

RESEARCH ARTICLE

Whole genome sequence analysis of *Salmonella* Typhi provides evidence of phylogenetic linkage between cases of typhoid fever in Santiago, Chile in the 1980s and 2010–2016

Mailis Maes¹*, Michael J. Sikorski^{2,3}*, Megan E. Carey¹, Ellen E. Higginson¹, Zoe A. Dyson^{1,4,5,6}, Alda Fernandez⁷, Pamela Araya⁷, Sharon M. Tennant², Stephen Baker¹, Rosanna Lagos⁸, Juan Carlos Hormazábal⁷, Myron M. Levine^{2,†}, Gordon Dougan^{1,6,‡}

1 Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Department of Medicine, University of Cambridge, Cambridge, United Kingdom, **2** Center for Vaccine Development and Global Health (CVD), University of Maryland School of Medicine, Baltimore, Maryland, United States of America, **3** Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland, United States of America, **4** Department of Infectious Diseases, Monash University, Melbourne, Australia, **5** London School of Hygiene & Tropical Medicine, London, United Kingdom, **6** Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, United Kingdom, **7** Bacteriología, Subdepartamento de Enfermedades Infecciosas, Departamento de Laboratorio Biomédico, Instituto de Salud Pública de Chile (ISP), Santiago, Chile, **8** Centro para Vacunas en Desarrollo-Chile (CVD-Chile), Hospital de Niños Roberto del Río, Santiago, Chile

* These authors contributed equally to this work.

† MML and GD also contributed equally to this work.

* mm2032@cam.ac.uk



OPEN ACCESS

Citation: Maes M, Sikorski MJ, Carey ME, Higginson EE, Dyson ZA, Fernandez A, et al. (2022) Whole genome sequence analysis of *Salmonella* Typhi provides evidence of phylogenetic linkage between cases of typhoid fever in Santiago, Chile in the 1980s and 2010–2016. PLoS Negl Trop Dis 16(6): e0010178. <https://doi.org/10.1371/journal.pntd.0010178>

Editor: Travis J. Bourret, Creighton University, UNITED STATES

Received: January 19, 2022

Accepted: June 8, 2022

Published: June 29, 2022

Copyright: © 2022 Maes 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: Raw sequence data have been submitted to the European Nucleotide Archive (ENA) and accession numbers of individual isolates are listed in [S1 Table](#).

Funding: This work was supported by a grant [OPP1161058] (to MML) from the Bill & Melinda Gates Foundation, this work was also supported by the Gates Foundation (TyVAC) [OPP1151153] (to GD) and the Wellcome Trust (STRATAA) [106158]

Abstract

Typhoid fever epidemiology was investigated rigorously in Santiago, Chile during the 1980s, when *Salmonella enterica* serovar Typhi (*S. Typhi*) caused seasonal, hyperendemic disease. Targeted interventions reduced the annual typhoid incidence rates from 128–220 cases/10⁵ population occurring between 1977–1984 to <8 cases/10⁵ from 1992 onwards. As such, Santiago represents a contemporary example of the epidemiologic transition of an industrialized city from amplified hyperendemic typhoid fever to a period when typhoid is no longer endemic. We used whole genome sequencing (WGS) and phylogenetic analysis to compare the genotypes of *S. Typhi* cultured from acute cases of typhoid fever occurring in Santiago during the hyperendemic period of the 1980s ($n = 74$) versus the nonendemic 2010s ($n = 80$) when typhoid fever was rare. The genotype distribution between “historical” (1980s) isolates and “modern” (2011–2016) isolates was similar, with genotypes 3.5 and 2 comprising the majority of isolations, and 73/80 (91.3%) of modern isolates matching a genotype detected in the 1980s. Additionally, phylogenomically ‘ancient’ genotypes 1.1 and 1.2.1, uncommon in the global collections, were also detected in both eras, with a notable rise amongst the modern isolates. Thus, genotypes of *S. Typhi* causing acute illness in the modern nonendemic era match the genotypes circulating during the hyperendemic 1980s.

and 098051].(to GD) MM and GD received funding by the NIHR Cambridge Biomedical Research Centre and NIHR AMR Research Capital Funding Scheme [NIHR200640]. ZAD was supported by a grant funded by the Wellcome Trust (STRATAA; 106158/Z/14/Z), and received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement TyphiNET (#845681). MJS received research support from the NIH NIAID grants F30AI156973 and U19AI110820 as well as NIDDK T32DK067872. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

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

The persistence of historical genotypes may be explained by chronic typhoid carriers originally infected during or before the 1980s.

Author summary

Studies of *Salmonella* Typhi (the cause of typhoid fever) rarely include isolates collected both before and after the interruption of hyperendemic transmission because this typically occurred decades before modern bacteria preservation methods. After substantial reduction in disease, it was assumed that sporadic cases and infrequent outbreaks were due to either chronic biliary carriers or importations, but this was difficult to characterize with low resolution bacterial typing methods. In Santiago, Chile, typhoid fever persisted at hyperendemic levels through the 1980s until organized control efforts in the 1980s and changes to wastewater policy in 1991 caused annual typhoid incidence to plummet.

In this study, we used whole genome sequencing (WGS) to investigate whether recent sporadic cases occurring in Santiago in the 2010s were genomically similar to *S. Typhi* circulating in the 1980s, or dissimilar, possibly representing importations of *S. Typhi* from outside of Chile. We found concordance amongst *S. Typhi* genotypes between the 1980s and 2010s, and differences from genotypes circulating in Southeast Asia and Africa where typhoid remains hyperendemic. Our findings suggest that a proportion of modern, rare typhoid cases in Santiago are autochthonous, and that chronic carriers or another unknown reservoir likely contribute. Broadly, our findings corroborate the epidemiologic importance of long-term reservoirs of typhoid fever decades after typhoid elimination.

Introduction

Typhoid fever is a systemic infection, caused by the bacterium *Salmonella enterica* subspecies *enterica* serovar Typhi (*S. Typhi*), which is spread through ingestion of food and water contaminated with human faeces containing *S. Typhi*. Typhoid fever in Santiago, Chile rose to hyperendemic levels in the mid-1970s with a striking annual incidence of 177–220 cases per 10^5 population per year from 1977–1983, partially declined during the years 1984–1991 (51–128 cases/ 10^5 per year), and finally plummeted below 8 cases/ 10^5 per year after 1991, reaching nonendemic levels of $\sim 0.5/10^5$ in 2012 [1]. During the hyperendemic years, the highest incidence rate was observed in school age children, and case numbers peaked in the rainless summer months [2]. To address this public health issue, in 1979 the Ministry of Health of Chile implemented a Typhoid Fever Control Program involving strategic preventive interventions including large-scale field trials of live oral typhoid vaccine Ty21a in >500,000 schoolchildren [3–6], and identification and treatment of chronic carriers among food handlers [7]. Collectively, these interventions led to a steady decline in typhoid cases between 1983 and 1990 [3–6,8].

