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

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

\(^a\)ASRT, Inc., Suwanee, Georgia, USA
\(^b\)Division of Foodborne, Waterborne, and Environmental Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, USA
\(^c\)Instituto de Biomedicina y Biotecnología de Cantabria, Universidad de Cantabria, Santander, Spain
\(^d\)Oak 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.

*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/nationalsurveillance/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, October 2022 Volume 11 Issue 10 10.1128/mra.00522-22]
TABLE 1 Summary information for 18 *Salmonella enterica* serotype Hadar (ST33) genomes from samples collected from patients in the United States

| Strain         | Collection yr | Biopsy Sample | Short-read Accession no. | Long-read Accession no. | GenBank Accession no. | PTU (plasmid replicon) | Mean read length (bp) | Mean read length (bp) | GC content (%) | Total size (bp) | Mean coverage (%) |
|----------------|---------------|---------------|--------------------------|--------------------------|------------------------|------------------------|----------------------|----------------------|-----------------|----------------|------------------|
| 2014AM-1331    | 2014          | SAMN05596322  | SRR4044556               | SRA376540                | CP093126               | aph(3')-Ib, aph(3')-Ia | 295.9                | 1,671,299            | 3                | 72.18           | 4,741,847         |
| 2014AM-2067    | 2014          | SAMN05596277  | SRA4044454               | SRA376539                | CP093122               | aph(3')-Ib, aph(3')-Ia | 286.7                | 1,838,042            | 4                | 72.22           | 4,777,204         |
| 2015AM0414     | 2015          | SAMN07268462  | SRA3740609               | SRA376530                | CP093120               | aph(3')-Ib, aph(3')-Ia | 278.3                | 1,430,061            | 2                | 52.22           | 4,805,578         |
| 2015AM0511     | 2015          | SAMN07415498  | SRA3768650               | SRA376529                | CP093121               | aph(3')-Ib, aph(3')-Ia | 274.7                | 1,566,516            | 7                | 52.21           | 4,805,332         |
| 2015AM0673     | 2015          | SAMN13512702  | SRA376528                | SRA376528                | CP093122               | aph(3')-Ib, aph(3')-Ia | 277.9                | 1,363,160            | 4                | 52.29           | 4,712,319         |
| 2016K-0377     | 2016          | SAMN05250424  | SRA3667804               | SRA376533                | CP093126               | aph(3')-Ib, aph(3')-Ia | 274.0                | 1,081,951            | 5                | 52.19           | 4,730,499         |
| 2017AM0493     | 2017          | SAMN17129770  | SRA3727812               | SRA376526                | CP093126               | aph(3')-Ib, aph(3')-Ia | 298.6                | 866,278              | 3                | 52.18           | 4,829,291         |
| 2021K-0017     | 2020          | SAMN17478013  | SRA376954                | SRA376531                | CP093126               | aph(3')-Ib, aph(3')-Ia | 272.7                | 662,153              | 7                | 52.27           | 4,711,128         |
| PNUSA0002131   | 2016          | SAMN04961843  | SRA499746                | SRA376527                | CP093126               | aph(3')-Ib, aph(3')-Ia | 289.4                | 2,242,755            | 4                | 52.17           | 4,807,169         |
| PNUSA018090    | 2017          | SAMN07247456  | SRA38014222              | SRA376524                | CP093126               | aph(3')-Ib, aph(3')-Ia | 294.5                | 977,909              | 7                | 52.17           | 4,764,800         |
| PNUSA021403    | 2017          | SAMN07521433  | SRA3951369               | SRA376525                | CP093126               | aph(3')-Ib, aph(3')-Ia | 284.6                | 828,218              | 6                | 52.17           | 4,837,812         |
| PNUSA037609    | 2018          | SAMN08151666  | SRA3961443               | SRA376523                | CP093126               | aph(3')-Ib, aph(3')-Ia | 243.3                | 960,837              | 3                | 52.24           | 4,775,081         |
| PNUSA039582    | 2018          | SAMN09011259  | SRA37091175              | SRA376538                | CP093126               | aph(3')-Ib, aph(3')-Ia | 276.7                | 2,708,767            | 5                | 52.27           | 4,908,883         |

(Continued on next page)
### TABLE 1 (Continued)

| Strain          | Accession no. | Collection yr | Short-read SRA | Long-read SRA | CollBank         | PTU (plasmid replicon<sup>a</sup>) | Antimicrobial resistance determinants                                                                 | Mean read length (bp) | No. of reads | Contig N<sub>50</sub> (bp) | Mean read length (bp) | No. of reads | No. of contigs | GC content (%) | Total size (bp) | Mean coverage (×) |
|-----------------|---------------|---------------|----------------|---------------|-----------------|------------------------------------|-----------------------------------------------------------------------------------------------|-----------------------|--------------|---------------------|-----------------------|--------------|----------------|----------------|-----------------|-------------------|
| PTU-NA (IncI<sub>1-Ig</sub>) | CP093089      | —             | —              | —             | CP093090        | PTU-NA (IncI<sub>1-Ig</sub>)<sup>b</sup> | —                                                                                | 272.5                 | 859,691      | 4,763,126           | 4,853.2               | 357,105       | 2               | 52.23          | 4,767,759        | 88                |
| PTU-E1 (ColE1)  | CP093091      | —             | —              | —             | CP093092        | —                    | —                                                                                | 230.7                 | 1,282,945    | 4,766,350           | 7,143.7               | 112,075       | 3               | 52.21          | 4,887,846        | 102               |
| PTU-E1 (ColE1)  | CP093074      | —             | —              | —             | CP093075        | PTU-E1 (IncI<sub>1</sub>) | —                                                                                | 263.0                 | 612,925      | 4,801,677           | 5,894.7               | 376,137       | 5               | 52.24          | 4,819,091        | 62                |
| PTU-E1 (ColE1)  | CP093087      | CP093075      | CP093075       | CP093075      | CP093076        | —                    | —                                                                                | 271.5                 | 755,934      | 4,719,084           | 5,320.8               | 289,486       | 2               | 52.2           | 4,813,237        | 75                |
| PTU-E1 (ColE1)  | CP093093      | CP093094      | CP093094       | CP093094      | CP093095        | —                    | —                                                                                | 129.2                 | 737,381      | 4,708,042           | 4,754.8               | 293,015       | 3               | 52.23          | 4,802,435        | 41                |

<sup>a</sup> PTU, plasmid taxonomic unit; PTU-NA, plasmid taxonomic unit not assigned; —, no information.

<sup>b</sup> Plasmid assigned using updated version of COPLA.

<sup>c</sup> Plasmid replicon missing from long-read assembly but present in short reads.
Sint-Martens-Latem, Belgium] quality control metrics: quality score, ≥30; coverage, ≥30× (C2)
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 IncI1-I 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).

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
1. Nichols M, Gollarza 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.
2. 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.
3. Tolar B, Joseph LA, Schroeder MN, Stroika S, Ribot EM, Hise KB, Gerner-Smidt P. 2019. An overview of PulseNet USA databases. Foodborne Pathog Dis 16:457–462. https://doi.org/10.1089/fpd.2019.2637.
4. 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.
5. 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, Domsealaw 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.
6. Redondo-Salvo S, Bartomeu-Peñalver R, Vieki 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.
7. 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.
8. 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.
9. Mikheenko A, Prijibelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics 34:1142–1150. https://doi.org/10.1093/bioinformatics/bty266.
10. 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.
11. Tatusova T, DiCuccio M, Badredin A, Chetverin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Prutit KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.
12. 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.