Genomic surveillance of *Salmonella enterica* serotype Minnesota strains from Brazilian poultry products imported into South Africa

Vishnu Raghuram[1,2,3,4], Thendo Mafuna[5], Vignesh Ramnath[1,2,3,4], Hadrien Gourlé[1,2,3,4], Josefin Blom[1,2,3,4], Kudakwashe Magwedere[6], Laura M. Carroll[1,2,3,4], Itumeleng Matle[7,#]

[1] Department of Clinical Microbiology, SciLifeLab, Umeå University, Umeå, Sweden

[2] Laboratory for Molecular Infection Medicine Sweden (MIMS), Umeå University, Umeå, Sweden

[3] Umeå Centre for Microbial Research (UCMR), Umeå University, Umeå, Sweden

[4] Integrated Science Lab (IceLab), Umeå University, Umeå, Sweden

[5] Department of Biochemistry, University of Johannesburg, Auckland Park, South Africa

[6] Unaffiliated Global Health Security Research Scholar, Pretoria, South Africa

[7] Department of Agriculture and Animal Health, University of South Africa, Science Campus, Florida, South Africa

[#]Corresponding author: Itumeleng Matle, ematlei@unisa.ac.za

Keywords: *Salmonella* Minnesota, CTX-M-8, beta-lactamase, whole-genome sequencing, antimicrobial resistance, poultry

Repositories: Whole-genome sequencing (WGS) data generated in this study is available under NCBI BioProject accession PRJNA1230142, with BioSample and Sequence Read Archive (SRA) accessions for each genome available in Supplementary Table S1.

Abstract

Salmonella enterica serotype Minnesota (*S.* Minnesota) has recently emerged as a predominant serotype in Brazilian poultry farming operations. Genomic surveillance efforts concentrated primarily in Europe have been used to evaluate food safety risks associated with *S.* Minnesota in imported Brazilian poultry/poultry products. However, the burden imposed by *S.* Minnesota on consumers in sub-Saharan Africa is not understood. Here, we used whole-genome sequencing (WGS) to characterize 36 *S.* Minnesota strains from Brazilian raw poultry imported into South Africa, specifically: (i) 11 strains isolated at port-of-entry, and (ii) 25 strains from imported poultry in South African supermarkets. While all 36 strains belonged to the same sequence type (ST548), multiple ST548 lineages were present among Brazilian poultry products. Comparison of the 36 strains sequenced here to all publicly available, high-quality ST548 genomes (*n* = 228, from Enterobase) identified several closely related public genomes (<30 core SNPs), including strains isolated previously from Brazilian poultry imported into the United Kingdom. Notably, a cluster consisting of 14 highly similar genomes sequenced here (0 core SNPs) uniquely possessed $bla_{\text{CTX-M-8}}$. A search of plasmids in public databases, alongside antimicrobial resistance (AMR) genes from >1.9 million bacterial genomes, revealed that this cluster harbored $bla_{\text{CTX-M-8}}$ on an IncI1 plasmid, which we hypothesize was acquired recently, within Brazil, from *Escherichia coli*. Overall, our study provides insight into the intercontinental dissemination of *S.* Minnesota and its associated AMR determinants via the global poultry trade.

46 **Impact Statement**

47 Brazil's status as a world-leading poultry exporter has disseminated *S.* Minnesota
48 internationally. Several countries (primarily in Europe) have thus used WGS to
49 characterize *S.* Minnesota strains from imported Brazilian poultry. Many African countries
50 also import significant quantities of Brazilian poultry, and Brazil is the world's largest
51 exporter of poultry into South Africa. However, very little is known about *S.* Minnesota in
52 South Africa, let alone sub-Saharan Africa. Here, we detect multiple *S.* Minnesota ST548
53 lineages at two control points in South Africa (port-of-entry and supermarkets), including
54 emerging AMR lineages. Notably, we found that all *S.* Minnesota genomes from South
55 Africa were closely related to poultry/food-associated *S.* Minnesota genomes from the
56 United Kingdom (UK) or Brazil. Recent WGS-based studies from the UK have posited
57 that *S.* Minnesota from Brazilian poultry is unlikely to cause illness, should it reach UK
58 consumers. While no links to clinical cases in South Africa were observed here, this could
59 be due to data gaps, as the vast majority of *Salmonella* WGS efforts are concentrated in
60 Europe and North America. Our study highlights the important role that genomic
61 surveillance plays in mitigating food safety risks associated with the global agro-food
62 trade and showcases the importance of local pathogen surveillance initiatives.

63 **Data Summary**

64 WGS data generated in this study is available under NCBI BioProject accession
65 PRJNA1230142. NCBI BioSample and Sequence Read Archive (SRA) accessions for
66 newly sequenced genomes, as well as for the publicly available genomes used in this
67 study, are available in the Supplementary Material. All code for running software is
68 described in the article, and additional data analysis code is available via GitHub
69 (https://github.com/VishnuRaghuram94/SEPI). Supplementary Tables S1-S3 are
70 available alongside the article. All larger supplementary datasets (Dataset S1-S17) and
71 intermediate files are available via doi: 10.5281/zenodo.15063662.

72

73 **INTRODUCTION**

74 Foodborne zoonotic pathogen *Salmonella enterica* imposes a massive burden on global
75 public health [1, 2]. Each year, *Salmonella* is estimated to cause 93.8 million cases of
76 gastroenteritis globally, the vast majority of which (>85%) are attributed to the
77 consumption of contaminated food [3]. Contaminated poultry in particular is widely
78 regarded to be a major source of *Salmonella* illnesses [1, 4–6]. In the United States (US),
79 for example, chicken and turkey products have been estimated to account for nearly a
80 quarter of foodborne *Salmonella* illness cases [7], with meat and poultry products serving
81 as the most common sources of fatal foodborne illness cases (caused primarily by *Listeria*
82 or *Salmonella*) [8, 9].

83 While poultry can serve as a reservoir for a range of *Salmonella* serotypes [10–12],
84 *Salmonella enterica* subspecies *enterica* serotype Minnesota (*S.* Minnesota) has recently
85 emerged as a serotype of concern in Brazilian poultry farming operations [4, 13, 14]. First
86 reportedly isolated from a turkey in Minnesota (US) in 1936 [15], *S.* Minnesota is a
87 relatively rare serotype globally [4, 13]. However, in Brazil, *S.* Minnesota has recently
88 emerged as a predominant serotype in poultry [4, 14, 16, 17]. It has been hypothesized
89 that a combination of factors contributed to *S.* Minnesota's rapid rise in Brazil, including
90 the 2003 introduction of a vaccine against *S.* Enteritidis (the predominant serotype in
91 Brazilian poultry at the time), antimicrobial use, and antimicrobial resistance (AMR) [4,
92 13, 18, 19].

93 From a genomic surveillance perspective, a great deal of effort has gone into monitoring
94 the spread of S. Minnesota and its associated AMR determinants in Brazil [4, 14, 17, 20].

95 However, due to Brazil's status as a leading global poultry exporter, S. Minnesota can be
96 disseminated to other countries via imported Brazilian poultry/poultry products [4, 13]. As
97 such, several other countries have employed whole-genome sequencing (WGS) to
98 characterize S. Minnesota strains from Brazilian poultry imports [4, 13, 21, 22]. Notably,
99 these efforts have been largely confined to Europe (e.g., the United Kingdom [UK],
100 Portugal) [4, 21], and only recently, Saudi Arabia [13]. Very little is known about S.
101 Minnesota in sub-Saharan Africa, and WGS-based studies of S. Minnesota from imported
102 poultry have not been conducted in Africa, despite the fact that many African countries
103 also import significant quantities of Brazilian poultry/poultry products [23].

104 Port-of-entry pathogen surveillance efforts in South Africa have highlighted the potential
105 food safety risks associated with imported poultry products [24]. Given that (i) S.
106 Minnesota has become relatively prevalent in imported poultry products tested at South
107 African port-of-entry [24], and (ii) Brazil has been the largest exporter of poultry meat to
108 South Africa since 2001 [25], it is apparent that poultry products imported from Brazil may
109 serve as a vehicle by which S. Minnesota can enter South Africa's domestic food system.
110 However, the extent to which this has occurred has not been studied. Here, we used
111 WGS to characterize 36 S. Minnesota strains isolated from raw poultry imported into
112 South Africa from Brazil. By comparing our novel genomes to hundreds of relevant high-
113 quality, publicly available S. Minnesota genomes from around the world, we provide
114 insight into the dissemination of S. Minnesota and its associated AMR determinants in
115 South Africa.

116 **METHODS**

117 **Study design and source of isolates**

118 This retrospective study was conducted using laboratory-confirmed *Salmonella* isolates
119 stored at the Agricultural Research Council – Onderstepoort Veterinary Research (ARC-
120 OVR) General Bacteriology Laboratory. The isolates were recovered from culture
121 samples obtained during routine-controlled inspections at ports of entry (POEs), routine
122 diagnostic services, and research studies at supermarkets. Sampling was not part of this
123 study; however, isolates were sent to the ARC-OVR General Bacteriology Laboratory
124 following the protocol recommended in Veterinary Procedural Notification (VPN) 56,
125 accompanied by all necessary documentation and metadata. This included details such
126 as the source of isolation, country or location of origin, and date of isolation or sample
127 collection, among other relevant information.

128 **Sample isolation, DNA extraction, and WGS**

129 Bacterial isolation and identification were carried out following the guidelines outlined in
130 "Microbiology of the Food Chain: Horizontal Method for the Detection, Enumeration, and
131 Serotyping of *Salmonella*" (ISO 6579-1:2017) [26]. Briefly, genomic DNA was extracted
132 from overnight cultures using the High Pure PCR Template Preparation Kit (Roche,
133 Germany) in accordance with the manufacturer's instructions. WGS of the isolates was
134 conducted at the Biotechnology Platform of the Agricultural Research Council (ARC),
135 Onderstepoort, South Africa. DNA libraries were prepared using TruSeq DNA Library

136 Preparation Kits (Illumina, San Diego, CA, USA) and sequenced on the HiSeq 2500
137 platform (Illumina, San Diego, CA, USA), according to the manufacturer's instructions.

138 **Quality control, pre-processing, and assembly of WGS data**

139 Trimming/filtering of raw Illumina paired-end reads was performed using fastp v0.23.4
140 with default parameters [27]. The resulting trimmed reads were assembled using Shovill
141 v1.1.0 with the `--assembler skesa` parameter (otherwise default parameters;
142 https://github.com/tseemann/shovill) [28]. Quality control of the resulting assemblies was
143 performed using: (i) QUAST v5.3.0 [29], with the complete *Salmonella enterica* sequence
144 type 548 (ST548) str. ATCC 49284 genome as the reference (`-r` option, NCBI Assembly
145 accession GCA_000486855.2; see section "*In silico* serotyping and sequence typing"
146 below for sequence typing details) [30]; (ii) CheckM v1.2.2 [31], using the `lineage_wf`
147 workflow with default parameters (**Dataset S1**).

148 ***In silico* serotyping and sequence typing**

149 *In silico* serotyping of assembled genomes (see section "Quality control, pre-processing,
150 and assembly of WGS data" above) was performed using (i) SISTR v1.1.1 (default
151 parameters) [32] and (ii) SeqSero2 v1.3.1 (`-m k -t 1`; otherwise default parameters) [33].
152 *In silico* seven-gene multi-locus sequence typing (MLST) was performed using mlst v2.9
153 (https://github.com/tseemann/mlst) with default parameters (**Dataset S1**), assigning all
154 genomes sequenced here to ST548.

155 **Acquisition of publicly available WGS data and metadata**

156 All publicly available ST548 genomes (see section "*In silico* serotyping and sequence
157 typing" above) and associated metadata were retrieved from Enterobase (*n* = 595 ST548
158 genomes, accessed 26 October 2022) [34, 35]. Samples were filtered to include only
159 those associated with a BioProject [36] and for which a NCBI Assembly accession and
160 year of isolation were available. In addition, only assemblies with (i) N50 > 50 Kbp, (ii)
161 number of contigs (>=200bp) < 200, (iii) SISTR1 serotype = SeqSero2 serotype =
162 "Minnesota", and (iv) species percentage >= 85% were retained (assembly statistics
163 obtained via Enterobase). This search resulted in a total of 229 publicly available
164 genomes, which were downloaded from NCBI's Assembly database (accessed 18 July
165 2024; **Dataset S2**).

166 To conduct a rough assessment of overall genome similarity among the 229 publicly
167 available ST548 genomes, a Mash distance-based tree [37] was constructed using
168 mashtree v1.4.6 [38], using `--genome-size 4716019` (i.e., the average length of the 229
169 publicly available genomes). Inspection of the resulting Mash distance-based tree
170 revealed a single outlier genome (NCBI Assembly accession GCA_011665355.1); this
171 outlier genome was relatively distant from all other publicly available ST548 genomes
172 and was thus not included in further analyses (**Fig S1**). The remaining 228 publicly
173 available ST548 genomes were combined with the 36 newly sequenced genomes (see
174 section "Sample isolation, DNA extraction, and WGS" above) to form the full dataset used
175 in subsequent analyses (*n* = 264 total ST548 genomes, referred to hereafter as the "full
176 ST548 dataset"; **Table S1**).