Epidemiologic and environmental bacteriology studies in the 1980s undertaken by the Chilean Typhoid Fever Control Program also implicated the use of raw untreated sewage water for irrigation of vegetable crops during the rainless summer months as a key risk factor responsible for amplified transmission of *S. Typhi* in Santiago [8,9]. In April 1991, the government of Chile abruptly prohibited the practice of irrigating crops with untreated sewage wastewater in Santiago [1,8,10,11]. This action was taken following the explosive epidemics of El Tor cholera in 1991 that occurred in neighbouring countries to the North, including in Peru, beginning in late January [12,13], followed by Ecuador in February [13], and Colombia in March [13]. The pivotal stimulus

for the intervention was a small outbreak of cholera in Santiago, Chile, in April, 1991 [1,8,10,11]. Following imposition and strict enforcement of the prohibition of use of untreated sewage water for irrigation, the Santiago cholera epidemic quickly came to an end [10,11], and the incidence of typhoid dropped precipitously in subsequent years [1,8,10]. From 2000–2014, the annual incidence of confirmed typhoid fever in Santiago has ranged from 0.20–1.18 cases/10⁵ population. The current rare sporadic cases of typhoid in Santiago include illnesses among: travellers who visit countries with high incidence rates of typhoid; immigrants from typhoid-endemic countries; and transmission from Chilean chronic carriers who were originally infected decades earlier (pre-1991) during the era of hyperendemic transmission.

Phylogenetic analysis of whole genome sequencing (WGS) data has become the gold standard for inferring relatedness between *S. Typhi* isolates [14]. WGS can also be used to assign genotypes, identify and characterize plasmids, compare gene content and homology, and identify molecular determinants of antibiotic resistance and virulence [15,16]. Whereas WGS data from many global isolates are available, few bacterial collections include isolates from both before and after the interruption of amplified hyperendemic typhoid fever. Also, WGS-derived information concerning the genotypes isolated from typhoid fever cases in South American countries is scarce. The majority of publicly available sequenced isolates come from hyperendemic regions in South Asia and East Africa where H58 (4.3.1) genotypes dominate [17].

Herein we describe a genomic analysis of *S. Typhi* isolates from Santiago, Chile from two distinct epidemiologic periods: the early 1980s (N = 74), when typhoid was hyperendemic, and from 2010–2016 (N = 80), a quarter century after the interruption of amplified transmission of *S. Typhi* in Santiago. We show that the genotypic composition of the population structure is similar between these two periods, even though the local epidemiology has changed drastically. In addition, we describe a potential carrier-related cluster of the “ancient” 1.1 clade, and the identification of a new variant of plasmid pHCM2 carried by 2 isolates of another ancient genotype (1.2.1).

Material and methods

Study design

In total, 74 “historical” isolates of *S. Typhi* from cases of acute typhoid fever in Santiago, Chile in the period 1981 to 1986, during the era of hyperendemic typhoid, were obtained from the culture collection of the Center for Vaccine Development and Global Health of the University of Maryland, Baltimore. “Modern” *S. Typhi* isolates (n = 80) were strains collected from acute cases of typhoid fever and referred to the Instituto de Salud Pública de Chile (ISP, Institute of Public Health) by clinical microbiology laboratories from 2011–2016; these referred strains were generally accompanied by only limited demographic and clinical information.

DNA extraction and sequencing

DNA was isolated from the 154 *S. Typhi* isolates using the Norgen DNA extraction kit as per the manufacturer’s instructions. Genomic DNA was sequenced by Illumina HiSeq at the Wellcome Sanger Institute, generating 125 bp paired reads. Raw read data were deposited in the European Nucleotide Archive and individual accession numbers are listed in [S1 Table](#).

Assembly, mapping, SNP-calling and genotype assignment

Raw Illumina reads were assembled using Velvet v1.2 via the Wellcome Sanger Institute automated analysis pipeline [18]. Sequenced reads were mapped and annotated, single nucleotide polymorphisms (SNPs) were called against the *S. Typhi* CT18 reference genome [19] using

SMALT v0.7.4 and subsequent variant calling was carried out as previously described [20]. Genotypes were assigned using GenoTyphi (v1.9.0) [14,16].

Phylogenetic analysis

Previously defined [21–23] recombinant regions such as prophages and plasmids were excluded manually, and any remaining recombinant regions were filtered using Gubbins (v2.5.0) [24]. The resultant core genome alignment was used to infer Maximum Likelihood (ML) phylogenies using RAxML (v8.2.8) [25], specifying 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 pseudoreplicates used to assess branch support. SNP distances for the core genome alignment of all the strains have also been calculated from this alignment using snp-dists package (available at: <https://github.com/tseemann/snp-dists>). Resulting phylogenies were visualized with iTOL [26] and are available for interactive inspection on Microreact [27] (<https://microreact.org/project/j25cGgoVX8ixQCRstPVhgo-chile-1980-2010>).

Molecular determination of AMR, plasmids and virulence genes

Raw read data were screened using SRST2 v0.2.0 [15] with the ARG-ANNOT [28], PlasmidFinder [29], and *Salmonella* Virulence Factor DB databases [30] to detect known and putative antimicrobial resistance (AMR) genes, known plasmid replicons, and known virulence-associated genes, respectively. Homology cutoffs of 90% nucleotide similarity were used for all SRST2 screens.

Plasmid pHCM2 sequences were assembled with Unicycler [31], and the circularised sequences extracted from *de bruijn* assembly graphs with Bandage (v0.8.1) [32]. These were then compared with the pHCM2 reference sequence (accession no: AL513384) using Artemis-ACT [33] and visualized with EasyFig [34].

Results

S. Typhi genotypes in Santiago, Chile

Genotypic analysis indicated that most of the isolates from Santiago in this study belonged to genotypes 2 (n = 43, 27.9%), 3.5 (n = 51, 33.1%) and 1.1 (n = 22, 14.3%) (Table 1). While these

Table 1. S. Typhi genotypes detected in isolates from acute clinical cases in Santiago, Chile during the hyperendemic 1980s versus the nonendemic 2010s.

