

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

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

Short title: Lingering *Salmonella* Typhi genotypes in Santiago, Chile

Mailis Maes^{1¶}, Michael J. Sikorski^{2,3¶}, Megan E. Carey¹, Ellen E. Higginson¹, Zoe A. Dyson^{1,4,5,8}, Alda Fernandez⁶, Pamela Araya⁶, Sharon M. Tennant², Stephen Baker¹, Rosanna Lagos⁷, Juan Carlos Hormazábal⁶, Myron M. Levine^{2&}, and Gordon Dougan^{1, 8 &}

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 Bacteriología, Subdepartamento de Enfermedades Infecciosas, Departamento de Laboratorio Biomédico, Instituto de Salud Pública de Chile (ISP), Santiago, Chile

⁷ Centro para Vacunas en Desarrollo-Chile (CVD-Chile), Hospital de Niños Roberto del Río, Santiago, Chile

⁸ Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, United Kingdom

Corresponding author:

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
E-mail: mm2032@cam.ac.uk

37

38 ¶These authors contributed equally to this work.

39 &These authors contributed equally to this work.

40

41 **Keywords:** *S. Typhi*, Chile, phylogenetics, historical lineages, South America

42

43 **Abstract**

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

62

63 **Author summary**

64 Studies of *Salmonella* Typhi (the cause of typhoid fever) rarely include isolates collected
65 both before and after the interruption of hyperendemic transmission because this typically
66 occurred decades before modern bacteria preservation methods. After interruption, it was
67 assumed that sporadic cases and infrequent outbreaks were due to either chronic biliary
68 carriers or importations, but this was difficult to characterize with low resolution bacterial
69 typing methods. In Santiago, Chile, typhoid fever persisted at hyperendemic levels through
70 the 1980s until organized control efforts in the 1980s and changes to wastewater policy in
71 1991 caused annual typhoid incidence to plummet.

72 In this study, we used whole genome sequencing (WGS) to investigate whether recent
73 sporadic cases occurring in Santiago in the 2010s were genomically similar to *S. Typhi*
74 circulating in the 1980s, or dissimilar, possibly representing importations of *S. Typhi* from
75 outside of Chile. We found concordance amongst *S. Typhi* genotypes between the 1980s and
76 2010s, and differences from genotypes circulating in Southeast Asia and Africa where
77 typhoid remains hyperendemic. Our findings suggest that a proportion of modern, rare
78 typhoid cases in Santiago are autochthonous, and that chronic carriers or another unknown
79 reservoir likely contribute. Broadly, our findings corroborate the epidemiologic importance
80 of long-term reservoirs of typhoid fever decades after typhoid elimination.

81

82 Introduction

83 Typhoid fever is a systemic infection, caused by the bacterium *Salmonella enterica*
84 subspecies *enterica* serovar Typhi (*S. Typhi*), which is spread through ingestion of food and
85 water contaminated with human faeces containing *S. Typhi*. Typhoid fever in Santiago, Chile
86 rose to hyper-endemic levels in the mid-1970s with a striking annual incidence of 177-220
87 cases per 10⁵ population per year from 1977-1983, partially declined during the years 1984-
88 1991 (51-128 cases/10⁵ per year), and finally plummeted below 8 cases/10⁵ per year after
89 1991, reaching nonendemic levels of ~0.5/10⁵ in 2012 [1]. During the hyperendemic years,
90 the highest incidence rate was observed in school age children, and case numbers peaked in
91 the rainless summer months [2]. To address this public health issue, in 1979 the Ministry of
92 Health of Chile implemented a Typhoid Fever Control Program involving strategic
93 preventive interventions including large-scale field trials of live oral typhoid vaccine Ty21a
94 in >500,000 schoolchildren [3-6], and identification and treatment of chronic carriers among
95 food handlers [7]. Collectively, these interventions led to a steady decline in typhoid cases
96 between 1983 and 1990 [3-6, 8].

97
98 Epidemiologic and environmental bacteriology studies in the 1980s undertaken by the
99 Chilean Typhoid Fever Control Program also incriminated the use of raw untreated sewage
100 water for irrigation of vegetable crops during the rainless summer months as a key risk factor
101 responsible for amplified transmission of *S. Typhi* in Santiago [8, 9]. In April 1991, the
102 government of Chile abruptly prohibited the practice of irrigating crops with untreated
103 sewage wastewater in Santiago [1, 8, 10, 11]. This action was taken following the enormous
104 explosive epidemics of El Tor cholera in 1991 that occurred in neighbouring countries to the
105 North, including in Peru, beginning in late January [12, 13], followed by Ecuador in February
106 [13], and Colombia in March [13]. The pivotal stimulus for the intervention was a small

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

116

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

127

128 Herein we describe a genomic analysis of *S. Typhi* isolates from Santiago, Chile from two
129 distinct epidemiologic periods including the early 1980s (N=74), when typhoid was
130 hyperendemic, and from 2010-2016 (N=80), a quarter century after the interruption of
131 amplified transmission of *S. Typhi* in Santiago. We show that the genotypic composition of

132 the population structure is similar between these two periods, even though the local
133 epidemiology has changed drastically. In addition, we describe a potential carrier-related
134 cluster of the “ancient” 1.1 clade, and the identification of a new variant of plasmid pHCM2
135 carried by 2 isolates of another ancient genotype (1.2.1).

136 **Material and methods**

137

138 **Study design**

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

147

148 **DNA extraction and sequencing**

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

153

154 **Assembly, mapping, SNP-calling and genotype assignment**

155 Raw Illumina reads were assembled using Velvet v1.2 via the Wellcome Sanger Institute
156 automated analysis pipeline [18]. Sequenced reads were mapped and annotated, single
157 nucleotide polymorphisms (SNPs) were called against the *S. Typhi* CT18 reference genome
158 [19] using SMALT v0.7.4 and subsequent variant calling was carried out as previously
159 described [20]. Genotypes were assigned using GenoTyphi (v1.9.0) [14, 16].

160

161 **Phylogenetic analysis**

162 Previously defined [21-23] recombinant regions such as prophages and plasmids were
163 excluded manually, and any remaining recombinant regions were filtered using Gubbins
164 (v2.5.0) [24]. The resultant core genome alignment was used to infer Maximum Likelihood
165 (ML) phylogenies using RAxML (v8.2.8) [25], specifying a generalized time-reversible
166 model and a gamma distribution to model site-specific rate variation (GTR+ Γ substitution
167 model; GTRGAMMA in RAxML) with 100 bootstrap pseudoreplicates used to assess branch
168 support. SNP distances for the core genome alignment of all the strains have also been
169 calculated from this alignment using snp-dists package (available at:
170 <https://github.com/tseemann/snp-dists>). Resulting phylogenies were visualized with iTOL
171 [26] and are available for interactive inspection on Microreact [27]
172 (<https://microreact.org/project/j25cGgoVX8ixQCRstPVhgo-chile-1980-2010>).

173

174 **Molecular determination of AMR, plasmids and virulence genes**

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

180

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

185 Results

186 *S. Typhi* genotypes in Santiago, Chile

187 Genotypic analysis indicated that most of the isolates from Santiago in this study belonged to
 188 genotypes 2 (n=43, 27.9%), 3.5 (n=51, 33.1%) and 1.1 (n=22, 14.3%) (Table 1). While these
 189 three genotypes were the most prevalent, we also detected 11 additional distinct genotypes
 190 across the two time periods. These include three previously described importations of isolates
 191 belonging to H58 genotypes 4.3.1.1 and 4.3.1.2 [35].

