**Supplementary Text. Detailed methodology applied the study.**

## Methods

### *Epidemiological information*

In Zambia, all medical and laboratory records of suspected typhoid cases are collected at the University Teaching Hospital (UTH) and the 41 peri-urban health centres in Lusaka and transferred to the Ministry of Health surveillance unit in the Department of Public Health and Research. Data associated with a typhoid fever outbreak covering a period from January 2010 to September 2012 were extracted for further analysis.

### *Samples and bacterial isolates*

Stool samples were subjected to pre-enrichment in selenite broth (0.023g / ml) (Oxoid Ltd, Hampshire, England) and incubated at 37˚C for 18-24 hours followed by plating onto selective media; MacConkey and Xylose Lysine Deoxycholate agar (Oxoid Ltd, Hampshire, England) and further incubated at 37˚C for 18-24 hours.

BACTEC blood culture system (BD Bactec, Sparks, USA) were applied for the blood samples. Positive samples were plated onto 5% sheep blood agar, chocolate agar, and MacConkey agar (Oxoid Ltd, Hampshire, United Kingdom) and incubated at 37˚C for 18-24 hours.

All presumptive positive suspected *S.* Typhi colonies were identified using biochemical tests followed by serogrouping using slide agglutination with hyperimmune polyvalent O antisera (Remel Europe Ltd, Kent, United Kingdom). Isolates belonging to serogroup O:9 were subsequently subjected to a *S.* Typhi specific PCR assay for confirmation and performed according to previously described publication (13).

A collection of 94 *S.* Typhi isolates were included the study and sent to the Technical University of Denmark, National Food Institute (DTU-Food) for further characterization.

***Antimicrobial susceptibility testing***

Minimum inhibitory concentration (MIC) determination was performed on all 94 *S.* Typhi isolates at DTU-Food, Denmark using a commercially prepared dehydrated panel (Sensititre; TREK Diagnostic Systems Ltd., East Grinstead,England) and according to previously published methodology (7). Furthermore, susceptibility to azithromycin was determined using agar dilution technique and E-tests according to CLSI guidelines and manufacturer (2,3). The tested antimicrobials / classes in this study have been listed in Table 2. Clinical and Laboratory Standards Institute (CLSI) (2) clinical breakpoints interpretative criteria for resistance (R) were used except for the following antimicrobials; azithromycin, ceftiofur, colistin, florfenicol, neomycin, spectinomycin, and streptomycin, respectively where epidemiological cut-off values according to EUCAST recommendations was used (http://www.eucast.org). Apramycin was interpreted according to research results from DTU-Food. Quality control was performed by using reference strain *E. coli* ATCC 25922 according to CLSI guidelines (2,3).

***Whole genome sequencing, multilocus sequence typing, antimicrobial resistance genes, and, plasmid replicons.***

Genomic DNA was extracted from the 33 isolates using an Invitrogen Easy-DNA™ Kit (Invitrogen, Carlsbad, CA, USA) and DNA concentrations were determined using the Qubit dsDNA BR assay kit (Invitrogen). The genomic DNA was prepared for Illumina pair-end sequencing using the Illumina (Illumina, Inc., San Diego, CA) NexteraXT® Guide 150319425031942 following the protocol revision C (http://support.illumina.com/downloads/nextera_xt_sample_preparation_guide_15031942.html). A sample of the pooled NexteraXT Libraries was loaded onto a Illumina MiSeq reagent

