Global diversity and evolution of *Salmonella* Panama, an understudied serovar causing gastrointestinal and invasive disease worldwide: a genomic epidemiology study

Authors

Caisey V. Pulford[1*], Blanca M. Perez-Sepulveda[1*], Danielle J. Ingle[2], Rebecca J. Bengtsson[1], Rebecca J. Bennett[1], Ella V. Rodwell[3], Maria Pardos de la Gandara[4], Charlotte Chong[5], P. Malaka De Silva[5], Magali Ravel[4], Véronique Guibert[4], Elisabeth Njamkepo[4], Neil Hall[6], Marie A. Chattaway[3], Benjamin P. Howden[7,8], Deborah A Williamson[8], Jay C. D. Hinton[1#], François-Xavier Weill[4#§], Kate S. Baker[1,5#§]

[1]Clinical Infection, Microbiology, and Immunology, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, United Kingdom.

[2]Department of Microbiology and Immunology, The University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Melbourne, Australia.

[3]Gastrointestinal Bacteria Reference Unit, UK Health Security Agency, London, United Kingdom.

[4]Institut Pasteur, Université Paris Cité, Unité des Bactéries pathogènes entériques, Paris, France.

[5]Department of Genetics, University of Cambridge, Cambridge, United Kingdom

[6]Earlham Institute, Norwich Research Park, Norwich, UK

[7]Microbiological Diagnostic Unit Public Health Laboratory, Department of Microbiology and Immunology University of Melbourne at The Peter Doherty Institute for Infection and Immunity, Melbourne, VIC, Australia.

[8]Centre for Pathogen Genomics, The University of Melbourne, Melbourne, Australia.

[*]These authors contributed equally

[#]These authors contributed equally

[§]Corresponding authors:

Email: kb827@cam.ac.uk

Email: francois-xavier.weill@pasteur.fr Tel : +33 145688345, +33 643247997

Abstract

Background

Nontyphoidal *Salmonella* (NTS) is a globally important bacterial pathogen, typically associated with foodborne gastrointestinal infection. Some NTS serovars can also colonise normally sterile sites in humans to cause invasive NTS (iNTS) disease. One understudied *Salmonella enterica* serovar which is responsible for a significant number of cases of iNTS disease is Panama. Despite global dissemination, numerous outbreaks, and a reported association with iNTS disease, *S. enterica* serovar Panama (*S.* Panama) has not been investigated in detail.

Methods

Using combined epidemiological and whole genome sequencing data we analysed 836 *S.* Panama isolates from all inhabited continents collected between 1931 and 2019. A combination of phylodynamic approaches were used to determine population structure & evolutionary history, and to infer geo-temporal dissemination. Subsequently, we characterised geographic and clade-specific trends in antimicrobial resistance (AMR), and genetic markers for invasiveness using epidemiological and bioinformatic approaches.

Findings

We identified the presence of multiple geographically linked *S.* Panama clades, and regional trends in antimicrobial resistance profiles. Most isolates were pan-susceptible to antibiotics and belonged to clades circulating in the United States of America, Latin America, and the Caribbean. Multidrug resistant (MDR) isolates belonged to two phylogenetic clades circulating in Europe and Asia/Oceania, which exhibited the highest invasiveness indices based on the conservation of 196 extra-intestinal predictor genes.

Interpretation

55 This first large-scale phylogenetic analysis of *S.* Panama revealed important information about
56 population structure, AMR, global ecology, and genetic markers of invasiveness of the identified
57 genomic subtypes. Our findings provide an important baseline for understanding *S.* Panama
58 infection in the future. The presence of MDR clades with an elevated invasiveness index should be
59 monitored by ongoing surveillance as such clades may pose an increased public health risk.
60

61 **Funding**

62 The project was supported by both a Global Challenges Research Fund (GCRF) data & resources
63 grant BBS/OS/GC/000009D and the BBSRC Core Capability Grant to the Earlham Institute
64 BB/CCG1720/1. JCDH received a Wellcome Trust Investigator award (grant number 222528/Z/21/Z).
65 CVP was supported by the John Lennon Memorial Scholarship from the University of Liverpool and a
66 Fee Bursary Award from the Institute of Integrative Biology at the University of Liverpool. FXW was
67 awarded the *Institut Pasteur*; *Santé publique France*; the *Fondation Le Roch-Les Mousquetaires*; and
68 the French government's *Investissement d'Avenir* programme*, Laboratoire d'Excellence* 'Integrative
69 Biology of Emerging Infectious Diseases' (grant number ANR-10- LABX-62-IBEID). MDS was
70 supported by a BBSRC grant (KSB: BB/V009184/1). RJBengtsson was supported by an MRC grant
71 (KSB: MR/R020787/1). RJBennett was supported by a BBSRC DTP (BB/M011186/1). DJI was
72 supported by a National Health and Medical Research Council (NHMRC) Emerging Leadership
73 Fellowship (GNT1195210). BPH was supported by NHMRC Leadership Fellowship (GNT1196103). KSB
74 and EVR are affiliated to the National Institute for Health Research Health Protection Research Unit
75 (NIHR HPRU) in Gastrointestinal Infections at University of Liverpool in partnership with the United
76 Kingdom Health Security Agency, in collaboration with University of Warwick. The views expressed
77 are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health
78 and Social Care or UKHSA. MAC is affiliated to the National Institute for Health Research Health
79 Protection Research Unit (NIHR HPRU) (NIHR200892) in Genomics and Enabling Data at University of
80 Warwick in partnership with the UK Health Security Agency (UKHSA), in collaboration with University

of Cambridge and Oxford. MAC is based at UKHSA. The views expressed are those of the author(s) and not necessarily those of the NIHR, the Department of Health and Social Care or the UK Health Security Agency.

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

No payment was received by any pharmaceutical company or other agency to write this article.

The authors were not precluded from accessing data in the study, and accept responsibility to submit for publication.

Research in context

Evidence before this study

Salmonella Panama has consistently been reported as a frequently isolated *Salmonella* serovar in national surveillance datasets, causing both sporadic cases and larger outbreaks. However, this picture has been masked due to most of the focus being placed on the top two serovars associated with nontyphoidal *Salmonella* (NTS) and invasive NTS disease: Typhimurium and Enteritidis. Previous works on *S.* Panama have determined transmission to include human faeces and breast milk, as well as non-human sources such as environmental reservoirs and animals (reptiles and pigs used for food). In contrast to most of the *Salmonella* serovars causing gastroenteritis, *S.* Panama has also been associated with a range of systemic infections including septicaemia, meningitis, and osteomyelitis, mostly affecting infants. The epidemiological picture of this pathogen was not well known with mixed antimicrobial resistance profiles depending on the studies.

To date, 228 studies on *S.* Panama are available on NCBI's PubMed (searched on 11th October 2023: '((Salmonella) AND (Panama[Title/Abstract])) AND (Panama[Text Word])').

Added value of this study

Here we conducted a large-scale study of 836 globally relevant *S.* Panama isolates to understand global population structure and disease ecology. The *S.* Panama genomes used in this study were sourced from combination of historical collections (including the first ever isolated strain of this serovar dating from 1931), national surveillance datasets, and publicly available data. Some of these isolates had been linked to travel, allowing an inferred location for understanding population structure. Overall, this assembled *S.* Panama collection spanned a range of 88 years and covered 45 countries and regions of all inhabited continents. We applied bacterial phylodynamic approaches to determine population structure, evolutionary history, and clade-specific trends of invasiveness and antimicrobial resistance correlated with geo-temporal parameters.

Implications of all the available evidence

This study revealed the population structure and global ecology of *S.* Panama, indicating that multidrug resistant *S.* Panama circulates in European and Asian regions, and presents an increased number of genetic markers for extra-intestinal invasiveness. The transmission and expansion of antimicrobial-resistant *S.* Panama strains presented in this study highlight the significance of viewing control and monitoring efforts from an international perspective.

