



Genome Sequences of 18 *Salmonella enterica* Serotype Hadar Strains Collected from Patients in the United States

 Hattie E. Webb,^{a,b} Justin Y. Kim,^{a,b} Kaitlin A. Tagg,^{a,b}  Fernando de la Cruz,^c Arancha Peñil-Celis,^c Beth Tolar,^b Zachary Ellison,^{b,d} Colin Schwensohn,^b Joshua Brandenburg,^{b,d} Megin Nichols,^b  Jason P. Folster^b

^aASRT, Inc., Suwanee, Georgia, USA

^bDivision of Foodborne, Waterborne, and Environmental Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

^cInstituto de Biomedicina y Biotecnología de Cantabria, Universidad de Cantabria, Santander, Spain

^dOak Ridge Institute for Science and Education, Oak Ridge, Tennessee, USA

ABSTRACT Despite being linked to a number of recent poultry-associated outbreaks in the United States, few reference genomes are available for *Salmonella enterica* serotype Hadar. Here, we address this need by reporting 18 *Salmonella* Hadar genomes from samples collected from patients in the United States between 2014 and 2020.

S *Salmonella enterica* serotype Hadar infections in humans in the United States increased in 2020 and 2021, compared with previous years, despite an overall decline in reported salmonellosis cases (1). Many infections occurred as part of recent outbreaks linked to either backyard poultry flocks (e.g., chickens and ducks) or consumption of ground turkey, but isolates linked to these different sources demonstrated a high degree of core genome relatedness (1, 2). Exploring the accessory genome may improve strain differentiation, as well as our understanding of the recent increase and evolution of this serotype. Here, we generated assemblies for 18 *S.* Hadar isolates collected from U.S. patients to serve as references for future investigations.

Briefly, isolates originated from clinical diagnostic laboratories or public health laboratories (PHLs) as part of the Centers for Disease Control and Prevention (CDC) national passive *Salmonella* surveillance (<https://www.cdc.gov/national-surveillance/salmonella-surveillance.html>); therefore, isolation methods varied by site (3). Isolates underwent short-read sequencing (<https://www.cdc.gov/pulsenet/pathogens/wgs.html>), and serotypes were confirmed using SeqSero2 v0.1 (4). Genomes were screened for resistance determinants and plasmids using the ResFinder database (downloaded 30 July 2020) (90% identity and a 50% cutoff value), the PointFinder scheme for *Salmonella* spp. (downloaded 30 August 2019) implemented in staramr v0.4.0 (5), a modified PlasmidFinder database (90% identity and 60% coverage) (<https://cge.cbs.dtu.dk/services>), and COPLA (6). Sequence types (STs) were determined using staramr v0.4.0 (with MLST software [<https://github.com/tseemann/mlst>] and PubMLST [7]). This report is a product of activities approved by the CDC internal review board (approval number 7172).

Isolates were selected for long-read sequencing based on diverse accessory genome content. Genomic DNA was extracted (Wizard genomic DNA purification kit [Promega, Madison, WI, USA], with a modification of the manufacturer's protocol) from cultures that had been incubated on tryptic soy agar-sheep blood overnight at 37°C. Libraries were prepared using the rapid barcoding kit (SQK-RBK004; Oxford Nanopore Technologies [ONT], Oxford, UK) according to the manufacturer's protocol and sequenced for 72 h on a GridION sequencing platform (R9.4.1 flow cells; ONT). Reads were base called using Guppy v4.2.2 and filtered for quality using MinKNOW (ONT). Hybrid assemblies were generated, polished, circularized, and rotated using Unicycler v0.4.8 (conservative option) (8); corresponding Illumina short reads that had been previously generated at the PHL (BioNumerics v7.6 [Applied Maths NV,

Editor David Rasko, University of Maryland School of Medicine

Copyright © 2022 Webb 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 Jason P. Folster, jpfolster@cdc.gov.

The authors declare no conflict of interest.

Received 22 June 2022

Accepted 16 August 2022

Published 29 August 2022

TABLE 1 Summary information for 18 *Salmonella enterica* serotype Hadar (ST33) genomes from samples collected from patients in the United States

