

RESEARCH ARTICLE

Whole genome sequence analysis of *Salmonella* Typhi in Papua New Guinea reveals an established population of genotype 2.1.7 sensitive to antimicrobials

Zoe Anne Dyson^{1,2,3,4*}, Elisheba Malau⁵, Paul F. Horwood⁶, Rebecca Ford⁷, Valentine Siba⁷, Mitton Yoannes⁷, William Pomat⁷, Megan Passey⁸, Louise M. Judd², Danielle J. Ingle⁹, Deborah A. Williamson^{9,10}, Gordon Dougan³, Andrew R. Greenhill^{5‡*}, Kathryn E. Holt^{1,2‡}

1 London School of Hygiene & Tropical Medicine, London, United Kingdom, **2** Department of Infectious Diseases, Central Clinical School, Monash University, Melbourne, Australia, **3** Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Department of Medicine, University of Cambridge, Cambridge, United Kingdom, **4** Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, United Kingdom, **5** School of Science, Psychology and Sport, Federation University, Churchill, Australia, **6** College of Public Health, Medical and Veterinary Sciences, James Cook University, Townsville, Australia, **7** Papua New Guinea Institute of Medical Research, Goroka, Papua New Guinea, **8** University Centre for Rural Health, Faculty of Medicine and Health, University of Sydney, Lismore, Australia, **9** Department of Microbiology and Immunology, The University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Melbourne, Australia, **10** Microbiological Diagnostic Unit Public Health Laboratory, Department of Microbiology and Immunology, The University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Melbourne, Australia

‡ These authors contributed equally as senior authors.

* zoe.dyson@lshtm.ac.uk (ZAD); andrew.greenhill@federation.edu.au (ARG)



OPEN ACCESS

Citation: Dyson ZA, Malau E, Horwood PF, Ford R, Siba V, Yoannes M, et al. (2022) Whole genome sequence analysis of *Salmonella* Typhi in Papua New Guinea reveals an established population of genotype 2.1.7 sensitive to antimicrobials. *PLoS Negl Trop Dis* 16(3): e0010306. <https://doi.org/10.1371/journal.pntd.0010306>

Editor: Abiola Senok, Mohammed Bin Rashid University of Medicine and Health Sciences, UNITED ARAB EMIRATES

Received: September 26, 2021

Accepted: March 5, 2022

Published: March 28, 2022

Copyright: © 2022 Dyson et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Genomic data are submitted to the European Nucleotide Archive (ENA) and accession numbers are available in the supplementary tables.

Funding: KEH was supported by a grant funded by the Wellcome Trust (STRATAA; 106158/Z/14/Z) to employ ZAD. ZAD received funding from the European Union's Horizon 2020 research and innovation programme under the Marie

Abstract

Background

Typhoid fever, a systemic infection caused by *Salmonella enterica* serovar Typhi, remains a considerable public health threat in impoverished regions within many low- and middle-income settings. However, we still lack a detailed understanding of the emergence, population structure, molecular mechanisms of antimicrobial resistance (AMR), and transmission dynamics of *S. Typhi* across many settings, particularly throughout the Asia-Pacific islands. Here we present a comprehensive whole genome sequence (WGS) based overview of *S. Typhi* populations circulating in Papua New Guinea (PNG) over 30 years.

Principle findings

Bioinformatic analysis of 86 *S. Typhi* isolates collected between 1980–2010 demonstrated that the population structure of PNG is dominated by a single genotype (2.1.7) that appears to have emerged in the Indonesian archipelago in the mid-twentieth century with minimal evidence of inter-country transmission. Genotypic and phenotypic data demonstrated that the PNG *S. Typhi* population appears to be susceptible to former first line drugs for treating typhoid fever (chloramphenicol, ampicillin and co-trimoxazole), as well as fluoroquinolones,

Skłodowska-Curie grant agreement TyphiNET (#845681). DJI was supported by a National Health and Medical Research Council (NHMRC) Emerging Leadership Fellowship (GNT1195210). DAW is supported by an NHMRC Emerging Leadership Fellowship (GNT1174555). KEH was supported by a Senior Medical Research Fellowship from the Viertel Foundation of Australia, and the Bill and Melinda Gates Foundation, Seattle (grant #OPP1175797). This work was supported, in whole or in part, by the Bill & Melinda Gates Foundation [OPP1175797]. Under the grant conditions of the Foundation, a Creative Commons Attribution 4.0 Generic License has already been assigned to the Author Accepted Manuscript version that might arise from this submission. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

third generation cephalosporins, and macrolides. PNG genotype 2.1.7 was genetically conserved, with very few deletions, and no evidence of plasmid or prophage acquisition. Genetic variation among this population was attributed to either single point mutations, or homologous recombination adjacent to repetitive ribosomal RNA operons.

Significance

Antimicrobials remain an effective option for the treatment of typhoid fever in PNG, along with other intervention strategies including improvements to water, sanitation and hygiene (WaSH) related infrastructure and potentially the introduction of Vi-conjugate vaccines. However, continued genomic surveillance is warranted to monitor for the emergence of AMR within local populations, or the introduction of AMR associated genotypes of *S. Typhi* in this setting.

Author summary

Typhoid fever, caused by *Salmonella enterica* serovar Typhi, is a systemic infection common to many low- to middle-income settings. While the population structure of *S. Typhi* has been genetically characterised using whole genome sequencing in many endemic countries throughout Sub-Saharan Africa and South Asia, we are still lacking a detailed understanding for many regions including those among the Asia-Pacific islands. Genomic surveillance of isolates spanning 30 years demonstrated a population structure of *S. Typhi* in Papua New Guinea (PNG) dominated by a single genotype (2.1.7) that emerged in the mid-twentieth century, is genetically homogeneous, and sensitive to a wide range of antibiotics commonly used in the treatment of typhoid. There was little evidence of inter-country transmission and the setting appeared free of *S. Typhi* genotypes commonly associated with AMR e.g. H58 (genotype 4.3.1). These data suggest that former first line drugs (chloramphenicol, ampicillin and co-trimoxazole), fluoroquinolones, third generation cephalosporins and macrolides all remain viable options for controlling typhoid in addition to the introduction of Vi-conjugate vaccines and improvements to water, sanitation and hygiene (WaSH) related infrastructure. Routine molecular surveillance is necessary to monitor for introduced or emerging AMR to inform treatment guidelines and intervention strategies.

Introduction

Typhoid fever is a systemic infection of the bacterium *Salmonella enterica* serovar Typhi[1]. Each year more than 10,000,000 cases of typhoid fever occur worldwide, of which more than 100,000 result in death[2]. *S. Typhi* is transmitted faeco-orally, usually through contaminated food and water and therefore constitutes a major public health threat in many low- to middle-income countries (LMIC), particularly in impoverished regions where hygiene and sanitation infrastructure is limited[3].

Antimicrobial chemotherapy has become a mainstay in the treatment of typhoid fever, with death rates declining from 20–30% to 1% following the introduction of antimicrobials[4,5]. However, antimicrobial resistance (AMR) has arisen multiple times in *S. Typhi*, with multi-drug-resistance (MDR; resistance to former first line antimicrobials chloramphenicol, co-trimoxazole, and ampicillin) emerging from the 1970s onwards[6], followed by resistance to

fluoroquinolones commonly mediated by mutations in the quinolone resistance determining region (QRDR) of genes *gyrA*, *gyrB*, and *parC*[7–9]. More recently, extensively drug-resistant (XDR) *S. Typhi* populations resistant to all orally administered antimicrobials except for azithromycin have emerged in Pakistan outbreaks[10,11], and mutations mediating azithromycin resistance have been observed in non-XDR *S. Typhi* populations[12–15].

Papua New Guinea (PNG) has been recognised as a high burden setting for typhoid since the mid-1990s when the incidence rate was reported as 1,208 cases per 100,000 population, amongst the highest rates in the world at that time[16]. Although there is a paucity of recent data, one diagnostic study demonstrated that typhoid remains a common diagnosis in febrile patients in the PNG highlands[17]. Vaccines for typhoid fever are not currently part of the routine schedule in PNG, nor have they been in the past.

Early molecular studies [18–20] revealed that several ribotypes of *S. Typhi* appeared endemic to PNG, with many appearing unique to the country, and some also observed in Southeast Asian countries. Further, 80.6% of isolates from 1986–1989 belonged to a single phage type (D2) [18,20], which was also observed at lower frequency in other countries including Malaysia and Thailand. Previous reports[18], have also noted that *S. Typhi* from PNG appeared free of both detectable AMR genes and plasmids, and have suggested that genome plasticity might be common. However, despite these early observations, only two locally collected *S. Typhi* sequences from PNG have been reported on in detail[21]. A further 45 isolates collected from travellers returning to Australia from PNG were sequenced as a part of a global overview study[9], but this study did not report on the specific dynamics and characteristics of the PNG pathogen population. Subsequently, there is currently no comprehensive whole genome sequence-based overview of the population structure, transmission dynamics, or evolutionary history of *S. Typhi* in PNG. Here we combine data on 41 novel isolates collected locally in PNG between 1992 and 2010, with those data sequenced previously from return travellers, to provide a 30-year overview (1980–2010) of *S. Typhi* in PNG.

