



microLife

**Salmonella enterica serovar Typhimurium ST313 sublineage  
2.2 has emerged in Malawi with a characteristic gene  
expression signature and a fitness advantage**

Journal:	<i>microLife</i>
Manuscript ID	Draft
Manuscript Type:	Research article
Date Submitted by the Author:	n/a
Complete List of Authors:	<p>Kumwenda, Benjamin; Kamuzu University of Health Sciences, School of Life Sciences and Health Professions; University of Liverpool, Institute of Infection, Veterinary &amp; Ecological Sciences; Malawi-Liverpool-Wellcome Trust Clinical Research Programme, Blantyre</p> <p>Canals, Rocío; University of Liverpool, Institute of Infection, Veterinary &amp; Ecological Sciences</p> <p>Predeus, Alexander; University of Liverpool, Institute of Infection, Veterinary &amp; Ecological Sciences</p> <p>Zhu, Xiaojun; University of Liverpool, Institute of Infection, Veterinary &amp; Ecological Sciences</p> <p>Kroger, Carsten; Trinity College Dublin School of Genetics and Microbiology, Department of Microbiology; University of Liverpool, Institute of Infection, Veterinary &amp; Ecological Sciences</p> <p>Pulford, Caisey; University of Liverpool, Institute of Infection, Veterinary &amp; Ecological Sciences</p> <p>Wenner, Nicolas; University of Liverpool, Institute of Infection, Veterinary &amp; Ecological Sciences</p> <p>Lacharme-Lora, Lizeth ; University of Liverpool, Institute of Infection, Veterinary &amp; Ecological Sciences</p> <p>Li, Yan; University of Liverpool, Institute of Infection, Veterinary &amp; Ecological Sciences</p> <p>Owen, Siân; University of Liverpool, Institute of Infection, Veterinary &amp; Ecological Sciences</p> <p>Everett, Dean; Khalifa University, College of Medicine and Health Sciences</p> <p>Hokamp, Karsten; Trinity College Dublin School of Genetics and Microbiology, Department of Genetics</p> <p>Heyderman, Robert; University College London, Research Department of Infection, Division of Infection &amp; Immunity; Malawi-Liverpool-Wellcome Trust Clinical Research Programme, Blantyre</p> <p>Ashton, Philip M.; Malawi-Liverpool-Wellcome Trust Clinical Research Programme, Blantyre</p> <p>Gordon, Melita; University of Liverpool, Institute of Infection, Veterinary &amp; Ecological Sciences; Malawi-Liverpool-Wellcome Trust Clinical Research Programme, Blantyre</p> <p>Msefula, Chisomo; Kamuzu University of Health Sciences, School of Life Sciences and Health Professions; Malawi-Liverpool-Wellcome Trust Clinical Research Programme, Blantyre</p>

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

	Hinton, Jay; University of Liverpool, Institute of Infection, Veterinary & Ecological Sciences
Keywords:	transcriptomics, Salmonella, evolution, AMR, genomics, sublineage



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1 ***Salmonella enterica* serovar Typhimurium ST313 sublineage 2.2**  
2 **has emerged in Malawi with a characteristic gene expression**  
3 **signature and a fitness advantage**  
4

---

5  
6  
7 Benjamin Kumwenda<sup>1,2,3, #</sup>, Rocío Canals<sup>2, #</sup>, Alexander V. Predeus<sup>2, #</sup>, Xiaojun  
8 Zhu<sup>2</sup>, Carsten Kröger<sup>2,4</sup>, Caisey Pulford<sup>2</sup>, Nicolas Wenner<sup>2</sup>, Lizeth Lacharme  
9 Lora<sup>2</sup>, Yan Li<sup>2</sup>, Siân V. Owen<sup>2</sup>, Dean Everett<sup>5</sup>, Karsten Hokamp<sup>6</sup>, Robert S.  
10 Heyderman<sup>3,7</sup>, Philip M. Ashton<sup>3</sup>, Melita A. Gordon<sup>2,3</sup>, Chisomo L. Msefula<sup>1,3</sup>,  
11 Jay C. D. Hinton<sup>2\*</sup>.

12  
13 <sup>1</sup>School of Life Sciences and Allied Health Professions, Kamuzu University of Health  
14 Sciences Blantyre, Blantyre, MALAWI

15 <sup>2</sup>Institute of Infection, Veterinary & Ecological Sciences, University of Liverpool, Liverpool,  
16 UNITED KINGDOM,

17 <sup>3</sup>Malawi–Liverpool–Wellcome Programme, Blantyre, MALAWI.

18 <sup>4</sup>Moyne Institute of Preventive Medicine, School of Genetics and Microbiology, Trinity College  
19 Dublin, Dublin, IRE.

20 <sup>5</sup>College of Medicine and Health Sciences, Khalifa University, Abu Dhabi, UAE.

21 <sup>6</sup>Smurfit Institute of Genetics, School of Genetics and Microbiology, Trinity College Dublin,  
22 Dublin, IRE.

23 <sup>7</sup>Research Department of Infection, Division of Infection & Immunity, University College  
24 London, UK.

25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37 **Key words:** transcriptomics, comparative genomics, lineage evolution, gene expression,  
38 antibiotic resistance

39  
40  
41  
42  
43 #BK, RC and AP contributed equally to this work.

44  
45 \* **Corresponding author:** E-mail: [jay.hinton@liverpool.ac.uk](mailto:jay.hinton@liverpool.ac.uk); Tel: +441517954573; Institute  
46 of Infection, Veterinary & Ecological Sciences, University of Liverpool, Liverpool, United  
47 Kingdom.

48 **Abstract**

49 Invasive non-typhoidal *Salmonella* (iNTS) disease is a serious bloodstream infection that  
50 targets immune-compromised individuals, and causes significant mortality in sub-Saharan  
51 Africa. *Salmonella enterica* serovar Typhimurium ST313 causes the majority of iNTS in  
52 Malawi. We performed an intensive comparative genomic analysis of 608 *S. Typhimurium*  
53 ST313 isolates dating between 1996 and 2018 from Blantyre, Malawi. We discovered that  
54 following the arrival of the well-characterised *S. Typhimurium* ST313 lineage 2 in 1999, two  
55 multidrug-resistant variants emerged in Malawi in 2006 and 2008, designated sublineage 2.2  
56 and 2.3 respectively. The majority of *S. Typhimurium* isolates from human bloodstream  
57 infections in Malawi now belong to sublineage 2.2 or 2.3. To understand the emergence of the  
58 prevalent ST313 sublineage 2.2, we studied two representative strains, D23580 (lineage 2)  
59 and D37712 (sublineage 2.2). The chromosome of ST313 lineage 2 and sublineage 2.2 only  
60 differed by 29 SNPs/small indels and a 3kb deletion of a Gifsy-2 prophage region including  
61 the *sseI* pseudogene. Lineage 2 and sublineage 2.2 had distinctive plasmid profiles. The  
62 transcriptome was investigated in 15 infection-relevant *in vitro* conditions and within  
63 macrophages. During growth in physiological conditions that do not usually trigger *S.*  
64 *Typhimurium* SPI2 gene expression, the SPI2 genes of D37712 were transcriptionally active.  
65 We identified down-regulation of flagellar genes in D37712 compared with D23580. Following  
66 phenotypic confirmation of transcriptomic differences, we discovered that sublineage 2.2 had  
67 increased fitness compared with lineage 2 during mixed-growth in minimal media. We  
68 speculate that this competitive advantage is contributing to the emergence of sublineage 2.2  
69 in Malawi.

70



## 71 Introduction

72 Non-typhoidal *Salmonella* (NTS) is a key bacterial pathogen that threatens people across the  
73 world. Typhimurium and Enteritidis are the two serovars of *Salmonella enterica* responsible  
74 for the highest levels of self-limiting gastrointestinal disease in Europe, the USA and other  
75 high-income countries (Zhang *et al.*, 2003). In the industrialised world, NTS has been  
76 associated with intensive food production, animal husbandry, and global distribution systems  
77 (Majowicz *et al.*, 2010). The *S. Typhimurium* sequence types responsible for gastroenteritis  
78 globally include ST19, ST34 and monophasic 1,4,[5],12:i:- variants (Branchu *et al.*, 2018).  
79 The diarrhoeal NTS disease is termed dNTS, and is mainly foodborne, posing a significant  
80 burden to public health with approximately 153 million cases and 57,000 deaths per annum  
81 worldwide (Kirk *et al.*, 2015; Chirwa *et al.*, 2023).

82 In contrast, a lethal systemic disease called invasive non-typhoidal Salmonellosis (iNTS) has  
83 emerged in recent decades in low- and middle-income countries in sub-Saharan Africa. iNTS  
84 targets immunocompromised individuals such as adults with HIV, and children under five  
85 years of age with malaria, malnutrition or severe anaemia (Feasey *et al.*, 2012). In some  
86 countries of sub-Saharan Africa, *Salmonella* causes more cases of community-onset  
87 bloodstream infections than any other bacterial pathogen (Marchello *et al.*, 2019). In 2017,  
88 535,000 cases of iNTS disease were estimated worldwide, with about 80% of cases and  
89 77,000 deaths occurring in sub-Saharan Africa (Stanaway *et al.*, 2019)

90 Clinically, the treatment of iNTS is complicated by multi-drug (MDR) resistance which limits  
91 therapeutic options (Crump *et al.*, 2015). Widespread resistance of iNTS pathogens to first-  
92 line drugs such as chloramphenicol, ampicillin and cotrimoxazole has been seen in many  
93 countries (Kariuki *et al.*, 2006). This MDR phenotype may be one of the reasons the case  
94 fatality rate associated with iNTS is amongst the highest in comparison to any infectious  
95 disease (15%) (Marchello *et al.*, 2022). Resistance to second-line drugs such as ceftriaxone,  
96 ciprofloxacin and azithromycin has been reported in a few African countries (Tack *et al.*,  
97 2020). Clearly, the challenge posed by MDR *Salmonella* must be addressed urgently  
98 (Gilchrist and MacLennan, 2019).

99 The African iNTS epidemic is mainly caused by two *Salmonella* pathovariants, *S.*  
100 Typhimurium sequence type 313 (ST313) and specific clades of *S. Enteritidis* (Kingsley *et al.*,  
101 2009; Okoro *et al.*, 2012; Feasey *et al.*, 2016). *S. Typhimurium* ST313 is responsible for about  
102 two-thirds of clinical iNTS cases that have been reported in Africa (Gilchrist and MacLennan,  
103 2019).

104 It is not certain how these pathogens are transmitted, but there is increasing evidence from  
105 case-control studies that ST313 strains are human-associated but not animal-associated  
106 within households (Post *et al.*, 2019; Koolman *et al.*, 2022). A recent summary concludes that  
107 the available data are consistent with iNTS disease being transmitted person-to-person  
108 (Chirwa *et al.*, 2023). Global efforts to combat iNTS infections are currently focused on

109 vaccine development, which has now progressed to Phase 1 clinical trials (Piccini and  
110 Montomoli, 2020; Skidmore *et al.*, 2023).

111 Since 1998, continuous sentinel surveillance for fever and bloodstream infections among  
112 adults and children has been undertaken at Queen Elizabeth Central Hospital (QECH). This  
113 tertiary referral hospital in Blantyre, Malawi, serves an urban population of about 920,000 with  
114 a high incidence of malaria, HIV and malnutrition (Musicha *et al.*, 2017). Following blood-  
115 culture of samples collected from patients of all ages presenting with fever, whole genome  
116 sequencing identified the ST313 variant of *S. Typhimurium* (Kingsley *et al.*, 2009).

117 Phylogenetic analysis revealed that the chloramphenicol-sensitive ST313 lineage 1 was  
118 clonally-replaced in Malawi by the chloramphenicol-resistant lineage 2 (Okoro *et al.*, 2012).  
119 More recently, a ST313 sublineage II.1 (2.1) emerged from lineage 2 in Democratic Republic  
120 of Congo (DRC) in Central Africa. Sublineage 2.1 had altered phenotypic properties including  
121 biofilm formation and metabolic capacity and resistance to azithromycin (Van Puyvelde *et al.*,  
122 2019). An elegant genomic analysis that provides insight regarding the diversity of *S.*  
123 *Typhimurium* ST19 clades in the context of ST313 lineage 2 clades is also available (Van  
124 Puyvelde *et al.*, 2023).

125 The initial suggestion that ST313 lineage 2 was undergoing evolutionary change in East  
126 Africa came from a small study that identified several *S. Typhimurium* ST313 Malawian  
127 isolates, dated between 2006 and 2008, that differed from lineage 2 by 22 core-genome  
128 single nucleotide polymorphisms (SNPs) (Msefula *et al.*, 2012).

129 To examine the evolutionary trajectory of *S. Typhimurium* in Malawi at a large scale, we  
130 conducted a comparative genomic analysis study focused on 680 isolates dating between  
131 1996 and 2018 (Pulford *et al.*, 2021). We previously reported that ST313 lineage 1 (L1) was  
132 replaced by lineage 2 (here designated L2.0), and discovered an antibiotic-sensitive lineage 3  
133 (L3) that emerged in 2016 (Pulford *et al.*, 2021). We have now performed a more intensive  
134 phylogenetic analysis of the same collection of *S. Typhimurium* ST313 isolates, most of which  
135 caused bloodstream infections in Malawi over two decades. We discovered two novel  
136 sublineages named 2.2 (L2.2) and 2.3 (L2.3) that emerged 2006 - 2008, and have been  
137 replacing L2.0.

138 Here we present a comprehensive comparative genomic analysis of the most prevalent  
139 ST313 L2.2 sublineage, and report the results of a functional genomic approach that identified  
140 key phenotypic characteristics that distinguish L2.2 from L2.0.

141

## 142 **Results**

### 143 **Identification of *S. Typhimurium* ST313 sublineages 2.2 and 2.3 in Malawi**

144 The emergence of the ST313 lineage 2 genotype in Malawi in 2002 prompted us to  
145 hypothesise that subsequent evolution would select for variants with increased fitness,

1  
2  
3 146 leading to the clonal expansion of one or more sublineages by outcompeting previously  
4 147 dominant genotypes. We investigated this hypothesis by conducting a detailed core-gene  
5 148 SNP-based maximum likelihood (ML) phylogenetic analysis to investigate the population  
6 149 structure of *S. Typhimurium* ST313 L2.0 (Fig. S1). As well as identifying members of the  
7 150 antibiotic-sensitive lineage 3, reported previously (Pulford *et al.*, 2021), we discovered that  
8 151 ST313 L2 comprised three phylogenetically-distinct sublineages that differed by a total of 39  
9 152 SNPs. The *S. Typhimurium* ST313 reference strain D23580 (Kingsley *et al.*, 2009) belongs to  
10 153 the ST313 L2.0 lineage (Fig. 1A). As ST313 sublineage L2.1 had been defined previously  
11 154 (Van Puyvelde *et al.*, 2019), the new sublineages which belonged to different hierBAPS level  
12 155 2 clusters were designated L2.2 and L2.3 (Fig. 1A and Fig. S1). In total, we identified 151  
13 156 L2.2 isolates, 74 L2.3 isolates and 350 L2.0 isolates.

14  
15  
16  
17  
18  
19 157 In Blantyre, Malawi, *S. Typhimurium* ST313 L2.2 was first detected in 2006, and L2.3 was  
20 158 initially observed in 2008 (Fig. 1BC). Both L2.2 and L2.3 increased in prevalence at the  
21 159 Queen Elizabeth Central Hospital in Blantyre in subsequent years. By 2018, L2.2 and L2.3  
22 160 had largely replaced L2.0 (Fig. 1BC). Our published Bayesian (BEAST) analysis (Pulford *et*  
23 161 *al.*, 2021) estimated that the Most Recent Common Ancestor (MRCA) of ST313 lineage 2  
24 162 dates back to 1948 (95% HPD = 1929-1959).

25  
26  
27  
28 163 To understand the accessory gene complement of L2.2 and L2.3, we compared the genomes  
29 164 of seven L2.2 isolates and four L2.3 isolates with 17 L2.0 isolates, ST313 L1 and ST19 and  
30 165 the results are shown in Fig. 1A and Table S1. *S. Typhimurium* strain D23580 is the  
31 166 representative strain of L2.0 (Kingsley *et al.*, 2009), for which we previously used long-read  
32 167 sequencing and other approaches to thoroughly characterise the chromosomal and plasmid  
33 168 complement (Canals *et al.*, 2019b).

34  
35  
36  
37  
38 169

### 39 170 **Antimicrobial Resistance**

40  
41 171 MDR variants of *S. Typhimurium* with resistance to ampicillin and cotrimoxazole were  
42 172 detected at an early stage of the iNTS epidemic, from 1997 onwards (Gordon *et al.*, 2008).  
43 173 Multidrug-resistant variants of *S. Typhimurium* ST313 that were no longer susceptible to  
44 174 chloramphenicol, ampicillin and cotrimoxazole subsequently emerged in Malawi (Gordon *et*  
45 175 *al.*, 2008) and have been reported elsewhere in sub-Saharan Africa by the GEMS study  
46 176 (Kasumba *et al.*, 2021). The *S. Typhimurium* ST313 L2.0, L2.2 and L2.3 isolates shared the  
47 177 same MDR profiles (resistance to chloramphenicol, ampicillin and cotrimoxazole), and carried  
48 178 identical IS21-associated antimicrobial gene cassettes within the pSLT-BT plasmid (Fig. 2B).

49  
50  
51  
52  
53 179

54  
55 180

### 56 57 181 **Comparative genomics of *S. Typhimurium* ST313 sublineage 2.2**

58  
59 182 Because *S. Typhimurium* ST313 L2.2 was the predominant novel sublineage in Blantyre,

1  
2  
3 183 Malawi in 2018, we focused on L2.2 for the remainder of this study. We used the phylogeny  
4 184 (Fig. 1A) to select strain D37712 as a representative of L2.2. D37712 was isolated from the  
5 185 blood of an HIV-positive Malawian male child and has been deposited in the National  
6 186 Collection of Type Cultures as [NCTC 14678](#). The draft genome sequence of D37712 was  
7 187 obtained in 2012 with Illumina technology, an assembly that comprised 27 individual contigs  
8 188 (Msefula *et al.*, 2012). To generate a reference-quality genome, we resequenced D37712 with  
9 189 both long-read PacBio and Illumina short-read technologies. Our hybrid strategy generated a  
10 190 complete genome assembly that included one circular chromosome and three plasmids (see  
11 191 Materials & Methods; GenBank CP060165, CP060166, CP060167 and CP060168). This  
12 192 high-quality genome sequence allowed us to conduct a detailed comparison between the  
13 193 genomes of L2.2 strain D37712 and L2.0 strain D23580 (accession number FN424405),  
14 194 summarised in Fig. 2 and Table S2.

15 195 Overall, the gene content of the two strains was largely equivalent. The D23580 annotation  
16 196 contains 4,823 protein-coding and pseudogenes and 287 small RNA (sRNA) genes that we  
17 197 identified previously (Canals *et al.*, 2019b), while D37712 contains 4,821 protein-coding and  
18 198 pseudogenes and the same 287 sRNAs. In total, the D37712 and D23580 genomes shared  
19 199 4,729 orthologous protein-coding genes and pseudogenes. The 104 protein genes that differ  
20 200 are encoded by the pSLT<sup>D37712</sup>, pBT1<sup>D37712</sup>, and pCol1B9<sup>D37712</sup> plasmids.

### 201 **Overview of D23580 and D37712 genomes**

202 The chromosomes of D23580 and D37712 are 4,879,402 and 4,876,060 bp, respectively,  
203 about the same size as other *S. Typhimurium* genomes (Kingsley *et al.*, 2009; Branchu *et al.*,  
204 2018). The D23580 and D37712 strains share an identical prophage profile, with both strains  
205 206 carrying five prophages (BTP1, Gifsy-2, ST64B, Gifsy-1, and BTP5) located at the same  
positions on the chromosome (Fig. 2A).

### 207 **Comparison of D23580 and D37712 chromosomes**

208 The detailed genomic comparison of D37712 with D23580 showed that the sizes of the two  
209 210 chromosomes varied by only 3,342 bp. Overall, the only differences between the genomes of  
the L2.0 and L2.2 strains were 26 chromosomal SNPs and small indels, plus one large  
211 212 deletion, and an inversion of the *hin* switch. In-depth annotation of the nucleotide variants  
213 214 identified 3 putative loss-of-function mutations (2 stop mutations, 1 frameshift insertion), 1  
disruptive in-frame deletion, 4 synonymous mutations, 13 missense mutations, and 5  
215 216 intergenic variants, summarised in Fig. 2A. None of the SNP differences that distinguished  
D37712 from D23580 were located within 150 nucleotides of a Transcriptional Start Site  
(Canals *et al.*, 2019b), and so would not be predicted to modulate gene expression.

217 The 3,358 bp-long deletion of a Gifsy-2 prophage-associated region that spanned the *sseI*  
218 219 pseudogene of D23580 (STMMW\_10631) removed two coding sequences (STM1050-51;  
STMMW\_10611-STMW\_10631), and substantially truncated the STM1049  
220 (STMMW\_10601) gene (Fig. 2E). The *sseI* gene encodes a cysteine hydrolase effector

221 protein that modulates the directional migration of dendritic cells during systemic infection  
222 (Brink *et al.*, 2018). In strain D23580, the insertion of an IS26 transposable element  
223 inactivated the *sseI* gene (Kingsley *et al.*, 2009), causing increased dendritic cell-mediated  
224 dissemination of strain D23580 during infection (Carden *et al.*, 2017). We used an  
225 independent PCR-based approach to confirm that the 3,358 bp deletion had removed the *sseI*  
226 gene from the chromosome of strain D37712 (Fig. S2).

### 227 **Comparison of D23580 and D37712 plasmids**

228 Here we put the genetic features of the representative strains for ST313 L2.0 and L2.2 into  
229 context with other isolates belonging to the Lineage 2 sublineages. ST313 L2.0 strain D23580  
230 carries four plasmids, pSLT-BT, pBT1, pBT2 and pBT3 (Kingsley *et al.*, 2009). In contrast,  
231 ST313 L2.2 has a distinct plasmid complement (Fig. 1A, Fig. 2BCD). The plasmid profiles of  
232 D23580 and D37712 were confirmed by a combination of Illumina (short-read) and PacBio  
233 (long-read) sequencing (Materials and Methods).

234 In summary, strain D37712 carried the pSLT-BT, pBT2 and pCol1B9 plasmids as detailed  
235 below. Both D23580 and D37712 strains carried a variant of the pSLT-BT virulence plasmid  
236 (Kingsley *et al.*, 2009) that contains a Tn21-like transposable element with five antibiotic  
237 resistance genes. The D37712 version of pSLT-BT only differs from the pSLT-BT of D23580  
238 in two important ways (Fig. 2B). Firstly, the Tn21-like element is inserted in the opposite  
239 direction with regards to the rest of the plasmid, suggesting that the transposable element  
240 remains active. Secondly, three nucleotide variants were identified in the pSLT-BT carried by  
241 D37712, two deletions in noncoding regions, and one frameshift insertion that generates a  
242 pseudogene of *spvD*.

243 Plasmid pCol1B9 was of particular interest because it was absent from D23580, but was  
244 present in *S. Typhimurium* ST19 strain 4/74 (Richardson *et al.*, 2011; Fig. 1A). 4/74 is the  
245 parent of *S. Typhimurium* SL1344, a strain that has been used extensively for the study of *S.*  
246 *Typhimurium* pathogenesis and gene regulation in recent decades (Kröger *et al.*, 2012;  
247 Rankin & Taylor, 1966). Our new annotation of the pCol1B9-like plasmid identified 95 distinct  
248 protein-coding genes, while the previously published annotation of pCol1B9<sup>4/74</sup> assigned 101  
249 protein-coding genes. Some of these represent annotation discrepancies, while others  
250 represent true genetic differences (Fig. S3).

251 Following careful examination, we identified 14 genes unique to pCol1B9<sup>D37712</sup>, and 20 genes  
252 unique to pCol1B9<sup>4/74</sup>. There were 81 genes carried by both plasmids. Interestingly,  
253 pCol1B9<sup>D37712</sup> lacked the colicin toxin-antitoxin system that both gave pCol1B9 its name, and  
254 provides *Salmonella* with a competitive advantage in the gut (Nedialkova *et al.*, 2014). The  
255 pCol1B9<sup>D37712</sup> plasmid carried a locus that was absent from pCol1B9<sup>4/74</sup>, namely the *impC-*  
256 *umuCD* operon (Fig. S3) which encodes the error-prone DNA polymerase V responsible for  
257 the increased mutation rate linked to the SOS stress response in *E. coli* (Sikand *et al.*, 2021).

### 258 **Comparison of pseudogene status of D23580 and D37712**



259 Our comparative genomic analysis focused on the pseudogenes found in strains 4/74,  
260 D23580, and D37712 (Fig. 2F, Table S3). The pseudogenisation of several D23580 genes,  
261 compared with strain 4/74, have been linked to the invasive phenotype of African *Salmonella*  
262 ST313 (Kingsley *et al.*, 2009). We found that the pseudogene complement of D23580 was  
263 largely conserved in D37712, consistent with inheritance from a common ancestor. We have  
264 recently reported the role of the MacAB-TolC macrolide efflux pump in the virulence of *S.*  
265 Typhimurium ST313, and showed experimentally that *macB* was an inactive pseudogene in  
266 D23580 (Honeycutt *et al.*, 2020). Interestingly, the *macB* gene is functional in D37712.  
267 Compared with D23580, three additional D37712 genes were pseudogenised (*spvD*, *yadE*,  
268 and STMMW\_42692, as detailed in Table S3). *YadE* is a predicted polysaccharide  
269 deacetylase lipoprotein. The functional impact of these pseudogenes on L2.2 remains to be  
270 established.

271 Overall the chromosomes of ST313 lineage 2 and sublineage 2.2 were highly-conserved and  
272 differed by just 29 SNPs/ small indels, and a 3kb deletion in the Gifsy-2 prophage region. The  
273 ST313 lineage 2 and sublineage 2.2 have distinct plasmid profiles.

#### 274 **Transcriptional landscape of *S. Typhimurium* ST313 sublineage L2.2**

275 Previously, we characterized the primary transcriptome of two other *S. Typhimurium* strains,  
276 4/74 and D23580, using a combination of multi-condition RNA-seq and differential RNA-seq  
277 (dRNA-seq) techniques (Canals *et al.*, 2019b; Kröger *et al.*, 2013). To identify the  
278 transcriptional start sites (TSS) of strain D37712, we analysed a pooled sample containing  
279 RNA from 15 *in vitro* conditions by dRNA-seq and RNA-seq as detailed previously (Kröger *et al.*  
280 *et al.*, 2013). The high similarity between the D23580 and D37712 chromosomes allowed us to  
281 map the curated set of TSS that were previously defined for D23580 (Hammarlöf *et al.*, 2018)  
282 onto a combined D37712/D23580 reference genome. To allow individual TSS to be examined  
283 in particular chromosomal or plasmid regions, data from both the dRNA-seq and pooled RNA-  
284 seq experiments can be visualised in our online genome browser  
285 ([http://hintonlab.com/jbrowse/index.html?data=Combo\\_D37/data](http://hintonlab.com/jbrowse/index.html?data=Combo_D37/data)).

#### 286 **Preliminary gene expression profiling of *S. Typhimurium* ST313 sublineage** 287 **L2.2**

288 Given the high level of similarity between the genomes of L2.2 and L2.0, we went on to  
289 identify differences at the transcriptional level. We performed a multi-condition RNA-seq-  
290 based transcriptomic analysis of gene expression profiles of L2.2 strain D37712 without  
291 biological replicates.

292 This comparative transcriptomic screen was based on our published approach (Canals *et al.*,  
293 2019b). Specifically, we used 15 individual infection-relevant *in vitro* conditions (Kröger *et al.*,  
294 2013) and did intra-macrophage transcriptome profiling using the protocol previously  
295 established for *S. Typhimurium* ST19 (Srikumar *et al.*, 2015). The RNA-seq samples were  
296 mapped to a combined reference genome, which included the annotated D23580  
297 chromosome (Canals *et al.*, 2019b), as well as all the plasmids described earlier (pSLT-BT,

298 pBT1, pBT3 and pCol1B9; see Methods). The initial RNA-seq assessment (detailed in  
299 Methods) involved 2-4M non-rRNA/tRNA reads per sample, allowing gene signatures specific  
300 for each *in vitro* condition to be identified. Although single replicate RNA-seq experiments of  
301 this type cannot be used for statistically-robust differential gene expression analysis, they do  
302 provide a useful screening approach for identifying growth conditions to be used for follow-up  
303 experiments. The individual RNA-seq experiments showed broad condition-specific  
304 similarities in gene expression between strains 4/74, D37712, and D23580 (Fig. 3A). The  
305 gene expression values from each profiled condition are available as raw counts and TPMs in  
306 Tables S4 and S5.

