bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 # Invasive atypical non-typhoidal *Salmonella* serovars in The Gambia
2

3 Abdoulie Kanteh[1*], Abdul Karim Sesay[1], Nabil-Fareed Alikhan[5], Usman Nurudeen
4 Ikumapayi[1], Rasheed Salaudeen[1], Jarra Manneh[1] Yekini Olatunji[1], Andrew J
5 Page[5] and Grant Mackenzie[1,2,3,4]
6
7 [1]Medical Research Council Unit The Gambia at the London School of Hygiene and
8 Tropical Medicine, Atlantic Boulevard, Fajara, PO Box 273, Banjul, The Gambia.
9 [2]Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical
10 Medicine, Keppel Street, London, WC1E 7HT, United Kingdom.
11 [3]Murdoch Children's Research Institute, Royal Children's Hospital Flemington Road,
12 Parkville Victoria 3052 Australia
13 [4]Department of Paediatrics, University of Melbourne, Melbourne, Australia
14 [5]Quadram Institute Bioscience, Norwich Research Park, Norwich, Norfolk, NR4 7UA,
15 UK.
16
17 * Corresponding author: akanteh@mrc.gm
18

19

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

20 **Abstract**

21 **Background**

22 Invasive non-typhoidal *Salmonella* (iNTS) disease continues to be a significant public
23 health problem in sub-Saharan Africa. Common clinical misdiagnosis, antimicrobial
24 resistance, high case fatality and lack of a vaccine make iNTS a priority for global
25 health research. Using whole genome sequence analysis of 164 invasive *Salmonella*
26 isolates obtained through population-based surveillance between 2008 and 2016, we
27 conducted genomic analysis of the serovars causing invasive *Salmonella* diseases in
28 rural Gambia.

29 **Results**

30 The incidence of iNTS varied over time. The proportion of atypical serovars causing
31 disease increased over time from 40% to 65% compared to the typical serovars
32 Enteritidis and Typhimurium decreasing from 30% to 12%. Overall iNTS case fatality
33 was 10% with 10% fatality in cases of atypical iNTS. Genetic virulence factors were
34 identified in 14/70 (20%) typical serovars and 45/68 (66%) of the atypical serovars and
35 were associated with: invasion, proliferation and/or translocation (Clade A); and host
36 colonization and immune modulation (Clade G). Among Enteritidis isolates, 33/40
37 were resistant to ≥4 the antimicrobials tested, except for ciprofloxacin, to which all
38 isolates were susceptible. Resistance was low in Typhimurium isolates, however, all16
39 isolates were resistant to gentamicin.

40 **Conclusion**

41 The increase in incidence and proportion of iNTS disease caused by atypical serovars
42 is concerning. The increased proportion of atypical serovars and the high associated
43 case fatality may be related to acquisition of specific genetic virulence factors. These
44 factors may provide a selective advantage to the atypical serovars. Investigations

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

45 should be conducted elsewhere in Africa to identify potential changes in the
46 distribution iNTS serovars and the extent of these virulence elements.

47 **Keywords:** Invasive non-typhoidal salmonella, Whole genome sequencing,
48 Cytolethal distending toxin gene, atypical serovar

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

49 **Introduction**

50 The species *Salmonella enterica* (*S. enterica*) is a phenotypically diverse Gram-
51 negative bacterial species, consisting of more than 2,600 serovars. Some serovars
52 are implicated in life-threatening systemic infections and are host-restricted to
53 humans[1]. These include *Salmonella enterica* serovar Typhi and *Salmonella enterica*
54 serovar Paratyphi (*S. Paratyphi* A-C). In contrast, non-typhoidal *Salmonella* species
55 infect both humans and animals[2]; *Salmonella enterica* serovar Typhimurium and
56 *Salmonella enterica* serovar Enteritidis are the most commonly reported in association
57 with *Salmonella* gastroenteritis[3]. Globally, these serovars are responsible for circa 75
58 million cases and 27,000 deaths annually[3].

59 In sub-Saharan Africa, in addition to causing gastroenteritis, non-typhoidal *Salmonella*
60 (NTS) cause life-threatening infections including septicaemia, pneumonia and
61 meningitis[4]. Circa 3.4 million cases of invasive *Salmonella* caused by NTS (iNTS) are
62 reported annually, with Typhimurium and Enteritidis being responsible for 80 - 90% of
63 these cases[5]. The majority of these infections affect children, and are often associated
64 with Human Immunodeficiency Virus (HIV) infection, prior malarial infection, severe
65 anaemia or malnutrition, and case fatality of up to 25%[6–9]. In adults, HIV infection is
66 associated with iNTS disease and case fatality up to 50% has been reported[7–9]. In
67 some parts of Africa, the burden of iNTS disease is higher than that of pneumococcus,
68 infecting tens of thousands of people[7–9]. In The Gambia, iNTS disease in children
69 ranks third after *Streptococcus pneumoniae* and *Staphylococcus aureus* as a cause
70 of invasive bacterial disease[10]. Despite the burden of this disease in our setting, the
71 genomic epidemiology of NTS is still poorly understood.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

72 Susceptibility to invasive *Salmonella* disease could be attributed to host genetic
73 background and immunological status[4]. However, some serovars are known to cause
74 bacteraemia more frequently than others, signifying the importance of pathogen
75 characteristics. For example, a high burden of invasive disease caused by a specific
76 genotype of *S.* Typhimurum has been associated with host adaptation as a result of
77 extensive genomic degradation and acquisition of resistance genes[11]. In addition, the
78 virulence factor cytolethal distending toxin gene *(CdtB)* is known to contribute to
79 variation in disease severity in some NTS serovars[12]. The *CdtB* gene, which was
80 thought to be unique to *Salmonella* Typhi, has been associated with increased host
81 colonization, tumorigenesis, neoplastic lesions[13] and DNA damage similar to that
82 caused by serovar Typhi[13]*.* The presence of the gene in Typhi is associated with host
83 immune modulation as well as persistence of the pathogen in *vivo*[12]. Recently, the
84 presence of *CdtB* has also been documented in NTS serovars and is believed to be
85 clade associated[12]. Thus, the presence of this virulence gene in NTS serovars could
86 influence the virulence of these strains.

87 During population-based invasive bacterial disease surveillance in rural Gambia
88 between 2008 and 2016, we observed changes in the incidence, case fatality, and
89 distribution of iNTS serovars. Surveillance in the same location from 2000 to 2004
90 documented Enteritidis and Typhimurium as the dominant iNTS serovars[14]. Although
91 shifts in *Salmonella* serovar prevalence and dominance have been documented in
92 The Gambia and elsewhere in the world[14,15,16], the genomic characteristics and
93 epidemiological factors responsible for this shift are unclear. We used whole genome
94 sequencing and bioinformatic analyses to investigate changes in pathogen
95 characteristics between 2008 and 2016.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

96 **Material and methods**

97 **Disease surveillance**

98 The surveillance methodology has been previously described[17]. We conducted
99 population-based surveillance for invasive bacterial disease in individuals aged 2
100 months and older resident in the Basse Health and Demographic Surveillance System
101 in Upper River Region, The Gambia[17]. We used standardised criteria to identify and
102 investigate patients presenting with suspected pneumonia, septicaemia, or meningitis
103 to all health facilities in the study area between May 12, 2008 and December 31, 2016.
104 Blood, cerebrospinal fluid (CSF), and lung aspirates (LA) were collected according to
105 standardised criteria and we used conventional microbiological methods to culture and
106 identify bacterial pathogens. Gram negative isolates were identified as *Salmonella*
107 biochemically using a commercial kit (Analytic Profile Index 20E) and antimicrobial
108 susceptibility testing was done using the disk diffusion method and following CLSI
109 reference thresholds[18].

110 **Domestic animal ownership**

111 Given that NTS also infects domestic animals, they can represent an important route
112 of transmission. Data from the Global Enteric Multicentre Study[19] collected in the study
113 area between 2007 and 2012 were used to compare changes in the prevalence of
114 domestic animal ownership and invasive *Salmonella* over time..

115 **Sample population**

116 We analysed 164 *Salmonella* genomes from isolates obtained from blood, CSF or LA
117 samples collected during the surveillance. We extracted genomic DNA from the

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

118 isolates that was sent to the Wellcome Sanger Institute, United Kingdom for whole
119 genome sequencing.