126 ## Introduction

127 Nontyphoidal *Salmonella* (NTS) disease poses a significant burden to public health globally, causing
128 approximately 153 million cases and 57,000 deaths per annum. As well as being responsible for
129 gastroenteritis in humans, NTS can also invade normally sterile sites resulting in bacteraemia and
130 meningitis. There were estimated to be 535,000 cases of invasive NTS (iNTS) disease in 2017, causing
131 77,500 deaths worldwide[1]. In humans, the clinical presentation depends upon a combination of host
132 immune factors and bacterial features that are specific to individual *Salmonella* pathovariants[2]. The
133 high levels of iNTS disease caused by *Salmonella enterica* serovars Typhimurium and Enteritidis in
134 sub-Saharan Africa have generated a significant research focus[1]. However, little is known about
135 other serovars, such as *S. enterica* serovar Panama (hereafter referred to as *S.* Panama), which are
136 also associated with iNTS disease and responsible for numerous outbreaks, as reviewed previously[3].
137
138 Whilst the majority of cases of iNTS have been identified in sub-Saharan Africa[1,4], *S.* Panama has a
139 very different geographical distribution. *S.* Panama has consistently been reported as a leading cause
140 of NTS disease in Martinique, Guadeloupe, and French Guiana and is significantly associated with
141 causing invasive infections in children[3,5,6]. Although the proportion of salmonellosis cases caused by
142 *S.* Panama is high in these locations, the prevalence of antimicrobial resistance (AMR) is reportedly
143 low[3,5]. In contrast, other regions have seen higher levels of AMR in *S.* Panama[7]. For example, in the
144 1970s and 1980s *S.* Panama caused public health concern in Europe after spreading through the pork
145 industry and causing multiple hospital outbreaks[3]. *S.* Panama maintained its ranking as one of the
146 top 20 most frequently isolated serovars in the European Union until 2017[8] and is a frequent cause
147 of invasive disease, with 7% of cases historically presenting as extraintestinal infection in England.
148 Similarly, *S.* Panama has been an important cause of iNTS in Asia, linked with high levels of AMR. Up
149 to 83% of domestic and imported *S.* Panama isolates in Tokyo were reported to be resistant to three
150 or more antimicrobial classes (multidrug resistant (MDR)) by the early 2000s[3]. In Oceania, *S.* Panama
151 is one of the top three serovars isolated between 2007 and 2016 in Queensland, Australia[9].

152

153 In the current era of genomic pathogen surveillance, characterising the molecular epidemiology and
154 evolutionary history of *S.* Panama will provide a vital baseline of understanding for this globally
155 important serovar. Here, we used whole genome-based approaches on historical and contemporary
156 global isolates to determine the genomic epidemiology of *S.* Panama using geographically and
157 temporally diverse datasets. We characterised four geographically-associated *S.* Panama clades and
158 identified regional trends in AMR. Our findings prompted a phylodynamic investigation which
159 revealed the evolution of the serovar. Finally, we determined genetic markers of invasiveness among
160 *S.* Panama clades.

161

162 Results and Discussion

163 Establishing an informative collection of *S.* Panama isolates

164 To create a relevant dataset, we sequenced the genomes of 552 *S.* Panama isolates, and included an
165 additional 284 *S.* Panama genomes from a combination of historical collections, national surveillance
166 datasets, and publicly available data, including the first isolate obtained in Panama in 1931 during a
167 food poisoning outbreak in US soldiers stationed in the Canal zone (Figure 1, Table S1).

168

169 For public health datasets, an isolate was deemed to be 'travel associated' if a case had reported
170 travel up to 28 days (UKHSA) or seven days (Institut Pasteur and MDU PHL) prior to symptom onset.
171 However, information on travel was not available for all cases and is likely under-reported. It is also
172 possible that those cases who reported travel, did not necessarily acquire their infection whilst
173 travelling. The 'inferred region' of each isolate was set as the travel destination where available (n =
174 78) or kept as the isolation region (n = 757). All the travel-associated isolates came from the Institut
175 Pasteur's national surveillance dataset (n = 43) and UKHSA (n = 42), leading to the inclusion of a
176 further 12 countries not previously represented by isolates.

177

178 Overall, the collection of 836 *S.* Panama included 771 human isolates of which 403 were isolated
179 from stool, 290 were derived from extra-intestinal sites (blood and cerebrospinal fluid), and 78 were
180 taken from other body sites (Figure 1, Table S1). The majority (67% *n* = 559) of the isolates were
181 sourced from the Unité des Bactéries pathogènes entériques, Institut Pasteur, Paris, France.

182

183 **Global population structure of *S.* Panama**

184 To define population structure, we explored the phylogenetic relationship among the 836 *S.* Panama
185 genomes and defined population groupings using rHeir Bayesian Analysis of Population Structure
186 (BAPS)[10] (methods, Figure 2). The BAPS groupings reflected four monophyletic clades with greater
187 than 95% bootstrap support containing variable numbers of isolates. We defined the four clades as
188 C1 (*n* = 338), C2 (*n* = 124), C3 (*n* = 131), and C4 (*n* = 104). The remaining 139 genomes in this study
189 were polyphyletic, representing further diversity of the *S.* Panama serovar.

190

191 To investigate the relationship between clades and geographic region, location data was correlated
192 with phylogenetic position (Figure 2). Where information on travel was available, we used the
193 destination of travel as an inferred location. Taken together, our findings revealed evidence of
194 regional clustering throughout the phylogenetic tree with relevance to the genome-based
195 surveillance of future outbreaks. The majority (*n* = 295/338) of genomes clustering in clade C1 were
196 from Latin America & the Caribbean, including the first *S.* Panama isolate from 1931. Within Clade
197 C1, there were five geographically-linked monophyletic clusters: three were dominated by isolates
198 from Guadeloupe, Martinique, and French Guiana respectively, and two contained isolates from
199 multiple geographical locations in Latin America & the Caribbean (Figure 2). In three of these
200 clusters, we identified isolates from both Guadeloupe and Martinique, suggesting that inter-island
201 transmission of C1 is likely occurring. Clade C1 was named 'Latin America & the Caribbean' reflecting

202 its geographical composition. A further clade, Clade C3 consisted of isolates from Martinique (*n* =
203 125/131), and was named the 'Martinique clade'.
204
205 Clade C2 was mainly composed of isolates obtained from Europe, and included distinct subclades in
206 which isolates from metropolitan France and the UK were overrepresented. Specifically, while the
207 majority (84% *n* = 107/124) of isolates in each of these subclades came from the UK (*n* = 60/124) and
208 metropolitan France (*n* = 47/124), C2 also contained isolates from less well sampled European
209 countries including Germany, Malta, Ireland, and the Netherlands. These findings are consistent with
210 European-wide circulation, and so C2 was designated as the 'European clade'.
211
212 Clade C4 included isolates mainly from Asia and Oceania (*n* = 63/104). We identified three C4
213 subclades, of which two were dominated by isolates obtained from Asia and Oceania and a third
214 comprised of isolates obtained in the USA. Accordingly, C4 was termed the 'Asia/Oceania clade'.
215
216 **Antimicrobial Resistance Trends**
217 To determine AMR trends in *S.* Panama, the draft genomes of all 836 isolates were examined for
218 genes and mutations that confer reduced susceptibility to antimicrobials. Previous studies have
219 demonstrated that genome-based AMR analysis accurately predicted phenotype for 97·8% of
220 *Salmonella* isolates[11]. In total, 14·5% (*n* = 121/836) of isolates were predicted to be resistant to at
221 least one antimicrobial class, with the remaining 85·5% (*n* = 715/836) predicted to be pan-
222 susceptible, lacking any known antibiotic resistance genes or mutations. Most resistant isolates (*n* =
223 113/121) fell within C2 (European clade) and C4 (Asia/Oceania clade) (Figure 3). These findings are
224 supported by reports of an increase in *S.* Panama isolates from Asia that were resistant to multiple
225 antibiotics, including cotrimoxazole, ampicillin, streptomycin, kanamycin, and gentamicin[12].
226

227 The most common resistance profile, occurring in 35·5% (*n* = 43/121) of resistant isolates, was
228 against streptomycin (*aadA1, aadA2*), ampicillin (bla_{TEM-1B}), chloramphenicol (*cmlA1*), trimethoprim
229 (*dfrA12*), sulfisoxazole (*sul3*), and tetracycline (*tetA*). Detailed analysis of the draft genomes showed
230 that most of the resistance genes (*aadA1*, *aadA2*, *cmlA1*, *dfrA12* and *tetA*) were encoded by an MDR
231 cassette located on a single contiguous sequence (contig; Figure 3B). A BLAST search of the contig
232 revealed 99·8% sequence identity to a 21·2 kb region of an 83·2kb plasmid (GenBank accession
233 number CP044301) which had previously been identified in an *E. coli* isolate obtained from pork in
234 Cambodia[13].
235
236 The second most common AMR profile, found in 12·4% (15/121) of resistant isolates was against
237 streptomycin (*aph(3'')-Ib* and *aph(6)-Id*), ampicillin (bla_{TEM-1B}), trimethoprim (*dfrA14*), sulfisoxazole
238 (*sul2*), and tetracycline (*tetA*). All genes were carried by an MDR cassette located on a single contig
239 (Figure 3B). The contig shared 99·9% sequence identity with a 28·3 kb region of a 50·9 kb *S. enterica*
240 plasmid (GenBank accession number CP028173), which has been identified in multiple strains
241 including an *S.* Enteritidis isolate from Ghana[14].
242
243 The remaining resistant *S.* Panama isolates had a variety of susceptibility profiles, including one
244 extensively drug resistant (XDR) isolate 201209115 (GenBank accession number SAMEA6142076),
245 which was phylogenetically situated within C4 (Asia/Oceania clade) and was isolated in 2012 from
246 the bloodstream of a case who had reported recent travel to Thailand. The nomenclature used to
247 describe the XDR isolate is in line with previous definitions of resistance to three first-line drugs
248 (chloramphenicol, ampicillin, and trimethoprim-sulfamethoxazole) as well as fluoroquinolones and
249 third-generation cephalosporins[15]. The isolate carried AMR determinants against eight different
250 antimicrobial classes, including: fluoroquinolones, polymyxins, and extended spectrum beta-
251 lactamase (ESBL) mediated resistance to third generation cephalosporins ($bla_{CTX-M-55}$). The resistance
252 genes in the XDR isolate were distributed across five contigs (Figure S1), with the kanamycin