307 To select the ideal environmental conditions to use for subsequent experiments, we assessed  
308 the expression profiles of known *Salmonella* pathogenicity islands which were broadly similar  
309 in strains D37712, and D23580. Although the expression profile of the SPI2 pathogenicity  
310 island was broadly similar between D37712, D23580 and 4/74 in most growth conditions, the  
311 SPI2 genes of D37712 were highly up-regulated in a single growth condition, NonSPI2 (Fig.  
312 3B-C). NonSPI2 is a minimal medium with a neutral pH and a relatively high level of  
313 phosphate, in which *S. Typhimurium* does not usually express the SPI2 pathogenicity island  
314 (Löber *et al.*, 2006; Kröger *et al.*, 2013). This intriguing observation prompted us to perform a  
315 more discriminating set of transcriptomic experiments, as described below.

316

### 317 **Differential gene expression analysis of *S. Typhimurium* D37712 versus D23580 in four** 318 ***in vitro* conditions with multiple biological replicates**

319 To define the transcriptional signature of strain D37712 more accurately, we generated RNA-  
320 seq data from D37712 grown in four *in vitro* conditions that stimulate expression of the  
321 majority of virulence genes: ESP, anaerobic growth, NonSPI2 and InSPI2, with multiple (3-4)  
322 biological replicates. The combination of acidity (pH 5.8) and low phosphate (0.4 mM Pi) in  
323 the InSPI2 media stimulates transcription of SPI2 genes in *S. Typhimurium* (Löber *et al.*,  
324 2006; Kröger *et al.*, 2013). The NonSPI2 condition is based on the same PCN media recipe  
325 as InSPI2 media, but is neutral (pH 7.4), and contains higher levels of phosphate (25 mM Pi)  
326 (Löber *et al.*, 2006; Kröger *et al.*, 2013).

327 We compared the results with our published transcriptomic data for *S. Typhimurium* strains  
328 4/74 and D23580 (Canals *et al.*, 2019b; Kröger *et al.*, 2013). Differential expression analysis  
329 with DEseq2, with conservative cut-offs (fold change  $\geq 2$ , FDR  $\leq 0.001$ ), showed that the gene  
330 expression profiles of D37712 and D23580 were broadly similar, and shared key differences  
331 to the transcriptional profile of strain 4/74 under each of the four *in vitro* conditions (Fig. 4A).  
332 The differential expression results are summarized in Table S6.

333 We specifically investigated transcription of the *pgtE* gene, which encodes the outer-  
334 membrane protease previously linked to the ability of African *Salmonella* ST313 to resist  
335 human serum killing (Hammarlöf *et al.*, 2018). Compared to 4/74, the *pgtE* gene of both the

1  
2  
3 336 D23580 and D37712 strains showed a similar pattern of up-regulation by a factor of 7 to 18  
4 337 across all conditions. This finding is consistent with the fact that D37712 carries the same T  
5 338 nucleotide in the -10 region of the *pgtE* promoter that is responsible for increased expression  
6 339 of the *pgtE* transcript in strain D23580 (Hammarlöf *et al.*, 2018).

9 340 There were no statistically-significant changes in expression of the majority (92%) of the  
10 341 4,729 orthologous coding genes shared by D37712 and D23580. We identified a total of 364  
11 342 genes that were differentially expressed in at least one growth condition between D37712 and  
12 343 D23580 as follows: ESP (69 differentially-expressed genes), anaerobic growth (214  
13 344 differentially-expressed genes), NonSPI2 (88 differentially-expressed genes) and InSPI2 (17  
14 345 differentially-expressed genes; Fig. 4B).

18 346 Overall, the differentially-expressed genes that distinguished D37712 from D23580 only  
19 347 showed expression differences in a single growth condition rather than across all conditions.  
20 348 The differentially expressed genes included flagellar genes (down-regulated), SPI2-  
21 349 associated genes (up-regulated), and genes involved in general and anaerobic metabolism  
22 350 (down-regulated).

26 351 SPI2 pathogenicity island genes play a key role in the intracellular replication of *S.*  
27 352 Typhimurium, and encode the type III secretion system that is responsible for translocation of  
28 353 key effector proteins into mammalian cells (Jennings *et al.*, 2017). The RNA-seq data showed  
29 354 that SPI2 genes were expressed at similarly high levels in both D37712 and D23580 strains  
30 355 following induction (InSPI2 media; Fig. 4B), and confirmed that the key SPI2 expression  
31 356 difference was only seen in strain D37712 under non-inducing growth conditions (NonSPI2  
32 357 media). It is important to put this differential SPI2 expression into context. D37712 expresses  
33 358 SPI2 genes at about a 10-fold higher level than D23580 during growth in non-inducing  
34 359 NonSPI2 media, but the actual level of expression was 20-fold less than the level stimulated  
35 360 by growth in SPI2-inducing conditions (InSPI2 medium).

41 361 The up-regulation of *fljA* and *fljB* and the down-regulation of *fliC* in D37712, compared to  
42 362 D23580 in all four growth conditions likely reflects the opposite orientation of the *hin* switch in  
43 363 the D37712 genome compared to D23580. This type of *hin* inversion occurs frequently in *S.*  
44 364 Typhimurium (Johnson and Simon, 1985).

47 365 Another gene that was up-regulated in D37712 across all profiled conditions was the  
48 366 chromosomally-encoded *cysS<sup>chr</sup>*, that encodes cysteine-tRNA synthetase. Previously, we  
49 367 reported that transcription of the *cysS<sup>chr</sup>* of strain D23580 was uniformly down-regulated  
50 368 compared to 4/74. This down-regulation was compensated by the presence of a pBT1  
51 369 plasmid-encoded cysteine-tRNA synthetase (Canals *et al.*, 2019a). Accordingly, the increased  
52 370 expression of the chromosomal *cysS* gene in D37712 was consistent with the absence of the  
53 371 pBT1 plasmid. Our comparative transcriptomic analysis showed that expression levels of  
54 372 *cysS* were similar in D37712 and 4/74 under all growth conditions.



3 373 Numerous virulence genes and operons were differentially expressed between D23580 and  
4 374 D37712. The SPI-16-associated *gtrABCa* operon (STM0557, STM0558, STM0559) is  
5 375 responsible for adding glucose residues to the O-antigen subunits of LPS that enhance the  
6 376 long-term colonisation of the mammalian gastrointestinal tract by *S. Typhimurium* ST19  
7 377 (Bogomolnaya *et al.*, 2008). We found that the *gtrABCa* genes were significantly up-regulated  
8 378 in several conditions in D37712, compared to both D23580 and 4/74.

9 379 The *spvABCD* operon of D37712 was up-regulated under non-SPI2-inducing growth  
10 380 conditions, compared to D23580. A signature pseudogene of ST313 L2.2 is the frameshift  
11 381 insertion in the *spvD* gene that generates a truncated version of the SpvD protein. The H1991  
12 382 mutation at position 199 and the associated 17 amino acid truncation is predicted to ablate  
13 383 the activity of the SpvD cysteine protease (Grabe *et al.*, 2016). SpvD negatively regulates the  
14 384 NF- $\kappa$ B signaling pathway and promotes virulence of *S. Typhimurium* in mice. The functional  
15 385 consequences of the *spvD* variant of ST313 L2.2 strain D37712 and the up-regulation of  
16 386 expression of the *spvABCD* operon remain to be established experimentally.  
17 387

### 18 388 **The SalComD37712 community transcriptional data resource**

19 389 To allow scientists to gain their own biological insights from analysis of this rich transcriptomic  
20 390 dataset, the transcriptomic and gene expression data generated in this study are presented  
21 391 online in a new community resource, [SalComD37712](#). The data resource shows the  
22 392 expression levels of all D37712 coding and non-coding genes, including both chromosomal  
23 393 and plasmid-encoded transcripts. The SalComD37712 website complements our existing  
24 394 SalComD23580 (<https://tinyurl.com/SalComD23580>) resource, and adds an inter-strain  
25 395 comparison of gene expression profiles between D37712 and D23580 as well as normalized  
26 396 gene expression values (TPM), using an intuitive heat map-based approach. [SalComD37712](#)  
27 397 included our published RNA-seq data (Canals *et al.*, 2019b), re-analysed with an updated  
28 398 bioinformatic pipeline and a combined reference genome (see Methods). This online resource  
29 399 facilitates the intuitive interrogation of transcriptomic data as described previously (Perez-  
30 400 Sepulveda and Hinton, 2018).

31 401 Additionally, we generated a unified genome-level browser that provides access to the *S.*  
32 402 *Typhimurium* L2.2 D37712 transcriptome, in the context of our previously published RNA-seq  
33 403 data for the L2.0 strain D23580 and the ST19 strain 4/74. This novel “combo” browser is  
34 404 available at [http://hintonlab.com/jbrowse/index.html?data=Combo\\_D37/data](http://hintonlab.com/jbrowse/index.html?data=Combo_D37/data).

35 405

**406 Identification of phenotypes that distinguish ST313 sublineage L2.2 from L2.0.**

407 To explore the phenotypic impact of the transcriptomic signature of L2.2 (D37712), we  
408 performed a series of motility experiments, fluorescence-based gene expression experiments  
409 and mixed-growth assays.

410 D33712 showed a significantly decreased level of motility on NonSPI2 minimal media,  
411 compared with both the ST19 strain 4/74 and the L2 D23580 strain (Fig. 5A). This finding was  
412 consistent with the transcriptomic data, which showed down-regulation of D37712 flagellar  
413 genes compared with D23580 in the NonSPI2 condition (Fig. 4). In contrast, no differential  
414 expression of flagellar genes was seen between D33712 and D23580 in the InSPI2 growth  
415 condition (Fig. 4). The decreased motility phenotype may be linked to the inversion of the *hin*  
416 element detailed above. The flagella system encodes a distinct type III secretion apparatus  
417 responsible for the dual functions of bacterial motility and activation of the mammalian innate  
418 immune system via TLR5 (Lai *et al.*, 2013).

419 A key transcriptomic finding for strain D33712 was the expression of SPI2 genes during  
420 growth in an unusual environmental condition (NonSPI2) (Fig. 3B-C and Fig. 4B). NonSPI2  
421 media differs from InSPI2 media by having a higher pH (pH7.4 versus pH5.8) and a higher  
422 level of phosphate (Löber *et al.*, 2006). This apparent differential expression of SPI2 genes at  
423 the transcriptomic level under non-inducing conditions led us to investigate the expression of  
424 SPI2 at a single cell level using fluorescence transcriptional fusions. First, we introduced an  
425 *ssaG*-GFP<sup>+</sup> transcriptional fusion into the chromosome of strains D33712 and D23580  
426 (Methods; Table S8) to interrogate expression of the key SPI2 operon with flow cytometry.  
427 Figure 5B shows that in NonSPI2 media, the *ssaG* promoter was expressed at a 62% higher  
428 level in D33712 than in D23580 confirming the results of the transcriptomic analysis.

429 Because only a proportion of *S. Typhimurium* cells express certain pathogenicity island-  
430 encoded genes during *in vitro* growth (Ackermann *et al.*, 2008; Hautefort *et al.*, 2003), we  
431 determined whether the increased level of expression of SPI2 genes (Fig. 4B) was caused by  
432 a higher proportion of D33712 cells expressing SPI2 than D23580 cells. Using derivatives of  
433 the two strains that carried the *ssaG*-GFP<sup>+</sup> construct, we determined the numbers of  
434 fluorescent and non-fluorescent cells with flow cytometry (Methods). Under non-inducing  
435 conditions, slightly more D37712 cells expressed the *ssaG* SPI2 promoter than D23580 cells  
436 (65% vs 60%, respectively) (Fig. 5C). Although this small difference was statistically  
437 significant (t-test:  $P < 0.001$ ,  $n = 3$ ), it did not account for the 62% increased level of non-induced  
438 SPI2 expression seen in Fig. 5B.

439 SPI2 expression is controlled by a complex regulatory system that operates at both a  
440 negative and positive level, involving silencing via H-NS (Lucchini *et al.*, 2006), activation by  
441 SlyA and SsrB (Fass and Groisman, 2009; Walthers *et al.*, 2011) as well as input from OmpR  
442 and Fis under non-inducing conditions (Osborne and Coombes, 2011). The mechanistic basis  
443 of the aberrant SPI2 expression in strain D37712 is worthy of further study. Possible  
444 explanations include the incomplete silencing of SPI2 transcription or the partial activation of

1  
2  
3 445 the SPI2 virulence genes under non-inducing growth conditions by an unknown regulatory  
4 446 factor.

5  
6 447

7  
8 448 **Increased fitness of *S. Typhimurium* ST313 sublineage L2.2 compared with L2.0 in**  
9 449 **minimal media.**

10  
11 450 It has become increasingly clear that distinct *Salmonella* pathovariants have evolved  
12 451 particular phenotypic properties that confer fitness advantages during infection of particular  
13 452 avian or mammalian hosts (Branchu *et al.*, 2018). Because *S. Typhimurium* ST313 L2.2  
14 453 appeared to have displaced *S. Typhimurium* ST313 L2.0 in Malawi, we speculated that *S.*  
15 454 *Typhimurium* ST313 L2.2 might have the competitive edge in some situations. Accordingly,  
16 455 we determined bacterial fitness using a mixed-growth competition assay (Wiser and Lenski,  
17 456 2015; Lian *et al.*, 2023). The competitive index was calculated in three different growth media  
18 457 using pair-wise combinations of strains D37712 and D23580. Two independent approaches  
19 458 were used to phenotypically distinguish the two strains, one based on antibiotic resistance  
20 459 (Fig. 5D) and the other based on fluorescent tagging (Fig. S5).

21  
22 460 To confirm that strains engineered to be kanamycin-resistant or gentamicin-resistant did not  
23 461 impact on fitness (Methods), we first verified that the tagged variants of D37712 or D23580  
24 462 did not confer a growth advantage in LB or NonSPI2 media (Fig. S7). Next, we used a mixed-  
25 463 growth assay to investigate fitness of *S. Typhimurium* ST313 L2.0 strain D23580 or *S.*  
26 464 *Typhimurium* ST313 L2.2 strain D37712 during growth in LB, or InSPI2 or NonSPI2 minimal  
27 465 media. The data show that both strains grew at similar levels following overnight mixed-  
28 466 growth in nutrient-rich LB media, but D37712 had a competitive advantage during mixed-  
29 467 growth in InSPI2 media (CI = 1.79;  $P < 0.05$ ) and a greater competitive edge in NonSPI2 media  
30 468 (CI = 2.20;  $P < 0.0001$ ).

31  
32 469 We then used an independent fluorescence-based approach to assess the fitness of strains  
33 470 D23580 and D37712 during mixed-growth in NonSPI2 media. This time, the strains were  
34 471 engineered to carry either mScarlet or sGFP2 proteins and the mixed-growth experiments  
35 472 involved pair-wise comparisons of reciprocally-tagged strains. The flow cytometric data  
36 473 showed that in both cases D37712 had a significant competitive advantage in NonSPI2 media  
37 474 (Fig. S5 and S6).

38  
39 475 This combination of antibiotic resistance-based and fluorescence-based competitive index  
40 476 experiments lead us to conclude that *S. Typhimurium* ST313 L2.2 strain D37712 had a clear  
41 477 fitness advantage over *S. Typhimurium* ST313 L2.0 strain D23580 during mixed-growth in two  
42 478 formulations of minimal media. The molecular basis of this fitness advantage remains to be  
43 479 established.

44  
45 480  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

481 **Perspective**

482 Here, we report that *S. Typhimurium* ST313 L2.0 has been clonally replaced by the ST313  
483 sublineages L2.2 and L2.3 as a cause of bloodstream infection in Blantyre, Malawi. In 2018,  
484 L2.2 represented the majority of the ST313 strains isolated from hospitalised patients in  
485 Malawi at the Queen Elizabeth Central Hospital. Our comparative genomic analysis of ST313  
486 L2.3 identified 30 chromosomal alterations, one of which generated a deletion of the *sseI*  
487 effector gene.

488 Our RNA-seq-based analysis of ST313 L2.2 involved a detailed comparison versus ST313  
489 L2.0 which revealed a key difference involving SPI2 expression. Following initially  
490 observations at the transcriptomic level in the ST313 L2 and L2.2 strains grown in a pH-  
491 neutral minimal medium (NonSPI2), the increased expression of SPI2 was confirmed at the  
492 single cell level using an *ssaG* transcriptional fusion.

493 A series of experiments showed that the ST313 L2.2 strain D37712 had a competitive  
494 advantage over L2 strain D23580 during mixed-growth in minimal media. We propose that  
495 this increased fitness of *S. Typhimurium* ST313 L2.2 has contributed to the replacement of  
496 ST313 L2.0 in Malawi in recent years.

497 Previously, we compared three virulence properties of the *S. Typhimurium* ST313 L2.0  
498 D23580 and ST313 L2.2 D37712 strains. First, experiments involving Mucosal Invariant T  
499 (MAIT) cells showed that both D37712 and D23580 fail to elicit the high level of activation of  
500 MAIT cells that characterises infection by *S. Typhimurium* ST19 4/74 (Preciado-Llanes *et al.*,  
501 2020). Second, the D37712 and D23580 strains stimulate similar levels of up-regulation of  
502 IL10 gene expression upon infection of human dendritic cells (Aulicino *et al.*, 2022). Third, we  
503 showed that both D37712 and D23580 express similarly high levels of the PgtE virulence  
504 factor that is responsible for the ability of *S. Typhimurium* ST313 to survive human serum-  
505 killing (Hammarlöf *et al.*, 2018). These findings lead us to conclude that the comparative  
506 genomic and transcriptomic differences that distinguish *S. Typhimurium* ST313 L2.0 strain  
507 D23580 from ST313 L2.2 D37712 (Fig. 4) do not modulate the ability of the pathogens to  
508 activate human MAIT cells or dendritic cells, or to influence the PgtE-mediated serum survival  
509 phenotype of *S. Typhimurium* ST313.

510 Ideally, the implications of the competitive advantage of ST313 L2.2 would be determined in  
511 the context of pathogenesis. However, we lack an informative infection model for *S.*  
512 *Typhimurium* ST313 (Lacharme-Lora *et al.*, 2019), and it is not yet possible to experimentally  
513 determine whether the improved fitness of L2.2 significantly enhances the success of ST313  
514 during infection of humans.

515 Here we have investigated the intricate interplay of gene function that underpins the success  
516 of *S. Typhimurium* ST313 L2.2. Our hope is that our findings could contribute to future  
517 therapeutic or prophylactic strategies for combatting iNTS infections in the African setting.

## 518 **Materials and methods**

### 519 **Bacterial strains**

520 To investigate the evolutionary dynamics of *S. Typhimurium* ST313 L2 in Malawi over a 22  
521 year period, we focused on the large collection of 8,000 *S. Typhimurium* isolates derived from  
522 bloodstream infection in hospitalised patients at the Queen Elizabeth Central Hospital,  
523 Blantyre, Malawi (Feasey *et al.*, 2015). The collection was assembled by the Malawi–  
524 Liverpool–Wellcome Trust Clinical Research Programme (MLW) between 1996 and 2018; the  
525 precise annual numbers of isolates are shown in Fig. 1C. A random sub-sampling strategy  
526 was used to select 608 isolates for whole-genome sequencing, which included 549 *S.*  
527 *Typhimurium* ST313 isolates (Pulford *et al.*, 2021).

528

529 The two *S. Typhimurium* ST313 strains that are the focus of this study are D23580 and  
530 D37712. D23580 was isolated from a Malawian 26-month-old child with malaria and anaemia  
531 in 2004. D37712 was isolated from the blood of an HIV-positive Malawian male child in 2006.  
532 These two African *Salmonella* strains have been deposited in the National Collection of Type  
533 Cultures (NCTC). The D23580 (lineage 2.0) strain is available as [NCTC 14677](#). The ST313  
534 sublineage 2.2 strain D37712 is available as [NCTC 14678](#). All bacterial strains are detailed in  
535 Table S8.

### 536 **Genome sequencing**

537 The assembled genome and annotation of D23580 (Kingsley *et al.*, 2009; Canals *et al.*,  
538 2019b) (L2.0) was obtained from the European Nucleotide Archive (ENA) repository (EMBL-  
539 EBI) under accession PRJEB28511 (<https://www.ebi.ac.uk/ena/data/view/PRJEB28511>). For  
540 genome sequencing of D37712 (L2.2), DNA was extracted using the Bioline mini kit, and  
541 quality was assessed using gel electrophoresis (0.5% agarose gel, at 30 volts for 18 h). The  
542 genome was generated by a combination of long read sequencing with a PacBio RS II and  
543 short-read sequencing on an Illumina HiSeq machine at the Center for Genome Research,  
544 University of Liverpool, United Kingdom.

545 Sequence reads were quality checked using FastQC version 0.11.9 (Andrews, 2010) and  
546 MultiQC version 1.8 (Ewels *et al.*, 2016), trimmed using Trimmomatic (Bolger *et al.*, 2014).  
547 Hybrid assembly of the Illumina and PacBio sequence reads was done with Unicycler v0.4.7  
548 (Wick *et al.*, 2017).

549 The assembled genome of *S. Typhimurium* SDT313 L2.2 strain D37712 was deposited in  
550 Genbank (GCA\_014250335.1, assembly ASM1425033v1). Raw sequencing reads were  
551 deposited for both PacBio and Illumina, under BioProject ID PRJNA656698. Sequence Read  
552 Archive (SRA) database IDs are: SRR12444880 for Illumina and SRR12444881 for PacBio.

### 553 **Comparative genomic analyses**



1  
2  
3 554 To generate the data summarised in Fig. 1C, sequencing data of 29 *S. Typhimurium* ST313  
4 555 strains (Msefula *et al.*, 2012) were downloaded from EMBL-EBI database  
5 556 (<https://www.ebi.ac.uk>, accession number ERA015722). Sequence reads were assembled  
6 557 using Unicycler v0.4.8 (Wick *et al.*, 2017). The quality of the assemblies was assessed by  
7 558 Quast v5.0.2 (Gurevich *et al.*, 2013). The N50 value of all assemblies was >20kb, and the  
8 559 number of contigs was <600.

9  
10  
11  
12 560 To construct the phylogenetic tree (Fig. 1C), *Salmonella* Typhimurium strains D23580,  
13 561 D37712, LT2 (GCA\_000006945.2), DT104 (GCA\_000493675.1), 4/74 (GCA\_000188735.1),  
14 562 and A130 (GCA\_902500285.1) were added as contextual genomes. Roary was used to make  
15 563 the core gene alignment, construct the gene presence/absence matrix and identify  
16 564 orthologous genes (Page *et al.*, 2015). Phylogenetic trees were constructed using  
17 565 Randomized Accelerated Maximum Likelihood (RAxML) (Stamatakis *et al.*, 2005), and were  
18 566 visualised with the interactive Tree of Life online tool (iTOL) (Letunic and Bork, 2006).

19  
20  
21  
22  
23 567 The assembled genome and annotation of *S. Typhimurium* ST19 representative strain 4/74  
24 568 (Richardson *et al.*, 2011) were obtained from GenBank (Accession number  
25 569 GCF\_000188735.1), while the raw sequencing data of 27 *S. Typhimurium* ST313 strains  
26 570 described in a previous study (Msefula *et al.*, 2012) were downloaded from EMBL-EBI  
27 571 database (<https://www.ebi.ac.uk>, accession number ERA015722). The raw reads were  
28 572 assembled using Unicycler v0.4.8 (Wick *et al.*, 2017). The quality of the assemblies was  
29 573 assessed by Quast v5.0.2 (Gurevich *et al.*, 2013). The N50 value of all assemblies  
30 574 was >20kb, and the number of contigs was <600.

31  
32  
33  
34  
35 575 To identify SNPs, Snippy v4.4.0 (<https://github.com/tseemann/snippy>) was used to map the  
36 576 raw reads against the 4/74 genome. To detect pseudogene-associated SNPs/indels in each  
37 577 sub-lineage, the SNPs/indels that caused nonsense or frameshifted mutations were filtered.  
38 578 The identifications and names of the disrupted genes were summarised, then the wild type  
39 579 gene sequences were extracted from the 4/74 genome. To validate the pseudogene-  
40 580 associated SNPs/indels, the wild type gene sequences were used to make a BLAST  
41 581 database with BLAST 2.9.0+ (Camacho *et al.*, 2009). The 29 genome assemblies were  
42 582 queried against the databases, using the BLASTn algorithm to confirm the nonsense and  
43 583 frameshifted mutations in all isolates.

#### 44 45 46 47 48 584 **Phylogenetic analysis of African *Salmonella* Typhimurium isolates dating from 1966 -** 49 585 **2018**

50  
51 586 To examine the overall population structure of *Salmonella* Typhimurium responsible for blood  
52 587 infection in Malawi (Fig. 1AB and Fig. S1), the raw reads of 707 published genome  
53 588 sequences were downloaded (Table S7). Trimmomatic v0.36 (Bolger, A. M., Lohse, 2014)  
54 589 was used to trim adapters and Seqtk v1.2-r94 (<https://github.com/lh3/seqtk>) was used to trim  
55 590 low-quality regions using the trimfq flag. Fastqc v0.11.5 (<https://www>.  
56 591 [bioinformatics.babraham.ac.uk/projects/fastqc/](http://bioinformatics.babraham.ac.uk/projects/fastqc/)) and multiqc v1.0 (<http://multiqc.info>) were  
57 592 used to pass sequence reads according to the following criteria: passed basic quality

1  
2  
3 593 statistics, per base sequence quality, per base N content, adapter content and an average  
4 594 GC content of between 47% and 57%. Only high-quality reads were used in the downstream  
5 595 analysis. Sequence reads were aligned to the *S. Typhimurium* D23580 genome using Snippy  
6 596 v4.4.0 with parameter “- - mincov 5”. The recombination sites of the alignment were removed  
7 597 by Gubbins (Croucher *et al.*, 2015), and the phylogenetic tree was built with Raxml-ng (Kozlov  
8 598 et al., 2019) using GTR\_G models ad 100 bootstraps. The tree was rooted on *Salmonella*  
9 599 Typhi strain CT18 (GCA\_000195995.1) as the outgroup. The tree was visualised with the  
10 600 interactive Tree of Life online tool (iTOL) (Letunic and Bork, 2006). The sub-lineages were  
11 601 identified with rHierBAPS (Tonkin-Hill *et al.*, 2018). The stacked-area chart and the bar chart  
12 602 showing the percentage and number of isolates from each sub-lineage were made in MS  
13 603 Excel.

#### 19 604 **RNA purification and growth conditions**

20  
21 605 Initially, a screen of transcriptomic gene expression was performed without biological  
22 606 replicates. Total RNA was purified using TRIzol from *S. Typhimurium* D37712 grown in 15  
23 607 different conditions as described previously (Kröger *et al.*, 2013). To generate statistically-  
24 608 robust gene expression profiles, total RNA was subsequently purified using TRIzol from *S.*  
25 609 *Typhimurium* D37712 grown in four *in vitro* growth conditions (ESP, anaerobic growth,  
26 610 NonSPI2, InSPI2) with three biological replicates as described previously (Kröger *et al.*,  
27 611 2013). RNA was isolated from intra-macrophage D37712 following infection of RAW264.7  
28 612 murine macrophages using our published protocol (Srikumar *et al.*, 2015).

#### 33 613 **RNA-seq of *S. Typhimurium* strain D37712 using Illumina technology**

34  
35 614 For transcriptomic analyses, cDNA samples were prepared from *S. Typhimurium* RNA by  
36 615 Vertis Biotechnologie AG (Freising, Germany). RNA was first treated with DNase and purified  
37 616 using the Agencourt RNAClean XP kit (Beckman Coulter Genomics). RNA samples were  
38 617 sheared using ultrasound, treated with antarctic phosphatase and re-phosphorylated with T4  
39 618 polynucleotide kinase. RNA fragments were poly(A)-tailed using poly(A) polymerase and an  
40 619 RNA adapter was ligated to the 5'- phosphate of the RNA. First-strand cDNA synthesis was  
41 620 performed using an oligo(dT)-adapter primer and M-MLV reverse transcriptase. The resulting  
42 621 cDNA was PCR-amplified to about 10-20 ng/μl. The cDNA was purified using the Agencourt  
43 622 AMPure XP kit. The cDNA samples were pooled using equimolar amounts and size  
44 623 fractionated in the size range of 200-500 bp using preparative agarose gels. The cDNA pool  
45 624 was sequenced on an Illumina NextSeq 500 system using 75 bp read length.

46  
47 625 For the biological replicates of the four growth conditions (ESP, anaerobic growth  
48 626 (abbreviated as NoO<sub>2</sub>), NonSPI2, and InSPI2) and the intra-macrophage RNA, cDNA  
49 627 samples were generated as above with some improvements in library preparation. First, after  
50 628 fragmentation with ultrasound, an oligonucleotide adapter was ligated to the 3' end of the  
51 629 RNA molecules. Second, first-strand cDNA synthesis was performed using M-MLV reverse  
52 630 transcriptase and the 3' adapter as primer, and, after purification, the 5' Illumina TruSeq  
53 631 sequencing adapter was ligated to the 3' end of the antisense cDNA. Sequencing of the

632 cDNA was performed as described above. All raw sequencing reads were deposited to the  
633 Gene Expression Omnibus (GEO) database under accession GSE161403.