120 **Quality Control, Assembly and Resistance genes**
121

122 Extracted DNA was sequenced using the Illumina Hiseq 2500 platform, to produce
123 sequencing reads of 125 base pairs in FASTQ format [20]**,** with a minimum target depth
124 coverage of 50X. The reads and genomes were quality checked using FASTQC
125 (v0.11.5) and an in-house pipeline, with manual review. The reads were of high quality
126 with an average Phred score of 30 and thus did not require any trimming. Spades
127 (v3.13.1) was used to perform *de novo* assembly with default settings[21] to produce
128 draft assemblies in FASTA format. Quast (v5.0.2)[22] was used to assess the quality of
129 assemblies. Contigs shorter than 300bp were removed from the assemblies as per
130 Page *et al.*,[23]. Four genomes were significantly larger (six Mbases) than the rest of the
131 genomes indicating contamination and were therefore removed from the analysis.
132

133 We used Abricate (v0.9.8) to identify antimicrobial resistance genes, plasmids and
134 virulence genes for each assembly using the comprehensive antimicrobial resistance
135 database (CARD)[24] (downloaded 24-10-2019), Resfinder[25] (downloaded 10-9-2019),
136 PlasmidFinder[26] (downloaded 10-9-2019) and the virulence factor database (VFDB)[27]
137 (downloaded 18-09-2019). A minimum nucleotide identity and coverage of 98% was
138 used for all databases. Virulence factors universally present in *Salmonella* were
139 excluded. The multilocus sequence type (MLST) of each draft genome was predicted
140 using mlst (v2.8) with default settings against the *Salmonella enterica* MLST scheme
141 in the PubMLST database[28].

142 **Phylogenetic analysis**

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

143 Sequencing reads were mapped to the *Salmonella enterica* serovar Typhimurium LT2
144 reference genome (accession number GCF_000006945.2) using Snippy (v4.0.7) with
145 default settings. Single nucleotide polymorphisms (SNPs) from the core genome
146 alignment were used to construct a maximum likelihood phylogenetic tree using the
147 general time-reversible model with IQTREE (v1.3.11.1)[29] and 1000 bootstrap for
148 branch length. Interactive Tree of Life (ITOL) (v5)[30] was used to visualise and annotate
149 the phylogenetic tree. Where particular serovars appeared to have developed into an
150 outbreak they were analysed phylogenetically with other isolates from outside our
151 study. In addition, when genotypes (or STs) were identified that were known to be
152 resitricted elsewhere in the world, phylogenetic comparisons were made to determine
153 whether they were related.

154 **Pan and accessory genome analysis**

155 We used Prokka (v1.13.3)[31] to annotate and predict coding genes from the assembled
156 genomes using *S.* Typhimurium LT2 protein sequences from GenBank to provide high
157 quality species-specific gene name annotation. The resulting GFF3 files were used as
158 input to Roary (v3.13.2)[32] to generate a pan-genome, producing an analysis of the
159 core and accessory genome.

160

161 **Statistical analysis**

162 Summary statistics were prepared using proportions for categorical and
163 mean/median/range for continuous variables including demographic and baseline
164 characteristics. We used Fisher's exact test for associations between categorical
165 variables. All data management and statistical analyses were performed using the R
166 statistical package.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

167

168 **Results**

169 **Demographic data**

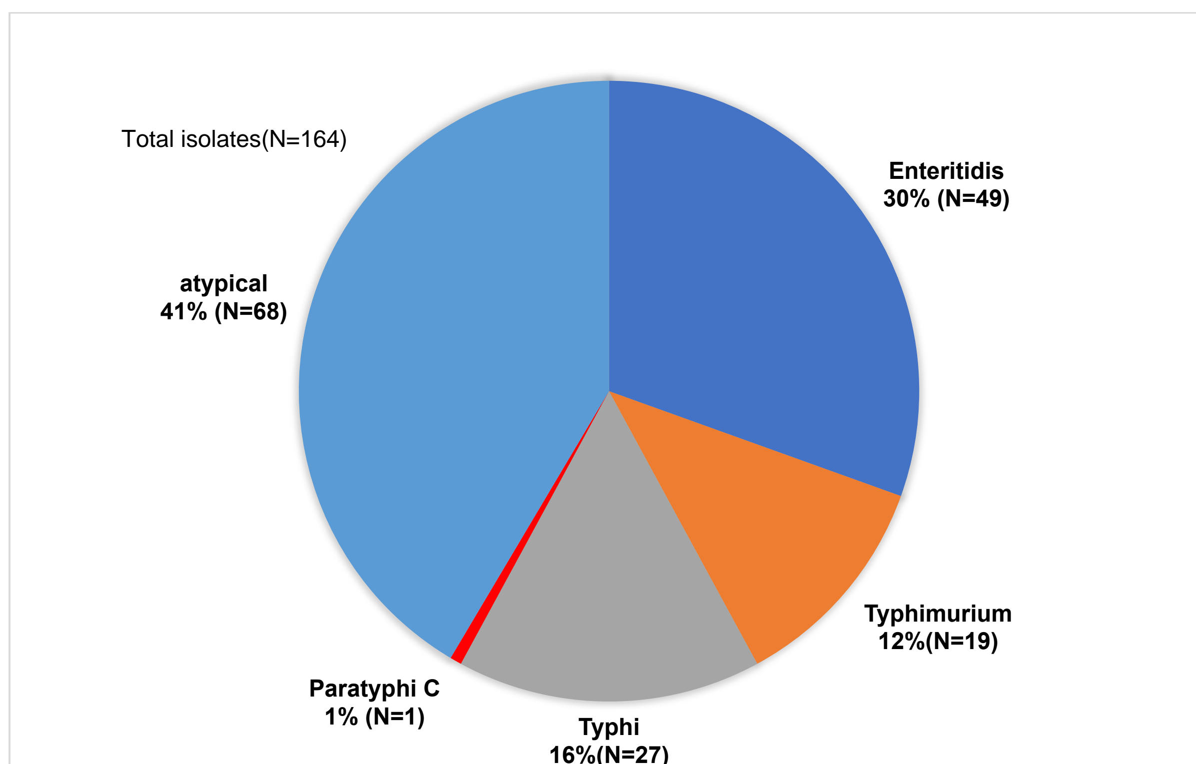
170 Between 2008 and 2016, 22,305 patients were enrolled in the surveillance with 20,199
171 microbiological cultures, an average 2,244 per year (range: 1,047 – 2,370) (Table 1).
172 Patient characteristics are shown in Table 2. From all cultures collected, 164
173 *Salmonella* isolates were obtained from 157 patients. Patient age ranged from 3 days
174 to 42 years with children aged <5 years representing more than 90% (n=145) of the
175 cases. By sample type, 157 isolates were from blood, six from CSF and one from LA.
176 Six patients had isolates detected from more than one clinical sample type.

177 **Genomic analysis**

178 MLST analysis revealed 31 distinct serovars and 45 sequence types (ST). We
179 detected 27 serovars that were not Enteritidis, Typhimurium, Typhi or Paratyphi. We
180 grouped these isolates and called them atypical serovars. A considerable proportion,
181 41% (n=68) of isolates were atypical. The atypical serovars most commonly isolated
182 were Dublin (n=14) Virchow (n=7) and Poona (n=5). Enteritidis, Typhimurium and
183 Typhi constituted 30% (n=49), 12% (n=19) and 16% (n=27) of the isolates,
184 respectively. Only one isolate was S*almonella enterica* serovar Paratyphi C of ST 3039
185 (Figure 1).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

186



187

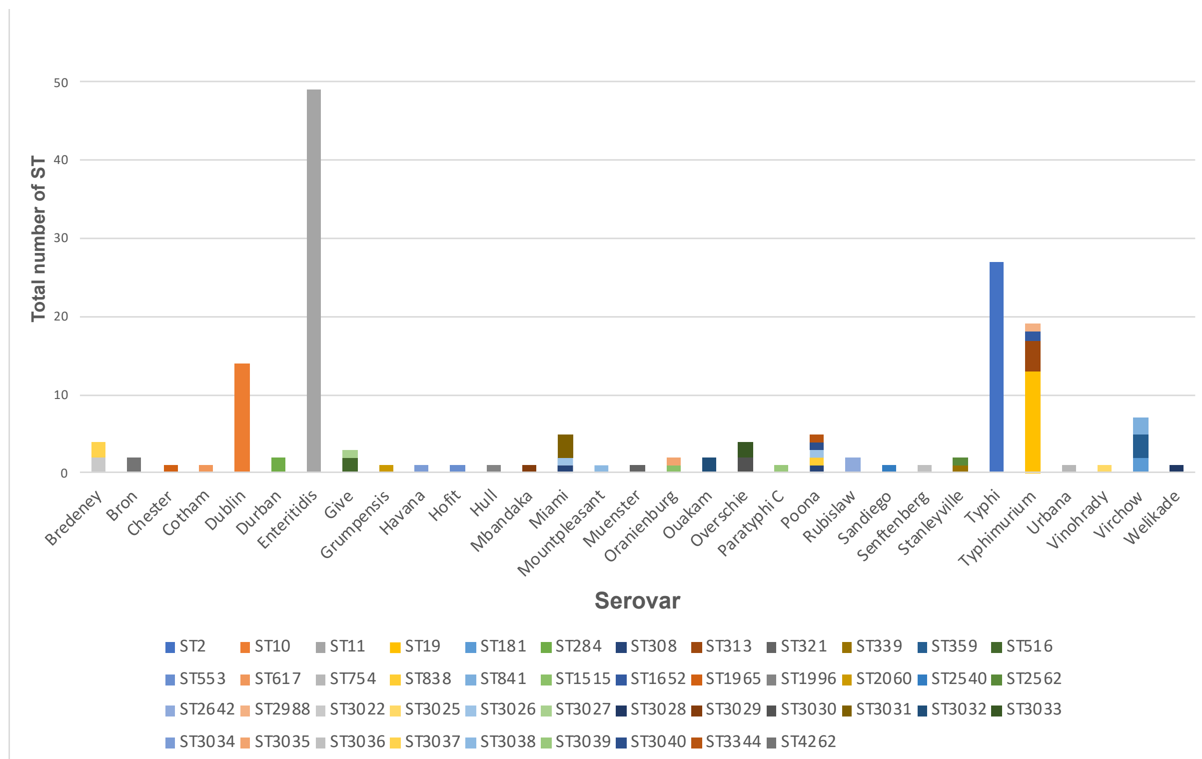
188 **Figure 1.** Breakdown of invasive *Salmonella* serovars isolated between 2008 and
189 2016 from patients in rural Gambia.

