**High-quality genomes of *Salmonella* Typhi 4.3.1 with IncHI1 plasmids encoding for limited resistance determinants**

**Authors**

Allan Zuza1, Alexander M Wailan1,2, Philip M Ashton1, Catherine Anscombe1, Nicholas A Feasey1,3,4, Eva Heinz3,5,6

1 Malawi Liverpool Clinical Research Programme, Kamuzu University of Health Sciences, Blantyre, Malawi

2 Wellcome Sanger Institute, Parasites and Microbes Program, Hinxton, UK

3 Liverpool School of Tropical Medicine, Department of Clinical Sciences, Liverpool, UK

4 University of St Andrews, School of Medicine, St Andrews, UK

5 Liverpool School of Tropical Medicine, Department of Vector Biology, Liverpool, UK

6 Strathclyde Institute of Pharmacy & Biomedical Sciences, University of Strathclyde, Glasgow, UK

For correspondence: eva.heinz@strath.ac.uk

**Keywords**

Antimicrobial resistance, Typhoid, Malawi, drug resistance, gene loss

**Abstract**

Typhoid fever is a significant public health problem endemic in Southeast Asia and Sub-Saharan Africa. Antimicrobial treatment of typhoid is however threatened by the increasing prevalence of antimicrobial resistant (AMR) *S.* Typhi, especially in the globally successful lineage (4.3.1) which has rapidly spread in East and Southern Africa. AMR elements can be found either on plasmids or in one of the three chromosomal integration sites, and there is variability of this across the lineage. Several previous studies with Malawian isolates indicated a clonal, locally spreading lineage with chromosomally integrated resistance genes. In a recent study however we noted three isolates with unusual predicted resistance profiles, and we here present the resolved genomes of these isolates using long- and short-read sequencing. Our work shows that these isolates are potentially imported cases, most closely related to the recently described sub-lineage 4.3.1.EA1, and encode IncHI1 plasmids with reduced resistance profiles compared to the major reference sequence of these plasmids spreading in East Africa. Similar reduced resistance plasmids were reported in a recent large-scale study in five isolates from Tanzania, highlighting the urgency for better coverage of the African continent in genome studies to better understand the dynamics of these potentially co-circulating plasmids.

**Data availability statement**

Raw sequence data is accessible under project ID PRJNA1127853 at SRA and are outlined in [Table 1](#Table1).

Assemblies were submitted to GenBank and are accessible under project ID PRJNA1127853, the individual accessions are presented in [Table 1](#Table1)

**Introduction**

*Salmonella enterica* subspecies *enterica* serovar Typhi (*S*. Typhi) is estimated to cause over 11 million typhoid fever cases annually(1). Typhoid caused by *S*. Typhi is endemic in Southeast Asia and Sub-Saharan Africa and a major cause of under-five mortality(1). Treatment of infections by *S*. Typhi relies on antimicrobial treatment; however, the prevalence of *S.* Typhi resistant to first-line antibiotics is rising (2). The spread of antimicrobial resistance (AMR) has been tracked in unprecedented resolution using large-scale whole-genome sequencing surveillance and can clearly be traced to specific lineages that acquire resistance elements and then spread clonally across large areas. Across Africa, there are two main lineages (3); in Eastern Africa the spread of resistant *S.* Typhi is mainly driven by the highly successful lineage 4.3.1 (formerly haplotype H58; (4, 5)) which is often associated with plasmid IncHI1 and which rapidly spread across the African continent from 2010 (6), whilst in Western Africa, the genotypes 2.3.2 and 3.1.1 dominate (3).

The resistance genes commonly found in *S*. Typhi confer resistance to aminopenicillins (*bla*TEM-1), chloramphenicol (*catA1*), co-trimoxazole (*dfrA, sul1*/*sul2*), and streptomycin (*strAB*) as well as tetracycline (*tetA*) and can be located either on plasmids or on one of the three chromosomal sites through the integration of a Tn2670 like element (3, 6-8), and there is variability of this across the lineage. Previous studies with Malawian isolates have shown the presence of AMR genes in the 4.3.1 haplotype through chromosomal integration (7, 9) and a clonal, locally spreading lineage. The IncHI1 plasmid is known to have substantial variability and is hypothesized to have acquired several resistance gene cassettes through step-wise insertions in an insertion hot-spot (10, 11). Most 4.3.1 isolates in recent large-scale sequencing studies encoding for an IncHI1 plasmid include the core set of resistance genes as described above, which would support the hypothesis that antimicrobial pressure selects for the most resistant lineages which outcompete ones encoding for IncHI1 plasmids encoding for fewer resistances (3, 10).