253 (*aph(3')-Ia*) and gentamicin (*aac(3)-IId*) genes being carried on separate contigs. Genes encoding
254 resistance to streptomycin (*aadA1, aadA2*), chloramphenicol (*cmlA1*), trimethoprim (*dfrA12*), and
255 sulfisoxazole (*sul3*) were co-located on the same contig, and the resistance genes against
256 ciprofloxacin (*qnrS1*), ampicillin, and ceftriaxone ($bla_{CTX\text{-}M\text{-}55}$) were also co-located on a separate
257 contig.
258
259 These findings of clustered AMR elements on the genome are consistent with plasmid-mediated
260 resistance in *S.* Panama, as has been seen in previous studies[12]. The plasmid complement of the *S.*
261 Panama serovar has rarely been studied in the past, but it has been reported that *S.* Panama does
262 not commonly carry the large virulence plasmids that have previously been characterised in other
263 *Salmonella* serovars[16]. Rather, *S.* Panama have previously been found to carry a heterogeneous
264 population of plasmids (Table S1)[17]. Our short read-based plasmid prediction analysis identified that
265 only 15·7% (*n* = 131/836) of *S.* Panama isolates in this study carried plasmids, including 64·9% (*n* =
266 85/131) of resistant isolates and 35·1% (*n* = 46/131) susceptible isolates. The main plasmid
267 incompatibility groups identified were IncFIA(HI1) (*n* = 40/131) and IncFIB(K) (*n* = 35/131) restricted
268 to C4 (Asia/Oceania clade), with the main AMR profile streptomycin, ampicillin, chloramphenicol,
269 trimethoprim, sulfisoxazole, and tetracycline (*n* = 27/35). IncN plasmids were identified mostly in C2
270 (European clade) (*n* = 28/131) that shared the same AMR profile as IncFIA(HI1) and IncFIB(K)
271 plasmids, but includes kanamycin resistance instead of chloramphenicol resistance.
272
273 To explore AMR context and evolution within C2 (which shows high levels of resistance), plasmid
274 DNA of three MDR isolates (isolates 61-66, 86-66, and 376-66; GenBank accession numbers
275 ERR12253897, ERR12253905, and ERR12253883 respectively) collected during hospital outbreaks
276 caused by *S.* Panama in France during the 1960s and 1970s were extracted and sequenced using
277 long-read sequencing technology (Oxford Nanopore). These isolates were obtained in France in 1966
278 and clustered within C2 (European clade). Sequence analysis showed a 45 kb IncN plasmid (pPan376-

279 IncN; GenBank accession number OR797034) encoding resistance to tetracycline (*tetA*), and a small
280 8 kb plasmid (pPan86; GenBank accession number OR797033) encoding resistance to ampicillin
281 (bla_{TEM-1}). Functional annotation of the plasmids identified in isolate 61-66 showed integration of the
282 bla_{TEM-1} gene with a 104 kb IncI1 plasmid (pPan61-IncI1; GenBank accession number OR797031), and
283 a 48 kb IncN plasmid (pPan61-IncN; GenBank accession number OR797032) encoding tetracycline
284 and ampicillin resistance, potentially from the fusion of pPan376-IncN and pPan86 (Figure S2).
285
286 To understand whether these IncN plasmids are driving resistance in recent European clade C2
287 isolates, we screened for the presence of plasmids in the genomes of all C2 isolates. This indicated
288 that 23% (*n* = 27/124) of isolates contained pPan376-IncN, 26% (n = 32/124) contained pPan86, 2%
289 (*n* = 3/124) pPan61-IncI1, and 23% (*n* = 28/124) pPan61-IncN (details in Figure S2 and Table S1). We
290 constructed phylogenies of the two most prevalent IncN plasmids to explore coevolution and
291 determine whether these plasmids are continuing to drive current AMR in *S.* Panama (Figure S2).
292 The results show temporal clustering, indicating that the older and contemporary IncN plasmids are
293 distinct.
294
295 **Phylodynamic and evolutionary timescales of *S.* Panama**
296 To investigate the evolutionary dynamics of the global *S*. Panama population, we inferred the times
297 to most recent common ancestor (MRCA) of each clade using BactDating (Figure 4)[18]. Analysis of the
298 chronograph predicted the MRCA of all isolates in this study to be year 1555 with a highest
299 probability density interval (HPD) of 1458 to 1634. C1 (Latin America & the Caribbean clade) had an
300 MRCA of 1899 (HPD 1880-1911) with separate introductions into Guadeloupe, Martinique, and
301 French Guiana, followed by clonal expansion in the 1930s and 1950s in each of the three regions.
302 The MRCA of C2 (European clade) was 1879 (HPD 1852-1894), followed by expansion of more
303 contemporary subclades in the 1970s. We speculate that C2 arose from an outbreak that involved
304 the continental spread of *S.* Panama across Europe, associated with the food production industry in

305 the 1970s and 1980s[3]. It is possible that these European *S.* Panama were exposed to a significant
306 levels of antibiotics in humans and/or food animals, which selected for resistance to multiple
307 antibiotic agents via acquisition of mobile genetic elements[3], consistent with the indicators of
308 mobilisation above (Figure 3). The emergence of C2 (European clade) in the 1970s is consistent with
309 this hypothesis, coinciding with the propagation of *S.* Panama in Europe via the pig industry. The
310 MRCA of C3 (Martinique clade) is 1880 (HPD 1862-1916), and of C4 (Asia/Oceania clade) is 1879
311 (HPD 1854-1905). These predicted timeframes overlap with European attempts to, and the final
312 construction of, the Panama Canal.
313
314 **The invasiveness index of *S.* Panama**
315 Although *S.* Panama causes extraintestinal disease globally, genetic determinants of invasiveness
316 have never been investigated. In this context, invasiveness reflects the ability of the pathogen to
317 cause systemic spread and extraintestinal infection in humans[19]. Invasive infections cause much
318 higher levels of mortality than gastroenteritis; for example, the case-fatality rate of iNTS is 20%[1],
319 whereas the case-fatality rate of *S.* Typhimurium-associated gastroenteritis is 0·6%[20]. It has been
320 observed that functions required for escalating growth within the inflamed gut are often lost among
321 iNTS variants in a series of gene degradation/pseudogenisation events[21].
322
323 To quantify the relative invasiveness of *S.* Panama clades, we calculated the 'invasiveness index'[22] for
324 all 836 isolates in this study (Table S1). The invasiveness index model relies upon nucleotide
325 differences within 196 genes identified as predictors for invasiveness and has been validated for
326 other *Salmonella* serovars[22]. The mean invasiveness index of *S.* Panama was 0·2294 (SD = 0·01790),
327 slightly higher than that described for serovars with a typically broader host range (*S.* Agona, *S.*
328 Enteritidis, *S.* Heidelberg, *S.* Newport, and *S.* Typhimurium)[22] (Figure S3A). In contrast, the
329 invasiveness index of *S.* Panama is significantly higher that that of host-restricted serovars (Typhi,
330 Paratyphi, Gallinarum, Pullorum, Dublin, and Choleraesuis)[22] (Figure S3A). Within *S.* Panama, C2 had