190

191 Of all the STs, ST11 was dominant, representing 30% (n=49) of the isolates, followed
192 by ST2 which accounted for 16% (n=27). ST10 and ST19 represented 9% (n=14) and
193 8% (n=13) of the isolates respectively. Other STs included ST313 (n=4), ST3031 (n=3)
194 and ST359 (n=3). Isolates of Typhimurium were represented by four STs: ST19,
195 ST313, ST2988 and ST165. Serovars Virchow and Poona were represented by three
196 and four STs, respectively. Some atypical serovars, including Bredeney, Give, Miami,
197 Oranienburg, Overschie, Poona, Stanleyville and Virchow, were represented by two
198 or more STs each. In contrast, serovars Enteritidis, Typhi and Dublin were
199 represented by only one ST each: ST11, ST2 and S10 respectively (Figure 2).

200

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



201

202 **Figure 2.** Representation of STs amongst invasive *Salmonella* serovars isolated
203 between 2008 and 2016 from patients in rural Gambia.

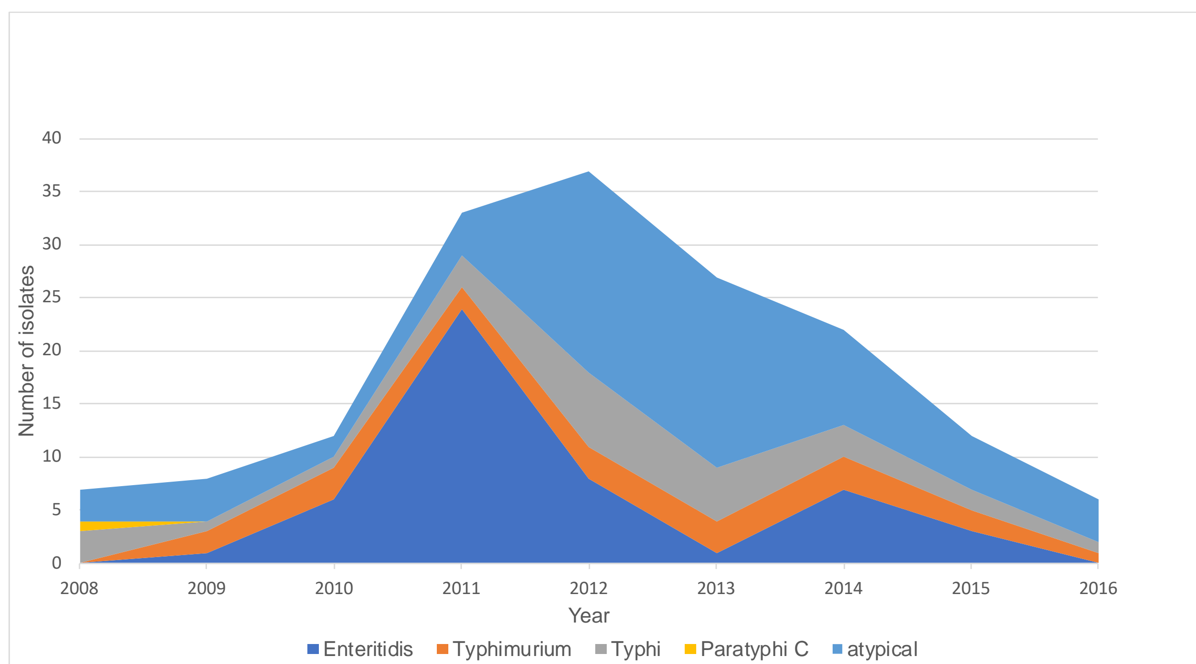
204

205 **Distribution of *Salmonella* serovars over time**

206 During 2000 to 2004 serovars Enteritidis (81%) and Typhimurium (8%) were the
207 dominant iNTS serovars [14]. Over the study period, we observed an increase in the
208 proportion of atypical serovars (Figure 3). In 2008 and 2009, invasive *Salmonella*
209 infection caused by atypical serovars accounted for the majority of cases compared
210 with infection caused by Enteritidis and Typhimurium. However, this trend changed in
211 2011 when Enteritidis became predominant and accounted for about 80% of all
212 *Salmonella* cases. A high proportion of atypical serovars was then observed between
213 2012 and 2014. Overall, from 2012 to 2014, atypical serovars were responsible for
214 almost 50% of *Salmonella* infections. The major serovars within this group included
215 Dublin, Bredeney, Miami and Overchie. From 2015 to 2016, we observed a further

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

216 decline in the proportion of Enteritidis and Typhimurium serovars in the population,
217 while atypical serovars were associated with over 50% of cases.



218

219 **Figure 3.** Case counts of each type of invasive *Salmonella* serovar in Basse, rural
220 Gambia between 2008 and 2016.

221

222 **Incidence and case fatality rate**

223 Amongst all cases of invasive *Salmonella* disease, case fatality rate was 10%
224 (16/157). Case fatality for atypical serovars was 10% (7/68) and 12% (6/49) for
225 Enteritidis. Typhi, Typhimurium and Paratyphic C were associated with only one death
226 each. Amongst hospitalised patients, Enteritidis and atypical serovars accounted for
227 42% (32/77) and 31% (24/77) of cases while Typhi and Typimurium accounted for
228 16% (12/77) and 13% (10/77) of cases, respectively. Amongst atypical serovars, those
229 with the cytolethal toxin gene *CdtB* were responsible for 10% (3/31) of all deaths while
230 atypical seovars without the toxin gene accounted for 11% (4/37) of all deaths.

231 The majority of the patients (59%) had suspected pneumonia or septicaemia (29%).
232 Of the 46 patients with septicaemia, 26 (56%) were infected with atypical serovars;

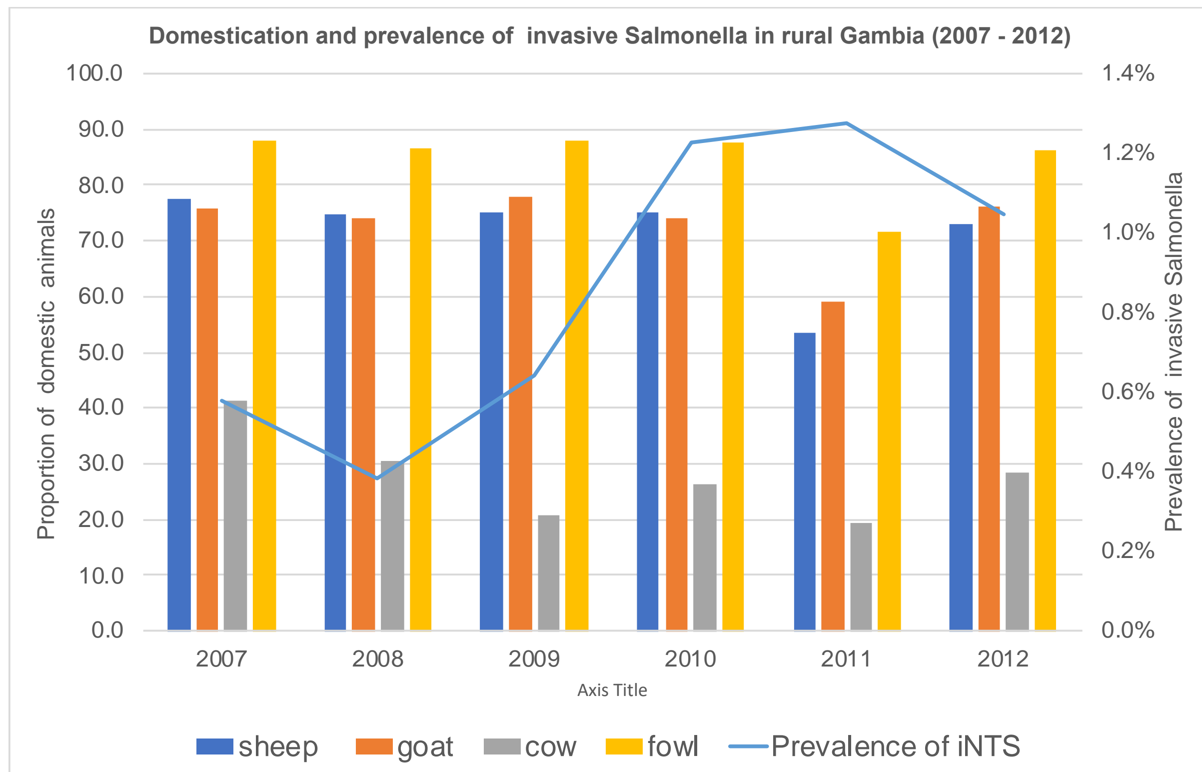
bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

