.7 million isolates belonging to 81 different pathogen species, primarily obtained from Illumina short-read sequencers, has been collected, stored, and processed in NCBI. The NCBI PD web portal (<https://www.ncbi.nlm.nih.gov/pathogens/>) provides public access to this extensive collection of WGS data. This portal offers high resolution strain typing, outbreak investigation, and surveillance capabilities. However, the NCBI PD database currently contains only around 500 closed genomes indicating a shortage of high-quality complete genomes.

Salmonella, one of the leading foodborne pathogens, causes widespread illnesses and poses a significant threat to public health worldwide. To address the

*Correspondence:

Yan Luo

Yan.Luo@fda.hhs.gov

¹U.S. Food and Drug Administration, Human Foods Program, College Park, MD, USA

²U.S. Department of Agriculture, Agriculture Research Service, Beltsville, MD, USA



This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

need for more comprehensive genomic information, between 2018 and 2021 we focused on sequencing and completely closing a diverse set of *Salmonella* genomes using PacBio technology. This effort provides higher resolution, accuracy, and precision to various analyses, including the detection of structural variations, gene annotations, phylogenetic analysis, and comparative genomics. These analyses which are often used during foodborne events require critical information that is often lacking in the short-read assemblies. The 217 *Salmonella* isolates used in this collection were collected from various food sources, clinical, and environment. Both the closed genomes and raw long reads have been deposited in NCBI (Table 1), contributing to the expansion of the available genomic resources for *Salmonella* surveillance.

Data description

NCBI PD processes the WGS data from bacterial and fungal pathogen genomes and places them into clusters based on the relatedness between genomes. Specifically for *Salmonella*, approximately 711,000 genomes were assigned to more than 30,000 clusters as of January 20, 2025. Within each cluster, a reference genome was used to construct a reference-based SNP matrix, which in turn was used to infer a phylogenetic tree. Having a high-quality closed reference genome would greatly help to identify lineage or serovar-specific *Salmonella* fragments, enhance the accuracy of the phylogenetic analysis, SNP annotation, and other related analyses [2, 3]. This enhanced resolution helps trace the source of the outbreaks more accurately, determine the pathways of contamination, and understand the genetic factors involved in virulence and resistance [4–6]. Consequently, it will lead to more

effective and targeted interventions to control and prevent future outbreaks, ensuring better food safety and public health protection [7]. For our study we carefully selected 217 *Salmonella* isolates for PacBio sequencing. Each isolate represents either a distinct NCBI SNP cluster or is a unique isolate that did not cluster with any other genomes.

Each isolate was cultured in Luria-Bertani (LB) medium at 35 °C overnight. DNA was extracted using the Maxwell RSC Cultured Cell DNA kit. The sequencing libraries were generated using the SMRTbell Template Prep Kit 1.0 following the manufacturer's recommended microbial multiplexing protocol. Each set of 4 samples was sequenced using the Pacific Biosciences (PacBio) Sequel platform (v2.1 Chemistry, Sequel SMRT cell 1 M v2, 10-hour movie). The PacBio raw reads were de novo assembled using the PacBio Hierarchical Genome Assembly Process (HGAP) 4.0. The assembled genomes were circularized using Circulator [8]. If the corresponding Illumina short reads were available in NCBI, the closed genomes were further polished with the Illumina short reads using Pilon [9]. In cases where short reads were unavailable, the closed genomes were polished with the PacBio raw reads using the Resequencing module provided by PacBio SMRTLink (from version 5.1 to 8.0) (Pacific Biosciences, Menlo Park, CA).

Among the 217 closed genomes, CFSAN017963 and CFSAN000871 were sequenced on PacBio RS II, assembled using HGAP 3.0, and published previously [10, 11]. They were re-sequenced on PacBio Sequel which, compared to RS II, provided higher throughput and higher consensus accuracy, assembled using HGAP 4.0, and included in the current dataset for comparison purpose.