Methods

Ethics statement

Ethical approval for this study was granted by the PNG Institute of Medical Research Institutional Review Board (1609), the PNG Medical Research Advisory Board (MRAC 16.43) and the Federation University Human Research Ethics Committee (A17-074).

Study setting

PNG is a tropical country situated approximately 6° south of the equator in the Western Pacific region. It is a LMIC, with a large proportion of its ~9 million inhabitants living in regional and remote areas of the country. Health and development indicators are low by regional and global standards, with ~40% of people having access to improved water supply and <20% with access to adequate sanitation[22,23]. In 2019, PNG's human development index of 0.55 ranked it 155 out of 189 countries and territories[24].

The PNG Institute of Medical Research (PNGIMR) is headquartered in Goroka, the capital of Eastern Highlands Province. The institute has been responsible for the majority of studies that have been conducted on typhoid fever in PNG.

Isolate collection and culture processing

Isolates of *S. Typhi* that were sequenced as part of this study (n = 41) were obtained from the PNGIMR culture collection. All isolates had been preserved by freeze-drying in glass ampules.

The content of each ampule was resuspended in nutrient broth, and inoculated onto nutrient agar (incubated for 18–24 hours at 35°C). PCR was conducted on isolates using previously described methods [17] to confirm their identity as *S. Typhi*.

Antimicrobial susceptibility testing and phage typing

Antimicrobial susceptibility testing (AST) was carried out for $n = 40$ (88.9%) of travel-associated cases for chloramphenicol, ampicillin, sulfamethoxazole, trimethoprim, nalidixic acid, tetracycline, kanamycin, gentamycin, and spectinomycin as detailed in **S1 Table**. AST was performed on *S. Typhi* isolates using agar breakpoint dilution and interpreted using Clinical and Laboratory Standards Institute (CLSI) breakpoints as described previously [25]. Phage types were extracted from the National Enteric Pathogens Surveillance Scheme (NEPSS), a dedicated surveillance system for human and nonhuman enteric pathogens (including *Salmonella*) that has been operated by the Microbiological Diagnostic Unit Public Health Laboratory (MDU PHL) [25].

DNA extraction and whole genome sequencing

Isolated colonies (1–3) were suspended in sterile water to 0.5 McFarland, and 100 μ l of the suspension was used for DNA extraction with the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany), following the manufacturer's instructions. Extracted genomic DNA (gDNA) was then subjected to indexed whole genome sequencing on an Illumina HiSeq 2500 platform at the Wellcome Sanger Institute to generate paired end reads of 150 bp in length as described previously [9].

Three isolates were selected for additional long-read sequencing using the Nanopore MinION R9 device which was carried out as described previously [26]. To generate high molecular weight gDNA (>60 kbp) suited to Oxford Nanopore sequencing without further size selection the following protocol was used; firstly, isolates were grown overnight at 37°C on Luria-Bertani (LB) plates, before single colonies were picked for overnight culture at 37°C in LB broth medium. Bacterial cell pellets from 3.0 ml LB culture were then generated by centrifugation at 15,000 g for 5 minutes. gDNA was extracted from these pellets using Agencourt GenFind V2 (Beckman Coulter) with modifications as follows. Cell pellets were resuspended in 400 μ l lysis buffer containing 9 μ l Proteinase K (96 mg/ml; Beckman Coulter) and 1 μ l RNase A (100 mg/ml; Sigma Aldrich) by tip mixing. Samples were lysed at 37°C for 30 minutes, and gDNA was extracted from the lysed samples by completing the remaining steps of the GenFind V2 for 200 μ l of blood/serum from the binding step onwards as per the manufacturer's instructions.

Read mapping, Single Nucleotide Variant (SNV) and phylogenomic analyses

For Single Nucleotide Variant (SNV) analysis, paired end reads for $n = 41$ *S. Typhi* isolates collected in this study together with $n = 45$ from a previous study [27] (**S1 Table**), were aligned to the CT18 reference genome (GenBank accession AL513382) [28] using the RedDog mapping pipeline (V1.beta10.3; available at <https://github.com/katholt/reddog>). RedDog maps reads to the reference genome with Bowtie2 (v2.2.9) [29], before using SAMtools (v1.3.1) [30] to identify high quality single nucleotide variant (SNV) calls as previously described [8]. A single sequence (sample MDUST355) was found to be a mixture of *S. Typhi* genotypes (~80% 2.1.7.1 and ~20% 2.1.7.2) and was subsequently removed from further phylogenomic analyses. A core SNV alignment was generated by concatenating alleles with high-quality consensus base calls (phred scores >20), for all SNV sites meeting this criteria in >95% of genomes (representing the 95% 'soft' core of the *S. Typhi* genome). This alignment was filtered to exclude SNV called

in previously defined [9,31,32] prophage regions and repetitive sequences (354 kb; ~7.4% of bases) in the CT18 reference chromosome, and Paratyphi A sequence AKU_12601 (GenBank accession FM200053) alleles were also included in the alignment for outgroup rooting of the phylogenetic tree. The resultant 657 bp SNV alignment then was used to infer a whole genome pseudoalignment using the CT18 reference sequence, and any remaining recombination was removed using Gubbins (v2.3.2)[33]. The resultant 620 bp recombination-filtered alignment was used as input for RAXML (v8.2.8)[34], which was used to infer a maximum likelihood (ML) phylogeny with a generalized time-reversible model and a Gamma distribution to model site-specific rate variation (GTR + Γ substitution model; GTRGAMMA in RAXML), with 100 bootstrap pseudo-replicates used to assess branch support (S1 Fig).

GenoTyphi (v1.9.1; available at <https://github.com/katholt/genotypyphi/>) was used together with RedDog Bowtie2 alignment (BAM) files to infer genotypes for all sequences analysed in this study using SNVs that were defined in the extended *S. Typhi* genotyping framework [27,35]. These data revealed that our sampling of the PNG *S. Typhi* population was dominated by a single genotype (2.1.7, $n = 84$, 98%; S1 Fig), necessitating the completion of a reference genome for the PNG genotype 2.1.7 population to allow for high resolution phylogenomic analyses. Three PNG genotype 2.1.7 genomes were selected for Nanopore long read sequencing and hybrid assembly (see S1 Table), with the oldest of these (MDUST348 from 1985), selected as a reference sequence for lineage-focused SNV analyses and phylogenetic inference using the mapping-based approach described above. For this newly completed PNG genotype 2.1.7 reference genome (MDUST348, deposited under accession GCA_001173525), repetitive regions were identified using the `ncmer` command in MUMmer (v3.23) [36] and integrated prophage sequences detected using the PHAge Search Tool Enhanced Release (PHASTER; available at <https://phaster.ca/>)[37]. These regions (detailed in S2 Table) were then filtered from the inferred PNG genotype 2.1.7 alignment, with any further recombination removed with Gubbins as described above. The resultant recombination filtered alignment of PNG genotype 2.1.7 *S. Typhi* sequences ($n = 83$) also included alleles from representative non-genotype 2.1.7 sequences as outgroups for phylogenetic tree rooting (detailed in S3 Table), and this final alignment (258 bp; 143 sequences) was then used to infer the final phylogenies shown in Figs 1 and S2. A global context phylogeny including an additional $n = 50$ sequences of clade 2.1 ($n = 133$ total; S1 Table) from previous studies[27,38] was inferred in the same manner with an alignment of 671 bp in length (Fig 2; 184 sequences including outgroups). An interactive version of the resulting global context clade 2.1 phylogeny is available at: https://microreact.org/project/icGPbF_O8.

Unique SNVs defining PNG genotype 2.1.7 *S. Typhi* sublineages were manually extracted from RedDog allele tables using R. SNVs causing non-synonymous mutations in highly conserved genes were prioritized for sublineage definitions. SNV distances were calculated from the core genome SNV alignment using `snp-dists` (v0.7.0; available at <https://github.com/tseemann/snp-dists>).

Assessments of the gene conservation across PNG genotype 2.1.7 *S. Typhi* (S4 Fig) were derived from the gene presence/absence matrix output from the RedDog mapping pipeline using the MDUST348 PNG 2.1.7 *S. Typhi* reference genome completed in this study.