331 the greatest mean invasiveness index (European clade; mean = 0·2435, SD = 0·01278), followed by
332 C3 (Martinique clade; mean = 0·2311, SD = 0·01087), C4 (Asia/Oceania clade; mean = 0·2302, SD =
333 0·01526), and C1 (Latin America & the Caribbean clade; mean = 0·2270, SD = 0·02079). Importantly,
334 the significantly higher (based on Kruskall-Wallis test) invasiveness index of C2 (European clade)
335 combined with the high proportion of AMR isolates, raises the potential of positioning this clade as a
336 public health issue (Figure S3B).
337
338 **Perspective**
339 As the first large-scale phylogenetic analysis of the *S.* Panama serovar, we have revealed important
340 information about population structure, AMR, global ecology, and invasiveness. We have identified
341 four *S*. Panama clades that share a most recent common ancestor in the 1500s and have undergone
342 contemporary introductions into different geographical regions. The majority of *S.* Panama in this
343 study were pan-susceptible to antibiotics, however there are a number of MDR isolates in C2
344 (European clade) and C4 (Asia/Oceania clade), of which the former presented genomic predictors of
345 elevated invasiveness.
346
347 It will be important to monitor the incidence of clades C2 and C4 in ongoing genome-based
348 surveillance, and to understand the movement of mobile genetic elements responsible for antibiotic
349 resistance. It is possible that the evolutionary trajectory of *S*. Panama will parallel the *Salmonella*
350 serovars Enteritidis and Typhimurium, which contain specific lineages and pathovariants with an
351 increased invasiveness index and are now causing significant levels of human bloodstream
352 infections[1,22,23]. Future work on *S.* Panama can be derived from the current knowledge of serovars
353 Enteritidis and Typhimurium and should focus on understanding the molecular basis of systemic
354 invasion, the development of a robust typing system, and a defined animal infection model.
355

356 ## Methods

357 **Assembly of a *S.* Panama collection with global relevance**

358 *S.* Panama isolates were derived from three archived public health collections: the Unité des
359 Bactéries pathogènes entériques (UBPE), French National Reference Center for *Escherichia coli*,
360 *Shigella*, and *Salmonella*, Institut Pasteur (Paris, France), the National Infection Service, UK Health
361 Security Agency (London, UK), and the Microbiological Diagnostic Unit Public Health Laboratory
362 (MDU PHL) (Melbourne, Australia).

363

364 The isolates were sampled from the Institut Pasteur UBPE's archives which hold all *S.* Panama
365 collected from France and French overseas territories between 1947 and 2018 (>5000 isolates) and
366 also reference strains, such as the 73K strain isolated in Panama in 1931. Our strategy for sampling
367 the Institut Pasteur's collection was designed to capture the temporal and global context and
368 include human (n = 542), environmental (n = 10), and animal isolates (n = 7). Specifically, we selected
369 497 isolates collected for public health purposes between 1993 and 2018 (hereafter termed
370 'national surveillance' isolates), and 62 isolates collected between 1931 and 1997 (hereafter referred
371 to as 'historical' isolates).

372

373 To eliminate redundancy, 'national surveillance' Institut Pasteur isolates taken from the same
374 patient prior to sampling were discarded. All isolates associated with travel outside of France and
375 French overseas territories were selected to capture the maximum global context. All 'extra-
376 intestinal' (defined as blood and cerebrospinal fluid) isolates from French overseas territories and a
377 small number of isolates (n = 90) from Metropolitan France were included. A random selection
378 (using the MS Excel random number generator tool) of one stool isolate per year from each sampling
379 location available was incorporated as non-invasive *S.* Panama isolates.

380

381 Isolates from the Institut Pasteur collection were stored as stab cultures. To resuscitate isolates, a
382 sterile loop was used to transfer culture into 10 mL of Tryptic Soy Broth which was grown overnight
383 at 37°C. To check for potential contaminants, each isolate was also streaked onto Drigalski agar
384 plates, which were visually inspected following overnight growth at 37°C. If colonies were pure on
385 Drigalski agar, 100 µL of culture was aliquoted into a FluidX 2D Sequencing Tube (FluidX Ltd, UK).
386
387 The dataset was complemented by 25 genomes generated by the Microbiological Diagnostic Unit
388 Public Health Laboratory (MDU PHL) at The University of Melbourne, Australia, collected between
389 2005 and 2019[24].
390
391 The United Kingdom Health Security Agency (UKHSA), London, UK. UKHSA (formerly Public Health
392 England) has routinely genome-sequenced every *Salmonella* isolate received by the *Salmonella*
393 reference service since 2015 in addition to a small number of pre-2015 isolates[25]. The resulting
394 genomic data are regularly deposited in publicly available repositories[25]. The UKHSA dataset includes
395 whole genome sequence data from 147 publicly available *S.* Panama collected between 2012 and
396 2019.
397
398 Finally, a set of 105 additional *S.* Panama publicly available genomes obtained through EnteroBase
399 (http://enterobase.warwick.ac.uk) were included, which were collected between 1931 and 2018.
400 The selection was made from 1,266 *S.* Panama genomes (*i.e.* assigned to the hierarchical cluster
401 HC400_369 using the EnteroBase "cgMLST V2 + HierCC V1" tool) on October 2020.
402
403 A description of all the 836 isolates is available with all metadata and genome accession numbers in
404 Table S1. Metadata collected from surveillance questionnaires was used to assign and infer
405 geographical location to each isolate. When a recent trip overseas was reported, the travel
406 destination was considered to be the inferred isolate location[26,27], and travel not reported was

407 considered as the sampling location to be the inferred location. Locations were grouped
408 geographically into six major areas based on the United Nations classification: Africa; Asia; Europe;
409 Latin America & the Caribbean; Northern America, and Oceania.
410
411 **Whole genome sequencing and annotation of *S.* Panama isolates**
412 Cultures from the Institut Pasteur stored in FluidX tubes (*n* = 315) were heat killed in a 95°C water
413 bath for 20 min to produce thermolysates before DNA extraction (MagAttract kit, Qiagen) and
414 whole-genome sequencing, as part of the 10,000 *Salmonella* Genomes Project[28]. Illumina Nextera XT
415 DNA Libraries were prepared (Illumina, FC-131-1096) and sequenced (Illumina HiSeq 4000) in
416 multiplex (768) as 150 bp paired-end reads. The rest of the isolates (*n* = 237) were grown overnight
417 in tryptic soy broth (TSB) at 37°C, DNA was extracted using the MagNA Pure DNA isolation kit (Roche
418 Molecular Systems, Indianapolis, IN, USA) at the Mutualized Platform for Microbiology (P2M) at
419 Institut Pasteur, Paris. The libraries were prepared with the Nextera XT kit (Illumina, San Diego, CA,
420 USA) and sequencing was performed with the NextSeq 500 system (Illumina) generating 150 bp
421 paired-end reads.
422
423 To prepare genome sequence data for downstream analysis, reads were assembled, subject to
424 quality control and annotated. Adapter content and raw read quality was assessed using FastQC
425 v0.11.5 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and MultiQC v1.0
426 (http://multiqc.info). Adapter trimming was then performed using Trimmomatic[29] v0.36 and low-
427 quality regions were trimmed using Seqtk v1.2-r94 (https://github.com/lh3/seqtk) with the trimfq
428 flag. Adapter and quality trimmed reads were assessed for quality using FastQC v0.11.5 and MultiQC
429 v1.0. Quality metrics used to pass sequencing reads were as follows: passed basic quality statistics,
430 per base sequence quality, per base N content and adapter content. GC content distributions were
431 visually inspected to determine any isolates with unusual peaks. Only high-quality reads were used
432 in downstream analysis. Unicycler[30] v0.3.0 was used to assemble genomes and QUAST[31] v4.6.3 was