192

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

193

194 The genomic population structure of *S. Typhi* in Santiago appeared highly consistent between
 195 the hyperendemic and post-hyperendemic eras (Fig 1). Both rare “ancient” genotypes 1.1 and
 196 1.2.1, as well as more commonly found genotypes 2 and 3.5, were frequently present across
 197 both study periods. Despite several decades having passed since the cessation of the era of
 198 hyperendemic transmission, most of the post-2010 sporadic *S. Typhi* isolates shared

199 bootstrap-supported phylogenetic clusters with the 1980 hyperendemic isolates. Only within
200 genotype 1.1 did recent and historical sequences form independent clusters. The
201 intermingling of historical and recent sequences suggests that the 2010s isolates originated
202 from progenitors within the same pool of *S. Typhi* circulating in the 1980s.
203
204 Amongst the 151 non-H58 isolates, few antimicrobial resistance mechanisms were detected.
205 One genotype 2.0.0 isolate carried a *tetB* gene, known to encode a mechanism of resistance to
206 tetracyclines (albeit not an antibiotic of clinical relevance for treatment of typhoid fever).
207 Two genotype 2.2 isolates and one genotype 4.1 isolate had single non-synonymous point
208 mutations in the Quinolone Resistance Determining Region (QRDR) of genes *gyrA* and *gyrB*,
209 associated with reduced susceptibility to fluoroquinolones [36]; in the late 1980s,
210 ciprofloxacin was the antibiotic of choice for treating patients with typhoid fever. Only 17
211 candidate plasmids were detected across all the isolates, including a Colicin plasmid, three
212 related to plasmid pHCM2 and 13 of incompatibility type IncX1 (Fig 1).

213

214 **Global transmission dynamics**

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

222

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

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

234

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

242

243 The other ancient clade present amongst the Santiago isolates was genotype 1.2.1. This was
244 the only clade in our datasets (except from one recently imported H58 isolate) in which
245 isolates harboured pHCM2 plasmids. pHCM2 plasmids are found in ~15% of *S. Typhi* and
246 are generally highly conserved (Fig 4A). To our knowledge based on the data available in the
247 *S. Typhi* public collection on Pathogenwatch:

248 (<https://pathogen.watch/collection/1y3cdsjq55hf-public-genomes>- accessed Dec 2021)

249 pHCM2 plasmids have not been detected in ancient clades 0.0 to 1.2 previously. Two
250 pHCM2 plasmids were found among our isolates from genotype 1.2.1; one in a 1984 isolate
251 (ERR2710544) and the other from 2013 (ERR4289185). These plasmids were highly similar
252 to each other, but significantly different to the pHCM2 found on *S. Typhi* reference CT18
253 (accession no: AL513384), and to another novel pHCM2 found in a 1973 isolate from
254 Thailand [37]. The pHCM2 sequences described here differed by >2000 SNPs (98.4%
255 nucleotide identity, 89% coverage) from the reference (Fig 4).

256

257 **Discussion**

258 This study utilizes WGS analyses to compare *S. Typhi* isolates from patients who developed
259 acute typhoid fever in Santiago during the modern era (i.e., 2011-2016, decades post-
260 interruption of amplified typhoid transmission in 1991) with isolates from acute cases during
261 the hyper-endemic era of the 1980s. Using phylogenetics and genome-wide screening, we
262 compared the genotypes, AMR determinants, plasmid content of *S. Typhi* isolates between
263 these two time periods. Whilst comparisons of isolates from the 1980's with isolates from
264 1990 (post vaccinating interventions-but pre-irrigation ban, when incidences were still high
265 47 cases /10⁵ [1]) have been undertaken previously with pulse field gel electrophoresis
266 (PFGE) [38], to our knowledge, this is the first WGS-based analysis comparing *S. Typhi*
267 collected during a period of hyper-endemicity and decades after typhoid control.

268

269 The observation that the 1980s and 2010s collections of *S. Typhi* isolates do not cluster
270 independently within our commonly found genotypes (2, 3.5 and 1.2.1) suggests that many of
271 the sporadic cases detected in the 2010s share common ancestry with *S. Typhi* cases from the
272 1980s. These may, in turn, be linked to shedding of *S. Typhi* from chronic typhoid carriers.

273 This is expected decades after hyper-endemic typhoid is controlled, as chronic carriers will
274 continue to shed sporadically until they are treated or die out of the population [39].

275

276 Although there have been several detected introductions of the globally dominant H58 *S.*
277 Typhi clade (4.3.1) into Santiago, as occurs with most industrialized countries in North
278 America, Europe, and Australia, this was not a frequent genotype found in the sporadic
279 modern Chilean cases. This is distinct from countries such as the United Kingdom where
280 imported cases of H58 *S. Typhi* have become common among travel-associated typhoid cases
281 in recent years [40, 41].

282

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

292

293 One of the more interesting findings from this dataset was the cluster of 20 highly related
294 strains of genotype 1.1 in 2011-2012. These isolates differed by a median of 8 SNPs in the
295 core genome. For context, the average core genome mutation rate in *S. Typhi* is
296 approximately 1 substitution/genome/year every year [21, 43-45]. As such, it is unlikely that
297 these isolates originated from a clonal source and sequentially acquired this genetic variation

298 in <2 years. Conversely, these strains are clearly highly related and were all collected during
299 a short time span. A potential explanation to account for this cluster is that it was an outbreak
300 caused by bacteria shed from a long-term carrier. *S. Typhi* colonizing the gallbladder or other
301 long-term physiologic niche of a chronic typhoid carrier may evolve and mutate over time
302 into a heterogeneous, yet closely related, community of bacteria [46]. Thus, the closely
303 related nature of these isolates occurring mainly over a single year and during a period when
304 typhoid was no longer hyperendemic in Santiago could be explained by an outbreak
305 instigated by a chronic carrier who was shedding multiple highly related strains. If this
306 hypothetical carrier was working as a food handler, intermittently contaminating food
307 vehicles consumed by subjects during 2011-12, this could result in the cluster of closely
308 related isolates we observe. The cessation of cases could then be explained by a change of
309 employment or residence, a cholecystectomy procedure, a course of antibiotics that
310 concentrate in the gallbladder (e.g., ciprofloxacin) prescribed for another infection or surgery,
311 or death of the hypothetical carrier at the end of 2012.

312
313 Notably, the only detectable difference in virulence factor-associated genes from this
314 potential carrier-related 2011-2012 cluster is the *ssrA* missense mutation . The *ssrA/B* two
315 component response regulator is a principle regulator of *Salmonella* pathogenicity island-2
316 (SPI-2) [47, 48]. SPI-2 has been shown to be insignificant in *S. Typhi* macrophage survival
317 capability and *S. Typhi*'s infection efficiency of humanized mice [49, 50]. The *ssrA* gene has
318 been linked to biofilm formation in *Salmonella* and persistence on gallstones [51, 52]. This
319 observation may be worth additional investigation to investigate the role of SPI-2 and *ssrA* in
320 Typhoid carriage.

321

322 Our analysis identified a novel variant of cryptic plasmid pHCM2 in two sequences from
323 genotype 1.2.1. A pHCM2 type plasmid was present in the *S. Typhi* reference isolate CT18
324 from Vietnam [53] and the plasmid family was described in some detail [54]. Interestingly,
325 any role for pHCM2 in *S. Typhi* biology remains elusive. We know that almost all of the
326 pHCM2 plasmids in the global *S. Typhi* collection are highly conserved in terms of overall
327 nucleotide diversity. However, the Chilean pHCM2 from clade 1.2.1 is substantially different
328 from most other family members. The only other reported pHCM2 plasmid with a similar
329 level of variation is harboured in a *S. Typhi* Thailand isolate from 1973 [55].