233 Dublin, Overchie, Bredeney and Poona accounted for most of these cases. Overall,
234 we did not find a statistically significant association between malnutrition and any
235 specific serovar though this should be interpreted with caution due to small numbers.
236 However, comparing typical vs atypical serovars, the proportion of children with severe
237 acute malnutrition 19/32 (59%) appeared to be higher in the atypical group compared
238 to Enteriditis 6/32 (18%), Typhimurium 3/32 (9%) or Typhi 4/32 (12%),p-value=0.05.
239

240 **Domestic animal ownership and prevalence of NTS over time**

241 The prevalence of invasive *Salmonella* increased from 2007 to 2010 while domestic
242 animal ownership by households remained constant throughout this period (Figure 4).

243

244


bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

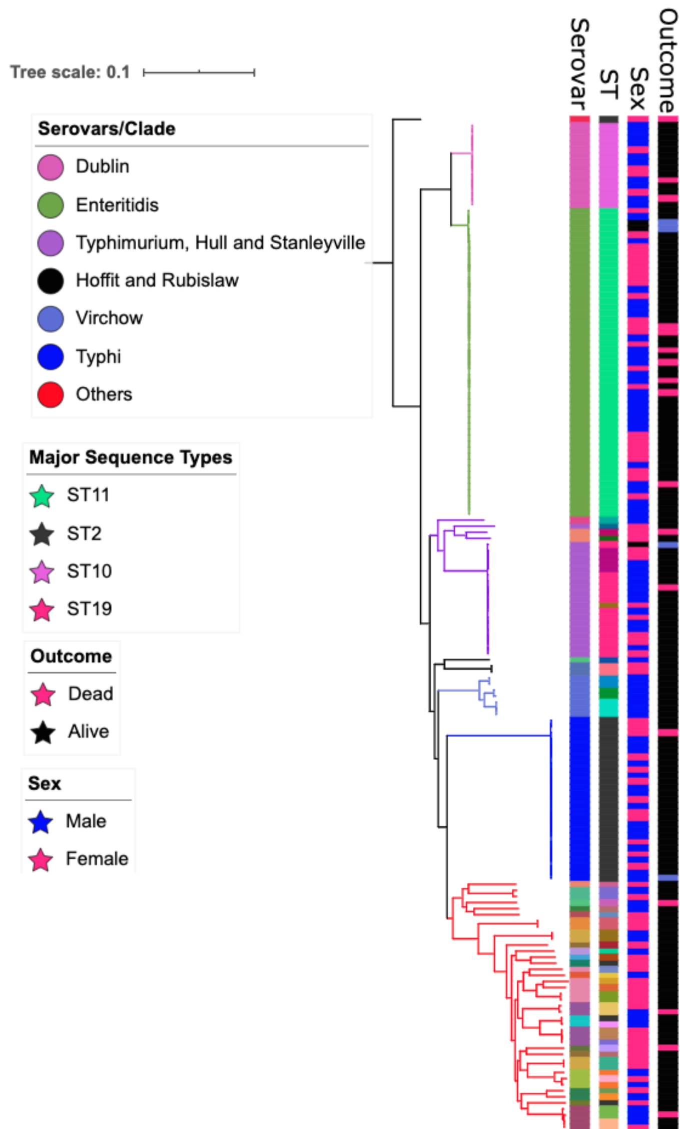
245 **Figure 4**. Relationship between invasive Salmonella disease incidence (blue line) and
246 the proportion of different species of domestic animals reared in rural Gambia between
247 2007 and 2012.

248

249 **Phylogenetic analysis**

250 We constructed a pan-*Salmonella* phylogenetic tree using single nucleotide
251 polymorphisms (SNPs) generated from 3,331 sites in the core genome, excluding
252 repeated regions and transposable elements. The tree resolved seven distinct clades.
253 We named these clades A-G. Clade A and B were comprised of Dublin and Enteritidis
254 serovars, respectively. Typhimurium clustered with Hull and Stanleyville in clade C.
255 Clade D included serovars Hofit and Rubislaw while clade E was comprised only of
256 Virchow isolates. All the Typhi isolates formed a distinct clade (clade F) and the
257 remaining serovars formed a separate clade, clade G (Figure 5).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



258

259

260

261 **Figure 5.** Maximum likelihood phylogenetic tree of 164 *Salmonella* genomes isolated
262 from patients in rural Gambia between 2008 and 2016. Seven distinct clades were
263 resolved from the tree and denoted by different colours (see legend). Metadata is
264 shown alongside the phylogenetic and includes host sex and disease status. The
265 serovars and most prevalent sequence types are annotated on the tree and denoted
266 using different colours. The tree was rooted on the *Salmonella* Paratyphi C isolate.