433 used to evaluate assembly quality to standards consistent with Enterobase[32]. Specifically, N50 > 20
434 kb, 600 or fewer contiguous sequences, total number of bases between 4 and 5·8 Mbp. Prokka[33]
435 v1.12 was used to annotate the genomes. The *Salmonella in Silico* Typing Resource (SISTR)[34] v1.0.2
436 was used to confirm serovar designations.
437
438 **Long-read sequencing and plasmid reconstruction**
439 The *S.* Panama isolates 61-66, 86-66, and 376-66 were cultured overnight at 37°C in alkaline nutrient
440 agar (20 g/L casein meat peptone E2 (Organotechnie), 5 g/L sodium chloride (Sigma), and 15 g/L
441 Bacto agar (Difco), adjusted to pH 8.4). Single colonies were used to inoculate 20 mL of Brain-Heart-
442 Infusion (BHI) broth, and incubated at 37°C with shaking (200 rpm —Thermo Scientific MaxQ 6800)
443 until a final OD_{600} of 0.8 was reached. The bacterial cells were harvested by centrifugation and the
444 DNA extraction was performed by the used of Genomic-tip 100/G column (Qiagen) according to the
445 manufacturer's protocol. The library was prepared according to the instructions of the 'Native
446 barcoding genomic DNA (with EXP-NBD114 and SQK-LSK109)' procedure provided by Oxford
447 Nanopore Technology. Sequencing was performed on a MinION Mk1C apparatus (Oxford Nanopore
448 Technologies). The genomic sequences of the isolates were assembled from long and short reads,
449 with a hybrid approach and Unicycler v0.4.8[30]. A polishing step was performed with Pilon[35] v1.23 to
450 generate a high-quality sequence composed of chromosomal and plasmid sequences. The sequences
451 are available via the folloqing GenBank accession numbers OR797034 (pPan376-IncN), OR797033
452 (pPan86), OR797031 (pPan61-IncI1), and OR797032 (pPan61-IncN).
453
454 **Phylogenetic analysis**
455 Trimmed sequencing data was mapped against *S.* Panama ATCC7378 as reference genome
456 (accession number CP012346) using Snippy v4.6.0 (https://github.com/tseemann/snippy) with
457 minimum coverage of 4 and base quality of 25. QualiMap[36] v2.0 was used to assess mapping quality
458 and only isolates with greater than 10x genome coverage were included in downstream analysis. The

459 mean coverage depth across all isolates was 45·13x. Snippy-core v4.6.0 was used to build a
460 reference-based pseudogenome for each isolate which was passed through snippy-clean v4.6.0 to
461 replace non-ACTG characters with 'N'. Gubbins[37] v2·2 was used to remove recombinant regions and
462 invariable sites. The resultant multiple sequence alignment of reference based pseudogenomes
463 (24,843 variant sites) was used to infer a maximum likelihood phylogeny using RAxML-NG[38] v1.0.3
464 with 100 bootstrap replicates to assess support. To assign clusters, rhierBAPs[10] was used specifying
465 two cluster levels, 20 initial clusters and infinite extra rounds.
466
467 To determine the evolutionary history of *S.* Panama, a chronogram was produced using BactDating[18]
468 v1.1. The maximum likelihood phylogenetic tree was used as an input, accounting for invariant sites.
469 BactDating's model comparison was used to determine the most appropriate clock model (mixed
470 arc). Effective sample size (ESS) for mu = 501, sigma = 501, and alpha = 435. The final model was run
471 for 10,000,000 iterations to reach convergence. All phylogenies were visualised using the Interactive
472 Tree of Life (iTOL)[39] v4.2.
473
474 **Antimicrobial resistance conferring gene identification & plasmid analysis**
475 Genetic determinants for AMR were identified using staramr v0.5.1 (https://github.com/phac-
476 nml/staramr) against the ResFinder[40] and PointFinder[41] databases with default parameters (98%
477 identity threshold and 60/95% overlap threshold respectively). Common resistance patterns were
478 identified and the respective contigs were extracted from the assembled genomes. To understand
479 the genetic context of regions carrying AMR determinants, the contigs were manually inspected
480 using Artemis[42] v10.2. The Prokka-based annotations were confirmed and updated using BLASTx[43]
481 v2.10.1 to compare all coding regions against the non-redundant protein sequence database.
482 PlasmidFinder[44] was used to predict contigs containing plasmids, and the results were compared
483 against contigs containing AMR determinants. Finally BLASTn[43] v2.10.1 was used to understand the
484 wider context of the genetic region.

485

486 Snippy v4.6.0 (https://github.com/tseemann/snippy) was used for mapping all C2 isolates to
487 pPan376-IncN, pPan86, pPan61-IncI1, and pPan61-IncN using minimum coverage of 4 and base
488 quality of 25. Snippy-core v4.6.0 was used to generate a multiple sequence alignment, and snippy-
489 clean v4.6.0 to replace non-ACTG characters with 'N'. Coverage was calculated using SAMtools[45]
490 v1.14, and presence of plasmid was considered with over 50% coverage. A maximum likelihood tree
491 was constructed for pPan61-IncN (alignment length = 47,695) and pPan376-IncN (alignment length =
492 44,651) with isolates that had >50% coverage using RAxML-NG[38] v.1.0.3 with 100 bootstrap
493 replicates to assess support.

494

495 **Determination of invasiveness index**

496 The invasiveness index of each isolate was calculated using previously defined methods, using the
497 196 top predictor genes for measuring invasiveness of *S. enterica*[22]. We used the same pre-trained
498 model to compare the median invasiveness index for *S.* Panama with the invasiveness index of
499 validation strains previously described[22]. The distribution of invasiveness index values for each
500 cluster were compared using the Kruskall-Wallis test implemented through GraphPhad Prism v10.1.
501 A custom-made database of the top 196 invasiveness predictor genes was then created from the
502 multi-fasta file provided in Wheeler *et al.* (2018)[22] (https://github.com/Gardner-
503 BinfLab/invasive_salmonella). SRST2[46] v0.2.0 was used to identify variations between the 196 genes
504 in study isolates.

505

506 ## Role of the funding source

507 The funders of the study had no role in study design, data collection, data analysis, data
508 interpretation, or writing of the report.

509

Acknowledgments

We thank all the corresponding laboratories of the French National Reference Center for *Escherichia coli*, *Shigella*, and *Salmonella*. We thank Dr Edward Cunningham-Oakes and Dr Nicolas Wenner for useful discussions.

Declaration of interests

The authors declare no conflict of interest.

References

1 Stanaway JD, Parisi A, Sarkar K, *et al.* The global burden of non-typhoidal salmonella invasive disease: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet Infectious Diseases* 2019; **19**: 1312–24.

2 Gilchrist JJ, MacLennan CA. Invasive Nontyphoidal Salmonella Disease in Africa. *EcoSal Plus* 2019; **8**. DOI:10.1128/ecosalplus.ESP-0007-2018.

3 Pulford CV, Perez-Sepulveda BM, Rodwell EV, Weill F, Baker KS, Hinton JCD. *Salmonella enterica* Serovar Panama, an Understudied Serovar Responsible for Extraintestinal Salmonellosis Worldwide. *Infection and Immunity* 2019; **87**: 1–10.

4 Marchello CS, Dale AP, Pisharody S, Rubach MP, Crump JA. A Systematic Review and Meta-analysis of the Prevalence of Community-Onset Bloodstream Infections among Hospitalized Patients in Africa and Asia. *Antimicrobial agents and chemotherapy* 2019; **64**. DOI:10.1128/AAC.01974-19.

5 Guyomard-Rabenirina S, Muanza B, Bastian S, *et al.* Salmonella enterica serovars Panama and Arechavaleta: Risk Factors for Invasive Non-Typhoidal Salmonella Disease in Guadeloupe, French West Indies. *The American journal of tropical medicine and hygiene* 2018; **99**: 584–9.

6 Gay N, Le Hello S, Weill F-X, de Thoisy B, Berger F. Salmonella serotypes in reptiles and humans, French Guiana. *Veterinary microbiology* 2014; **170**: 167–71.

7 Williamson DA, Lane CR, Easton M, *et al.* Increasing Antimicrobial Resistance in Nontyphoidal Salmonella Isolates in Australia from 1979 to 2015. *Antimicrobial Agents and Chemotherapy* 2018; **62**: 10.1128/aac.02012-17.

8 EFSA, ECDC. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. *EFSA Journal* 2018; **16**: 262.

9 Parisi A, Crump JA, Stafford R, Glass K, Howden BP, Kirk MD. Increasing incidence of invasive nontyphoidal Salmonella infections in Queensland, Australia, 2007-2016. *PLOS Neglected Tropical Diseases* 2019; **13**: e0007187.

10 Tonkin-Hill G, Lees JA, Bentley SD, Frost SDW, Corander J. RhierBAPS: An R implementation of the population clustering algorithm hierBAPS. *Wellcome open research* 2018; **3**: 93.

11 Neuert S, Nair S, Day MR, *et al.* Prediction of Phenotypic Antimicrobial Resistance Profiles From Whole Genome Sequences of Non-typhoidal Salmonella enterica. *Frontiers in microbiology* 2018; **9**: 592.

12 Lee H-Y, Yang Y-J, Su L-H, Hsu C-H, Fu Y-M, Chiu C-H. Genotyping and antimicrobial susceptibility of Salmonella enterica serotype Panama isolated in Taiwan. *Journal of microbiology, immunology, and infection = Wei mian yu gan ran za zhi* 2008; **41**: 507–12.

13 Nadimpalli ML, Stegger M, Viau R, *et al.* Plugging the leaks: antibiotic resistance at human–animal interfaces in low-resource settings. *Frontiers in Ecology and the Environment* 2023; published online Sept 19. DOI:10.1002/fee.2639.