330
331 In conclusion, this genomics analysis suggests that asymptomatic typhoid carriers (or other
332 unknown reservoirs) are likely responsible for continuing sporadic typhoid cases in Chile.
333 The ancient clades present in the country have persisted. We recommend that all sporadic
334 typhoid cases and clusters should be evaluated using epidemiological outbreak investigations
335 and WGS as routine surveillance. This will help ensure the detection of any potential
336 introduction or development of drug-resistant *S. Typhi* and may lead to the identification and
337 treatment of carriers, which would further limit ongoing disease transmission.

338

339 **Acknowledgements**

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

343

344

345 **References**

- 346 1. Marco C, Delgado I, Vargas C, Munoz X, Bhutta ZA, Ferreccio C. Typhoid Fever in
347 Chile 1969-2012: Analysis of an Epidemic and Its Control. *Am J Trop Med Hyg.*
348 2018;99(3_Suppl):26-33. doi:10.4269/ajtmh.18-0125
- 349 2. Levine MM BR, Ferreccio C, Clements ML, Lanata C, Sear S, Morris JG, Cisneros L,
350 Germanier R. Interventions to control endemic typhoid fever: field studies in Santiago, Chile.
351 *Control and Eradication of Infectious Diseases, An International Symposium.* 1985;Series
352 1:37-53.
- 353 3. Black RE, Levine MM, Ferreccio C, Clements ML, Lanata C, Rooney J, et al.
354 Efficacy of one or two doses of Ty21a *Salmonella typhi* vaccine in enteric-coated capsules in
355 a controlled field trial. *Vaccine.* 1990;8(1):81-4. doi:10.1016/0264-410X(90)90183-M
- 356 4. Levine MM, Ferreccio C, Ortiz E, Cryz S. Comparison of enteric-coated capsules and
357 liquid formulation of Ty21a typhoid vaccine in randomised controlled field trial. *The Lancet.*
358 1990;336(8720):891-4. doi:10.1016/0140-6736(90)92266-K
- 359 5. Levine MM, Ferreccio C, Black RE, Germanier R. Large-scale field trial of Ty21a
360 live oral typhoid vaccine in enteric-coated capsule formulation. *Lancet.* 1987;1(8541):1049-
361 52. doi:10.1016/s0140-6736(87)90480-6
- 362 6. Ferreccio C, Levine MM, Rodriguez H, Contreras R. Comparative efficacy of two,
363 three, or four doses of TY21a live oral typhoid vaccine in enteric-coated capsules: a field trial
364 in an endemic area. *J Infect Dis.* 1989;159(4):766-9. doi:10.1093/infdis/159.4.766
- 365 7. Ferreccio C, Morris JG, Jr., Valdivieso C, Prenzel I, Sotomayor V, Drusano GL, et al.
366 Efficacy of ciprofloxacin in the treatment of chronic typhoid carriers. *J Infect Dis.*
367 1988;157(6):1235-9. doi:10.1093/infdis/157.6.123
- 368 8. Gauld JS, Hu H, Klein DJ, Levine MM. Typhoid fever in Santiago, Chile: Insights
369 from a mathematical model utilizing venerable archived data from a successful disease
370 control program. *PLoS Negl Trop Dis.* 2018;12(9):e0006759.
371 doi:10.1371/journal.pntd.0006759
- 372 9. Sears SD, Ferreccio C, Levine MM, Cordano AM, Monreal J, Black RE, et al. The
373 use of Moore swabs for isolation of *Salmonella typhi* from irrigation water in Santiago,
374 Chile. *J Infect Dis.* 1984;149(4):640-2. doi:10.1093/infdis/149.4.640
- 375 10. Levine M. South America: The return of cholera. *The Lancet.* 1991;338(8758):45-6.
376 doi:10.1016/0140-6736(91)90024-J
- 377 11. Medina E. [Epidemic of cholera in Chile; 1991]. *Rev Med Chil.* 1991;119(8):943-56.
- 378 12. Swerdlow DL, Mintz ED, Rodriguez M, Tejada E, Ocampo C, Espejo L, et al.
379 Waterborne transmission of epidemic cholera in Trujillo, Peru: lessons for a continent at risk.
380 *Lancet.* 1992;340(8810):28-33. doi:10.1016/0140-6736(92)92432-f
- 381 13. Tauxe RV, Blake PA. Epidemic cholera in Latin America. *Jama.* 1992;267(10):1388-
382 90.
- 383 14. Wong VK, Baker S, Connor TR, Pickard D, Page AJ, Dave J, et al. An extended
384 genotyping framework for *Salmonella enterica* serovar Typhi, the cause of human typhoid.
385 *Nature communications.* 2016;7(1):1-11. doi: 10.1038/ncomms12827
- 386 15. Inouye M, Dashnow H, Raven LA, Schultz MB, Pope BJ, Tomita T, et al. SRST2:
387 Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med.*
388 2014;6(11):90. doi:10.1186/s13073-014-0090-6
- 389 16. Dyson ZA, Holt KE. Five years of GenoTyphi: updates to the global *Salmonella*
390 Typhi genotyping framework. *The Journal of Infectious Diseases.* 2021.
391 doi:10.1093/infdis/jiab414

- 392 17. Argimon S, Yeats CA, Goater RJ, Abudahab K, Taylor B, Underwood A, et al. A
393 global resource for genomic predictions of antimicrobial resistance and surveillance of
394 Salmonella Typhi at pathogenwatch. *Nat Commun.* 2021;12(1):2879. doi:10.1038/s41467-
395 021-23091-2
- 396 18. Page AJ, De Silva N, Hunt M, Quail MA, Parkhill J, Harris SR, et al. Robust high-
397 throughput prokaryote de novo assembly and improvement pipeline for Illumina data. *Microb*
398 *Genom.* 2016;2(8):e000083. doi:10.1099/mgen.0.000083
- 399 19. Parkhill J, Dougan G, James K, Thomson N, Pickard D, Wain J, et al. Complete
400 genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18.
401 *Nature.* 2001;413(6858):848-52. doi:10.1038/35101607
- 402 20. Kariuki S, Mbae C, Van Puyvelde S, Onsare R, Kawai S, Wairimu C, et al. High
403 relatedness of invasive multi-drug resistant non-typhoidal Salmonella genotypes among
404 patients and asymptomatic carriers in endemic informal settlements in Kenya. *PLoS Negl*
405 *Trop Dis.* 2020;14(8):e0008440. doi:10.1371/journal.pntd.0008440
- 406 21. Britto CD, Dyson ZA, Duchene S, Carter MJ, Gurung M, Kelly DF, et al. Laboratory
407 and molecular surveillance of paediatric typhoidal Salmonella in Nepal: Antimicrobial
408 resistance and implications for vaccine policy. *PLoS neglected tropical diseases.*
409 2018;12(4):e0006408. doi: 10.1371/journal.pntd.0006408
- 410 22. Rahman SIA, Dyson ZA, Klemm EJ, Khanam F, Holt KE, Chowdhury EK, et al.
411 Population structure and antimicrobial resistance patterns of Salmonella Typhi isolates in
412 urban Dhaka, Bangladesh from 2004 to 2016. *PLoS neglected tropical diseases.*
413 2020;14(2):e0008036. doi: 10.1371/journal.pntd.0008036
- 414 23. Wong VK, Holt KE, Okoro C, Baker S, Pickard DJ, Marks F, et al. Molecular
415 Surveillance Identifies Multiple Transmissions of Typhoid in West Africa. *PLoS Negl Trop*
416 *Dis.* 2016;10(9):e0004781. doi:10.1371/journal.pntd.0004781
- 417 24. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid
418 phylogenetic analysis of large samples of recombinant bacterial whole genome sequences
419 using Gubbins. *Nucleic Acids Res.* 2015;43(3):e15. doi:10.1093/nar/gku1196
- 420 25. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses
421 with thousands of taxa and mixed models. *Bioinformatics.* 2006;22(21):2688-90.
422 doi:10.1093/bioinformatics/btl446
- 423 26. Letunic I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic
424 tree display and annotation. *Bioinformatics.* 2007;23(1):127-8.
425 doi:10.1093/bioinformatics/btl529
- 426 27. Argimón S, Abudahab K, Goater RJE, Fedosejev A, Bhai J, Glasner C, et al.
427 Microreact: visualizing and sharing data for genomic epidemiology and phylogeography.
428 *Microb Genom.* 2016;2(11):e000093. doi:10.1099/mgen.0.000093
- 429 28. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, et
430 al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in
431 bacterial genomes. *Antimicrob Agents Chemother.* 2014;58(1):212-20.
432 doi:10.1128/aac.01310-13
- 433 29. Carattoli A, Zankari E, Garcia-Fernandez A, Voldby Larsen M, Lund O, Villa L, et al.
434 In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus
435 sequence typing. *Antimicrob Agents Chemother.* 2014;58(7):3895-903.
436 doi:10.1128/AAC.02412-14
- 437 30. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, et al. VFDB: a reference database for
438 bacterial virulence factors. *Nucleic Acids Res.* 2005;33(Database issue):D325-8.
439 doi:10.1093/nar/gki008