Table 1 Complete DNA sequences (.Fasta) for the 217 Salmonella genomes

Label	Name of data file/ data set	Data repository and identifier (DOI or accession number)
Data set 1	S. Enteritidis CFSAN000045	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338875.1 [12]
Data set 2	S. Bareilly CFSAN000224	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338855.1 [12]
Data set 3	S. Heidelberg CFSAN000445	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338435.1 [12]
Data set 4	S. 66:z41:- CFSAN000510	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338735.1 [12]
Data set 5	S. Anatum CFSAN000511	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338835.1 [12]
Data set 6	S. Choleraesuis CFSAN000514	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338815.1 [12]
Data set 7	S. Dublin CFSAN000516	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338715.1 [12]
Data set 8	S. Dublin CFSAN000517	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338555.1 [12]
Data set 9	S. Indiana CFSAN000520	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_007232625.2 [12]
Data set 10	S. Paratyphi B CFSAN000535	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_010732035.2 [12]
Data set 11	S. Paratyphi B CFSAN000540	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_003869075.2 [12]
Data set 12	S. Paratyphi B CFSAN000541	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006519455.2 [12]
Data set 13	S. Paratyphi B CFSAN000542	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006518555.2 [12]
Data set 14	S. Paratyphi B CFSAN000545	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338795.1 [12]
Data set 15	S. Paratyphi B CFSAN000546	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338775.1 [12]
Data set 16	S. Paratyphi B CFSAN000548	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338755.1 [12]
Data set 17	S. Paratyphi B CFSAN000549	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338695.1 [12]
Data set 18	S. 40:z4,z24:- CFSAN000554	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338515.1 [12]
Data set 19	S. Derby CFSAN000565	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338675.1 [12]
Data set 20	S. Newport CFSAN000598	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338635.1 [12]
Data set 21	S. Gallinarum CFSAN000606	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338655.1 [12]
Data set 22	S. Saintpaul CFSAN000614	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338535.1 [12]
Data set 23	S. Typhimurium CFSAN000630	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_003871895.2 [12]
Data set 24	S. Typhimurium CFSAN000646	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_007762235.2 [12]
Data set 25	S. Typhimurium CFSAN000648	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_004222645.2 [12]
Data set 26	S. Wien CFSAN000656	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338415.1 [12]
Data set 27	S. Abaetetuba CFSAN000658	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338375.2 [12]
Data set 28	S. Panama CFSAN000713	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338495.1 [12]
Data set 29	S. Thompson CFSAN000736	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338475.1 [12]
Data set 30	S. Bareilly CFSAN000755	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338355.1 [12]
Data set 31	S. Newport CFSAN000827	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_003581245.2 [12]
Data set 32	S. Newport CFSAN000871	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338455.2 [12]
Data set 33	S. Javiana CFSAN000905	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338395.1 [12]
Data set 34	S. Typhimurium CFSAN000982	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_008916265.2 [12]
Data set 35	S. Typhi CFSAN001004	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_011117275.2 [12]
Data set 36	S. Brandenburg CFSAN001008	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_007230335.2 [12]
Data set 37	S. II 47:d: z39 CFSAN001010	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005899965.2 [12]
Data set 38	S. II 48:d: z6 CFSAN001011	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006914355.2 [12]
Data set 39	S. II 50:b: z6 CFSAN001013	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006912375.2 [12]
Data set 40	S. II 13,22:z29:e, n,x CFSAN001015	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005900065.2 [12]
Data set 41	S. II 4:b:- CFSAN001016	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_003873385.2 [12]
Data set 42	S. IIIa 41:z4,z23:- CFSAN001018	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_004222725.2 [12]
Data set 43	S. 40:z35:- CFSAN001045	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_014338425.3 [12]
Data set 44	S. Senftenberg CFSAN001066	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_009221265.2 [12]
Data set 45	S. Montevideo CFSAN001068	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_009221365.2 [12]
Data set 46	S. Javiana CFSAN001070	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338275.1 [12]
Data set 47	S. London CFSAN001081	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338335.1 [12]
Data set 48	S. Havana CFSAN001082	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338295.1 [12]
Data set 49	S. Cubana CFSAN001083	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338315.1 [12]
Data set 50	S. Eastbourne CFSAN001084	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338255.1 [12]
Data set 51	S. Nchanga CFSAN001091	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340155.1 [12]
Data set 52	S. Saintpaul CFSAN001195	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_000486545.2 [12]
Data set 53	S. Meleagridis CFSAN001281	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339875.1 [12]