14 Kudirkiene E, Andoh LA, Ahmed S, *et al.* The Use of a Combined Bioinformatics Approach to Locate Antibiotic Resistance Genes on Plasmids From Whole Genome Sequences of Salmonella enterica Serovars From Humans in Ghana. Frontiers in Microbiology. 2018; **9**: 1010.

15 Klemm EJ, Shakoor S, Page AJ, *et al.* Emergence of an Extensively Drug-Resistant *Salmonella enterica* Serovar Typhi Clone Harboring a Promiscuous Plasmid Encoding Resistance to Fluoroquinolones and Third-Generation Cephalosporins. *mBio* 2018; **9**: 10.1128/mbio.00105-18.

16 Carneiro MRP, Patrício MI de A, Jain S, Rodrigues DDP, Fracalanzza SEL. Meningitis caused by Salmonella enterica serotype Panama in Brazil: first case reported. *Revista da Sociedade Brasileira de Medicina Tropical* 2018; **51**: 244–6.

17 Stanley J, Baquar N, Burnens A. Molecular subtyping scheme for Salmonella panama. *Journal of clinical microbiology* 1995; **33**: 1206–11.

18 Didelot X, Croucher NJ, Bentley SD, Harris SR, Wilson DJ. Bayesian inference of ancestral dates on bacterial phylogenetic trees. *Nucleic Acids Research* 2018; **46**: e134.

19 Langridge GC, Wain J, Nair S. Invasive Salmonellosis in Humans. *EcoSal Plus* 2012; **5**. DOI:10.1128/ecosalplus.8.6.2.2.

20 Jones TF, Ingram LA, Cieslak PR, *et al.* Salmonellosis outcomes differ substantially by serotype. *The Journal of infectious diseases* 2008; **198**: 109–14.

21 Pulford CV, Perez-Sepulveda BM, Canals R, *et al.* Stepwise evolution of *Salmonella* Typhimurium ST313 causing bloodstream infection in Africa. *Nature Microbiology* 2021; **6**: 327–38.

22 Wheeler NE, Gardner PP, Barquist L. Machine learning identifies signatures of host adaptation in the bacterial pathogen Salmonella enterica. *PLoS genetics* 2018; **14**: e1007333.

23 Van Puyvelde S, Pickard D, Vandelannoote K, *et al.* An African Salmonella Typhimurium ST313 sublineage with extensive drug-resistance and signatures of host adaptation. *Nature Communications* 2019; **10**: 4280.

24 Ingle DJ, Andersson P, Valcanis M, *et al.* Genomic Epidemiology and Antimicrobial Resistance Mechanisms of Imported Typhoid in Australia. *Antimicrobial Agents and Chemotherapy* 2021; **65**: 10.1128/aac.01200-21.

25 Ashton PM, Nair S, Peters TM, *et al.* Identification of Salmonella for public health surveillance using whole genome sequencing. *PeerJ* 2016; **4**: e1752.

26 Baker KS, Dallman TJ, Field N, *et al.* Genomic epidemiology of Shigella in the United Kingdom shows transmission of pathogen sublineages and determinants of antimicrobial resistance. *Scientific Reports* 2018; **8**: 7389.

27 Ingle DJ, Nair S, Hartman H, *et al.* Informal genomic surveillance of regional distribution of Salmonella Typhi genotypes and antimicrobial resistance via returning travellers. *PLOS Neglected Tropical Diseases* 2019; **13**: e0007620.

28 Perez-Sepulveda BM, Heavens D, Pulford CV, *et al.* An accessible, efficient and global approach for the large-scale sequencing of bacterial genomes. *Genome Biology* 2021; : 2020.07.22.200840-2020.07.22.200840.

29 Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* 2014; **30**: 2114–20.

30 Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS computational biology* 2017; **13**: e1005595.

31 Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics (Oxford, England)* 2013; **29**: 1072–5.

32 Alikhan N-F, Zhou Z, Sergeant MJ, Achtman M. A genomic overview of the population structure of Salmonella. *PLoS genetics* 2018; **14**: e1007261.

33 Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics (Oxford, England)* 2014; **30**: 2068–9.

34 Yoshida CE, Kruczkiewicz P, Laing CR, *et al.* The Salmonella In Silico Typing Resource (SISTR): An Open Web-Accessible Tool for Rapidly Typing and Subtyping Draft Salmonella Genome Assemblies. *PloS one* 2016; **11**: e0147101.

35 Walker BJ, Abeel T, Shea T, *et al.* Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLOS ONE* 2014; **9**: e112963.

36 García-Alcalde F, Okonechnikov K, Carbonell J, *et al.* Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* 2012; **28**: 2678–9.

37 Croucher NJ, Page AJ, Connor TR, *et al.* Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Research* 2014; **43**: e15–e15.

38 Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 2019; **35**: 4453–5.

39 Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic acids research* 2016; **44**: W242-5.

40 Zankari E, Hasman H, Cosentino S, *et al.* Identification of acquired antimicrobial resistance genes. *The Journal of antimicrobial chemotherapy* 2012; **67**: 2640–4.

41 Zankari E, Allesøe R, Joensen KG, Cavaco LM, Lund O, Aarestrup FM. PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. *Journal of Antimicrobial Chemotherapy* 2017; **72**: 2764–8.

42 Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics (Oxford, England)* 2012; **28**: 464—469.

43 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of molecular biology* 1990; **215**: 403–10.

44 Carattoli A, Zankari E, García-Fernández A, *et al.* In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrobial agents and chemotherapy* 2014; **58**: 3895–903.

45 Li H, Handsaker B, Wysoker A, *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–9.

46 Inouye M, Dashnow H, Raven L-A, *et al.* SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome medicine* 2014; **6**: 90.

Figure and Table legends

Figure 1. The sample collection used to construct the *S.* Panama population structure.

(A) Total number of isolates per collection. Stacked colour bars on the left are proportional to the number of isolates from that source, and colours act as a key for histograms to indicate source. Airplane symbols show the number of travel-associated isolates obtained by the Institut Pasteur and United Kingdom Health Security Agency (UKHSA). Three stacked histograms show the number of isolates from UKHSA, Institut Pasteur, and MDU PHL (upper graph), publicly available data (middle graph), and historical collections (lower graph). (B) Number of isolates per region (ambiguous or outdated locations registered as "Caribbean", "Asia", "Antilles", and "Yugoslavia" are omitted from the map), with regions shaded according to the inlaid key. Map created with MapChart (http://mapchart.net). (C) Number of isolates categorised by source and, for human isolates, specimen type. Three isolates of unknown source are indicated in grey. Shape sizes are proportional to the number of isolates.

Figure 2. Population structure and geographical distribution of *S.* Panama isolates

Maximum likelihood phylogenetic tree of 837 isolates (including reference isolate ATCC7378 marked by a purple arrow) showing clade (first column) and subclade (second column) assignment based on rhierBAPS, inferred region (third column), and number of isolates (fourth column) as a graph with countries (x-axis) clustered and coloured by the United Nations geographical regions. Data for two 'other' clades (asterisked) representing polyphyletic diversity were combined into a single (asterisked) histogram marked in grey.

Figure 3. Antimicrobial resistance trends in *S.* Panama

(A) Maximum likelihood phylogenetic tree showing clades (first column), inferred region (second column) coloured based on UN geographical regions, and AMR (third column) presence of at least one resistance conferring gene (red) or absence (grey). (B) Number of isolates (coloured by clade)

with different AMR profiles. Resistance to different antimicrobials is shown as present (dark circles) or absent (light circles). Dark circles are connected to highlight a variety of unique AMR profiles. Functional annotation of contiguous sequences with the two predominant AMR profiles (respectively indicated by grey arrows pointing to sequence schematics). Arrows indicate predicted reading frames highlighting genes encoding AMR (red) and transposases (light blue).

Figure 4. Phylodynamic and evolutionary timescales of *S.* Panama

Chronograph for estimation of most recent common ancestors for *S*. Panama clades. Arrows indicate MCRA for each clade, with year and highest probability density (95%) interval shown as text.

Supplementary data:

Table S1: *S.* Panama metadata and accession numbers

Figure S1: Functional annotation of five AMR gene cassettes in XDR isolate 201209115

The predicted phenotype conferred (left) and total length (right) for each contiguous sequence that carries an AMR gene cassette is displayed. Arrows indicate open reading frames, and AMR genes are highlighted in red. GenBank accession number SAMEA6142076.