- 440 31. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome
441 assemblies from short and long sequencing reads. *PLoS Comput Biol*. 2017;13(6):e1005595.
442 doi:10.1371/journal.pcbi.100559
- 443 32. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de
444 novo genome assemblies. *Bioinformatics*. 2015;31(20):3350-2.
445 doi:10.1093/bioinformatics/btv383
- 446 33. Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J.
447 ACT: the Artemis Comparison Tool. *Bioinformatics*. 2005;21(16):3422-3.
448 doi:10.1093/bioinformatics/bti553
- 449 34. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer.
450 *Bioinformatics*. 2011;27(7):1009-10. doi:10.1093/bioinformatics/btr0
- 451 35. Maes M, Dyson ZA, Higginson EE, Fernandez A, Araya P, Tennant SM, et al.
452 Multiple Introductions of *Salmonella enterica* Serovar Typhi H58 with Reduced
453 Fluoroquinolone Susceptibility into Chile. *Emerg Infect Dis*. 2020;26(11):2736-40.
454 doi:10.3201/eid2611.201676
- 455 36. Dyson ZA, Klemm EJ, Palmer S, Dougan G. Antibiotic resistance and typhoid.
456 *Clinical Infectious Diseases*. 2019;68(Supplement 2):S165-S70. doi: 10.1093/cid/ciy111
- 457 37. Dyson ZA, Thanh DP, Bodhidatta L, Mason CJ, Srijan A, Rabaa MA, et al. Whole
458 genome sequence analysis of *Salmonella* Typhi isolated in Thailand before and after the
459 introduction of a national immunization program. *PLoS neglected tropical diseases*.
460 2017;11(1):e0005274. doi: 10.1371/journal.pntd.0005274
- 461 38. Fica AE, Prat-Miranda S, Fernandez-Ricci A, D'Ottone K, Cabello FC. Epidemic
462 typhoid in Chile: analysis by molecular and conventional methods of *Salmonella* typhi strain
463 diversity in epidemic (1977 and 1981) and nonepidemic (1990) years. *J Clin Microbiol*.
464 1996;34(7):1701-7. doi:10.1128/jcm.34.7.1701-1707.1996
- 465 39. Ames WR, Robins M. Age and Sex as Factors in the Development of the Typhoid
466 Carrier State, and a Method for Estimating Carrier Prevalence. *Am J Public Health Nations*
467 *Health*. 1943;33(3):221-30. doi:10.2105/ajph.33.3.221
- 468 40. Chattaway MA, Gentle A, Nair S, Tingley L, Day M, Mohamed I, et al.
469 Phylogenomics and antimicrobial resistance of *Salmonella* Typhi and Paratyphi A, B and C
470 in England, 2016-2019. *Microb Genom*. 2021;7(8). doi:10.1099/mgen.0.000633
- 471 41. Ingle DJ, Nair S, Hartman H, Ashton PM, Dyson ZA, Day M, et al. Informal genomic
472 surveillance of regional distribution of *Salmonella* Typhi genotypes and antimicrobial
473 resistance via returning travellers. *PLoS Negl Trop Dis*. 2019;13(9):e0007620.
474 doi:10.1371/journal.pntd.0007620
- 475 42. Guevara PD, Maes M, Thanh DP, Duarte C, Rodriguez EC, Montañó LA, et al. A
476 genomic snapshot of *Salmonella enterica* serovar Typhi in Colombia. *PLoS Negl Trop Dis*.
477 2021;15(9):e0009755. doi:10.1371/journal.pntd.0009755
- 478 43. Wong VK, Baker S, Pickard DJ, Parkhill J, Page AJ, Feasey NA, et al.
479 Phylogeographical analysis of the dominant multidrug-resistant H58 clade of *Salmonella*
480 Typhi identifies inter-and intracontinental transmission events. *Nature genetics*.
481 2015;47(6):632-9. doi: 10.1038/ng.3281
- 482 44. Dyson ZA, Malau E, Horwood PF, Ford R, Siba V, Yoannes M, et al. Whole genome
483 sequence analysis of *Salmonella* Typhi in Papua New Guinea reveals an
484 established population of genotype 2.1.7 sensitive to antimicrobials. medRxiv.
485 2021:2021.09.27.21264209.
- 486 45. Kariuki S, Dyson ZA, Mbae C, Ngetich R, Kawai SM, Wairimu C, et al. Multiple
487 introductions of multidrug-resistant typhoid associated with acute infection and
488 asymptomatic carriage, Kenya. *Elife*. 2021;10. doi:10.7554/eLife.67852