Table 1 (continued)

Label	Name of data file/ data set	Data repository and identifier (DOI or accession number)
Data set 54	S. Give CFSAN001314	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_000487215.2 [12]
Data set 55	S. Derby CFSAN001315	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340175.1 [12]
Data set 56	S. Soerenga CFSAN001321	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340135.1 [12]
Data set 57	S. Cerro CFSAN001330	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_000487275.2 [12]
Data set 58	S. Miami CFSAN001416	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339975.1 [12]
Data set 59	S. Javiana CFSAN001999	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_014447715.2 [12]
Data set 60	S. Typhimurium CFSAN002003	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_019457675.1 [12]
Data set 61	S. Anatum CFSAN003983	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339915.1 [12]
Data set 62	S. Anatum CFSAN003985	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_010542265.2 [12]
Data set 63	S. Saintpaul CFSAN004079	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340095.1 [12]
Data set 64	S. Saintpaul CFSAN004085	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340075.1 [12]
Data set 65	S. Saintpaul CFSAN004090	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339995.1 [12]
Data set 66	S. Saintpaul CFSAN004112	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339955.1 [12]
Data set 67	S. Saintpaul CFSAN004114	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339895.1 [12]
Data set 68	S. Give CFSAN004343	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339855.2 [12]
Data set 69	S. Muenster CFSAN004344	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339495.2 [12]
Data set 70	S. Minnesota CFSAN006156	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339775.1 [12]
Data set 71	S. Saintpaul CFSAN006166	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339795.1 [12]
Data set 72	S. Rubislaw CFSAN006171	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339755.1 [12]
Data set 73	S. Anatum CFSAN006191	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339715.1 [12]
Data set 74	S. Saintpaul CFSAN006195	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339515.1 [12]
Data set 75	S. Braenderup CFSAN006197	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339535.1 [12]
Data set 76	S. Senftenberg CFSAN006211	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339695.1 [12]
Data set 77	S. Braenderup CFSAN006217	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339675.1 [12]
Data set 78	S. Javiana CFSAN006238	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339655.1 [12]
Data set 79	S. Saintpaul CFSAN006241	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339595.1 [12]
Data set 80	S. Saintpaul CFSAN006246	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339475.1 [12]
Data set 81	S. Agona CFSAN006264	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339615.1 [12]
Data set 82	S. Javiana CFSAN006269	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339635.1 [12]
Data set 83	S. Typhimurium CFSAN008081	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001756025.2 [12]
Data set 84	S. Enteritidis CFSAN008104	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001713515.2 [12]
Data set 85	S. Buzu CFSAN008300	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339555.1 [12]
Data set 86	S. Paratyphi B CFSAN008708	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_003878535.2 [12]
Data set 87	S. Hartford CFSAN008725	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_014432465.2 [12]
Data set 88	S. Hartford CFSAN008740	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339455.1 [12]
Data set 89	S. Sharon CFSAN008783	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006949785.2 [12]
Data set 90	S. Inverness CFSAN008785	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005411755.2 [12]
Data set 91	S. Inverness CFSAN008789	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006949085.2 [12]
Data set 92	S. Muenchen CFSAN008798	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_014431865.2 [12]
Data set 93	S. Inverness CFSAN008810	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006948385.2 [12]
Data set 94	S. Inverness CFSAN008812	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005534535.2 [12]
Data set 95	S. Inverness CFSAN008813	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005991765.2 [12]
Data set 96	S. Schwarzengrund CFSAN008848	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340505.1 [12]
Data set 97	S. Kentucky CFSAN011778	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340485.1 [12]
Data set 98	S. Montevideo CFSAN012296	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340465.1 [12]
Data set 99	S. Infantis CFSAN012496	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_008025235.2 [12]
Data set 100	S. Oranienburg CFSAN012497	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_010651985.2 [12]
Data set 101	S. Sandiego CFSAN012498	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_011436695.2 [12]
Data set 102	S. Sandiego CFSAN012509	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_007626775.2 [12]
Data set 103	S. Give CFSAN012622	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_000732045.2 [12]
Data set 104	S. Derby CFSAN013765	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006651685.2 [12]
Data set 105	S. Meleagridis CFSAN017807	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_007635275.2 [12]
Data set 106	S. Minnesota CFSAN017963	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_002196535.1 [12]