1 cartridge using MiSeq Reagent Kit v2 and 500 cycles with a Standard Flow Cell.The libraries
2 were sequenced using an Illumina platform and MiSeq Control Software 2.3.0.3. Twenty-
3 four isolates were pair-end and six isolates were single-end sequenced. Pair-end sequences
4 ranged in insert size from 11 to 129 with an average of 68. The read depth of the sequences
5 was between 147 to 497 with an average of 259.
6 Five previously published genomic sequences of haplotype H58; AG3, E02-2759, ISP-04-
7 06979, E03-9804, ISP-03-07467 were obtained from GenBank and Sanger Institute (accessed
8 5/4/2013). Full genomic information is shown in Supplementary Table 1A.
9
10 Raw sequence data have been submitted to the European Nucleotide Archive
11 (http://www.ebi.ac.uk/ena) under study accession no.: PRJEB7179 and PRJEB7182. The raw
12 reads were assembled using the Assemble pipeline (version 1.0) available from the Center for
13 Genomic Epidemiology (CGE) http://cge.cbs.dtu.dk/services/all.php which is based on the
14 Velvet algorithms for *de novo* short reads assembly. A complete list of genomic sequence
15 data is available in the Supplementary Table 1A. The assembled sequences were analyzed to
16 identify the MLST sequence type (ST) for *Salmonella enterica*, plasmid replicons, and
17 acquired antimicrobial resistance genes using the pipelines; MLST (version 1.7),
18 PlasmidFinder (version 1.2), and ResFinder (version 2.1) available from CGE (1,11,27).
19
20 ***Transferability of incQ1 plasmid replicon by conjugation and electroporation***
21 Plate-mating experiments were attempted with four *S.* Typhi isolates; # 31, #34, #54, and
22 #71harbouring the *inc*Q1plasmid replicon as donors and plasmid-free, rifampicin and
23 nalidixic acid resistant *E. coli* MT102RN as recipients (21). The strains were grown to both
24 late exponential as well as stationary phase, mixed (1:1) and incubated on solid blood agar at
25 both room temperature and at 37°C for 18 h.

1 The four *S.* Typhi isolates; # 31, #34, #54, and #71harbouring the *inc*Q1were subjected to
2 plasmid purification using Qiagen tip-100 and plasmid DNA using QIAGEN Plasmid Mini
3 Kit as described by the manufacture (Qiagen, Hilden, Germany). The isolates were attempted
4 electroporated into electrocompetent *E. coli* DH10B cells.
5
6 ***Identification of the chromosomally translocated region***
7 Due to an unsuccessful conjugation and electroporation, Genomic DNA was extracted from
8 the isolates # 31, #34, #54, and #71 using an Invitrogen Easy-DNA™ Kit. The genomic DNA
9 was prepared for Illumina mate pair sequencing using the Gel-Free version of the Illumina
10 Nextera® Mate Pair Sample Preparation Kit strictly following the protocol revision D and
11 sequenced using an Illumina platform.
12 Prior to assembly, the data quality was assessed using FastQC quality control tool
13 http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ and reads with quality score of
14 below 20 were filtered out. The genomes of the four strains were assembled using
15 SOAPdenovo2 software (17) by combining both paired-end and mate pair raw reads.
16 Subsequently, the two integration points of the chromosomally translocated region were
17 verified by PCR amplification and Sanger sequencing. All PCR amplifications were
18 performed using 0.1 µl taq polymerase and corresponding buffer (VWR International,
19 Radnor, PA, USA) in a final reaction volume of 50 mL. The following primers were used;
20 integration point 1: Chr_pla_forward_P5: 5'TCTGTTTTATATCCGCTGGG'3 /
21 Chr_pla_reverse_P6: 5'CTGAACGGTCTGGTTATAGG'3 and for integration point 2:
22 Pla_chr_forward_P5: 5'TTAGCGTGCTTTATTTTCCG'3 / Pla_chr_reverse_P6:
23 5'TCTACCCTATTGTTCCAGGT'3 as well as the cycling conditions for all reactions: 3
24 minutes at 94°C; 25 cycles of 1 minute at 94°C, 1 minute at the appropriate annealing
25 temperature 52°C, and 1minute at 72°C; 10 minutes at 72°C. Only amplicons for the four *S.*

Typhi isolates; # 31, #34, #54, and #71were selected for sequencing. Prior to sequencing, the amplicons were purified using the 20 U and 2 U of Exonuclease I (Exo I) and FastAP™ thermosensitive alkaline phosphatase, respectively (Fisher Scientific, Pittsburgh PA, USA).The DNA was shipped to Macrogen Inc., Amsterdam, The Netherlands for sequencing using the same primers as in the PCR analysis. The genomes of the four strains were re-assembled using CLC Bio Workbench by combining the Sanger sequences with the previously assembled scaffold.

Raw sequence data have been submitted to the European Nucleotide Archive (http://www.ebi.ac.uk/ena) under accession no. In progress.

Open Reading Frames (ORFs) were predicted on the scaffolds using Prodigal software (9) and were subsequently functionally annotated by constructing functional profiles for all proteins using the PanFunPro tool (16).