Temporal analysis

To investigate the temporal signal and emergence of genotype 2.1.7 *S. Typhi* in PNG, we used multiple methods. Initially we used TempEst (v1.5.1) [39] to assess the temporal structure of the data (i.e. signals of clock-like behaviour) by conducting a regression analysis of root-to-tip branch distances for the PNG genotype 2.1.7 ML tree inferred using the MDUST348 reference

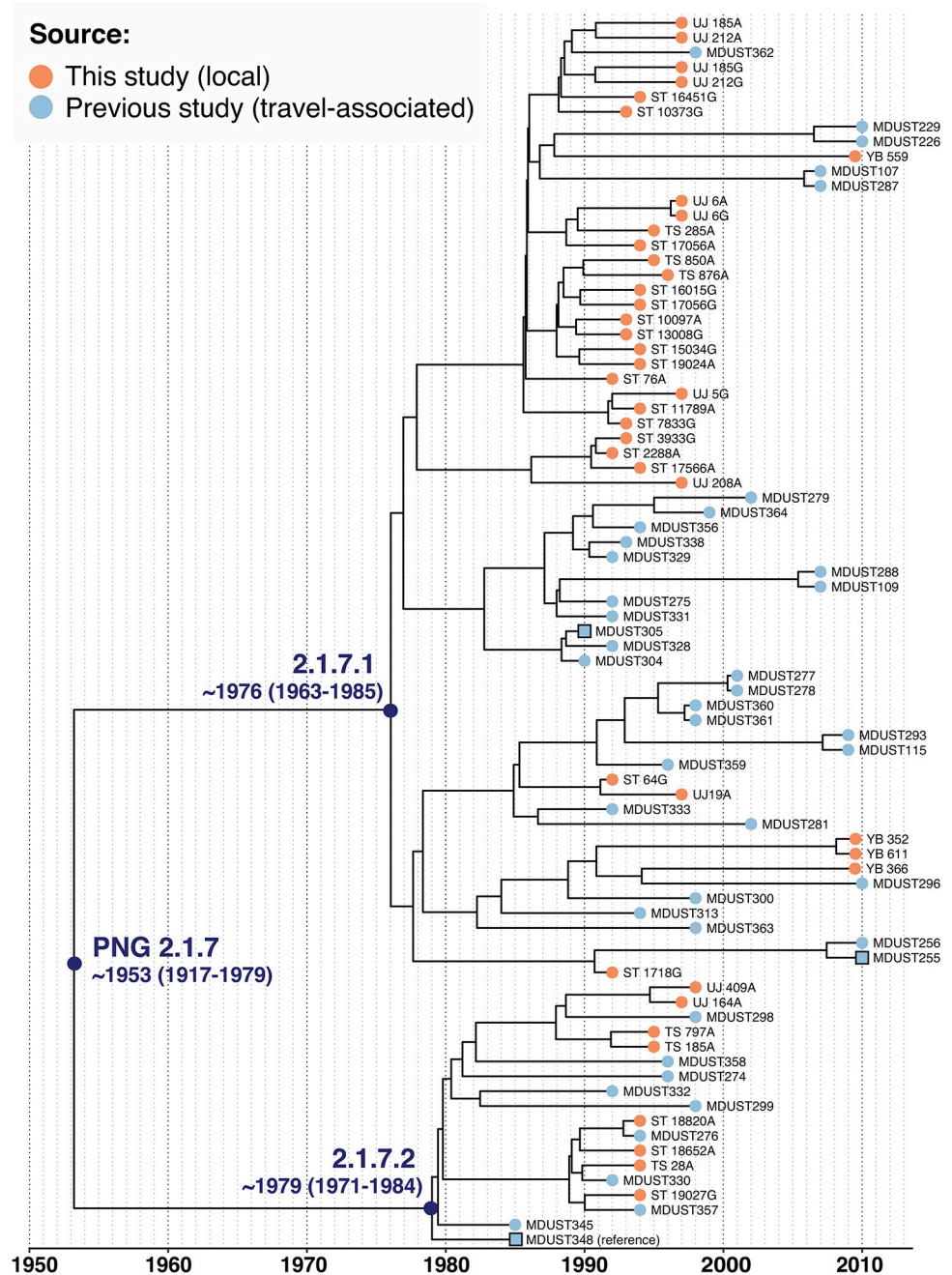


Fig 1. Maximum clade credibility tree inferred from all available (n = 83) PNG 2.1.7 *S. Typhi* using BEAST2. Tip colours indicate isolate source as per inset legend, branches are labelled by *S. Typhi* genotype. Square nodes indicate the reference sequence and other completed genomes. Divergence dates for PNG genotype 2.1.7 and sublineages are labelled (95% highest posterior density in brackets). Interactive phylogeny available at: <https://microreact.org/project/CYcTnSmjT>.

<https://doi.org/10.1371/journal.pntd.0010306.g001>

genome (S2 Fig) as a function of sampling time. The heuristic residual mean squared method was used for the regression analysis with the best-fitting root option selected. The resultant data from these analyses were visualised in R.

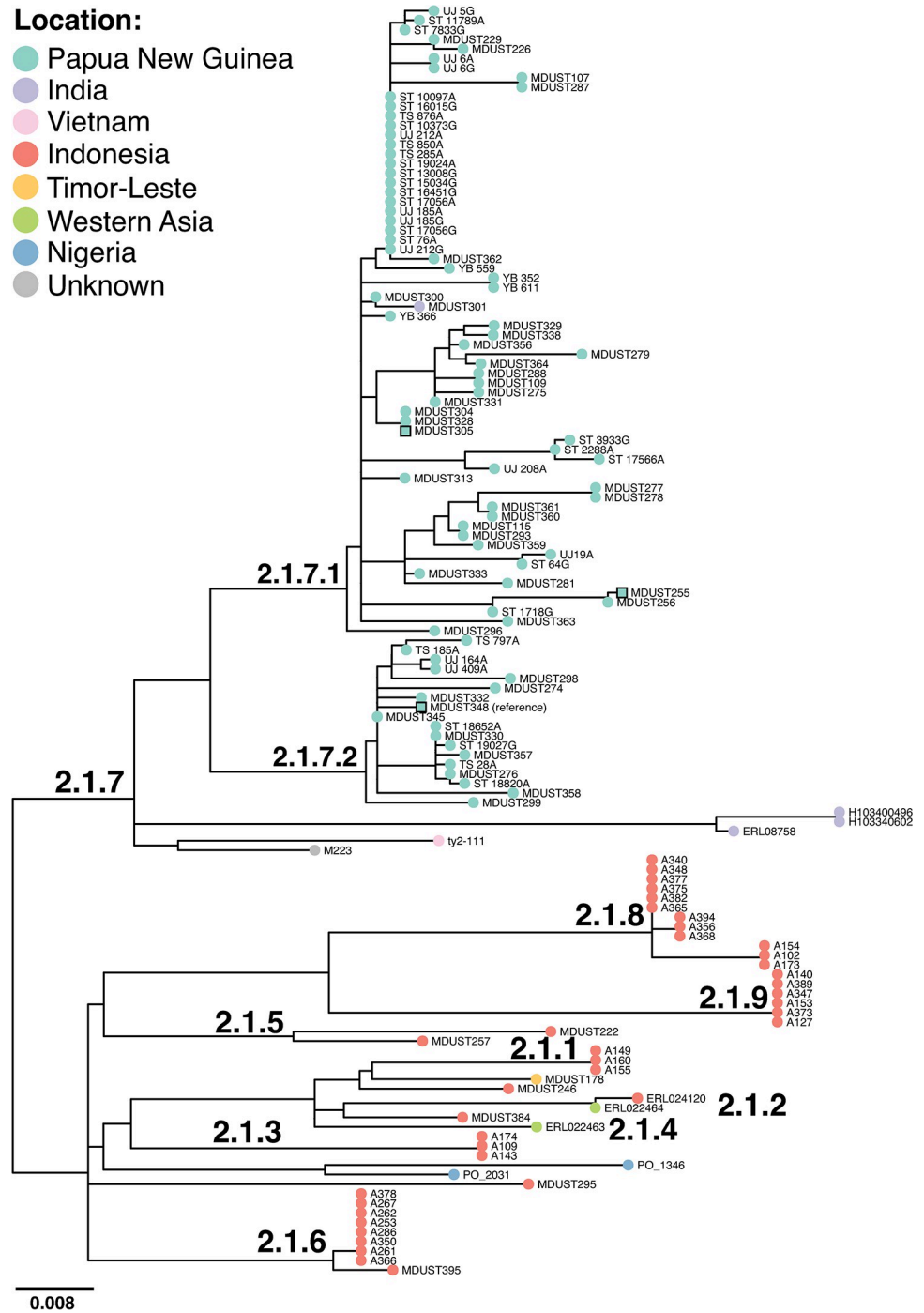


Fig 2. Maximum likelihood phylogeny inferred from all available (n = 133) clade 2.1 *S. Typhi*. Tips are coloured by location of origin as per inset legend (for traveller isolates, this is country of travel rather than country of isolation), branches are labelled by *S. Typhi* genotype. Square nodes indicate the position of the reference sequence and other completed genomes. Phylogeny is outgroup rooted with a selection of non-clade 2.1 taxa (see **Methods**). Interactive phylogeny available at: https://microreact.org/project/icGPbF_O8.

<https://doi.org/10.1371/journal.pntd.0010306.g002>

In order to estimate divergence times we analysed the sequence data with BEAST2 [40] (v2.4.7), using the GTR + Γ substitution model and sampling times (tip dates) defined by the year of isolation to calibrate the molecular clock. For n = 4 sequences (YB 352, YB 366, YB

559, and YB 611; [S1 Table](#)), the precise year of isolation was not available, however, as they were known to have been collected between 2009–2010 their date of isolation was estimated with BEAST2 using a tip date sampling prior specifying a uniform distribution between 2009–2010. Models were fit using both constant-coalescent population size and Bayesian skyline tree priors, together with either a strict clock model or a relaxed (uncorrelated lognormal distribution) clock model, to identify the model combination that best fit the data. For all model and tree prior combinations, a chain length of 10,000,000 steps sampling every 5,000 steps was selected[41]. Preliminary analyses highlighted the constant-coalescent population size and relaxed (uncorrelated lognormal) clock models to best fit the data. However, this model combination yielded an implausible divergence date (~1914, 95% HPD = 1835–1969), compared to those inferred from other model combinations (~1953, 95% HPD = 1917–1979), and analyses of $n = 43$ genotype 2.1.7 from the previous study alone [27] suggested the Bayesian skyline tree prior was better suited to the data. Therefore, the analyses presented here are those using the coalescent Bayesian skyline model together with a relaxed (uncorrelated log normal) clock model. To test the temporal signal we conducted a date-randomisation test whereby sampling times were randomly shuffled, and the analysis re-run 20 times[41,42]. These tests ([S3 Fig](#)) indicated that the data displayed temporal structure [41].