Table 1 (continued)

Label	Name of data file/ data set	Data repository and identifier (DOI or accession number)
Data set 107	<i>S. Typhimurium</i> CFSAN018746	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_003030285.2 [12]
Data set 108	<i>S. Newport</i> CFSAN022621	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_008606975.2 [12]
Data set 109	<i>S. Newport</i> CFSAN022622	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_008606925.2 [12]
Data set 110	<i>S. Choleraesuis</i> CFSAN022623	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_010713575.2 [12]
Data set 111	<i>S. Gallinarum</i> CFSAN022627	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_007626135.2 [12]
Data set 112	<i>S. Choleraesuis</i> CFSAN022628	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_010396085.2 [12]
Data set 113	<i>S. Choleraesuis</i> CFSAN022631	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_010794875.2 [12]
Data set 114	<i>S. Newport</i> CFSAN022633	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_008607025.2 [12]
Data set 115	<i>S. Dublin</i> CFSAN022635	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_008607005.2 [12]
Data set 116	<i>S. Enteritidis</i> CFSAN022640	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006651425.2 [12]
Data set 117	<i>S. Gallinarum</i> CFSAN022642	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_007500235.2 [12]
Data set 118	<i>S. Gallinarum</i> CFSAN024144	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005676255.2 [12]
Data set 119	<i>S. Gaminara</i> CFSAN024152	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005936085.2 [12]
Data set 120	<i>S. Bovismorbificans</i> CFSAN024184	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006846945.2 [12]
Data set 121	<i>S. Bareilly</i> CFSAN024219	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005935865.2 [12]
Data set 122	<i>S. Give</i> CFSAN024230	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006849305.2 [12]
Data set 123	<i>S. Newport</i> CFSAN024414	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340235.1 [12]
Data set 124	<i>S. Newport</i> CFSAN024415	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340215.1 [12]
Data set 125	<i>S. Newport</i> CFSAN024417	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340425.1 [12]
Data set 126	<i>S. Tennessee</i> CFSAN024439	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_037201925.1 [12]
Data set 127	<i>S. Gaminara</i> CFSAN024441	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_008730735.2 [12]
Data set 128	<i>S. Newport</i> CFSAN024495	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_037347515.1 [12]
Data set 129	<i>S. Newport</i> CFSAN024515	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340355.1 [12]
Data set 130	<i>S. Muenchen</i> CFSAN024517	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340335.1 [12]
Data set 131	<i>S. Muenchen</i> CFSAN024522	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340275.1 [12]
Data set 132	<i>S. Newport</i> CFSAN024541	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340315.1 [12]
Data set 133	<i>S. Weltevreden</i> CFSAN024549	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340295.1 [12]
Data set 134	<i>S. Pomona</i> CFSAN024552	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_037345635.1 [12]
Data set 135	<i>S. Newport</i> CFSAN024555	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340255.1 [12]
Data set 136	<i>S. Johannesburg</i> CFSAN024562	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340525.1 [12]
Data set 137	<i>S. Saintpaul</i> CFSAN024564	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340195.1 [12]
Data set 138	<i>S. Irumu</i> CFSAN024579	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339415.1 [12]
Data set 139	<i>S. Rubislaw</i> CFSAN024580	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339375.1 [12]
Data set 140	<i>S. Minnesota</i> CFSAN024581	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338975.1 [12]
Data set 141	<i>S. Rubislaw</i> CFSAN024587	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339055.1 [12]
Data set 142	<i>S. Sandiego</i> CFSAN024590	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339355.1 [12]
Data set 143	<i>S. Newport</i> CFSAN024599	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339175.1 [12]
Data set 144	<i>S. Newport</i> CFSAN024608	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339255.1 [12]
Data set 145	<i>S. Waycross</i> CFSAN024659	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006858285.2 [12]
Data set 146	<i>S. Livingstone</i> CFSAN024717	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_000807055.1 [12]
Data set 147	<i>S. Senftenberg</i> CFSAN024719	NCBI: https://www.ncbi.nlm.nih.gov/nuccore/CP074261.1 [12]
Data set 148	<i>S. Senftenberg</i> CFSAN024722	NCBI: https://www.ncbi.nlm.nih.gov/nuccore/CP074260.1 [12]
Data set 149	<i>S. Anatum</i> CFSAN024763	NCBI: https://www.ncbi.nlm.nih.gov/nuccore/CP074259.1 [12]
Data set 150	<i>S. Brandenburg</i> CFSAN024765	NCBI: https://www.ncbi.nlm.nih.gov/nuccore/CP074258.1 [12]
Data set 151	<i>S. Infantis</i> CFSAN024778	NCBI: https://www.ncbi.nlm.nih.gov/nuccore/CP074261.1 [12]
Data set 152	<i>S. Infantis</i> CFSAN024781	NCBI: https://www.ncbi.nlm.nih.gov/nuccore/CP074256.1 [12]
Data set 153	<i>S. Enteritidis</i> CFSAN026631	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339215.1 [12]
Data set 154	<i>S. Enteritidis</i> CFSAN026633	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339155.1 [12]
Data set 155	<i>S. Rubislaw</i> CFSAN027379	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339115.1 [12]
Data set 156	<i>S. Braenderup</i> CFSAN027384	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339075.1 [12]
Data set 157	<i>S. Kentucky</i> CFSAN027385	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339035.1 [12]
Data set 158	<i>S. Kentucky</i> CFSAN027395	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338955.1 [12]
Data set 159	<i>S. Brandenburg</i> CFSAN027396	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338995.1 [12]