A functional profile is the combination of all non-repeating functional domains in each ORF. The profiles were created by using InterProScan software to scan the annotated proteins against the collections PfamA, TIGRFAM and Superfamily based on Hidden Markov Models (HMMs) to identify non-overlapping functional domains with an E-value below 0.001 (16). Through this annotation and analysis, the position of the *inc*Q1 replicon fragment was identified in every strain. The respective scaffolds containing this fragment were further compared to the complete genome and plasmid pHCM1 of our reference strain *S.* Typhi CT18 (National Center for Biotechnology Information, accession: AL513382, length of 4,809,037 bp) in order to determine the exact insertion site and homology between the strains.

***Screenings for mutations in DNA Gyrase and DNA topoisomerase IV genes***

1 Each genome was examined for mutation in *gyrA*, *gyrB*, *parC* and *parE* genes(23) by
2 determining SNP from the position of those genes in *S*. Typhi str. CT18 (18). Additionally,
3 the *gyrA* sequences of quinolone resistant strains (#269, #748) were extracted and translated.
4 DNA and protein sequences were compared to the sequences of *S*. Typhi str. CT18 by using
5 multiple alignments in CLC Bio Workbench to localize the QRDR nucleotide and the
6 consequent and amino acid changes.
7
8 ***Phylogenetic structure of S. Typhi using Single Nucleotide Polymorphisms, calculation of***
9 ***dN/dS, identification of S. Typhi haplotypes, and genomic deletions***
10 Single nucleotide polymorphisms (SNPs) were determined using the pipeline; SnpTree
11 version 1.1 available on the Center for Genomic Epidemiology
12 https://cge.cbs.dtu.dk/services/snpTree/ (12). Fundamentally, each of the assembled genomes
13 or contigs were aligned against the reference genome (*S*. Typhi str. CT18 (National Center
14 for Biotechnology Information, accession: AL513382, length of 4,809,037 bp)) using the
15 application "Nucmer" of MUMmer version 3.23 (4). SNPs were identified from the
16 alignments using "Show-snps" (using option "-Cl1rT") from MUMmer. Subsequently, SNPs
17 were selected when meeting the following criteria: 1) a minimum distance of 20 bps between
18 each SNP, and 2) all indels were excluded. The selected SNPs from assembled genomes were
19 confirmed by SNPs being called by mapping raw reads to the reference genome using
20 BWA(14) and SAMTools (15).
21 The qualified SNPs from each genome were concatenated to a single alignment
22 corresponding to position of the reference genome using an in-house Perl script. In case SNPs
23 were absent in the reference genome, they were interpreted as not being a variation and the
24 relatively base from the reference genome was expected (12). The concatenated sequences
25 were subjected to multiple alignments using MUSCLE from MEGA5(25). The final

phylogenetic SNP tree was computed by MEGA5 using the maximum likelihood method of 1,000 bootstrap replicates (5) using Tamura-Nei model for inference (24). All 415 SNPs related to the outbreak isolates of haplotype H58B are listed in the Supplementary Table 1B.

The non-synonymous SNP/synonymous SNP ratio (dN/dS) is a measurement of stabilizing selection (8). A ratio of 1 is expected in the absence of selection, a low ratio (dN/dS<1) indicates stabilizing selection, while a high ratio (dN/dS>1) indicates positive selection (19). The dN/dS ratio, was calculated for each core gene using codeML from the package PAML (26). The approximation of the dN/dS ratio was an average of dN/dS from all core genes.

The haplotypes from each genome were assigned based on biallelic polymorphisms positions (BiP) previously described (20). Positions based on WGST are listed in the Supplementary Table 1C. Additionally, node B of haplotype H58 was determined from SNP position 1,193,220 (10).

A BLAST atlas based on BLASTP (6) was used to predict putative deletions in a comparison of the genomes against the reference genome; CT18. The putative deletions were aligned against Zambian genomes using execrate (22). The hit score was calculated by multiplying percent identify with deletion’s alignment length and dividing with deletion’s sequence length. The presence of deletions in the Zambian genomes was confirmed based on the hit score with a threshold of at least 95%. The presence and absence of the deletions were finally visualized in a heatmap. Details of the genomic deletions detected in this study are listed in the Supplementary Table 1D.