For the final analysis reported here, a single independent run conducted with a chain length of 10,000,000 states, sampling every 5,000 iterations, was subjected to removal of the first 10% of steps as burn-in with LogCombiner (v2.4.7)[40]. Maximum-clade credibility (MCC) trees were then generated with ‘common ancestor heights’ specified for node heights using TreeAnnotator (v2.4.6)[40]. Effective sample sizes were estimated to be >200 for all parameters reported. All phylogenies were visualised using the R package ggtree (v3.1.4.992) [43] (R script for this and other analyses available at: https://github.com/zadyson/PNG_STyphi_2022). The online tool Microreact [44] was used to create an online interactive dated phylogeny is available at <https://microreact.org/project/CYcTnSmjT>.

Reference genome assembly and annotation

Hybrid Illumina-Nanopore genome assemblies were generated with Unicycler (v0.4.7) [45] for three PNG genotype 2.1.7 isolates MDUST305, MDUST348, and MDUST255 (see [S2 Fig and S1 Table](#)). The resultant *de Bruijn* assembly graphs were then resolved into single contigs where necessary through manual curation with Bandage (v0.8.1)[46], and annotated with PROKKA (v1.14.0)[47]. Nucleotide pairwise comparisons of the completed genomes were carried out using Mauve [48] (vSnapshot_2015-02-25 1) and visualised using the genoPlotR (v0.8.11) package for R[49].

For analysis of prophage regions detected with PHASTER [37] in hybrid assembled sequences, these were aligned with their closest genetic relative in CT18 (prophage ST35) [50] using Mauve to confirm the prophage region boundaries and insertion site. Prophage regions were then extracted manually from the bacterial chromosome using seqret from the EMBOSS toolkit[51]. Annotations were then transferred to our PNG sequences from the CT18 reference sequence for prophage ST35 using the Rapid Annotation Transfer Tool (RATT)[52], and then manually curated together with PROKKA derived annotations. BLASTn comparisons were then carried out, and visualized using genoPlotR[49].

AMR detection

The mapping based allele typer SRST2 [53] (v0.1.8) was used to detect the presence of known acquired AMR genes and plasmid replicons using both the ARG-annot [54] and PlasmidFinder [55] databases, respectively. SNVs in the Quinolone Resistance Determining Regions

(QRDR) of genes *gyrA*, *gyrB*, and *parC* that are known to reduce susceptibility to fluoroquinolones in *S. Typhi* [7], and *acrB* mutations associated with azithromycin resistance in typhoidal *Salmonella* [12], were detected using GenoTyphi (**described above**).

Nucleotide sequence accession numbers

Raw read sequences for the 41 novel isolates from PNG have been deposited in the European Nucleotide Archive under project PRJEB20541; and individual accession numbers are listed in **S1 Table**. Genome assemblies for completed sequences of traveller isolates MDUST348, MDUST255, and MDUST305 were deposited in GenBank under accession numbers GCA_001173525, GCA_001122625, and GCA_001134605, respectively (see **S1 Table** for details).

Results

Population structure and antimicrobial resistance of *S. Typhi* in PNG

We combined sequence data from $n = 41$ *S. Typhi* collected locally in PNG and sequenced for this study, with $n = 45$ sequences from a previous study [27] that originate from returned travellers who acquired their infections in PNG. Together these data span 30 years (1980–2010). Travel-associated and locally-collected isolate sequences were intermingled in the inferred phylogeny (**S1 Fig**), and revealed that the population structure of *S. Typhi* during this period was dominated by a single genotype, 2.1.7 ($n = 84$, 98%; isolated 1985 to 2010) which has been present for at least 25 years. Only one other genotype was detected; clade 4.1 ($n = 2$, 2.3%; isolated 1980 and 1986) (**S1 Fig**).

The population structure of genotype 2.1.7 in PNG can be subdivided into two main lineages (**S2 Fig**), members of which are separated by a median pairwise distance of ~ 33 SNVs. We designated these as lineage I (genotype 2.1.7.1) and lineage II (2.1.7.2). In our data, genotype 2.1.7.1 was more common with $n = 66$ (78.6%) members, compared to $n = 18$ (21.4%) 2.1.7.2 members. Interestingly, 2.1.7.2 was not observed after 1998. Genotype 2.1.7.1 can be identified by the presence of a synonymous marker SNV C1503T in gene STY4417 (position 4286788 in CT18), and genotype 2.1.7.2 can be identified by the presence of a synonymous marker SNV C540T in gene STY4106 (position 3967063 in CT18). These genotypes have been added to the GenoTyphi scheme available at <http://github.com/katholt/genotypyphi> [27,35,56].

No known acquired AMR genes, nor mutations associated with either decreased ciprofloxacin susceptibility or azithromycin resistance (see **Methods**) were detected among any of the *S. Typhi* sequences from PNG, indicating this population is sensitive to commonly used antimicrobials. Susceptibility testing data was available for $n = 40$ (88.9%) traveller isolates, all of which tested susceptible to all drugs tested (see **Methods** and **S1 Table**).

Emergence and transmission dynamics of PNG genotype 2.1.7 sublineages

We applied Bayesian phylodynamic analysis to the set of PNG 2.1.7 sequences to estimate the dates of emergence for genotype 2.1.7, and its sublineages 2.1.7.1 and 2.1.7.2 (**Fig 1**, interactive phylogeny available at: <https://microrreact.org/project/CYcTnSmjT>). This analysis yielded a local substitution rate of 0.49 SNVs per genome per year (95% HPD 0.32–0.66) or 1.1×10^{-7} genome-wide substitutions per site per year (95% HPD 7.3×10^{-8} – 1.5×10^{-7}). These data showed temporal structure [41] to support these results (see **Methods** and **S3 Fig**), which were slightly lower than published estimates for *S. Typhi* for genotype 4.3.1 in Nepal (1.7×10^{-7} genome-wide substitutions per site per year, 95% HPD 1.1×10^{-7} – 2.4×10^{-7}) [8], Kenya (1.9×10^{-7} , 95% HPD 1.5×10^{-7} – 2.2×10^{-7}) [57], and globally (1.42×10^{-7} , 95% HPD 1.0×10^{-7} to 1.8×10^{-7}) [9]. We

estimate that the most recent common ancestor (MRCA) for all PNG genotype 2.1.7 existed circa ~1953 (95% HPD = 1917–1979) as shown in **Fig 1**. PNG 2.1.7 sublineages appear to have emerged contemporaneously in PNG in the mid-late 1970s, with the MRCA of 2.1.7.1 estimated circa ~1976 (95% HPD 1963–1985) and the MRCA of 2.1.7.2 around the same time or potentially slightly later, circa ~1979 (95% HPD 1971–1984).

In order to better understand the relationship of PNG genotype 2.1.7 *S. Typhi* to its nearest relatives in the pathogen population, we inferred a phylogeny of all available genome sequences belonging to its parent clade 2.1 (**Fig 2**, interactive phylogeny available at: https://microreact.org/project/icGPbF_O8). Genotype 2.1.7 isolates were mostly from PNG (n = 84, 93.3%). Within genotype 2.1.7, PNG sequences formed a distinct cluster (further subdivided into lineages 2.1.7.1 and 2.1.7.2), with a small number of distant relatives in India (three genotype 2.1.7 isolates separated from PNG isolates by a median pairwise distance of ~64 SNVs) and Vietnam (one sequence separated by median ~41 SNVs). The only non-PNG isolate clustering within the PNG sublineages was a single 2.1.7.1 sequence from India (MDUST301 from 1988; **Fig 2 and S1 Table**), which was intermingled in the phylogeny with the PNG isolates. This sequence was separated by a distance of ~3 SNV from its closest PNG relative suggestive of transmission of this genotype from PNG to India on at least one occasion. The other (i.e. non-2.1.7) isolates from clade 2.1 originated almost exclusively from other islands in the Indonesian archipelago (total n = 40 isolates, 90.1%; see **Fig 2**), suggesting that the PNG 2.1.7 group emerged from a local pathogen population that had been established in the region for some time.

Genetic variation among PNG genotype 2.1.7 *S. Typhi*

As described in **Methods**, we completed three PNG genotype 2.1.7 genomes to facilitate high resolution analysis of this clone. Three 2.1.7 genomes from PNG *S. Typhi* were selected for completion; two were the oldest sequences from each sublineage available (genotype 2.1.7.1 sequence MDUST305 isolated in 1990, and genotype 2.1.7.2 sequence MDUST348 isolated in 1985). We also completed a recent sequence of the most common genotype 2.1.7.1 (sequence MDUST255 isolated in 2010). The three completed genomes were markedly similar in size, differing by 4.4 kbp (**Fig 3A**). Further, pairwise nucleotide sequence comparisons revealed high levels of genetic similarity (99–100% coverage, 100% identity), and evidence of genome plasticity with multiple large-scale inversions observed adjacent to repetitive ribosomal RNA (*rrn*) operons (**Fig 3B**).