#### 634 **RNA-seq and dRNA-seq read processing and visualization**

635 RNA-seq data from *S. Typhimurium* 4/74 and D23580 were extracted from previously  
636 published experiments (Kröger *et al.*, 2013; Srikumar *et al.*, 2015; Canals *et al.*, 2019b; GEO  
637 dataset GSE119724). A combined reference genome was generated that contained the  
638 D23580 chromosome plus plasmids pBT1, pBT2, pBT3, pSLT-BT (from D23580) and the  
639 D37712 plasmid pCol1B9<sup>D37712</sup>. All reads were aligned and quantified using Bacpipe v0.8a  
640 (<https://github.com/apredeus/multi-bacpipe>). Briefly, basic read quality control was performed  
641 with FastQC v0.11.8. RNA-seq reads were aligned to the genome sequence using STAR  
642 v2.6.0c using “--alignIntronMin 20 --alignIntronMax 19 --outFilterMultimapNmax 20” options. A  
643 combined GFF file was generated by Bacpipe, where all features of interest were listed as a  
644 “gene”, with each gene identified by a D37712 locus tag. Subsequently, read counting was  
645 done by featureCounts v1.6.4, using options “-O -M --fraction -t gene -g ID -s 1”. For  
646 visualization, scaled gedGraph files were generated using bedtools genomecov with a scaling  
647 coefficient of  $10^9/(\text{number of aligned bases})$ , separately for sense and antisense DNA  
648 strands. Bedgraph files were converted to bigWig using bedGraphToBigWig utility  
649 ([http://hgdownload.soe.ucsc.edu/admin/exe/linux.x86\\_64/](http://hgdownload.soe.ucsc.edu/admin/exe/linux.x86_64/)). Coverage tracks, annotation, and  
650 genome sequence were visualized using JBrowse v1.16.6. Transcripts Per Million (TPM)  
651 were calculated for all samples and used as absolute expression values (Table S5). A  
652 conservative cut-off was used to distinguish between expressed (TPM >10) and not  
653 expressed (TPM ≤10), as we previously described (Kröger *et al.*, 2013). Relative expression  
654 values were calculated by dividing the TPM value for one condition in one strain by the TPM  
655 value for the same condition in a different strain. Before the calculation, all TPM values below  
656 10 were set up to 10. A conservative fold-change cut-off of 3 was used to highlight differences  
657 in expression between strains.

#### 658 **Differential gene expression analysis with multiple biological replicates**

659 For differential expression analysis of *S. Typhimurium* strains 4/74, D23580, and D37712, the  
660 raw counts (Table S4) from 3-5 biological replicates in four growth conditions were used  
661 (ESP, anaerobic growth (abbreviated as NoO<sub>2</sub>), NonSPI2, and InSPI2). Differential  
662 expression analysis was done using DESeq2 v1.24.0 with default settings. A gene was  
663 considered to be differentially expressed if the absolute value of its log<sub>2</sub> fold change was at  
664 least 1 (i.e. fold change > 2), and adjusted p-value was < 0.001.

#### 665 **The SalComD37712 community data resource, and the associated Jbrowse genome 666 browser**

667 SalCom provides a user-friendly Web interface that allows the visualisation and comparison of  
668 gene expression values across multiple conditions and between strains. Particular genes can  
669 be selected through pre-defined lists of interest, such as all sRNAs or all genes belonging to a



670 specific pathogenicity island. The resulting heatmap-style display highlights expression  
671 differences, and provides access to the rich, manually curated annotation of strains D37712  
672 and D23580. The actual values behind the display can be downloaded for further processing,  
673 and a link connects the current view to a genome browser interface.

674 Visualisation of all the RNA-seq and dRNA-seq (TSS) coverage tracks in JBrowse 1.16.6  
675 shows sequence reads mapped against the combined reference genome described above.  
676 Overall, the genomic distance between strains 4/74 and D23580 (approximately 1000 SNPs,  
677 or ~1 SNP per 5000 nucleotides), and between D37712 and D23580 (approximately 30  
678 SNPs, ~1 SNP per 150,000 nucleotides) allowed the alignment of RNA-seq reads to the  
679 simplified combined reference genome without significant loss of reads. The combined  
680 reference genome facilitated a direct comparison of gene coverage as well as transcriptional  
681 start sites. The unified browser is hosted at  
682 [http://hintonlab.com/jbrowse/index.html?data=Combo\\_D37/data](http://hintonlab.com/jbrowse/index.html?data=Combo_D37/data).

### 683 **Phenotypic and mixed competitive growth experiments**

684 The swimming motility of *S. Typhimurium* strains D37712, D23580 and 4/74 was determined  
685 by a plate assay (Canals *et al.*, 2019b), which involved spotting 3  $\mu$ L overnight culture onto  
686 0.3% LB agar. Relative motility of the three strains was assessed by migration diameter after  
687 4h and 8h of incubation at 37°C.

688 Relative expression of the *ssaG* SPI2 promoter in strains D23580 and D37712 was measured  
689 at the single cell level via GFP fluorescence. Following the construction of a kanamycin-  
690 sensitive derivative of D23580 (strain JH4235), a *PssaG::gfp*<sup>+</sup> transcriptional fusion was  
691 incorporated into the chromosome of JH4235 and D37712 by inserting the *gfp*<sup>+</sup> gene  
692 downstream of the *ssaG* gene, under the control of the *PssaG* promoter. The *PssaG::gfp*<sup>+</sup>  
693 D23580 derivative (JH4692), and the *PssaG::gfp*<sup>+</sup> D37712 derivative (JH4693) are listed in  
694 Table S8.

695 The strains JH4692 and JH4693 were genome sequenced to confirm the integrity of the  
696 transcriptional fusions, and to verify that unintended nucleotide changes had not arisen.  
697 Following growth in 25 mL non-inducing NonSPI2 media in a 250 mL flask at 37°C with  
698 shaking at 220 rpm for approximately 8 hours until OD<sub>600</sub>=0.3, fluorescence was determined  
699 with a BD FACSAria Flow Cytometer. The relative fluorescence of the two strains JH4692 and  
700 JH4693, and the numbers of individual fluorescent bacteria that expressed the *PssaG::gfp*<sup>+</sup>  
701 promoter, were determined with FlowJo VX software.

702 The relative fitness of *S. Typhimurium* strains D37712 and D23580 was assessed in two  
703 independent mixed-growth experiments. First, kanamycin-resistant derivatives of each strain  
704 were constructed by inserting the *aph* kanamycin resistance gene into the chromosome at the  
705 intergenic region between the *STM4196* and *STM4197* genes, a region that we have  
706 previously shown to be transcriptionally silent (Canals *et al.*, 2019b). The strains were  
707 designated D23580::Km<sup>R</sup> JH3794 and D37712::Km<sup>R</sup>, JH4232. Mixed cultures of wild-type or

1  
2  
3 708 kanamycin-resistant derivatives of each strain were grown overnight in LB, InSPI2 and  
4 709 NonSPI2 media in a 250 mL flask at 37°C with shaking at 220 rpm. Following plating on LB  
5 710 agar or LB + kanamycin, colonies were counted and the ratio of bacterial strains was  
6 711 determined. To confirm that the insertion of kanamycin resistance at the intergenic region  
7 712 between *STM4196* and *STM4197* did not impact upon fitness, a mixed-growth experiment  
8 713 was done in both LB and NonSPI2 media (Fig. S7).  
9  
10  
11  
12 714 Second, to independently assess relative fitness, Tn7-based plasmids (Schlechter and  
13 715 Remus-Emsermann, 2019) were used to construct chromosomal sGFP2 and mScarlet  
14 716 derivatives of *S. Typhimurium* strains D23580 (sGFP2 derivative: JH4694; mScarlet  
15 717 derivative: JH4695) and D37712 (sGFP2 derivative: JH4696; mScarlet derivative: JH4697).  
16 718 The gene cassettes were inserted into the *S. Typhimurium*Tn7 insertion site between the  
17 719 gene *STMMW\_38451* and *glmS*. Mixed cultures of pairs of fluorescently-labelled strains were  
18 720 grown in NonSPI2 media at 37°C with shaking at 220 rpm for approximately 8 hours until  
19 721  $OD_{600}=0.3$ . Levels of green and red fluorescence were determined with a BD FACSAria Flow  
20 722 Cytometer.  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

723 **Figure Legends**

724 **Fig. 1. Emergence of *S. Typhimurium* ST313 sublineages L2.2 and L2.3 in Malawi.** (A)  
 725 Evolutionary dynamics of *S. Typhimurium* lineages in Blantyre, Malawi from 1996 to 2018. A  
 726 maximum likelihood tree constructed with 1000 bootstraps using the GTRGAMMA model in  
 727 RaxML rooted on ST19, LT2. The genomes of 549 *S. Typhimurium* ST313 isolates from  
 728 bacteraemic patients at the Queen Elizabeth Hospital in Blantyre, Malawi were used for this  
 729 analysis. The proportions of the five lineages/sublineages are shown. (B) The total number of  
 730 isolates of each lineage/sublineage per year. (C) Phylogenetic comparison between  
 731 representative strains of *S. Typhimurium* ST19 and four ST313 lineages/sublineages (L1, L2.0,  
 732 L2.2, L2.3) showing the presence and absence of plasmids, prophages and the *spvD*  
 733 pseudogene. The complete phylogenetic analysis of 707 *S. Typhimurium* genomes is shown in  
 734 Fig.S1.

735 **Fig. 2. Key genetic similarities and differences between the chromosome and plasmid**  
 736 **profiles of D23580 (lineage 2) and D37712 (L2.2).** (A) A comparison of the D23580 (L2.0)  
 737 and D37712 (L2.2) chromosomes. The dots around the chromosome are different kinds of  
 738 SNPs identified. Phages and *Salmonella* pathogenicity islands are shown in blue and red  
 739 respectively. (B) Plasmid profile of D37712 versus D23580. The pSLT-BT virulence plasmid is  
 740 present in both D37712 and D23580, and carries the Tn-21 transposable element; (C) pCol1B9  
 741 is present in D37712 and absent from D23580 (D) pBT3 is present in both D37712 and D23580.  
 742 (E) Absence of *sseI* gene and the STM1050 coding sequence in L2.2 (D37712), as compared  
 743 to *S. Typhimurium* ST19 4/74 and *S. Typhimurium* ST313 L2.0 (D23580). (F) List of  
 744 pseudogenes in D37712 and D23580, with reference to 4/74. The colour blue means  
 745 pseudogene/disrupted gene while grey indicates functional genes. *macB* is a pseudogene in  
 746 D23580 (L2.0) but not in L2.2, while *spvD* is a pseudogene in L2.2 but not in L2.0. All L2.2  
 747 strains share similar pseudogenes.

748 **Fig. 3. General comparison of expression profiles of strains 4/74, D23580, and D37712**  
 749 **under 17 different *in vitro* conditions.** (A) Principal component analysis (PCA) plot of the  
 750 individual RNA-seq samples, indicating the overall similarity in gene expression between the  
 751 three strains. The 17 growth conditions have been defined previously (Kröger *et al.*, 2013).  
 752 (B) Visualization of SPI-2 pathogenicity island expression with the Jbrowse genomic browser,  
 753 at mid-exponential phase (MEP), InSPI2, and NonSPI2 *in vitro* conditions, which can be  
 754 accessed [here](#). (C) Boxplot visualization of SPI-2 gene expression at mid-exponential phase  
 755 (MEP), InSPI2, and NonSPI2 *in vitro* conditions. The y-axis shows the combined log TPM  
 756 values for 45 genes located in the SPI2 pathogenicity island, namely *ssaU*, *ssaT*, *ssaS*, *ssaR*,  
 757 *ssaQ*, *ssaP*, *ssaO*, *ssaN*, *ssaV*, *ssaM*, *ssaL*, *ssaK*, *STnc1220*, *STM1410*, *ssaJ*, *ssal*, *ssaH*,  
 758 *ssaG*, *sseG*, *sseF*, *sscB*, *sseE*, *sseD*, *sseC*, *sscA*, *sseB*, *sseA*, *ssaE*, *ssaD*, *ssaC*, *ssaB*, *ssrA*,  
 759 *ssrB*, *orf242*, *orf319*, *orf70*, *ttrR*, *ttrS*, *ttrC*, *ttrB*, *ttrA*, *orf408*, *orf245*, *orf32*, and *orf48*. The  
 760 elevated expression of SPI-2 genes in strain D37712 cultured in NonSPI2 media is  
 761 highlighted in a red box.

1  
2  
3 762 **Fig. 4. Differential gene expression of *S. Typhimurium* 4/74, D37712, and D23580 under**  
4 763 **4 in vitro conditions. (A)** Boxplots indicating the number of differentially-expressed genes  
5 764 identified in the following *in vitro* growth conditions: early stationary phase, ESP; anaerobic  
6 765 growth, NoO<sub>2</sub>; SPI-2 inducing medium, InSPI2; SPI-2 non-inducing minimal medium, NonSPI2.  
7 766 Multiple (3 to 5) biological replicates were used for comparison. DESeq2 was used for  
8 767 differential analysis; only genes with  $|\log_2FC| \geq 1$  and with adjusted  $p$ -value  $\leq 0.001$  were  
9 768 retained. **(B)** Heatmap of the genes differentially expressed between D23580 and D37712.  
10 769 Functional groups and operons of interest are highlighted on the right of Panel B.

11  
12  
13  
14  
15 770 **Fig. 5. Phenotypes that distinguish ST313 L2.2 from ST313 L2.0. (A)** Swimming motility  
16 771 assay of strains D23589, D37712 and 4/74, with a representative plate shown on the left.  
17 772 Average migration diameters were measured after 4 and 8 hours. Each bar represents the  
18 773 mean of three biological replicates, with *error bars* showing standard deviation. Significant  
19 774 difference (\*\*\*) indicates  $P$  value ( $t$  test)  $< 0.001$ . In Panels B & C, comparison of *ssaG*  
20 775 expression by flow cytometry using D23580 and D37712 derivatives containing a chromosomal  
21 776 *ssaG*-GFP<sup>+</sup> transcriptional fusion, strains SZS008 and SZS032, respectively. Cells were  
22 777 collected at 8 hours after inoculation in NonSPI2 media. Ten thousand events were acquired  
23 778 for each sample. **(B)** Mean fluorescent intensity signal of *ssaG*-GFP<sup>+</sup> for D23580 (SZS008, dark  
24 779 grey) and D37712 (SZS032, grey) grown in NonSPI2 media. Significant difference (\*\*\*)  
25 780 indicates  $P$  value ( $t$  test)  $< 0.001$ . **(C)** The proportions of bacterial cells that expressed *ssaG*-  
26 781 GFP<sup>+</sup> during growth in NonSPI2 media was determined. Percentage of GFP-expressing (green)  
27 782 and non-fluorescent cells (white) for D23580 (SZS008) and D37712 (SZS032) is shown. Each  
28 783 bar represents the mean of three biological replicates, error bars show standard deviation.  
29 784 Significant difference (\*\*\*) indicates  $P$  value ( $t$  test)  $< 0.001$ . **(D)** Relative fitness of wild-type  
30 785 D23580 and D37712 and their kanamycin-resistant derivatives. Bacterial numbers were  
31 786 determined after overnight culture of a 1:1 mixture (wild-type versus Km<sup>R</sup>) in LB (left), InSPI2  
32 787 (middle) and NonSPI2 (right) media. Each dot represents the log-transformed mean  
33 788 competitive index of three biological replicates with *error bars* representing 95% confidence  
34 789 interval from standard deviation. A log number higher than 0 reflects the increased fitness of  
35 790 kanamycin-resistant derivatives.  $P$  values were determined by  $t$  test (\*\*\*:  $P < 0.001$ ; \*\*:  $P <$   
36 791  $0.01$ ; \*:  $P < 0.05$ ; ns: not significant).

792

### 793 Supporting information

794 **Fig. S1, Maximum-likelihood phylogeny of 707 African *S. Typhimurium* isolates.** All  
795 genome sequences have been published (Msefula *et al.*, 2012, Pulford *et al.*, 2021, Canals *et*  
796 *al.*, 2019b). Raw sequence reads were aligned to the *S. Typhimurium* D23580 genome  
797 (FN424405) using Snippy. The recombination sites of the alignment were removed by Gubbins,  
798 and the phylogenetic tree was built with Raxml-ng. The tree is rooted on *Salmonella* Typhi strain  
799 CT18 as the outgroup. The MLST sequence types, HierBAPS level 1 and level 2 clusters are  
800 shown in coloured concentric rings as indicated. The *S. Typhimurium* ST313 isolates are

1  
2  
3 801 categorised as Lineage 1, Lineage 2 or Lineage 3 according to HierBAPS level 1 clustering.  
4 802 ST313 Lineage 2 was then sub- divided into 3 sub-lineages according to HierBAPS level 2  
5 803 clustering: ST313 L2.0, ST313 L2.2 and ST313 L2.3. The metadata and lineage designations  
6 804 of all the *S. Typhimurium* isolates are in Table S7.

9 805 **Fig. S2.** PCR-based confirmation of the deletion of the *ssel* gene from *S. Typhimurium* L2.2  
10 806 D37712. Arrows from left to right show the forward strand while the left strand is shown by  
11 807 arrows from right to left. However, *ssel* gene in D23580 is a pseudogene with a SNP  
12 808 indicated as a red line.

15 809 **Fig. S3.** Genomic comparison of plasmids pCol1B9<sup>4/74</sup> and pCol1B9<sup>D37712</sup> using Artemis  
16 810 Comparison Tool (ACT). Bottom panel details the differences observed in the most divergent  
17 811 regions, including colicin toxin-antitoxin system (in pCol1B9) and *impC-umuC-umuD* operon (in  
18 812 pCol1B9).

21 813 **Fig. S4. RDAR Phenotypes of 4/74, D23580, D37712 and BKQZM9.** The top panel shows  
22 814 the RDAR morphology assay and the bottom panel shows a complementary experiment that  
23 815 involves the induction of biofilm formation on 1% tryptone agar (MacKenzie *et al.*, 2019).  
24 816 Strain 4/74 was used as a RDAR-positive control, which has concentric rings and a wrinkled  
25 817 appearance (Pulford *et al.*, 2021). The *S. Typhimurium* ST313 L3 strain BKQZM9 is shown for  
26 818 comparative purposes.

29 819 **Fig. S5. Competitive index analysis of D23580 and D37712 using fluorescently-tagged *S.***  
30 820 **Typhimurium strains (A)** Km<sup>R</sup>-sGFP2 and Gm<sup>R</sup>-mScarlet were inserted into the transposon  
31 821 Tn7 site of D23580 or D37712. Bent arrows represent promoters and directional arrows  
32 822 represent genes. **(B)** A 1:1 mix of Km<sup>R</sup>-sGFP2 and Gm<sup>R</sup>-mScarlet marked strain was inoculated  
33 823 in NonSPI2 media, followed by an overnight incubation in 37°C. The percentage of sGFP2  
34 824 (green) and mScarlet (Red) -marked cells was determined by flow cytometry. Raw data are  
35 825 shown in Figure S7, 10,000 events were acquired for each sample. **(C)** Competitive index  
36 826 analysis of Km<sup>R</sup>-sGFP2 and Gm<sup>R</sup>-mScarlet marked strain. Bacterial numbers were determined  
37 827 by counting CFU for overnight culture of a 1:1 mixture in NonSPI2 media. Each dot represents  
38 828 a single biological replicate and the lane represents mean value. A competitive index of 1  
39 829 indicates the equal fitness of two strains, while a number higher than 1 reflects an increased  
40 830 fitness of D37712.

43 831 **Fig. S6. Raw flow cytometric data related to Fig. S5B. (A)** JH4695 + JH4698 and **(B)**  
44 832 JH4696 + JH4697. A 1:1 mix of the Km<sup>R</sup>-sGFP2 and Gm<sup>R</sup>-mScarlet marked strains were  
45 833 inoculated in NonSPI2 media, followed by growth at 37°C until OD<sub>600</sub> = 0.3. The X-axis  
46 834 (labelled FITC) shows the GFP level and the Y-axis (labelled PE Yell-Gm) indicates the  
47 835 mScarlet level. Quadrant gates were used to separate four populations, and the black  
48 836 numbers indicate the percentage of events in each quadrant. In total, 10,000 events were  
49 837 acquired for each sample.



838 **Fig. S7. The insertion of GFP-Km or RFP-Gm did not impact on fitness.** A 1:1 mix of  
 839 Km<sup>R</sup>-sGFP2 and Gm<sup>R</sup>-mScarlet marked strains were inoculated in LB or NonSPI2 media,  
 840 followed by overnight incubation in 37°C. The competitive index (CI) was calculated using the  
 841 formula  $(CFU_{Gm})/(CFU_{Km})$ . Each dot represents the CI from a single replicate and the  
 842 horizontal bars indicate the mean of each dataset.

#### 843 **Supplementary data**

844 **Table S1:** SNP and indel variants that differentiate L2.2 (strain D37712) and L2.3 (strain  
 845 D49679).

846 **Table S2:** SNP and indel variants that differentiate L2.2 (strain D37712) and L2.0 (strain  
 847 D23580).

848 **Table S3:** Pseudogenes carried by ST19 and ST313 L2.0 and L2.2 (strains 4/74, D23580 and  
 849 D37712).

850 **Table S4:** Raw read counts for all processed RNA-seq samples shown in Figures 3 and 4  
 851 (strains 4/74, D23580, and D37712).

852 **Table S5:** TPM values for all processed RNA-seq samples shown in Figures 3 and 4 (strains  
 853 4/74, D23580, and D37712).

854 **Table S6:** DESeq2-based differential gene expression analysis for strains D23580 vs D37712  
 855 grown in four *in vitro* conditions.

856 **Table S7:** Metadata and lineage designations of the 708 *S. Typhimurium* isolates used to  
 857 generate the maximum likelihood phylogeny (Fig. S1).

858 **Table S8:** Bacterial strains used in this study.

#### 859 **Acknowledgements**

860  
 861 The authors thank Brian Coombes and Rob Kingsley for their constructive comments during  
 862 the peer review process. We are grateful to present and former members of the Hinton  
 863 laboratory for helpful discussions, and to Paul Loughnane for his expert technical assistance.  
 864

865 This work was supported by a Wellcome Trust Investigator award [grant numbers  
 866 106914/Z/15/Z and 222528/Z/21/Z] to J.C.D.H., and by the Malawi-Liverpool-Wellcome  
 867 Research Centre Director's Fund. B.K. was funded by an AESA-RISE fellowship from the  
 868 African Academy of Sciences [Grant Number: RPDF-18-04]. For the purpose of open access,  
 869 the authors have applied a CC BY public copyright licence to any Author Accepted  
 870 Manuscript version arising from this submission.  
 871

#### 872 **Author contributions**

873  
 874 **Conceptualization:** B.K., R.H., M.A.G., C.L.M. and J.C.D.H.  
 875

876 **Data curation:** B.K., R.C., A.V.P., C.V.P., P.A.  
 877

878 **Formal analysis:** B.K., R.C., C.V.P., A.V.P., X.Z., C.K., S.V.O., Y.L., P.A., A.D. and J.C.D.

*Kumwenda et. al.*

25

879

880 **Funding acquisition:** B.K., R.C., A.V.P., X.Z. and J.C.D.H.

881

882 **Investigation:** B.K., R.C.A., A.V.P., X.Z. and J.C.D.H.

883

884 **Methodology:** B.K., R.H., M.A.G., C.L.M. and J.C.D.H.

885

886 **Project administration:** B.K. and J.C.D.H.

887

888 **Resources:** B.K., R.H., M.A.G, C.L.M. and J.C.D.H.

889

890 **Software:** B.K. and A.V.P.

891

892 **Supervision:** M.A.G., C.L.G., C.L.M. and J.C.D.H.

893

894 **Validation:** B.K., R.C., A.V.P., X.Z. and J.C.D.H.

895

896 **Visualization:** B.K., R.C., Y.L., C.V.P., A.V.P. and J.C.D.H.

897

898 **Writing original draft:** B.K., R.C. and J.C.D.H

899

900 **Writing reviews and editing:** B.K., R.C., A.V.P., X.Z., C.K., S.V.O., A.D., R.H., M.G and  
901 J.C.D.H.

902

903 **Equal contribution:** Authors B.K., R.C. and A.V.P. made equal contributions to this work.

904

**References**

- 905  
906
- 907 Ackermann M, Stecher B, Freed N, Songhet P, Hardt W, and Doebeli M (2008)  
908 Self-destructive cooperation mediated by phenotypic noise. *454:987–990*.
- 909 Andrews S (2010) FastQC: a quality control tool for high throughput sequence  
910 data. Available online at:  
911 <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- 912 Aulicino A, Antanaviciute A, Frost J, Sousa Geros A, Mellado E, Attar M,  
913 Jagielowicz M, Hublitz P, Sinz J, Preciado-Llanes L, Napolitani G, Bowden R,  
914 Koohy H, Drakesmith H, and Simmons A (2022) Dual RNA sequencing reveals  
915 dendritic cell reprogramming in response to typhoidal *Salmonella* invasion.  
916 *Commun Biol* 5:111.
- 917 Bogomolnaya LM, Santiviago CA, Yang H-J, Baumler AJ, and Andrews-  
918 Polymenis HL (2008) 'Form variation' of the O12 antigen is critical for  
919 persistence of *Salmonella* Typhimurium in the murine intestine. *Mol Microbiol*  
920 70:1105–1119.
- 921 Bolger AM, Lohse M, and Usadel B (2014) Trimmomatic: a flexible trimmer for  
922 Illumina sequence data. *Bioinformatics* 30:2114–2120.
- 923 Branchu P, Bawn M, and Kingsley RA (2018) Genome Variation and Molecular  
924 Epidemiology of *Salmonella enterica* Serovar Typhimurium Pathovariants.  
925 *Infect Immun* 86:e00079-18.
- 926 Brink T, Leiss V, Siegert P, Jehle D, Ebner JK, Schwan C, Shymanets A, Wiese  
927 S, Nürnberg B, Hensel M, Aktories K, and Orth JHC (2018) *Salmonella*  
928 Typhimurium effector Ssel inhibits chemotaxis and increases host cell survival  
929 by deamidation of heterotrimeric Gi proteins. *PLOS Pathog* 14:e1007248.
- 930 Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, and  
931 Madden TL (2009) BLAST+: architecture and applications. *BMC Bioinformatics*  
932 10:421.
- 933 Canals R, Chaudhuri RR, Steiner RE, Owen SV, Quinones-Olvera N, Gordon  
934 MA, Baym M, Ibba M, and Hinton JCD (2019a) The fitness landscape of the  
935 African *Salmonella* Typhimurium ST313 strain D23580 reveals unique  
936 properties of the pBT1 plasmid. *PLOS Pathog* 15:e1007948.
- 937 Canals R, Hammarlöf DL, Kröger C, Owen SV, Fong WY, Lacharme-Lora L,  
938 Zhu X, Wenner N, Carden SE, Honeycutt J, Monack DM, Kingsley RA,  
939 Brownridge P, Chaudhuri RR, Rowe WPM, Predeus AV, Hokamp K, Gordon  
940 MA, and Hinton JCD (2019b) Adding function to the genome of African  
941 *Salmonella* Typhimurium ST313 strain D23580. *PLOS Biol* 17:e3000059.
- 942 Carden SE, Walker GT, Honeycutt J, Lugo K, Pham T, Jacobson A, Bouley D,  
943 Idoyaga J, Tsolis RM, and Monack D (2017) Pseudogenization of the Secreted



- 1  
2  
3 944 Effector Gene *ssel* Confers Rapid Systemic Dissemination of *S. Typhimurium*  
4 945 ST313 within Migratory Dendritic Cells. *Cell Host Microbe* 21:182–194.  
5  
6 946 Chirwa E, Dale, H, Gordon, MA, and Ashton, PM (2023) What is the Source of  
7 947 Infections Causing Invasive Nontyphoidal Salmonella Disease? *Open Forum*  
8 948 *Infect Diseases* February 2023.  
9  
10  
11 949 Crump JA, Sjölund-Karlsson M, Gordon MA, and Parry CM (2015)  
12 950 Epidemiology, Clinical Presentation, Laboratory Diagnosis, Antimicrobial  
13 951 Resistance, and Antimicrobial Management of Invasive Salmonella Infections.  
14 952 *Clin Microbiol Rev* 28:901–937.  
15  
16 953 Ewels P, Magnusson M, Lundin S, and Käller M (2016) MultiQC: summarize  
17 954 analysis results for multiple tools and samples in a single report. *Bioinformatics*  
18 955 32:3047–3048.  
19  
20  
21 956 Fass E, and Groisman E (2009) Control of Salmonella pathogenicity island-2  
22 957 gene expression. *Curr Opin Microbiol* 12:199–204.  
23  
24 958 Feasey NA, Dougan G, Kingsley RA, Heyderman RS, and Gordon MA (2012)  
25 959 Invasive non-typhoidal salmonella disease: an emerging and neglected tropical  
26 960 disease in Africa. *The Lancet* 379:2489–2499.  
27  
28  
29 961 Feasey NA, Hadfield J, Keddy KH, Dallman TJ, Jacobs J, Deng X, Wigley P,  
30 962 Barquist L, Langridge GC, Feltwell T, Harris SR, Mather AE, Fookes M, Aslett  
31 963 M, Msefula C, Kariuki S, Maclennan CA, Onsare RS, Weill F-X, Le Hello S,  
32 964 Smith AM, McClelland M, Desai P, Parry CM, Cheesbrough J, French N,  
33 965 Campos J, Chabalgoity JA, Betancor L, Hopkins KL, Nair S, Humphrey TJ,  
34 966 Lunguya O, Cogan TA, Tapia MD, Sow SO, Tennant SM, Bornstein K, Levine  
35 967 MM, Lacharme-Lora L, Everett DB, Kingsley RA, Parkhill J, Heyderman RS,  
36 968 Dougan G, Gordon MA, and Thomson NR (2016) Distinct Salmonella Enteritidis  
37 969 lineages associated with enterocolitis in high-income settings and invasive  
38 970 disease in low-income settings. *Nat Genet* 48:1211–1217.  
39  
40  
41 971 Gilchrist J, and MacLennan C (2019) Invasive Nontyphoidal Salmonella  
42 972 Disease in Africa. *Cell Mol Biol E Coli Salmonella Enterobact*, doi: doi:10.1128/  
43 973 ecosalplus.ESP-0007-2018.  
44  
45  
46 974 Grabe GJ, Zhang Y, Przydacz M, Rolhion N, Yang Y, Pruneda JN, Komander  
47 975 D, Holden DW, and Hare SA (2016) The Salmonella Effector SpvD Is a  
48 976 Cysteine Hydrolase with a Serovar-specific Polymorphism Influencing Catalytic  
49 977 Activity, Suppression of Immune Responses, and Bacterial Virulence. *J Biol*  
50 978 *Chem* 291:25853–25863.  
51  
52  
53 979 Gurevich A, Saveliev V, Vyahhi N, and Tesler G (2013) QUASt: quality  
54 980 assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075.  
55  
56 981 Hammarlöf DL, Kröger C, Owen SV, Canals R, Lacharme-Lora L, Wenner N,  
57 982 Schager AE, Wells TJ, Henderson IR, Wigley P, Hokamp K, Feasey NA,  
58 983 Gordon MA, and Hinton JCD (2018) Role of a single noncoding nucleotide in  
59  
60