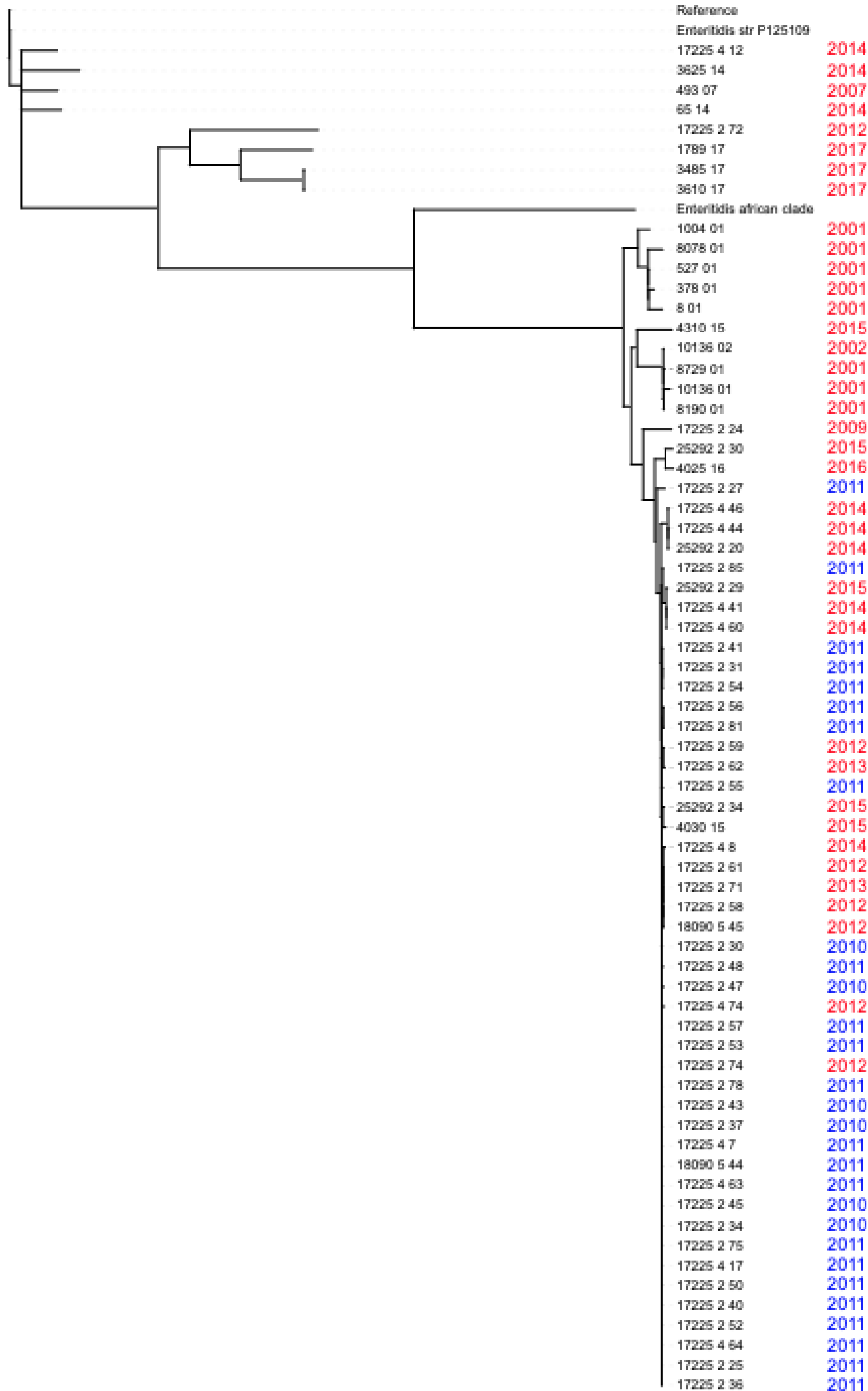
267

268 **Genomic analysis of Enteritidis isolates**

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

269 To understand the reason for the high proportion of Enteritidis between 2010 and 2011
270 we used phylogenetic analysis to compare the 2010 and 2011 Enteritidis genomes in
271 our dataset with Enteritidis genomes collected in The Gambia before and after 2010.
272 This analysis indicated a potential outbreak (Figure 6) with more than 70% (21/29) of
273 the Enteritidis isolates collected during the surveillance in 2010 and 2011 clustered
274 closely on the tree with short branch lengths, suggesting closely related strains
275 circulating during this time frame.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



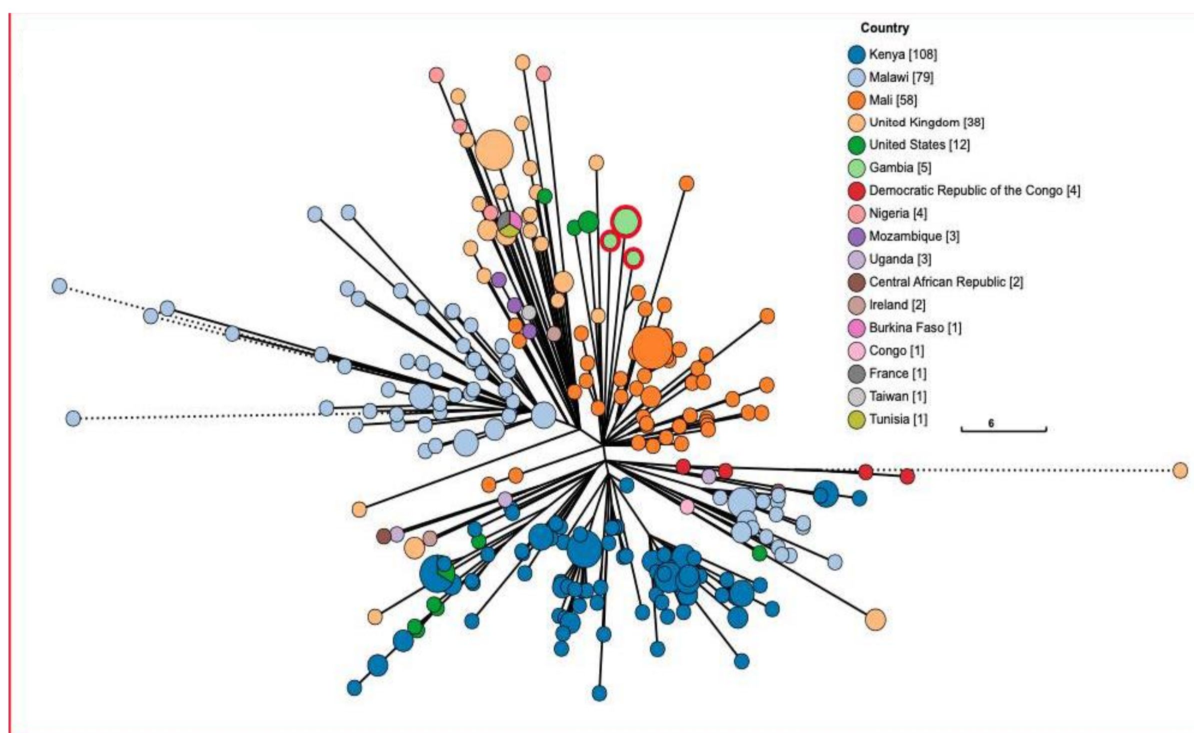
276

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

277 **Figure 6.** Phylogenetic tree of 49 *Salmonella* Enteritidis isolates collected during the
278 surveillance period and 16 other isolates collected from The Gambia (both within the
279 surveillance area and outside) at different time points. Isolates collected in the present
280 study between 2010 and 2011 are colored blue and those collected before or after the
281 surveillance period are coloured red. The tree is rooted on the *Salmonella*
282 Typhimurium LT2 reference genome.
283

284 **Genomic analysis of S Typhimurium ST313 isolates**

285 We found that five isolates had the ST313 genotype, which has been implicated as
286 the causative agent of invasive *Salmonella* disease in Kenya and Malawi. For this
287 reason, we used phylogenetic analysis to compare the ST313 isolates in our study
288 with other global strains in Enterobase[33]. We found that the isolates circulating in The
289 Gambia are of the lineage 1 type and different from the type circulating in Kenya and
290 Malawi which are of the lineage 2 (Figure 7).
291



292
293

294 **Figure 7.** Phylogenetic tree of five *Salmonella* Typhimurium ST313 isolates from our
295 study and all ST313 isolates from other countries (as indicated in the legend).
296 Isolates from our study are highlighted in green with a red ring and are clustered away
297 from the Kenyan (dark blue) and Malawian (sky blue) ST313 strains.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

298

299

300 **Distribution of virulence, resistance and plasmid genes**

301 A total of 124 virulence genes within and outside the *Salmonella* pathogenicity islands
302 (SPI) were detected. The distribution of virulence genes detected and how they
303 grouped based on the loci present can be found in Supplementary Table 1. Some
304 virulence genes were conserved in the *Samonella* isolates evaluated while others
305 were only present in some serovars. For example, SPI-7 which encodes *vex* and *tvi*
306 genes was found in Typhi serovars only while SPI-11, which encodes the *CdtB* gene
307 was found in several serovars within the atypical group.

308

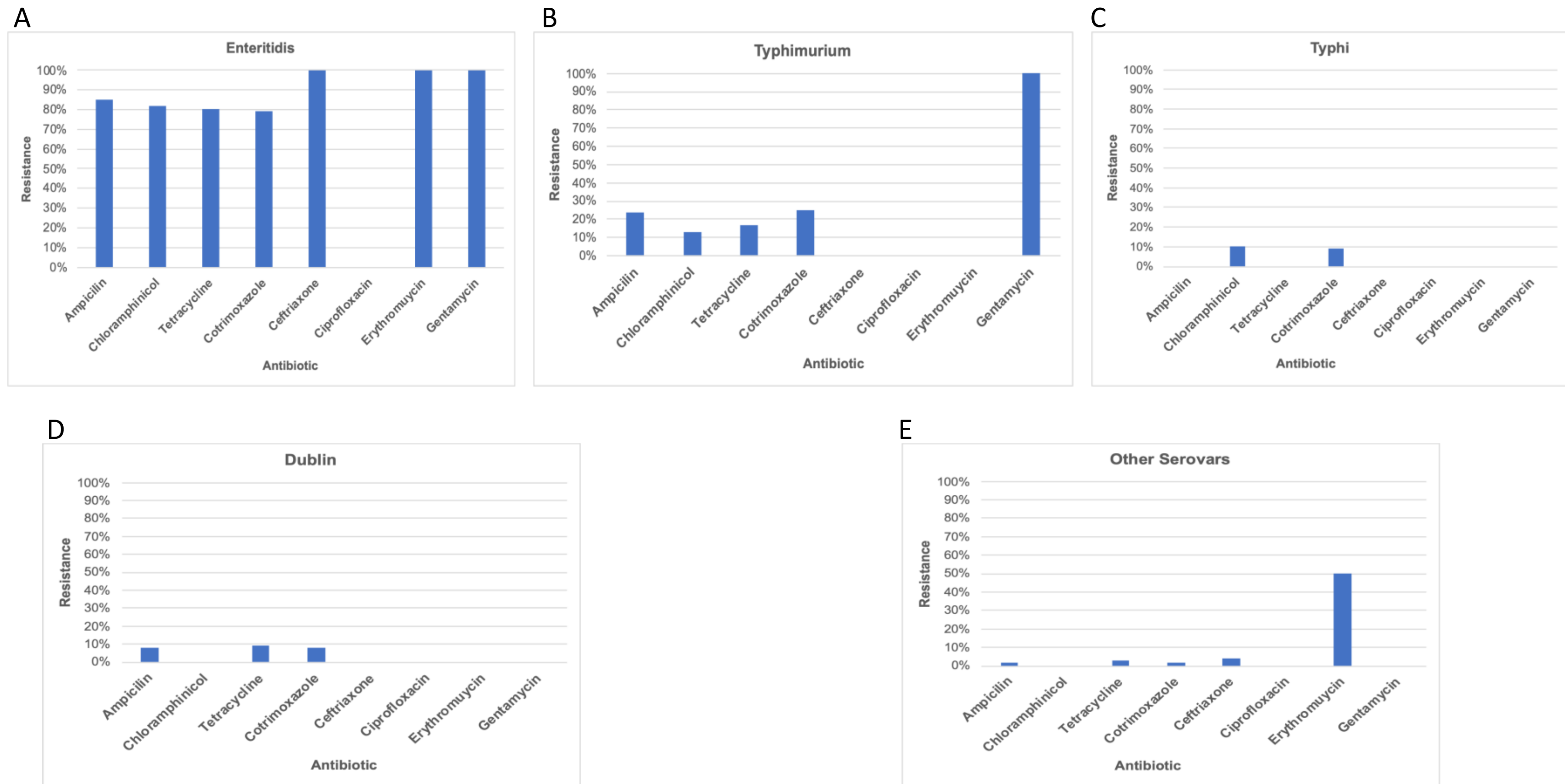
309 Some genes found outside the SPI, including fimbriae and adhesion encoding genes
310 as well as the type 1 fimbriae, were conserved in all isolates. Most of the genes that
311 were variable in their distribution were found residing outside the pathogenicity
312 islands. These genes included Gifsy-1 found in Typhimurium and Paratyphi C
313 serovars only, and Gifsy-2 effector genes found only in Bron, Dublin, Enteritidis,
314 Paratyphi C and Typhimurium isolates. Interestingly, we found 42% (31/68) of
315 serovars in the atypical group had the virulence gene *cdtB* and that this gene was
316 present in all our Typhi isolates.