In a recent study we noted three isolates with unusual resistance profiles; these three were furthermore predicted to carry IncHI1 plasmids which is unusual in the Malawi context (9). We aimed to resolve the structure of these IncHI1 plasmids by performing long-read sequencing and establish whether these were likely imported cases or whether a part of the lineage spreading in Blantyre carried the IncHI1 plasmid. We report a detailed description of the genome and plasmid structure of four isolates, including two with unusual resistance profiles detected and two as comparisons for the insertion site, as well as comparisons of the plasmids with reference sequences for the *S*. Typhi IncHI1 plasmid commonly found in 4.3.1. These data will be a valuable resource for researchers working on comparative genome analyses of *S*. Typhi, in particular in southern Africa, and to improve our understanding of competitive advantages between lineages with different resistance plasmids.

**Methods**

**Long-fragment DNA extraction**

These bacteria isolated from the blood of febrile patients attending Queen Elizabeth Central Hospital in Blantyre, Malawi, as part of the quality assured routine diagnostic microbiology service supported by the Malawi Liverpool Wellcome Programme as described previously (7). ERS327391 and ERS207185 were previously sequenced by Feasey et al and ERS1509723 and ERS1509734 were sequenced by Gauld et al (7, 9) using short-read sequencing only. Isolates were recovered from the MLW sample archive and then plated on nutrient agar (Thermo Scientific™ Oxoid™, United Kingdom) and incubated at 37 ℃ for 18 hours. A single colony was then transferred into 5ml of nutrient broth (Thermo Scientific™ Oxoid™, United Kingdom) and incubated for 18 hours at 37 ℃. The bacteria cells were then concentrated by centrifugation at 3500rpm for 30 minutes. Total nucleic acids were extracted using the MasterPure complete DNA and RNA Purification Kit (Bioresearch Technologies, United Kingdom) according to the manufacturer's instructions.

**Sequencing**

The double-stranded DNA was quantified using the Qubit 4.0 (ThermoFisher Scientific inc.) fluorometer and normalized to 400ng in 7.5 ul using UltraPure™ Distilled Water (Invitrogen, Life Technologies Limited). The normalised DNA was used for library preparation using the rapid barcoding kit (SBQ-RBK004, Oxford Nanopore Technologies plc) following the manufacturer's instructions. The prepared library was sequenced on an Oxford Nanopore r9.4.1 flow cell on the MinION Mk1C sequencer (Oxford Nanopore Technologies plc). Data acquisition was done by the MinKNOWTM software (version 20.10.6, Oxford Nanopore Technologies plc). Base calling and demultiplexing using Guppy Basecalling Software, (C) Oxford Nanopore Technologies plc. (Version 6.0.7+c7819bc). We used Guppy with the default settings and the “dna\_r9.4.1\_450bps\_sup.cfg” configuration file for the guppy\_barcoder. We used the command line argument “--trim\_barcodes” for the debarcoding step using guppy\_barcoder. For guppy\_barcoder we passed “SQK-RBK004” to the “--barcode\_kits” command line argument.

**De novo assembly**

We retrieved the short reads from ENA and performed trimming and filtering using fastp (version 0.23.1--9f2e255, RRID:SCR\_016962) (12, 13). The long-reads derived from nanopore sequencing were trimmed and filtered using Filtlong (version 0.2.0--0c4cbe3, RRID:SCR\_024020) with options “--min\_length 1000” and “--keep\_percent 90” to discard all reads less than 1kb and remove the worst 10% reads (14). We used FastQC (v0.11.9\_cv7, RRID:SCR\_014583) and NanoPlot (version 1.38.1--e303519) with the default settings to check the quality of the filtered short and long reads respectively (15, 16).