177 **AMR determinant and plasmid replicon detection**

178 All 264 genomes in the full ST548 dataset (see section "Acquisition of publicly available
179 WGS data and metadata" above) were screened for AMR determinants using
180 AMRFinderPlus v3.12.8 [39, 40], with the `-O Salmonella --plus` options and database
181 version 2024-05-02.2 (otherwise default parameters; **Dataset S3**). ABRicate v1.0.1
182 (https://github.com/tseemann/abricate) was used with the PlasmidFinder database (`--db
183 plasmidfinder`, otherwise default parameters) to identify plasmid replicons in each
184 genome (PlasmidFinder version 2023-Nov-4; **Dataset S4**) [41].

185 **Pan-genome analysis**

186 All 264 genomes in the full ST548 dataset (see section "Acquisition of publicly available
187 WGS data and metadata" above) were annotated using Bakta v1.8.2 with database
188 version 5.0 [42]. The resulting .gff3 files were used as input for the pan-genome estimation
189 software Panaroo v1.3.4 [43] with the following parameters: `-f 0.5 --core_threshold 0.95
190 --remove-invalid-genes --clean-mode strict -a pan`. The resulting
191 `gene_presence_absence.Rtab` file was imported into R v4.4.0 (https://www.R-
192 project.org/) and filtered to contain only intermediate genes (i.e., genes present in <95%
193 and >5% of the population; **Dataset S5**). The intermediate gene presence/absence matrix
194 was supplied as input to the R package umap v0.2.10.0 (https://CRAN.R-
195 project.org/package=umap), and the resulting Uniform Manifold Approximation and
196 Projection for Dimension Reduction (UMAP) [44] was visualized using ggplot2 v3.5.1
197 [45].

198 **Core SNP calling and phylogeny construction**

199 Pairwise average nucleotide identity (ANI) values were calculated between all 264
200 genomes in the full ST548 dataset (see section "Acquisition of publicly available WGS
201 data and metadata" above) using skani v0.2.0 [46] with the `--min-af 0` option (otherwise
202 default parameters; **Dataset S6**). The genome with the highest average ANI relative to
203 all other genomes was selected as the reference genome for reference-based SNP
204 calling (NCBI Assembly accession GCA_010125045.1). A core genome alignment was
205 generated using Snippy v4.6.0 (default parameters; https://github.com/tseemann/snippy).
206 Recombinant regions in the core genome alignment were masked using Gubbins v3.3.1
207 [47] with the options `--tree-builder iqtree-fast --first-tree-builder iqtree-fast` (otherwise
208 default parameters). snp-sites v2.5.1 [48] with `-c` (otherwise default parameters) was
209 used to extract variable sites in the recombination-masked alignment, resulting in a final
210 recombination-free core SNP alignment. The final recombination-free core SNP
211 alignment was used to calculate all-vs-all pairwise SNP distances using snp-dists v0.8.2
212 with default parameters (https://github.com/tseemann/snp-dists; **Dataset S7**). The above
213 workflow was repeated to generate all-vs-all pairwise SNP distances for the 36 newly
214 sequenced genomes alone, with publicly available genomes omitted (**Dataset S8**).
215 Pairwise SNP distances were visualized using the R package pheatmap v1.0.12
216 (https://CRAN.R-project.org/package=pheatmap) with average linkage hierarchical
217 clustering performed by hclust (from R package stats v4.3.2).

218 Phylogenetic inference was performed in IQ-TREE v2.2.5 [49], using the final
219 recombination-free core SNP alignment for the full ST548 dataset as input and the

220 following parameters: (i) the GTR+R nucleotide substitution model [50–52]; (ii) 1,000
221 ultrafast bootstrap replicates [53]; (iii) the number of constant sites in the alignment
222 (obtained using snp-sites v2.5.1 with the `-C` option), specified using the `-fconst`
223 parameter (i.e., `1078511,1182908,1166762,1071765` for A, C, G, and T, respectively);
224 (iv) `--seed` was set to 1000. The resulting unrooted maximum likelihood (ML) phylogeny
225 was imported into R v4.4.0, using the package ggtree v3.11.1 [54], and subsequently
226 midpoint rooted using the `midpoint_root` function in phytools v2.1-1 (referred to hereafter
227 as the "midpoint-rooted ML phylogeny"; **Fig S3**) [55].

228 Two rooted, time-scaled phylogenies were additionally constructed using Rlsd2 v2.4.4
229 (https://github.com/tothuhien/Rlsd2) [56], with the unrooted ML phylogeny from IQ-TREE
230 supplied as input, and tip dates corresponding to the year of isolation reported for each
231 genome: (i) a fixed-rate ML phylogeny, with the options `seqLen=4708764,estimateRoot
232 = "as",confidenceInterval = 1000,givenRate = 1.27e-7` (**Fig S4**) [14]; (ii) a LSD2-
233 estimated-rate phylogeny, constructed as described in (i), but without supplying
234 `givenRate = 1.27e-7` to Rlsd2 (**Fig S5**). The LSD2-estimated-rate phylogeny indicated
235 that a temporal signal may be present in the full ST548 dataset, although confidence
236 intervals for the estimated rate and time to most recent common ancestor (tMRCA) were
237 relatively large: LSD2 rate = 1.88268e-07 substitutions/site/year, confidence interval
238 [6.9163e-08, 3.03092e-07]; tMRCA = 1886.88 [1685.09, 1936.07]. However, subsequent
239 analysis of the unrooted ML phylogeny in TempEst v1.5.3 [57] did not reveal any evidence
240 of a temporal signal (using the "best-fitting root" option: slope = 0, tMRCA = 2028, R-
241 squared = 0.341) [57]. As such, Bayesian time-scaled phylogeny construction was not

242 performed, and all three ML phylogenies (i.e., the midpoint-rooted, LSD2-fixed-rate, and
243 LSD2-estimated-rate ML phylogenies) are provided (**10.5281/zenodo.15063662**).

244 **Core SNP calling and ML phylogeny construction for the high-AMR group**

245 Additional ML phylogenies were constructed for a well-supported clade within the full
246 ST548 dataset, which harbored many AMR determinants (i.e., the "high-AMR group").
247 Briefly, genome names for members of the high-AMR group were extracted from the
248 larger ST548 phylogeny using the `mrca` function in the R package ape v5.7-1 (*n* = 108
249 high-AMR group genomes) [58]. The Snippy/Gubbins/snp-sites pipeline described above
250 was used to identify core SNPs among all 108 high-AMR group genomes (see section
251 "Core SNP calling and phylogeny construction"). The resulting recombination-free core
252 SNPs were then used to build (i) midpoint-rooted (**Fig S6A**) and (ii) LSD2-estimated-rate
253 ML phylogenies (**Fig S6B**) as described above (see section "Core SNP calling and
254 phylogeny construction"). Evidence of a temporal signal was observed for the high-AMR
255 group using both LSD2 and TempEst: LSD2 rate = 5.98136e-07 [3.97087e-07; 8.28845e-
256 07], tMRCA = 2009.89 [2004.52; 2012.52]; TempEst best-fitting root slope = 7.6933e-07,
257 tMRCA = 2009.3, R-squared = 0.5046 for the unrooted ML phylogeny
258 (**10.5281/zenodo.15063662**). As such, a Bayesian time-scaled phylogeny was
259 constructed for the high-AMR group (described below).

260 **Bayesian phylogenetic reconstruction of the high-AMR group**

261 BEAUti2 v2.7.6 and BEAST2 v2.7.6 [59] were used to construct a Bayesian time-scaled
262 phylogeny of the high AMR group, using the high-AMR group recombination-free core

263 SNP alignment as input and tip dates corresponding to each genome's year of isolation
264 (see section "Core SNP calling and ML phylogeny construction for the high-AMR group"
265 above). A combination of (i) strict clock (starting clock rate = 1.27e-7
266 substitutions/site/year, with a lognormal prior on the clockRate parameter: mean = 2.77e-
267 7, standard deviation = 1.25, lower and upper bounds = [0, ∞]) [14] and (ii) coalescent
268 Bayesian skyline model [60] were selected as the clock and population model,
269 respectively, as these models have been identified previously as the optimal
270 clock/population model combination for *S.* Minnesota [14]. The nucleotide substitution
271 model was inferred using Bayesian model averaging (transitions and transversions split)
272 with bModelTest v1.3.3 [61], and an ascertainment bias correction was used to correct
273 for the lack of constant sites in the input core SNP alignment (number of constant sites
274 obtained as described above in "Core SNP calling and ML phylogeny construction";
275 https://groups.google.com/g/beast-users/c/QfBHMOqImFE, accessed 13 March 2025).

276 Three independent BEAST2 runs were performed (using seeds 1000, 1001, and 1002),
277 with chain lengths of 2 x 1e8 generations and parameters logged every 1e4 generations.
278 The log file for each independent BEAST2 run was loaded into Tracer v1.7.2 to assess
279 mixing (using 10% burn-in) and to confirm that effective sample size (ESS) values for all
280 parameters were >200. The resulting log and tree files were combined using
281 LogCombiner v2.7.6 with 10% burn-in. The final maximum clade credibility (MCC) tree
282 was generated using TreeAnnotator v2.7.6 with common ancestor node heights. This
283 MCC tree was loaded into R for visualisation as described above (see section "Core SNP
284 calling and ML phylogeny construction"). The median effective population size over time

285 was estimated using Bayesian skyline analysis in Tracer v1.7.2, using default parameters
286 and the most recent sampling date set to `2022.5`.

287 To validate the observed evolutionary rate and tree height (i.e., to ensure that they were
288 the result of a true temporal signal and not an artifact), a date randomization test was
289 performed [62]. Briefly, sampling dates were randomized using the R package
290 TipDatingBeast v1.1.0 [63], and the above BEAST2 workflow was performed for 10
291 independent sets of date-randomized samples. For these 10 randomized-date sample
292 sets, models/parameters were identical to those described above, except for upper and
293 lower clock-rate bounds on the “clockRate” parameter, which were set to 1e-3 and 1e-10,
294 respectively (due to a ‘positive infinite posterior’ error, which occurred when bounds were
295 not set, suggesting that our clock rate was escaping to extreme values in the date-
296 randomized datasets). The date randomization test confirmed that there was no overlap
297 between the 95% highest posterior density (HPD) intervals of the true clock rate/tree
298 height and the clock rates/tree heights of the 10 independent date-randomized datasets
299 (**Fig S7AB**). A prior-only run without any sequence data (‘sample from prior’ enabled in
300 BEAUti2) was also performed to confirm minimal overlap in the resulting parameters
301 between our prior-only analysis and our analysis with sequence data. Log files, XML files,
302 and tree files for all BEAST2 runs described above are provided in the Zenodo repository
303 (10.5281/zenodo.15063662).

304 **Identification of plasmid-mediated AMR genes**

305 To determine which AMR genes were likely harbored on plasmids, the `end-to-end`
306 command in geNomad v1.7.1 [64] was used to classify all contigs in the full ST548 dataset

307 as “chromosomal”, “plasmid”, or “viral” (see section “AMR determinant and plasmid
308 replicon detection” above; **Dataset S9**). “Aggregated-classification” scores were
309 compared between AMR gene-harboring (AMR+) contigs and contigs without AMR genes
310 (AMR-; as determined by AMRFinderPlus) using the Kruskal-Wallis test in R
311 (`kruskal.test` function). The “aggregated-classification” scores for the AMR+ contigs
312 were then grouped by AMR gene and plotted in R using ggplot2 v3.5.1.

313 The plasmid database PLSDB version ‘2024_05_31_v2’ [65] was downloaded and
314 converted into a BLAST v2.15.0 database using makeblastdb [66]. SeqKit v2.7.0 [67] was
315 used to extract AMR+ contigs from the corresponding assemblies, and the contigs were
316 then queried against the PLSDB BLAST database using megablast (`blastn` command
317 with default parameters). Contigs that aligned to plasmid sequences in the PLSDB
318 database at >95% coverage and >95% identity were considered to be candidate AMR
319 gene-harboring plasmids (**Dataset S10**).

320 **Analysis of *bla*$_{\text{CTX-M-8}}$-harboring contigs**

321 Beta-lactamase *bla*$_{\text{CTX-M-8}}$ was detected in several closely related ST548 genomes
322 sequenced here (referred to hereafter as “ZA SNP cluster 0” genomes; described in detail
323 in the “Results” section below). All *bla*$_{\text{CTX-M-8}}$-harboring (*bla*$_{\text{CTX-M-8}}$+) contigs underwent
324 further investigation. Specifically, all *bla*$_{\text{CTX-M-8}}$+ contigs matched a ~90Kbp IncI pST113
325 plasmid from *Eschericha coli* (*E. coli*) in the PLSDB BLAST database with >95% coverage
326 and >95% identity (NCBI Nucleotide accession CP134392.1, see section “Identification
327 of plasmid-mediated AMR genes” above; **Dataset S10**). The accession number
328 CP134392.1 was supplied as the query for a NCBI Nucleotide BLAST search against the

329 `core_nt` database, using `blastn` as the selected program and the search limited to
330 organism `Salmonella enterica (taxid:28901)` (https://blast.ncbi.nlm.nih.gov/Blast.cgi,
331 search conducted on 21 March 2025; **Dataset S11**) [68]. Aligned sequences for all hits
332 were downloaded and underwent AMR determinant detection using AMRFinderPlus with
333 the `-O Salmonella --plus` options and database version 2024-05-02.2 (otherwise default
334 parameters). This was done to identify similar plasmids in *Salmonella enterica*, and to
335 determine if the hits also contained $bla_{\text{CTX-M-8}}$. The only $bla_{\text{CTX-M-8}}$+ hits were in two
336 plasmids, both having >99% identity and >99% query coverage to the IncI pST113
337 plasmid from *E. coli*: (i) a ~90Kbp plasmid previously identified in *S.* Mbandaka (NCBI
338 Nucleotide accession CP146619.1) [69], and (ii) a recently uploaded ~150Kbp plasmid
339 sequence from *S.* Enteritidis (NCBI Nucleotide accession CP183502.1; **Dataset S12**).
340 As the *S.* Mbandaka complete plasmid sequence was of similar size to the *E. coli* IncI
341 plasmid (both ~90Kbp) and associated with a publication [68], it was chosen as the
342 reference sequence for further analysis.