317

318 Genomic analysis indicated more antimicrobial resistance genes in Enteritidis than
319 any other serovar. Analysis of phenotypic data showed a similar pattern where 80%
320 to 100% (n=40) of Enteriditis isolates were resistant to all the antimicrobials tested
321 except ciprofloxacin. 100% (n=40) sensitivity was observed in all Enteriditis isolates
322 tested against ciprofloxacin (Figure 8A). Some of the resistance genes present in
323 Enteritidis were also found in Typhimurium ST313 isolates, but were present in only

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

324 few of the atypical serovars. All Typhimurim isolates (n=16) tested were resistant to
325 gentamycin. We found only few plasmid genes in our dataset. This was more
326 pronounced in some serovars such as Dublin, Enteritidis and Typhimurium. In fact,
327 none of the Typhi strains had a plasmid gene and only a few of the atypical serovars
328 had one or two plasmids. We found that some plasmids were specific to particular
329 serovars. For example IncX1 was found only in Dublin isolates. IncFIIB was common
330 in Typhimurium isolates while Incl1 and IncQ were found in all Enteritidis isolates (see
331 Table 5 for full summary).
332
333



334
335 **Figure 8**. Antibiotic resistance patterns in invasive *Salmonella* serovars isolated in rural Gambia between 2008 and 2016.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

336 **Discussion**

337 In The Gambia, NTS is an important cause of invasive bacterial infections especially
338 in children[14,34–37]. Using population-based epidemiological data and whole genome
339 sequencing, we found an increase in the proportion of atypical NTS serovars causing
340 invasive disease in rural Gambia between 2008 and 2016. We also observed changes
341 in the incidence of disease over time. We identified sets of virulence genes in atypical
342 serovar isolates that may be responsible for the increased prevalence of these
343 serovars.

344 Few studies have described the distribution of non-typhoidal *Salmonella* serovars in
345 The Gambia[14,37]. Between 2000 and 2004, Ikumapayi *et al.*, reported Enteritidis as the
346 major cause of invasive disease in rural Gambia while Typhimurium and other
347 serovars accounted for only few cases[14]. Interestingly, the present study showed a
348 significant reduction in the proportion of invasive *Salmonella* disease caused by
349 Enteritidis. To identify serovars, Ikumapayi *et al.*, used conventional antisera
350 agglutination methods while polymerase chain reaction (PCR) methods were used for
351 MLST typing[14]. This could underestimate the proportion of some serovars as antisera-
352 based methods are limited in their ability to distinguish between closely-related and
353 polyphyletic serovars[38]. By exploiting the advantages of whole genome sequencing,
354 we identified 31 different serovars and thus a greater diversity of *Salmonella* serovars
355 causing invasive disease. Between 2005-2015, Kwambana-Adams *et al.*, reported
356 Typhimurium to be the predominant invasive serovar in the coastal parts of The
357 Gambia[37], with 25% of isolates being serovars other than Typhi, Typhimurium, or
358 Enteritidis. In comparison, our data show temporal and/or regional differences in the
359 prevalence of *Salmonella* which could be attributed to many factors including host and
360 pathogen genetic characteristics.

361

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

362 Globally, Typhimurium and Enteritidis are the two major serovars associated with
363 invasive *Salmonella* disease[39,40]. However, this trend was different in rural Gambia
364 where atypical serovars including Dublin, Virchow and Poona are increasing in
365 prevalence. Studies have shown that genetic factors and immune status predispose
366 individuals to invasive *Salmonella* disease[4]. For example, malnutrition and HIV have
367 been associated with increased susceptibility to invasive *Salmonella* disease[41].
368 However, in The Gambia, the prevalence of malnutrition and HIV has not changed
369 over the years suggesting that the increased incidence of invasive *Salmonella* disease
370 may be attributable to other environmental factors or the genetic characteristics of the
371 pathogen. We observed an increase in atypical serovars with the majority of cases
372 occuring between 2012 and 2014. However, genomic analysis revealed various
373 virulence factors implicated in invasion, proliferation and or translocation by Type III
374 secretion systems in all Dublin isolates. Between 2012 and 2014, Dublin was the most
375 common serovar isolated within the atypical group. Studies have reported that Dublin
376 is associated with more severe disease and more frequently the cause of invasive
377 disease than other types of non-Typhi *Salmonella*[42,43]. The present study reported two
378 deaths associated with the Dublin serovar ranking second in mortality after Enteritidis.
379 Moreover, this study identified the cytolethal distending toxin gene (*CdtB*) in the
380 majority of atypical serovars (Clade G). This gene encodes cytolethal distending toxin
381 (CDT) which activates host DNA damage and thus leads to G_2/M phase arrest[12].
382 Analysis of all *Salmonella* genome assemblies in RefSeq (accessed 26-03-2020)
383 showed overall prevalence of *cdtB* to be 35% (3832/10882), and when Typhi is
384 excluded, this falls to 14% (1628/8678). This shows an uncommonly high level of *CdtB*
385 in our atypical serovars. Experimental studies show that populations of HeLa cells
386 infected with *cytolethal distending toxin* (CDT)-positive NTS serovars have a

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

387 significantly larger proportion of cells with DNA damage response protein (53BP1) and
388 γH2AX foci than CDT negative serotypes[12]. More importantly, *in vivo* analysis showed
389 increased colonization of the host by CDT-producing pathogens that was associated
390 with tumorigenesis and neoplastic lesions that led to chronic infections[12]. Thus, we
391 speculate that increased prevalence of *cdtB* genes in our study may provide these
392 serovars with a fitness advantage over Enteritidis and Typhimurium, potentially
393 contributing to the shift we observed.
394 In contrast, we observed a high proportion of Enteritidis between 2010 and 2011. This
395 period coincided with heavy rains resulted in severe flooding in the Upper River
396 Region. Subsequent high rates of malaria infection may have influenced the
397 population's susceptibility to iNTS disease. Phylogenetic analysis of the Enteritidis
398 isolates suggests a potential outbreak. All Enteritidis isolates recovered during this
399 period were isolated within the Basse area with similar virulence and antimicrobial
400 resistance patterns. Outbreaks of *Salmonella* Enteritidis as a result of consumption of
401 contaminated food or animal products have been reported elsewhere[44]. Although this
402 theory could be true, a study in Mali highlighted that, in contrast to *Salmonella*
403 Typhimurium, iNTS disease caused by *Salmonella* Enteritidis started to increase from
404 2008 with the highest peak seen in 2010 and 2011[16]. The finding in Mali corresponds
405 with our observed increase in Enteritidis in 2010 and 2011 suggesting the potential
406 combination of a regional increase in Enteritidis exacerbated by the impact of the flood
407 in our setting.
408
409 Antibiotic resistance in some *Salmonella* serotypes has been reported in many parts
410 of Africa including The Gambia[14,45]. Our Enteritidis serovars had more resistance
411 genes than other serovars. Similar findings were also reported in previous studies

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

412 done in The Gambia which showed high percentages of multidrug resistance among
413 *Salmonella* Enteritidis isolates[14]. However, five of our Typhimurium isolates of the
414 ST313 genotype had resistance genes similar to those found in Enteritidis. In Kenya
415 and Malawi, a distinct genotype of Typhimurium ST313 was reported to have a
416 multidrug resistance gene located on a virulence plasmid[11]. Genomic analysis of all
417 ST313 isolates in our study and those found in Enterobase suggest that this unique
418 Typimurium ST313 is restricted to eastern Africa. Nonetheless, continued monitoring
419 of these genotypes in other parts of Africa is vital. It is, however, reassuring that many
420 of the atypical serovars did not acquire resistance genes, although continued
421 monitoring is essential as antiomicrobial resistance (AMR) is increasing, and has a
422 high global health burden. We found only one Dublin isolate with resistance genes.