- 489 46. Thanh Duy P, Thieu NTV, Nguyen Thi Nguyen T, Ngoc Dan Thanh H, Dongol S,
490 Karkey A, et al. Gallbladder carriage generates genetic variation and genome degradation in
491 Salmonella Typhi. *PLoS Pathog.* 2020;16(10):e1008998. doi:10.1371/journal.ppat.1008998
492 47. Johnson R, Mylona E, Frankel G. Typhoidal Salmonella: Distinctive virulence factors
493 and pathogenesis. *Cell Microbiol.* 2018;20(9):e12939. doi:10.1111/cmi.12939
494 48. Murret-Labarthe C, Kerhoas M, Dufresne K, Daigle F. New Roles for Two-
495 Component System Response Regulators of Salmonella enterica Serovar Typhi during Host
496 Cell Interactions. *Microorganisms.* 2020;8(5). doi:10.3390/microorganisms8050722
497 49. Forest CG, Ferraro E, Sabbagh SC, Daigle F. Intracellular survival of Salmonella
498 enterica serovar Typhi in human macrophages is independent of Salmonella pathogenicity
499 island (SPI)-2. *Microbiology (Reading).* 2010;156(Pt 12):3689-98.
500 doi:10.1099/mic.0.041624-0
501 50. Karlinsey JE, Stepien TA, Mayho M, Singletary LA, Bingham-Ramos LK, Brehm
502 MA, et al. Genome-wide Analysis of Salmonella enterica serovar Typhi in Humanized Mice
503 Reveals Key Virulence Features. *Cell Host Microbe.* 2019;26(3):426-34.e6.
504 doi:10.1016/j.chom.2019.08.001
505 51. Desai SK, Winardhi RS, Periasamy S, Dykas MM, Jie Y, Kenney LJ. The
506 horizontally-acquired response regulator SsrB drives a Salmonella lifestyle switch by
507 relieving biofilm silencing. *Elife.* 2016;5. doi:10.7554/eLife.10747
508 52. Marshall JM, Flechtner AD, La Perle KM, Gunn JS. Visualization of extracellular
509 matrix components within sectioned Salmonella biofilms on the surface of human gallstones.
510 *PLoS One.* 2014;9(2):e89243. doi:10.1371/journal.pone.0089
511 53. Parkhill J, Dougan G, James KD, Thomson NR, Pickard D, Wain J, et al. Complete
512 genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18.
513 *Nature.* 2001;413(6858):848-52. doi:10.1038/35101607
514 54. Kidgell C, Pickard D, Wain J, James K, Diem Nga LT, Diep TS, et al.
515 Characterisation and distribution of a cryptic Salmonella typhi plasmid pHCM2. *Plasmid.*
516 2002;47(3):159-71. doi:10.1016/s0147-619x(02)00013-
517 55. Dyson ZA, Thanh DP, Bodhidatta L, Mason CJ, Srijan A, Rabaa MA, et al. Whole
518 Genome Sequence Analysis of Salmonella Typhi Isolated in Thailand before and after the
519 Introduction of a National Immunization Program. *PLoS Negl Trop Dis.*
520 2017;11(1):e0005274. doi:10.1371/journal.pntd.0005
521

522

523 **Authorship declaration**

524 All the Co-authors contributed to manuscript editing

525 **Authors contributions**

526 **Conception and design of experiments**

527 Myron M. Levine, Rosanna Lagos,

528 **Sample collection and DNA extractions**

529 Alda Fernandez, Pamela Araya, Rosanna Lagos, Juan Carlos Hormazabal

530 **Analyzed the data**

531 Mailis Maes, Megan E. Carey, Ellen E. Higginson,

532 **Contributed reagents/materials/analysis tools**

533 Zoe A. Dyson, Stephen Baker, Michael J. Sikorski, Sharon M. Tennant,

534 **Funding acquisition,**

535 Myron M. Levine, Gordon Dougan

536 **Writing original draft**

537 Mailis Maes, Michael J. Sikorski, Megan E. Carey, , Ellen E. Higginson

538 **Writing review and editing**

539 Mailis Maes, Michael J. Sikorski, Megan E. Carey, Ellen E. Higginson, , Myron M. Levine,

540 Zoe A. Dyson, Sharon M. Tennant

541

542 **Funding**

543 This work was supported by a grant [OPP1161058] from the Bill & Melinda Gates

544 Foundation, this work was also supported by the Gates Foundation (TyVAC) [OPP1151153]

545 and the Wellcome Trust (STRATAA) [106158 and 098051].

546 MM was funded by the NIHR Cambridge Biomedical Research Centre and NIHR AMR

547 Research Capital Funding Scheme [NIHR200640]. ZAD was supported by a grant funded by

548 the Wellcome Trust (STRATAA; 106158/Z/14/Z), and received funding from the European
549 Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-
550 Curie grant agreement TyphiNET (#845681). MJS received research support from the NIH
551 NIAID grants F30AI156973 and U19AI110820 as well as NIDDK T32DK067872. The
552 views expressed are those of the authors and not necessarily those of the NIHR or the
553 Department of Health and Social Care. The funders had no role in the design and conduct of
554 the study; collection, management, analysis, and interpretation of the data; preparation,
555 review, or approval of the manuscript; and decision to submit the manuscript for publication.

556

557 **Data availability**

558 Raw sequence data have been submitted to the European Nucleotide Archive (ENA) and
559 accession numbers of individual isolates are listed in the appendix.