343 **Reconstruction of $bla_{\text{CTX-M-8}}$ + plasmid from ZA SNP cluster 0 genomes**

344 Trimmed forward and reverse reads from all ZA SNP cluster 0 genomes (see section
345 "Analysis of $bla_{\text{CTX-M-8}}$-harboring contigs" above) were combined to create pooled forward
346 and reverse reads. The pooled ZA SNP cluster 0 reads were then fed to the
347 `plasmidspades` command in SPAdes v4.0.0 [70], which was used to assemble plasmids
348 among all pooled reads using the `--only-assembler` parameter. The resulting
349 `contigs.fasta` file, the *S.* Mbandaka $bla_{\text{CTX-M-8}}$+ IncI plasmid (NCBI Nucleotide accession
350 CP146619.1), and the *E. coli* $bla_{\text{CTX-M-8}}$+ IncI plasmid (NCBI Nucleotide accession

351 CP134392.1) were annotated using Bakta v1.8.2 (see sections “AMR determinant and
352 plasmid replicon detection” and “Analysis of *bla*$_{CTX-M-8}$-harboring contigs” above). Using
353 the corresponding Bakta .gbk files for these three plasmids, plasmid comparison and
354 visualisation was performed using Clinker v0.0.31 [71] with the `-i = 0.5` option. According
355 to plasmidSPAdes, all ZA SNP cluster 0 contigs that mapped to CP146619.1 belonged
356 to the same “plasmid component” (i.e., were predicted to be on the same plasmid). For
357 the purposes of visualisation, contigs from the ZA SNP cluster 0 plasmidSPAdes
358 assembly that did not align to the *S.* Mbandaka/*E. coli* reference IncI plasmids (i.e.,
359 assembled contigs assigned to a different “plasmid component”) were excluded from the
360 figure. ABRicate v1.0.1 was used with the PlasmidFinder database to detect plasmid
361 replicons in the plasmidSPAdes contigs as described above (see section “AMR
362 determinant and plasmid replicon detection”; **Dataset S13**). We would like to note that
363 while our PlasmidFinder results did not show the presence of IncI1 in any of the ZA SNP
364 cluster 0 genomes individually (**Fig S3, Dataset S4,** see “AMR determinant and plasmid
365 replicon detection” section above), IncI1 was detected in the pooled plasmidSPAdes
366 assembly.

367 **Identification of *bla*$_{CTX-M-8}$+ and other ZA ST548 genomes in the AllTheBacteria**
368 **dataset**

369 To assess the prevalence of *bla*$_{CTX-M-8}$ in other bacterial lineages (e.g., other species,
370 other *Salmonella* serotypes and sequence types, other ST548 genomes sequenced after
371 our Enterobase data freeze), AMRFinderPlus results from AllTheBacteria (ATB) v0.2
372 (https://allthebacteria.readthedocs.io/en/latest/amr.html) were downloaded [72]. The ATB

373 AMRFinderPlus results were then searched for “bla-CTX-M-8” (**Dataset S14**). All *bla*$_{CTX-M-8}$-harboring ATB genomes classified as “*Salmonella enterica*” were downloaded (*n* = 64), and the resulting genomes underwent *in silico* seven-gene MLST using mlst v2.9 with default parameters. Sixteen genomes were assigned to ST548; these 16 genomes, plus the 14 ZA SNP cluster 0 genomes (see section “Analysis of *bla*$_{CTX-M-8}$-harboring contigs” above), underwent all-vs-all pairwise SNP distance estimation as described above (see section “Core SNP calling and phylogeny construction”; **Dataset S15**). Metadata for all 354 *bla*$_{CTX-M-8}$-harboring genomes from ATB were extracted using Entrez Direct v22.4 (**Dataset S16**) [73].

382 In addition to the 16 ST548 *bla*$_{CTX-M-8}$-harboring genomes, we identified 10 other ST548 genomes from South Africa in the ATB dataset (all 10 were *bla*$_{CTX-M-8}$-negative). These 10 genomes were present in Enterobase but did not have an associated NCBI assembly and were thus not included in our original data freeze. All-vs-all pairwise SNP distance estimation as described above (see section “Core SNP calling and phylogeny construction”) was performed with these 10 publicly available South African ST548 genomes and our 36 newly sequenced genomes. (**Dataset S17**).

389 **RESULTS**

390 **Multiple *S.* Minnesota ST548 lineages can be detected in South African poultry**
391 **imports of Brazilian origin**

392 A total of 36 *Salmonella enterica* strains isolated from frozen raw poultry samples
393 collected between 2020-2022 underwent WGS (see section "Sample isolation, DNA
394 extraction, and WGS" above). While all 36 strains were isolated within South African
395 borders, the frozen raw poultry from which each strain was isolated originated from Brazil.
396 More specifically, (i) 11 out of 36 genomes (31%) were isolates originating from Brazilian
397 poultry meat collected for regulatory compliance, at port-of-entry (POE); (ii) the remaining
398 25 isolates (69%) were from imported Brazilian poultry meat placed for sale in South
399 African supermarkets (referred to hereafter as "**POE**" and "**supermarket**" genomes,
400 respectively, with the total set of 36 genomes referred to as "**SEPI**" genomes [*Salmonella*
401 *enterica* Poultry Imports]; **Table S1**).

402 Using *in silico* seven-gene MLST, all SEPI genomes belonged to ST548. *In silico*
403 serotyping identified 35 genomes as serotype Minnesota (antigenic formula 21:b:e,n,x)
404 by both SISTR and SeqSero2. One genome (strain 73, NCBI BioSample accession
405 SAMN47166240) was assigned a serotype of "4:r:e,n,x" by SeqSero2 and "-:b:e,n,x" by
406 SISTR. However, using core genome MLST (via SISTR), all SEPI genomes were
407 identified as serotype Minnesota and will thus be referred to hereafter as *S.* Minnesota
408 (see section "*In silico* serotyping and sequence typing" in the Methods above).

409 Despite belonging to the same ST, and despite sharing a similar origin (i.e., frozen raw
410 poultry imported into South Africa from Brazil), considerable genomic diversity could be
411 observed across SEPI genomes. ANI values calculated between SEPI genomes ranged
412 from 99.84% - 100%. After removing recombination, the SEPI genomes differed by 0-85
413 core SNPs (mean 46, median 55; **Fig 1A**). The range of core SNP distances was not
414 affected by the specific source of each genome, i.e., the core SNP distance range was
415 approximately 0-85, even when comparing supermarket genomes only or POE genomes
416 only (**Fig 1B**). Using average linkage hierarchical clustering of the pairwise core SNP
417 distances, SEPI genomes could be partitioned into 7 SNP clusters at a threshold of 50
418 SNPs, to 21 SNP clusters at a threshold of 5 SNPs (**Fig 1C**). Altogether, these results
419 indicate that multiple *S.* Minnesota ST548 lineages are present in our collection of isolates
420 from Brazilian frozen poultry imports.

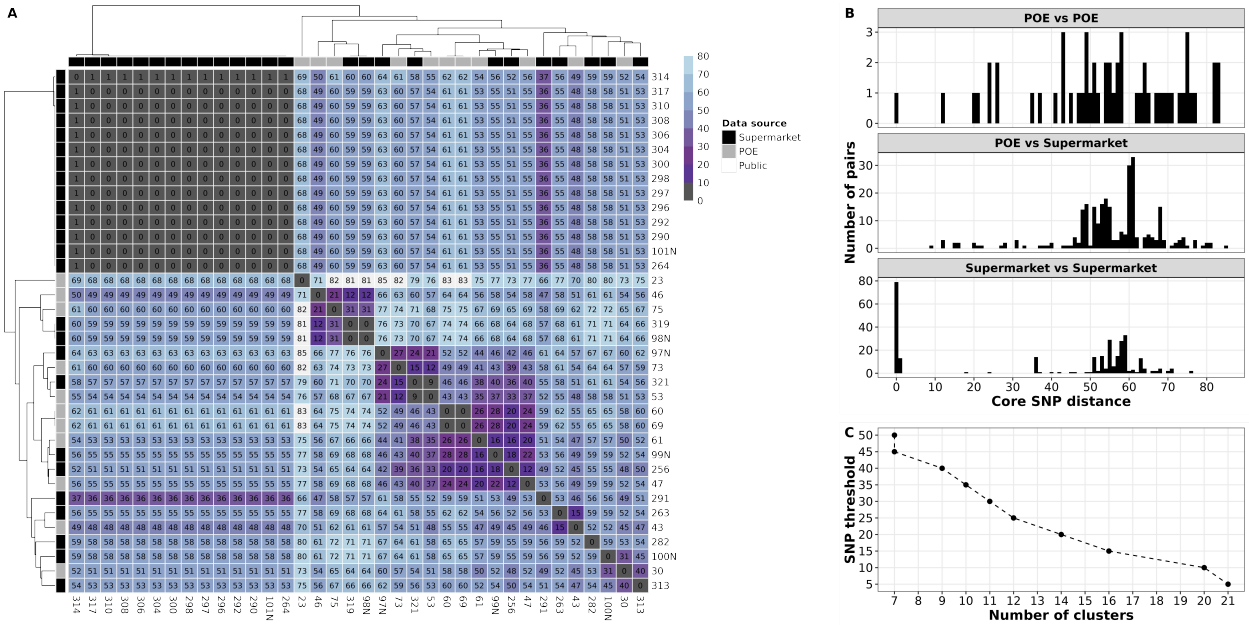


Fig 1. Core genome SNP distances show that multiple *S.* Minnesota ST548 lineages can be detected in Brazilian poultry imported into South Africa.

(A) Heatmap showing pairwise core SNP distances calculated between the 36 ST548 genomes sequenced here (i.e., SEPI genomes). For the heatmap cells, darker shading indicates lower core SNP distances (i.e., more closely related genomes), while lighter shading indicates higher core SNP distances. Dendrograms were produced via average linkage hierarchical clustering of core SNP distances. Color strips above and to the left of the heatmap denote the source of isolation for each genome ("Data source"). POE - port-of-entry; Supermarket - imported poultry from supermarkets in South Africa. (B) Histogram of pairwise core SNP distances among the SEPI genomes separated by specific source of isolation. "POE vs POE" - comparisons within port-of-entry genomes only; "POE vs Supermarket" - comparisons between port-of-entry and supermarket genomes; "Supermarket vs Supermarket" - comparisons within supermarket genomes only. (C) Core SNP distance-based clustering of SEPI genomes. The X-axis shows the number of clusters at the corresponding core SNP threshold in the Y-axis. Clusters were obtained via average linkage hierarchical clustering (see "Core SNP calling and phylogeny construction" in the Methods above).

439 ***S.* Minnesota ST548 strains detected in South African supermarkets are closely**
440 **related to strains isolated from Brazilian poultry imported into the United Kingdom**

441 To gain insight into the evolution of South African ST548 within the context of the global
442 ST548 population, we compared our 36 SEPI genomes to 228 high-quality, publicly
443 available ST548 genomes (*n* = 264 total ST548 genomes; see section "Acquisition of
444 publicly available WGS data and metadata" in the Methods above). Using Snippy, a total
445 of 4,432 core SNP sites were identified among all 264 ST548 genomes, with SNP
446 distances ranging from 0-322 (mean 160.5, median 193; **Fig S2A**).