423 **Conclusion**

424 Overall, this study has shown a wide distribution of invasive *Salmonella* serovars
425 circulating in The Gambia. More importantly, an increase over time in atypical serovars
426 with high case fatality rates was also documented. The study highlighted the potential
427 effect of some virulence genes in contributing to the shift we observed. However,
428 experimental and functional studies could shed more light on the role of such virulence
429 genes and the evolutionary pressures on these serovars. The shift in serovar
430 prevalence could have implications for vaccine development and thus represent a
431 public health concern. Therefore, investigations should be made to identify potential
432 changes in the distribution of iNTS serovars elsewhere in Africa and the prevalence of
433 these virulence elements.

434

435 **Authors and contributors**

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

436 AK and GM conceived the research idea and AK wrote the first draft of the manuscript.
437 AK, AP and NFA did the bioinformatics analysis. UNI, RS and JM did the microbiology.
438 GD and team did the sequencing. AKS supervised AK and reviewed the manuscript.
439 All authors have read and approved the final version of the manuscript.
440

441 **Conflicts of interest**

442 The author(s) declare that there are no conflicts of interest
443

444 **Funding information**

445 The surveillance study was sponsored by GAVI's Pneumococcal vaccines
446 Accelerated Development and Introduction Plan (PneumoADIP), the Bill & Melinda
447 Gates Foundation, and the UK Medical Research Council
448

449 AJP and NFA gratefully acknowledge the support of the Biotechnology and Biological
450 Sciences Research Council (BBSRC); this research was funded by the BBSRC
451 Institute Strategic Programme Grant Microbes in the Food Chain BB/R012504/1 and
452 its constituent project(s) BBS/E/F/000PR10348 and BBS/E/F/000PR10352. AK was
453 partially supported by a BBSRC Impact Acceleration Account award.
454

455 **Ethical approval**

456 The parent project consented participants before enrolling them in the study.
457 Therefore, this study does not require any ethical approval.
458

459 **Data availability**

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

460 The raw sequencing data is publicly available from the European Nucleotide Archive
461 under BioProject PRJEB39996.
462

463 **Acknowledgements**
464 We thank the following people: Gordon Dougan and team for sequencing the isolates.
465 Thanh Le Viet for bioinformatics support, Abdul Khalid Muhammad and Nuredin
466 Mohammed for statistics support and Jahangir Hussain for providing GEMS domestic
467 data. AK also wishes to acknowledge the management of MRCG at LSHTM for partly
468 funding his internship at Quadram while working on this project.
469

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

470 **Reference:**

471 1. Kingsley, R. A. & Bäumler, A. J. Host adaptation and the emergence of infectious
472 disease: The Salmonella paradigm. *Mol. Microbiol.* **36**, 1006–1014 (2000).

473 2. Popoff, M. Y., Bockemühl, J. & Gheesling, L. L. Supplement 2002 (no. 46) to the
474 Kauffmann-White scheme. *Res. Microbiol.* **155**, 568–570 (2004).

475 3. Kirk, M. D. *et al.* World Health Organization Estimates of the Global and Regional
476 Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A
477 Data Synthesis. *PLoS Med.* **12**, 1–21 (2015).

478 4. Gordon, M. a. Europe PMC Funders Group Invasive Non-typhoidal Salmonella
479 Disease – epidemiology , pathogenesis and diagnosis. *Curr. Opin. Infect. Dis.* **24**, 484–
480 489 (2012).

481 5. Uche, I. V., MacLennan, C. A. & Saul, A. A Systematic Review of the Incidence, Risk
482 Factors and Case Fatality Rates of Invasive Nontyphoidal Salmonella (iNTS) Disease
483 in Africa (1966 to 2014). *PLoS Negl. Trop. Dis.* **11**, 1–28 (2017).

484 6. Gordon, M. A. *et al.* Epidemics of Invasive Salmonella enterica Serovar Enteritidis
485 and S. enterica Serovar Typhimurium Infection Associated with Multidrug Resistance
486 among Adults and Children in Malawi . *Clin. Infect. Dis.* **46**, 963–969 (2008).

487 7. Berkley, J. A. *et al.* Bacteremia among children admitted to a rural hospital in Kenya.
488 *N. Engl. J. Med.* **352**, 39–47 (2005).

489 8. Kiehlbauch, J. A. *et al.* Pathogens and Predictors of Fatal Septicemia Associated with
490 Human Immunodeficiency Virus Infection in Ivory Coast, West Africa. *J. Infect. Dis.*
491 **168**, 564–570 (1993).

492 9. Gordon, M. A. Salmonella infections in immunocompromised adults. *J. Infect.* **56**,
493 413–422 (2008).

494 10. Darboe, S. *et al.* Community-acquired Invasive Bacterial Disease in Urban Gambia,
495 2005-2015: A Hospital-based Surveillance. *Clin. Infect. Dis.* **69**, S105–S113 (2019).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

496 11. Kingsley, R. A. *et al.* Epidemic multiple drug resistant Salmonella Typhimurium
497 causing invasive disease in sub-Saharan Africa have a distinct genotype. *Genome Res.*
498 **19**, 2279–2287 (2009).
499 12. Miller, R. A. & Wiedmann, M. The cytolethal distending toxin produced by
500 nontyphoidal Salmonella serotypes javiana, montevideo, oranienburg, and mississippi
501 induces DNA damage in a manner similar to that of serotype Typhi. *MBio* **7**, 1–11
502 (2016).
503 13. Ge, Z. *et al.* Cytolethal distending toxin is essential for Helicobacter hepaticus
504 colonization in outbred Swiss Webster mice. *Infect. Immun.* **73**, 3559–3567 (2005).
505 14. Ikumapayi, U. N. *et al.* Molecular epidemiology of community-acquired invasive non-
506 typhoidal Salmonella among children aged 2-29 months in rural Gambia and discovery
507 of a new serovar, Salmonella enterica Dingiri. *J. Med. Microbiol.* **56**, 1479–1484
508 (2007).
509 15. Le Bacq, F., Louwagie, B. & Verhaegen, J. Salmonella typhimurium and Salmonella
510 enteritidis: Changing epidemiology from 1973 until 1992. *Eur. J. Epidemiol.* **10**, 367–
511 371 (1994).
512 16. Bornstein, K. *et al.* Genetic changes associated with the temporal shift in invasive non-
513 typhoidal salmonella serovars in Bamako Mali. *PLoS Negl. Trop. Dis.* **13**, 1–15
514 (2019).
515 17. Mackenzie, G. A. *et al.* Monitoring the introduction of pneumococcal conjugate
516 vaccines into west africa: Design and implementation of a population-based
517 surveillance system. *PLoS Med.* **9**, (2012).
518 18. Dolinsky, A. L. *M100 Performance Standards for Antimicrobial Susceptibility Testing*.
519 *Journal of Services Marketing* vol. 8 (2017).
520 19. Kotloff, K. L. *et al.* The incidence, aetiology, and adverse clinical consequences of less

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

521 severe diarrhoeal episodes among infants and children residing in low-income and
522 middle-income countries: a 12-month case-control study as a follow-on to the Global
523 Enteric Multicenter St. *Lancet Glob. Heal.* **7**, e568–e584 (2019).
524 20. Cock, P. J. A., Fields, C. J., Goto, N., Heuer, M. L. & Rice, P. M. The Sanger FASTQ
525 file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants.
526 *Nucleic Acids Res.* **38**, 1767–1771 (2009).
527 21. Bankevich, A. *et al.* SPAdes: A new genome assembly algorithm and its applications
528 to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
529 22. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: Quality assessment tool
530 for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).
531 23. Page, A. J. *et al.* Robust high-throughput prokaryote de novo assembly and
532 improvement pipeline for Illumina data. *Microb. genomics* **2**, e000083 (2016).
533 24. Jia, B. *et al.* CARD 2017: Expansion and model-centric curation of the comprehensive
534 antibiotic resistance database. *Nucleic Acids Res.* **45**, D566–D573 (2017).
535 25. Zankari, E. *et al.* Identification of acquired antimicrobial resistance genes. *J.*
536 *Antimicrob. Chemother.* **67**, 2640–2644 (2012).
537 26. Carattoli, A. *et al.* In Silico detection and typing of plasmids using plasmidfinder and
538 plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* **58**, 3895–3903
539 (2014).
540 27. Liu, B., Zheng, D., Jin, Q., Chen, L. & Yang, J. VFDB 2019: A comparative
541 pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* **47**,
542 D687–D692 (2019).
543 28. Jolley, K. A. & Maiden, M. C. J. BIGSdb: Scalable analysis of bacterial genome
544 variation at the population level. *BMC Bioinformatics* **11**, (2010).
545 29. Nguyen, L. T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