	1983–1986	2011–2016	TOTAL
1.1	2 (2.7%)	20 (25%)	22 (14.3%)
1.2.1	1 (1.4%)	5 (6.3%)	6 (3.9%)
2	30 (40.5%)	13 (16.3%)	43 (27.9%)
2.0.2	11 (14.9%)	5 (6.3%)	16 (10.4%)
2.2	-	3 (3.8%)	3 (1.9%)
2.3.2	1 (1.4%)	1 (1.3%)	2 (1.3%)
2.3.3	1 (1.4%)	1 (1.3%)	2 (1.3%)
2.3.4	1 (1.4%)	-	1 (0.6%)
3.1	1 (1.4%)	-	1 (0.6%)
3.3	-	1 (1.3%)	1 (0.6%)
3.5	25 (33.8%)	26 (32.5%)	51 (33.1%)
4.1	1 (1.4%)	2 (2.5%)	3 (1.9%)
4.3.1.1	-	2 (2.5%)	2 (1.3%)
4.3.1.2	-	1 (1.3%)	1 (0.6%)
TOTAL	74	80	154

<https://doi.org/10.1371/journal.pntd.0010178.t001>

three genotypes were the most prevalent, we also detected 11 additional distinct genotypes across the two time periods. These include three previously described importations of isolates belonging to H58 genotypes 4.3.1.1 and 4.3.1.2 [35].

The genomic population structure of *S. Typhi* in Santiago appeared highly consistent between the hyperendemic and post-hyperendemic eras (Fig 1). Both rare “ancient” genotypes 1.1 and 1.2.1, as well as more commonly found genotypes 2 and 3.5, were frequently present across both study periods. Despite several decades having passed since the cessation of the era of hyperendemic transmission, most of the post-2010 sporadic *S. Typhi* isolates shared bootstrap-supported phylogenetic clusters with the 1980 hyperendemic isolates. Only within genotype 1.1 did recent and historical sequences form independent clusters. The intermingling of historical and recent sequences suggests that the 2010s isolates originated from progenitors within the same pool of *S. Typhi* circulating in the 1980s.

Amongst the 151 non-H58 isolates, few antimicrobial resistance mechanisms were detected. One genotype 2.0.0 isolate carried a *tetB* gene, known to encode a mechanism of resistance to tetracyclines (albeit not an antibiotic of clinical relevance for treatment of typhoid fever). Two genotype 2.2 isolates and one genotype 4.1 isolate had single non-synonymous point mutations in the Quinolone Resistance Determining Region (QRDR) of genes *gyrA* and *gyrB*, associated with reduced susceptibility to fluoroquinolones [36]; in the late 1980s, ciprofloxacin was the antibiotic of choice for treating patients with typhoid fever. Only 17 candidate plasmids were detected across all the isolates, including a Colicin plasmid, three related to plasmid pHCM2 and 13 of incompatibility type IncX1 (Fig 1).

Global transmission dynamics

To better understand how this local population structure fits into the global context of *S. Typhi*, a maximum likelihood phylogenetic tree was generated using a total of 2,013 non-H58 *S. Typhi* genomes from approximately 70 countries (Fig 2). The *S. Typhi* isolates from Santiago of genotypes 3.5, 2, 1.1 and 1.2.1 clustered together and were genetically distinct from those sequenced previously from other countries. The other genotypes present in Santiago were mostly found on long branches, indicating significant differentiation from their closest relatives in the global context phylogeny.

“Ancient” *S. Typhi* clades endemic to Chile

Two of the clades common in our dataset, 1.1 and 1.2, are found closer to the root of the global *S. Typhi* tree and thus can be considered “ancient” clades, which are detected with very low global frequency. Of the 1.1 isolates ($n = 22$), two were from the 1980s collection (Fig 3). Only six other isolates, 0.3% of the global collection described by Wong et al [14], belonged to clade 1.1. Interestingly, a cluster of closely related genotype 1.1 *S. Typhi* isolates was detected in the years 2011 (1 isolate) and 2012 (19 isolates). These isolates also exhibit lower genetic variation (median SNP distance of 8) than those observed among the two 1980s 1.1 isolates (119 SNPs). The smallest SNP distance between a 1980s 1.1 isolate and a 2010s 1.1 isolate was 79 SNPs. This suggests the 2011–2012 isolates have a recent common ancestor.

We sought to determine if there was a specific virulence gene associated with the 2011–2012 genotype 1.1 cluster, using SRST2 with the *Salmonella* VFDB. Although there were no additional or missing virulence factor-associated genes detected, this screening did reveal that all of the genotype 1.1 isolates from 2011–2012 cluster harboured a SNP in the *ssrA* gene. The alignment of this gene revealed the same missense mutation (G to A at position 1648968 of CT18, causing a glycine to glutamic acid amino acid change) in the *ssrA* gene (*ssrA*-G639E), (Fig 3).