We performed long-read-first de novo assembly using the Trycycler (v0.5.4) assembly pipeline (17). Here is a shortened version of the assembly with Trycycler: We created 16 read subsamples using Trycycler subsample. We created draft assemblies on each of the simulated reads using either Flye (version 2.9.1-b1780, RRID:SCR\_017016), Minipolish (version 0.1.2) and Raven (version 1.5.0) assemblers (18-20). We used the “--nano-hq” command line argument when generating the Flye assemblies. We run the rest of the assemblers using default settings. We used Any2fasta (version 0.4.2) to convert the gfa formatted output from Minipolish to a fasta file format (21). We used the generated assemblies to create groups of per replicon clusters using trycycler cluster (17). We then performed manual curation steps using trycycler reconcile, trycycler msa, trycycler partition and trycycler consensus following instructions on (<https://github.com/rrwick/Trycycler/wiki/How-to-run-Trycycler>) (17).

We polished the assemblies using medaka (1.8.1), Polypolish (version 0.5.0) and POLCA (MaSuRCA version 4.1.0, RRID:SCR\_010691) (22-24). The assembly was first polished with the long reads using medaka with “r941\_min\_sup\_g507” passed to the -m command line argument (22). This polished assembly was further polished using short read data using Polypolish and finally using POLCA. Both Polypolish and POLCA were run using default settings (23, 24).

**Genotyping, plasmid calling and AMR gene detection**

Bacterial genotyping and plasmid detection was done using the GenoTyphi bacterial typing framework (4, 25). We implemented Genotyphi using Mykrobe v0.12.1 with the default settings and the Typhi panel (version 20221208) (25, 26). The json output files were parsed into a csv document using the script “parse\_typhi\_mykrobe.py” provided at the GenoTyphi repository (26).

We then used AMRFinderPlus (Software version 3.11.26 and Database version 2023-11-15.1) to assess the presence of AMR genes and plasmids in the assembled contigs (27). AMRFinder Plus was run by passing the ” –nucleotide –protein –gff –organism –plus” flags. We passed “Salmonella” to the –organism flag. AMRFinderPlus outputs a file with AMR genes and their location on the contigs. The AMR genes are presented in [Figure 2](#Figure2). We used Ariba (version 2.14.6, RRID:SCR\_015976) using the ARG-ANNOT database to detect AMR genes in the raw reads to annotate the phylogenetic trees (27, 28).

**Bacteriophage detection**

To detect bacteriophages, we used PHASTER web tool (<https://phaster.ca/>) (29, 30). We loaded the polished assemblies into the web tool and exported the results as a summary.txt file containing the position of phages in the CT18 chromosome, shown in [Figure 1](#dkz39epkffzp) and [Figure 2A](#sevmy4av7h9r).

**Annotation**

We performed annotation of the hybrid assemblies using Prokka (version 1.14.6, RRID:SCR\_014732) with the default settings and performed a genus-specific annotation with “Salmonella” passed as an argument to the “-–genus” command line option (31).

**Sequence comparison**

We used the ACT: Artemis Comparison Tool (ACT Release 18.2.0, RRID:SCR\_004507) to perform a comparison of assemblies for the isolates sequenced on the MinION platform (32). We visualised the comparison in the BRIG: Blast Ring Image Generator (Version 0.95, RRID:SCR\_007802) (33). Both of these programs used the Basic Local Alignment Tool (BLAST) (Nucleotide-Nucleotide BLAST 2.15.0+) to perform nucleotide sequence comparisons (34). For all sequence comparisons, we used the “blastn” option with the default setting for comparing 2 sequences. We used the CT18 reference genome (AL513382.1) as a reference for the BRIG comparison of the chromosome and used the incHI1 plasmid from the same isolate (AL513383.1) as a reference for the plasmid comparison (35). To assess the presence of AMR genes on the plasmids and parts of the chromosome previously known to contain the insertion of AMR genes, we used the GenoPlotR v 0.8.11 R package which gives a visual comparison of the coding sequences in these regions (34).