546 effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol.*
547 *Biol. Evol.* **32**, 268–274 (2015).
548 30. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v4: Recent updates and new
549 developments. *Nucleic Acids Res.* **47**, 256–259 (2019).
550 31. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–
551 2069 (2014).
552 32. Page, A. J. *et al.* Roary: Rapid large-scale prokaryote pan genome analysis.
553 *Bioinformatics* **31**, 3691–3693 (2015).
554 33. Alikhan, N. F., Zhou, Z., Sergeant, M. J. & Achtman, M. A genomic overview of the
555 population structure of Salmonella. *PLoS Genet.* **14**, 1–13 (2018).
556 34. Enwere, G. *et al.* Epidemiologic and clinical characteristics of community-acquired
557 invasive bacterial infections in children aged 2-29 months in The Gambia. *Pediatr.*
558 *Infect. Dis. J.* **25**, 700–705 (2006).
559 35. Okomo, U. A. *et al.* Bacterial Isolates and Antibiotic Sensitivity among Gambian
560 Children with Severe Acute Malnutrition. *Int. J. Pediatr.* **2011**, 1–8 (2011).
561 36. Hill, P. C. *et al.* Bacteraemia in patients admitted to an urban hospital in West Africa.
562 *BMC Infect. Dis.* **7**, 1–8 (2007).
563 37. Kwambana-Adams, B. *et al.* Salmonella infections in The Gambia, 2005-2015. *Clin.*
564 *Infect. Dis.* **61**, S354–S362 (2015).
565 38. Shi, C., Singh, P., Ranieri, M. L., Wiedmann, M. & Moreno Switt, A. I. Molecular
566 methods for serovar determination of Salmonella. *Crit. Rev. Microbiol.* **41**, 309–325
567 (2015).
568 39. Galanis, E. *et al.* Web-based surveillance and global Salmonella distribution, 2000-
569 2002. *Emerg. Infect. Dis.* **12**, 381–388 (2006).
570 40. Stanaway, J. D. *et al.* The global burden of non-typhoidal salmonella invasive disease:

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

571 a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Infect. Dis.*
572 **19**, 1312–1324 (2019).
573 41. Bronzan, R. N. *et al.* Bacteremia in Malawian children with severe malaria:
574 Prevalence, etiology, HIV coinfection, and outcome. *J. Infect. Dis.* **195**, 895–904
575 (2007).
576 42. Suez, J. *et al.* Virulence Gene Profiling and Pathogenicity Characterization of Non-
577 Typhoidal Salmonella Accounted for Invasive Disease in Humans. *PLoS One* **8**,
578 (2013).
579 43. Feldman, M. J. B. and R. A. Salmonella Bacteremia : Reports to the Centers for
580 Disease Control , 1968-1979 Author (s): Martin J . Blaser and Roger A . Feldman
581 Published by : Oxford University Press Stable URL :
582 http://www.jstor.org/stable/30113298. *J. Infect. Dis.* **143**, 743–746 (1981).
583 44. Inns, T. *et al.* Prospective use of whole genome sequencing (WGS) detected a multi-
584 country outbreak of Salmonella Enteritidis. *Epidemiol. Infect.* **145**, 289–298 (2017).
585 45. Feasey, N. A. *et al.* Distinct Salmonella Enteritidis lineages associated with
586 enterocolitis in high-income settings and invasive disease in low-income settings. *Nat.*
587 *Genet.* **48**, 1211–1217 (2016).
588

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1. Numbers of patients enrolled, blood cultures collected, and *Salmonella* isolates detected each year

Year	Total enrolled	Total blood cultures taken	Enteritidis	Typhimurium	Typhi	Atypical	Paratyphi	Total
2008	1212	1047	0	0	3	3	1	7
2009	2099	1898	1	2	1	4	0	8
2010	1869	1605	6	3	1	2	0	12
2011	2688	2385	23	2	4	4	0	33
2012	2899	2592	7	3	7	20	0	37
2013	2580	2200	2	3	5	17	0	27
2014	2707	2536	7	3	3	9	0	22
2015	3742	3566	3	2	2	5	0	12
2016	2509	2370	0	1	1	4	0	6
Total	**22305**	**20199**	**49**	**19**	**27**	**68**	**1**	**164**

Table 2. Summary of baseline patient characteristics.

Variable	Characteristic	N (%)
Sex	Male	84 (53.5)
	Female	73 (46.5)
Diagnosis	Pneumonia	93 (59.2)
	Meningitis	11 (7.0)
	Septicaemia	46 (29.3)
	Other focal sepsis	6 (3.8)
	Other	1 (0.6)
Disease Outcome	Dead	16 (10.2)
	Discharged and/or recovered	111 (70.7)
	Not admitted	22 (14.0)
	Absconded	1 (0.6)
	Transfered	6 (3.8)
	Missing	1 (0.6)
Age range	0-5 yrs	144 (91.7)
	6-15 yrs	7 (4.5)
	>15 yrs	6 (3.8)
Nutritional status	Acute malnutrition	51 (32.5)
	Moderate acute malnutrition	32 (20.4)
	Well nourished	64 (40.8)
	Missing	10 (6.4)
Reside with the surveillance area	Yes	136(86.6)

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

	No	21(13.4)
Sample type	Blood	157 (95.7)
	Cerebospinal fluid	6 (3.7)
	Lung Aspirate	1 (0.6)
Infection rate by serotype	Enteritidis	47 (29.9)
	Typhimurium	18 (11.5)
	Typhi	27 (17.2)
	Paratyphi C	1 (0.6)
	Atypical	64 (40.8)

Table 3. Summary of resistance and plasmid genes in each serovar.

Clade	Serovar	Gene Name	Total (%)	Plasmid genes	Total (%)
A	Dublin	fosA7_1	1/14 (7.1)	IncFII(S)_1	14/14 (100)
				IncI1_1_Alpha	1/14 (7.1)
				IncX1_1	14/14 (100)
B	Enteritidis	aph(3'')-Ib_5	45/49 (91.8)	ColpVC_1	1/49 (2.1)
		aph(6)-Id_1	45/49 (91.8)	IncFIB(S)_1	2/49 (4.1)
		blaTEM-1B_1	49/49 (100)	IncFII(S)_1	2/49 (4.1)
		catA1_1	46/49 (93.8)	IncI1_1_Alpha	47/49 (95.9)
		dfrA7_5	46/49 (93.8)	IncQ1_1	45/49 (91.8)
		sul1_5	46/49 (93.8)	rep21_9_rep(pKH12)	2/49 (4.1)
		sul2_6	45/49 (91.8)		
		tet(B)_2	46/49 (93.8)		
C	Typhimurium	aph(3'')-Ib_5	3/19 (15.8)	IncFIB(S)_1	18/19 (94.7)
		aph(6)-Id_1	3/19 (15.8)	IncFII(S)_1	18/19 (94.7)
		blaTEM-1B_1	3/19 (15.8)	IncQ1_1	1/19 (5.3)
		catA1_1	3/19 (15.8)		
		dfrA7_5	3/19 (15.8)		
		sul1_5	3/19 (15.8)		
		sul1_3	3/19 (15.8)		
		fosA7_1	1/19 (5.3)		
	Stanleyvelle	fosA7_1	2/2 (100)		
D	Hofit	fosA7_1	1/1 (100)	IncFIB(S)_1	1/1 (100)
				IncFII(S)_1	1/1 (100)
E	Virchow	no gene		pSL483_1	1/7 (14.3)
F	Typhi	catA1_1	2/27 (7.4)	no plasmid	
		dfrA7_5	2/27 (7.4)		

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.18.431831; this version posted February 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

		sul1_5	2/27 (7.4)		
G	Others:				
	Mountpleasant	fosA7_1	1/41 (100)	IncFII(S)_1	2/41 (4.8)
	Senftenberg	fosA7_1	1/41 (100)	IncFII(pCoo)_1_pCoo	1/41 (2.4)
	Grumpensis	fosA7_1	1/41 (100)		
	Paratyphi C	fosA7_1	1/1 (100)	IncFIB(S)_1	1/1 (100)
				IncFII(S)_1	1/1 (100)