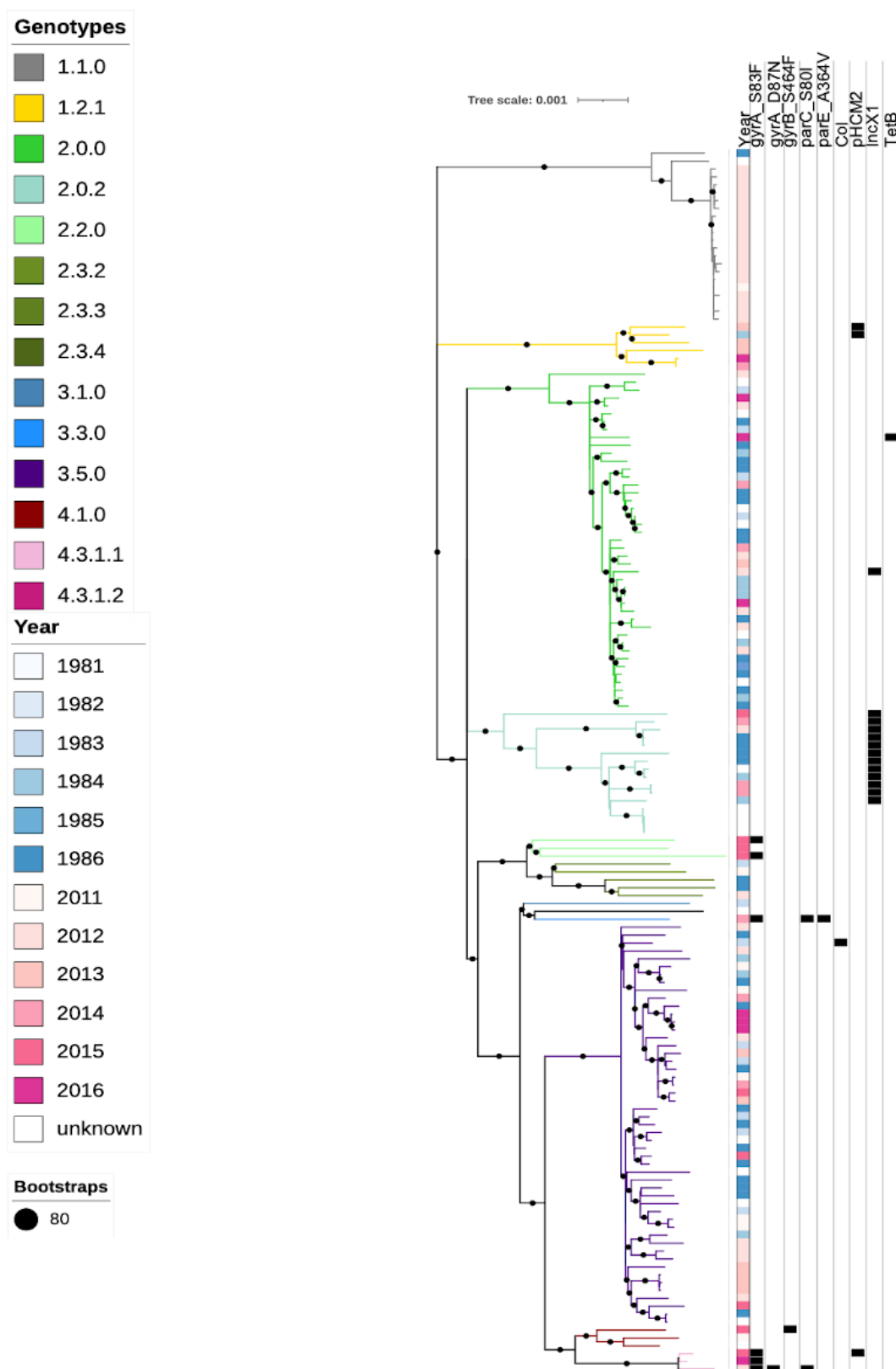
560

561 **Competing interest**

562 The authors have declared that no competing interests exist

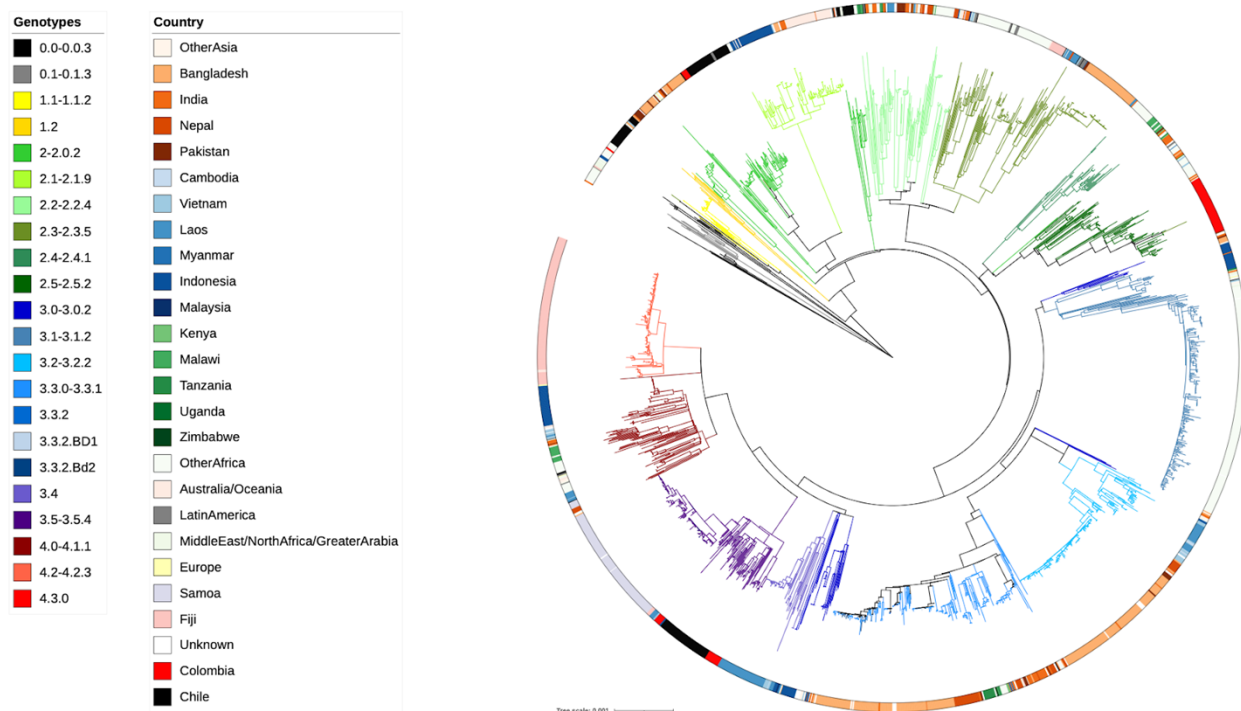
563

564 Figures



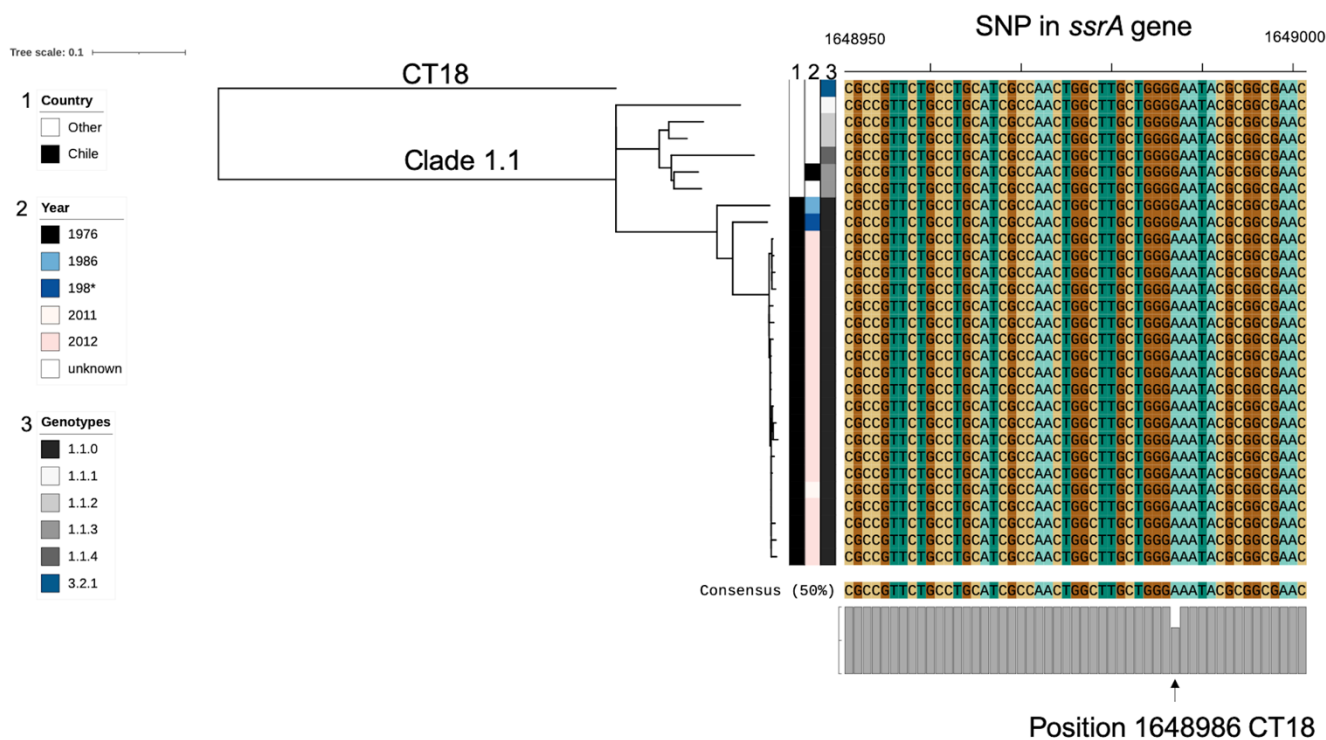
565

566 **Fig. 1. Genotypes and molecular determinates of AMR and plasmids in the Santiago**
567 **collection.** SNP based maximum likelihood core genome phylogeny indicating the genotypes
568 of all the Santiago *S. Typhi* isolates is shown by the branch colour as per the inset legend. A
569 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,
570 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
571 3.5. Year of isolation, SNPs in the QRDR region associated with reduced susceptibility to
572 fluoroquinolones as well as plasmids and detected resistance genes are shown in the columns
573 adjacent to the tree. Bootstrap confidence of 80 or above is indicated with a black circle on
574 the branch.