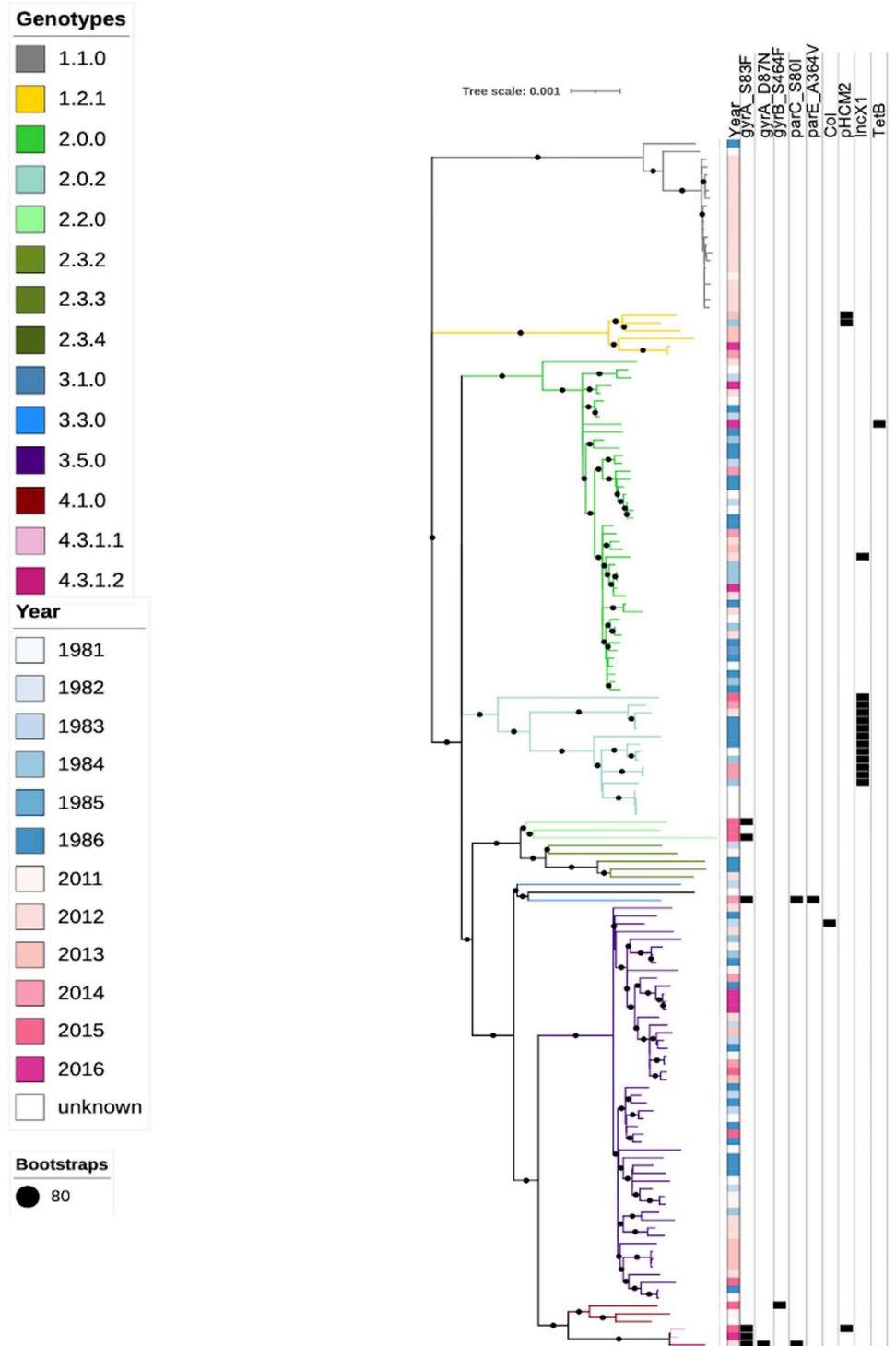


Fig 1. Genotypes and molecular determinates of AMR and plasmids in the Santiago collection. SNP based maximum likelihood core genome phylogeny indicating the genotypes of all the Santiago *S. Typhi* isolates is shown by the branch colour as per the inset legend. A diverse range of 14 genotypes were detected including 1.1, 1.2.1, 2, 2.0.2, 2.2, 2.3.2, 2.3.3, 2.3.4, 3.1, 3.3, 3.5, 4.1, 4.3.1.1, 4.3.1.2. Most of the isolates belonged to genotype 1.1, 2 and 3.5. Year of isolation, SNPs in the QRDR region associated with reduced susceptibility to fluoroquinolones as well as plasmids and detected resistance genes are shown in the columns adjacent to the tree. Bootstrap confidence of 80 or above is indicated with a black circle on the branch.

<https://doi.org/10.1371/journal.pntd.0010178.g001>

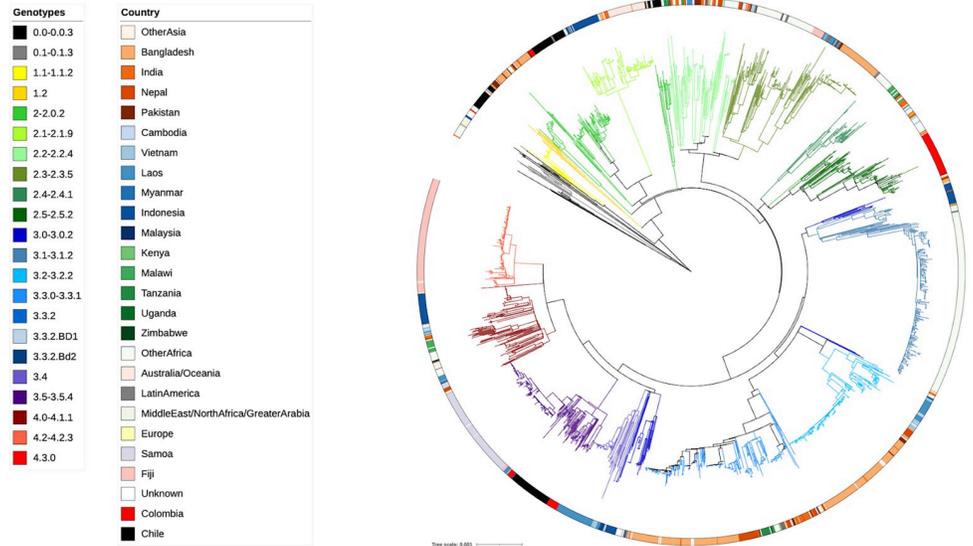


Fig 2. A global context maximum likelihood core genome phylogeny incorporating publicly available non-H58 and Santiago *S. Typhi* isolates. SNP based core genome maximum likelihood tree of 2013 global isolates, branches are coloured by genotype as per the inset legend; the ring indicates country of isolation as per the inset legend; Santiago isolates are highlighted in black and the Colombian isolates in red.

<https://doi.org/10.1371/journal.pntd.0010178.g002>

The other ancient clade present amongst the Santiago isolates was genotype 1.2.1. This was the only clade in our datasets (except from one recently imported H58 isolate) in which isolates harboured pHCM2 plasmids. pHCM2 plasmids are found in ~15% of *S. Typhi* and are

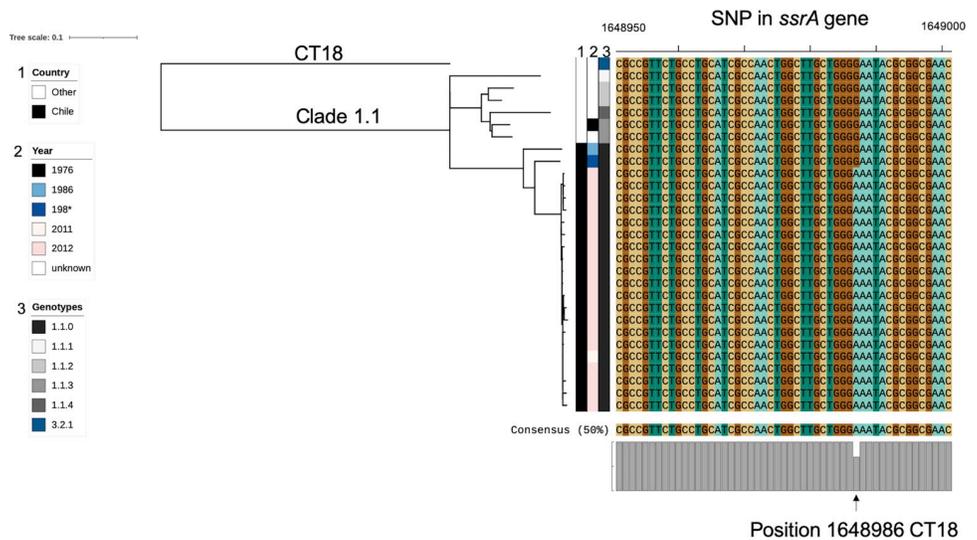


Fig 3. *S. Typhi* clade 1.1 phylogeny showing sequences carrying *ssrA* mutation (*ssrA*-G639E). Phylogenetic core genome analysis of all *S. Typhi* isolates of genotype 1.1 identified from the global collection. The majority of these isolates were collected in Santiago (black bar); two isolates belonged to the 1980s collection appearing ancestral to the 2010s collection. The twenty isolates from 2011–2021 showed an interesting cluster with a median SNP distance of 8 SNPs which is unlikely to have evolved among acute cases, where the mutation rate is calculated to be < 1 SNP/genome/year, in one year. Also shown is an alignment of a section of the *ssrA* gene, showing the conserved SNP at position 1648986 of the *S. Typhi* CT18 reference genome. * the precise isolation year of one of the 1980s isolates was unclear.