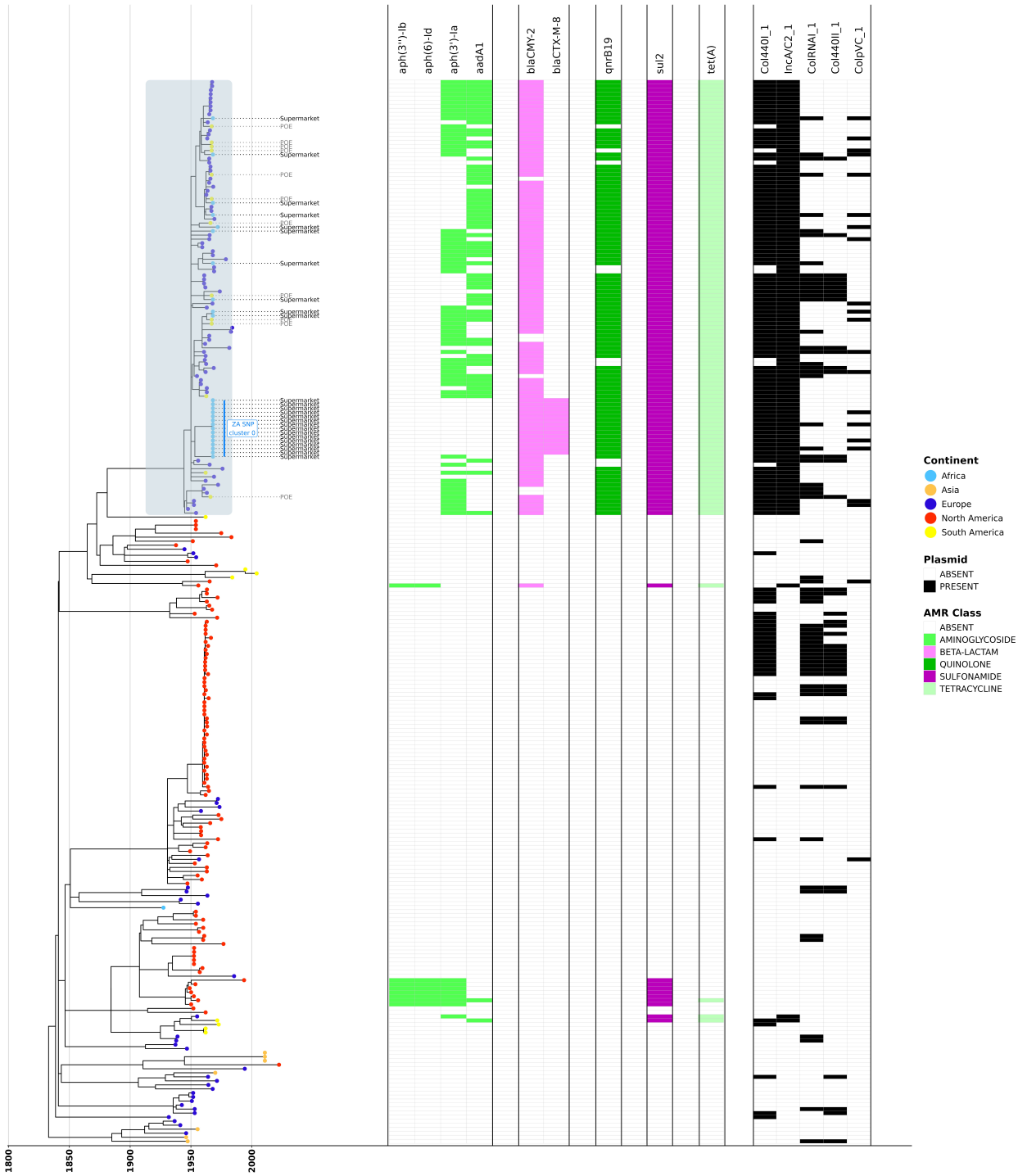
447 For each SEPI genome sequenced here, we identified the corresponding publicly
448 available genome with the lowest SNP distance. This approach identified 16 distinct
449 publicly available genomes (ties were kept), which differed from SEPI genomes by 2-27
450 core SNPs (**Table S2**). Notably, 18 SEPI genomes (17 supermarket genomes) were 19-
451 27 SNPs apart from one Brazilian poultry-associated public genome (NCBI Assembly
452 accession GCA_006209225.1). This suggests that nearly 70% of our supermarket
453 genomes (17 out of 25 genomes) may share a recent common ancestor with a poultry-
454 associated isolate from Brazil. In addition, 11 other ST548 strains from a Public Health
455 England collection (NCBI BioProject accession PRJNA248792) were 5-50 core SNPs
456 apart from all SEPI genomes. All 11 of these strains, though isolated in the UK, were
457 associated with Brazilian imported poultry. This links SEPI isolates from South Africa with
458 previously sequenced, poultry-associated strains from the UK, all of which share a
459 common Brazilian origin. Collectively, these results suggest that while some South
460 African ST548 genomes are closely related (<10 core SNPs) to previously sequenced

461 food-associated genomes, others may represent previously unsampled lineages of
462 similar origin (>20 core SNPs).

463 ***S.* Minnesota ST548 strains from South African poultry imports of Brazilian origin**
464 **are confined to a largely AMR clade**

465 A ML phylogeny constructed using the 4,432 core SNPs detected among all 264 ST548
466 genomes revealed a clear separation in line with geographical origin, AMR profile, and
467 core SNP distances (**Fig 2, Fig S2AB**). One clade comprising 108 genomes of European,
468 African, and Brazilian origin (ultrafast bootstrap support = 99%; referred to hereafter as
469 the "high-AMR group") was found to contain 88% of all AMR genes detected in our total
470 ST548 sample set (937 of 1,063 AMR genes; **Fig 2, S3,S4, S5**). Further, all 108 genomes
471 in the high-AMR group were predicted to be multidrug-resistant (MDR), i.e., harbored
472 AMR genes conferring resistance to ≥3 antibiotic classes. The high-AMR group harbored
473 67% (70 of 104) of all European, 96% (25 of 26) of all African, and 59% (13 of 22) of all
474 South American ST548 genomes (**Fig S5, Fig S6**). Notably, the high-AMR group also
475 included all 36 of our SEPI genomes. The remaining genomes from Europe, Africa, and
476 South America (*n* = 40), along with all North American and Asian genomes (*n* = 106 and
477 7, respectively) belonged to other clades within the ST548 phylogeny and contained only
478 ~12% of all AMR genes detected (126 of 1,063 AMR genes; *n* = 154 total genomes,
479 referred to hereafter as the "low-AMR group"). A predicted MDR phenotype was observed
480 among low-AMR group genomes only sporadically (13 of 154 low-AMR group genomes,
481 8.4%; **Fig S5**).

482


aph(3'')-Ib
aph(6)-Id
aph(3')-Ia
aadA1
blaCMY-2
blaCTX-M-8
qnrB19
sul2
tet(A)
Col440I_1
IncA/C2_1
ColRNAI_1
Col440II_1
ColpVC_1
Supermarket
POE
ZA SNP cluster 0
1800
1850
1900
1950
2000
Continent
Africa
Asia
Europe
North America
South America
Plasmid
ABSENT
PRESENT
AMR Class
ABSENT
AMINOGLYCOSIDE
BETA-LACTAM
QUINOLONE
SULFONAMIDE
TETRACYCLINE

483

484 **Fig 2. LSD2-estimated-rate maximum likelihood (ML) phylogeny of *S.* Minnesota**
485 **ST548 shows that the majority of AMR genes are confined to a single clade** (*n* =
486 264; 36 SEPI genomes, plus 228 high-quality, publicly available *S.* Minnesota ST548
487 genomes).
488 The "high-AMR group" is denoted in the phylogeny by light blue shading. Tree tips for the
489 36 SEPI genomes are annotated with the isolation source ("POE" for strains isolated at
490 port-of-entry; "Supermarket" for poultry samples from supermarkets in South Africa).
491 Unlabelled tips denote publicly available genomes from Enterobase. Tips are colored by
492 continent of isolation ("POE" and "Supermarket" genomes sequenced in this study were
493 classified as "South American" and "African" isolates, respectively). Heatmaps to the right
494 of the phylogeny denote (from left to right): (i) presence/absence of AMR genes detected
495 by AMRFinderPlus, colored by AMR class; (ii) presence/absence of plasmid replicons
496 (detected using ABRicate/PlasmidFinder). For readability, only select AMR genes and
497 plasmid replicons are shown (see Fig S5 for the full version of this figure). The ML
498 phylogeny was constructed using IQ-TREE, and the evolutionary rate was estimated by
499 LSD2, with branch lengths reported in years (to view phylogenies constructed using
500 different rooting/scaling methods, see Fig S3, S4).

501 The separation between the high- and low-AMR groups was also evident when ST548
502 genomes were clustered by accessory gene presence/absence (i.e., genes present
503 in >5% and <95% of ST548 per Panaroo). As was the case with the core genome ML
504 phylogeny, all SEPI genomes clustered with publicly available genomes from the high-
505 AMR group (**Fig 3**). Interestingly, two genomes from the low-AMR group clustered with
506 high-AMR group genomes (NCBI Assembly accessions GCA_011578495.1 and
507 GCA_022318025.1; **Fig 3**). These two genomes belonged to the small group of sporadic
508 AMR+ isolates in the low-AMR group and were positive for aminoglycoside, sulfonamide,
509 and tetracycline resistance genes (**Fig S5**).

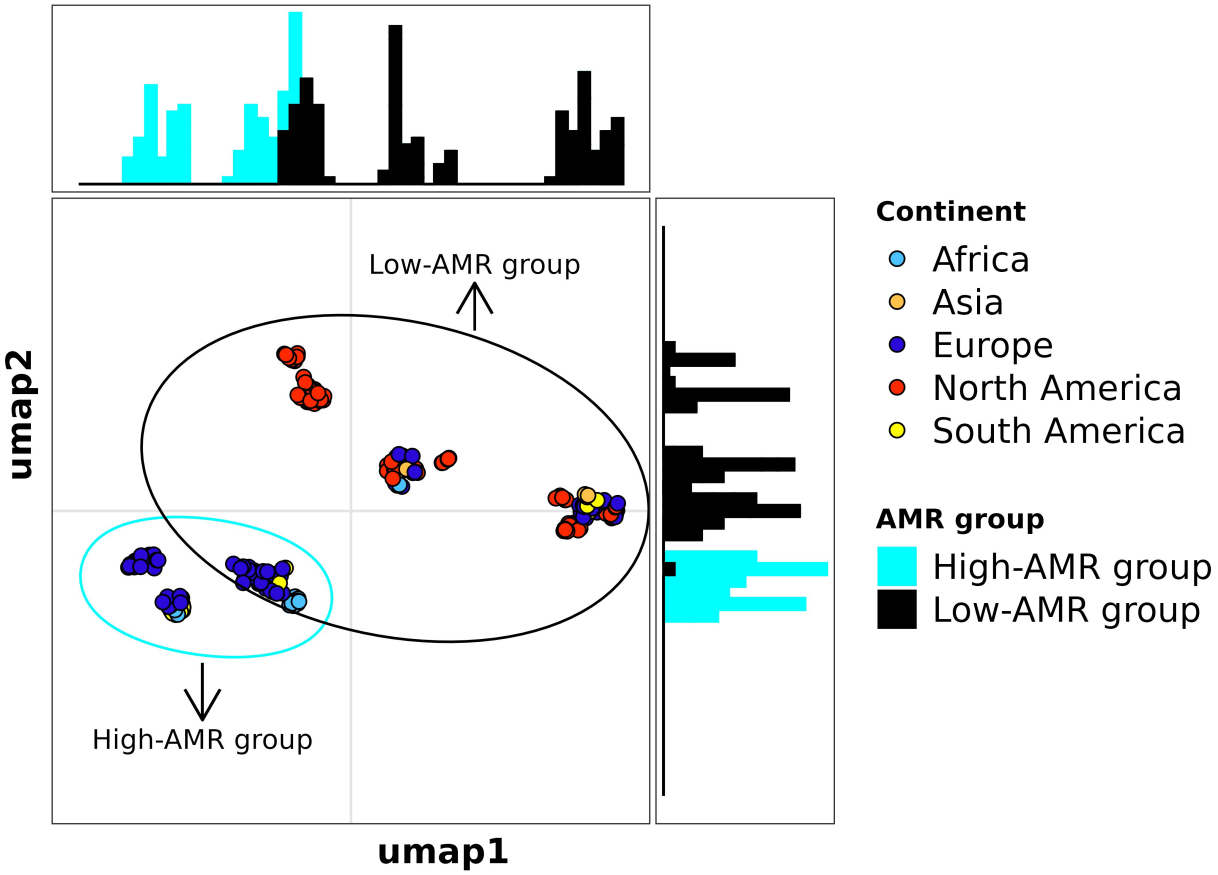


Fig 3. UMAP constructed using accessory gene presence/absence reveals separation between high-AMR and low-AMR groups.
The UMAP was constructed from a presence/absence matrix of intermediate genes detected among 264 *S.* Minnesota ST548 genomes (i.e., genes present in >5% and < 95% of the population according to Panaroo). Each point represents a genome, colored by the reported continent of isolation. Ellipses encompass all genomes from the high-AMR group (cyan) or low-AMR group (black). Histograms along the X and Y axes of the plot represent the distribution of high-AMR group and low-AMR group genomes along the corresponding axis in the UMAP.

520 To further investigate the emergence of the high-AMR group, we constructed a Bayesian
521 time-scaled phylogeny of the 108 high-AMR group genomes (see section “Bayesian
522 phylogenetic reconstruction of the high-AMR group” in the Methods above). The
523 evolutionary rate for the high-AMR group was estimated to be 6.262e-07
524 substitutions/site/year (95% HPD [5.428e-07, 7.102e-07]), with a tMRCA of 2010.485
525 (95% HPD [2009.02, 2011.83]; **Fig 4**). Bayesian skyline analysis estimated that the
526 effective population size of the high-AMR group underwent an increase in ~2013 and
527 then subsequently a decrease post 2020 (**Fig S7C**). However, such sharp population
528 changes can be an artifact of sampling and population structure as demonstrated
529 previously [74, 75].