Table 1 (continued)

Label	Name of data file/ data set	Data repository and identifier (DOI or accession number)
Data set 160	S. Montevideo CFSAN028508	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018339015.1 [12]
Data set 161	S. Kentucky CFSAN028527	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338935.1 [12]
Data set 162	S. Weltevreden CFSAN028546	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_010699505.2 [12]
Data set 163	S. Oranienburg CFSAN028548	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_003872475.2 [12]
Data set 164	S. Newport CFSAN028549	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_008029255.2 [12]
Data set 165	S. Shamba CFSAN029516	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_007473075.2 [12]
Data set 166	S. Potsdam CFSAN029622	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018338915.1 [12]
Data set 167	S. Worthington CFSAN029662	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005899025.2 [12]
Data set 168	S. Gaminara CFSAN029843	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005606435.2 [12]
Data set 169	S. Saintpaul CFSAN029855	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006016765.2 [12]
Data set 170	S. Rubislaw CFSAN029856	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006868345.2 [12]
Data set 171	S. Rubislaw CFSAN029865	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006868405.2 [12]
Data set 172	S. Reading CFSAN029868	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005946705.2 [12]
Data set 173	S. Muenchen CFSAN029871	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_004223785.2 [12]
Data set 174	S. Rubislaw CFSAN029872	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005532935.2 [12]
Data set 175	S. Mississippi CFSAN029878	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006868905.2 [12]
Data set 176	S. Inverness CFSAN029879	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005532895.2 [12]
Data set 177	S. Gaminara CFSAN029882	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006870005.2 [12]
Data set 178	S. Inverness CFSAN029891	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006869965.2 [12]
Data set 179	S. Muenchen CFSAN029894	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005606535.2 [12]
Data set 180	S. Inverness CFSAN029895	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_003872625.2 [12]
Data set 181	S. Newport CFSAN029927	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006873825.2 [12]
Data set 182	S. Rubislaw CFSAN029939	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_003872215.2 [12]
Data set 183	S. Bareilly CFSAN029943	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006090915.2 [12]
Data set 184	S. Mississippi CFSAN029945	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_003872395.2 [12]
Data set 185	S. Typhimurium CFSAN029958	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006019175.2 [12]
Data set 186	S. Illb 61:l, v:z35 CFSAN030538	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_003872565.2 [12]
Data set 187	S. Oranienburg CFSAN030601	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_005532075.2 [12]
Data set 188	S. Typhimurium CFSAN033950	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001652385.2 [12]
Data set 189	S. Illb 60:r, e:n, x:z15 CFSAN044865	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001831555.2 [12]
Data set 190	S. I 38:k- CFSAN044875	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001831715.2 [12]
Data set 191	S. Meleagridis CFSAN044885	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001831915.2 [12]