<https://doi.org/10.1371/journal.pntd.0010178.g003>

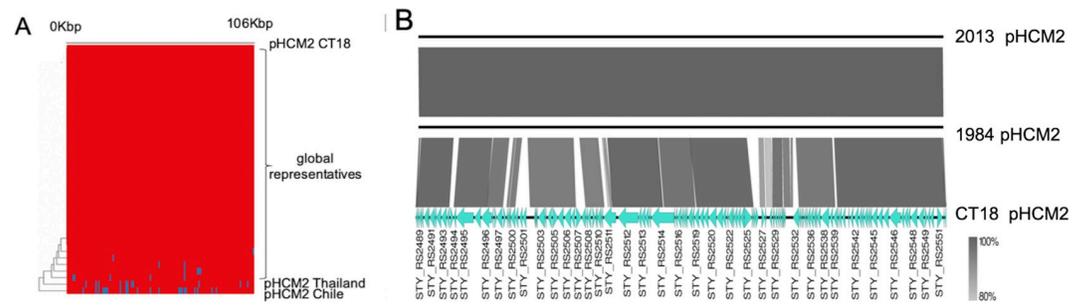


Fig 4. Genetic comparison of cryptic plasmid pHCM2-like sequences. Global representative collection of pHCM2 aligned to the reference pHCM2 from CT18. (A) Presence and absence of genes located on pHCM2. Red indicates presence of the genes located on the reference pHCM2 sequence from CT18, while blue indicates absence of the genes compared to the reference. Nucleotide BLAST comparison of the 2 Chilean assembled pHCM2-like plasmids (ERR271054, ERR4289185) against the reference pHCM2 plasmid from *S. Typhi* CT18 (accession no:AL513384) (B). Gene annotations are shown for the reference pHCM2 from CT18, with the shading intensity indicating percent nucleotide homology between the plasmid sequences.

<https://doi.org/10.1371/journal.pntd.0010178.g004>

generally highly conserved (Fig 4A). To our knowledge, based on the data available in the *S. Typhi* public collection on Pathogenwatch: (<https://pathogen.watch/collection/1y3cdsjq55hf-public-genomes>- accessed Dec 2021), pHCM2 plasmids have not been detected in ancient clades 0.0 to 1.2 previously. Two pHCM2 plasmids were found among our isolates from genotype 1.2.1; one in a 1984 isolate (ERR2710544) and the other from 2013 (ERR4289185). These plasmids were highly similar to each other, but significantly different from the pHCM2 found on *S. Typhi* reference CT18 (accession no: AL513384), and from another novel pHCM2 found in a 1973 isolate from Thailand [37]. The pHCM2 sequences described here differed by >2000 SNPs (98.4% nucleotide identity, 89% coverage) from the reference (Fig 4).

Discussion

This study utilizes WGS analyses to compare *S. Typhi* isolates from patients who developed acute typhoid fever in Santiago during the modern era (i.e., 2011–2016, decades post-interruption of amplified typhoid transmission in 1991) with isolates from acute cases during the hyper-endemic era of the 1980s. Using phylogenetics and genome-wide screening, we compared the genotypes, AMR determinants and plasmid content of *S. Typhi* isolates between these two time periods. Whilst comparisons of isolates from the 1980's with isolates from 1990 (post vaccinating interventions-but pre-irrigation ban, when incidences were still high 47 cases /10⁵ [1]) have been undertaken previously with pulse field gel electrophoresis (PFGE) [38], to our knowledge, this is the first WGS-based analysis comparing *S. Typhi* collected during a period of hyper-endemicity and decades after typhoid control.

The observation that the 1980s and 2010s collections of *S. Typhi* isolates do not cluster independently within our commonly found genotypes (2, 3.5 and 1.2.1) suggests that many of the sporadic cases detected in the 2010s share common ancestry with *S. Typhi* cases from the 1980s. These may, in turn, be linked to shedding of *S. Typhi* from chronic typhoid carriers. This is expected decades after hyper-endemic typhoid is controlled, as chronic carriers will continue to shed sporadically until they are treated or die out of the population [39].

Although there have been several detected introductions of the globally dominant H58 *S. Typhi* clade (4.3.1) into Santiago, as occurs with most industrialized countries in North America, Europe, and Australia, this was not a frequent genotype found in the sporadic modern Chilean cases. This is distinct from countries such as the United Kingdom where imported

cases of H58 *S. Typhi* have become common among travel-associated typhoid cases in recent years [40,41].

S. Typhi WGS data from a recent study in Colombia found genotypes 2, 2.5 and 3.5 to be the local endemic genotypes [42]. Genotype 2.5 was not detected among the Santiago isolates; however, genotype 3.5 and 2 were detected in Santiago, and were the nearest genomic neighbours of the Colombian isolates on the global tree. Even though the sequenced isolates from Colombia and Santiago, Chile both contained isolates from the 2010s and historical isolates, they clustered separately on the global tree, suggesting that there has not been obvious transmission between the two countries. More representative WGS data from other Central and South American countries are needed to determine if this is the case with other unsampled potentially endemic countries in the area.

One of the more interesting findings from this dataset was the cluster of 20 highly related strains of genotype 1.1 in 2011–2012. These isolates differed by a median of 8 SNPs in the core genome. For context, the average core genome mutation rate in *S. Typhi* is approximately 1 substitution/genome/year every year [21,43–45]. As such, it is unlikely that these isolates originated from a clonal source and sequentially acquired this genetic variation in <2 years. Conversely, these strains are clearly highly related and were all collected during a short time span. A potential explanation to account for this cluster is that it was an outbreak caused by bacteria shed from a long-term carrier. *S. Typhi* colonizing the gallbladder or other long-term physiologic niche of a chronic typhoid carrier may evolve and mutate over time into a heterogeneous, yet closely related, community of bacteria [46]. Thus, the closely related nature of these isolates occurring mainly over a single year and during a period when typhoid was no longer hyperendemic in Santiago could be explained by an outbreak instigated by a chronic carrier who was shedding multiple highly related strains. If this hypothetical carrier was working as a food handler, intermittently contaminating food or water consumed by subjects during 2011–12, this could result in the cluster of closely related isolates we observe. The cessation of cases could then be explained by a change of employment or residence, a cholecystectomy procedure, a course of antibiotics that concentrate in the gallbladder (e.g., ciprofloxacin) prescribed for another infection or surgery, or death of the hypothetical carrier at the end of 2012.