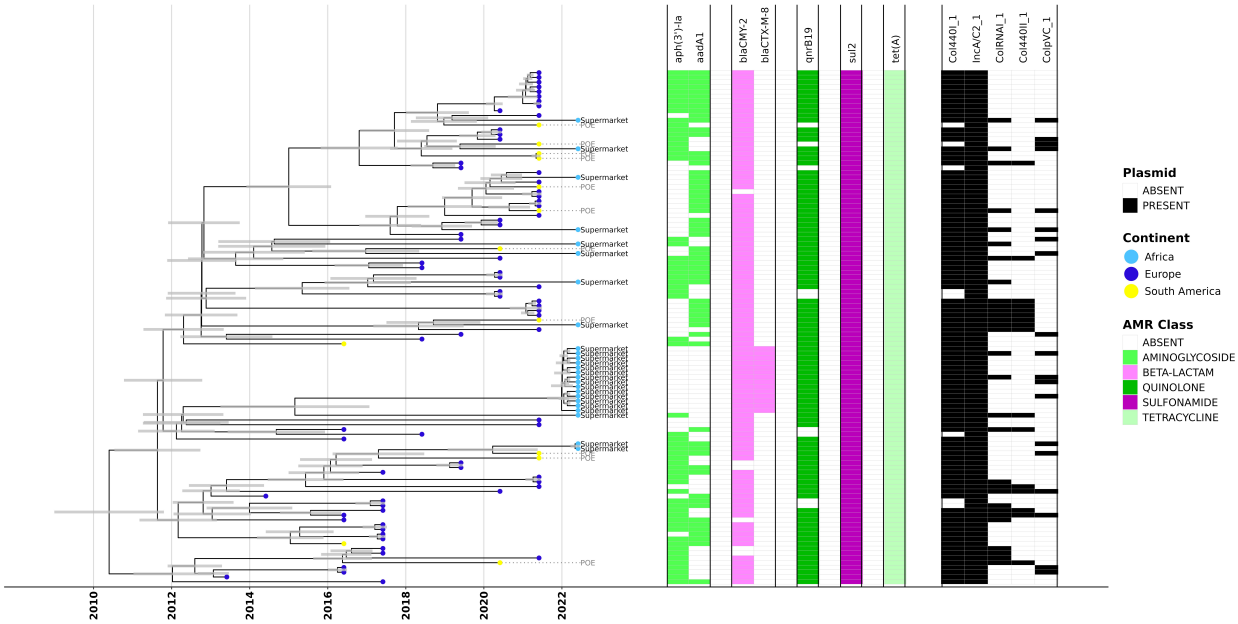


Fig 4. Bayesian phylogenetic reconstruction shows recent emergence of the high-AMR group (comprising 108 isolates, including all 36 SEPI genomes). The X-axis denotes time in years, and the grey node bars represent 95% highest posterior density (HPD) intervals for node heights. Tree tips for the 36 SEPI genomes are annotated with the isolation source ("POE" for strains isolated at port-of-entry; "Supermarket" for poultry samples from supermarkets in South Africa). Unlabelled tips denote publicly available genomes from Enterobase. Tree tips are colored by the continent of isolation ("POE" and "Supermarket" isolates sequenced in this study were classified as "South American" and "African" isolates, respectively). Heatmaps to the right of the phylogeny denote (from left to right): (i) presence/absence of AMR genes detected by AMRFinderPlus, colored by AMR class; (ii) presence/absence of plasmid replicons (detected using ABRicate/PlasmidFinder). Only select AMR genes and plasmid replicons are shown (to view the full phylogeny, see Fig S6).

544 **Antimicrobial resistance genes in *S.* Minnesota ST548 are largely plasmid-**
545 **associated**

546 To investigate whether AMR genes from the high-AMR group were plasmid-associated,
547 geNomad was used to predict the origin (chromosomal, plasmid, or viral) of assembled
548 contigs in the full ST548 set (**Fig 5**). geNomad predicted plasmid scores for contigs
549 harboring AMR genes (AMR+ contigs) were then compared to contigs without AMR genes
550 (AMR- contigs) as detected by AMRFinderPlus. Overall, plasmid scores were significantly
551 greater for AMR+ contigs compared to AMR- contigs (P < 2.22e-16, Mann Whitney U
552 test, **Fig S8**). Using blastn, each AMR+ contig aligned to ≥1 plasmid in PLSDB with >95%
553 coverage and >95% identity (see "Identification of plasmid-mediated AMR genes" in the
554 Methods above). Contigs harboring $bla_{\text{CMY-2}}$, *tet*(A), and *sul2* aligned to >400 distinct ~100
555 Kbp plasmids, each with coverage and identity >95%. Moreover, 70 of these ~100Kbp
556 plasmids were found to harbor all three AMR genes. This suggests that MDR-conferring
557 megaplasmids are in circulation in this lineage, as has been noted previously [17]. While
558 we were unable to pinpoint the exact plasmid due to the fragmented nature of short-read
559 assemblies, we identified reference sequences for potential candidates (NCBI Nucleotide
560 accessions NZ_CP012923.1, NZ_CP060509.1, or NZ_CP080427.1).

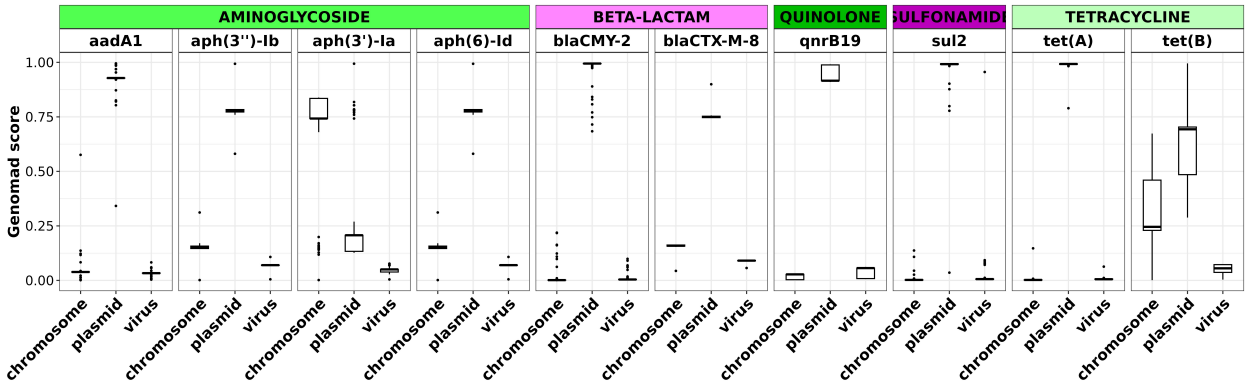


Fig 5. geNomad scores for ST548 AMR+ contigs show that most AMR genes are plasmid-associated.

Boxplots showing geNomad aggregated classification scores for AMR gene-harboring (AMR+) contigs present among 264 *S.* Minnesota ST548 genomes. Black horizontal lines within each boxplot denote the median geNomad score, and boxes show the interquartile range. Whiskers represent values up to 1.5x the first (lower) or third (higher) quartile, with black points showing outliers beyond the whiskers' range. Each facet shows the geNomad chromosome, plasmid, and phage score for the corresponding AMR gene. AMR genes are grouped by AMR class.

571 **A *bla*$_{CTX-M-8}$-harboring plasmid from *E. coli* has entered South Africa via Brazilian**
572 **poultry imports**

573 Among the 36 SEPI genomes sequenced here, the largest SNP cluster consisted of 14
574 isolates, all 0 core SNPs apart (referred to hereafter as "ZA SNP cluster 0"; **Fig 2 and Fig**
575 **S1**). All 14 ZA SNP cluster 0 genomes were isolated from poultry meat sold in
576 supermarkets in South Africa (**Fig 2**). A unique feature of genomes within ZA SNP cluster
577 0 was that they harbored *bla*$_{CTX-M-8}$, which was otherwise absent in our full *S.* Minnesota
578 ST548 dataset (**Fig 2**). This was in addition to *bla*$_{CMY-2}$, *sul2*, and *tet*(A), which were, as
579 stated previously, potentially carried by a megaplasmid (see section "Antimicrobial
580 resistance genes in *S.* Minnesota ST548 are largely plasmid associated" above).

581 To assess the prevalence of *bla*$_{CTX-M-8}$ in other *Salmonella enterica* lineages, we searched
582 a catalog of AMR determinants detected in >1.9 million bacterial genomes (i.e.,
583 AMRFinderPlus results from the AllTheBacteria [ATB] dataset) [72]. Out of 644,699 ATB
584 *Salmonella enterica* genomes, only 64 possessed *bla*$_{CTX-M-8}$ (0.01%), suggesting that
585 *bla*$_{CTX-M-8}$ is not common in *Salmonella enterica* (see "Identification of *bla*$_{CTX-M-8}$+ and
586 other ZA ST548 genomes in the AllTheBacteria dataset" in the Methods above). Out of
587 these 64 *bla*$_{CTX-M-8}$-harboring (*bla*$_{CTX-M-8}$+) ATB *Salmonella enterica* genomes, 16
588 belonged to ST548 (25%). These 16 *bla*$_{CTX-M-8}$+ ST548 ATB genomes were not included
589 in our initial dataset because there was no associated GenBank assembly available at
590 the time of our data freeze, or the existing assembly did not pass our quality filters (see
591 section "Acquisition of publicly available WGS data and metadata" in the Methods above).
592 Further investigation of these 16 ATB genomes revealed all 16 (100%) differed from ZA

593 SNP cluster 0 by >50 core SNPs (**Fig S8**), indicating that ZA SNP cluster 0 may not share
594 a common source with any of the *bla*$_{CTX-M-8}$+ ST548 genomes from ATB. Further, two of
595 these 16 *bla*$_{CTX-M-8}$+ ST548 ATB genomes (12.5%) differed from ZA SNP cluster 0
596 by >250 core SNPs, meaning they are likely not part of the high-AMR group itself (**Fig S9,**
597 **S2B**). Interestingly, one of these 16 *bla*$_{CTX-M-8}$+ ST548 ATB genomes was reported to be
598 from South Africa (NCBI BioSample accession SAMEA14452848), isolated in 2022 from
599 animal meat along with 10 other *bla*$_{CTX-M-8}$- samples. Each of these 11 South African
600 samples were 23-65 core SNPs apart from at least one SEPI genome sequenced here
601 (**Table S3**). From the remaining 15 *bla*$_{CTX-M-8}$+ ST548 ATB genomes, five were from
602 Brazilian poultry, isolated between 2016 and 2022; nine were reported to be isolated from
603 "food" in Portugal and the UK. The one remaining sample from the UK was reported to be
604 isolated from a human (NCBI BioSample accession SAMN30444884). Taken together,
605 our results suggest that, while *bla*$_{CTX-M-8}$ is relatively rare in *Salmonella enterica*, ST548
606 has been subjected to multiple independent introductions of *bla*$_{CTX-M-8}$.

607 All ST548 *bla*$_{CTX-M-8}$+ contigs (i.e., from ZA SNP cluster 0, as well as the 16 ST548
608 genomes from ATB) matched an IncI pST113 plasmid from *E. coli* with >95% coverage
609 and >95% identity (NCBI Nucleotide accession CP134392.1). A similar *bla*$_{CTX-M-8}$+
610 plasmid has previously been reported in *Salmonella enterica* serotype Mbandaka (*S.*
611 Mbandaka; NCBI Nucleotide accession CP146619.1) [69]. In an attempt to assemble the
612 complete plasmid, we pooled all reads from the ZA SNP cluster 0 genomes and ran
613 plasmidSPAdes. We then annotated the resulting contigs with Bakta and compared the
614 resulting gene clusters with the *S.* Mbandaka and *E. coli bla*$_{CTX-M-8}$+ IncI plasmids using

615 Clinker (**Fig 6**). This showed that, in addition to an existing MDR-conferring megaplasmid
616 commonly found in the high-AMR group of *S.* Minnesota ST548, 14 SEPI genomes have
617 an additional plasmid with $bla_{\text{CTX-M-8}}$. Using AMRFinderPlus results from ATB, we found
618 a total of 354 genomes with $bla_{\text{CTX-M-8}}$; of these, 248 were detected in *E. coli* (70.1%),
619 including the earliest reported $bla_{\text{CTX-M-8}}$+ sample in the ATB dataset (NCBI BioSample
620 accession SAMN33422834, isolated in 2005 from a rectal swab of a dog in the United
621 Arab Emirates; see "Identification of $bla_{\text{CTX-M–8}}$+ and other ST548 genomes in the
622 AllTheBacteria dataset" in the Methods above). Overall, this suggests that the $bla_{\text{CTX-M-8}}$+
623 IncI plasmid was potentially transferred to ST548 from *E. coli*.

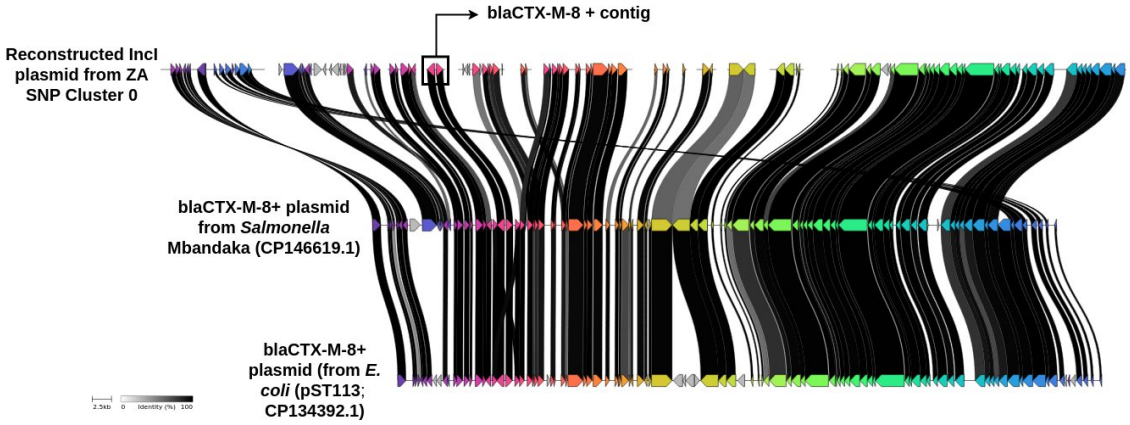


Fig 6. *bla*$_{CTX-M-8}$-harboring contigs from ZA SNP cluster 0 are highly similar to an IncI1 plasmid from *S.* Mbandaka and *E. coli*.
Plot showing the synteny and relatedness of *bla*$_{CTX-M-8}$-harboring IncI1 plasmids detected in genomes sequenced here (i.e., ZA SNP cluster 0) and the most closely related plasmids, detected in *S.* Mbandaka (CP146619.1) and *E. coli* (CP134392.1). The bands connecting each gene (arrows) indicate amino acid sequence identity, with darker bands indicating higher identities. Sequences with <50% identity are not connected. The color of each gene represents the functional annotation group.