Data set 192	S. Rubislaw CFSAN044888	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001831985.2 [12]
Data set 193	S. Saintpaul CFSAN044909	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001833225.2 [12]
Data set 194	S. Inverness CFSAN044911	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001834735.2 [12]
Data set 195	S. Rubislaw CFSAN044921	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001833255.2 [12]
Data set 196	S. 60:r, e:n, x:z15 CFSAN044923	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001832255.2 [12]
Data set 197	S. Saintpaul CFSAN044925	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001833305.2 [12]
Data set 198	S. Rubislaw CFSAN044945	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001834965.2 [12]
Data set 199	S. Enteritidis CFSAN051827	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001973785.2 [12]
Data set 200	S. Enteritidis CFSAN051882	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001972625.2 [12]
Data set 201	S. Enteritidis CFSAN051890	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_001972545.2 [12]
Data set 202	S. Senftenberg CFSAN056662	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006344865.2 [12]
Data set 203	S. Eastbourne CFSAN059883	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_002266205.2 [12]
Data set 204	S. Dublin CFSAN059898	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_002266085.2 [12]
Data set 205	S. Infantis CFSAN059939	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340565.1 [12]
Data set 206	S. Infantis CFSAN059940	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_018340545.1 [12]
Data set 207	S. Give CFSAN060804	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_002507915.2 [12]
Data set 208	S. Uganda CFSAN060807	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_002507875.2 [12]
Data set 209	S. I 1,3,19:l, v:1,2 CFSAN060808	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_002507745.2 [12]
Data set 210	S. Manchester CFSAN060809	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_002507865.2 [12]
Data set 211	S. Anatum CFSAN064276	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_002507715.2 [12]
Data set 212	S. Javiana CFSAN073546	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006199305.2 [12]

Table 1 (continued)

Label	Name of data file/ data set	Data repository and identifier (DOI or accession number)
Data set 213	S. Oranienburg CFSAN073549	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_007484345.2 [12]
Data set 214	S. Saintpaul CFSAN073553	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006315305.2 [12]
Data set 215	S. Javiana CFSAN073554	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006199285.2 [12]
Data set 216	S. Muenchen CFSAN073555	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_007634455.2 [12]
Data set 217	S. Mbandaka CFSAN082805	NCBI: https://identifiers.org/ncbi/insdc.gca:GCA_006520715.2 [12]

Abbreviations

FDA-CFSAN	U.S. Food and Drug Administration's Center for Food Safety and Applied Nutrition
NCBI	National Center for Biotechnology Information
SNP	Single Nucleotide Polymorphism
SRA	Short-read Archive
WGS	Whole Genome Sequencing

Author contributions

M.H. designed the study. J.J. and M.H. performed all sequencing and wet lab work. Y.L. assembled the sequencing data and closed the genomes. M.H. and M.B. submitted the closed genomes and raw reads to NCBI, respectively. Y.L. wrote the manuscript. All authors read, provided feedback, and approved the final manuscript.