## Reference List

1. **Carattoli, A., E. Zankari, A. Garcia-Fernandez, L. M. Volby, O. Lund, L. Villa, F. M. Aarestrup, and H. Hasman**. 2014. PlasmidFinder and pMLST: in silico detection and typing of plasmids. Antimicrob.Agents Chemother.

2. Clinical and Laboratory Standards Institute. Performance Standards for Antimicrobial Susceptibility Testing. **M100-S24**. 24th Informational Supplement. 2014. Wayne, PA, USA.

3. Clinical and Laboratory Standards Institute. Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically. **M07-A9**. 9th Edition[Approved Standard]. 2012. Wayne, PA, USA.

4. **Delcher, A. L., A. Phillippy, J. Carlton, and S. L. Salzberg**. 2002. Fast algorithms for large-scale genome alignment and comparison. Nucleic Acids Res. **30**:2478-2483.

5. **Felsenstein J**. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution **39**:783-791.

6. **Hallin, P. F., T. T. Binnewies, and D. W. Ussery**. 2008. The genome BLASTatlas-a GeneWiz extension for visualization of whole-genome homology. Mol.Biosyst. **4**:363-371.

7. **Hendriksen, R. S., H. S. Le, V. Bortolaia, C. Pulsrikarn, E. M. Nielsen, S. Pornruangmong, P. Chaichana, C. A. Svendsen, F. X. Weill, and F. M. Aarestrup**. 2012. Characterization of isolates of Salmonella enterica serovar Stanley, a serovar endemic to Asia and associated with travel. J.Clin.Microbiol. **50(3)**:709-720.

8. **Holt, K. E., J. Parkhill, C. J. Mazzoni, P. Roumagnac, F. X. Weill, I. Goodhead, R. Rance, S. Baker, D. J. Maskell, J. Wain, C. Dolecek, M. Achtman, and G. Dougan**. 2008. High-throughput sequencing provides insights into genome variation and evolution in Salmonella Typhi. Nat.Genet. **40**:987-993.

9. **Hyatt, D., G. L. Chen, P. F. Locascio, M. L. Land, F. W. Larimer, and L. J. Hauser**. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC.Bioinformatics. **11:119.**:119.

10. **Kariuki, S., G. Revathi, J. Kiiru, D. M. Mengo, J. Mwituria, J. Muyodi, A. Munyalo, Y. Y. Teo, K. E. Holt, R. A. Kingsley, and G. Dougan**. 2010. Typhoid in Kenya is associated with a dominant multidrug-resistant Salmonella enterica serovar Typhi haplotype that is also widespread in Southeast Asia. J.Clin.Microbiol. **48**:2171-2176.

11. **Larsen, M. V., S. Cosentino, S. Rasmussen, C. Friis, H. Hasman, R. L. Marvig, L. Jelsbak, T. Sicheritz-Ponten, D. W. Ussery, F. M. Aarestrup, and O. Lund**. 2012. Multilocus sequence typing of total-genome-sequenced bacteria. J.Clin.Microbiol. **50(4)**:1355-1361.

1 12. **Leekitcharoenphon, P., R. S. Kaas, M. C. Thomsen, C. Friis, S. Rasmussen, and F.**
2 **M. Aarestrup**. 2012. snpTree--a web-server to identify and construct SNP trees from
3 whole genome sequence data. BMC.Genomics. **13 Suppl 7:S6. doi: 10.1186/1471-**
4 **2164-13-S7-S6. Epub;%2012 Dec 13.**:S6-13.

5 13. **Levy, H., S. Diallo, S. M. Tennant, S. Livio, S. O. Sow, M. Tapia, P. I. Fields, M.**
6 **Mikoleit, B. Tamboura, K. L. Kotloff, R. Lagos, J. P. Nataro, J. E. Galen, and M.**
7 **M. Levine**. 2008. PCR method to identify Salmonella enterica serovars Typhi,
8 Paratyphi A, and Paratyphi B among Salmonella Isolates from the blood of patients
9 with clinical enteric fever. J.Clin.Microbiol. **46**:1861-1866.

10 14. **Li, H. and R. Durbin**. 2009. Fast and accurate short read alignment with Burrows-
11 Wheeler transform. Bioinformatics. **25**:1754-1760.

12 15. **Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G.**
13 **Abecasis, and R. Durbin**. 2009. The Sequence Alignment/Map format and SAMtools.
14 Bioinformatics. **25**:2078-2079.