575

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



581

582

583 **Fig. 3. *S. Typhi* clade 1.1 phylogeny showing sequences carrying *ssrA* mutation (*ssrA*-**

584 **G639E).** Phylogenetic core genome analysis of all *S. Typhi* isolates of genotype 1.1

585 identified from the global collection. The majority of these isolates were collected in

586 Santiago (black bar), two isolates belonged to the 1980s collection appearing ancestral to the

587 2010s collection. The twenty isolates from 2011-2012 showed an interesting cluster with a

588 median SNP distance of 8 SNPs which is unlikely to have evolved among acute cases, where

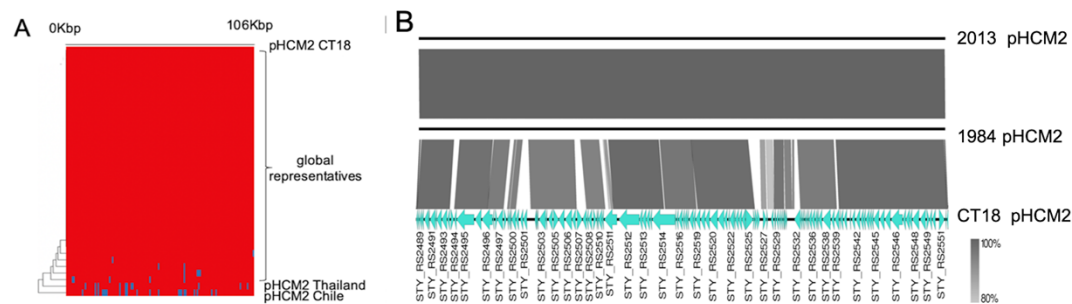
589 the mutation rate is calculated to be < 1 SNP/genome/year, in one year. Also shown is an

590 alignment of a section of the *ssrA* gene, showing the conserved SNP at position 1648986 of

591 the *S. Typhi* CT18 reference genome. * the precise isolation year of one of the 1980s isolates

592 was unclear.

593



594

595

596 **Fig. 4. Genetic comparison of cryptic plasmid pHCM2-like sequences.** Global

597 representative collection of pHCM2 aligned to the reference pHCM2 from CT18. (A)

598 Presence and absence of genes located on pHCM2. Red indicates presence of the genes

599 located on the reference pHCM2 sequence from CT18, while blue indicates absence of the

600 genes compared to the reference. Nucleotide BLAST comparison of the 2 Chilean assembled

601 pHCM2-like plasmids (ERR271054, ERR4289185) against the reference pHCM2 plasmid

602 from *S. Typhi* CT18 (accession no:AL513384) (B). Gene annotations are shown for the

603 reference pHCM2 from CT18, with the shading intensity indicating percent nucleotide

604 homology between the plasmid sequences.