Funding

This study was supported by funding from the FDA Human Foods Program Intramural Funds.

Data availability

Data is provided within the manuscript (Table 1), which have been deposited in NCBI with the primary accession number PRJNA186035.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 13 November 2024 / Accepted: 31 January 2025

Published online: 28 February 2025

References

1. Allard MW, Strain E, Melka D, Bunning K, Musser SM, Brown EW, Timme R. Practical value of food pathogen traceability through building a whole-genome sequencing network and database. *J Clin Microbiol.* 2016;54(8):1975–83. <https://doi.org/10.1128/JCM.00081-16>

2. Fraser CM, Eisen JA, Nelson KE, Paulsen IT, Salzberg SL. The value of complete microbial genome sequencing (you get what you pay for). *J Bacteriol.* 2002;184(23):6403–5. <https://doi.org/10.1128/JB.184.23.6403-6405.2002>

3. Korlach J. Returning to more finished genomes. *Genom Data.* 2014;2:46–8. <https://doi.org/10.1016/j.gdata.2014.02.003>

4. Gilchrist CA, Turner SD, Riley MF, Petri WA Jr, Hewlett EL. Whole-genome sequencing in outbreak analysis. *Clin Microbiol Rev.* 2015;28(3):541–63. <https://doi.org/10.1128/CMR.00075-13>

5. Zhou Y, Ren M, Zhang P, Jiang D, Yao X, Luo Y, Yang Z, Wang Y. Application of nanopore sequencing in the detection of foodborne microorganisms. *Nanomaterials.* 2022;12(9):1534. <https://doi.org/10.3390/nano12091534>

6. Sikorski MJ, Hazen TH, Desai SN, Nimarota-Brown S, Tupua S, Sialeipata M, Rambocus S, Ingle DJ, Duchene S, Ballard SA, Valcanis M, Zufan S, Ma J, Sahl JW, Maes M, Dougan G, Thomsen RE, Robins-Browne RM, Levine MM. Rasko DA, 2022. Persistence of rare *Salmonella* Typhi genotypes susceptible to first-line antibiotics in the remote islands of Samoa. *mBio.* 2022;13(5):e0192022. <https://doi.org/10.1128/mbio.01920-22>

7. Stevens EL, Carleton HA, Beal J, Tillman GE, Lindsey RL, Lauer AC, Pightling A, Jarvis KG, Ottesen A, Ramachandran P, Hintz L, Lee S, Katz, Jason P, Folster JM, Whichard E, Trees RE, Timme P, McDermott B, Wolpert M, Bazaco S, Zhao S, Lindley BB, Bruce PM, Griffin M, Hoffmann M, Wise R, Tauxe P, Gerner-Smidt M, Musser, Chris Braden. Use of whole genome sequencing by the federal inter-agency collaboration for genomics for food and feed safety in the United States. *J Food Prot.* 2022;85(5):755–722. <https://doi.org/10.4315/JFP-21-437>

8. Hunt M, De Silva N, Otto TD, Parkhill J, Keane J, Harris SR. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol.* 2015;16:294. <https://doi.org/10.1186/s13059-015-0849-0>

9. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Earl AM. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE.* 2014;9(11):e112963. <https://doi.org/10.1371/journal.pone.0112963>

10. Wang W, Hoffmann M, Laasri A, Jacobson AP, Melka D, Curry PE, Hammack TS, Zheng J. Complete genome sequence of *Salmonella enterica* subsp. *Enterica* Serovar Minnesota strain. *Genome Biol Evol.* 2017;9(10):2727–31. <https://doi.org/10.1093/gbe/evx209>

11. Zheng J, Luo Y, Reed E, Bell R, Brown EW, Hoffmann M. Whole-genome comparative analysis of *Salmonella enterica* Serovar Newport strains reveals lineage-specific divergence. *Genome Biol Evol.* 2017;9(4):1047–50. <https://doi.org/10.1093/gbe/evx065>

12. The 217 genomes were under the same BioProject number. PRJNA186035 - NCBI. <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA186035>

Publisher’s note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.