Figure S2: Conservation and functional annotation of plasmids within C2 ascertained with long-read sequencing technology

(A) Maximum likelihood phylogenetic tree of C2 showing subclades (column one), inferred region (column two) of isolation coloured by UN geographical regions in Figure 2, year (column three), AMR (column four) presence (red) or absence (grey), and percentage coverage (>50%) of pPan61-IncI1 (column five), pPan61-IncN (column six), pPan376-IncN (column seven), and pPan86 (column eight). (B) Pairwise comparison of linearised plasmids showing genomic similarity (red) of pPan376-IncN and pPan86 with pPan61-IncN. GenBank accession numbers are shown below the plasmid name. (C)

Phylogenetic trees of pPan376-IncN and pPan61-IncN showing year of isolation (first column) and AMR (following column(s)). The AMR profile of all isolates in pPan376-IncN tree was streptomycin, ampicillin, chloramphenicol, trimethoprim, sulfisoxazole, tetracycline. The sequenced plasmids that acted as reference sequences are indicated by a star. (D) Functional annotation of genes on pPan61-IncI1, pPan61-IncN, pPan376-IncN, and pPan86.

Figure S3: Invasiveness index of *S.* Panama clades in comparison with common *S. enterica* serovars

(A) Median invasiveness index of different *Salmonella* serovars as calculated in Wheeler et al (2018) (black dots) with *S.* Panama (this study) highlighted in pink. (B) Distribution of invasiveness index of all 836 isolates clustered by clade. Overlying box and whiskers plots indicate median and interquartile range, which are enumerated below in the table. Comparisons between groups are indicated by lines uppermost with *p*-values <0.001 indicated by four asterisks.

Figure 1: The sample collection used to construct the *S.* Panama population structure

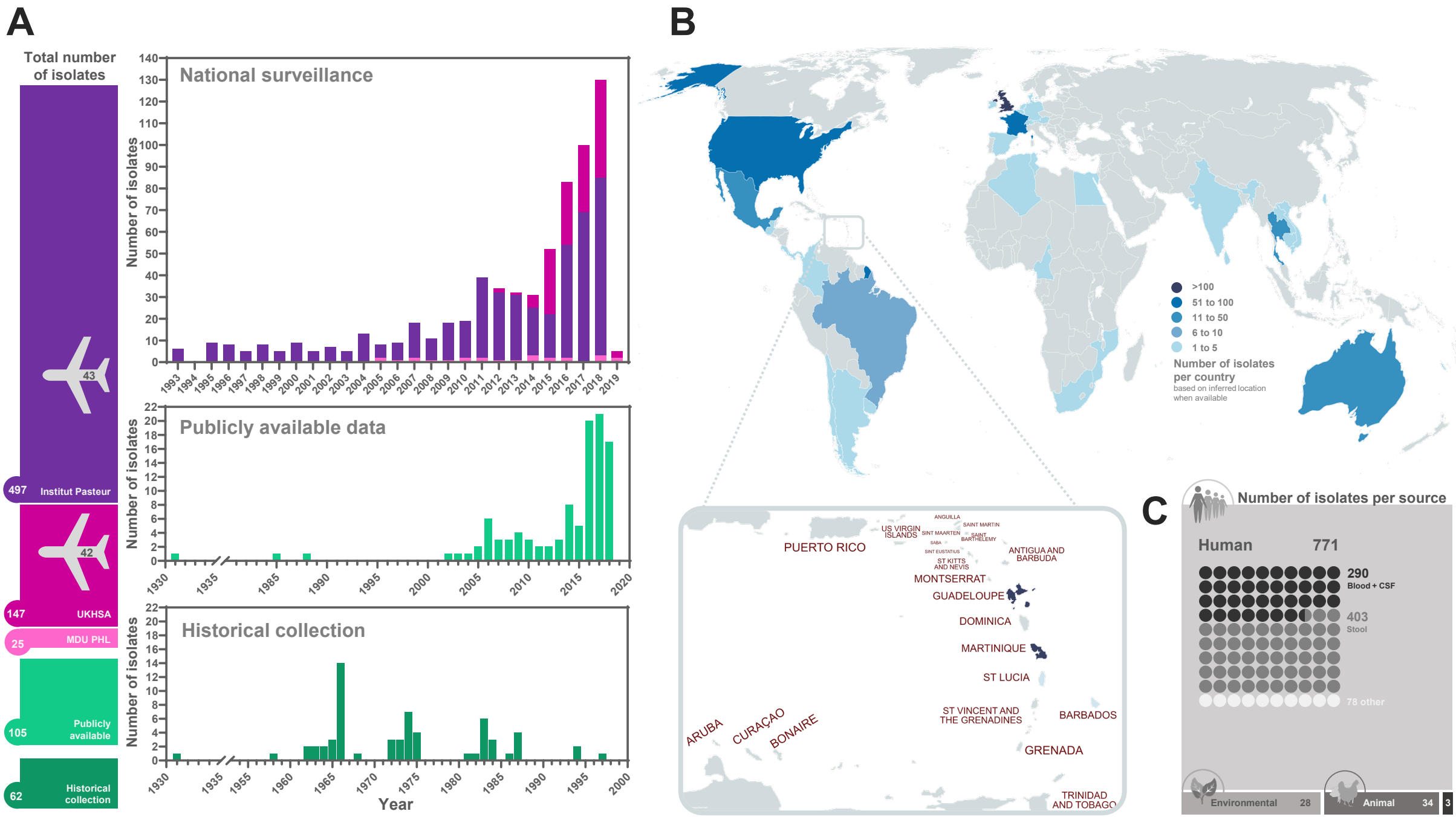

Figure 2. Population structure and geographical distribution of *S.* Panama isolates