**Constructing the phylogenetic tree**

We constructed a core snp maximum likelihood phylogenetic tree to put the newly assembled isolates in the context of other isolates from Malawi. The isolates used in the phylogenic analysis are from the paper by Gaud et al, where the assembled isolates were first described [(7)](https://web.endnote.%3D%3D). We used Snippy (version 4.6.0, RRID:SCR\_023572) to perform variant calling and IQ-TREE (RRID:SCR\_017254) to construct the phylogenetic tree. We used AL513382.1 as the reference genome for all variant calling steps. We used snippy-core to create a multiple sequence alignment file with all the variants from the previous step. Then we used snp-sites (version 2.5.1) with “-c” argument to extract sites containing exclusively ACGT from the alignment. We also used snp-sites with “-C” to calculate the number of the constant sites which we then used for the “-fconst” argument in IQ-TREE. We run iqtree on the alignment with ACGT only, with the command line parameters “-m GTR+G -bb 1000”.

**Results**

**Isolate selection**

We performed long-read whole genome sequencing of 4 isolates (ERS327391, ERS207185, ERS1509723, ERS1509734) to better understand unusual antimicrobial resistance patterns in 3 isolates from Malawi as observed previously (9). The years of isolation for the isolates were between 2010 and 2016 (9). Isolate ERS1509723 (BKQT8S) and ERS1509734 (BKQU3X) were selected because they had a distinct AMR pattern to the rest of the isolates from Malawi from the same period (9). These two isolates are two of only three isolates with incHI1 plasmids from the same study (9). Isolate ERS327391 (A58390) and ERS207185 (1017142) were selected because they represented isolates at the beginning of the *S*. Typhi outbreak in Blantyre, Malawi (7). Strain A58390 represents one of the first H58 isolates causing illnesses during the S. Typhi outbreal. Strain 1017142 represents non-H58 isolates from the same period (7).

**Genome analyses**

The nanopore reads for all samples had a mean Phred Quality score of 13.5 with a mean of 29K (21K - 34K) reads per isolate. We generated a mean 275Mb reads (158Mb-338Mb) with a mean N50 of 12K (9k - 13K). We generated hybrid assemblies of the four isolates, accession numbers and their metadata are presented in [Table 1](#7b38p493ij0e). Two isolates (A58390 and 1017142) assembled to 1 contig each, and the other two (BKQT8S and BKQU3X) assembled to 2 contigs each. All isolates had a 4.5Mb contig which represented the chromosome of the isolate. The two isolates with 2 contigs, each had a 185Kb contig that was identified as an IncHI1 plasmid.

The assembled genomes had a mean GC% of 52.0 ([Table 2](#Table2)). Three of the four isolates with long reads belong to the lineage 4.3.1 (H58 haplotype); the two isolates with IncHI1 plasmid were assigned sublineage 4.3.1.1EA1 (36); the remaining isolate belonging to the lineage 4.1.1 ([Table 2](#Table2)). A pairwise comparison of the assemblies from this study to the *Salmonella* Typhi reference strain CT18 (AL513382.1) shows a high level of sequence conservation as seen in [Figure 1](#dkz39epkffzp).

We assessed the known chromosomal insertion sites for AMR genes. The insertion sites at 3815480 and 1690327 of the CT18 strain were conserved across the isolates (supplementary Figure S1 and S2). The insertion site at position 3472059 of the CT18 strain was conserved in 1 isolate. The other three isolates had insertion of IS1 insertion sequences, [Figure](#Figure2) 2.