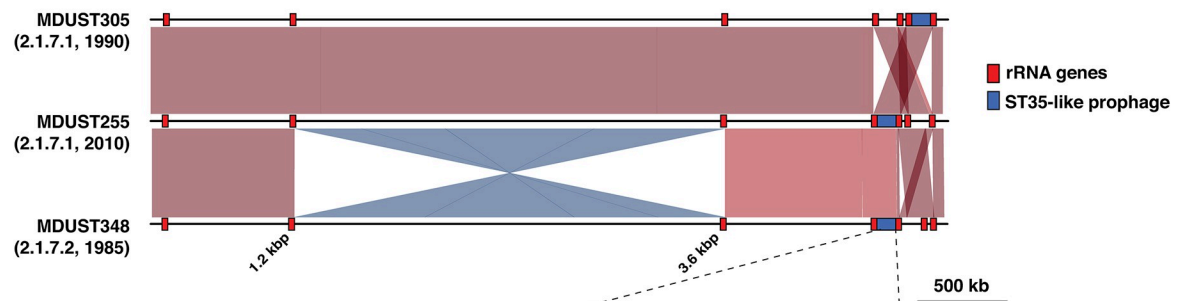
Gene content was highly conserved among both PNG 2.1.7 sublineages (96–100% coverage; **S4 Fig**). Very few deletions were observed, and were not restricted to a particular sublineage. Three sequences, two of genotype 2.1.7.1 (MDUST329 and UJ 212G) and one of 2.1.7.2 (MDUST332) appeared to have lost the *Salmonella* pathogenicity island 7 (SPI7) locus known to encode genes for Vi polysaccharide capsule formation, type IVB pili, and bacteriophage SopE. Such deletions are known to occur in the laboratory and did not appear to affect the phage types previously determined for the PNG 2.1.7 traveller isolates as all were of phage type D2. Besides the deletion of prophage SopE within SPI7, no other prophage deletions were apparent.

Examination of detected prophage elements amongst the three completed prophage sequences revealed a smaller ~1.8 kbp inversion within an ST35-like prophage in the genome sequence of MDUST255 (**Fig 3C**). The inversion was located in the prophage tail gene module carrying homologues of prophage ST35 genes (from reference sequence CT18) encoding for a tail fibre (STY3691), tail fibre assembly chaperone (STY3692), alternate tail fibre tips (STY3693 and STY3694), a site-specific invertase (STY3695), and major tail tube filament

A

Sequence	Year	Sample type	Genotype	Genome size (bp)	G+C mol %	Prophage regions	Phage type
MDUST305	1990	Blood	2.1.7.1	4,796,602	52.04	6	D2
MDUST255	2010	Blood	2.1.7.1	4,797,669	52.04	6	D2
MDUST348	1985	Unknown	2.1.7.2	4,793,333	52.04	6	D2

B



C

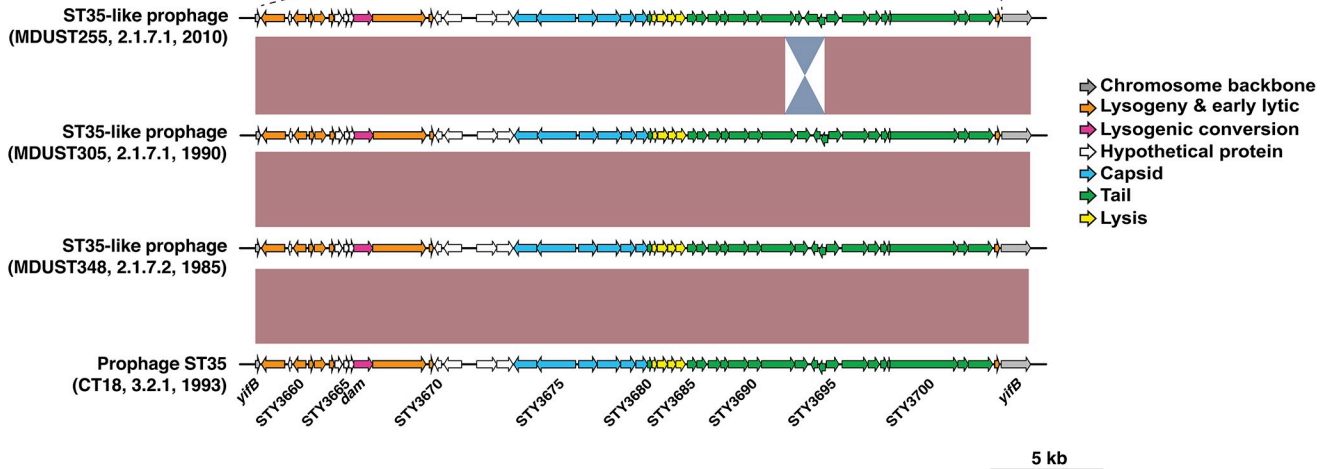


Fig 3. Genome plasticity among PNG *S. Typhi*. (A) Summary of genetic features of completed PNG 2.1.7 *S. Typhi* genomes. (B) Pairwise nucleotide alignments of PNG 2.1.7 *S. Typhi*. Shaded regions indicate nucleotide sequence homology, red blocks indicate ribosomal RNA (rRNA) operons, and blue blocks indicate the ST35-like prophage as per the inset legend. (C) Pairwise alignment of prophage ST35-like elements in PNG *S. Typhi* and the CT18 reference sequence. Shaded regions indicate nucleotide sequence homology, and genes are coloured by putative function as per the inset legend.

<https://doi.org/10.1371/journal.pntd.0010306.g003>

structural gene (STY3696)[50]. The inversion, likely mediated by the site-specific invertase (STY3695), appears to have resulted in the formation of an alternative tail fibre protein; a fusion of the N-terminal region of tail fibre protein STY3691 with alternative tail fibre tip protein STY3694. Examination of alignments of read data to reference sequence MDUST348

suggested that this locus was present in both orientations across the PNG 2.1.7 population in line with previous reports[50].

Discussion

These findings demonstrate that the vast majority of typhoid fever cases in PNG are caused by a single lineage of *S. Typhi* (genotype 2.1.7) which emerged in the Indonesian archipelago in the mid-twentieth century before differentiating into two sublineages in the mid-late 1970s, with genotype 2.1.7.2 no longer appearing common after 1998. These findings correlate well with the broad epidemiology of typhoid fever in PNG. Typhoid fever was sporadically reported in PNG in the 1960s, becoming more commonly diagnosed in the 1970s and then becoming endemic in the PNG highlands in the 1980s [16,58]. Thus, the divergence dates for PNG *S. Typhi* 2.1.7 sublineages correspond to the approximate time when typhoid fever became endemic. *S. Typhi* genotype 2.1.7 appears to have persisted within PNG since this time, and was dominant among both returning travellers and locally detected cases over the course of 30 years of genomic surveillance. We observed little evidence of genetic variation among the genotype 2.1.7 population, with infrequent gene deletions (S4 Fig), no obvious evidence of prophage or plasmid acquisitions (Figs 3 and S4), and a lower substitution rate compared to the widespread H58 (4.3.1) genotypes associated with intercontinental transmission and AMR [8,9,57]. Notably, our data suggest that AMR has not emerged among the *S. Typhi* population in this setting, with both genotype and phenotype data revealing all isolates to be sensitive to former first line drugs (chloramphenicol, ampicillin and co-trimoxazole), as well as fluoroquinolones, third generation cephalosporins, and macrolides. The AMR sensitivity of the PNG 2.1.7 genotype and the long term stability of this population is in sharp contrast to many other endemic settings, e.g. South Asia where infections are commonly driven by H58 genotypes[9], both MDR and reduced susceptibility to fluoroquinolones are common [7,8,59,60], and resistance to third generation cephalosporins[10,11], and azithromycin [12–15,61] have also emerged.

We observed limited evidence of inter-country transmission, with no obvious recent introductions into PNG from outside. A small number of genotype 2.1.7 isolates were found outside PNG in India and Vietnam, but these were mostly distantly related to the PNG lineages and we observed only one possible case of transmission of PNG genotype 2.1.7 *S. Typhi* into India (travel-associated case of lineage 2.1.7.1 isolated in 1988 Fig 2). However, it is possible that this individual may also have visited PNG as well as India in their travels with only the later recorded. The limited international movement of *S. Typhi* in and out of PNG may partially explain why AMR has not been observed in this setting, in the form of the arrival of H58 or other drug resistant genotypes of *S. Typhi*.

Among PNG 2.1.7 populations, the lack of emergence of AMR may be influenced by a lack of local selective pressures related to antimicrobial usage. Tracking antimicrobial usage is challenging in PNG, but evidence of sub-optimal administration of antibiotics has been reported [62]. Approximately 73% of outpatients with non-malarial febrile illness are administered antibiotics, commonly contrary to the relevant prescription guidelines. Amoxicillin, trimethoprim-sulfamethoxazole and benzylpenicillin collectively accounted for ~75% of antibiotic administrations, demonstrating the high dependence on β -lactams in PNG, but little exposure to cephalosporins. Data pertaining to other Gram-negative pathogens in PNG are limited. A recent investigation of *Shigella* from PNG and the Pacific region revealed this pathogen to have resistance to ampicillin, chloramphenicol, tetracycline, and trimethoprim-sulfamethoxazole, with resistance more common in bacteria isolated since 2010. Resistance to cephalosporins (ceftriaxone) and ciprofloxacin was not detected in *Shigella* from PNG[63]. On the basis

of observations in *Shigella*, resistance to first-line antibiotics could emerge in other gastrointestinal pathogens in PNG, but exposure, and therefore resistance to, cephalosporins and second-generation quinolones (e.g. ciprofloxacin) is unlikely.

