

Draft genome sequences of three poultry *Salmonella* Shamba isolates from South Africa

Musafiri Karama,¹ Opeyemi U. Lawal,² Valeria R. Parreira,² Mitra Soni,² Yanhong Chen,² Beniamino T. Cenci-Goga,^{1,3} Luca Grispoldi,³ Janita Greyling,⁴ Lawrence Goodridge²

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT Nontyphoidal *Salmonella enterica* serovars are foodborne pathogens commonly transmitted through poultry products. Draft genome sequences of three *Salmonella enterica* subsp. *enterica* serovar Shamba isolates which were obtained from poultry house dust in South Africa are reported herein.

KEYWORDS *Salmonella* Shamba, sequence, draft genome, poultry house dust

Nontyphoidal Salmonellosis is at the top of the list among bacterial foodborne diseases transmitted through poultry products, worldwide. More than 2,600 *Salmonella* serovars have been identified so far, with more than 500,000 *Salmonella* whole-genome sequences deposited in public databases (<https://www.ncbi.nlm.nih.gov/datasets/genome/?taxon=590>). However, whole-genome sequences of a number of *Salmonella* serovars such as Shamba remain rare or unavailable. *Salmonella* Shamba has been previously associated with food poisoning and isolated from symptomless animal and human carriers (1, 2). This work presents the draft genome sequences of three *Salmonella* Shamba poultry isolates. The isolates were cultured from poultry house dust samples which were collected in 1999, 2000, and 2001 in South Africa.

Salmonella was isolated as described previously (3). Briefly, 25 g of poultry house dust was cultured in 225 mL buffered peptone water for 16–18 hours at 37°C. A 100 µL aliquot of buffered peptone water was inoculated and incubated into 10 mL Rappaport-Vassiliadis medium at 42°C for 24 hours before plating on XLD medium for 24 hours at 37°C. Presumptive *Salmonella* colonies were initially identified by Gram staining, catalase, oxidase, and spot indole tests. The *Salmonella* status of presumptive isolates was verified and confirmed by the API10S and PCR (4). *Salmonella* was serotyped by the Kauffman-White Scheme (5).

Before whole-genome sequencing, frozen *Salmonella* isolates were resuscitated into tryptic soy broth for 24 hours at 35°C, followed by streaking on *Salmonella Shigella* Agar and purification on Brilliance *Salmonella* Agar Base. Purified colonies were identified by VITEK (Biomérieux, Canada). DNA was extracted from pure colonies by the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). Sequencing was carried out as described previously (3, 6). Briefly, DNA libraries were prepared using the Illumina DNA Prep Tagmentation Kit and Integrated DNA Technologies for Illumina DNA/RNA unique dual indexes (3, 6). The Illumina MiniSeq system was used to perform paired-end (2 × 150 bp) sequencing. Default parameters were used in all bioinformatics tools. Pre-processing of raw reads was performed by FastQC v0.11.9 (<https://github.com/s-andrews/FastQC>) and Trimmomatic v0.39 (7). SKESA v2.4.0 was used for *de novo* assembly of reads with quality scores above 20 (8). Assembly quality was assessed with QUAST v5.2 (9), and genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline v6.6 (10). The isolates were serotyped and the sequence types were identified by SISTR v1.0 (11) and the PubMLST scheme, respectively (12). Antimicrobial resistance encoding

Editor Vincent Michael Bruno, University of Maryland School of Medicine, Baltimore, Maryland, USA

Address correspondence to Musafiri Karama, musafiri.karama@up.ac.za.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 25 March 2024

Accepted 16 July 2024

Published 31 July 2024

Copyright © 2024 Karama 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/).

TABLE 1 Summary of sequence metrics of three *Salmonella* Shamba isolates recovered from poultry house dust in South Africa

ID	Collection date and origin	Number of contigs	Number of reads	Genome size	Protein coding genes	% GC	N_{50} (bp)	Coverage	Plasmid type	tRNAs	ncRNAs	AMR gene profile	Assembly and SRA accession numbers
<i>Salmonella</i> Shamba SE-110	1999 Poultry house dust	26	1,626,621	4,596,393	4,203	52.3	395,585	94.08×	None	70	9	<i>aac(6)-laa</i> , <i>mdtK</i> , <i>sdiA</i> , <i>fosA</i> , <i>golS</i> , <i>mdsABC</i>	JAWDKU000000000.1 SRR26197580
<i>Salmonella</i> Shamba SE-135	2000 Poultry house dust	29	1,568,060	4,598,475	4,202	52.3	477,344	89.91×	None	72	9	<i>aac(6)-laa</i> , <i>mdtK</i> , <i>sdiA</i> , <i>fosA</i> , <i>golS</i> , <i>mdsABC</i>	JAWDKT000000000.1 SRR26197579
<i>Salmonella</i> Shamba SE-138	2001 Poultry house dust	29	1,640,609	4,703,673	4,314	52.2	329,177	91.03×	Incl-gamma/K	71	9	<i>aac(6)-laa</i> , <i>mdtK</i> , <i>sdiA</i> , <i>fosA</i> , <i>golS</i> , <i>mdsABC</i> , <i>TEM-1</i> , <i>sul2</i>	JAWDKS000000000.1 SRR26197578

genes and plasmid types were identified using the CARD database (13), and MOB-suite v3.1 (14), respectively.

Genome sequencing revealed that the three *Salmonella* isolates belonged to serotype *Salmonella* Shamba. Genome annotation data regarding genome size, number of reads and contigs, protein-coding genes, tRNAs and ncRNAs, coverage, % GC content, plasmid presence, and antimicrobial resistance gene profiles for the three *Salmonella* isolates are summarized in Table 1. All three *Salmonella* Shamba isolates possessed one CRISPR Array each.