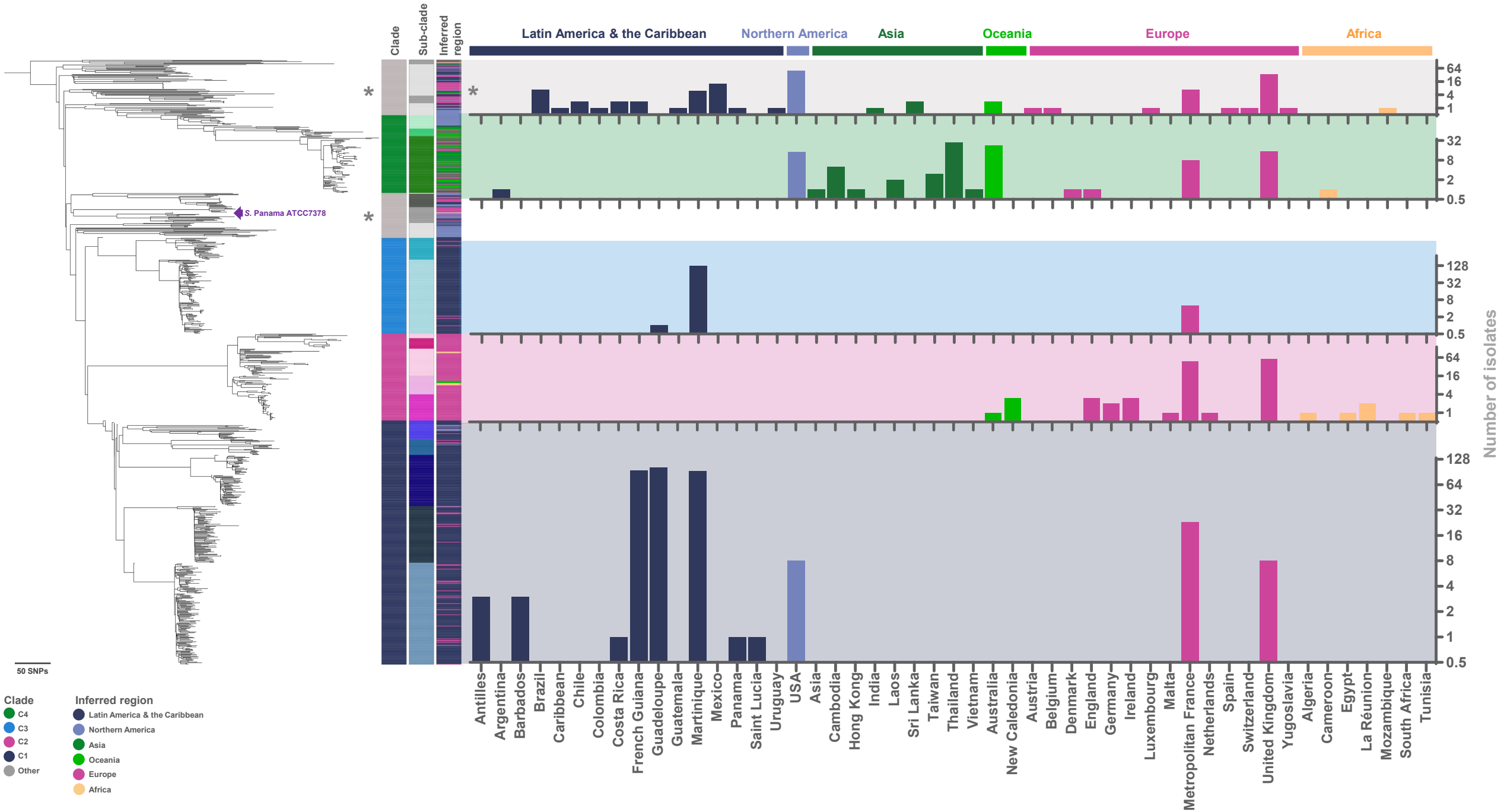

Figure 3. Antimicrobial resistance trends in *S.* Panama

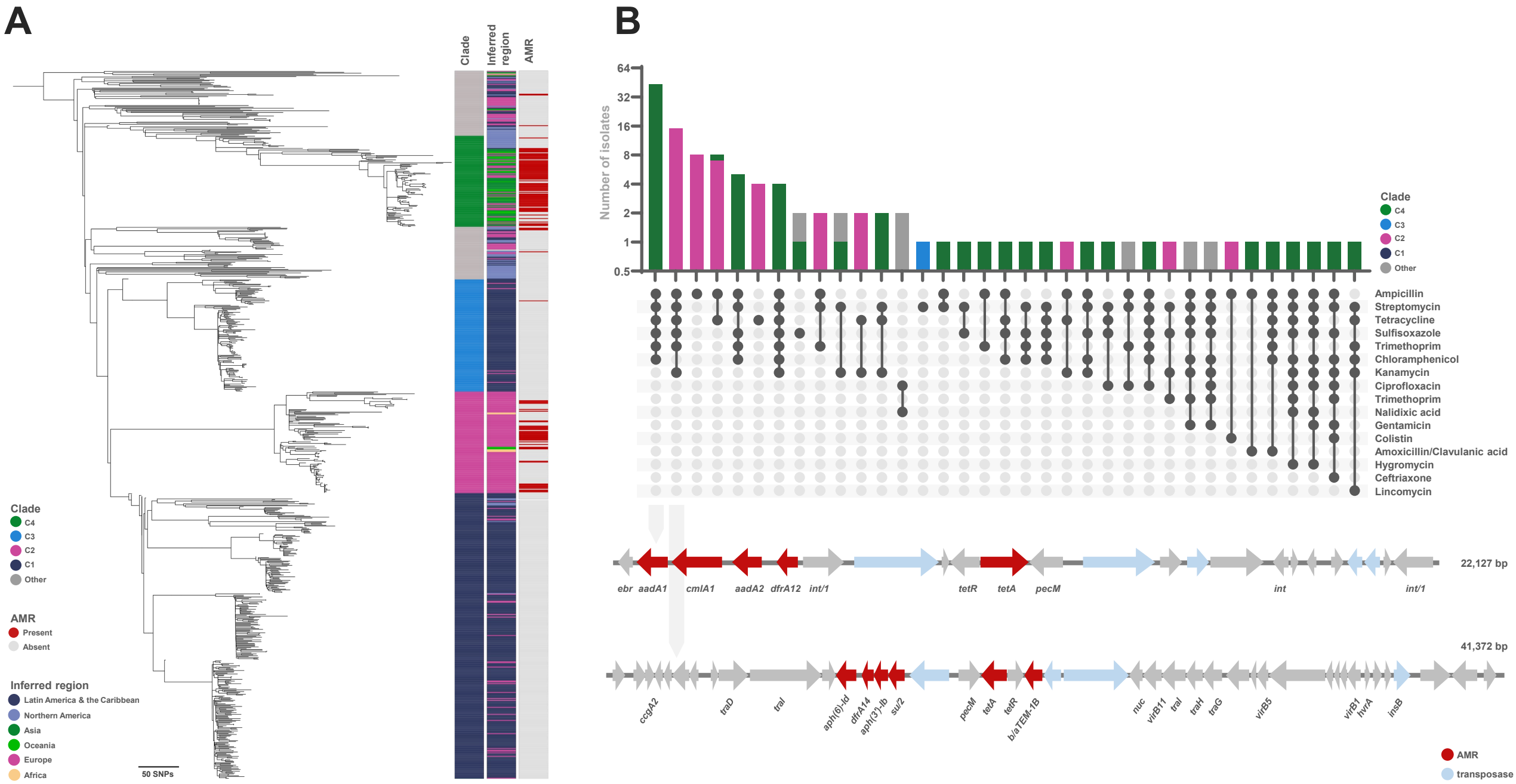
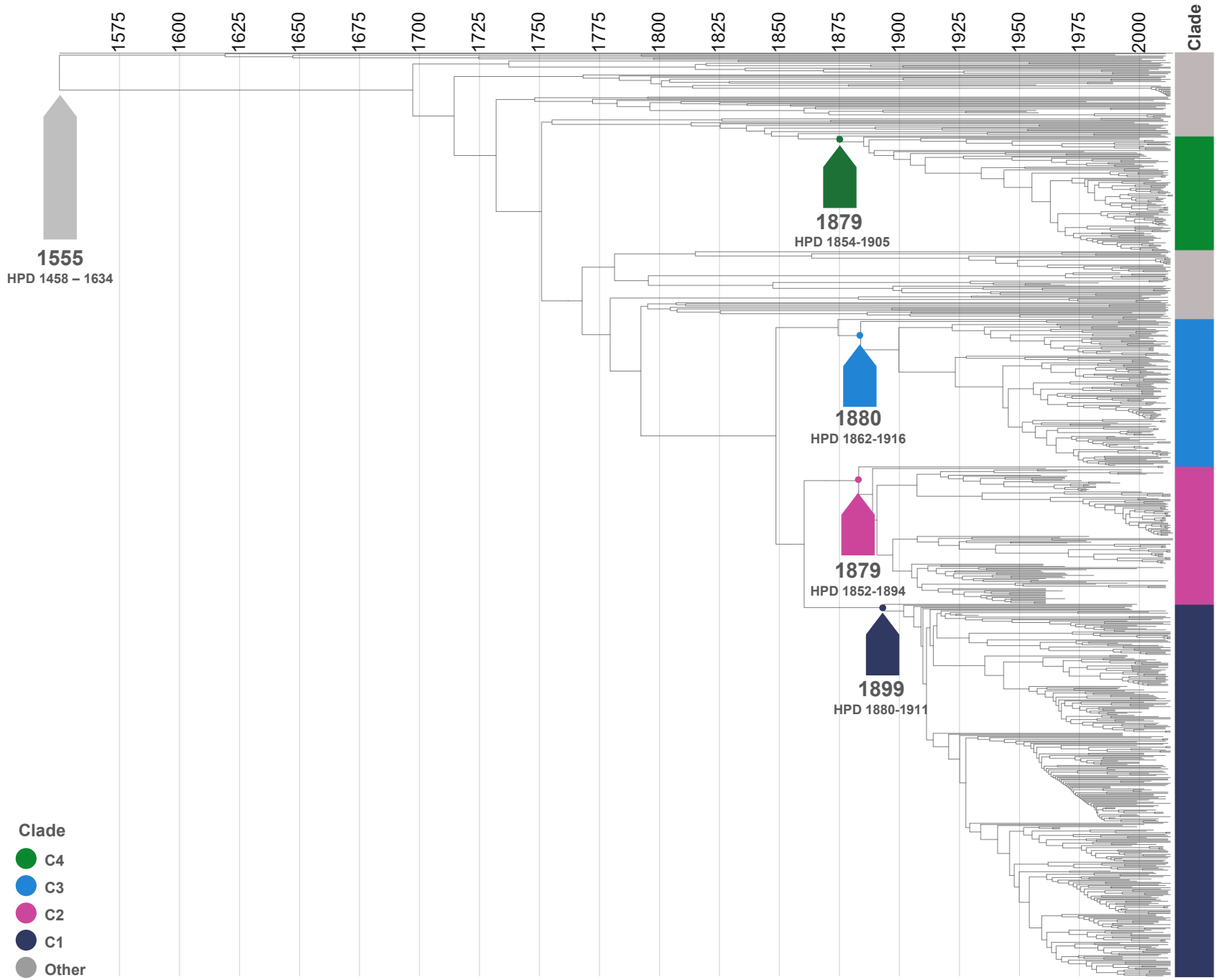
Figure 4 Phylodynamic and evolutionary timescales of S. Panama