**Plasmids, AMR and bacteriophages**

The plasmids in the isolates BKQT8S and BKQU3X have an average GC content of 45.94% for each isolate. Both isolates had 3 plasmid replicons of the IncHI1 type (IncFIAHI1, IncHI1A and IncHI1BR27) [Table 2](#Table2). The plasmids typed as IncHI1 PST2 using pubMLST (37). A blastn comparison using Megablast shows high levels of similarity between the plasmids ([Figure 3A](#sevmy4av7h9r)). Comparing the plasmids to the pHCM1 plasmid isolated from the CT18 strain (AL513383.1) indicates a loss of coding sequences in an area with antimicrobial resistance determinants ([Figure 3A](#sevmy4av7h9r)). The two plasmid sets have a notable difference to pHCM1 in coding sequences in two regions. The first region is located between HCM1.149 and HCM1.175 of the pHCM1 plasmid. This region has a loss of  17 coding sequences in the plasmids from Malawi. Five of these coding sequences are associated with a mercury resistance operon and 1 AMR gene (dhrA14) has relocated to a different area of the Malawi plasmid. The second region is between HCM1.194 and HCM1.252. This region has been inverted in the Malawian plasmids. The region with the inversion has seen a loss of 13 coding sequences, reducing in size from ~40kb to ~20kb. Some of the lost genes include mercury resistance genes and the two AMR genes *catA1* and *aph(6)−Id*. These two AMR genes are responsible for resistance to chloramphenicol and aminoglycosides respectively. Our plasmid maintained aminoglycosides resistance genotype by maintaining a copy of *aph(3'')−Ib.* The plasmid in our collection has managed to maintain its resistance genotype while removing the metal resistance genes which are unnecessary for infection in humans.These changes in the structure of the plasmid could indicate adaptation of the plasmid in the host specific pathogen which could lead to a Typhi population which persistently carry the plasmid without relocating the resistance cassette to the chromosome (38). The rearrangements in the plasmid that relocated all the AMR genes to a single location, allow for centralized control of transcription of these genes which may lead to better control over the AMR phenotype for isolates carrying this particular plasmid. The AMR genes are highlighted in the ring showing the coding sequences of the reference plasmid ([Figure 3A](#sevmy4av7h9r)). [Figure 3B](#w7479xngefdd) further highlights the region in comparison to the pHCM1 plasmid.

Neither lineage 4.1.1 isolate, 1017142 nor the H58 isolate A58390 carried any acquired AMR genes, however there was a *gyrA\_S83F* point mutation in A58390 which is associated with reduced susceptibility to fluoroquinolones. This was the only isolate with mutations in the DNA gyrase. The other two isolates each carried acquired resistance genes (*blaTEM-1d*, *sul2*, *aph(3”)-lb*, *dfrA14* and t*et(B)*) on the plasmid which are known to cause resistance to beta-lactams, sulfonamides, aminoglycoside, trimethoprim and tetracyclines.

Running PHASTER on the assemblies reveals at least 4 intact bacteriophages, 1 questionable phage call and some incomplete bacteriophages per isolate Figure 4. Isolate BKQU3X and BKQT8S show similarity in the number of bacteriophage species per region and the completeness of the bacteriophages in these regions ([Figure](#pkd0dyujoi21) 4). Both isolates contain phages similar in composition and size (Figure 4B).

Phylogenetic analysis

A snp based phylogenetic tree reveals that the islates in tehe collection cluster into distinct clades based on their amr and olasmid prifiles. However a few isolates, four isolates from the lineage 4.3.1.1 and 3 isolates of the lineage 4.3.1.2 are located within the larger 4.3.1.1.EA1 lineage. This may suggest a presence of snps in thes isolates which are not included in the current genotype typing schemem making them to be incorectly typed as such.

Performing a phylogenetic analysis reveals that the H58 lineage 4.3.1.1.EA1 has diverse plasmid and AMR profiles. Most of the 4.3.1.1.EA1 isolates do not carry plasmids. Those which carry plasmids seem to be in two plasmid combinations, with isolates from Malawi carrying the plasmid replicon inc\_fiahi1 which is missing in the isolates from Kenya, while those from Kenya carry the plasmid replicon inc\_hi1\_st6 which is not available in the isolates from Malawi. Isolates with plasmids in the Kenyan collection also have additional AMR genes compared to the rest of the 4.3.1.1.EA1 lineage. Isolates which are not 4.3.1.1.EA1 are clustered in distinct clades with 1 clade per lineage. These isolates also have uniform AMR and plasmid profiles. There are distinct differences in acquired AMR genes between isolates of the lineage 4.3.1.1.EA1 and the other H58 lineages. The acquired AMR genes present in the lineage 4.3.1.1.EA1 are absent in the other H58 lineages and vice versa. This might mean that the mobile genetic elements driving resistance for the 4.3.1.1.EA1 are different from those in other H58 lineages.