605

606

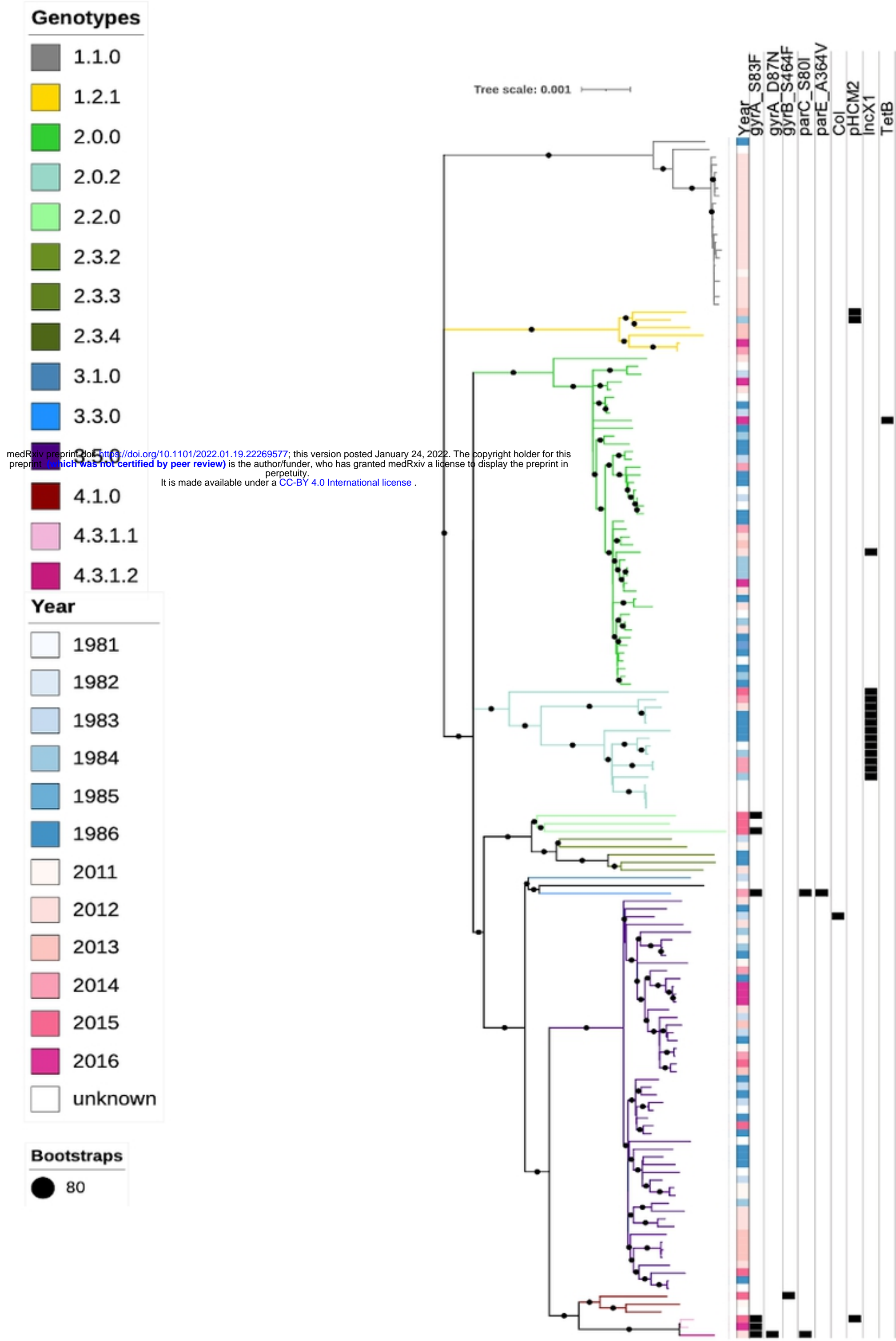


Figure 1

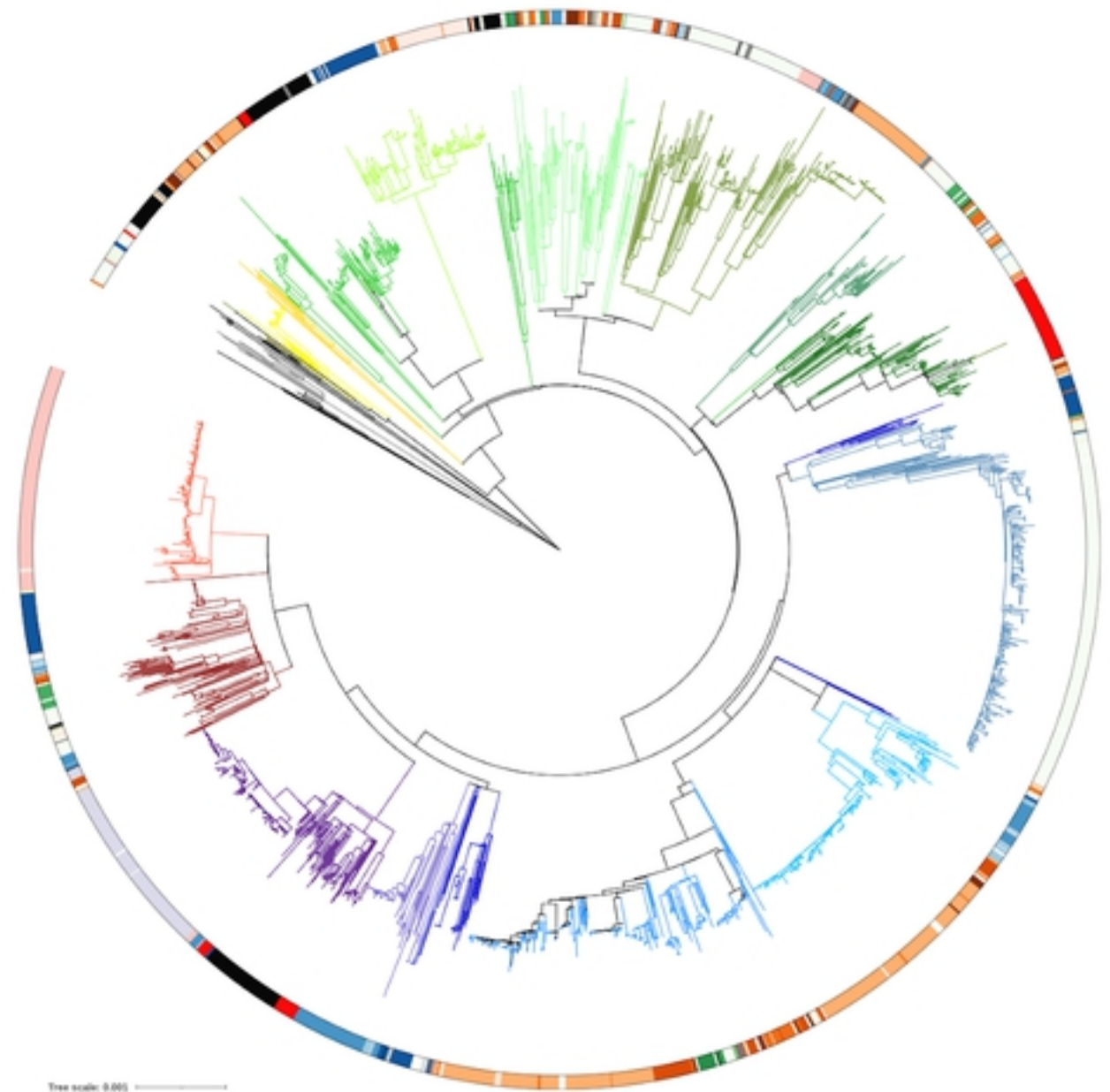
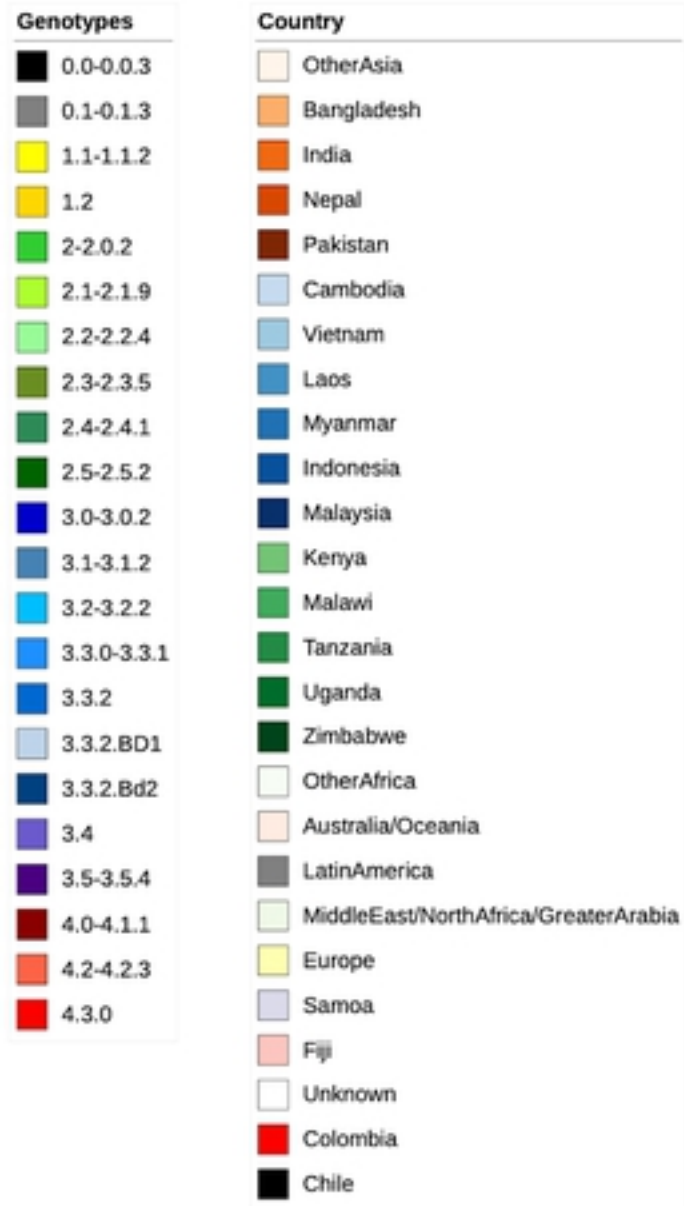


Figure 2

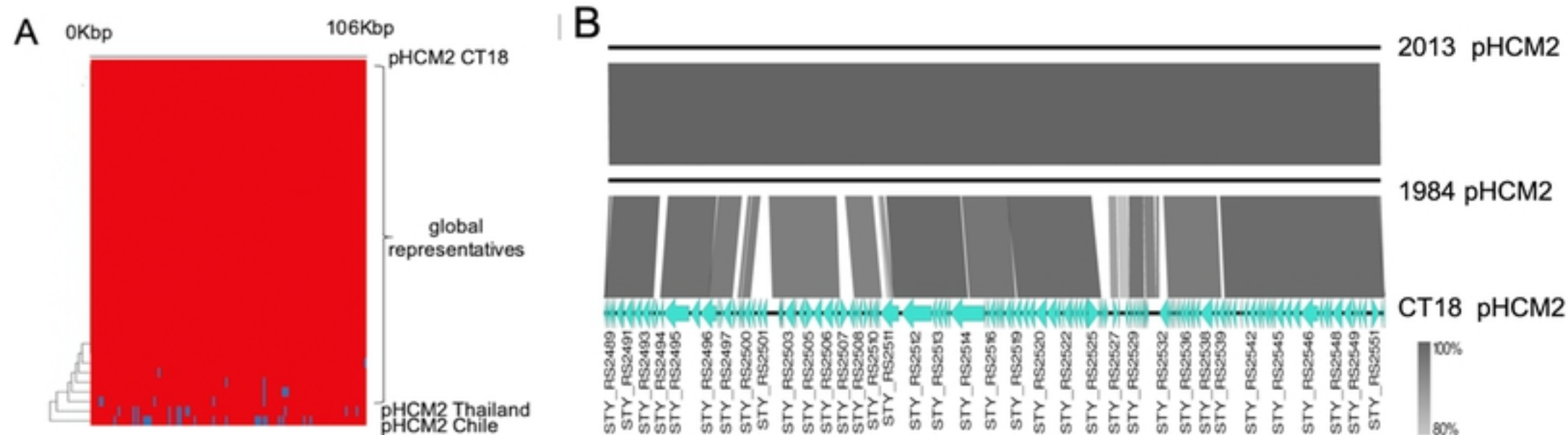


Figure 4