The two *S. Typhi* genomes sequenced previously from PNG by Baddam and colleagues [21] were genotyped as 2.1.7 under the new scheme, consistent with our finding that most PNG isolates belong to this genotype. Genome variation reported previously in the PNG *S. Typhi* population [18] appears to be primarily mediated by large-scale re-arrangements adjacent to repetitive *rrn* operons (Fig 3B), rather than a diverse range of genotypes or the acquisition of mobile genetic elements such as prophages or plasmids consistent with previous reports [18,19]. Rearrangements between the *rrn* operons have been described previously in *S. Typhi* and are driven by homologous recombination [64,65]. This has previously been postulated to cause sequence variation (e.g. diverse ribotypes) among PNG *S. Typhi* [19,20], and our data support this hypothesis. Recent long-read data suggests that rearrangements around the *rrn* operons are relatively common in *S. Typhi* [66], and appear to be associated with host persistence and carriage in the gallbladder [67]. We also observed inversions in the tail morphogenesis gene module of the ST35-like prophage. These inversions are also well described and are likely a mechanism for generating variation in tail fibre genes that are known to be involved with the attachment stage of viral replication [68]. Such variation is known to cause alteration of phage host ranges [50,69].

This study is not without limitations, primarily relating to the challenges associated with obtaining clinical isolates of *S. Typhi*. There is no ongoing routine surveillance conducted in PNG, nor is there routine blood culture for diagnosis of febrile illness. Consequently, locally collected isolates are obtained primarily through targeted studies. Most isolates analysed in this study originated from Eastern Highlands Province, where the PNG Institute of Medical Research is headquartered and is thus the site of most studies. Eastern Highlands Province is one of 22 provinces in the country (albeit the second most populated province), thus there is a lack of geographical distribution of isolates. Given most isolates originate from targeted studies, the temporal distribution of isolates is intermittent. Moreover, the lack of ongoing surveillance precluded contemporary isolates (from the past 10 years) being collected and thus included in this study. However, as locally collected and travel associated cases both demonstrated a lack of AMR and similar genotype frequencies, in line with limited previously published sequences [21], these combined data appear representative of major trends among the *S. Typhi* population in this setting.

Overall, our data highlight that unlike other recently examined settings in Kenya, Zimbabwe and other regions of Sub-Saharan Africa, as well as South Asian settings in India, Nepal, Bangladesh and Pakistan [8,10,57,59,60,70–73] the *S. Typhi* populations in PNG appear sensitive to a diverse range of antimicrobials with former first line drugs, fluoroquinolones, third generation cephalosporins and macrolides all remaining viable treatment options prior to the introduction of Vi-conjugate vaccines and improvements to water, sanitation and hygiene (WaSH) related infrastructure. However, despite this, further ongoing surveillance is required to monitor for the potential emergence or introduction of AMR *S. Typhi* in this setting.

Supporting information

S1 Fig. Maximum likelihood phylogeny for all available *S. Typhi* originating in PNG. Tip colours indicate the source of the sequence as per the inset legend. Branches are labelled by *S. Typhi* genotype. Square nodes indicate the position of the reference sequence and other completed genomes.

(PDF)

S2 Fig. Maximum likelihood phylogeny for all available PNG genotype 2.1.7 *S. Typhi*. Tip colours indicate the source of the sequence as per the inset legend, and branches are labelled by genotype. Square nodes indicate the position of the reference sequence and other completed genomes. A Bayesian dated tree inferred using the same alignment and year of isolation is shown in [Fig 1](#).

(PDF)

S3 Fig. Temporal Analysis of PNG genotype 2.1.7 *S. Typhi*. (A) **Tempest regression of root-to-tip distances as a function of time**, with the root of the tree selected using heuristic residual mean squared. Each point represents a tip in the maximum likelihood phylogenetic tree shown in [S2 Fig](#). The slope is a crude estimate of the annual substitution rate for the SNV alignment, the x-intercept corresponds to the age of the root node, and R^2 is a measure of clock-like behaviour among the data. (B) **Date randomisation test results** with the right most box plot showing the posterior substitution rate estimate from the SNV alignment of the data with the correct sampling times, and the remaining 20 boxplots showing the posterior distributions of the rate estimate from replicate runs where the dates were subjected to randomisation. The data are considered to have strong temporal structure if the estimates using the correct dates do not overlap with those where the dates were randomised.

(PDF)

S4 Fig. Genetic conservation among PNG 2.1.7 *S. Typhi*. Branches are labelled by genotype. Coloured square nodes indicate the position of the reference sequence and other completed genomes. Heatmap shows the presence or absence of all annotated genes in the reference sequence as per the inset legend. Features of interest are highlighted with red boxes and labels.

(PDF)

S1 Table. Accession numbers and data for *S. Typhi* sequences used in this study.

(XLSX)

S2 Table. Excluded repeat and phage regions in PNG MDUST348 2.1.7.2 completed reference sequence.

(XLSX)

S3 Table. Outgroups used for phylogenetic tree rooting.

(XLSX)

Acknowledgments

The authors wish to thank Sebastian Duchene for useful discussions regarding phylodynamic analyses.

Author Contributions

Conceptualization: Paul F. Horwood, Andrew R. Greenhill, Kathryn E. Holt.

Data curation: Zoe Anne Dyson, Deborah A. Williamson.

Formal analysis: Zoe Anne Dyson.

Funding acquisition: Zoe Anne Dyson, Gordon Dougan, Kathryn E. Holt.

Investigation: Zoe Anne Dyson, Elisheba Malau, Valentine Siba, Mition Yoannes, Louise M. Judd, Danielle J. Ingle, Deborah A. Williamson, Kathryn E. Holt.

Methodology: Zoe Anne Dyson, Kathryn E. Holt.

Project administration: Andrew R. Greenhill.

Resources: Rebecca Ford, William Pomat, Megan Passey, Deborah A. Williamson, Andrew R. Greenhill, Kathryn E. Holt.

Software: Zoe Anne Dyson, Kathryn E. Holt.

Supervision: Rebecca Ford, William Pomat, Gordon Dougan, Kathryn E. Holt.

Visualization: Zoe Anne Dyson.

Writing – original draft: Zoe Anne Dyson, Paul F. Horwood, Andrew R. Greenhill, Kathryn E. Holt.

Writing – review & editing: Zoe Anne Dyson, Elisheba Malau, Paul F. Horwood, Rebecca Ford, Valentine Siba, Mitton Yoannes, William Pomat, Megan Passey, Louise M. Judd, Danielle J. Ingle, Deborah A. Williamson, Gordon Dougan, Andrew R. Greenhill, Kathryn E. Holt.

References

1. Parry C, Hien T, Dougan G, White N, Farrar J. Typhoid fever. *N Engl J Med*. 2002; 347: 1770–1782. <https://doi.org/10.1056/NEJMra020201> PMID: 12456854
2. GBD 2017 Typhoid and Paratyphoid Collaborators. The global burden of typhoid and paratyphoid fevers: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Infect Dis*. 2019; 19: 369–381. [https://doi.org/10.1016/S1473-3099\(18\)30685-6](https://doi.org/10.1016/S1473-3099(18)30685-6) PMID: 30792131
3. Connor B, Schwartz E. Typhoid and paratyphoid fever in travellers. *Lancet Infect Dis*. 2005; 5: 623–628. [https://doi.org/10.1016/S1473-3099\(05\)70239-5](https://doi.org/10.1016/S1473-3099(05)70239-5) PMID: 16183516
4. Butler T, Knight J, Nath S, Speelman P, Roy S, Azad M. Typhoid fever complicated by intestinal perforation: a persisting fatal disease requiring surgical management. *Rev Infect Dis*. 1985; 7: 244–256. <https://doi.org/10.1093/clinids/7.2.244> PMID: 3890097
5. Crump J. Progress in Typhoid Fever Epidemiology. *Clin Infect Dis*. 2019; 68: S4–S9. <https://doi.org/10.1093/cid/ciy846> PMID: 30767000
6. Dyson Z, Klemm E, Palmer S, Dougan G. Antibiotic Resistance and Typhoid. *Clin Infect Dis*. 2019; 68: S165–S170. <https://doi.org/10.1093/cid/ciy1111> PMID: 30845331
7. Thanh D, Karkey A, Dongol S, Thi N, Thompson C, Rabaa M, et al. A novel ciprofloxacin-resistant sub-clade of H58 *Salmonella* Typhi is associated with fluoroquinolone treatment failure. *Elife*. 2016; 5: e14003. <https://doi.org/10.7554/eLife.14003> PMID: 26974227
8. Britto C, Dyson Z, Duchene S, Carter M, Gurung M, Kelly D, et al. Laboratory and molecular surveillance of paediatric typhoidal *Salmonella* in Nepal: Antimicrobial resistance and implications for vaccine policy. *PLoS Negl Trop Dis*. 2018; 12: e0006408. <https://doi.org/10.1371/journal.pntd.0006408> PMID: 29684021
9. Wong V, Baker S, Pickard D, Parkhill J, Page A, Feasey N, et al. Phylogeographical analysis of the dominant multidrug-resistant H58 clade of *Salmonella* Typhi identifies inter- and intracontinental transmission events. *Nat Genet*. 2015; 47: 632–639. <https://doi.org/10.1038/ng.3281> PMID: 25961941
10. Klemm E, Shakoor S, Page A, Qamar F, Judge K, Saeed D, et al. Emergence of an Extensively Drug-Resistant *Salmonella enterica* Serovar Typhi Clone Harboring a Promiscuous Plasmid Encoding Resistance to Fluoroquinolones and Third-Generation Cephalosporins. *mBio*. 2018; 9. <https://doi.org/10.1128/mBio.00105-18> PMID: 29463654
11. Rasheed F, Saeed M, Alikhan N, Baker D, Khurshid M, Ainsworth E, et al. Emergence of Resistance to Fluoroquinolones and Third-Generation Cephalosporins in *Salmonella* Typhi in Lahore, Pakistan. *Microorganisms*. 2020; 8. <https://doi.org/10.3390/microorganisms8091336> PMID: 32883020
12. Hooda Y, Sajib M, Rahman H, Luby S, Bondy-Denomy J, Santosham M, et al. Molecular mechanism of azithromycin resistance among typhoidal *Salmonella* strains in Bangladesh identified through passive pediatric surveillance. *PLoS Negl Trop Dis*. 2019; 13: e0007868. <https://doi.org/10.1371/journal.pntd.0007868> PMID: 31730615
13. Carey M, Jain R, Yousuf M, Maes M, Dyson Z, Thu T, et al. Spontaneous Emergence of Azithromycin Resistance in Independent Lineages of *Salmonella* Typhi in Northern India. *Clin Infect Dis*. 2021; 72: e120–e127. <https://doi.org/10.1093/cid/ciaa1773> PMID: 33515460