**Discussion**

We noticed three incidences of predicted IncHI1 plasmids present in Malawi *S*. Typhi isolates amongst a set of 335 isolates, with different resistance gene repertoires. These three isolates carry *tetB* and *dfrA14* in addition to *sul2, strA,* and *strB* genes which are found in all other isolates carrying AMR genes, but lack the *cat1*, *sul1* and *tetD* genes usually seen on IncHI plasmids carried by *S*. Typhi. Initial analyses of the assemblies confirmed a single chromosome assembled for the two control isolates, whereas the two isolates with predicted IncHI replicons assembled as two molecules, the chromosome and one plasmid, as expected.

The two unusual isolates were further investigated given their reduced resistance gene repertoire compared to the main IncHI1 plasmid and we identified that these carry the PST2 IncHI1 plasmids, which is not commonly seen in the 4.3.1 haplotype (39). Two studies previously observed PST2 in non-4.3.1 haplotypes among travelers from West Africa to Europe during the same period as our isolates (38, 40), as did a collection from Tanzania. This IncHI variant, also with a reduced set of resistance genes and the plasmid replicons as encoded in our isolates (including a truncated IncFI replicon) were reported initially for the IncHI plasmid R27 in *S*. Typhi (cite), which is likely also the one reported for the isolates from Tanzania (cite).

Initial speculations on S. Typhi plasmid diversity considered whether this reduced plasmid might be in the course of being replaced. It is thus interesting to see that, whilst lacking the mercury operons and several resistance determinants, this reduced IncHI plasmid type is extant in parts of south-East Africa and embedded in the highly successful 3.4.1.EA1 lineage. This indicates that selection for the larger resistance plasmid might not be as strong as initially assumed (cite) or might even be disadvantageous in specific settings given its larger size.

Our isolates show *S*. Typhi encoding AMR genes either on plasmids or chromosomes isolated from the same region, and we thus investigated the known chromosomal insertion sites in the isolates with IncHI1 plasmids for scars or IS elements (8, 41). We found the insertion of coding sequences and transposases in isolates of the 4.3.1 haplotype. This might suggest the beginning of chromosomal integration of AMR genes previously on the plasmid or represent a scar of lost resistances in this region, although no isolates with acquired AMR genes both on plasmids and the chromosome were identified, which would represent a snapshot of this process in progress.

**Author contributions**

Conceptualization: EH, AMW, NAF; Data curation: EH, AZ, PMA; Formal analysis: AZ; Funding acquisition: NAF, EH; Investigation: AZ, AMW, EH; Methodology: AZ, EH, PMA; Project administration: EH, NAF; Resources: CA, NAF; Software: AZ, EH; Supervision: NAF, EH; Validation: AZ; Visualization: AZ; Writing-original draft: AZ; Writing-review and editing: AZ, NAF, EH. All authors read and approved the final manuscript.

**Conflict of interest**

The authors declare that there are no conflicts of interest.

**Funding information**

EH acknowledges funding from Wellcome (217303/Z/19/Z) and BBSRC (BB/V011278/1). NAF acknowledges support by the Bill & Melinda Gates Foundation (Investment OPP1128444) and the Wellcome Programme (grant 206454).

**Acknowledgments**

We want to thank the MLW core informatics support team for expert technical support.

# **Figure legends:**

**Table 1:** Metadata and accessions for the data presented in this study.

**Table** **2:** Antimicrobial resistance genes and plasmid replicons.

**Figure 1:** Pairwise comparison of the hybrid assemblies to the CT18 reference strain. The figure also shows regions containing bacteriophages.

**Figure 2**: Gene map showing the introduction of insertion sequences near *adenylate cyclase* gene of CT18 strain.

**Figure 3: A)** BLAST comparison of the plasmids from this study to the pHCM1 plasmid from the CT18 reference genome of *Salmonella* Typhi. From inside to outside ring: reference sequence pHCM1, plasmid of BKQT8S, plasmid BKQU3X, coding sequences in pHCM1, AMR genes in pHCM1, transposable elements in pHCM1 and bacteriophage region on pHCM1. **B)** BLAST comparison between plasmid of isolate BKQT8S and pHCM1 highlighting the location of IS elements and AMR genes. This region is an expansion of the variable region on Figure 2A. CDS, Coding Sequences; AMR, Antimicrobial Resistance.