633 **DISCUSSION**

634 **Port-of-entry genomic surveillance provides insight into the global dissemination**
635 **of pathogens via agro-food trade**

636 Bacterial pathogens endemic to a given geographical region can be exposed to unique,
637 local selection pressures, which shape their evolution and population structure (e.g.,
638 exposure to antimicrobials in agricultural, environmental, or clinical settings, dictated by
639 local regulations or the lack thereof) [76–79]. For foodborne zoonotic pathogens like
640 *Salmonella enterica*, the global agro-food trade serves as a mechanism by which endemic
641 pathogens can be disseminated to other world regions, potentially infecting humans
642 and/or animals [4, 5, 80–84]. As such, many countries routinely test agro-food imports for
643 the presence of pathogens, with the goal of preventing contaminated imports from
644 entering domestic food systems [84–86]. In South Africa, under the Meat Safety Act, 2000
645 (Act No.40 of 2000) and related Veterinary Procedural Notifice (VPN), microbiological
646 testing is required for compliance monitoring of both domestic and imported meat
647 products. Upon arrival and during storage, random samples are collected from selected
648 consignment packages and sent to approved laboratories for bacteriological testing,
649 including *Salmonella* detection [24]. Since 2018, VPN 56 mandates that all *Salmonella*
650 strains isolated from meat samples in South Africa be forwarded to the ARC: OVR-
651 General Bacteriology Laboratory for storage and molecular typing monitoring [87].

652

653 Here, we used WGS to characterize 36 *S.* Minnesota ST548 strains isolated in South
654 Africa, from imported poultry of Brazilian origin. In addition to querying ST548 strains from
655 poultry meat sold in supermarkets, our sampling efforts included ST548 strains isolated
656 from poultry meat consignments held at port-of-entry. Notably, our WGS data showed
657 that *S*. Minnesota ST548 strains detected in South African supermarkets were closely
658 related to strains isolated in poultry meat exported from Brazil to the UK. All 36 of our
659 isolates were confined to a well-supported clade within the ST548 phylogeny (i.e., the
660 "high-AMR group"), which contained at least eleven other isolates known to be associated
661 with Brazilian poultry imports. These eleven isolates were reported in a previous study
662 from the UK, in which *S.* Minnesota genomes from UK poultry meat imports of Brazilian
663 origin formed a monophyletic clade with genomes from Brazil [4]. Further, seven of our
664 genomes were highly similar to genomes from the UK (≤10 core SNPs; **Table S2**),
665 indicating that some (but not all) European and African imports may share a common
666 source. We also found a set of 11 South African food-associated ST548 genomes from
667 ATB, which were 23-65 core SNPs apart from at least one of our SEPI genomes (**Table**
668 **S3**).

669 Overall, currently available WGS data suggests that several *S.* Minnesota ST548 lineages
670 have been introduced into South Africa through Brazilian poultry imports. The results
671 showcase the value of rich geographic metadata in pathogen surveillance studies,
672 particularly metadata conveying the provenance of imported foods. As shown here and
673 elsewhere [21, 22, 24, 88, 89], foodborne pathogens isolated at POE, from a known

674 exporter, are particularly useful, as they can capture pathogen migration events between
675 geographic regions with known directionality.

676

677 **Existing publicly available WGS data cannot link *S.* Minnesota ST548 from Brazilian**
678 **poultry imports to South African human clinical cases**

679 Some importers have expressed concern regarding *Salmonella* in Brazilian meat and
680 poultry [4, 25, 90, 91]. However, the increasing prevalence of *S.* Minnesota in poultry may
681 not necessarily translate to an increase in human salmonellosis cases. For example, in
682 the UK, there was no observable increase in human clinical cases caused by *S.*
683 Minnesota, despite the increasing prevalence of *S.* Minnesota in Brazilian poultry, and
684 the few observed human clinical cases could largely be explained by recent international
685 travel [4]. The European Centre for Disease Prevention and Control (ECDC) infectious
686 disease surveillance atlas shows only 354 total reported cases of human salmonellosis
687 caused by *S*. Minnesota from 2007-2023 in the EU (UK not included after 2019). Another
688 common Brazilian poultry-associated serotype, *S.* Heidelberg, showed 1,790 cases in
689 the same time period, though with a decreasing trend. These figures are dwarfed by
690 serotype *S.* Enteritidis, which showed 33,088 cases in the EU in 2023 alone
691 (https://atlas.ecdc.europa.eu/public/index.aspx). Overall, these data suggest that *S.*
692 Minnesota may not be a significant threat to human health compared to other *Salmonella*
693 serotypes in Europe.

694 Here, we found that all *S.* Minnesota strains isolated in South Africa are relatively closely
695 related to food/poultry associated strains from Brazil or the UK (<65 core SNPs; **Fig S2A,**
696 **Table S2, Table S3**). Therefore, it is possible that Brazilian *S.* Minnesota ST548 also
697 does not impose a significant disease burden on South African consumers, as evidence
698 indicates for the UK and EU [4]. However, a lack of links to clinical cases could rather be
699 the result of data gaps. Despite imposing a disproportionately high burden of illness on
700 sub-Saharan Africa [92, 93], *Salmonella* WGS efforts are largely concentrated in world
701 regions where the burden of salmonellosis is lower [92, 94, 95]. For example, of the 228
702 Enterobase *Salmonella* ST548 genomes included in our dataset, 188 (~82%) originated
703 from strains isolated in the US or the UK. While African-led pathogen surveillance efforts
704 are producing an unprecedented amount of data, African WGS data generators can still
705 face major hurdles when sharing, publishing, and/or disseminating pathogen WGS
706 (meta)data [96]. Thus, the lack of observed links between imported *S.* Minnesota and
707 human clinical cases in South Africa could be due to the lack of publicly available WGS
708 data from *S.* Minnesota strains in South Africa (e.g., from human clinical cases, as well
709 as domestic animals, foods, and environmental sources). Overall, the UK experience with
710 *S.* Minnesota, both in terms of surveillance and disease burden, cannot be applied directly
711 to South Africa, and our study highlights the importance of local pathogen surveillance
712 initiatives, beyond those conducted in Europe and North America.

713

714 **Multiple *bla*$_{CTX-M-8}$-harboring *S.* Minnesota ST548 lineages have been disseminated**
715 **via international poultry trade**

716 Antimicrobial use in livestock and food production environments can select for AMR
717 foodborne zoonotic pathogens, which can be subsequently disseminated via international
718 trade [97–100]. Consequently, farm management and antimicrobial use practices in a
719 single country, region, or even farm can impact human and animal health across the
720 globe [100, 101]. Here, we identified a closely related *S.* Minnesota ST548 lineage (i.e.,
721 ZA SNP cluster 0, differing by 0 core SNPs), which had acquired *bla*$_{CTX-M-8}$. *bla*$_{CTX-M-8}$ was
722 published as a novel CTX-M beta-lactamase in 2000 [102], when it was identified in three
723 strains of three species (*Enterobacter cloacae*, *Enterobacter aerogenes*, and *Citrobacter*
724 *amalonaticus*) isolated in 1996-1997 from intensive care unit patients in Rio de Janeiro,
725 Brazil. Since then, studies in Brazil and other South American countries have noted that
726 *bla*$_{CTX-M-8}$ appears to be relatively common among extended-spectrum beta-lactamase
727 (ESBL)-producing *E. coli* in the region [103, 104]. However, *bla*$_{CTX-M-8}$ has been detected
728 in numerous other countries, in a range of hosts (e.g., *E. coli*, *Shigella* spp., *Enterobacter*
729 spp., *Klebsiella* spp., *Salmonella enterica*), isolated not only from human clinical cases,
730 but apparently healthy humans (e.g., working as food handlers), animals/animal products
731 (e.g., chicken meat, chickens, cattle), wastewater, and wildlife (e.g., gulls, wild birds) as
732 well [103–115]. Most notably within the context of this study, *bla*$_{CTX-M-8}$ has previously
733 been detected in *E. coli* isolated from Brazilian poultry products imported into Japan [105,
734 116], indicating that this AMR gene has previously been disseminated intercontinentally
735 via Brazilian poultry.

736 In addition to ZA SNP cluster 0, we identified an additional 16 *bla*$_{CTX-M-8}$-harboring *S.*
737 Minnesota ST548 genomes, isolated between 2017-2022 (three samples had unreported
738 isolation dates) from 4 different countries (i.e., Brazil, Portugal, South Africa and the UK;
739 two samples had unreported isolation locations). Notably, two of these *bla*$_{CTX-M-8}$-
740 harboring *S.* Minnesota ST548 genomes (NCBI BioSample accessions SAMN30444884,
741 SAMEA12320398), one from Brazil and one from the UK, were 91 core SNPs apart
742 and >290 core SNPs apart from ZA SNP cluster 0. (**Fig S9**). This indicates that multiple
743 independent acquisition events are responsible for the emergence of *bla*$_{CTX-M-8}$ in ST548.
744 Previous studies have identified *bla*$_{CTX-M-8}$ on plasmids, including IncI1 plasmids harbored
745 by *E. coli* [103, 105, 113]. Notably, *bla*$_{CTX-M-8}$-harboring IncI1 plasmids have been
746 detected across multiple *E. coli* STs [103, 105, 113], indicating that multiple *E. coli*
747 lineages have acquired this AMR gene via IncI1 plasmids. Further, a *bla*$_{CTX-M-8}$-harboring
748 IncI1 plasmid was recently described in a *Salmonella* Mbandaka strain (isolated in 2022
749 from a broiler farm environment in Poland) [69].

750 Overall, *bla*$_{CTX-M-8}$ can be acquired within and across species boundaries via IncI1
751 plasmids. Based on the genomic analyses conducted here, we hypothesize that ZA SNP
752 cluster 0 acquired *bla*$_{CTX-M-8}$ recently, within Brazil, via an *E. coli* IncI1 plasmid, and was
753 subsequently disseminated to South Africa via imported poultry. However, whether *bla*$_{CTX-}$
754 $_{M-8}$-harboring *S.* Minnesota imposes a significant burden on humans and animals in South
755 Africa remains a mystery. Future WGS efforts targeting *S.* Minnesota from human and
756 non-human sources (e.g., poultry flocks, farm environments, poultry products) in South
757 Africa are needed. We anticipate that long-read sequencing methods in particular will

758 provide insight into AMR dynamics within *S.* Minnesota. This, combined with continued
759 surveillance of imported foods at port-of-entry, will allow for improved risk evaluation of
760 imported foods in South Africa.

761

762 **AUTHOR STATEMENTS**

763 **Conflicts of interest**

764 The authors declare that there are no conflicts of interest.

765 **Funding information**

766 V Raghuram, V Ramnath, HG, JB, and LMC were supported by the SciLifeLab &
767 Wallenberg Data Driven Life Science Program (grant: KAW 2020.0239), with additional
768 funding provided by the Swedish Research Council (grant: 2023-05212). Sequencing
769 was funded by the Department of Agriculture, Land Reform, and Rural Development
770 (DALRRD).

771 **Acknowledgments**

772 Computation was enabled by resources provided by (i) High Performance Computing
773 Center North (HPC2N; Umeå University, Umeå, Sweden), and (ii) the National Academic
774 Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish
775 Research Council through grant agreement no. 2022-06725. The bacterial isolates
776 sequenced in this study were supplied by the Agricultural Research Council –
777 Onderstepoort Veterinary Research.

Author contributions

V Raghuram performed all computational analyses with input from TM, V Ramnath, HG, JB, and LMC. IM and LMC conceptualized and funded the study. V Raghuram, LMC, KM, and IM co-wrote the manuscript with input from all co-authors.

REFERENCES

1. **WHO**. *Salmonella* (non-typhoidal). https://www.who.int/news-room/fact-sheets/detail/salmonella-(non-typhoidal) (accessed 16 February 2025).

2. **Havelaar AH, Kirk MD, Torgerson PR, Gibb HJ, Hald T, *et al.*** World Health Organization Global Estimates and Regional Comparisons of the Burden of Foodborne Disease in 2010. *PLOS Medicine* 2015;12:e1001923.

3. **Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, *et al.*** The global burden of nontyphoidal *Salmonella* gastroenteritis. *Clin Infect Dis* 2010;50:882–889.

4. **Alikhan N-F, Moreno LZ, Castellanos LR, Chattaway MA, McLauchlin J, *et al.*** Dynamics of *Salmonella* enterica and antimicrobial resistance in the Brazilian poultry industry and global impacts on public health. *PLOS Genetics* 2022;18:e1010174.

5. **Li S, He Y, Mann DA, Deng X**. Global spread of *Salmonella* Enteritidis via centralized sourcing and international trade of poultry breeding stocks. *Nature Communications* 2021;12:1–12.