14. Duy P, Dongol S, Giri A, Nguyen T, Thanh H, Nguyen Q, et al. The emergence of azithromycin-resistant *Salmonella* Typhi in Nepal. *JAC-AMR*. 2020; 2. <https://doi.org/10.1093/jacamr/dlaa109> PMID: [34223059](https://pubmed.ncbi.nlm.nih.gov/34223059/)
15. Iqbal J, Dehraj I, Carey M, Dyson Z, Garrett D, Seidman J, et al. A Race against Time: Reduced Azithromycin Susceptibility in *Salmonella enterica* Serovar Typhi in Pakistan. *mSphere*. 2020; 5. <https://doi.org/10.1128/mSphere.00215-20> PMID: [32699118](https://pubmed.ncbi.nlm.nih.gov/32699118/)
16. Passey M. The new problem of typhoid fever in Papua New Guinea: how do we deal with it? *PNG Med J*. 1995; 38: 300–304. PMID: [9522872](https://pubmed.ncbi.nlm.nih.gov/9522872/)
17. Siba V, Horwood P, Vanuga K, Wapling J, Sehuko R, Siba P, et al. Evaluation of serological diagnostic tests for typhoid fever in Papua New Guinea using a composite reference standard. *Clin Vaccine Immunol*. 2012; 19: 1833–1837. <https://doi.org/10.1128/CVI.00380-12> PMID: [22993409](https://pubmed.ncbi.nlm.nih.gov/22993409/)
18. Combs B, Passey M, Michael A, Pang T, Lightfoot D, Alpers M. Ribotyping of *Salmonella enterica* serovar Typhi isolates from Papua New Guinea over the period 1977 to 1996. *PNG Med J*. 2005; 48: 158–167. PMID: [17212062](https://pubmed.ncbi.nlm.nih.gov/17212062/)
19. Thong K, Goh Y, Yasin R, Lau M, Passey M, Winston G, et al. Increasing genetic diversity of *Salmonella enterica* serovar typhi isolates from Papua New Guinea over the period from 1992 to 1999. *J Clin Microbiol*. 2002; 40: 4156–4160. <https://doi.org/10.1128/JCM.40.11.4156-4160.2002> PMID: [12409390](https://pubmed.ncbi.nlm.nih.gov/12409390/)
20. Thong K, Passey M, Clegg A, Combs B, Yassin R, Pang T. Molecular analysis of isolates of *Salmonella typhi* obtained from patients with fatal and nonfatal typhoid fever. *J Clin Microbiol*. 1996; 34: 1029–1033. <https://doi.org/10.1128/jcm.34.4.1029-1033.1996> PMID: [8815078](https://pubmed.ncbi.nlm.nih.gov/8815078/)
21. Baddam R, Thong K, Avasthi T, Shaik S, Yap K, Teh C, et al. Whole-genome sequences and comparative genomics of *Salmonella enterica* serovar Typhi isolates from patients with fatal and nonfatal typhoid fever in Papua New Guinea. *J Bacteriol*. 2012; 194: 5122–5123. <https://doi.org/10.1128/JB.01051-12> PMID: [22933755](https://pubmed.ncbi.nlm.nih.gov/22933755/)
22. United Nations Children’s Fund. At a glance: Papua New Guinea. [Cited 2021 September 24]. Available from: <https://www.unicef.org/png>
23. United Nations Children’s Fund. Papua New Guinea. [Cited 2021 September 24]. Available from: <https://www.unicef.org/png/what-we-do/water-sanitation-and-hygiene>
24. United Nations Development Programme. Papua New Guinea. 2020 [Cited 2021 September 24]. Available from: http://hdr.undp.org/sites/all/themes/hdr_theme/country-notes/PNG.pdf
25. Williamson D, Lane C, Easton M, Valcanis M, Strachan J, Veitch M, et al. Increasing Antimicrobial Resistance in Nontyphoidal *Salmonella* Isolates in Australia from 1979 to 2015. *Antimicrob Agents Chemother*. 2018; 62. <https://doi.org/10.1128/AAC.02012-17> PMID: [29180525](https://pubmed.ncbi.nlm.nih.gov/29180525/)
26. Wick R, Judd L, Gorrie C, Holt K. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genom*. 2017; 3: e000132. <https://doi.org/10.1099/mgen.0.000132> PMID: [29177090](https://pubmed.ncbi.nlm.nih.gov/29177090/)
27. Wong V, Baker S, Connor T, Pickard D, Page A, Dave J, et al. An extended genotyping framework for *Salmonella enterica* serovar Typhi, the cause of human typhoid. *Nat Commun*. 2016; 7: 12827. <https://doi.org/10.1038/ncomms12827> PMID: [27703135](https://pubmed.ncbi.nlm.nih.gov/27703135/)
28. Parkhill J, Dougan G, James K, Thomson N, Pickard D, Wain J, et al. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature*. 2001; 413: 848–852. <https://doi.org/10.1038/35101607> PMID: [11677608](https://pubmed.ncbi.nlm.nih.gov/11677608/)
29. Langmead B, Salzberg S. Fast gapped-read alignment with Bowtie 2. *Nat Meth*. 2012; 9: 357–359. <https://doi.org/10.1038/nmeth.1923> PMID: [22388286](https://pubmed.ncbi.nlm.nih.gov/22388286/)
30. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25: 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352> PMID: [19505943](https://pubmed.ncbi.nlm.nih.gov/19505943/)
31. Holt K, Parkhill J, Mazzoni CJ, Roumagnac P, Weill F, Goodhead I, et al. High-throughput sequencing provides insights into genome variation and evolution in *Salmonella* Typhi. *Nat Genet*. 2008; 40: 987–993. <https://doi.org/10.1038/ng.195> PMID: [18660809](https://pubmed.ncbi.nlm.nih.gov/18660809/)
32. Ingle D, Nair S, Hartman H, Ashton P, Dyson Z, Day M, et al. Informal genomic surveillance of regional distribution of *Salmonella* Typhi genotypes and antimicrobial resistance via returning travellers. *PLoS Negl Trop Dis*. 2019; 13: e0007620. <https://doi.org/10.1371/journal.pntd.0007620> PMID: [31513580](https://pubmed.ncbi.nlm.nih.gov/31513580/)
33. Croucher N, Page A, Connor T, Delaney A, Keane J, Bentley S, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res*. 2015; 43: e15–e15. <https://doi.org/10.1093/nar/gku1196> PMID: [25414349](https://pubmed.ncbi.nlm.nih.gov/25414349/)
34. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014; 30: 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033> PMID: [24451623](https://pubmed.ncbi.nlm.nih.gov/24451623/)