- 1  
2  
3 984 the evolution of an epidemic African clade of *Salmonella*. *Proc Natl Acad Sci*  
4 985 115.
- 6 986 Hautefort I, Proença MJ, and Hinton JCD (2003) Single-Copy Green  
7 987 Fluorescent Protein Gene Fusions Allow Accurate Measurement of *Salmonella*  
8 988 Gene Expression In Vitro and during Infection of Mammalian Cells. *Appl Environ*  
9 989 *Microbiol* 69:7480–7491.
- 12 990 Honeycutt JD, Wenner N, Li Y, Brewer SM, Massis LM, Brubaker SW,  
13 991 Chairatana P, Owen SV, Canals R, Hinton JCD, and Monack DM (2020)  
14 992 Genetic variation in the MacAB-TolC efflux pump influences pathogenesis of  
15 993 invasive *Salmonella* isolates from Africa. *PLOS Pathog* 16:e1008763.
- 18 994 Jennings E, Thurston TLM, and Holden DW (2017) *Salmonella* SPI-2 Type III  
19 995 Secretion System Effectors: Molecular Mechanisms And Physiological  
20 996 Consequences. *Cell Host Microbe* 22:217–231.
- 22 997 Johnson R, and Simon M (1985) Hin-mediated site-specific recombination  
23 998 requires two 26 bp recombination sites and a 60 bp recombinational enhancer.  
24 999 *Cell* 41:781–791.
- 26 1000 Kariuki S, Revathi G, Kariuki N, Kiiru J, Mwituria J, Muyodi J, Githinji JW,  
27 1001 Kagendo D, Munyalo A, and Hart CA (2006) Invasive multidrug-resistant non-  
28 1002 typhoidal *Salmonella* infections in Africa: zoonotic or anthroponotic  
29 1003 transmission? *J Med Microbiol* 55:585–591.
- 32 1004 Kasumba IN, Pulford CV, Perez-Sepulveda BM, Sen S, Sayed N, Permala-  
33 1005 Booth J, Livio S, Heavens D, Low R, Hall N, Roose A, Powell H, Farag T,  
34 1006 Panchalingham S, Berkeley L, Nasrin D, Blackwelder WC, Wu Y, Tamboura B,  
35 1007 Sanogo D, Onwuchekwa U, Sow SO, Ochieng JB, Omore R, Oundo JO,  
36 1008 Breiman RF, Mintz ED, O'Reilly CE, Antonio M, Saha D, Hossain MJ,  
37 1009 Mandomando I, Bassat Q, Alonso PL, Ramamurthy T, Sur D, Qureshi S, Zaidi  
38 1010 AKM, Hossain A, Faruque ASG, Nataro JP, Kotloff KL, Levine MM, Hinton JCD,  
39 1011 and Tennant SM (2021) Characteristics of *Salmonella* Recovered From Stools  
40 1012 of Children Enrolled in the Global Enteric Multicenter Study. *Clin Infect Dis*  
41 1013 73:631–641.
- 44 1014 Kingsley RA, Msefula CL, Thomson NR, Kariuki S, Holt KE, Gordon MA, Harris  
45 1015 D, Clarke L, Whitehead S, Sangal V, Marsh K, Achtman M, Molyneux ME,  
46 1016 Cormican M, Parkhill J, MacLennan CA, Heyderman RS, and Dougan G (2009)  
47 1017 Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive  
48 1018 disease in sub-Saharan Africa have a distinct genotype. *Genome Res* 19:2279–  
49 1019 2287.
- 52 1020 Kirk MD, Pires SM, Black RE, Caipo M, Crump JA, Devleeschauwer B, Döpfer  
53 1021 D, Fazil A, Fischer-Walker CL, Hald T, Hall AJ, Keddy KH, Lake RJ, Lanata CF,  
54 1022 Torgerson PR, Havelaar AH, and Angulo FJ (2015) World Health Organization  
55 1023 Estimates of the Global and Regional Disease Burden of 22 Foodborne  
56 1024 Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. *PLOS Med*  
57 1025 12:e1001921.
- 59  
60

- 1  
2  
3 1026 Koolman L, Prakash R, Diness Y, Msefula C, Nyirenda TS, Olgemoeller F,  
4 1027 Perez-Sepulveda B, Hinton JCD, Owen SV, Feasey NA, Ashton PM, and  
5 1028 Gordon MA (2022) *Case-control investigation of invasive Salmonella disease*  
6 1029 *in Africa – comparison of human, animal and household environmental isolates*  
7 1030 *find no evidence of environmental or animal reservoirs of invasive strains,*  
8 1031 *Infectious Diseases (except HIV/AIDS).*
- 10  
11 1032 Kröger C, Colgan A, Srikumar S, Händler K, Sivasankaran SK, Hammarlöf DL,  
12 1033 Canals R, Grissom JE, Conway T, Hokamp K, and Hinton JCD (2013) An  
13 1034 Infection-Relevant Transcriptomic Compendium for *Salmonella enterica*  
14 1035 Serovar Typhimurium. *Cell Host Microbe* 14:683–695.
- 16  
17 1036 Kröger C, Dillon SC, Cameron ADS, Papenfort K, Sivasankaran SK, Hokamp  
18 1037 K, Chao Y, Sittka A, Hébrard M, Händler K, Colgan A, Leekitcharoenphon P,  
19 1038 Langridge GC, Lohan AJ, Loftus B, Lucchini S, Ussery DW, Dorman CJ,  
20 1039 Thomson NR, Vogel J, and Hinton JCD (2012) The transcriptional landscape  
21 1040 and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc Natl Acad*  
22 1041 *Sci* 109.
- 24  
25 1042 Lai MA, Quarles EK, López-Yglesias AH, Zhao X, Hajjar AM, and Smith KD  
26 1043 (2013) Innate Immune Detection of Flagellin Positively and Negatively  
27 1044 Regulates Salmonella Infection. *PLoS ONE* 8:e72047.
- 29  
30 1045 Letunic I, and Bork P (2006) Interactive Tree Of Life (iTOL): an online tool for  
31 1046 phylogenetic tree display and annotation. *Bioinformatics* 23:1283–1287.
- 33  
34 1047 Lian ZJ, Phan M, Hancock SJ, Nhu NTK, Paterson DL, and Schembri MA  
35 1048 (2023) Genetic basis of I-complex plasmid stability and conjugation. *PLOS*  
36 1049 *Genet* 19:e1010773.
- 37  
38 1050 Löber S, Jäckel D, Kaiser N, and Hensel M (2006) Regulation of *Salmonella*  
39 1051 pathogenicity island 2 genes by independent environmental signals. *Int J Med*  
40 1052 *Microbiol* 296:435–447.
- 42  
43 1053 Lucchini S, Rowley G, Goldberg MD, Hurd D, Harrison M, and Hinton JCD  
44 1054 (2006) H-NS Mediates the Silencing of Laterally Acquired Genes in Bacteria.  
45 1055 *PLoS Pathog* 2:e81.
- 47  
48 1056 Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, Jones TF,  
49 1057 Fazil A, and Hoekstra RM (2010) The Global Burden of Nontyphoidal  
50 1058 *Salmonella* Gastroenteritis. *Clin Infect Dis* 50:882–889.
- 52  
53 1059 Marchello C, Dale, A. P., Pisharody, S., Rubach, M. P., and Crump, J.A. (2019)  
54 1060 A Systematic Review and Meta-analysis of the Prevalence of Community-  
55 1061 Onset Bloodstream Infections among Hospitalized Patients in Africa and Asia.  
56 1062 *Antimicrobial Agents and Chemotherapy*. 64.
- 58  
59 1063 Marchello CS, Dale AP, Pisharody S, Rubach MP, and Crump JA (2019) A  
60 1064 Systematic Review and Meta-analysis of the Prevalence of Community-Onset  
61 1065 Bloodstream Infections among Hospitalized Patients in Africa and Asia.  
62 1066 *Antimicrob Agents Chemother* 64:e01974-19.

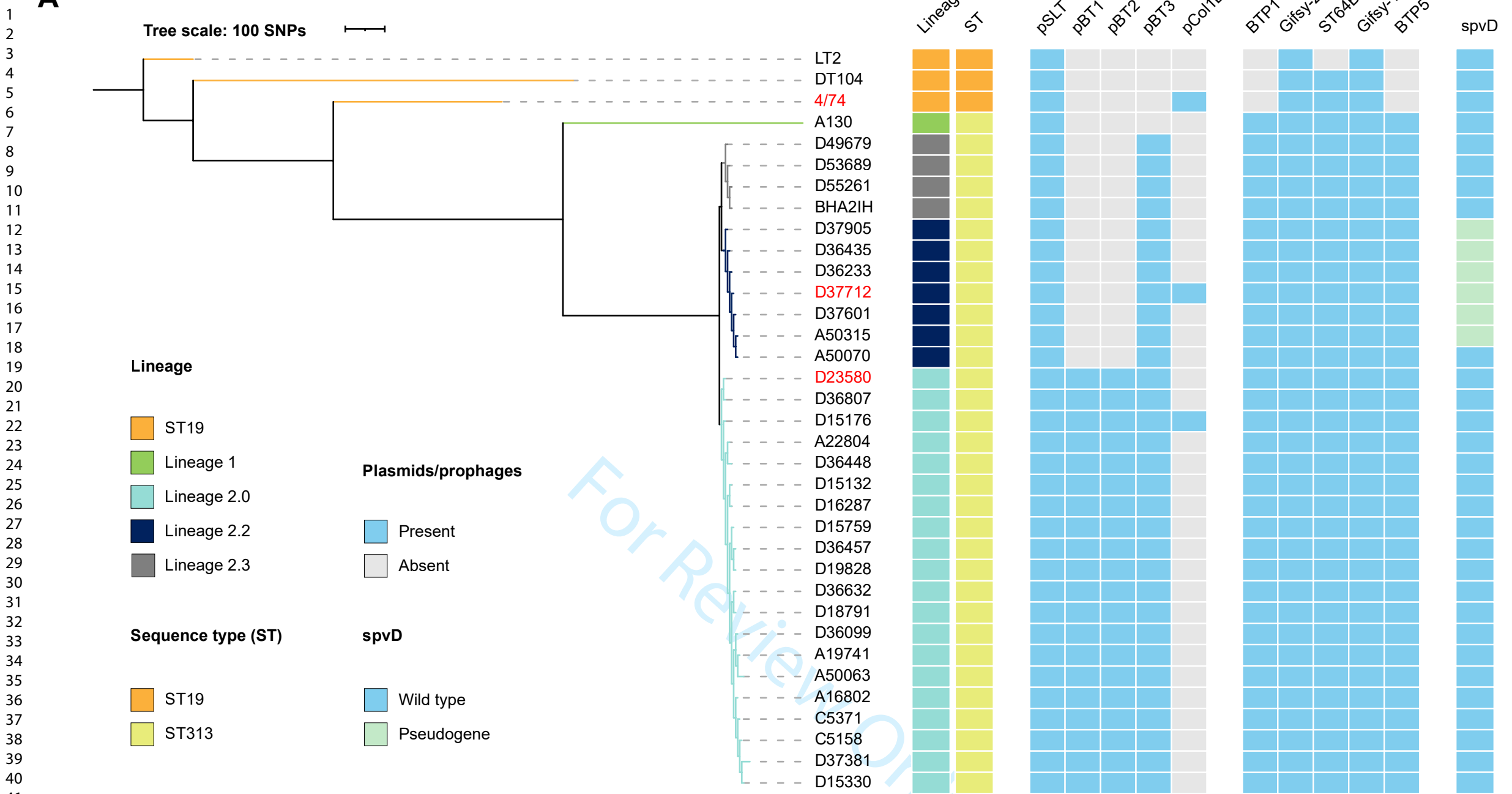
- 1  
2  
3 1067 Msefula CL, Kingsley RA, Gordon MA, Molyneux E, Molyneux ME, MacLennan  
4 1068 CA, Dougan G, and Heyderman RS (2012) Genotypic Homogeneity of  
5 1069 Multidrug Resistant *S. Typhimurium* Infecting Distinct Adult and Childhood  
6 1070 Susceptibility Groups in Blantyre, Malawi. *PLoS ONE* 7:e42085.
- 7  
8  
9 1071 Musicha P, Cornick JE, Bar-Zeev N, French N, Masesa C, Denis B, Kennedy  
10 1072 N, Mallewa J, Gordon MA, Msefula CL, Heyderman RS, Everett DB, and  
11 1073 Feasey NA (2017) Trends in antimicrobial resistance in bloodstream infection  
12 1074 isolates at a large urban hospital in Malawi (1998–2016): a surveillance study.  
13 1075 *Lancet Infect Dis* 17:1042–1052.
- 14  
15  
16 1076 Nedialkova LP, Denzler R, Koepfel MB, Diehl M, Ring D, Wille T, Gerlach RG,  
17 1077 and Stecher B (2014) Inflammation Fuels Colicin Ib-Dependent Competition of  
18 1078 *Salmonella* Serovar Typhimurium and *E. coli* in Enterobacterial Blooms. *PLoS*  
19 1079 *Pathog* 10:e1003844.
- 20  
21 1080 Okoro CK, Barquist L, Connor TR, Harris SR, Clare S, Stevens MP, Arends MJ,  
22 1081 Hale C, Kane L, Pickard DJ, Hill J, Harcourt K, Parkhill J, Dougan G, and  
23 1082 Kingsley RA (2015) Signatures of Adaptation in Human Invasive *Salmonella*  
24 1083 Typhimurium ST313 Populations from Sub-Saharan Africa. *PLoS Negl Trop Dis*  
25 1084 9:e0003611.
- 26  
27  
28 1085 Okoro CK, Kingsley RA, Connor TR, Harris SR, Parry CM, Al-Mashhadani MN,  
29 1086 Kariuki S, Msefula CL, Gordon MA, de Pinna E, Wain J, Heyderman RS, Obaro  
30 1087 S, Alonso PL, Mandomando I, MacLennan CA, Tapia MD, Levine MM, Tennant  
31 1088 SM, Parkhill J, and Dougan G (2012) Intracontinental spread of human invasive  
32 1089 *Salmonella* Typhimurium pathovariants in sub-Saharan Africa. *Nat Genet*  
33 1090 44:1215–1221.
- 34  
35  
36 1091 Owen SV, Wenner N, Canals R, Makumi A, Hammarlöf DL, Gordon MA,  
37 1092 Aertsen A, Feasey NA, and Hinton JCD (2017) Characterization of the  
38 1093 Prophage Repertoire of African *Salmonella* Typhimurium ST313 Reveals High  
39 1094 Levels of Spontaneous Induction of Novel Phage BTP1. *Front Microbiol* 8.
- 40  
41 1095 Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M,  
42 1096 Falush D, Keane JA, and Parkhill J (2015) Roary: rapid large-scale prokaryote  
43 1097 pan genome analysis. *Bioinformatics* 31:3691–3693.
- 44  
45  
46 1098 Perez-Sepulveda BM, and Hinton JCD (2018) Functional Transcriptomics for  
47 1099 Bacterial Gene Detectives. *Microbiol Spectr* 6:6.5.06.
- 48  
49 1100 Piccini G, and Montomoli E (2020) Pathogenic signature of invasive non-  
50 1101 typhoidal *Salmonella* in Africa: implications for vaccine development. *Hum*  
51 1102 *Vaccines Immunother* 16:2056–2071.
- 52  
53  
54 1103 Post AS, Diallo SN, Guiraud I, Lompo P, Tahita MC, Maltha J, Van Puyvelde S,  
55 1104 Mattheus W, Ley B, Thriemer K, Rouamba E, Derra K, Deborggraeve S, Tinto  
56 1105 H, and Jacobs J (2019) Supporting evidence for a human reservoir of invasive  
57 1106 non-Typhoidal *Salmonella* from household samples in Burkina Faso. *PLoS*  
58 1107 *Negl Trop Dis* 13:e0007782.
- 59  
60



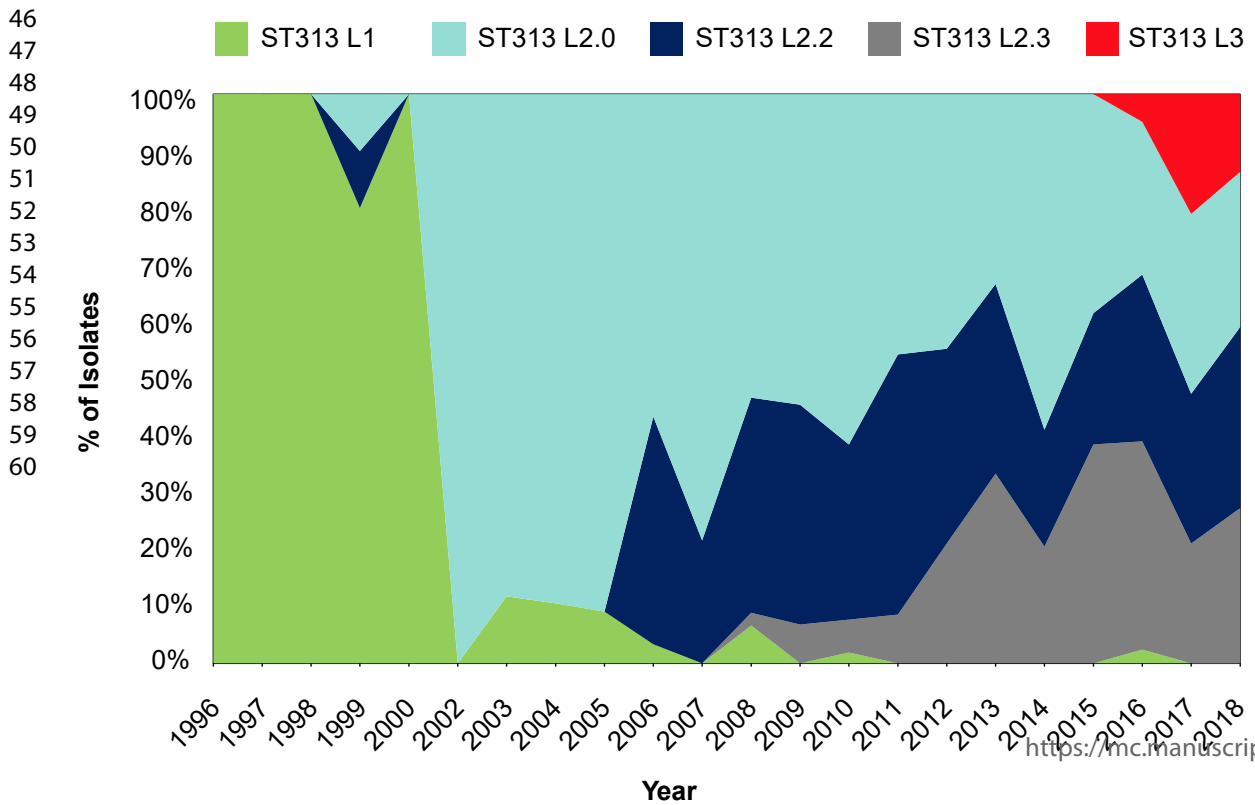
- 1  
2  
3 1108 Preciado-Llanes L, Aulicino A, Canals R, Moynihan PJ, Zhu X, Jambo N,  
4 1109 Nyirenda TS, Kadwala I, Sousa Gerós A, Owen SV, Jambo KC, Kumwenda B,  
5 1110 Veerapen N, Besra GS, Gordon MA, Hinton JCD, Napolitani G, Salio M, and  
6 1111 Simmons A (2020) Evasion of MAIT cell recognition by the African *Salmonella*  
7 1112 Typhimurium ST313 pathovar that causes invasive disease. *Proc Natl Acad Sci*  
8 1113 117:20717–20728.
- 9  
10  
11 1114 Pulford CV, Perez-Sepulveda BM, Canals R, Bevington JA, Bengtsson RJ,  
12 1115 Wenner N, Rodwell EV, Kumwenda B, Zhu X, Bennett RJ, Stenhouse GE,  
13 1116 Malaka De Silva P, Webster HJ, Bengoechea JA, Dumigan A, Tran-Dien A,  
14 1117 Prakash R, Banda HC, Alufandika L, Mautanga MP, Bowers-Barnard A,  
15 1118 Beliavskaia AY, Predeus AV, Rowe WPM, Darby AC, Hall N, Weill F-X, Gordon  
16 1119 MA, Feasey NA, Baker KS, and Hinton JCD (2021) Stepwise evolution of  
17 1120 *Salmonella* Typhimurium ST313 causing bloodstream infection in Africa. *Nat*  
18 1121 *Microbiol* 6:327–338.
- 19  
20  
21 1122 Rankin, J.D. and Taylor R.J. (1966) The estimation of doses of *Salmonella*  
22 1123 typhimurium suitable for the experimental production of disease in calves. *Vec*  
23 1124 *Rec* 78:706–7.
- 24  
25  
26 1125 Richardson EJ, Limaye B, Inamdar H, Datta A, Manjari KS, Pullinger GD,  
27 1126 Thomson NR, Joshi RR, Watson M, and Stevens MP (2011) Genome  
28 1127 Sequences of *Salmonella enterica* Serovar Typhimurium, Choleraesuis, Dublin,  
29 1128 and Gallinarum Strains of Well- Defined Virulence in Food-Producing Animals.  
30 1129 *J Bacteriol* 193:3162–3163.
- 31  
32  
33 1130 Schlechter R, and Remus-Emsermann M (2019) Delivering &quot;Chromatic  
34 1131 Bacteria&quot; Fluorescent Protein Tags to Proteobacteria Using Conjugation.  
35 1132 *BIO-Protoc* 9.
- 36  
37 1133 Sikand A, Jaszczur M, Bloom LB, Woodgate R, Cox MM, and Goodman MF  
38 1134 (2021) The SOS Error-Prone DNA Polymerase V Mutasome and  $\beta$ -Sliding  
39 1135 Clamp Acting in Concert on Undamaged DNA and during Translesion  
40 1136 Synthesis. *Cells* 10:1083.
- 41  
42  
43 1137 Skidmore PD, Canals R, and Ramasamy MN (2023). The iNTS-GMMA vaccine:  
44 1138 a promising step in non-typhoidal *Salmonella* vaccine development. *Expert Rev*  
45 1139 *Vaccines* (In Press) doi: 10.1080/14760584.2023.2270596.
- 46 1140  
47 1141 Srikumar S, Kröger C, Hébrard M, Colgan A, Owen SV, Sivasankaran SK,  
48 1142 Cameron ADS, Hokamp K, and Hinton JCD (2015) RNA-seq Brings New  
49 1143 Insights to the Intra-Macrophage Transcriptome of *Salmonella* Typhimurium.  
50 1144 *PLOS Pathog* 11:e1005262.
- 51  
52  
53 1145 Stamatakis A, Ludwig T, and Meier H (2005) RAxML-III: a fast program for  
54 1146 maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics*  
55 1147 21:456–463.
- 56  
57 1148 Stanaway JD, Parisi A, Sarkar K, Blacker BF, Reiner RC, Hay SI, Nixon MR,  
58 1149 Dolecek C, James SL, Mokdad AH, Abebe G, Ahmadian E, Alahdab F,  
59 1150 Alemnew BTT, Alipour V, Allah Bakeshei F, Animut MD, Ansari F, Arabloo J,

- 1  
2  
3 1151 Asfaw ET, Bagherzadeh M, Bassat Q, Belayneh YMM, Carvalho F, Daryani A,  
4 1152 Demeke FM, Demis ABB, Dubey M, Duken EE, Dunachie SJ, Eftekhari A,  
5 1153 Fernandes E, Fouladi Fard R, Gedefaw GA, Geta B, Gibney KB, Hasanzadeh  
6 1154 A, Hoang CL, Kasaeian A, Khater A, Kidanemariam ZT, Lakew AM,  
7 1155 Malekzadeh R, Melese A, Mengistu DT, Mestrovic T, Miazgowski B,  
8 1156 Mohammad KA, Mohammadian M, Mohammadian-Hafshejani A, Nguyen CT,  
9 1157 Nguyen LH, Nguyen SH, Nirayo YL, Olagunju AT, Olagunju TO, Pourjafar H,  
10 1158 Qorbani M, Rabiee M, Rabiee N, Rafay A, Rezapour A, Samy AM, Sepanlou  
11 1159 SG, Shaikh MA, Sharif M, Shigematsu M, Tessema B, Tran BX, Ullah I, Yimer  
12 1160 EM, Zaidi Z, Murray CJL, and Crump JA (2019) The global burden of non-  
13 1161 typhoidal salmonella invasive disease: a systematic analysis for the Global  
14 1162 Burden of Disease Study 2017. *Lancet Infect Dis* 19:1312–1324.
- 15  
16  
17  
18 1163 Tack B, Phoba M-F, Barbé B, Kalonji LM, Hardy L, Van Puyvelde S, Ingelbeen  
19 1164 B, Falay D, Ngonda D, van der Sande MAB, Deborggraeve S, Jacobs J, and  
20 1165 Lunguya O (2020) Non-typhoidal Salmonella bloodstream infections in Kisantu,  
21 1166 DR Congo: Emergence of O5-negative Salmonella Typhimurium and extensive  
22 1167 drug resistance. *PLoS Negl Trop Dis* 14:e0008121.
- 23  
24  
25 1168 Van Puyvelde S, Pickard D, Vandelannoote K, Heinz E, Barbé B, de Block T,  
26 1169 Clare S, Coomber EL, Harcourt K, Sridhar S, Lees EA, Wheeler NE, Klemm  
27 1170 EJ, Kuijpers L, Mbuyi Kalonji L, Phoba M-F, Falay D, Ngbonda D, Lunguya O,  
28 1171 Jacobs J, Dougan G, and Deborggraeve S (2019) An African Salmonella  
29 1172 Typhimurium ST313 sublineage with extensive drug-resistance and signatures  
30 1173 of host adaptation. *Nat Commun* 10:4280.
- 31  
32  
33 1174 Van Puyvelde S, de Block T, Sridhar S, Bawn M, Kingsley RA, Ingelbeen B,  
34 1175 Beale MA, Barbé B, Jeon HJ, Mbuyi-Kalonji L, Phoba MF, Falay D, Martiny D,  
35 1176 Vandenberg O, Affolabi D, Rutanga JP, Ceysens PJ, Mattheus W, Cuypers  
36 1177 WL, van der Sande MAB, Park SE, Kariuki S, Otieno K, Lusingu JPA, Mbwana  
37 1178 JR, Adjei S, Sarfo A, Agyei SO, Asante KP, Otieno W, Otieno L, Tahita MC,  
38 1179 Lompo P, Hoffman IF, Mvalo T, Msefula C, Hassan-Hanga F, Obaro S,  
39 1180 Mackenzie G, Deborggraeve S, Feasey N, Marks F, MacLennan CA, Thomson  
40 1181 NR, Jacobs J, Dougan G, Kariuki S, Lunguya O. (2023) A genomic appraisal of  
41 1182 invasive Salmonella Typhimurium and associated antibiotic resistance in sub-  
42 1183 Saharan Africa. *Nat Commun*. 14:6392.
- 43  
44  
45 1184 Wick RR, Judd LM, Gorrie CL, and Holt KE (2017) Unicycler: Resolving  
46 1185 bacterial genome assemblies from short and long sequencing reads. *PLOS*  
47 1186 *Comput Biol* 13:e1005595.
- 48  
49  
50 1187 Zhang S, Kingsley RA, Santos RL, Andrews-Polymenis H, Raffatellu M,  
51 1188 Figueiredo J, Nunes J, Tsolis RM, Adams LG, and Bäumlér AJ (2003)  
52 1189 Molecular Pathogenesis of *Salmonella enterica* Serotype Typhimurium-  
53 1190 Induced Diarrhea. *Infect Immun* 71:1–12.
- 54  
55 1191  
56  
57  
58  
59  
60