6. **CDC**. How *Salmonella* Infections Happen. *Salmonella Infection (Salmonellosis)*. https://www.cdc.gov/salmonella/spread/index.html (2024, accessed 16 February 2025).

7. **USDA-FSIS**. *Salmonella* By The Numbers. https://www.fsis.usda.gov/inspection/inspection-programs/inspection-poultry-products/reducing-salmonella-poultry/salmonella.

8. **Painter JA, Hoekstra RM, Ayers T, Tauxe RV, Braden CR, *et al.*** Attribution of foodborne illnesses, hospitalizations, and deaths to food commodities by using outbreak data, United States, 1998-2008. *Emerg Infect Dis* 2013;19:407–415.

9. **CDC**. Attribution of Foodborne Illness: Findings. https://www.cdc.gov/foodborneburden/attribution/attribution-1998-2008.html (2019, accessed 16 February 2025).

10. **Shaji S, Selvaraj RK, Shanmugasundaram R**. *Salmonella* Infection in Poultry: A Review on the Pathogen and Control Strategies. *Microorganisms* 2023;11:2814.

11. **Ferrari RG, Rosario DKA, Cunha-Neto A, Mano SB, Figueiredo EES, *et al.*** Worldwide Epidemiology of Serovars in Animal-Based Foods: a Meta-analysis. *Appl Environ Microbiol*;85. Epub ahead of print 15 July 2019. DOI: 10.1128/AEM.00591-19.

12. **Shah DH, Paul NC, Sischo WC, Crespo R, Guard J**. Population dynamics and antimicrobial resistance of the most prevalent poultry-associated *Salmonella* serotypes. *Poultry Science* 2017;96:687–702.

13. **Huang J, Alzahrani KO, Zhou G, Alsalman SA, Alsufyani AT, *et al.*** Genomic survey of multidrug resistant *Salmonella enterica* serovar Minnesota clones in chicken products. *npj Antimicrobials and Resistance* 2025;3:1–11.

14. **Kipper D, Carroll LM, Mascitti AK, Streck AF, Fonseca ASK, *et al.*** Genomic Characterization of Minnesota Clonal Lineages Associated with Poultry Production in Brazil. *Animals (Basel)*;10. Epub ahead of print 5 November 2020. DOI: 10.3390/ani10112043.

15. **Edwards PR, Bruner DW**. Two new *Salmonella* types isolated from fowls. *J Hyg (Lond)* 1938;38:716–720.

16. **Voss-Rech D, Vaz CSL, Alves L, Coldebella A, Leão JA, *et al.*** A temporal study of *Salmonella enterica* serotypes from broiler farms in Brazil. *Poult Sci* 2015;94:433–441.

17. **dos Santos AMP, Panzenhagen P, Ferrari RG, Portes AB, de Jesus AC, *et al.*** Genomic characterization of a clonal emergent *Salmonella* Minnesota lineage in Brazil reveals the presence of a novel megaplasmid of resistance and virulence. *Applied and Environmental Microbiology*. Epub ahead of print 20 November 2024. DOI: 10.1128/aem.01579-24.

18. **Tavechio AT, Fernandes SA, Neves BC, Dias AM, Irino K**. Changing patterns of *Salmonella* serovars: increase of Salmonella enteritidis in São Paulo, Brazil. *Rev Inst Med Trop Sao Paulo* 1996;38:315–322.

19. **Hernandez T, Sierra A, Rodriguez-Alvarez C, Torres A, Arevalo MP, *et al.*** *Salmonella enterica* serotypes isolated from imported frozen chicken meat in the Canary islands. *J Food Prot* 2005;68:2702–2706.

20. **Monte DF, Lincopan N, Berman H, Cerdeira L, Keelara S, *et al.*** Genomic Features of High-Priority *Salmonella enterica* Serovars Circulating in the Food Production Chain, Brazil, 2000–2016. *Scientific Reports* 2019;9:1–12.

21. **Silveira L, Nunes A, Pista Â, Isidro J, Belo Correia C, *et al.*** Characterization of Multidrug-Resistant Isolates of Serovars Heidelberg and Minnesota from Fresh Poultry Meat Imported to Portugal. *Microb Drug Resist* 2021;27:87–98.

22. **Campos J, Mourão J, Silveira L, Saraiva M, Belo Correia C, et al**. Imported poultry meat as a source of extended-spectrum cephalosporin-resistant CMY-2-producing *Salmonella* Heidelberg and *Salmonella* Minnesota in the European Union, 2014–2015. *International Journal of Antimicrobial Agents* 2018;51:151–154.

23. **Brazilian Association of Animal Protein**. Annual Report 2024. https://abpa-br.org/wp-content/uploads/2024/04/RA_2024_ABPA_ingles_avicultura.pdf (accessed 17 February 2025).

24. **Ndobeni TN, Magwedere K, Qekwana DN**. Prevalence of species and factors associated with contamination of mechanically recovered poultry meat imported into South Africa, 2016-2017. *Vet World* 2023;16:2236–2243.

25. **USDA**. https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=South+Africa+Suspends+Imports+from+Three+Brazilian+Establishments_Pretoria_South+Africa+-+Republic+of_3-23-2018.pdf (accessed 17 February 2025).

857 26. **ISO** 6579-1:2017. *ISO*. https://www.iso.org/standard/56712.html (2022, accessed
858 17 February 2025).

859 27. **Chen S, Zhou Y, Chen Y, Gu J**. fastp: an ultra-fast all-in-one FASTQ
860 preprocessor. *Bioinformatics* 2018;34:i884–i890.

861 28. **Souvorov A, Agarwala R, Lipman DJ**. SKESA: strategic k-mer extension for
862 scrupulous assemblies. *Genome Biology* 2018;19:1–13.

863 29. **Gurevich A, Saveliev V, Vyahhi N, Tesler G**. QUAST: quality assessment tool
864 for genome assemblies. *Bioinformatics* 2013;29:1072–1075.

865 30. **Kitts PA, Church DM, Thibaud-Nissen F, Choi J, Hem V, *et al.*** Assembly: a
866 resource for assembled genomes at NCBI. *Nucleic Acids Res* 2016;44:D73–80.

867 31. **Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW**. CheckM:
868 assessing the quality of microbial genomes recovered from isolates, single cells, and
869 metagenomes. *Genome Res* 2015;25:1043–1055.

870 32. **Yoshida CE, Kruczkiewicz P, Laing CR, Lingohr EJ, Gannon VPJ, *et al.*** The
871 *Salmonella* In Silico Typing Resource (SISTR): An Open Web-Accessible Tool for Rapidly
872 Typing and Subtyping Draft Salmonella Genome Assemblies. *PLoS One*
873 2016;11:e0147101.

874 33. **Zhang S, den Bakker HC, Li S, Chen J, Dinsmore BA, *et al.*** SeqSero2: Rapid
875 and Improved Serotype Determination Using Whole-Genome Sequencing Data. *Appl*
876 *Environ Microbiol*;85. Epub ahead of print 1 December 2019. DOI: 10.1128/AEM.01746-19.

877 34. **Zhou Z, Alikhan N-F, Mohamed K, Fan Y, Agama Study Group, *et al.*** The
878 EnteroBase user's guide, with case studies on transmissions, phylogeny, and core genomic
879 diversity. *Genome Res* 2020;30:138–152.

880 35. **Dyer NP, Päuker B, Baxter L, Gupta A, Bunk B, *et al.*** EnteroBase in 2025:
881 exploring the genomic epidemiology of bacterial pathogens. *Nucleic Acids Res*
882 2025;53:D757–D762.

883 36. **Barrett T, Clark K, Gevorgyan R, Gorelenkov V, Gribov E, *et al.*** BioProject
884 and BioSample databases at NCBI: facilitating capture and organization of metadata.
885 *Nucleic Acids Res* 2012;40:D57–63.

886 37. **Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, *et al.*** Mash:
887 fast genome and metagenome distance estimation using MinHash. *Genome Biology*
888 2016;17:1–14.

889 38. **Katz LS, Griswold T, Morrison SS, Caravas JA, Zhang S, *et al.*** Mashtree: a
890 rapid comparison of whole genome sequence files. *Journal of Open Source Software*
891 2019;4:1762.

892 39. **Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, *et al.*** Validating the
893 AMRFinder Tool and Resistance Gene Database by Using Antimicrobial Resistance
894 Genotype-Phenotype Correlations in a Collection of Isolates. *Antimicrob Agents*

895 *Chemother*;63. Epub ahead of print November 2019. DOI: 10.1128/AAC.00483-19.

896 40. **Feldgarden M, Brover V, Gonzalez-Escalona N, Frye JG, Haendiges J, *et al.***
897 AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic
898 links among antimicrobial resistance, stress response, and virulence. *Sci Rep*
899 2021;11:12728.

900 41. **Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, *et al.***
901 In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus
902 sequence typing. *Antimicrob Agents Chemother* 2014;58:3895–3903.

903 42. **Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, *et al.*** Bakta:
904 rapid and standardized annotation of bacterial genomes via alignment-free sequence
905 identification. *Microb Genom*;7. Epub ahead of print November 2021. DOI:
906 10.1099/mgen.0.000685.

907 43. **Tonkin-Hill G, MacAlasdair N, Ruis C, Weimann A, Horesh G, *et al.***
908 Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biology*
909 2020;21:1–21.

910 44. **McInnes L, Healy J, Melville J**. UMAP: Uniform Manifold Approximation and
911 Projection for Dimension Reduction. http://arxiv.org/abs/1802.03426 (2018, accessed 17
912 February 2025).

913 45. **Wickham H**. *ggplot2*. Springer International Publishing.

914 46. **Shaw J, Yu YW**. Fast and robust metagenomic sequence comparison through
915 sparse chaining with skani. *Nature Methods* 2023;20:1661–1665.

916 47. **Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, *et al.*** Rapid
917 phylogenetic analysis of large samples of recombinant bacterial whole genome sequences
918 using Gubbins. *Nucleic Acids Res* 2015;43:e15.

919 48. **Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, *et al.*** : rapid efficient
920 extraction of SNPs from multi-FASTA alignments. *Microb Genom* 2016;2:e000056.

921 49. **Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, *et al.*** IQ-
922 TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era.
923 *Mol Biol Evol* 2020;37:1530–1534.

924 50. **Tavaré S**. Some probabilistic and statistical problems in the analysis of DNA
925 sequences. https://api.semanticscholar.org/CorpusID:82212051 (1986, accessed 23
926 February 2025).

927 51. **Yang Z**. A space-time process model for the evolution of DNA sequences.
928 *Genetics* 1995;139:993–1005.

929 52. **Soubrier J, Steel M, Lee MSY, Der Sarkissian C, Guindon S, *et al.*** The
930 influence of rate heterogeneity among sites on the time dependence of molecular rates. *Mol*
931 *Biol Evol* 2012;29:3345–3358.

932 53. **Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS**. UFBoot2:
933 Improving the Ultrafast Bootstrap Approximation. *Mol Biol Evol* 2018;35:518–522.

934 54. **Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y**. ggtree: an r package for
935 visualization and annotation of phylogenetic trees with their covariates and other associated
936 data. *Methods in Ecology and Evolution* 2017;8:28–36.

937 55. **Revell LJ**. phytools 2.0: an updated R ecosystem for phylogenetic comparative
938 methods (and other things). *PeerJ* 2024;12:e16505.

939 56. **To T-H, Jung M, Lycett S, Gascuel O**. Fast Dating Using Least-Squares Criteria
940 and Algorithms. *Syst Biol* 2016;65:82–97.

941 57. **Rambaut A, Lam TT, Max Carvalho L, Pybus OG**. Exploring the temporal
942 structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol*
943 2016;2:vew007.

944 58. **Paradis E, Schliep K**. ape 5.0: an environment for modern phylogenetics and
945 evolutionary analyses in R. *Bioinformatics* 2019;35:526–528.

946 59. **Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, *et al.***
947 BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS*
948 *Comput Biol* 2019;15:e1006650.

949 60. **Drummond AJ, Rambaut A, Shapiro B, Pybus OG**. Bayesian coalescent
950 inference of past population dynamics from molecular sequences. *Mol Biol Evol*
951 2005;22:1185–1192.

952 61. **Bouckaert RR, Drummond AJ**. bModelTest: Bayesian phylogenetic site model
953 averaging and model comparison. *BMC Evolutionary Biology* 2017;17:1–11.

954 62. **Duchêne S, Duchêne D, Holmes EC, Ho SYW**. The Performance of the Date-
955 Randomization Test in Phylogenetic Analyses of Time-Structured Virus Data. *Mol Biol Evol*
956 2015;32:1895–1906.

957 63. **Rieux A, Khatchikian CE**. tipdatingbeast: an r package to assist the
958 implementation of phylogenetic tip-dating tests using beast. *Mol Ecol Resour*
959 2017;17:608–613.

960 64. **Camargo AP, Roux S, Schulz F, Babinski M, Xu Y, *et al.*** Identification of
961 mobile genetic elements with geNomad. *Nature Biotechnology* 2023;42:1303–1312.

962 65. **Schmartz GP, Hartung A, Hirsch P, Kern F, Fehlmann T, *et al.*** PLSDB:
963 advancing a comprehensive database of bacterial plasmids. *Nucleic Acids Res*
964 2022;50:D273–D278.

965 66. **Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, *et al.*** BLAST+:
966 architecture and applications. *BMC Bioinformatics* 2009;10:1–9.

967 67. **Shen W, Sipos B, Zhao L**. SeqKit2: A Swiss army knife for sequence and
968 alignment processing. *iMeta* 2024;3:e191.