| Strain | Accession no. | Collection yr | BioSample | Short-read SRA | Long-read SRA | GenBank | PTU (plasmid replicon) ^a | Antimicrobial resistance determinants | Short-read findings | | | Long-read findings | | | GC content (%) | Total size (bp) | Mean coverage (x) |
|--------------|---------------|---------------|--------------|----------------|---------------|--|--|---------------------------------------|-----------------------|--------------|-----------------------------|-----------------------|--------------|----------------|----------------|-----------------|-------------------|
| | | | | | | | | | Mean read length (bp) | No. of reads | Contig N ₅₀ (bp) | Mean read length (bp) | No. of reads | No. of contigs | | | |
| 2014AM-1331 | SAMN05596322 | 2014 | SAMN05596322 | SRR4044556 | SRR19768540 | CP093126 | — | <i>aph(3'')-Ib, aph(6)-Id, tet(A)</i> | 295.9 | 1,617,299 | 4,702,738 | 4,302 | 391,542 | 3 | 52.26 | 4,741,847 | 170 |
| 2014AM-2067 | SAMN05596277 | 2014 | SAMN05596277 | SRR4044454 | SRR19768539 | CP093127 CP093128 CP093122 | PTU-N3 PTU-E10 — | — | 286.7 | 1,838,942 | 4,763,043 | 5,225.5 | 165,516 | 4 | 52.22 | 4,777,204 | 192 |
| 2015AM-0414 | SAMN07268462 | 2015 | SAMN07268462 | SRR5740069 | SRR19768530 | CP093123 CP093124 CP093125 CP093120 | PTU-E22 PTU-E1 (ColE1) PTU-E19 (ColI56) ^b — | — | 278.3 | 1,430,061 | 4,801,674 | 5,580.5 | 174,520 | 2 | 52.25 | 4,805,578 | 148 |
| 2015AM-0511 | SAMN07415498 | 2015 | SAMN07415498 | SRR5868150 | SRR19768529 | CP093121 CP093140 | PTU-E19 (ColI56) — | <i>aph(3'')-Ib, aph(6)-Id, tet(A)</i> | 274.7 | 1,566,516 | 4,778,352 | 5,993 | 223,197 | 7 | 52.21 | 4,805,332 | 162 |
| 2016AM-0673 | SAMN13512702 | 2016 | SAMN13512702 | SRR10607636 | SRR19768528 | CP093141 CP093142 CP093143 CP093144 CP093145 CP093146 CP093116 | PTU-E3 [Col440II, Col(pHAD28)] PTU-E7 (ColI56) PTU-E1 (ColE1) PTU-E58 (ColE1) PTU-E58 PTU-E11 (ColpVC) — | <i>bla_{TBBK-1C}</i> | 277.9 | 1,363,160 | 4,703,663 | 5,196.8 | 324,412 | 4 | 52.29 | 4,712,319 | 143 |
| 2016K-0377 | SAMN05250424 | 2016 | SAMN05250424 | SRR3667804 | SRR19768533 | CP093117 CP093118 CP093119 CP093076 CP093077 CP093078 CP093079 CP093080 CP093109 CP093110 CP093111 CP093072 | PTU-E10 PTU-NA PTU-E11 — PTU-X1 (IncX1) PTU-E7 (ColI56) PTU-E3 [Col(pHAD28)] PTU-E11 (ColpVC) PTU-E22 PTU-E1 (ColE1) — | — | 274.0 | 1,081,951 | 4,685,556 | 5,555.7 | 234,342 | 5 | 52.19 | 4,730,499 | 102 |
| 2017AM-0493 | SAMN17129770 | 2017 | SAMN17129770 | SRR13277812 | SRR19768526 | CP093112 CP093113 CP093114 CP093115 | PTU-H1 (IncI1-H) PTU-X1 (IncX1) PTU-E10 (Col8282) [ColpVC] ^c | <i>aph(3'')-Ib, aph(6)-Id, tet(A)</i> | 298.6 | 866,278 | 4,819,027 | 5,104.1 | 104,476 | 3 | 52.18 | 4,829,291 | 90 |
| 2021K-0017 | SAMN17478013 | 2020 | SAMN17478013 | SRR13496954 | SRR19768531 | CP093112 CP093113 CP093114 CP093115 | PTU-H1 (IncI1-H) PTU-X1 (IncX1) PTU-E10 (Col8282) [ColpVC] ^c | — | 272.7 | 662,153 | 4,711,128 | 4,785.5 | 459,695 | 1 | 52.27 | 4,711,128 | 67 |
| PNUSA5002131 | SAMN04961841 | 2016 | SAMN04961841 | SRR3499746 | SRR19768527 | CP093096 CP093097 CP093098 CP093099 | PTU-X1 (IncX1) PTU-N3 PTU-NA [Col(pHAD28)] | — | 289.4 | 2,242,755 | 4,683,655 | 5,252.9 | 345,694 | 4 | 52.17 | 4,807,169 | 173 |
| PNUSA5018090 | SAMN07427456 | 2017 | SAMN07427456 | SRR6014222 | SRR19768524 | CP093097 CP093098 CP093100 CP093101 CP093102 CP093103 CP093104 | PTU-X1 (IncX1) PTU-N3 PTU-E19 (ColI56) PTU-E3 PTU-E11 (ColpVC) PTU-E11 (ColpVC) PTU-H1 (IncI1-H) | — | 294.5 | 977,909 | 4,685,429 | 5,517.2 | 241,397 | 7 | 52.17 | 4,764,800 | 88 |
| PNUSA5021403 | SAMN07521433 | 2017 | SAMN07521433 | SRR5951569 | SRR19768525 | CP093105 CP093106 CP093107 CP093108 CP093093 | PTU-X1 (IncX1) PTU-E3 PTU-E3 [Col(pHAD28)] PTU-E11 (ColpVC) — | <i>aadA2, ant(3'')-I, amlA1, sul3</i> | 284.6 | 828,218 | 4,685,480 | 4,125.7 | 280,865 | 6 | 52.17 | 4,837,812 | 80 |
| PNUSA037609 | SAMN08815166 | 2018 | SAMN08815166 | SRR6916443 | SRR19768523 | CP093105 CP093106 CP093107 CP093108 CP093093 | PTU-X1 (IncX1) PTU-E3 PTU-E3 [Col(pHAD28)] PTU-E11 (ColpVC) — | <i>aph(3'')-Ib, aph(6)-Id, tet(A)</i> | 243.3 | 960,837 | 4,771,555 | 5,272.1 | 147,247 | 3 | 52.24 | 4,775,081 | 90 |
| PNUSA5039582 | SAMN09011259 | 2018 | SAMN09011259 | SRR7093175 | SRR19768538 | CP093094 CP093095 CP093088 | PTU-E58 PTU-NA — | <i>aph(3'')-Ib, aph(6)-Id, tet(A)</i> | 276.7 | 2,708,767 | 4,801,678 | 6,900.7 | 314,036 | 5 | 52.27 | 4,908,883 | 252 |