15 16. **Lukjancenko O, Thomsen, M. C., Volby, Larsen M., and Ussery, D. W**. PanFunPro:
16 PAN-genome analysis based on FUNctional PROfiles. [2], 1-13. 2014. F1000Research.

17 17. **Luo, R., B. Liu, Y. Xie, Z. Li, W. Huang, J. Yuan, G. He, Y. Chen, Q. Pan, Y. Liu,**
18 **J. Tang, G. Wu, H. Zhang, Y. Shi, Y. Liu, C. Yu, B. Wang, Y. Lu, C. Han, D. W.**
19 **Cheung, S. M. Yiu, S. Peng, Z. Xiaoqian, G. Liu, X. Liao, Y. Li, H. Yang, J. Wang,**
20 **T. W. Lam, and J. Wang**. 2012. SOAPdenovo2: an empirically improved memory-
21 efficient short-read de novo assembler. Gigascience. **1**:18-1.

22 18. **Parkhill, J., G. Dougan, K. D. James, N. R. Thomson, D. Pickard, J. Wain, C.**
23 **Churcher, K. L. Mungall, S. D. Bentley, M. T. Holden, M. Sebaihia, S. Baker, D.**
24 **Basham, K. Brooks, T. Chillingworth, P. Connerton, A. Cronin, P. Davis, R. M.**
25 **Davies, L. Dowd, N. White, J. Farrar, T. Feltwell, N. Hamlin, A. Haque, T. T.**
26 **Hien, S. Holroyd, K. Jagels, A. Krogh, T. S. Larsen, S. Leather, S. Moule, P.**
27 **O'Gaora, C. Parry, M. Quail, K. Rutherford, M. Simmonds, J. Skelton, K.**
28 **Stevens, S. Whitehead, and B. G. Barrell**. 2001. Complete genome sequence of a
29 multiple drug resistant Salmonella enterica serovar Typhi CT18. Nature. **413**:848-852.

30 19. **Rocha, E. P., J. M. Smith, L. D. Hurst, M. T. Holden, J. E. Cooper, N. H. Smith,**
31 **and E. J. Feil**. 2006. Comparisons of dN/dS are time dependent for closely related
32 bacterial genomes. J.Theor.Biol. **239**:226-235.

33 20. **Roumagnac, P., F. X. Weill, C. Dolecek, S. Baker, S. Brisse, N. T. Chinh, T. A. Le,**
34 **C. J. Acosta, J. Farrar, G. Dougan, and M. Achtman**. 2006. Evolutionary history of
35 Salmonella typhi. Science. **314**:1301-1304.

36 21. **Sirichote, P., H. Hasman, C. Pulsrikarn, H. C. Schonheyder, J. Samulioniene, S.**
37 **Pornruangmong, A. Bangtrakulnonth, F. M. Aarestrup, and R. S. Hendriksen**.
38 2010. Molecular characterization of extended-spectrum cephalosporinase-producing
39 Salmonella enterica serovar Choleraesuis isolates from patients in Thailand and
40 Denmark. J.Clin.Microbiol. **48(3)**:883-888.

22. **Slater, G. S. and E. Birney**. 2005. Automated generation of heuristics for biological sequence comparison. BMC.Bioinformatics. **6:31.**:31.

23. **Song, Y., P. Roumagnac, F. X. Weill, J. Wain, C. Dolecek, C. J. Mazzoni, K. E. Holt, and M. Achtman**. 2010. A multiplex single nucleotide polymorphism typing assay for detecting mutations that result in decreased fluoroquinolone susceptibility in Salmonella enterica serovars Typhi and Paratyphi A. J.Antimicrob.Chemother. **65**:1631-1641.

24. **Tamura, K. and M. Nei**. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol.Biol.Evol. **10**:512-526.

25. **Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar**. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol.Biol.Evol. **28**:2731-2739.

26. **Yang, Z.** 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol.Biol.Evol. **24**:1586-1591.

27. **Zankari, E., H. Hasman, S. Cosentino, M. Vestergaard, S. Rasmussen, O. Lund, F. M. Aarestrup, and M. V. Larsen**. 2012. Identification of acquired antimicrobial resistance genes. J.Antimicrob.Chemother. **67**:2640-2644.