Notably, the only detectable difference in virulence factor-associated genes from this potential carrier-related 2011–2012 cluster is the *ssrA* missense mutation. The *ssrA/B* two component response regulator is a principle regulator of *Salmonella* pathogenicity island-2 (SPI-2) [47,48]. SPI-2 has been shown to be insignificant in *S. Typhi* macrophage survival capability and *S. Typhi*'s infection efficiency of humanized mice [49,50]. The *ssrA* gene has been linked to biofilm formation in *Salmonella* and persistence on gallstones [51,52]. This observation may be worth additional investigation to investigate the role of SPI-2 and *ssrA* in Typhoid carriage.

Our analysis identified a novel variant of cryptic plasmid pHCM2 in two sequences from genotype 1.2.1. A pHCM2 type plasmid was present in the *S. Typhi* reference isolate CT18 from Vietnam [19] and the plasmid family was described in some detail [53]. Interestingly, any role for pHCM2 in *S. Typhi* biology remains elusive. We know that almost all of the pHCM2 plasmids in the global *S. Typhi* collection are highly conserved in terms of overall nucleotide diversity. However, the Chilean pHCM2 from clade 1.2.1 is substantially different from most other family members. The only other reported pHCM2 plasmid with a similar level of variation is harboured in a *S. Typhi* Thailand isolate from 1973 [37].

The main limitation of this study is the limited sample size of *S. Typhi* genomes from both the historical hyperendemic and modern nonendemic periods in Santiago, Chile. Although not all isolates were stored during the hyperendemic period more than 40 years ago, we incorporated all available stored isolates from acute cases. For the modern time period we used a

similar number of available isolates exclusively from acute cases to represent circulating infection. Among the approximately 7 million inhabitants in Santiago, Chile, few blood culture isolates from acute cases of typhoid fever were recorded in the 2010s, ($<0.5/10^5$ cases per year) and not all isolates are sent to the ISP; thus, only a limited number of *S. Typhi* isolates were available for sequencing. Nevertheless, the study design, based on *S. Typhi* causing acute disease between epidemiologically-distinct eras, lends support to our primary finding: the same genotypes circulating in the hyperendemic 1980s are found circulating in the modern 2010s period.

There are especially limited genomic data available for *S. Typhi* circulating in neighbouring countries, presenting another limitation of this study. Nevertheless, efforts to foster sequencing infrastructure and data generated in South America through the PAHO regional networks have accelerated during the Covid-19 pandemic. It is expected that whole genome sequencing of other pathogens, like *S. Typhi*, will follow in this region.

In conclusion, this genomics analysis suggests that asymptomatic typhoid carriers (or other unknown reservoirs) are likely responsible for continuing sporadic typhoid cases in Chile. The ancient clades present in the country have persisted. We recommend that all sporadic typhoid cases and clusters should be evaluated using epidemiological outbreak investigations and WGS as routine surveillance. This will help ensure the detection of any potential introduction or development of drug-resistant *S. Typhi* and may lead to the identification and treatment of carriers, which would further limit ongoing disease transmission.

Supporting information

S1 Table. S1 Table includes from left to right: Id, lane accession, __latitude, __longitude, genotype, Ampicillin resistance, Cephalexin resistance, Chloramphenicol resistance, Ciprofloxacin resistance, Nalidixic acid resistance, Sulfamethoxazole resistance, Trimethoprim resistance, Trimethoprim/Sulfamethoxazole resistance, Tetracycline resistance, Azithromycin resistance, Colistin resistance, Meropenem resistance, *gyrA*_D87N mutation, *gyrB*_S464F mutation, *parC*_S80I mutation, *perE*_A364V mutation, presence of *tetA* (B) gene, presence of *tetR* (B) gene, presence of pHCM2 plasmid (any variant), year of isolation, *ssrA*_G639E mutation, presence of pHCM2 Chile variant. (CSV)

Acknowledgments

We express our thanks to the typhoid fever patients whose samples were analysed and included in this paper, as well as the local public health personnel in Chile. Furthermore, we wish to thank Nicholas Thomson for access to the Sanger analysis pipelines.

Author Contributions

Conceptualization: Rosanna Lagos, Myron M. Levine.

Data curation: Mailis Maes, Michael J. Sikorski, Rosanna Lagos, Juan Carlos Hormazábal.

Formal analysis: Mailis Maes, Michael J. Sikorski.

Funding acquisition: Myron M. Levine, Gordon Dougan.

Investigation: Mailis Maes, Megan E. Carey, Ellen E. Higginson, Alda Fernandez, Pamela Araya, Rosanna Lagos, Juan Carlos Hormazábal.

Methodology: Mailis Maes.

Resources: Zoe A. Dyson, Sharon M. Tennant, Stephen Baker, Myron M. Levine, Gordon Dougan.

Writing – original draft: Mailis Maes, Michael J. Sikorski.

Writing – review & editing: Mailis Maes, Michael J. Sikorski, Megan E. Carey, Ellen E. Higinson, Zoe A. Dyson, Sharon M. Tennant, Stephen Baker, Rosanna Lagos, Juan Carlos Hormazábal, Myron M. Levine, Gordon Dougan.