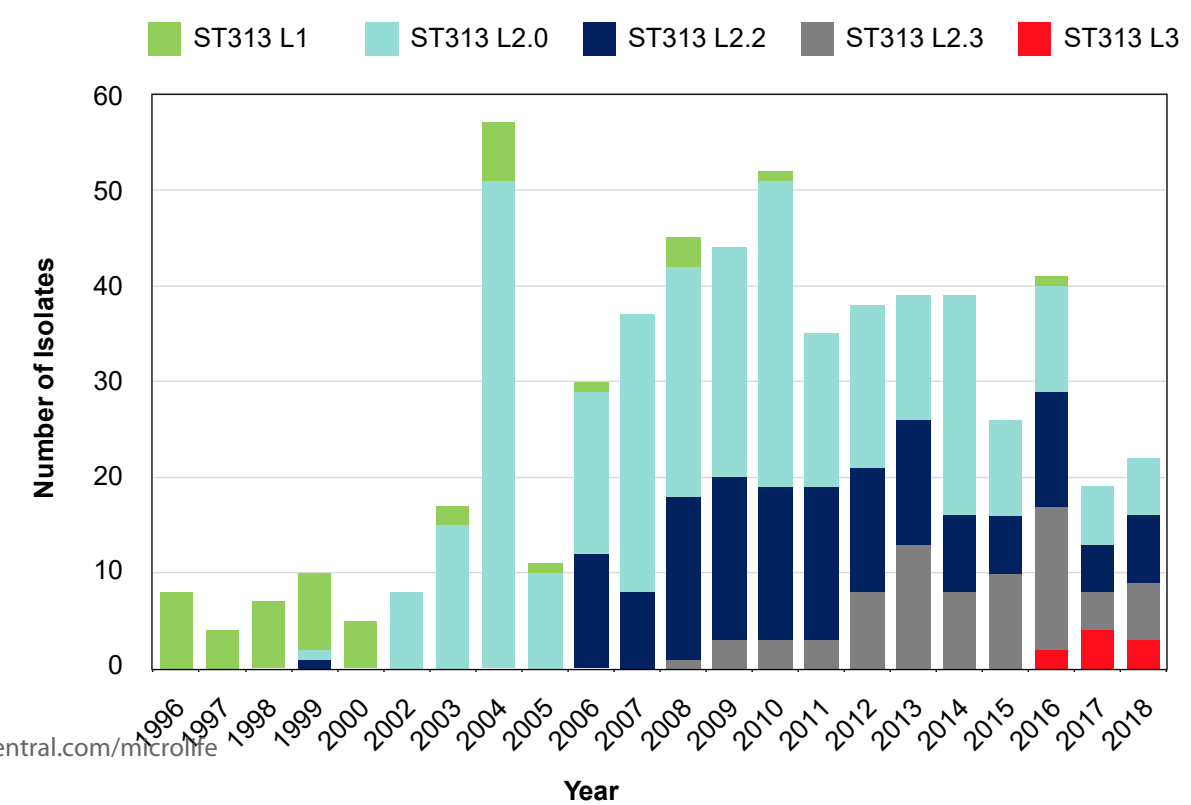
**A**



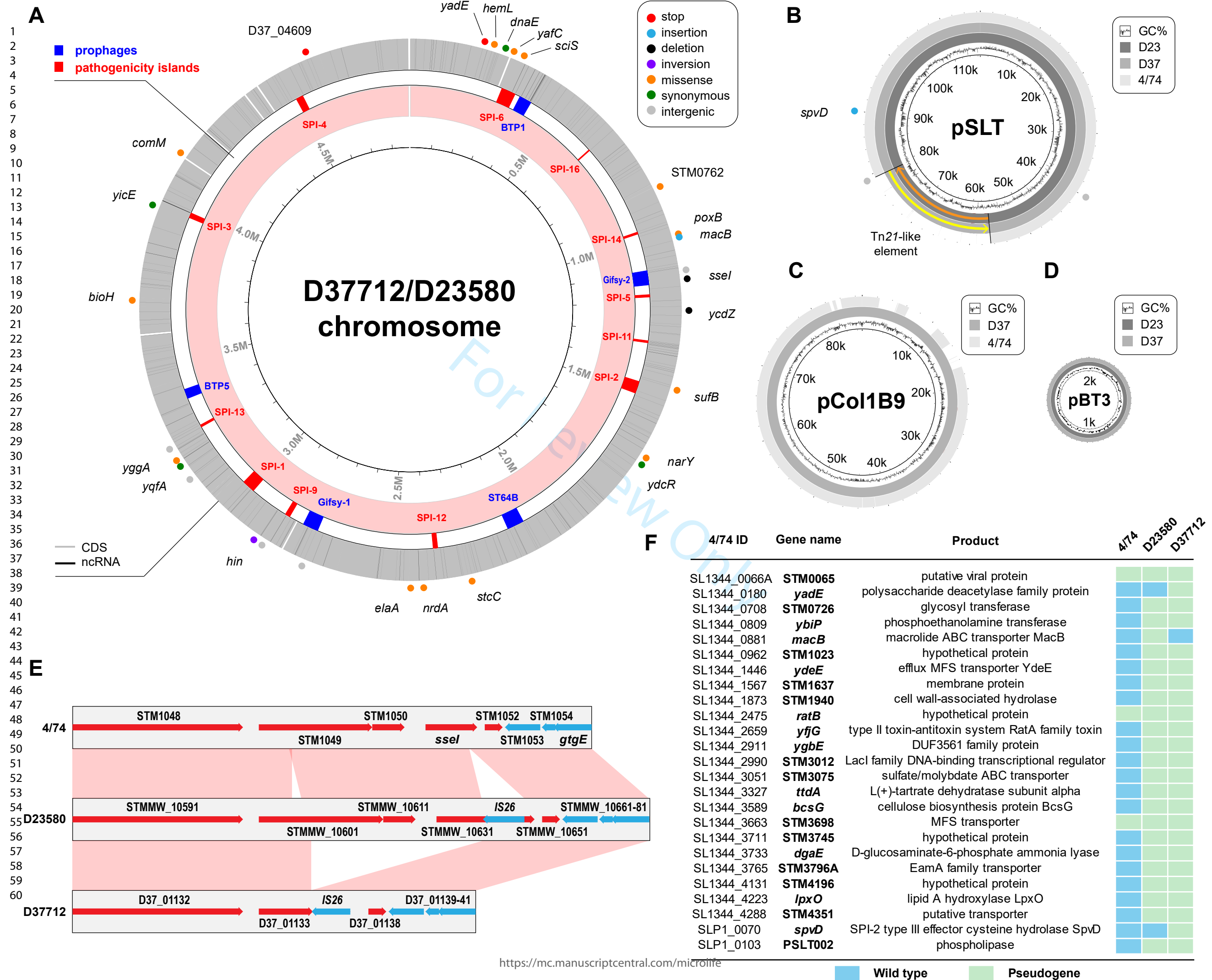
**B**



**C**

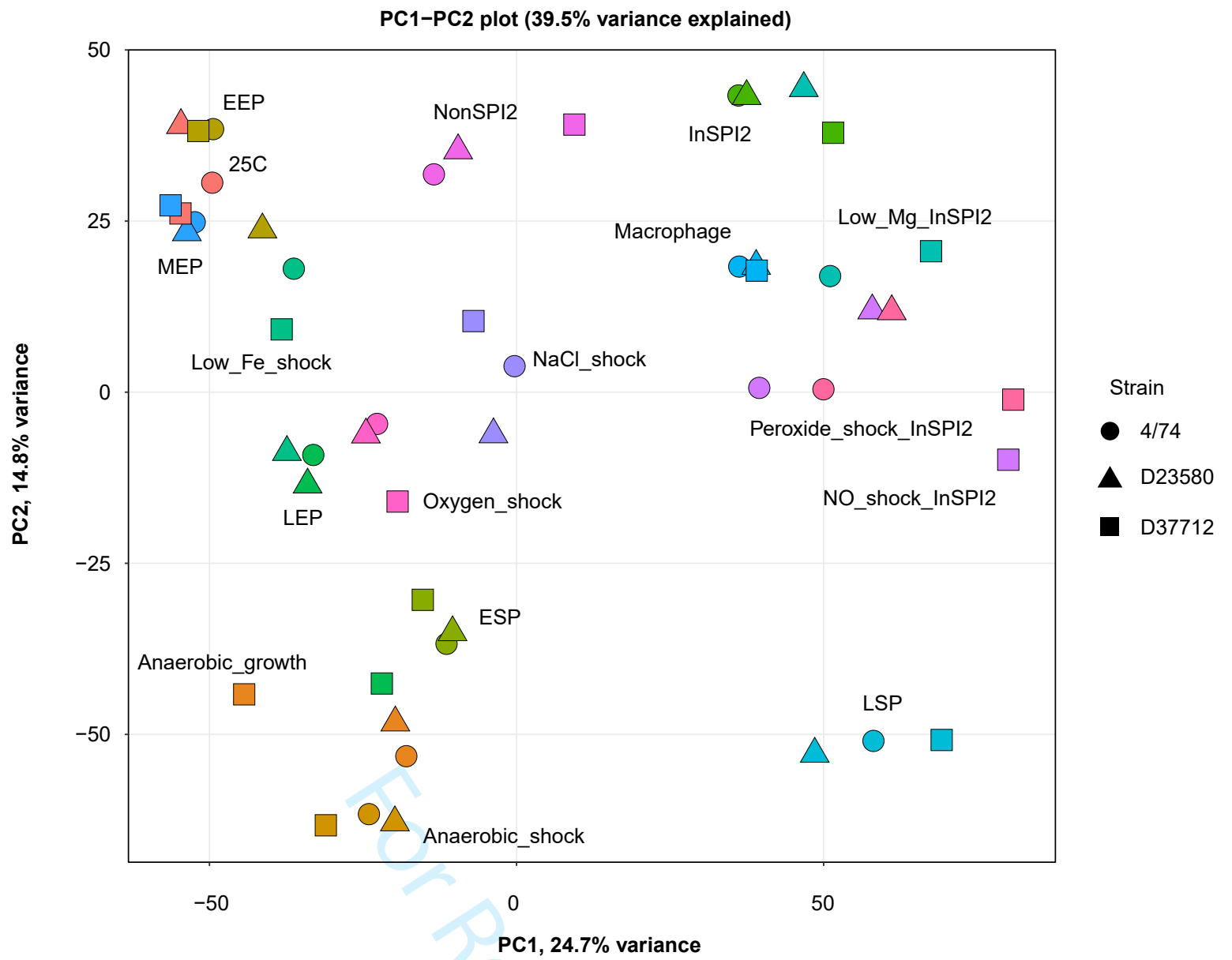


**Fig. 2**

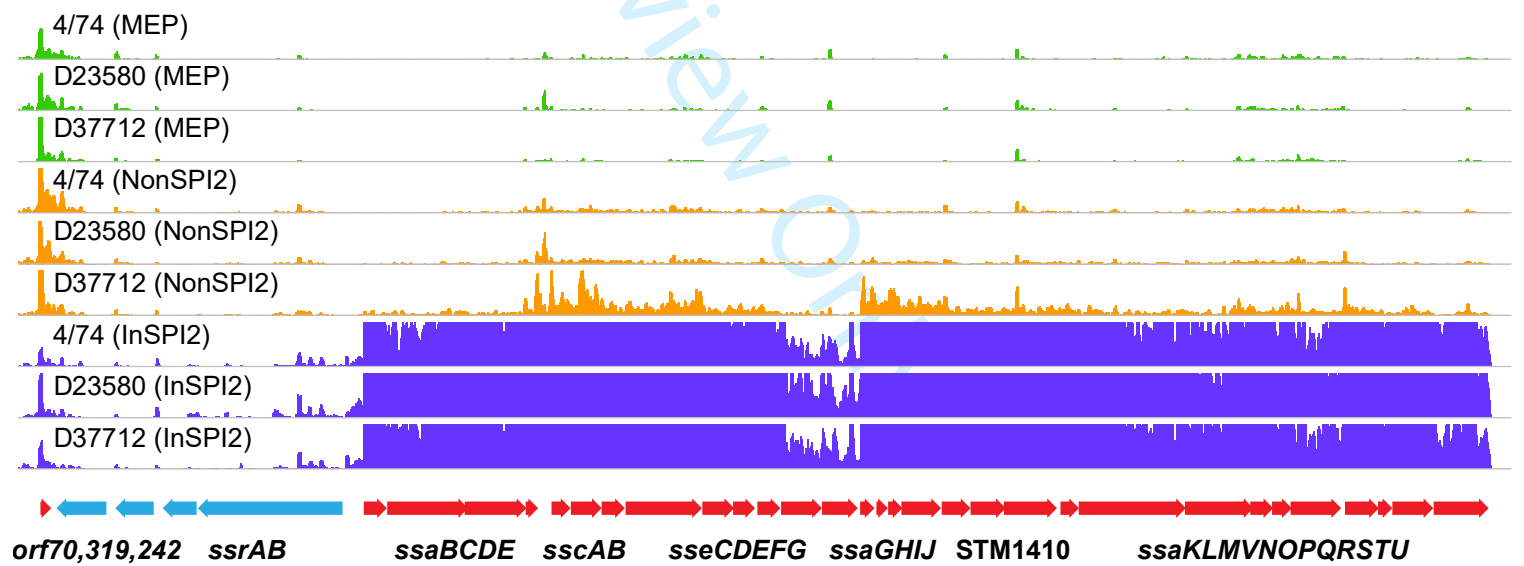




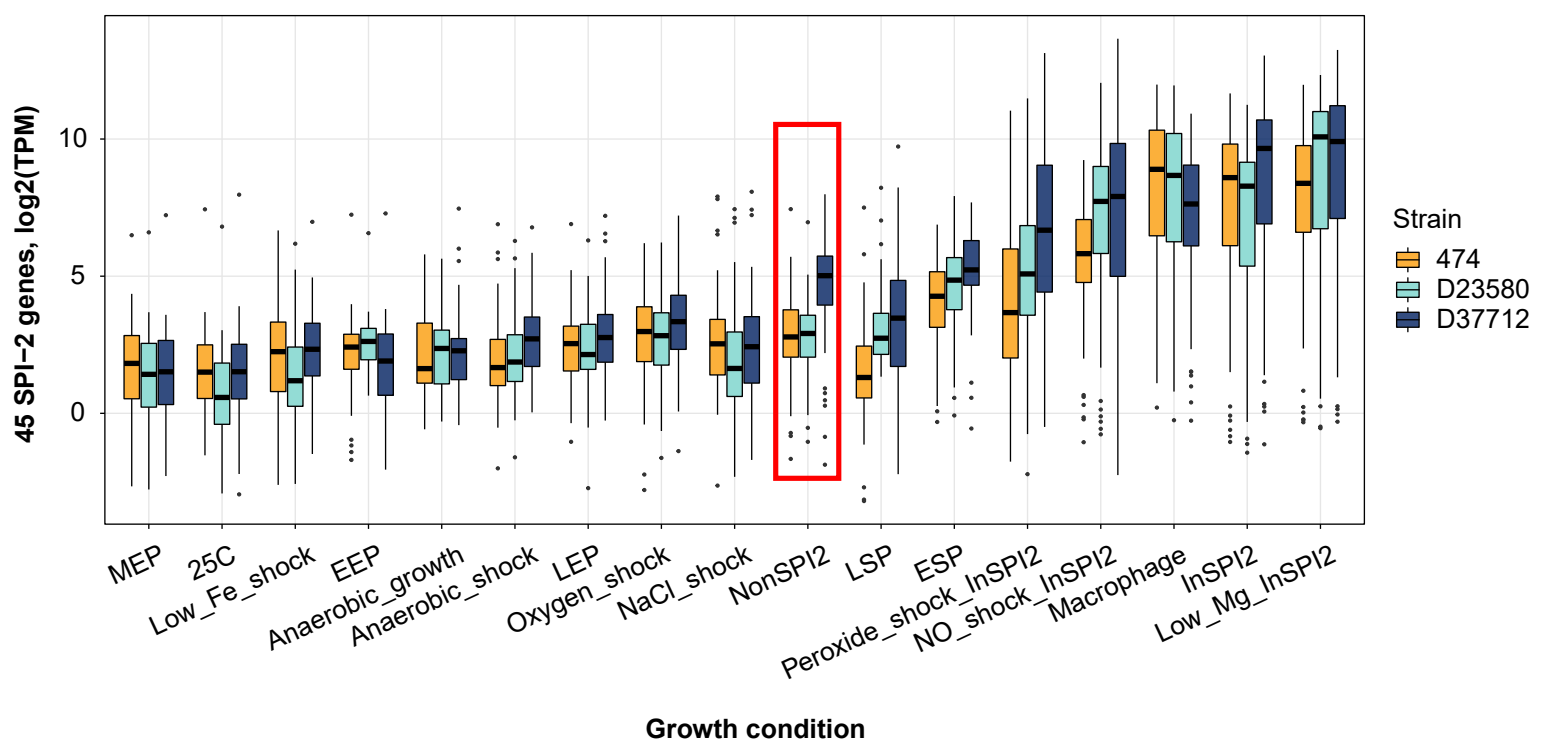
**A**



**B**

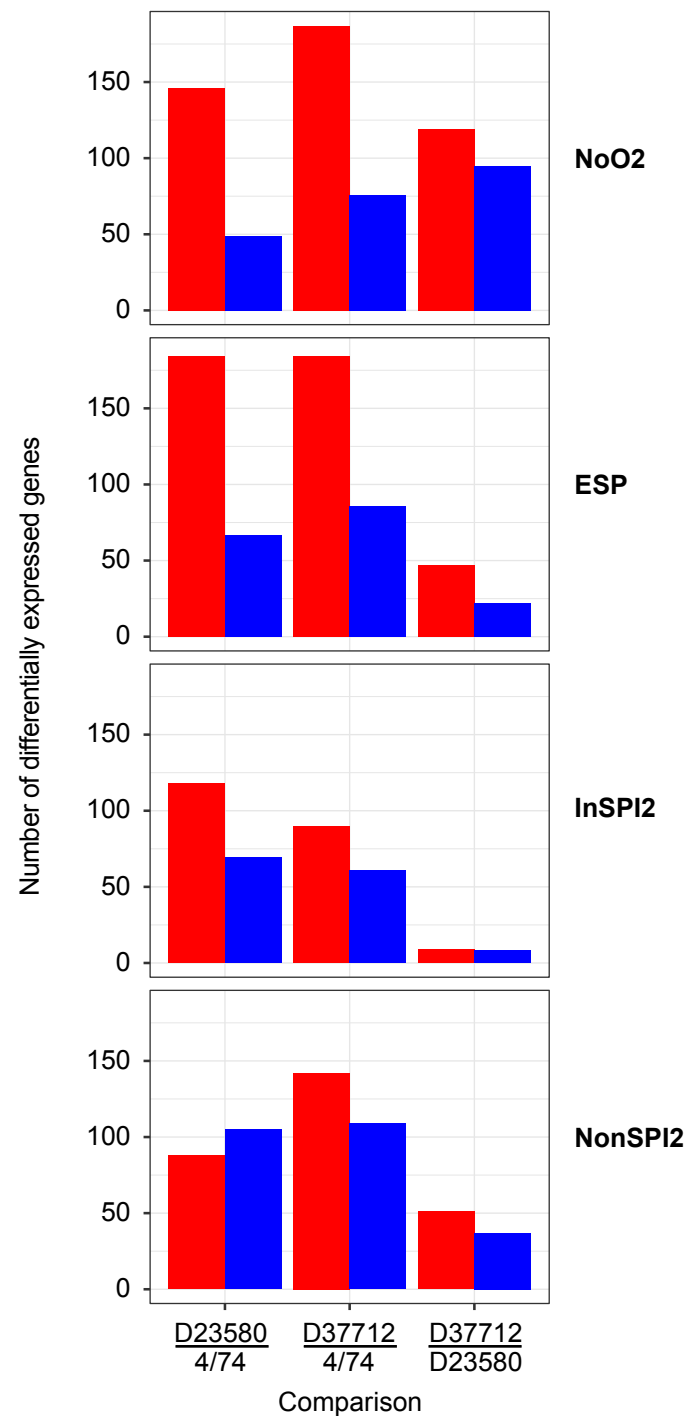


**C**



**A**

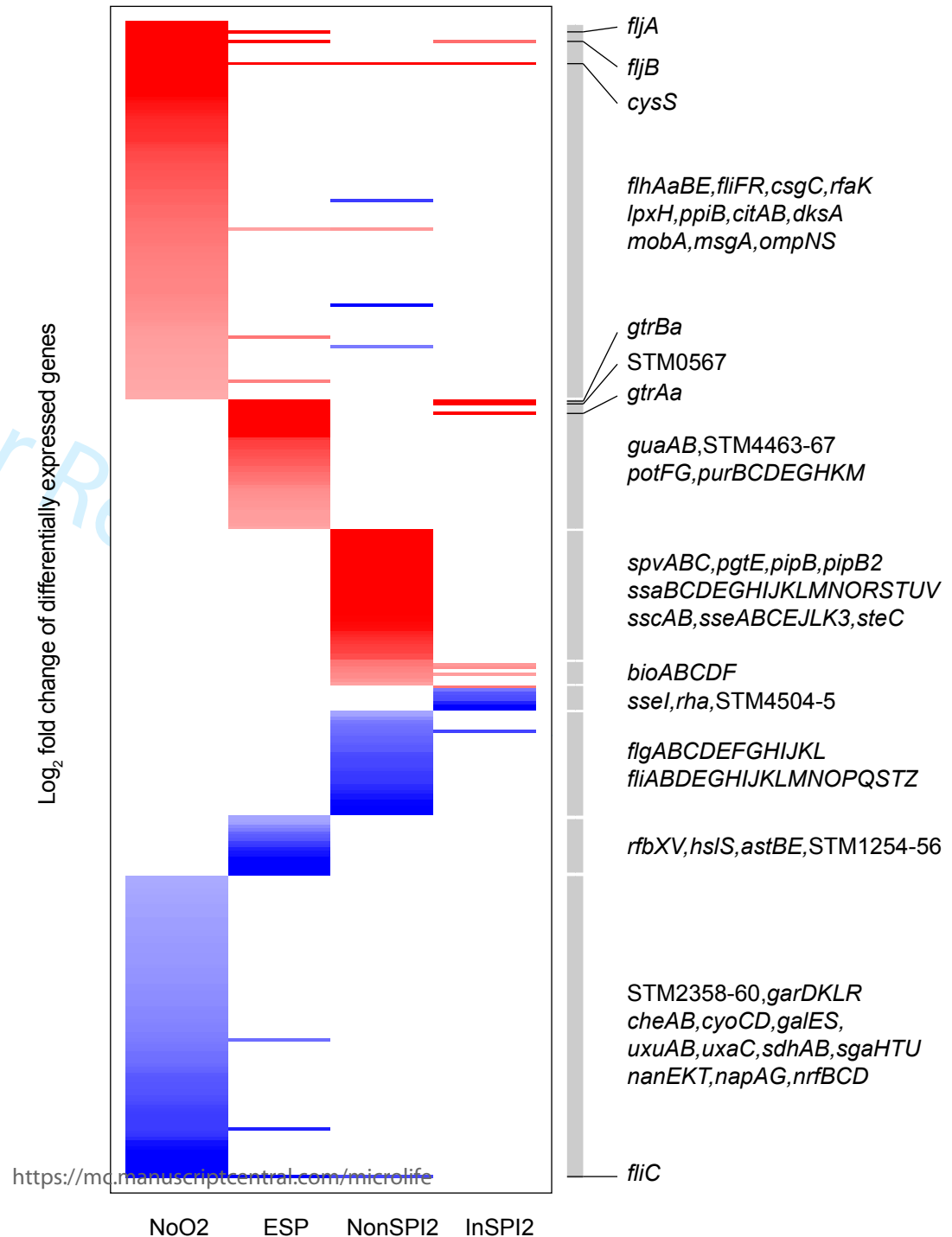
up-regulated dn-regulated



**B**

Log<sub>2</sub> fold change of differentially expressed genes

FOR REVIEW



**Fig. 4**

**Fig. 5**

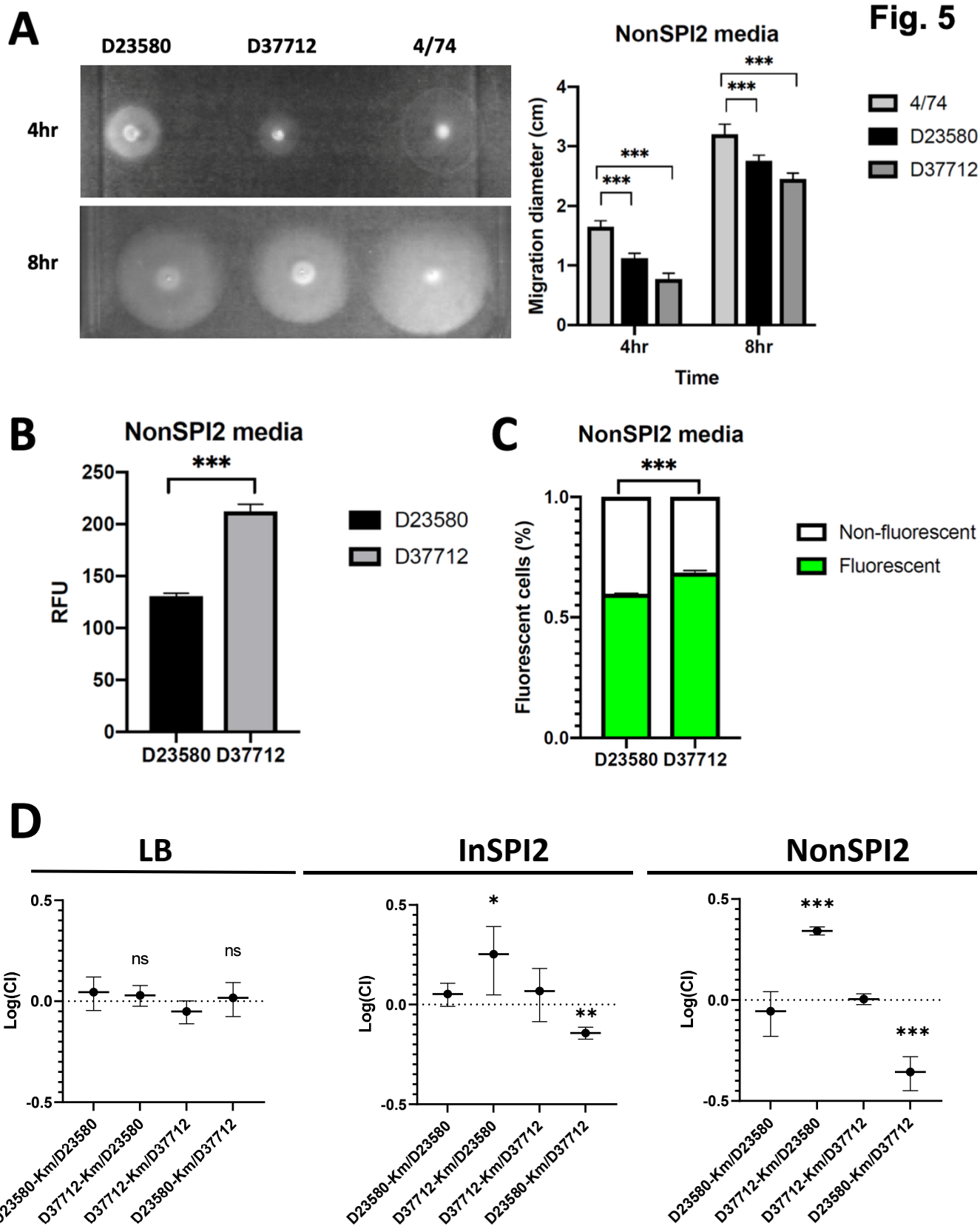


Fig. S1

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55

Free scale: 0.01

**MLST**

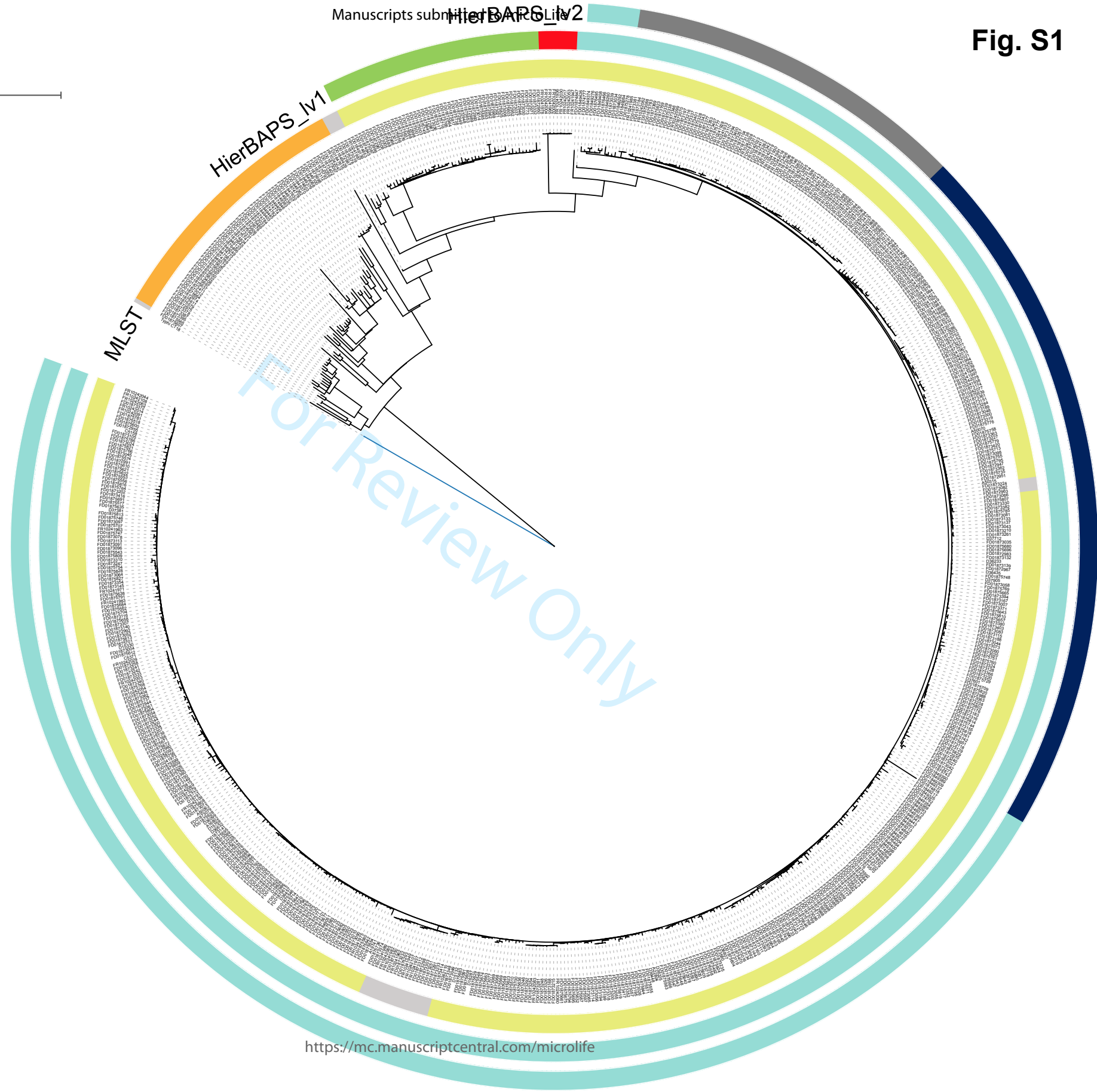
- ST313
- ST19
- Others

**HierBAPS\_iv1**

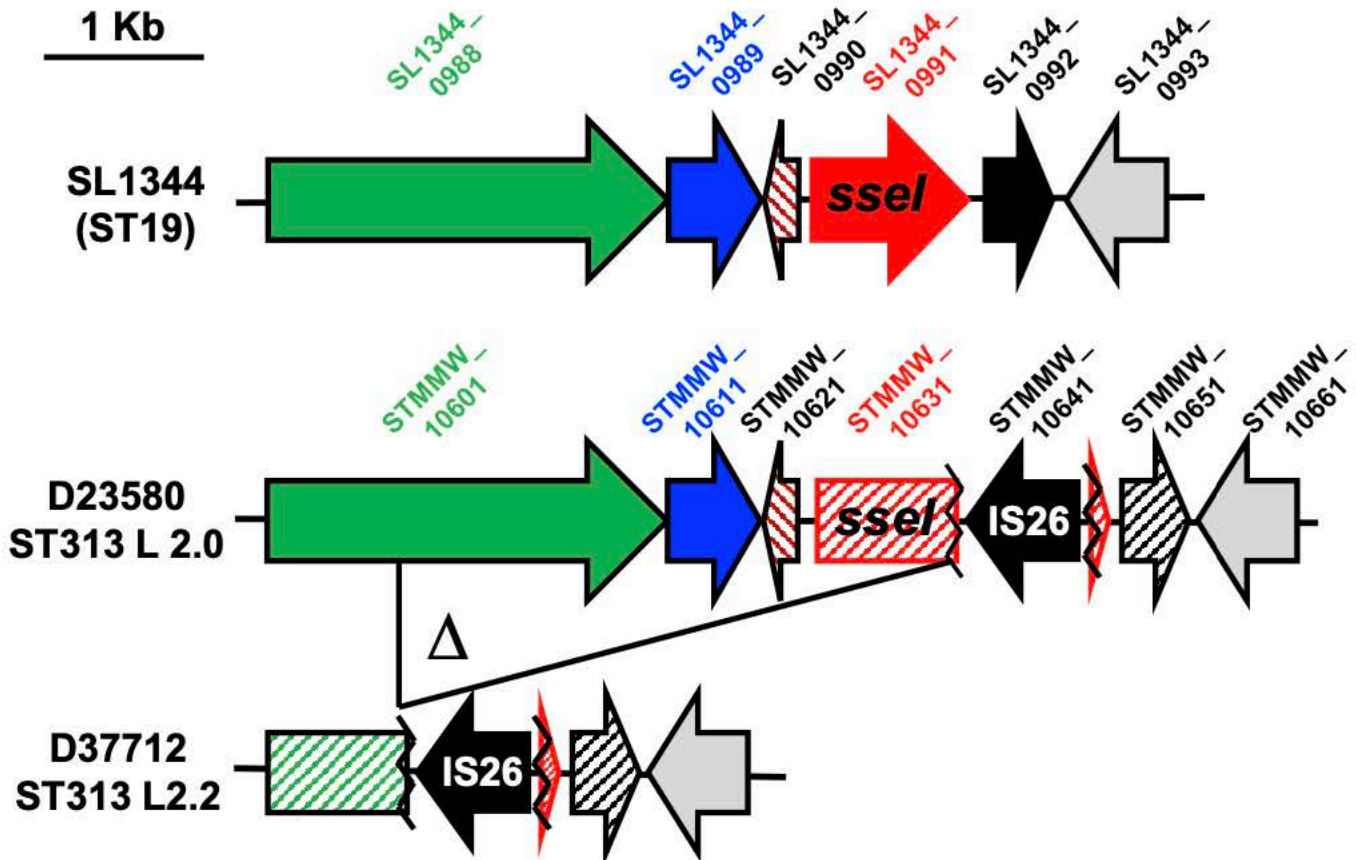
- ST313 L2
- ST313 L1
- ST313 L3

**HierBAPS\_iv2**

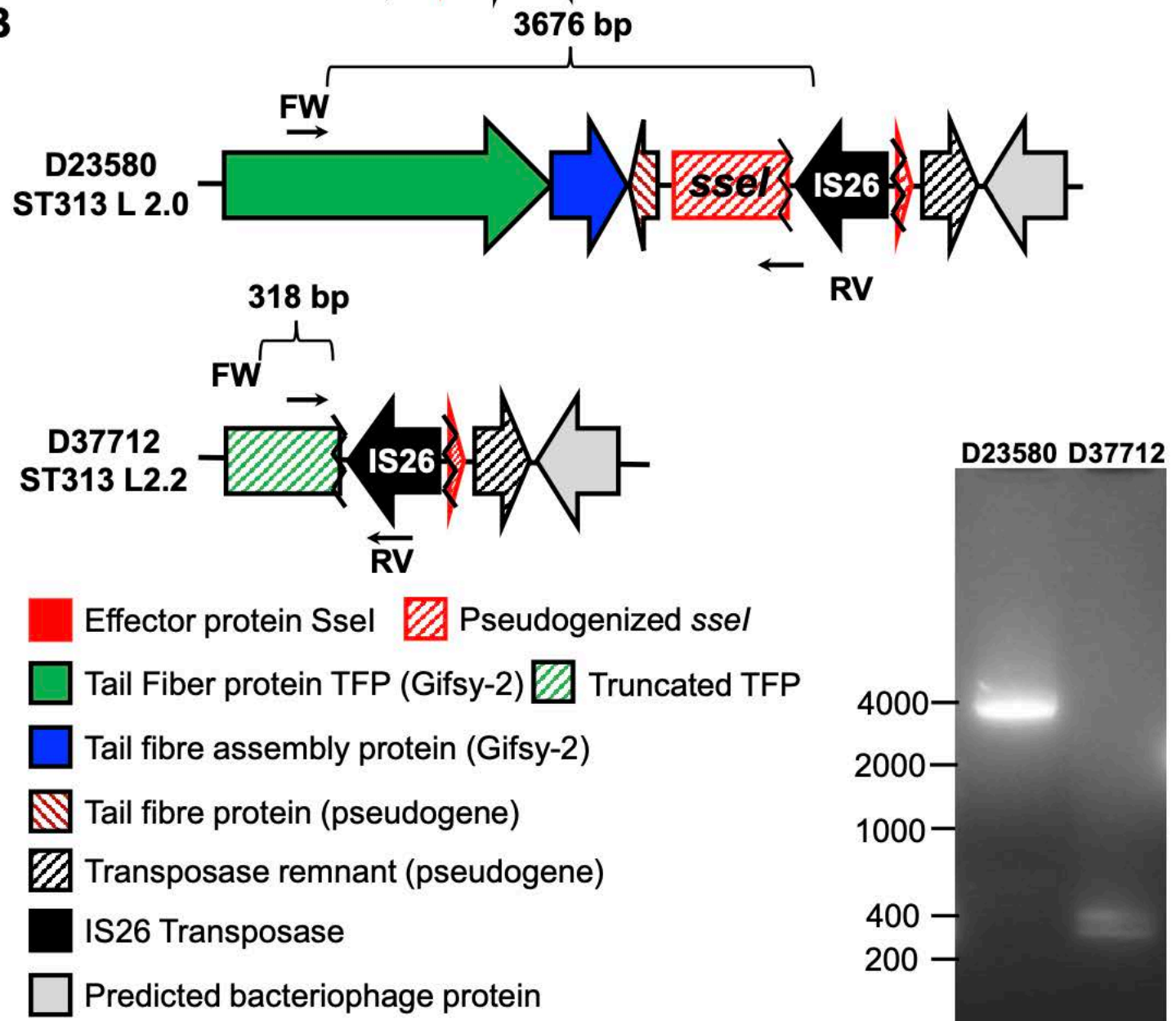