(Continued on next page)

TABLE 1 (Continued)

| Strain | Collection yr | BioSample | Accession no. | | Long-read SRA | GenBank | PTU (plasmid replicon) ^a | Antimicrobial resistance determinants | Short-read findings | | | Long-read findings | | | Total size (bp) | Mean coverage (X) |
|--------------|---------------|--------------|----------------|---------------|---------------|--|-------------------------------------|---------------------------------------|-----------------------|--------------|-----------------------------|-----------------------|--------------|----------------|-----------------|-------------------|
| | | | Short-read SRA | Long-read SRA | | | | | Mean read length (bp) | No. of reads | Contig N ₅₀ (bp) | Mean read length (bp) | No. of reads | No. of contigs | | |
| PNUSAS067730 | 2019 | SAMN11029788 | SRR8643922 | SRR19768537 | CP093089 | PTU-NA (IncI1-Hy) | aph(3'')-Ib, aph(6)-Id, tet(A) | 272.5 | 859,691 | 4,763,126 | 4,853.2 | 357,105 | 2 | 52.23 | 4,767,759 | 88 |
| | | | | | | PTU-E1 [Col(pHAD28)] ^b | | | | | | | | | | |
| | | | | | | PTU-E58 (ColE1) | | | | | | | | | | |
| | | | | | | PTU-E10 (ColE282) | | | | | | | | | | |
| PNUSAS074905 | 2019 | SAMN11618493 | SRR9040334 | SRR19768532 | CP093091 | aph(3'')-Ib, aph(6)-Id, tet(A) | 230.7 | 1,282,945 | 4,766,350 | 7,143.7 | 112,075 | 3 | 52.21 | 4,887,846 | 102 | |
| | | | | | CP093092 | | | | | | | | | | | |
| | | | | | CP093086 | | | | | | | | | | | |
| | | | | | CP093087 | | | | | | | | | | | |
| PNUSAS127695 | 2019 | SAMN13795856 | SRR10856733 | SRR19768536 | CP093073 | aac(3)-IId, ant(3'')-Ia, bla _{TEM-1b,19} , tet(A) | 263.0 | 612,925 | 4,801,677 | 5,894.7 | 376,137 | 5 | 52.24 | 4,819,091 | 62 | |
| | | | | | CP093074 | | | | | | | | | | | |
| | | | | | CP093075 | | | | | | | | | | | |
| | | | | | CP093135 | | | | | | | | | | | |
| PNUSAS147811 | 2020 | SAMN15239347 | SRR12017862 | SRR19768535 | CP093136 | aph(3'')-Ib, aph(6)-Id, tet(A) | 271.5 | 755,994 | 4,719,084 | 5,320.8 | 289,486 | 2 | 52.2 | 4,813,237 | 75 | |
| | | | | | CP093137 | | | | | | | | | | | |
| | | | | | CP093138 | | | | | | | | | | | |
| | | | | | CP093139 | | | | | | | | | | | |
| PNUSAS148096 | 2020 | SAMN15249649 | SRR12024307 | SRR19768534 | CP093084 | aph(3'')-Ib, aph(6)-Id, tet(A) | 129.2 | 737,381 | 4,708,042 | 4,754.8 | 293,015 | 3 | 52.23 | 4,802,435 | 41 | |
| | | | | | CP093085 | | | | | | | | | | | |
| | | | | | CP093081 | | | | | | | | | | | |
| | | | | | CP093082 | | | | | | | | | | | |
| | | | | | CP093083 | PTU-NA (IncI1-Hy) | | | | | | | | | | |

^a PTU, plasmid taxonomic unit; PTU-NA, plasmid taxonomic unit not assigned; —, no information.