35. Dyson Z, Holt K. Five years of GenoTyphi: updates to the global *Salmonella* Typhi genotyping framework. *J Infect Dis*. 2021. <https://doi.org/10.1093/infdis/jiab414> PMID: 34453548
36. Delcher A, Salzberg S, Phillippy A. Using MUMmer to identify similar regions in large sequence sets. *Curr Protoc Bioinformatics*. 2003;Chapter 10: Unit 10.3. <https://doi.org/10.1002/0471250953.bi1003s00> PMID: 18428693
37. Arndt D, Grant J, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res*. 2016; 44: W16–21. <https://doi.org/10.1093/nar/gkw387> PMID: 27141966
38. International Typhoid Consortium, Wong V, Holt K, Okoro C, Baker S, Pickard D, et al. Molecular Surveillance Identifies Multiple Transmissions of Typhoid in West Africa. *PLoS Negl Trop Dis*. 2016; 10: e0004781. <https://doi.org/10.1371/journal.pntd.0004781> PMID: 27657909
39. Rambaut A, Lam T, Carvalho L, Pybus O. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol*. 2016; 2: vew007. <https://doi.org/10.1093/ve/vew007> PMID: 27774300
40. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C, Xie D, et al. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol*. 2014; 10: e1003537. <https://doi.org/10.1371/journal.pcbi.1003537> PMID: 24722319
41. Duchene S, Holt K, Weill F, Hello S, Hawkey J, Edwards D, et al. Genome-scale rates of evolutionary change in bacteria. *Microb Genom*. 2016; 2: e000094. <https://doi.org/10.1099/mgen.0.000094> PMID: 28348834
42. Firth C, Kitchen A, Shapiro B, Suchard M, Holmes E, Rambaut A. Using time-structured data to estimate evolutionary rates of double-stranded DNA viruses. *Mol Biol Evol*. 2010; 27: 2038–2051. <https://doi.org/10.1093/molbev/msq088> PMID: 20363828
43. Yu G, Smith D, Zhu H, Guan Y, Lam T. ggtree: an rpackage for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol*. 2016; 8: 28–36. <https://doi.org/10.1111/2041-210x.12628>
44. Argimon S, Abudahab K, Goater R, Fedosejev A, Bhai J, Glasner C, et al. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb Genom*. 2016; 2: e000093. <https://doi.org/10.1099/mgen.0.000093> PMID: 28348833
45. Wick R, Judd L, Gorrie C, Holt K. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol*. 2017; 13: e1005595. <https://doi.org/10.1371/journal.pcbi.1005595> PMID: 28594827
46. Wick R, Schultz M, Zobel J, Holt K. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics*. 2015; 31: 3350–3352. <https://doi.org/10.1093/bioinformatics/btv383> PMID: 26099265
47. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014; 30: 2068–2069. <https://doi.org/10.1093/bioinformatics/btu153> PMID: 24642063
48. Darling A, Mau B, Blattner F, Perna N. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res*. 2004; 14: 1394–1403. <https://doi.org/10.1101/gr.2289704> PMID: 15231754
49. Guy L, Kultima J, Andersson S. genoPlotR: comparative gene and genome visualization in R. *Bioinformatics*. 2010; 26: 2334–2335. <https://doi.org/10.1093/bioinformatics/btq413> PMID: 20624783
50. Thomson N, Baker S, Pickard D, Fookes M, Anjum M, Hamlin N, et al. The role of prophage-like elements in the diversity of *Salmonella enterica* serovars. *J Mol Biol*. 2004; 339: 279–300. <https://doi.org/10.1016/j.jmb.2004.03.058> PMID: 15136033
51. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet*. 2000; 16: 276–277. [https://doi.org/10.1016/s0168-9525\(00\)02024-2](https://doi.org/10.1016/s0168-9525(00)02024-2) PMID: 10827456
52. Otto T, Dillon G, Degraeve W, Berriman M. RATT: Rapid Annotation Transfer Tool. *Nucleic Acids Res*. 2011; 39: e57. <https://doi.org/10.1093/nar/gkq1268> PMID: 21306991
53. Inouye M, Dashnow H, Raven L, Schultz M, Pope B, Tomita T, et al. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med*. 2014; 6: 90. <https://doi.org/10.1186/s13073-014-0090-6> PMID: 25422674
54. Gupta S, Padmanabhan B, Diene S, Lopez-Rojas R, Kempf M, Landraud L, et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother*. 2014; 58: 212–220. <https://doi.org/10.1128/AAC.01310-13> PMID: 24145532
55. Carattoli A, Zankari E, García-Fernández A, Larsen M, Lund O, Villa L, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother*. 2014; 58: 3895–3903. <https://doi.org/10.1128/AAC.02412-14> PMID: 24777092
56. Dyson Z, Holt K. katholt/genotyphi: GenoTyphi v1.9.1. Zenodo. 2021. <https://doi.org/10.5281/zenodo.4707614>

57. Kariuki S, Dyson Z, Mbae C, Ngetich R, Kavai S, Wairimu C, et al. Multiple introductions of multidrug-resistant typhoid associated with acute infection and asymptomatic carriage, Kenya. *Elife*. 2021; 10: e67852. <https://doi.org/10.7554/eLife.67852> PMID: 34515028
58. Richens J. Typhoid in the highlands of Papua New Guinea 1984–1990: a hospital-based perspective. *PNG Med J*. 1995; 38: 305–314. PMID: 9522873
59. Britto C, Dyson Z, Mathias S, Bosco A, Dougan G, Jose S, et al. Persistent circulation of a fluoroquinolone-resistant *Salmonella enterica* Typhi clone in the Indian subcontinent. *J Antimicrob Chemother*. 2020; 75: 337–341. <https://doi.org/10.1093/jac/dkz435> PMID: 31665304
60. Rahman S, Dyson Z, Klemm E, Khanam F, Holt K, Chowdhury E, et al. Population structure and antimicrobial resistance patterns of *Salmonella* Typhi isolates in urban Dhaka, Bangladesh from 2004 to 2016. *PLoS Negl Trop Dis*. 2020; 14: e0008036. <https://doi.org/10.1371/journal.pntd.0008036> PMID: 32106221
61. Sajib M, Tanmoy A, Hooda Y, Rahman H, Andrews J, Garrett D, et al. Tracking the Emergence of Azithromycin Resistance in Multiple Genotypes of Typhoidal *Salmonella*. *mBio*. 2021; 12: e03481–20. <https://doi.org/10.1128/mBio.03481-20> PMID: 33593966
62. Saweri O, Hetzel M, Mueller I, Siba P, Pulford J. The treatment of non-malarial febrile illness in Papua New Guinea: findings from cross sectional and longitudinal studies of health worker practice. *BMC Health Serv Res*. 2017; 17: 10. <https://doi.org/10.1186/s12913-016-1965-6> PMID: 28056949
63. Malau E, Ford R, Valcanis M, Jennison A, Mosse J, Bean D, et al. Antimicrobial sensitivity trends and virulence genes in *Shigella* spp. from the Oceania region. *Infect Genet Evol*. 2018; 64: 52–56. <https://doi.org/10.1016/j.meegid.2018.06.015> PMID: 29906636
64. Brüssow H, Canchaya C, Hardt W. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev*. 2004; 68: 560–602. <https://doi.org/10.1128/MMBR.68.3.560-602.2004> PMID: 15353570
65. Sanderson K, Liu S. Chromosomal rearrangements in enteric bacteria. *Electrophoresis*. 1998; 19: 569–572. <https://doi.org/10.1002/elps.1150190417> PMID: 9588803
66. Page A, Ainsworth E, Langridge G. socru: typing of genome-level order and orientation around ribosomal operons in bacteria. *Microb Genom*. 2020; 6: mgen000396. <https://doi.org/10.1099/mgen.0.000396> PMID: 32584752
67. Matthews T, Rabsch W, Maloy S. Chromosomal rearrangements in *Salmonella enterica* serovar Typhi strains isolated from asymptomatic human carriers. *mBio*. 2011; 2: e00060–11. <https://doi.org/10.1128/mBio.00060-11> PMID: 21652779
68. Guttman B, Raya R, Kutter E. Basic Phage Biology. In: Kutter E, Sulakvelidze A, editors. *Bacteriophages Biology and Applications*. Bacteriophages Biology and Applications; 2005. pp. 29–66.
69. Kutter E, Raya R, Carlson K. Molecular mechanisms of phage infection. In: Kutter E, Sulakvelidze A, editors. *Bacteriophages Biology and Applications*. Bacteriophages Biology and Applications; 2005. pp. 165–222.
70. Mashe T, Leekitcharoenphon P, Mtapuri-Zinyowera S, Kingsley R, Robertson V, Tarupiwa A, et al. *Salmonella enterica* serovar Typhi H58 clone has been endemic in Zimbabwe from 2012 to 2019. *J Antimicrob Chemother*. 2020; 75: 1160–1167. <https://doi.org/10.1093/jac/dkaa519> PMID: 33347558
71. Pragasam A, Pickard D, Wong V, Dougan G, Kang G, Thompson A, et al. Phylogenetic Analysis Indicates a Longer Term Presence of the Globally Distributed H58 Haplotype of *Salmonella* Typhi in Southern India. *Clin Infect Dis*. 2020; 71: 1856–1863. <https://doi.org/10.1093/cid/ciz1112> PMID: 32003431
72. Park S, Pham D, Boinett C, Wong V, Pak G, Panzner U, et al. The phylogeography and incidence of multi-drug resistant typhoid fever in sub-Saharan Africa. *Nat Commun*. 2018; 9: 5094. <https://doi.org/10.1038/s41467-018-07370-z> PMID: 30504848
73. da Silva K, Tanmoy A, Pragasam A, Iqbal J, Sajib M, Mutreja A, et al. The international and intercontinental spread and expansion of antimicrobial-resistant *Salmonella* Typhi. *medRxiv* [Preprint]. 2021 [cited 2021 September 24]. Available from: <https://www.medrxiv.org/content/10.1101/2021.09.03.21262852v1> PMID: 33655261