- ST313 L2.0
- ST313 L2.2
- ST313 L2.3



**A**



**B**



- Effector protein SseI     Pseudogenized *sseI*
- Tail Fiber protein TFP (Gifsy-2)     Truncated TFP
- Tail fibre assembly protein (Gifsy-2)
- Tail fibre protein (pseudogene)
- Transposase remnant (pseudogene)
- IS26 Transposase
- Predicted bacteriophage protein







Fig. S4

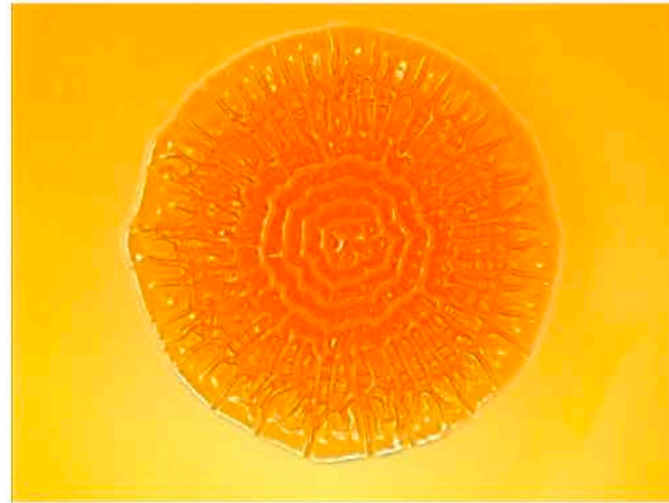
4/74

D23580

D37712

BKQZM9

Congo Red

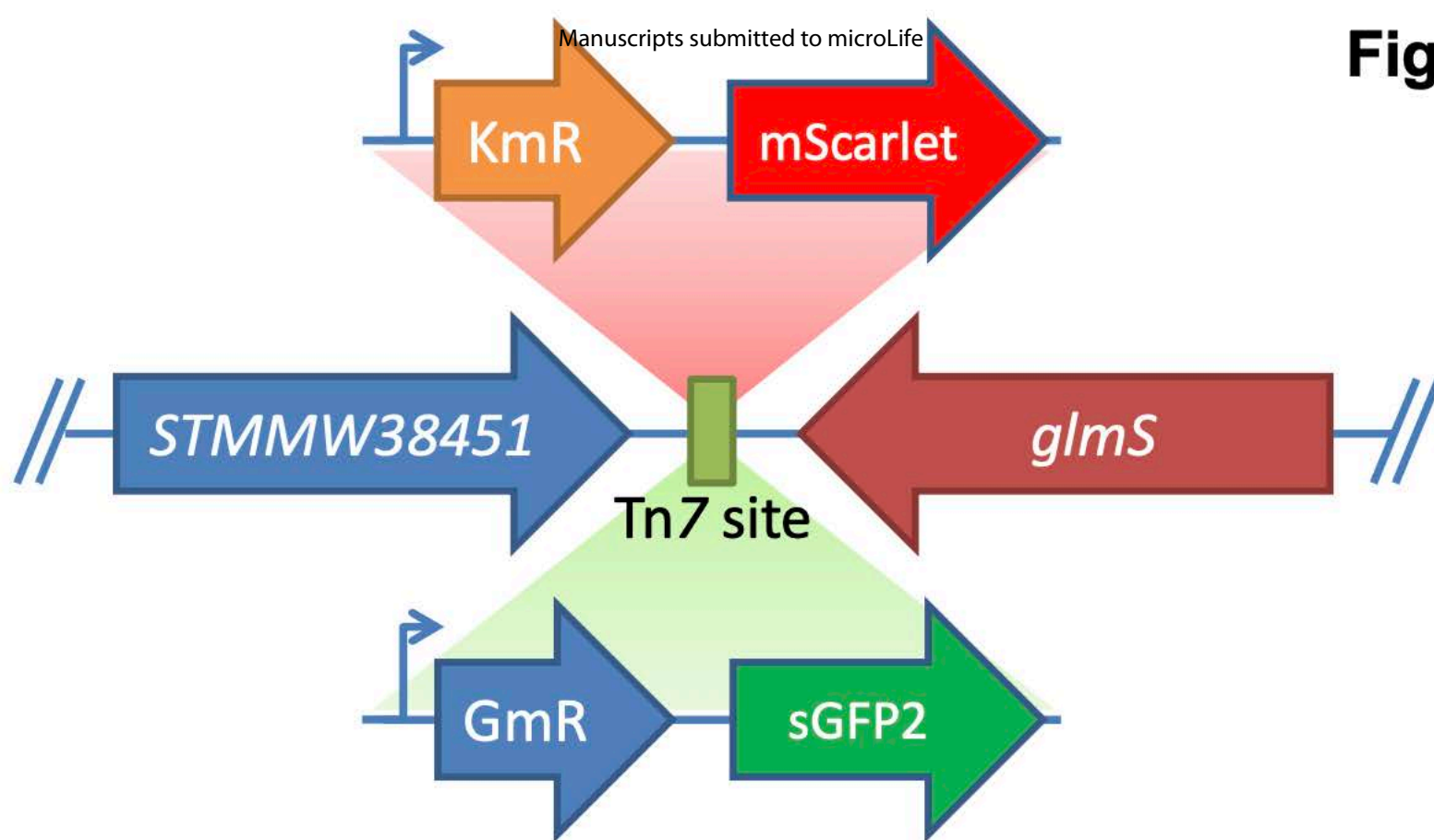


1% Tryptone  
agar



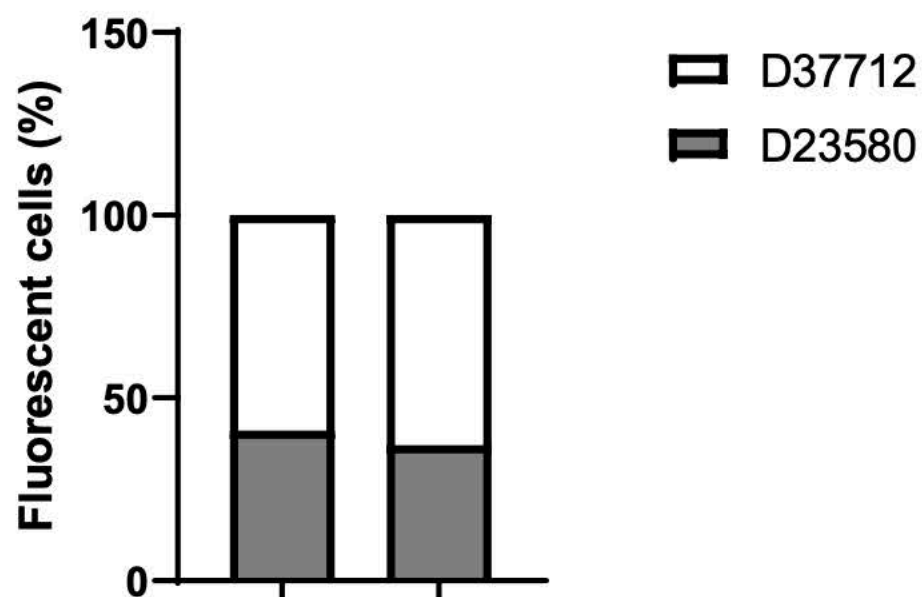
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41

A



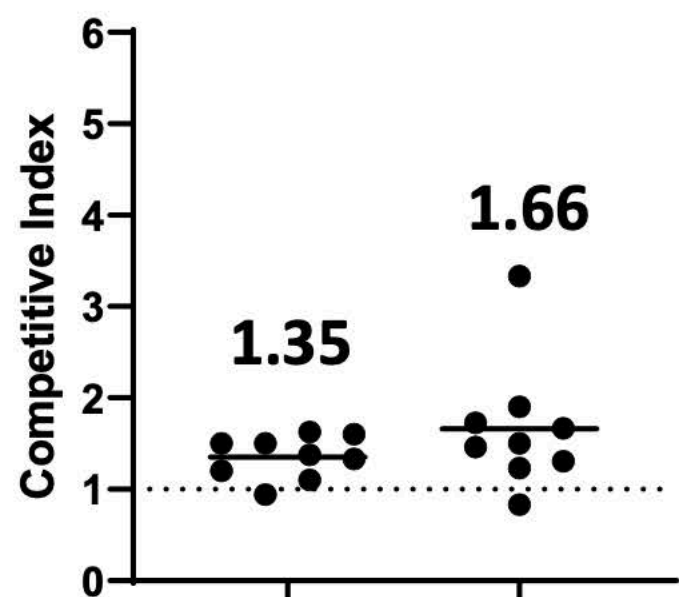
B

NonSPI2



C

NonSPI2





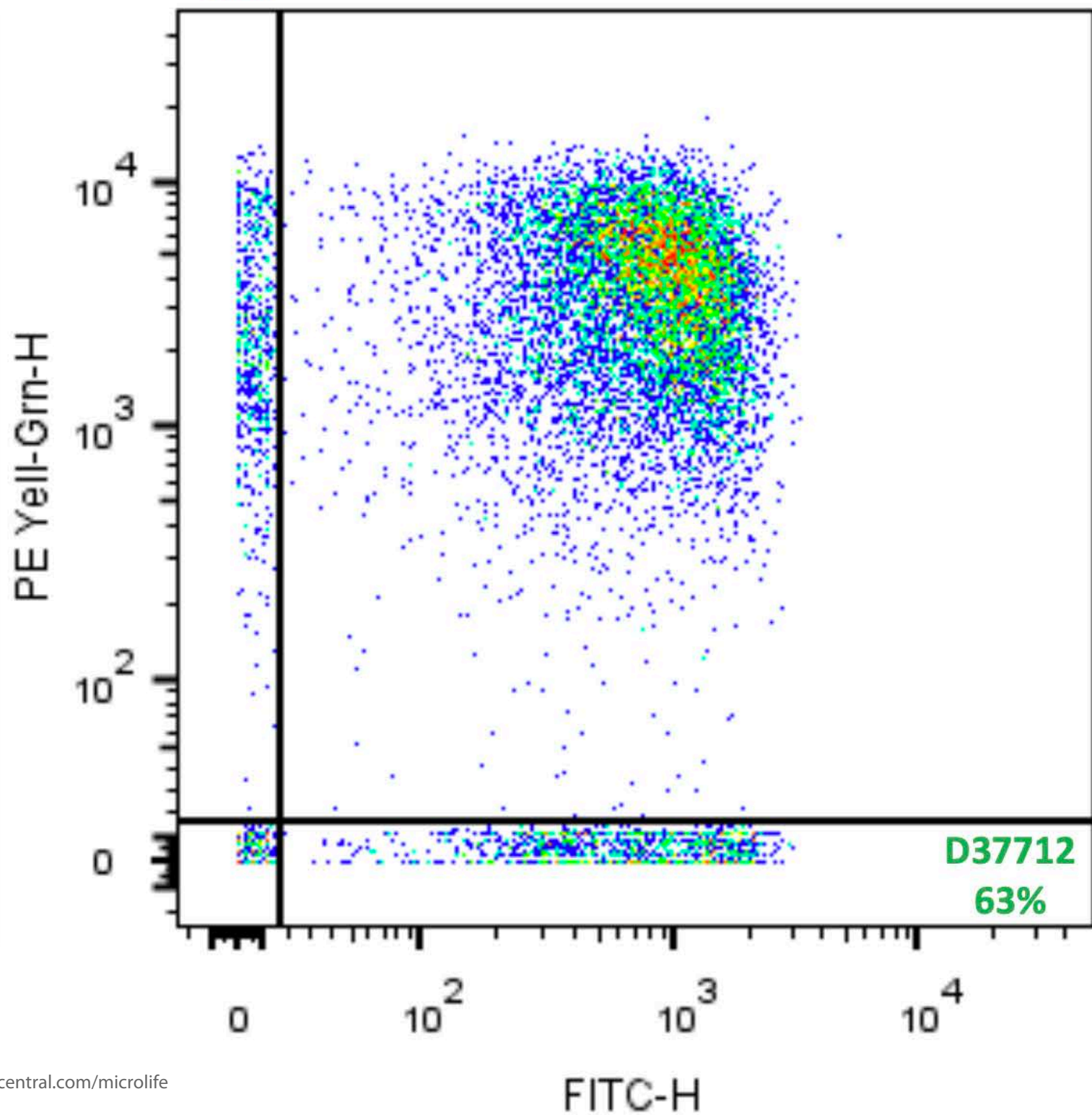
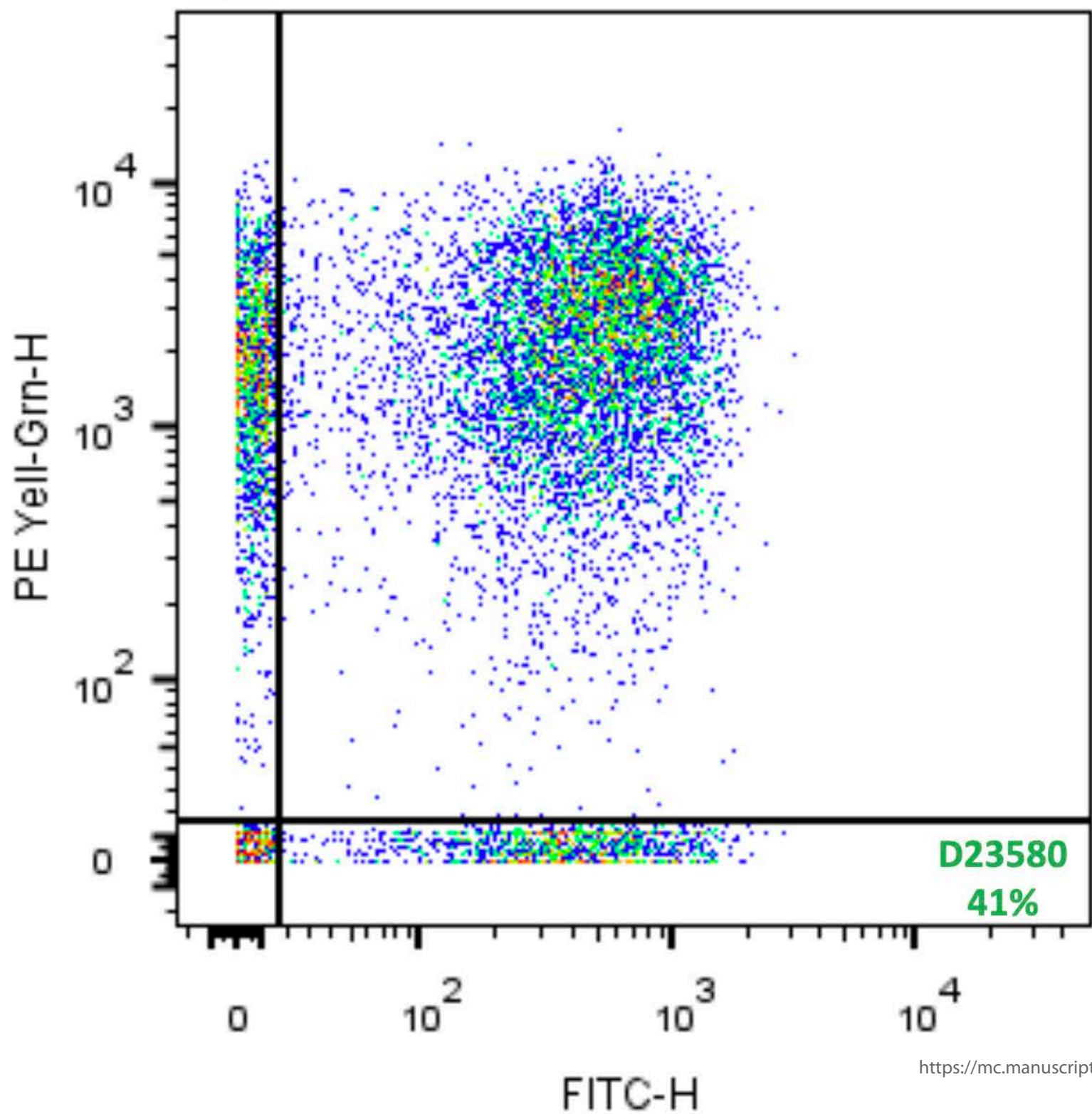
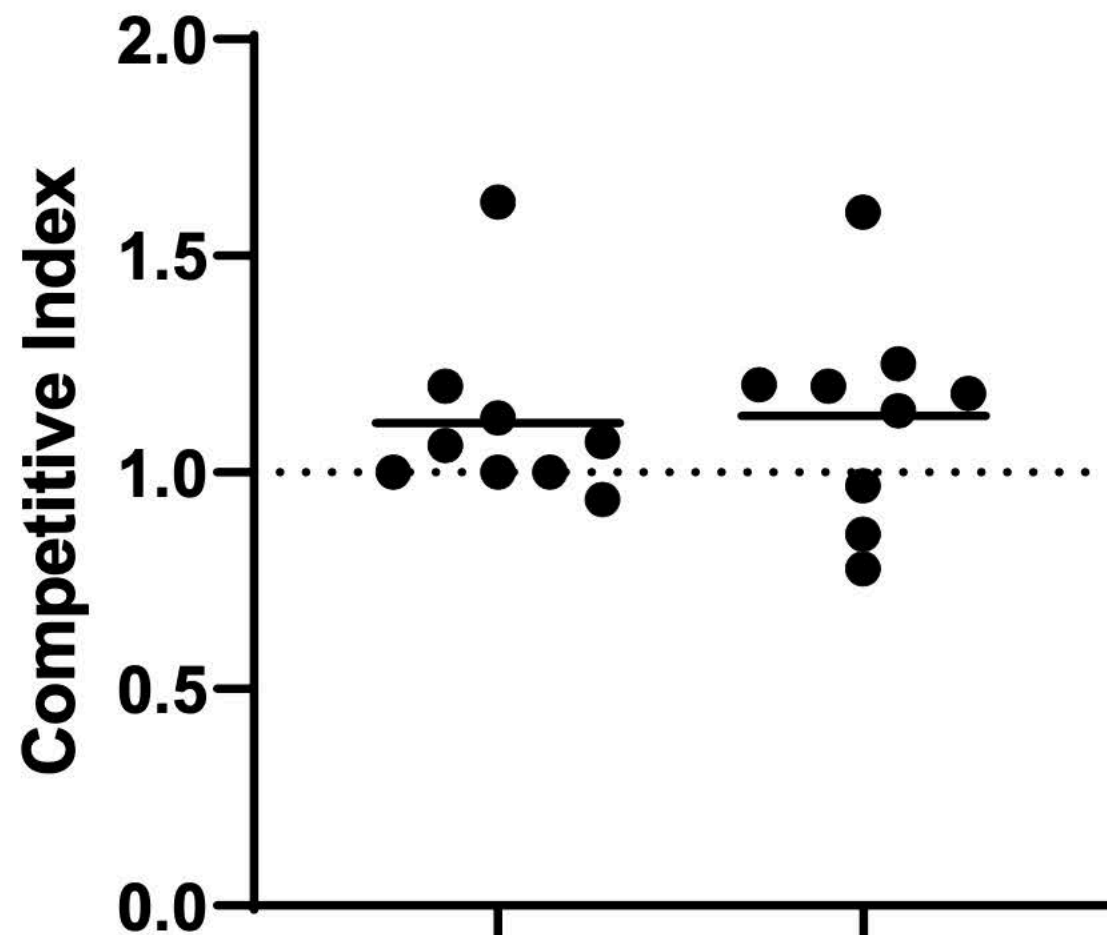
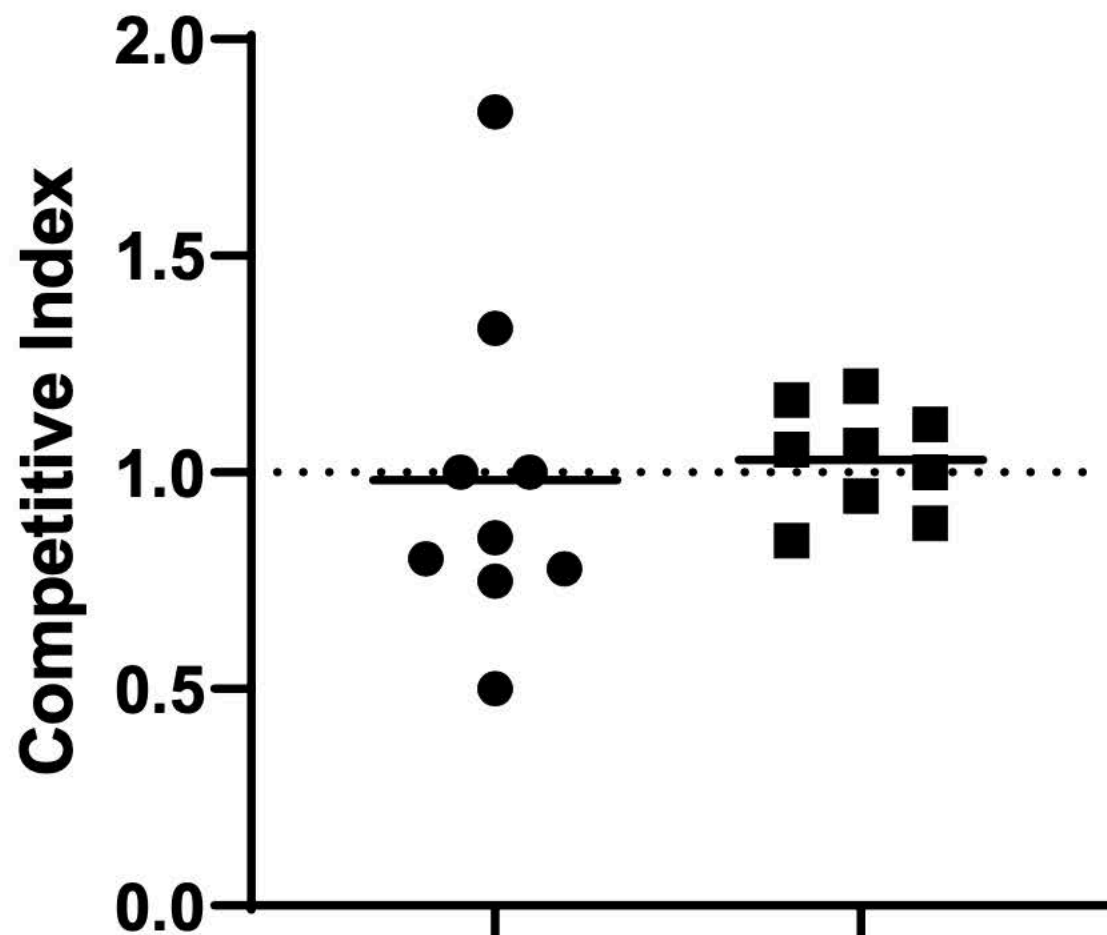