969 68. **Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, *et al.*** BLAST: a
970 more efficient report with usability improvements. *Nucleic Acids Res* 2013;41:W29–33.

971 69. **Zając M, Skarżyńska M, Lalak A, Iwan E, Wasyl D**. Detection of *Salmonella*
972 Mbandaka Carrying the blaCTX-M-8 Gene Located on IncI1 Plasmid Isolated from a Broiler
973 Flock Environment. *Pathogens* 2024;13:723.

974 70. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, *et al.*** SPAdes: a
975 new genome assembly algorithm and its applications to single-cell sequencing. *J Comput*
976 *Biol* 2012;19:455–477.

977 71. **Gilchrist CLM, Chooi Y-H**. clinker & clustermap.js: automatic generation of gene
978 cluster comparison figures. *Bioinformatics* 2021;37:2473–2475.

979 72. **Hunt M, Lima L, Anderson D, Hawkey J, Shen W, *et al.*** AllTheBacteria - all
980 bacterial genomes assembled, available and searchable. *bioRxiv* 2024;2024.03.08.584059.

981 73. **Kans J**. Entrez Direct: E-utilities on the Unix Command Line. In: *Entrez*
982 *Programming Utilities Help [Internet]*. National Center for Biotechnology Information (US);
983 2025.

984 74. **Heller R, Chikhi L, Siegismund HR**. The confounding effect of population
985 structure on Bayesian skyline plot inferences of demographic history. *PLoS One*
986 2013;8:e62992.

987 75. **Lapierre M, Blin C, Lambert A, Achaz G, Rocha EPC**. The Impact of Selection,
988 Gene Conversion, and Biased Sampling on the Assessment of Microbial Demography. *Mol*
989 *Biol Evol* 2016;33:1711–1725.

990 76. **Fenske GJ, Thachil A, McDonough PL, Glaser A, Scaria J**. Geography
991 Shapes the Population Genomics of *Salmonella enterica* Dublin. *Genome Biol Evol*
992 2019;11:2220–2231.

993 77. **Frost I, Van Boeckel TP, Pires J, Craig J, Laxminarayan R**. Global geographic
994 trends in antimicrobial resistance: the role of international travel. *J Travel Med*;26. Epub
995 ahead of print 23 December 2019. DOI: 10.1093/jtm/taz036.

996 78. **Shaw LP, Chau KK, Kavanagh J, AbuOun M, Stubberfield E, *et al.*** Niche and
997 local geography shape the pangenome of wastewater- and livestock-associated
998 Enterobacteriaceae. *Science Advances*. Epub ahead of print April 2021. DOI:
999 10.1126/sciadv.abe3868.

1000 79. **Jespersen ML, Munk P, Johansen J, Kaas RS, Webel H, *et al.*** Global within-
1001 species phylogenetics of sewage microbes suggest that local adaptation shapes
1002 geographical bacterial clustering. *Communications Biology* 2023;6:1–9.

1003 80. **Li H, Wu Y, Feng D, Jiang Q, Li S, *et al.*** Centralized industrialization of pork in
1004 Europe and America contributes to the global spread of *Salmonella enterica*. *Nature Food*
1005 2024;5:413–422.

1006 81. **Rohr JR, Barrett CB, Civitello DJ, Craft ME, Delius B, *et al.*** Emerging human

1007 infectious diseases and the links to global food production. *Nature Sustainability*
1008 2019;2:445–456.

1009 82. **Antunes P, Novais C, Peixe L**. Food-to-Humans Bacterial Transmission.
1010 *Microbiol Spectr*;8. Epub ahead of print January 2020. DOI: 10.1128/microbiolspec.MTBP-
1011 0019-2016.

1012 83. **Baker RE, Mahmud AS, Miller IF, Rajeev M, Rasambainarivo F, *et al.***
1013 Infectious disease in an era of global change. *Nature Reviews Microbiology*
1014 2021;20:193–205.

1015 84. **Jansen W, Müller A, Grabowski NT, Kehrenberg C, Muylkens B, *et al.***
1016 Foodborne diseases do not respect borders: Zoonotic pathogens and antimicrobial
1017 resistant bacteria in food products of animal origin illegally imported into the European
1018 Union. *Vet J* 2019;244:75–82.

1019 85. **Human Foods Program**. FDA Strategy for the Safety of Imported Food. *U.S.*
1020 *Food and Drug Administration*. https://www.fda.gov/food/importing-food-products-united-
1021 states/fda-strategy-safety-imported-food (2024, accessed 19 February 2025).

1022 86. **Manyi-Loh CE, Lues R**. A South African Perspective on the Microbiological and
1023 Chemical Quality of Meat: Plausible Public Health Implications. *Microorganisms*
1024 2023;11:2484.

1025 87. **DALRRD**. Veterinary Procedural Notice 56/2019-0.
1026 https://old.dalrrd.
1027 gov.za/vetweb/VPN%20&%20SO
1028 P/VPN56%202019%20-%20Requirements%20for%20food%20saf
1029 ety%20testing%20laboratories.pdf (accessed 16 March 2025).

1030 88. **Hoffmann M, Luo Y, Monday SR, Gonzalez-Escalona N, Ottesen AR, *et al.***
1031 Tracing Origins of the *Salmonella* Bareilly Strain Causing a Food-borne Outbreak in the
1032 United States. *J Infect Dis* 2016;213:502–508.

1033 89. **Allard MW, Strain E, Melka D, Bunning K, Musser SM, *et al.*** Practical Value of
1034 Food Pathogen Traceability through Building a Whole-Genome Sequencing Network and
1035 Database. *J Clin Microbiol* 2016;54:1975–1983.

1036 90. **EU**. COMMISSION IMPLEMENTING REGULATION (EU) 2018/700 of 8 May
1037 2018 amending the lists of third country establishments from which imports of specified
1038 products of animal origin are permitted, regarding certain establishments from Brazil.
1039 https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R0700 (accessed
1040 19 February 2025).

1041 91. **Wasley A**. Brazil sent one million *Salmonella*-infected chickens to UK in two
1042 years. *The Guardian*, 3 July 2019.
1043 https://www.theguardian.com/environment/2019/jul/03/brazil-one-million-salmonella-
1044 infected-chickens-uk (3 July 2019, accessed 19 February 2025).

1045 92. **Kirk MD, Pires SM, Black RE, Caipo M, Crump JA, *et al.*** World Health
1046 Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne

1047 Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. *PLoS Med*
1048 2015;12:e1001921.

1049 93. **GBD 2017 Non-Typhoidal *Salmonella* Invasive Disease Collaborators**. The
1050 global burden of non-typhoidal *Salmonella* invasive disease: a systematic analysis for the
1051 Global Burden of Disease Study 2017. *Lancet Infect Dis* 2019;19:1312–1324.

1052 94. **Inzaule SC, Tessema SK, Kebede Y, Ogwell Ouma AE, Nkengasong JN**.
1053 Genomic-informed pathogen surveillance in Africa: opportunities and challenges. *Lancet*
1054 *Infect Dis* 2021;21:e281–e289.

1055 95. **Carroll LM, Pierneef R, Mathole M, Matle I**. Genomic Characterization of
1056 Endemic and Ecdemic Non-typhoidal *Salmonella* enterica Lineages Circulating Among
1057 Animals and Animal Products in South Africa. *Front Microbiol* 2021;12:748611.

1058 96. **Mboowa G, Tessema SK, Christoffels A, Ndembi N, Kebede Tebeje Y, *et al.***
1059 Africa in the era of pathogen genomics: Unlocking data barriers. *Cell* 2024;187:5146–5150.

1060 97. **Van Boeckel TP, Glennon EE, Chen D, Gilbert M, Robinson TP, *et al.***
1061 Reducing antimicrobial use in food animals. *Science* 2017;357:1350–1352.

1062 98. **He Y, Yuan Q, Mathieu J, Stadler L, Senehi N, *et al.*** Antibiotic resistance
1063 genes from livestock waste: occurrence, dissemination, and treatment. *npj Clean Water*
1064 2020;3:1–11.

1065 99. **Larsson DGJ, Flach C-F**. Antibiotic resistance in the environment. *Nature*
1066 *Reviews Microbiology* 2021;20:257–269.

1067 100. **Jung D, Morrison BJ, Rubin JE**. A review of antimicrobial resistance in
1068 imported foods. *Can J Microbiol* 2021;1–15.

1069 101. **Olesen SW, Lipsitch M, Grad YH**. The role of 'spillover' in antibiotic resistance.
1070 *Proc Natl Acad Sci U S A* 2020;117:29063–29068.

1071 102. **Bonnet R, Sampaio JL, Labia R, De Champs C, Sirot D, *et al.*** A novel CTX-M
1072 beta-lactamase (CTX-M-8) in cefotaxime-resistant Enterobacteriaceae isolated in Brazil.
1073 *Antimicrob Agents Chemother* 2000;44:1936–1942.

1074 103. **Palmeira JD, Haenni M, Metayer V, Madec J-Y, Ferreira HMN**. Epidemic
1075 spread of IncI1/pST113 plasmid carrying the Extended-Spectrum Beta-Lactamase (ESBL)
1076 bla gene in Escherichia coli of Brazilian cattle. *Vet Microbiol* 2020;243:108629.

1077 104. **Naseer U, Sundsfjord A**. The CTX-M conundrum: dissemination of plasmids
1078 and Escherichia coli clones. *Microb Drug Resist* 2011;17:83–97.

1079 105. **Norizuki C, Wachino J-I, Suzuki M, Kawamura K, Nagano N, *et al.*** Specific
1080 /IncI1 Plasmid Transfer among Genetically Diverse Escherichia coli Isolates between
1081 Humans and Chickens. *Antimicrob Agents Chemother*;61. Epub ahead of print June 2017.
1082 DOI: 10.1128/AAC.00663-17.

1083 106.**Valenzuela X, Hedman H, Villagomez A, Cardenas P, Eisenberg JNS, et al.**

1084 Distribution of blaCTX-M-gene variants in E. coli from different origins in Ecuador. *Medicine*
1085 *in Microecology* 2023;18:100092.

1086 107. **Gundran RS, Cardenio PA, Villanueva MA, Sison FB, Benigno CC, *et al.***
1087 Prevalence and distribution of blaCTX-M, blaSHV, blaTEM genes in extended- spectrum β-
1088 lactamase- producing E. coli isolates from broiler farms in the Philippines. *BMC Veterinary*
1089 *Research* 2019;15:1–8.

1090 108. **Eller C, Leistner R, Guerra B, Fischer J, Wendt C, *et al.*** Emergence of
1091 extended-spectrum β-lactamase (ESBL) CTX-M-8 in Germany. *J Antimicrob Chemother*
1092 2014;69:562–564.

1093 109. **Lewis JS 2nd, Herrera M, Wickes B, Patterson JE, Jorgensen JH**. First report
1094 of the emergence of CTX-M-type extended-spectrum beta-lactamases (ESBLs) as the
1095 predominant ESBL isolated in a U.S. health care system. *Antimicrob Agents Chemother*
1096 2007;51:4015–4021.

1097 110. **Stedt J, Bonnedahl J, Hernandez J, Waldenström J, McMahon BJ, *et al.***
1098 Carriage of CTX-M type extended spectrum β-lactamases (ESBLs) in gulls across Europe.
1099 *Acta Veterinaria Scandinavica* 2015;57:1–8.

1100 111. **Kamijo K, Abe Y, Kagami T, Ugajin K, Mikawa T, *et al.*** Bacteriuria With CTX-
1101 M-8 Extended-Spectrum β-Lactamase-Producing in a Patient With Incomplete Kawasaki
1102 Disease. *Glob Pediatr Health* 2019;6:2333794X18821944.

1103 112. **Seribelli AA, Ballaben AS, Ferreira JC, Medeiros MIC, Darini ALC, *et al.*** IncU
1104 plasmid harbouring bla in multidrug-resistant Shigella sonnei in Brazil. *J Glob Antimicrob*
1105 *Resist* 2018;14:99–100.

1106 113. **Dropa M, Lincopan N, Balsalobre LC, Oliveira DE, Moura RA, *et al.*** Genetic
1107 background of novel sequence types of CTX-M-8- and CTX-M-15-producing Escherichia
1108 coli and Klebsiella pneumoniae from public wastewater treatment plants in São Paulo,
1109 Brazil. *Environ Sci Pollut Res Int* 2016;23:4953–4958.

1110 114. **Fernandes SA, Camargo CH, Francisco GR, Bueno MFC, Garcia DO, *et al.***
1111 Prevalence of Extended-Spectrum β-Lactamases CTX-M-8 and CTX-M-2-Producing
1112 *Salmonella* Serotypes from Clinical and Nonhuman Isolates in Brazil. *Microb Drug Resist*
1113 2017;23:580–589.

1114 115. **Ewbank AC, Fuentes-Castillo D, Sacristán C, Cardoso B, Esposito F, *et al.***
1115 Extended-spectrum β-lactamase (ESBL)-producing Escherichia coli survey in wild seabirds
1116 at a pristine atoll in the southern Atlantic Ocean, Brazil: First report of the O25b-ST131
1117 clone harboring bla. *Sci Total Environ* 2022;806:150539.

1118 116. **Nahar A, Awasthi SP, Hatanaka N, Okuno K, Hoang PH, *et al.*** Prevalence and
1119 characteristics of extended-spectrum β-lactamase-producing Escherichia coli in domestic
1120 and imported chicken meats in Japan. *J Vet Med Sci* 2018;80:510–517.

1121