**Figure 4**: Bacteriophage regions on the assemblies. **A)** Each column indicates the length of bacteriophage at the position in the CT18 genome colored by the completeness of the bacteriophage at the region**. B)** Each column indicates the length of bacteriophage at the position in the CT18 genome colored by the estimated species numbers.

Figure 5: A core-snp maximum likelihood phylogenetic tree of Isolates from Malawi put in a global context by including Isolates from Kenya from the same period. The tips are coloured by the genotype of the isolate according to the GenoTyphi typing scheme. The bars on the right are the year of isolation, and samples in this study, The next four bars show the presence of plasmid replicons in the samples, then the next four bars show quinolone point mutations and the rest of the bars are antimicrobial resistance genes in the isolates as identified by ARIBA.

[**Reference list**](https://web.endnote.com/reference-list/)

1. Stanaway JD, Reiner RC, Blacker BF, Goldberg EM, Khalil IA, Troeger CE, et al. The global burden of typhoid and paratyphoid fevers: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet Infectious Diseases. 2019;19(4):369-81.

2. Carey ME, Dyson ZA, Ingle DJ, Amir A, Aworh MK, Chattaway MA, et al. Global diversity and antimicrobial resistance of typhoid fever pathogens: Insights from a meta-analysis of 13,000 Salmonella Typhi genomes. eLife. 2023;12:e85867.

3. Park SE, Pham DT, Boinett C, Wong VK, Pak GD, Panzner U, et al. The phylogeography and incidence of multi-drug resistant typhoid fever in sub-Saharan Africa. Nature Communications. 2018;9(1):5094.

4. Wong VK, Baker S, Connor TR, Pickard D, Page AJ, Dave J, et al. An extended genotyping framework for Salmonella enterica serovar Typhi, the cause of human typhoid. Nature Communications 2016 7:1. 2016-10-05;7(1).

5. Dyson ZA, Holt KE. Five Years of GenoTyphi: Updates to the Global <i>Salmonella</i> Typhi Genotyping Framework. The Journal of Infectious Diseases. 2021;224(Supplement\_7):S775-S80.

6. Wong VK, Baker S, Pickard DJ, Parkhill J, Page AJ, Feasey NA, et al. Phylogeographical analysis of the dominant multidrug-resistant H58 clade of Salmonella Typhi identifies inter- and intracontinental transmission events. Nature Genetics. 2015;47(6):632-9.

7. Feasey NA, Gaskell K, Wong V, Msefula C, Selemani G, Kumwenda S, et al. Rapid Emergence of Multidrug Resistant, H58-Lineage Salmonella Typhi in Blantyre, Malawi. PLOS Neglected Tropical Diseases. 2015;9(4):e0003748.

8. Lima NCB, Tanmoy AM, Westeel E, De Almeida LGP, Rajoharison A, Islam M, et al. Analysis of isolates from Bangladesh highlights multiple ways to carry resistance genes in Salmonella Typhi. BMC Genomics. 2019;20(1).

9. Gauld JS, Olgemoeller F, Heinz E, Nkhata R, Bilima S, Wailan AM, et al. Spatial and Genomic Data to Characterize Endemic Typhoid Transmission. Clinical Infectious Diseases. 2022;74(11):1993-2000.

10. Phan M-D, Kidgell C, Nair S, Holt KE, Turner AK, Hinds J, et al. Variation in Salmonella enterica Serovar Typhi IncHI1 Plasmids during the Global Spread of Resistant Typhoid Fever. Antimicrobial Agents and Chemotherapy. 2009;53(2):716-27.

11. Parkhill J, Dougan G, James KD, Thomson NR, Pickard D, Wain J, et al. Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18. Nature. 2001;413(6858):848-52.

12. Chen S. Ultrafast one‐pass FASTQ data preprocessing, quality control, and deduplication using fastp. iMeta. 2023;2(2).

13. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884-i90.

14. Filtlong [Internet]. 2021. Available from: <https://github.com/rrwick/Filtlong>.

15. De Coster W, Rademakers R. NanoPack2: population-scale evaluation of long-read sequencing data. Bioinformatics. 2023;39(5).