Fig. S7

# LB

# NonSPI2



D23580-Gm/D23580-Km

D37712-Gm/D37712-Km

D23580-Gm/D23580-Km

D37712-Gm/D37712-Km

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1 ***Salmonella enterica* serovar Typhimurium ST313 sublineage 2.2**  
2 **has emerged in Malawi with a characteristic gene expression**  
3 **signature and a fitness advantage**  
4

---

5  
6  
7 Benjamin Kumwenda<sup>1,2,3, #</sup>, Rocío Canals<sup>2, #</sup>, Alexander V. Predeus<sup>2, #</sup>, Xiaojun  
8 Zhu<sup>2</sup>, Carsten Kröger<sup>2,4</sup>, Caisey Pulford<sup>2</sup>, Nicolas Wenner<sup>2</sup>, Lizeth Lacharme  
9 Lora<sup>2</sup>, Yan Li<sup>2</sup>, Siân V. Owen<sup>2</sup>, Dean Everett<sup>5</sup>, Karsten Hokamp<sup>6</sup>, Robert S.  
10 Heyderman<sup>3,7</sup>, Philip M. Ashton<sup>3</sup>, Melita A. Gordon<sup>2,3</sup>, Chisomo L. Msefula<sup>1,3</sup>,  
11 Jay C. D. Hinton<sup>2,\*, \_</sup>

12  
13 <sup>1</sup>School of Life Sciences and [Allied](#) Health Professions, Kamuzu University of Health  
14 Sciences Blantyre, Blantyre, MALAWI

15 <sup>2</sup>Institute of Infection, Veterinary & Ecological Sciences, University of Liverpool, Liverpool,  
16 UNITED KINGDOM,

17 <sup>3</sup>Malawi–Liverpool–Wellcome Programme, Blantyre, MALAWI.

18 <sup>4</sup>Moyne Institute of Preventive Medicine, [School of Genetics and Microbiology](#), Trinity College  
19 Dublin, Dublin, IRE.

20 <sup>5</sup>College of Medicine and Health Sciences, Khalifa University, Abu Dhabi, UAE.

21 <sup>6</sup>Smurfit Institute of Genetics, School of Genetics and Microbiology, Trinity College Dublin,  
22 Dublin, IRE.

23 <sup>7</sup>Research Department of Infection, Division of Infection & Immunity, University College  
24 London, UK.

25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44 **Key words:** transcriptomics, comparative genomics, lineage evolution, gene expression,  
45 antibiotic resistance

*Kumwenda et. al.*

2

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

50 #BK, RC and AP contributed equally to this work.

51

52 \* **Corresponding author:** E-mail: [jay.hinton@liverpool.ac.uk](mailto:jay.hinton@liverpool.ac.uk); Tel: +441517954573; Institute  
53 [of Infection, Veterinary & Ecological Sciences, University of Liverpool, Liverpool, United](#)  
54 [Kingdom.](#)

For Review Only



## 55 Abstract

56 Invasive non-typhoidal *Salmonella* (iNTS) disease is a serious bloodstream infection that  
57 targets immune-compromised individuals, and causes significant mortality in sub-Saharan  
58 Africa. *Salmonella enterica* serovar Typhimurium ST313 causes the majority of iNTS in  
59 Malawi. ~~W, and we~~ we performed an intensive comparative genomic analysis of 608 *S.*  
60 *Typhimurium* ST313 isolates ~~dating between 1996 and 2018 obtained from fever surveillance~~  
61 ~~at the Queen Elizabeth Hospital, Blantyre, Malawi between 1996 and 2018.~~ We discovered  
62 that following the ~~upsurge arrival~~ of the well-characterised *S. Typhimurium* ST313 lineage 2  
63 ~~from in 1999 onwards~~, two ~~new~~ multidrug-resistant ~~variants~~ ~~sublineages designated 2.2 and~~  
64 ~~2.3~~, emerged in Malawi in 2006 and 2008, ~~designated sublineage 2.2 and 2.3~~ respectively.  
65 The majority of *S. Typhimurium* isolates from human bloodstream infections in Malawi now  
66 belong to sublineage 2.2 or 2.3. To ~~identify factors that characterised~~ ~~understand~~ the  
67 emergence of the prevalent ST313 sublineage 2.2, we ~~performed genomic and functional~~  
68 ~~analysis of studied~~ two representative strains, D23580 (lineage 2) and D37712 (sublineage  
69 2.2). ~~Comparative genomic analysis showed that the~~ ~~The~~ chromosome of ST313 lineage 2  
70 and sublineage 2.2 ~~were broadly similar~~, only ~~differing~~ ~~differed~~ by 29 SNPs ~~and~~ ~~/~~ small indels  
71 and ~~a~~ ~~a~~ 3kb deletion ~~in the of a~~ Gifsy-2 prophage region ~~that spanned including~~ the *sseI*/  
72 pseudogene. ~~Lineage~~ ~~The of lineage 2~~ and sublineage 2.2 ~~have unique~~ ~~had distinctive~~  
73 plasmid profiles ~~that were verified by long read sequencing~~. The transcriptome ~~was~~ ~~was~~  
74 ~~initially explored~~ ~~investigated~~ in 15 infection-relevant *in vitro* conditions and within  
75 macrophages. ~~Differential gene expression was subsequently investigated in depth in the four~~  
76 ~~most important in vitro~~ ~~During growth in physiological conditions~~. ~~that do not usually trigger S.~~  
77 *Typhimurium* SPI2 gene expression, the SPI2 genes of D37712 were transcriptionally active.  
78 We identified ~~up-regulation of SPI2 genes in non-inducing conditions~~, and down-regulation of  
79 flagellar genes in D37712, compared ~~to~~ ~~with~~ D23580. Following phenotypic confirmation of  
80 ~~transcriptional~~ ~~transcriptomic~~ differences, we discovered that sublineage 2.2 had increased  
81 fitness compared with lineage 2 during mixed-growth in minimal media. We speculate that this  
82 competitive advantage is contributing to the ~~continuing presence~~ ~~emergence~~ of sublineage 2.2  
83 in Malawi.

84

## 85 Introduction

86 Non-typhoidal *Salmonella* (NTS) is a ~~majorkey bacterial~~ pathogen that threatens people  
87 across the world. Typhimurium and Enteritidis are the two serovars of *Salmonella enterica*  
88 ~~that causeresponsible for~~ the highest levels of self-limiting gastrointestinal disease in Europe,  
89 the USA and other high-income countries (Zhang *et al.*, 2003). In the industrialised world,  
90 NTS has ~~largely~~ been associated with intensive food production, animal husbandry, and  
91 global distribution systems (Majowicz *et al.*, 2010). ~~Globally, the most common sequence type~~  
92 ~~ofThe~~ S. Typhimurium ~~associated with sequence types responsible for~~ gastroenteritis is  
93 ~~globally include~~ ST19-~~Diarrhoeal, ST34 and monophasic 1,4.[5].12:i-~~ variants (Branchu *et*  
94 *al.*, 2018). The diarrhoeal NTS disease (is termed dNTS), and is mainly foodborne ~~and poses,~~  
95 ~~posing~~ a significant burden to public health ~~globally, causingwith~~ approximately 153 million  
96 cases and 57,000 deaths per annum ~~worldwide~~ (Kirk *et al.*, 2015; Chirwa *et al.*, 2023).

97 In contrast, a lethal systemic disease called invasive non-typhoidal Salmonellosis (iNTS) has  
98 emerged in recent decades in low- and middle-income countries in sub-Saharan Africa.  
99 ~~Cases of iNTS are characterized by bloodstream infections of immune-compromised targets~~  
100 ~~immunocompromised~~ individuals such as ~~adults with HIV, and~~ children under five years of  
101 age, ~~and HIV-positive adults. Anaemia with malaria,~~ malnutrition ~~and malaria are some of the~~  
102 ~~major risk factors or severe anaemia~~ (Feasey *et al.*, 2012). In some countries of sub-Saharan  
103 Africa, *Salmonella* causes more cases of community-onset bloodstream infections than any  
104 other bacterial pathogen (Marchello *et al.*, 2019). In 2017, 535,000 cases of iNTS disease  
105 were estimated worldwide, with about 80% of cases and 77,000 deaths occurring in sub-  
106 Saharan Africa (Stanaway *et al.*, 2019)

107 Clinically, the treatment of iNTS is complicated by multi-drug (MDR) resistance which limits  
108 therapeutic options (Crump *et al.*, 2015). Widespread resistance of iNTS pathogens to first-  
109 line drugs such as chloramphenicol, ampicillin and cotrimoxazole has been seen in many  
110 countries (Kariuki *et al.*, 2006). This MDR phenotype may be one of the reasons the case  
111 fatality rate associated with iNTS is amongst the highest in comparison to any infectious  
112 disease (15%) (Marchello *et al.*, 2022). Resistance to second-line drugs such as ceftriaxone,  
113 ciprofloxacin and azithromycin has been reported in a few African countries (Tack *et al.*,  
114 2020). Clearly, the ~~problem of challenge posed by~~ MDR *Salmonella* must be addressed  
115 urgently (Gilchrist and MacLennan, 2019).

116 The African iNTS epidemic is mainly caused by two *Salmonella* pathovariants, S.  
117 Typhimurium sequence type 313 (ST313) and specific clades of *S. Enteritidis* (Kingsley *et al.*,  
118 2009; Okoro *et al.*, 2012; Feasey *et al.*, 2016). *S. Typhimurium* ST313 is responsible for about  
119 two-thirds of clinical iNTS cases that have been reported in Africa (Gilchrist and MacLennan,  
120 2019).

121 It is not certain how these pathogens are transmitted, but there is increasing evidence from  
122 case-control studies that ST313 strains are human-associated but not animal-associated

1  
2  
3 123 within households (Post *et al.*, 2019; Koolman *et al.*, 2022). A recent summary concludes that  
4 124 the available data are consistent with ~~the~~the iNTS disease being transmitted person-to-person  
5 125 transmission hypothesis for iNTS disease (Chirwa *et al.*, 2023). Global efforts to combat iNTS  
6 126 infections are currently focused on vaccine development, which ~~is currently progressing~~has  
7 127 now progressed to Phase 1 clinical trials (Piccini and Montomoli, 2020; Skidmore *et al.*,  
8 128 2023).

9  
10  
11  
12 129 Since 1998, continuous sentinel surveillance for fever and bloodstream infections among  
13 130 adults and children has been undertaken at Queen Elizabeth Central Hospital (QECH). This  
14 131 tertiary referral hospital in Blantyre, Malawi, serves an urban population of about 920,000 with  
15 132 a high incidence of malaria, HIV and malnutrition (Musicha *et al.*, 2017). Following blood-  
16 133 culture of samples collected from patients of all ages presenting with fever, whole genome  
17 134 sequencing identified the ST313 variant of *S. Typhimurium* (Kingsley *et al.*, 2009).

18 135 Phylogenetic analysis revealed that the chloramphenicol-sensitive ST313 lineage ~~was~~was  
19 136 clonally-replaced in Malawi by the chloramphenicol-resistant lineage 2 (Okoro *et al.*, 2012).  
20 137 More recently, a ST313 sublineage II.1 (2.1) emerged from lineage 2 in Democratic Republic  
21 138 of Congo (DRC) in Central Africa. Sublineage 2.1 had altered phenotypic properties including  
22 139 biofilm formation and metabolic capacity and resistance to azithromycin (Van Puyvelde *et al.*,  
23 140 2019). An elegant genomic analysis that provides insight regarding the diversity of *S.*  
24 141 *Typhimurium* ST19 clades in the context of ST313 lineage 2 clades is also available (Van  
25 142 Puyvelde *et al.*, 2023).

26 143 ~~An~~The initial suggestion that ST313 lineage 2 was undergoing evolutionary change in East  
27 144 Africa came from a small study that identified ~~seven~~several *S. Typhimurium* ST313 Malawian  
28 145 isolates, dated between 2006 and 2008, that differed from lineage 2 by 22 core-genome  
29 146 single nucleotide polymorphisms (SNPs) (Msefula *et al.*, 2012).

30 147 ~~To begin to~~ examine the evolutionary trajectory of *S. Typhimurium* in Malawi at a large scale,  
31 148 we conducted a comparative genomic analysis study focused on 680 isolates dating between  
32 149 ~~1998~~1996 and 2018 (Pulford *et al.*, 2021). We previously ~~confirmed~~reported that ST313 lineage  
33 150 1 (L1) was replaced by lineage 2 (here designated L2.0), and discovered an antibiotic-sensitive  
34 151 lineage 3 (L3) that emerged in 2016 (Pulford *et al.*, 2021).

35 152 We have now performed a more intensive phylogenetic analysis of the same collection of *S.*  
36 153 *Typhimurium* ST313 isolates, most of which caused bloodstream infections in Malawi over  
37 154 two decades. We discovered two novel sublineages named 2.2 (L2.2) and 2.3 (L2.3) that  
38 155 emerged 2006 - 2008, and have been replacing L2.0 ~~since 2006~~.

39 156 Here we present a comprehensive comparative genomic analysis of the most prevalent  
40 157 ST313 L2.2 sublineage, and report the results of a functional genomic approach that identified  
41 158 key phenotypic characteristics that distinguish L2.2 from L2.0.

42 159

## 160 Results

### 161 Identification of *S. Typhimurium* ST313 sublineages 2.2 and 2.3 in Malawi

162 ~~The *S. Typhimurium* ST313 L2 (Lineage II) was originally identified as the major cause of iNTS~~  
163 ~~cases across sub-Saharan Africa in the early 2000's (Kingsley *et al.*, 2009; Okoro *et al.*, 2012)~~  
164 ~~(Okoro *et al.*, 2015). Subsequently, an azithromycin-resistant variant of *S. Typhimurium* ST313~~  
165 ~~was found in a single country, the Democratic Republic of Congo between 2008 and 2016, and~~  
166 ~~was designated sublineage L2.1 (Van Puyvelde *et al.*, 2019).~~

167 ~~The emergence of the ST313 lineage 2 genotype in Malawi in 2002 prompted us to~~  
168 ~~hypothesise that subsequent evolution would select for variants with increased fitness,~~  
169 ~~leading to the clonal expansion of one or more sublineages by outcompeting previously~~  
170 ~~dominant genotypes. We investigated this hypothesis by conducting a detailed core-gene~~  
171 ~~SNP-based maximum likelihood (ML) phylogenetic analysis to investigate the population~~  
172 ~~structure of *S. Typhimurium* ST313 L2.0 (Fig. S1). As well as identifying members of the~~  
173 ~~antibiotic-sensitive lineage 3, To investigate the evolutionary dynamics of *S. Typhimurium*~~  
174 ~~ST313 L2 in Malawi over a 22 year period, we focused on the large collection of 8,000 *S.*~~  
175 ~~*Typhimurium* isolates derived from bloodstream infection in hospitalised patients at the Queen~~  
176 ~~Elizabeth Central Hospital, Blantyre, Malawi (Feasey *et al.*, 2015). The collection was~~  
177 ~~assembled by the Malawi Liverpool Wellcome Trust Clinical Research Programme (MLW)~~  
178 ~~between 1996 and 2018; the precise annual numbers of isolates are shown in Fig 1B. A~~  
179 ~~random sub-sampling strategy was used to select 608 isolates selected for whole-genome~~  
180 ~~sequencing which included 549 *S. Typhimurium* ST313 isolates (Pulford *et al.*, 2021).~~

181 ~~Here, we used a core-gene SNP-based maximum likelihood (ML) phylogenetic tree to~~  
182 ~~investigate the population structure of *S. Typhimurium* ST313 L2.0 in more detail (Fig. S1). As~~  
183 ~~well as identifying members of the antibiotic-sensitive lineage 3 that we reported previously~~  
184 ~~(Pulford *et al.*, 2021), we discovered that ST313 L2 could be split into comprised three~~  
185 ~~phylogenetically-distinct sublineages that differed by a total of 39 SNPs. The *S. Typhimurium*~~  
186 ~~ST313 L2 reference strain D23580 (Kingsley *et al.*, 2009) belonged belongs to the first~~  
187 ~~sublineage, which we have now designated as ST313 L2.0 lineage (Fig 4C, 1A). As ST313~~  
188 ~~sublineage L2.1 has had been defined previously (Van Puyvelde *et al.*, 2019), the new~~  
189 ~~sublineages were designated as L2.2 and L2.3, and which belonged to different hierBAPS~~  
190 ~~level 2 clusters were designated L2.2 and L2.3 (Fig 4C, 1A and Fig. S1). We In total, we~~  
191 ~~identified 151 L2.2 isolates and, 74 L2.3 isolates, against a backdrop of and 350 L2.0 isolates.~~

192 In Blantyre, Malawi, *S. Typhimurium* ST313 L2.2 was first detected in 2006, and L2.3 was  
193 initially observed in 2008 (Fig. 4A-1BC). Both L2.2 and L2.3 increased in prevalence at the  
194 Queen Elizabeth Central Hospital in Blantyre in subsequent years. By 2018, L2.2 and L2.3  
195 had largely replaced L2.0 (Fig 4A-B, 1BC). Our published Bayesian (BEAST) analysis  
196 (Pulford *et al.*, 2021) estimated that the Most Recent Common Ancestor (MRCA) of ST313  
197 lineage 2 dates back to 1948 (95% HPD = 1929-1959).

198 To understand the accessory gene complement of L2.2 and L2.3, we compared the genomes  
199 of seven L2.2 isolates and four L2.3 isolates with 17 L2.0 isolates, ST313 L1 and ST19 ~~(and~~  
200 ~~the results are shown in Fig 1C, 1A and Table S1).~~ *S. Typhimurium* strain D23580 is the  
201 representative strain of L2.0 (Kingsley *et al.*, 2009), for which we previously used long-read  
202 sequencing and other approaches to thoroughly characterise the ~~chromosome~~ chromosomal  
203 and ~~the~~ plasmid complement (Canals *et al.*, 2019b).

## 204

### 205 Antimicrobial Resistance

206 AMRMDR variants of *S. Typhimurium* with resistance to ampicillin and cotrimoxazole were  
207 detected at an early stage of the iNTS epidemic, from 1997 onwards (Gordon *et al.*, 2008).  
208 Multidrug-resistant variants of *S. Typhimurium* ST313 that were no longer susceptible to  
209 chloramphenicol, ampicillin and cotrimoxazole subsequently emerged in Malawi (Gordon *et*  
210 *al.*, 2008) and have been reported elsewhere in sub-Saharan Africa by the GEMS study  
211 (Kasumba *et al.*, 2021). The *S. Typhimurium* ST313 L2.0, L2.2 and L2.3 isolates shared the  
212 same MDR profiles (resistance to chloramphenicol, ampicillin and cotrimoxazole), and carried  
213 identical IS21-AMR associated antimicrobial gene cassettes within the pSLT-BT plasmid. ~~(Fig.~~  
214 2B).

### 217 Comparative genomics of *S. Typhimurium* ST313 sublineage 2.2

218 Because *S. Typhimurium* ST313 L2.2 was the predominant novel sublineage in Blantyre,  
219 Malawi in 2018, we focused on L2.2 for the remainder of this study. We used the phylogeny  
220 (Fig 1C, 1A) to select strain D37712 as a representative ~~isolate~~ of L2.2. D37712 was isolated  
221 from the blood of an HIV-positive Malawian male child and has been deposited in the National  
222 Collection of Type Cultures ~~(NCTC). The initial as NCTC 14678. The draft~~ genome sequence  
223 of D37712 was obtained in 2012 with Illumina technology, an assembly that comprised 27  
224 individual contigs (Msefula *et al.*, 2012). To generate a reference-quality genome, we  
225 resequenced D37712 with both long-read PacBio and Illumina short-read technologies. Our  
226 hybrid strategy generated a complete genome assembly that included one circular  
227 chromosome and three plasmids (see Materials & Methods; GenBank CP060165, CP060166,  
228 CP060167 and CP060168). This high-quality genome sequence allowed us to conduct a  
229 detailed comparative genomic analysis ~~comparison between the genomes~~ of L2.2 strain  
230 D37712 ~~with~~ and L2.0 strain D23580 (accession number FN424405), summarised in Fig. 2  
231 and Table S2.

232 Overall, the gene content of the two strains ~~contain a similar number of genes. The D37712~~  
233 ~~and D23580 genomes shared 5,016 orthologous genes, including 4,729 protein-coding genes~~  
234 ~~and pseudogenes as well as the 287 small RNA (sRNA) genes that we identified~~  
235 previously was largely equivalent. The D23580 annotation contains 4,823 protein-coding and



236 pseudogenes and 287 ~~sRNAs~~ small RNA (sRNA) genes that we identified previously (Canals  
237 *et al.*, 2019b), while D37712 contains 4,821 protein-coding and pseudogenes and ~~287~~  
238 sRNAs ~~the same 287 sRNAs~~. In total, the D37712 and D23580 genomes shared 4,729  
239 orthologous protein-coding genes and pseudogenes. The 104 protein genes that differ are  
240 encoded by the pSLT<sup>D37712</sup>, pBT1<sup>D37712</sup>, and pCol1B9<sup>D37712</sup> plasmids.

#### 241 **Overview of D23580 and D37712 genomes**

242 The chromosomes of D23580 and D37712 are 4,879,402 and 4,876,060 bp, respectively, ~~and~~  
243 similar in about the same size to other *S. Typhimurium* genomes (Kingsley *et al.*, 2009;  
244 Branchu *et al.*, 2018). The D23580 and D37712 strains share ~~a similar~~ an identical prophage  
245 profile, with both strains carrying five prophages (BTP1, Gifsy-2, ST64B, Gifsy-1, and BTP5)  
246 ~~which were located at the same positions on the chromosome. Previously, we have~~  
247 established that just one of these prophages, BTP1, is functional (Owen *et al.*, 2017). The  
248 BTP1 prophage of D23580 encodes the novel BstA phage defence system (Owen *et al.*,  
249 2021) and a particularly high level of viable BTP1 phages is produced by spontaneous  
250 induction (Owen *et al.*, 2017) located at the same positions on the chromosome (Fig. 2A).

#### 251 **Comparison of D23580 and D37712 chromosomes**

252 The detailed genomic comparison of D37712 with D23580 showed that the ~~two genomes~~  
253 were remarkably similar. sizes of the two chromosomes varied by only 3,342 bp. Overall, the  
254 only differences between the genomes of the L2.0 and L2.2 strains were 26 chromosomal  
255 SNPs and small indels, plus one large deletion, and an inversion of the *hin* switch. In-depth  
256 annotation of the nucleotide variants identified 3 putative loss-of-function mutations (2 stop  
257 mutations, 1 frameshift insertion), 1 disruptive in-frame deletion, 4 synonymous mutations, 13  
258 missense mutations, and 5 intergenic variants, summarised in Fig. ~~2A~~. 2A. None of the SNP  
259 differences that distinguished D37712 from D23580 were located within 150 nucleotides of a  
260 Transcriptional Start Site (Canals *et al.*, 2019b), and so would not be predicted to modulate  
261 gene expression.

262 The 3,358 bp-long deletion of a Gifsy-2 prophage-associated region that spanned the *sseI*  
263 pseudogene of D23580 (STMMW\_10631) removed two coding sequences (STM1050-51;  
264 STMMW\_10611-STMMW\_10631), and substantially truncated the STM1049  
265 (STMMW\_10601) gene (Fig. 2E). The *sseI* gene encodes a cysteine hydrolase effector  
266 protein that modulates the directional migration of dendritic cells during systemic infection  
267 (Brink *et al.*, 2018). In strain D23580, the insertion of aan IS26 transposable element  
268 IS15DEV inactivated the *sseI* gene (Kingsley *et al.*, 2009), causing increased dendritic cell-  
269 mediated dissemination of strain D23580 during infection (Carden *et al.*, 2017). ~~To~~ We used  
270 an independent PCR-based approach to confirm that the 3,358 bp deletion had removed the  
271 *sseI* gene from the chromosome of strain D37712, ~~we used an independent PCR-based~~  
272 approach (Fig. S2).