References

1. Marco C, Delgado I, Vargas C, Munoz X, Bhutta ZA, Ferreccio C. Typhoid Fever in Chile 1969–2012: Analysis of an Epidemic and Its Control. *Am J Trop Med Hyg.* 2018; 99(3_Suppl):26–33. <https://doi.org/10.4269/ajtmh.18-0125> PMID: 30047371
2. Levine MM BR, Ferreccio C, Clements ML, Lanata C, Sear S, Morris JG, et al. Interventions to control endemic typhoid fever: field studies in Santiago, Chile. *Control and Eradication of Infectious Diseases, An International Symposium.* 1985;Series 1:37–53.
3. Black RE, Levine MM, Ferreccio C, Clements ML, Lanata C, Rooney J, et al. Efficacy of one or two doses of Ty21a *Salmonella typhi* vaccine in enteric-coated capsules in a controlled field trial. *Vaccine.* 1990; 8(1):81–4. [https://doi.org/10.1016/0264-410x\(90\)90183-m](https://doi.org/10.1016/0264-410x(90)90183-m) PMID: 2180234
4. Levine MM, Ferreccio C, Ortiz E, Cryz S. Comparison of enteric-coated capsules and liquid formulation of Ty21a typhoid vaccine in randomised controlled field trial. *The Lancet.* 1990; 336(8720):891–4. [https://doi.org/10.1016/0140-6736\(90\)92266-k](https://doi.org/10.1016/0140-6736(90)92266-k) PMID: 1976928
5. Levine MM, Ferreccio C, Black RE, Germanier R. Large-scale field trial of Ty21a live oral typhoid vaccine in enteric-coated capsule formulation. *Lancet.* 1987; 1(8541):1049–52. [https://doi.org/10.1016/s0140-6736\(87\)90480-6](https://doi.org/10.1016/s0140-6736(87)90480-6) PMID: 2883393
6. Ferreccio C, Levine MM, Rodriguez H, Contreras R. Comparative efficacy of two, three, or four doses of TY21a live oral typhoid vaccine in enteric-coated capsules: a field trial in an endemic area. *J Infect Dis.* 1989; 159(4):766–9. <https://doi.org/10.1093/infdis/159.4.766> PMID: 2647863
7. Ferreccio C, Morris JG Jr., Valdivieso C, Prenzel I, Sotomayor V, Drusano GL, et al. Efficacy of ciprofloxacin in the treatment of chronic typhoid carriers. *J Infect Dis.* 1988; 157(6):1235–9. <https://doi.org/10.1093/infdis/157.6.1235> PMID: 3373023
8. Gauld JS, Hu H, Klein DJ, Levine MM. Typhoid fever in Santiago, Chile: Insights from a mathematical model utilizing venerable archived data from a successful disease control program. *PLoS Negl Trop Dis.* 2018; 12(9):e0006759. <https://doi.org/10.1371/journal.pntd.0006759> PMID: 30188904
9. Sears SD, Ferreccio C, Levine MM, Cordano AM, Monreal J, Black RE, et al. The use of Moore swabs for isolation of *Salmonella typhi* from irrigation water in Santiago, Chile. *J Infect Dis.* 1984; 149(4):640–2. <https://doi.org/10.1093/infdis/149.4.640> PMID: 6373964
10. Levine M. South America: The return of cholera. *The Lancet.* 1991; 338(8758):45–6. [https://doi.org/10.1016/0140-6736\(91\)90024-J](https://doi.org/10.1016/0140-6736(91)90024-J)
11. Medina E. [Epidemic of cholera in Chile; 1991]. *Rev Med Chil.* 1991; 119(8):943–56. PMID: 1844781
12. Swerdlow DL, Mintz ED, Rodriguez M, Tejada E, Ocampo C, Espejo L, et al. Waterborne transmission of epidemic cholera in Trujillo, Peru: lessons for a continent at risk. *Lancet.* 1992; 340(8810):28–33. [https://doi.org/10.1016/0140-6736\(92\)92432-f](https://doi.org/10.1016/0140-6736(92)92432-f) PMID: 1351608
13. Tauxe RV, Blake PA. Epidemic cholera in Latin America. *Jama.* 1992; 267(10):1388–90. PMID: 1740864
14. 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):1–11. <https://doi.org/10.1038/ncomms12827> PMID: 27703135
15. Inouye M, Dashnow H, Raven LA, Schultz MB, Pope BJ, Tomita T, et al. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med.* 2014; 6(11):90. <https://doi.org/10.1186/s13073-014-0090-6> PMID: 25422674
16. Dyson ZA, Holt KE. Five years of GenoTyphi: updates to the global *Salmonella* Typhi genotyping framework. *The Journal of Infectious Diseases.* 2021. <https://doi.org/10.1093/infdis/jiab414> PMID: 34453548
17. Argimon S, Yeats CA, Goater RJ, Abudahab K, Taylor B, Underwood A, et al. A global resource for genomic predictions of antimicrobial resistance and surveillance of *Salmonella* Typhi at pathogenwatch. *Nat Commun.* 2021; 12(1):2879. <https://doi.org/10.1038/s41467-021-23091-2> PMID: 34001879

18. Page AJ, De Silva N, Hunt M, Quail MA, Parkhill J, Harris SR, et al. Robust high-throughput prokaryote de novo assembly and improvement pipeline for Illumina data. *Microb Genom.* 2016; 2(8):e000083. <https://doi.org/10.1099/mgen.0.000083> PMID: 28348874
19. 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. <https://doi.org/10.1038/35101607> PMID: 11677608
20. Kariuki S, Mbae C, Van Puyvelde S, Onsare R, Kawai S, Wairimu C, et al. High relatedness of invasive multi-drug resistant non-typhoidal *Salmonella* genotypes among patients and asymptomatic carriers in endemic informal settlements in Kenya. *PLoS Negl Trop Dis.* 2020; 14(8):e0008440. <https://doi.org/10.1371/journal.pntd.0008440> PMID: 32745137
21. Britto CD, Dyson ZA, Duchene S, Carter MJ, Gurung M, Kelly DF, et al. Laboratory and molecular surveillance of paediatric typhoidal *Salmonella* in Nepal: Antimicrobial resistance and implications for vaccine policy. *PLoS neglected tropical diseases.* 2018; 12(4):e0006408. <https://doi.org/10.1371/journal.pntd.0006408> PMID: 29684021
22. Rahman SIA, Dyson ZA, Klemm EJ, Khanam F, Holt KE, Chowdhury EK, et al. Population structure and antimicrobial resistance patterns of *Salmonella* Typhi isolates in urban Dhaka, Bangladesh from 2004 to 2016. *PLoS neglected tropical diseases.* 2020; 14(2):e0008036. <https://doi.org/10.1371/journal.pntd.0008036> PMID: 32106221
23. Wong VK, Holt KE, Okoro C, Baker S, Pickard DJ, Marks F, et al. Molecular Surveillance Identifies Multiple Transmissions of Typhoid in West Africa. *PLoS Negl Trop Dis.* 2016; 10(9):e0004781. <https://doi.org/10.1371/journal.pntd.0004781> PMID: 27657909
24. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* 2015; 43(3):e15. <https://doi.org/10.1093/nar/gku1196> PMID: 25414349
25. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* 2006; 22(21):2688–90. <https://doi.org/10.1093/bioinformatics/btl446> PMID: 16928733
26. Letunic I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics.* 2007; 23(1):127–8. <https://doi.org/10.1093/bioinformatics/btl529> PMID: 17050570
27. Argimón S, Abudahab K, Goater RJE, Fedosejev A, Bhai J, Glasner C, et al. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb Genom.* 2016; 2(11):e000093. <https://doi.org/10.1099/mgen.0.000093> PMID: 28348833
28. Gupta SK, Padmanabhan BR, Diene SM, 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(1):212–20. <https://doi.org/10.1128/AAC.01310-13> PMID: 24145532
29. Carattoli A, Zankari E, Garcia-Fernandez A, Voldby 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(7):3895–903. <https://doi.org/10.1128/AAC.02412-14> PMID: 24777092
30. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, et al. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res.* 2005; 33(Database issue):D325–8. <https://doi.org/10.1093/nar/gki008> PMID: 15608208
31. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol.* 2017; 13(6):e1005595. <https://doi.org/10.1371/journal.pcbi.1005595> PMID: 28594827
32. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics.* 2015; 31(20):3350–2. <https://doi.org/10.1093/bioinformatics/btv383> PMID: 26099265
33. Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. ACT: the Artemis Comparison Tool. *Bioinformatics.* 2005; 21(16):3422–3. <https://doi.org/10.1093/bioinformatics/bti553> PMID: 15976072
34. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics.* 2011; 27(7):1009–10. <https://doi.org/10.1093/bioinformatics/btr039> PMID: 21278367
35. Maes M, Dyson ZA, Higginson EE, Fernandez A, Araya P, Tennant SM, et al. Multiple Introductions of *Salmonella enterica* Serovar Typhi H58 with Reduced Fluoroquinolone Susceptibility into Chile. *Emerg Infect Dis.* 2020; 26(11):2736–40. <https://doi.org/10.3201/eid2611.201676> PMID: 33079054
36. Dyson ZA, Klemm EJ, Palmer S, Dougan G. Antibiotic resistance and typhoid. *Clinical Infectious Diseases.* 2019; 68(Supplement_2):S165–S70. <https://doi.org/10.1093/cid/ciy1111> PMID: 30845331
37. Dyson ZA, Thanh DP, Bodhidatta L, Mason CJ, Srijan A, Rabaa MA, et al. Whole genome sequence analysis of *Salmonella* Typhi isolated in Thailand before and after the introduction of a national