Overall, the draft genomes described in this report will be an indispensable resource for understanding the evolutionary and functional genomics of *Salmonella* Shamba.

ACKNOWLEDGMENTS

This research was funded by the National Research Foundation (NRF) of South Africa SARCHI COP Grant 120317 and the Canada First Research Excellence Fund.

We thank the staff of the Bacteriology Laboratory, Department of Veterinary Tropical Diseases, Faculty of Veterinary Science, University of Pretoria.

AUTHOR AFFILIATIONS

¹Veterinary Public Health Section, Department of Paraclinical Sciences, Faculty of Veterinary Science, University of Pretoria, Onderstepoort, South Africa

²Canadian Research Institute for Food Safety (CRIFS), Department of Food Science, University of Guelph, Guelph, Ontario, Canada

³Department of Veterinary Medicine, Laboratorio di Ispezione Degli Alimenti di Origine Animale, University of Perugia, Perugia, Italy

⁴Department of Veterinary Tropical Diseases, Faculty of Veterinary Science, University of Pretoria, Onderstepoort, South Africa

AUTHOR ORCIDs

Musafiri Karama  <http://orcid.org/0000-0002-6197-5309>

Opeyemi U. Lawal  <http://orcid.org/0000-0003-2352-2832>

Lawrence Goodridge  <http://orcid.org/0009-0008-9965-3509>

FUNDING

Funder	Grant(s)	Author(s)
National Research Foundation (NRF)	SARCHI COP Grant 120317	Musafiri Karama
Canada First Research Excellence Fund (CFREF)	(CFREF)	Lawrence Goodridge

AUTHOR CONTRIBUTIONS

Musafiri Karama, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing | Opeyemi U. Lawal, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing.

DATA AVAILABILITY

Whole-genome sequences for *Salmonella* Shamba SE-110, SE-135, and SE-138 were deposited in the DDBJ/ENA/GenBank database under the accession numbers [JAWDKU000000000.1](https://doi.org/10.1186/s13059-018-1540-z), [JAWDKT000000000.1](https://doi.org/10.1186/s13059-018-1540-z), [JAWDKS000000000.1](https://doi.org/10.1186/s13059-018-1540-z) and the SRA accession numbers [SRR26197580](https://doi.org/10.1186/s13059-018-1540-z), [SRR26197579](https://doi.org/10.1186/s13059-018-1540-z) and [SRR26197578](https://doi.org/10.1186/s13059-018-1540-z).

REFERENCES

- Contreras-Soto M, Medrano-Félix J, Ibarra-Rodríguez J, Martínez-Urtaza J, Chaidez Q, Castro-del Campo N. 2019. The last 50 years of *Salmonella* in Mexico: sources of isolation and factors that influence its prevalence and diversity. *Revista Bio Ciencias* 6. <https://doi.org/10.15741/revbio.06.nesp.e540>
- Lapage SP, Taylor J, Nicewonger CR, Phillips AG. 1966. New serotypes of *Salmonella* identified before 1964 at the *Salmonella* reference laboratory, Colindale. *Int J Syst Evol Microbiol* 16:253–298. <https://doi.org/10.1099/00207713-16-3-253>
- Karama M, Lawal OU, Parreira VR, Soni M, Chen Y, Cenci-Goga BT, Grispoldi L, Greyling J, Goodridge L. 2024. Draft genome sequences of two *Salmonella* Uzaramo isolates from poultry in South Africa. *Microbiol Resour Announc* 13:e0102623. <https://doi.org/10.1128/mra.01026-23>
- Soumet C, Ermel G, Rose V, Rose N, Drouin P, Salvat G, Colin P. 1999. Identification by a multiplex PCR-based assay of *Salmonella* Typhimurium and *Salmonella* Enteritidis strains from environmental swabs of poultry houses. *Lett Appl Microbiol* 29:1–6. <https://doi.org/10.1046/j.1365-2672.1999.00559.x>
- Popoff MY. 1997. *Antigenic formulas of the Salmonella serovars*. WHO collaborating center for reference and research on Salmonella.
- Lawal OU, Bryan N, Soni M, Chen Y, Precious M, Parreira VR, Goodridge L. 2024. Whole genome sequence of *Vibrio cholerae* NB-183 isolated from freshwater in Ontario, Canada harbors a unique gene repertoire. *BMC Genom Data* 25:18. <https://doi.org/10.1186/s12863-024-01204-2>
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Souvorov A, Agarwala R, Lipman DJ. 2018. SKESA: strategic k-mer extension for scrupulous assemblies. *Genome Biol* 19:153. <https://doi.org/10.1186/s13059-018-1540-z>
- Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
- Tatusova T, DiCuccio M, Badretdin A, Chetvermin 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>
- Yoshida CE, Kruczkiewicz P, Laing CR, Lingohr EJ, Gannon VPJ, Nash JHE, Taboada EN. 2016. The *Salmonella in silico* typing resource (SISTR): an open web-accessible tool for rapidly typing and subtyping draft *Salmonella* genome assemblies. *PLoS One* 11:e0147101. <https://doi.org/10.1371/journal.pone.0147101>
- Jolley KA, Maiden MCJ. 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 11:1–11. <https://doi.org/10.1186/1471-2105-11-595>
- Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen A-LV, Cheng AA, Liu S, et al. 2020. CARD 2020: antibiotic resistance surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res* 48:D517–D525. <https://doi.org/10.1093/nar/gkz935>
- Robertson J, Nash JHE. 2018. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microb Genom* 4:e000206. <https://doi.org/10.1099/mgen.0.000206>