16. Andrews S, Biggins L, Inglesfield S, Carr H, Montgomery J. FastQC. 0.12.0 ed2023. p. A quality control tool for high throughput sequence data.

17. Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, Vezina B, et al. Trycycler: consensus long-read assemblies for bacterial genomes. Genome Biology. 2021;22(1):266.

18. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nature Biotechnology. 2019;37(5):540-6.

19. Wick RR, Holt KE. Benchmarking of long-read assemblers for prokaryote whole genome sequencing. F1000Research. 2021;8:2138.

20. Vaser R, Šikić M. Raven: a de novo genome assembler for long reads. bioRxiv. 2021:2020.08.07.242461.

21. Seemann T. any2fasta: Convert various sequence formats to FASTA. Github2018.

22. Ltd. ONT. Medaka. Github2022.

23. Wick RR, Holt KE. Polypolish: Short-read polishing of long-read bacterial genome assemblies. PLOS Computational Biology. 2022;18(1):e1009802.

24. Zimin AV, Salzberg SL. The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. PLOS Computational Biology. 2020;16(6):e1007981.

25. Dyson ZA, Holt KE. Five Years of GenoTyphi: Updates to the Global Salmonella Typhi Genotyping Framework. The Journal of Infectious Diseases. 2021;224(Supplement\_7):S775-S80.

26. Holt K, Dyson ZA, Hawkey J, flashton2003. GenoTyphi. 2.0 ed. DOI 10.5281/zenodo.47076132022.

27. Feldgarden M, Brover V, Gonzalez-Escalona N, Frye JG, Haendiges J, Haft DH, et al. AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. Scientific Reports. 2021;11(1).

28. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, Rolain J-M. ARG-ANNOT, a New Bioinformatic Tool To Discover Antibiotic Resistance Genes in Bacterial Genomes. Antimicrobial Agents and Chemotherapy. 2014;58(1):212-20.

29. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 2016;44(W1):W16-21.

30. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: A Fast Phage Search Tool. Nucleic Acids Research. 2011;39(suppl):W347-W52.

31. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068-9.

32. Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. Bioinformatics. 2005;21(16):3422-3.

33. Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics. 2011;12(1):402.

34. Zhang Z, Schwartz S, Wagner L, Miller W. A Greedy Algorithm for Aligning DNA Sequences. Journal of Computational Biology. 2000;7(1-2):203-14.

35. Deng W, Liou S-R, Plunkett G, Mayhew GF, Rose DJ, Burland V, et al. Comparative Genomics of Salmonella enterica Serovar Typhi Strains Ty2 and CT18. Journal of Bacteriology. 2003;185(7):2330-7.

36. Kariuki S, Dyson ZA, Mbae C, Ngetich R, Kavai SM, Wairimu C, et al. Multiple introductions of multidrug-resistant typhoid associated with acute infection and asymptomatic carriage, Kenya. eLife. 2021;10:e67852.

37. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res. 2018;3:124.

38. Baltazar M, Ngandjio A, Holt KE, Lepillet E, Pardos de la Gandara M, Collard J-M, et al. Multidrug-Resistant Salmonella enterica Serotype Typhi, Gulf of Guinea Region, Africa. Emerging Infectious Disease journal. 2015;21(4):655.

39. Holt KE, Phan MD, Baker S, Duy PT, Nga TVT, Nair S, et al. Emergence of a Globally Dominant IncHI1 Plasmid Type Associated with Multiple Drug Resistant Typhoid. PLOS Neglected Tropical Diseases. 2011;5(7):e1245.

40. Ingle DJ, Nair S, Hartman H, Ashton PM, Dyson ZA, Day M, et al. Informal genomic surveillance of regional distribution of Salmonella Typhi genotypes and antimicrobial resistance via returning travellers. PLOS Neglected Tropical Diseases. 2019;13(9):e0007620.

41. Chiou C-S, Alam M, Kuo J-C, Liu Y-Y, Wang P-J. Chromosome-Mediated Multidrug Resistance in Salmonella enterica Serovar Typhi. Antimicrobial Agents and Chemotherapy. 2015;59(1):721-3.