- immunization program. PLoS neglected tropical diseases. 2017; 11(1):e0005274. <https://doi.org/10.1371/journal.pntd.0005274> PMID: 28060810
38. Fica AE, Prat-Miranda S, Fernandez-Ricci A, D'Ottone K, Cabello FC. Epidemic typhoid in Chile: analysis by molecular and conventional methods of *Salmonella typhi* strain diversity in epidemic (1977 and 1981) and nonepidemic (1990) years. J Clin Microbiol. 1996; 34(7):1701–7. <https://doi.org/10.1128/JCM.34.7.1701-1707.1996> PMID: 8784573
 39. Ames WR, Robins M. Age and Sex as Factors in the Development of the Typhoid Carrier State, and a Method for Estimating Carrier Prevalence. Am J Public Health Nations Health. 1943; 33(3):221–30. <https://doi.org/10.2105/ajph.33.3.221> PMID: 18015749
 40. Chattaway MA, Gentle A, Nair S, Tingley L, Day M, Mohamed I, et al. Phylogenomics and antimicrobial resistance of *Salmonella* Typhi and Paratyphi A, B and C in England, 2016–2019. Microb Genom. 2021; 7(8). <https://doi.org/10.1099/mgen.0.000633> PMID: 34370659
 41. 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 Negl Trop Dis. 2019; 13(9):e0007620. <https://doi.org/10.1371/journal.pntd.0007620> PMID: 31513580
 42. Guevara PD, Maes M, Thanh DP, Duarte C, Rodriguez EC, Montañó LA, et al. A genomic snapshot of *Salmonella enterica* serovar Typhi in Colombia. PLoS Negl Trop Dis. 2021; 15(9):e0009755. <https://doi.org/10.1371/journal.pntd.0009755> PMID: 34529660
 43. 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. <https://doi.org/10.1038/ng.3281> PMID: 25961941
 44. Dyson ZA, Malau E, Horwood PF, Ford R, Siba V, Yoannes M, et al. Whole genome sequence analysis of *Salmonella* Typhi in Papua New Guinea reveals an established population of genotype 2.1.7 sensitive to antimicrobials. medRxiv. 2021:2021.09.27.21264209.
 45. Kariuki S, Dyson ZA, Mbae C, Ngetich R, Kawai SM, Wairimu C, et al. Multiple introductions of multi-drug-resistant typhoid associated with acute infection and asymptomatic carriage, Kenya. Elife. 2021;10. <https://doi.org/10.7554/eLife.67852> PMID: 34515028
 46. Thanh Duy P, Thieu NTV, Nguyen Thi Nguyen T, Ngoc Dan Thanh H, Dongol S, Karkey A, et al. Gallbladder carriage generates genetic variation and genome degradation in *Salmonella* Typhi. PLoS Pathog. 2020; 16(10):e1008998. <https://doi.org/10.1371/journal.ppat.1008998> PMID: 33085725
 47. Johnson R, Mylona E, Frankel G. Typhoidal *Salmonella*: Distinctive virulence factors and pathogenesis. Cell Microbiol. 2018; 20(9):e12939. <https://doi.org/10.1111/cmi.12939> PMID: 30030897
 48. Murret-Labarthe C, Kerhoas M, Dufresne K, Daigle F. New Roles for Two-Component System Response Regulators of *Salmonella enterica* Serovar Typhi during Host Cell Interactions. Microorganisms. 2020; 8(5). <https://doi.org/10.3390/microorganisms8050722> PMID: 32413972
 49. Forest CG, Ferraro E, Sabbagh SC, Daigle F. Intracellular survival of *Salmonella enterica* serovar Typhi in human macrophages is independent of *Salmonella* pathogenicity island (SPI)-2. Microbiology (Reading). 2010; 156(Pt 12):3689–98. <https://doi.org/10.1099/mic.0.041624-0> PMID: 20817644
 50. Karlinsey JE, Stepien TA, Mayho M, Singletary LA, Bingham-Ramos LK, Brehm MA, et al. Genome-wide Analysis of *Salmonella enterica* serovar Typhi in Humanized Mice Reveals Key Virulence Features. Cell Host Microbe. 2019; 26(3):426–34.e6. <https://doi.org/10.1016/j.chom.2019.08.001> PMID: 31447308
 51. Desai SK, Winardhi RS, Periasamy S, Dykas MM, Jie Y, Kenney LJ. The horizontally-acquired response regulator SsrB drives a *Salmonella* lifestyle switch by relieving biofilm silencing. Elife. 2016;5. <https://doi.org/10.7554/eLife.10747> PMID: 26880544
 52. Marshall JM, Flechtner AD, La Perle KM, Gunn JS. Visualization of extracellular matrix components within sectioned *Salmonella* biofilms on the surface of human gallstones. PLoS One. 2014; 9(2):e89243. <https://doi.org/10.1371/journal.pone.0089243> PMID: 24551241
 53. Kidgell C, Pickard D, Wain J, James K, Diem Nga LT, Diep TS, et al. Characterisation and distribution of a cryptic *Salmonella typhi* plasmid pHCM2. Plasmid. 2002; 47(3):159–71. [https://doi.org/10.1016/S0147-619X\(02\)00013-6](https://doi.org/10.1016/S0147-619X(02)00013-6) PMID: 12151231