^b Plasmid assigned using updated version of COPLA.

^c Plasmid replicon missing from long-read assembly but present in short reads.

Sint-Martens-Latem, Belgium] quality control metrics: quality score, ≥ 30 ; coverage, $\geq 30\times$) were accessed through NCBI (Table 1). Assemblies were quality controlled using QUAST v5.0.2 (9) and BLASTn v2.9.0 (10) and were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v6.1 (11). Default parameters were used for all software unless otherwise specified.

All 18 *S. Hadar* strains were found to be ST33. Resistance determinants and plasmid types are summarized in Table 1. The most common resistance genes were *aph(3'')-Ib*, *aph(6)-Id*, and *tet(A)*, which were always located on the chromosome ($n = 13$). When present, other resistance genes were associated with Inc11- γ or Col(pHAD28) plasmids. High levels of small plasmids with no known resistance genes were observed, some of which had not been previously characterized, as indicated by small, circular genetic elements not containing a known plasmid replicon. More generally, the hybrid assembly method employed here recovered small plasmids at a higher rate than did long-read-only assembly methods (data not shown). For two genomes, however, some small plasmids were not recovered despite the use of a hybrid assembly method (Table 1), a known artifact of the long-read sequencing process (12).

Data availability. The sequences discussed here have been deposited in GenBank and the SRA under the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

This work was supported through the CDC.

The findings and conclusions of this article are those of the authors and do not necessarily represent the views of the CDC.

We acknowledge the state and local PHLs that participated in the National Antimicrobial Resistance Monitoring System (NARMS) and PulseNet.

REFERENCES

- Nichols M, Gollara L, Palacios A, Stapleton GS, Basler C, Hoff C, Low M, McFadden K, Koski L, Leeper M, Brandenburg J, Tolar B. 2021. *Salmonella* illness outbreaks linked to backyard poultry purchasing during the COVID-19 pandemic: United States, 2020. *Epidemiol Infect* 149:e234. <https://doi.org/10.1017/S0950268821002132>.
- Centers for Disease Control and Prevention. 2021. *Salmonella* outbreak linked to ground turkey. <https://www.cdc.gov/salmonella/hadar-04-21/index.html>. Accessed 31 January 2022.
- Tolar B, Joseph LA, Schroeder MN, Stroika S, Ribot EM, Hise KB, Gerner-Smith P. 2019. An overview of PulseNet USA databases. *Foodborne Pathog Dis* 16:457–462. <https://doi.org/10.1089/fpd.2019.2637>.
- Zhang S, den Bakker HC, Li S, Chen J, Dinsmore BA, Lane C, Lauer AC, Fields PI, Deng X. 2019. SeqSero2: rapid and improved *Salmonella* serotype determination using whole-genome sequencing data. *Appl Environ Microbiol* 85:e01746–19. <https://doi.org/10.1128/AEM.01746-19>.
- Bharat A, Petkau A, Avery BP, Chen JC, Folster JP, Carson CA, Kearney A, Nadon C, Mabon P, Thiessen J, Alexander DC, Allen V, El Bailey S, Bekal S, German GJ, Haldane D, Hoang L, Chui L, Minion J, Zahariadis G, Domselaar GV, Reid-Smith RJ, Mulvey MR. 2022. Correlation between phenotypic and in silico detection of antimicrobial resistance in *Salmonella enterica* in Canada using Staramr. *Microorganisms* 10:292. <https://doi.org/10.3390/microorganisms10020292>.
- Redondo-Salvo S, Bartomeus-Peñalver R, Vielva L, Tagg KA, Webb HE, Fernández-López R, de la Cruz F. 2021. COPLA, a taxonomic classifier of plasmids. *BMC Bioinformatics* 22:390. <https://doi.org/10.1186/s12859-021-04299-x>.
- Jolley KA, Maiden MC. 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 11:595. <https://doi.org/10.1186/1471-2105-11-595>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- Mikheenko A, Pribelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* 34:i142–i150. <https://doi.org/10.1093/bioinformatics/bty266>.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic Local Alignment Search Tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Tatusova T, DiCuccio M, Badretdin A, Chetverin 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>.
- Wick RR, Judd LM, Wyres KL, Holt KE. 2021. Recovery of small plasmid sequences via Oxford Nanopore sequencing. *Microb Genom* 7:000631. <https://doi.org/10.1099/mgen.0.000631>.