#### 273 **Comparison of D23580 and D37712 plasmids**



274 ~~Here we put the genetic features of the representative strains for ST313 L2.0 and L2.2 into~~  
275 ~~context with other isolates belonging to the Lineage 2 sublineages.~~ ST313 L2.0 strain D23580  
276 carries four plasmids, pSLT-BT, pBT1, pBT2 and pBT3 (Kingsley *et al.*, 2009). In contrast,  
277 ST313 L2.2 ~~carried~~has a distinct plasmid complement (Fig. 4C, 1A, Fig. 2BCD). ~~The plasmid~~  
278 ~~profiles of D23580 and D37712 were confirmed by a combination of Illumina (short-read) and~~  
279 ~~PacBio (long-read) sequencing (Materials and Methods).~~

280 In summary, strain D37712 carried ~~the~~ pSLT-BT, pBT2 and pCol1B9 ~~plasmids~~ as detailed  
281 below. Both ~~D23580 and D37712~~ strains ~~had~~carried a variant of the pSLT-BT virulence  
282 plasmid (Kingsley *et al.*, 2009) that contains a Tn21-like transposable element with five  
283 antibiotic resistance genes. The D37712 version of pSLT-BT ~~is similar to that~~only differs from  
284 ~~the pSLT-BT~~ of D23580, ~~with in~~ two important ~~difference~~st ways (Fig. 2B). Firstly, the Tn21-  
285 like element is inserted in the opposite direction with regards to the rest of the plasmid,  
286 suggesting that the transposable element remains active. Secondly, three nucleotide variants  
287 were identified in the pSLT-BT ~~variant~~carried by D37712, two deletions in noncoding regions,  
288 and one frameshift insertion that generates a pseudogene of *spvD*. ~~The SpvD effector protein,~~  
289 ~~a cysteine protease, is translocated by the SPI2 type 3 secretion system and suppresses the~~  
290 ~~NF- $\kappa$ B-mediated pro-inflammatory immune response and contributes to virulence in mice~~  
291 ~~(Grabe *et al.*, 2016).~~

292 Plasmid pCol1B9 was of particular interest because it was absent from D23580, but ~~is~~was  
293 present in *S. Typhimurium* ST19 strain 4/74 (Richardson *et al.*, 2011; Fig. 1A). 4/74 is the  
294 parent of ~~the~~ *S. Typhimurium* SL1344, ~~a~~ strain that has been used extensively for the study of  
295 *S. Typhimurium* pathogenesis and gene regulation ~~since 1986~~in recent decades (Kröger *et al.*,  
296 2012; Rankin & Taylor, 1966). Our ~~new~~ annotation of the pCol1B9-like plasmid  
297 ~~included~~identified 95 distinct protein-coding genes, while the previously published annotation  
298 of pCol1B9<sup>4/74</sup> assigned 101 protein-coding genes. Some of these represent annotation  
299 discrepancies, while others represent true genetic differences (Fig. S3). ~~Upon~~

300 ~~Following~~ careful examination, ~~we~~ identified 14 genes ~~were~~unique to pCol1B9<sup>D37712</sup>, ~~while~~and  
301 ~~20~~ ~~were~~genes unique to pCol1B9<sup>4/74</sup>. There were 81 genes carried by both plasmids.

302 Interestingly, pCol1B9<sup>D37712</sup> lacked the colicin toxin-antitoxin system that both gave pCol1B9  
303 its name, and provides *Salmonella* with a competitive advantage in the gut (Nedialkova *et al.*,  
304 2014). The pCol1B9<sup>D37712</sup> plasmid carried a locus that was absent from pCol1B9<sup>4/74</sup>, namely  
305 the *impC-umuCD* operon (Fig. S3) which encodes the error-prone DNA polymerase V  
306 responsible for the increased mutation rate linked to the SOS stress response in *E. coli*  
307 (Sikand *et al.*, 2021).

308 ~~An 85-kb plasmid carried by D23580, pBT1, was previously shown by our laboratory to play an~~  
309 ~~important role in *Salmonella* biology by encoding an orthologous *cysS* gene responsible for~~  
310 ~~expressing the essential cysteinyl tRNA-synthetase enzyme (Canals *et al.*, 2019b).~~ This pBT1  
311 ~~plasmid was completely absent from D37712, and from all isolates of sublineage L2.2 that were~~  
312 ~~examined (Fig. 1C).~~

1  
2  
3 §134  
5 §14 **Comparison of pseudogene status of D23580 and D37712**

6  
7 §15 Our comparative genomic analysis focused on the pseudogenes found in strains 4/74,  
8 §16 D23580, and D37712 (Fig. 2F, Table S3). The pseudogenisation of several D23580 genes,  
9  
10 §17 compared with strain 4/74, have been linked to the invasive phenotype of African *Salmonella*  
11 §18 ST313 (Kingsley *et al.*, 2009). We found that the pseudogene complement of D23580 was  
12 §19 largely conserved in D37712, consistent with inheritance from a common ancestor. We have  
13  
14 §20 recently reported the role of the MacAB-TolC macrolide efflux pump in the virulence of *S.*  
15 §21 Typhimurium ST313, and showed experimentally that *macB* was an inactive pseudogene in  
16 §22 D23580 (Honeycutt *et al.*, 2020). Interestingly, the *macB* gene is functional in D37712.  
17 §23 Compared with D23580, three additional D37712 genes were pseudogenised (*spvD*, *yadE*,  
18 §24 and STMMW\_42692), as detailed in Table S3. YadE is a predicted polysaccharide  
19 §25 deacetylase lipoprotein. The functional impact of these pseudogenes on L2.2 remains to be  
20 §26 established.

21  
22  
23  
24  
25 §27 Overall the chromosomes of ST313 lineage 2 and sublineage 2.2 were highly-conserved and  
26 §28 differed by just 29 SNPs/ small indels, and a 3kb deletion in the Gifsy-2 prophage region. The  
27 §29 ST313 lineage 2 and sublineage 2.2 have distinct plasmid profiles.

30 §30 **Transcriptional landscape of *S. Typhimurium* ST313 sublineage L2.2**

31 §31 Previously, we characterized the primary transcriptome of two other *S. Typhimurium* strains,  
32 §32 4/74 and D23580, using a combination of multi-condition RNA-seq and differential RNA-seq  
33 §33 (dRNA-seq) techniques (Canals *et al.*, 2019b; Kröger *et al.*, 2013). To identify the  
34 §34 transcriptional start sites (TSS) of strain D37712, we analysed a pooled sample containing  
35 §35 RNA from 15 *in vitro* conditions by dRNA-seq and RNA-seq as detailed previously (Kröger *et*  
36 §36 *al.*, 2013). The high similarity between the D23580 and D37712 chromosomes allowed us to  
37 §37 map the curated set of TSS that were previously defined for D23580 (HammarlöfHammarlöf  
38 §38 *et al.*, 2018) onto a combined D37712/D23580 reference genome. To allow individual TSS to  
39 §39 be examined in particular chromosomal or plasmid regions, data from both the dRNA-seq and  
40 §40 pooled RNA-seq experiments can be visualised in our online genome browser  
41 §41 ([http://hintonlab.com/jbrowse/index.html?data=Combo\\_D37/data](http://hintonlab.com/jbrowse/index.html?data=Combo_D37/data)).

42 §42 **Preliminary gene expression profiling of *S. Typhimurium* ST313 sublineage**  
43 §43 **L2.2**

44 §44 Given the high level of similarity between the genomes of L2.2 and L2.0, we went on to  
45 §45 identify differences at the transcriptional level. We performed a multi-condition RNA-seq-  
46 §46 based transcriptomic analysis of gene expression profiles of L2.2 strain D37712 without  
47 §47 biological replicates.

48 §48 This comparative transcriptomic screen was based on our published approach (Canals *et al.*,  
49 §49 2019b). Specifically, we used 15 individual infection-relevant *in vitro* conditions (Kröger *et al.*,  
50 §50 2013) and did intra-macrophage transcriptome profiling using the protocol previously  
51 §51  
52 §52  
53 §53  
54 §54  
55 §55  
56 §56  
57 §57  
58 §58  
59 §59  
60 §60

1  
2  
3 351 established for *S. Typhimurium* ST19 (Srikumar *et al.*, 2015). The RNA-seq samples were  
4 352 mapped to a combined reference genome, which included the annotated D23580  
5 353 chromosome (Canals *et al.*, 2019b), as well as all the plasmids described earlier (pSLT-BT,  
6 354 pBT1, pBT3 and pCol1B9; see Methods). The initial RNA-seq assessment (detailed in  
7 355 Methods) involved 2-4M non-rRNA/tRNA reads per sample, allowing gene signatures specific  
8 356 for each *in vitro* condition to be identified. Although single replicate RNA-seq experiments of  
9 357 this type cannot be used for statistically-robust differential gene expression analysis, they do  
10 358 provide a useful screening approach for identifying growth conditions to be used for follow-up  
11 359 experiments. The individual RNA-seq experiments showed broad condition-specific  
12 360 similarities in gene expression between strains 4/74, D37712, and D23580 (Fig. 3A). The  
13 361 gene expression values from each profiled condition are available as raw counts and TPMs in  
14 362 Tables S4 and S5.

15 363 To select the ideal environmental conditions to use for subsequent experiments, we assessed  
16 364 the expression profiles of known *Salmonella* pathogenicity islands which were broadly similar  
17 365 in strains D37712, and D23580. Although the expression profile of the SPI2 pathogenicity  
18 366 island was broadly similar between D37712, D23580 and 4/74 in most growth conditions, the  
19 367 SPI2 genes of D37712 were highly up-regulated in a single growth condition, NonSPI2 (Fig.  
20 368 3B-C). NonSPI2 is a minimal medium with a neutral pH and a relatively high level of  
21 369 phosphate, in which *S. Typhimurium* does not usually express the SPI2 pathogenicity island  
22 370 (Löber *et al.*, 2006; Kröger *et al.*, 2013). This intriguing observation prompted us to perform  
23 371 ~~thea~~ more discriminating set of transcriptomic experiments, as described below.

24 372

### 25 373 **Differential gene expression analysis of *S. Typhimurium* D37712 versus D23580 in four** 26 374 ***in vitro* conditions with multiple biological replicates**

27 375 To define the transcriptional signature of strain D37712 more accurately, we generated RNA-  
28 376 seq data from D37712 grown in four *in vitro* conditions that stimulate expression of the  
29 377 majority of virulence genes: ESP, anaerobic growth, NonSPI2 and InSPI2, with multiple (3-4)  
30 378 biological replicates. The combination of acidity (pH 5.8) and low phosphate (0.4 mM Pi) in  
31 379 the InSPI2 media stimulates transcription of SPI2 genes in *S. Typhimurium* (Löber *et al.*,  
32 380 2006; Kröger *et al.*, 2013). The NonSPI2 condition is based on the same PCN media recipe  
33 381 as InSPI2 media, but is neutral (pH 7.4), and contains higher levels of phosphate (25 mM Pi)  
34 382 (Löber *et al.*, 2006; Kröger *et al.*, 2013).

35 383 We compared the results with our published transcriptomic data for *S. Typhimurium* strains  
36 384 4/74 and D23580 (Canals *et al.*, 2019b; Kröger *et al.*, 2013). Differential expression analysis  
37 385 with DEseq2, with conservative cut-offs (fold change  $\geq 2$ , FDR  $\leq 0.001$ ), showed that the gene  
38 386 expression profiles of D37712 and D23580 were broadly similar, and shared key differences  
39 387 to the transcriptional profile of strain 4/74 under each of the four *in vitro* conditions (Fig. 4A).  
40 388 The differential expression results are summarized in Table S6.

389 We specifically investigated transcription of the *pgtE* gene, which encodes the outer-  
390 membrane protease previously linked to the ability of African *Salmonella* ST313 to resist  
391 human serum killing (Hammarlöf *et al.*, 2018). Compared to 4/74, the *pgtE* gene of both the  
392 D23580 and D37712 strains showed a similar pattern of up-regulation by a factor of 7 to 18  
393 across all conditions. This finding is consistent with the fact that D37712 carries the same T  
394 nucleotide in the -10 region of the *pgtE* promoter that is responsible for increased expression  
395 of the *pgtE* transcript in strain D23580 (Hammarlöf *et al.*, 2018).

396 ~~The~~ There were no statistically-significant changes in expression of the majority (92%) of the  
397 4,729 orthologous coding genes ~~of both shared by~~ D37712 and D23580 ~~were expressed at~~  
398 ~~similar levels~~. We identified a total of 364 genes that were differentially expressed in at least  
399 one growth condition between D37712 and D23580 as follows: ESP (69 differentially-  
400 expressed genes), anaerobic growth (214 differentially-expressed genes), NonSPI2 (88  
401 differentially-expressed genes) and InSPI2 (17 differentially-expressed genes; Fig. 4B).

402 Overall, the differentially-expressed genes that distinguished D37712 from D23580 ~~were~~  
403 ~~seen only showed expression differences~~ in a single growth condition ~~and rather than across~~  
404 ~~all conditions~~. The differentially expressed genes included flagellar genes (down-regulated),  
405 SPI2-associated genes (up-regulated), and genes involved in general and anaerobic  
406 metabolism (down-regulated).

407 ~~The~~ SPI2 pathogenicity island genes play a key role in the intracellular replication of *S.*  
408 Typhimurium, and encode the type III secretion system that is responsible for translocation of  
409 key effector proteins into mammalian cells (Jennings *et al.*, 2017). The RNA-seq data showed  
410 that SPI2 genes were expressed at similarly high levels in both D37712 and D23580 strains  
411 following induction (InSPI2 media; Fig. 4B), and confirmed that the key SPI2 expression  
412 difference was only seen in strain D37712 under non-inducing growth conditions (NonSPI2  
413 media). It is important to put this differential SPI2 expression into context. D37712 expresses  
414 SPI2 genes at about a 10-fold higher level than D23580 during growth in non-inducing  
415 NonSPI2 media, but the actual level of expression was 20-fold less than the level stimulated  
416 by growth in SPI2-inducing conditions (InSPI2 medium).

417 The up-regulation of *fljA* and *fljB* and the down-regulation of *fliC* in D37712, compared to  
418 D23580 in all four growth conditions likely reflects the opposite orientation of the *hin* switch in  
419 the D37712 genome compared to D23580. This type of *hin* inversion occurs frequently in *S.*  
420 Typhimurium (Johnson and Simon, 1985).

421 Another gene that was up-regulated in D37712 across all profiled conditions was the  
422 chromosomally-encoded *cysS<sup>chr</sup>*, that encodes cysteine-tRNA synthetase. Previously, we  
423 reported that transcription of the *cysS<sup>chr</sup>* of strain D23580 was uniformly down-regulated  
424 compared to 4/74, ~~a defect that~~. This down-regulation was compensated by the presence of a  
425 pBT1 plasmid-encoded cysteine-tRNA synthetase (Canals *et al.*, 2019a).

426 ~~Increased~~ Accordingly, the increased expression of the chromosomal *cysS* gene in D37712  
427 was consistent with the absence of the pBT1 plasmid. Our comparative

428 [transcriptomic](#) analysis showed that expression levels of *cysS* were similar in  
429 D37712 and 4/74 under all growth conditions.

430 Numerous virulence genes and operons were differentially expressed between D23580 and  
431 D37712. The SPI-16-associated *gtrABCa* operon (STM0557, STM0558, STM0559) is  
432 responsible for adding glucose residues to the O-antigen subunits of LPS that enhance the  
433 long-term colonisation of the mammalian gastrointestinal tract by *S. Typhimurium* ST19  
434 (Bogomolnaya *et al.*, 2008). We found that the *gtrABCa* genes were significantly up-regulated  
435 in several conditions in D37712, compared to both D23580 and 4/74.

436 The *spvABCD* operon of D37712 was up-regulated under non-SPI2-inducing growth  
437 conditions, compared to D23580. A signature pseudogene of ST313 L2.2 is the frameshift  
438 insertion in the *spvD* gene that generates a truncated version of the SpvD protein. The H1991  
439 mutation at position 199 and the associated 17 amino acid truncation is predicted to ablate  
440 the activity of the SpvD cysteine protease (Grabe *et al.*, 2016). [SpvD negatively regulates the](#)  
441 [NF- \$\kappa\$ B signaling pathway and promotes virulence of \*S. Typhimurium\* in mice.](#) The functional  
442 consequences of the *spvD* variant of ST313 L2.2 strain D37712 and the up-regulation of  
443 [expression of](#) the *spvABCD* operon remain to be established experimentally.

#### 445 **The SalComD37712 community transcriptional data resource**

446 To allow scientists to gain their own biological insights from analysis of this rich transcriptomic  
447 dataset, the transcriptomic and gene expression data generated in this study are presented  
448 online in a new community resource, [SalComD37712](#). The data resource shows the  
449 expression levels of all D37712 coding and non-coding genes, including both chromosomal  
450 and plasmid-encoded transcripts. The SalComD37712 website complements our existing  
451 SalComD23580 (<https://tinyurl.com/SalComD23580>) resource, and adds an inter-strain  
452 comparison of gene expression profiles between D37712 and D23580 as well as normalized  
453 gene expression values (TPM), using an intuitive heat map-based approach. [SalComD37712](#)  
454 included our published RNA-seq data (Canals *et al.*, 2019b), re-analysed with an updated  
455 bioinformatic pipeline and a combined reference genome (see Methods). This online resource  
456 facilitates the intuitive interrogation of transcriptomic data as described previously (Perez-  
457 Sepulveda and Hinton, 2018).

458 Additionally, we generated a unified genome-level browser that provides access to the *S.*  
459 *Typhimurium* L2.2 D37712 transcriptome, in the context of our previously published RNA-seq  
460 data for the L2.0 strain D23580 and the ST19 strain 4/74. This novel “combo” browser is  
461 available at [http://hintonlab.com/jbrowse/index.html?data=Combo\\_D37/data](http://hintonlab.com/jbrowse/index.html?data=Combo_D37/data).

462



### 463 Identification of phenotypes that distinguish ST313 sublineage L2.2 from L2.0.

464 To explore the phenotypic impact of the transcriptomic signature of L2.2 (D37712), we  
465 performed a series of motility experiments, fluorescence-based gene expression experiments  
466 and mixed-growth assays.

467 D33712 showed a significantly decreased level of motility on NonSPI2 minimal media,  
468 compared with both the ST19 strain 4/74 and the L2 D23580 strain (Fig. 5A). This finding was  
469 consistent with the transcriptomic data, which showed down-regulation of D37712 flagellar  
470 genes compared with D23580 in the NonSPI2 condition (Fig. 4). In contrast, no differential  
471 expression of flagellar genes was seen between D33712 and D23580 in the InSPI2 growth  
472 condition (Fig. 4). The decreased motility phenotype may be linked to the inversion of the *hin*  
473 element detailed above. The flagella system encodes a distinct type III secretion apparatus  
474 responsible for the dual functions of bacterial motility and activation of the mammalian innate  
475 immune system via TLR5 (Lai *et al.*, 2013).

476 A key transcriptomic finding for strain D33712 was the expression of SPI2 genes during  
477 growth in an unusual environmental condition (NonSPI2) (Fig. 3B-C and Fig. 4B). NonSPI2  
478 media differs from InSPI2 media by having a higher pH (pH7.4 versus pH5.8) and a higher  
479 level of phosphate (Löber *et al.*, 2006). This apparent differential expression of SPI2 genes at  
480 the transcriptomic level under non-inducing conditions led us to investigate the expression of  
481 SPI2 at a single cell level using fluorescence transcriptional fusions. First, we introduced an  
482 *ssaG*-GFP<sup>+</sup> transcriptional fusion into the chromosome of strains D33712 and D23580  
483 (Methods; Table S8) to interrogate expression of the key SPI2 operon with flow cytometry.  
484 Figure 5B shows that in NonSPI2 media, the *ssaG* promoter was expressed at a 62% higher  
485 level in D33712 than in D23580 confirming the results of the transcriptomic analysis.

486 Because only a proportion of *S. Typhimurium* cells express certain pathogenicity island-  
487 encoded genes during *in vitro* growth (Ackermann *et al.*, 2008; Hautefort *et al.*, 2003), we  
488 determined whether the increased level of expression of SPI2 genes (Fig. 4B) was caused by  
489 a higher proportion of D33712 cells expressing SPI2 than D23580 cells. Using derivatives of  
490 the two strains that carried the *ssaG*-GFP<sup>+</sup> construct, we determined the numbers of  
491 fluorescent and non-fluorescent cells with flow cytometry (Methods). Under non-inducing  
492 conditions, slightly more D37712 cells expressed the *ssaG* SPI2 promoter than D23580 cells  
493 (65% vs 60%, respectively) (Fig. 5C). ~~However, Although~~ this small difference was statistically  
494 significant (t-test: P<0.001, n=3), it did not account for the 62% increased level of non-induced  
495 SPI2 expression seen in Fig. 5B.

496 SPI2 expression is controlled by a complex regulatory system that operates at both a  
497 negative and positive level, involving silencing via H-NS (Lucchini *et al.*, 2006), activation by  
498 SlyA and SsrB (Fass and Groisman, 2009; Walthers *et al.*, 2011) as well as input from OmpR  
499 and Fis under non-inducing conditions (Osborne and Coombes, 2011). The ~~reason~~  
500 for mechanistic basis of the aberrant SPI2 expression in strain D37712 is worthy of further  
501 study. Possible explanations include the incomplete silencing of SPI2 transcription or the

1  
2  
3 502 partial activation of the SPI2 virulence genes under non-inducing growth conditions: by an  
4 503 unknown regulatory factor.  
5  
6  
7 504

8 505 **Increased fitness of *S. Typhimurium* ST313 sublineage L2.2 compared with L2.0 in**  
9 506 **minimal media.**

10  
11 507 It has become increasingly clear that distinct *Salmonella* pathovariants have evolved  
12 508 particular phenotypic properties that confer fitness advantages during infection of particular  
13 509 avian or mammalian hosts (Branchu *et al.*, 2018). Because *S. Typhimurium* ST313 L2.2  
14 510 appeared to have displaced *S. Typhimurium* ST313 L2.0 in Malawi, we speculated that *S.*  
15 511 *Typhimurium* ST313 L2.2 might have the competitive edge in some situations. Accordingly,  
16 512 we determined bacterial fitness using a mixed-growth competition assay (Wiser and Lenski,  
17 513 2015; Lian *et al.*, 2023). The competitive index was calculated in three different growth media  
18 514 using pair-wise combinations of strains D37712 and D23580. Two independent approaches  
19 515 were used to phenotypically distinguish the two strains, one based on antibiotic resistance  
20 516 (Fig. 5D) and the other based on fluorescent tagging (Fig. S5).

21  
22 517 To confirm that strains engineered to be kanamycin-resistant or gentamicin-resistant did not  
23 518 impact on fitness (Methods), we first verified that the tagged variants of D37712 or D23580  
24 519 did not confer a growth advantage in LB or NonSPI2 media (Fig. S7). Next, we used a mixed-  
25 520 growth assay to investigate fitness of *S. Typhimurium* ST313 L2.0 strain D23580 or *S.*  
26 521 *Typhimurium* ST313 L2.2 strain D37712 during growth in LB, or InSPI2 or NonSPI2 minimal  
27 522 media. The data show that both strains grew at similar levels following overnight mixed-  
28 523 growth in nutrient-rich LB media, but D37712 had a competitive advantage during mixed-  
29 524 growth in InSPI2 media (CI = 1.79;  $P < 0.05$ ) and a greater competitive edge in NonSPI2 media  
30 525 (CI = 2.20;  $P < 0.0001$ ).

31  
32 526 We then used an independent fluorescence-based approach to assess the fitness of strains  
33 527 D23580 and D37712 during mixed-growth in NonSPI2 media. This time, the strains were  
34 528 engineered to carry either mScarlet or sGFP2 proteins and the mixed-growth experiments  
35 529 involved pair-wise comparisons of reciprocally-tagged strains. The flow cytometric data  
36 530 showed that in both cases D37712 had a significant competitive advantage in NonSPI2 media  
37 531 (Fig. S5 and S6).

38  
39 532 This combination of antibiotic resistance-based and fluorescence-based competitive index  
40 533 experiments lead us to conclude that *S. Typhimurium* ST313 L2.2 strain D37712 had a clear  
41 534 fitness advantage over *S. Typhimurium* ST313 L2.0 strain D23580 during mixed-growth in two  
42 535 formulations of minimal media. The molecular basis of this fitness advantage remains to be  
43 536 established.  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56

57 537  
58  
59  
60

538 **Perspective**

539 Here, we report that *S. Typhimurium* ST313 L2.0 has been clonally replaced by the ST313  
540 sublineages L2.2 and L2.3 as a cause of bloodstream infection in Blantyre, Malawi. In 2018,  
541 L2.2 represented the majority of the ST313 strains isolated from hospitalised patients in  
542 Malawi at the Queen Elizabeth Central Hospital. Our comparative genomic analysis of ST313  
543 L2.3 identified 30 chromosomal alterations, one of which generated a deletion of the *sseI*  
544 effector gene.

545 Our RNA-seq-based analysis of ST313 L2.2 involved a detailed comparison versus ST313  
546 L2.0 which revealed a key difference involving SPI2 expression. Following initially  
547 observations at the transcriptomic level in the ST313 L2 and L2.2 strains grown in a pH-  
548 neutral minimal medium (NonSPI2), the increased expression of SPI2 was confirmed at the  
549 single cell level using an *ssaG* transcriptional fusion.

550 A series of experiments showed that the ST313 L2.2 strain D37712 had a competitive  
551 advantage over L2 strain D23580 during mixed-growth in minimal media. We propose that  
552 this increased fitness of *S. Typhimurium* ST313 L2.2 has contributed to the replacement of  
553 ST313 L2.0 in Malawi in recent years.

554 Previously, we compared three virulence properties of the *S. Typhimurium* ST313 L2.0  
555 D23580 and ST313 L2.2 D37712 strains. First, experiments involving Mucosal Invariant T  
556 (MAIT) cells showed that both D37712 and D23580 fail to elicit the high level of activation of  
557 MAIT cells that characterises infection by *S. Typhimurium* ST19 4/74 (Preciado-Llanes *et al.*,  
558 2020). Second, the D37712 and D23580 strains stimulate similar levels of up-regulation of  
559 IL10 gene expression upon infection of human dendritic cells (Aulicino *et al.*, 2022). Third, we  
560 showed that both D37712 and D23580 express similarly high levels of the PgtE virulence  
561 factor that is responsible for the ability of *S. Typhimurium* ST313 to survive human serum-  
562 killing (Hammarlöf *et al.*, 2018). These findings lead us to conclude that the comparative  
563 genomic and transcriptomic differences that distinguish *S. Typhimurium* ST313 L2.0 strain  
564 D23580 from ST313 L2.2 D37712 (Fig. 4) do not modulate the ability of the pathogens to  
565 activate human MAIT cells or dendritic cells, or to influence the PgtE-mediated serum survival  
566 phenotype of *S. Typhimurium* ST313.

567 Ideally, the implications of the competitive advantage of ST313 L2.2 would be determined in  
568 the context of pathogenesis. However, we lack an informative infection model for *S.*  
569 *Typhimurium* ST313 (Lacharme-Lora *et al.*, 2019), and it is not yet possible to experimentally  
570 determine whether the improved fitness of L2.2 significantly enhances the success of ST313  
571 during infection of humans.

572 Here we have investigated the intricate interplay of gene function that **is**  
573 **underpinning** the success of *S. Typhimurium* ST313 L2.2. **We** hope **is** that our  
574 findings **might** contribute to future therapeutic or prophylactic strategies for combatting  
575 iNTS infections in the African setting.

## 576 **Materials and methods**

### 577 **Bacterial strains**

578 To investigate the evolutionary dynamics of *S. Typhimurium* ST313 L2 in Malawi over a 22  
579 year period, we focused on the large collection of 8,000 *S. Typhimurium* isolates derived from  
580 bloodstream infection in hospitalised patients at the Queen Elizabeth Central Hospital,  
581 Blantyre, Malawi (Feasey *et al.*, 2015). The collection was assembled by the Malawi–  
582 Liverpool–Wellcome Trust Clinical Research Programme (MLW) between 1996 and 2018; the  
583 precise annual numbers of isolates are shown in Fig. 1C. A random sub-sampling strategy  
584 was used to select 608 isolates for whole-genome sequencing, which included 549 *S.*  
585 *Typhimurium* ST313 isolates (Pulford *et al.*, 2021).

586

587 The two *S. Typhimurium* ST313 strains that are the focus of this study are D23580 and  
588 D37712. D23580 was isolated from a Malawian 26-month-old child with malaria and anaemia  
589 in 2004. D37712 was isolated from the blood of an HIV-positive Malawian male child in 2006.  
590 These two African *Salmonella* strains have been deposited in the National Collection of Type  
591 Cultures (NCTC). The D23580 (lineage 2.0) strain is available as NCTC 14677. The ST313  
592 sublineage 2.2 strain D37712 is available as NCTC 14678. All bacterial strains are detailed in  
593 Table S8.

### 594 **Genome sequencing**

595 The assembled genome and annotation of D23580 (Kingsley *et al.*, 2009; Canals *et al.*,  
596 2019b) (L2.0) was obtained from the European Nucleotide Archive (ENA) repository (EMBL-  
597 EBI) under accession PRJEB28511 (<https://www.ebi.ac.uk/ena/data/view/PRJEB28511>). For  
598 genome sequencing of D37712 (L2.2), DNA was extracted using the Bioline mini kit, and  
599 quality was assessed using gel electrophoresis (0.5% agarose gel, at 30 volts for 18 h). The  
600 genome was generated by a combination of long read sequencing with a PacBio RS II and  
601 short-read sequencing on an Illumina HiSeq machine at the Center for Genome Research,  
602 University of Liverpool, United Kingdom.

603 Sequence reads were quality checked using FastQC version 0.11.9 (Andrews, 2010) and  
604 MultiQC version 1.8 (Ewels *et al.*, 2016), trimmed using Trimmomatic (Bolger *et al.*, 2014).  
605 Hybrid assembly of the Illumina and PacBio sequence reads was done with Unicycler v0.4.7  
606 (Wick *et al.*, 2017).

607 The assembled genome of *S. Typhimurium* SDT313 L2.2 strain D37712 was deposited in  
608 Genbank (GCA\_014250335.1, assembly ASM1425033v1). Raw sequencing reads were  
609 deposited for both PacBio and Illumina, under BioProject ID PRJNA656698. Sequence Read  
610 Archive (SRA) database IDs are: SRR12444880 for Illumina and SRR12444881 for PacBio.

### 611 **Comparative genomic analyses**

612 To generate the data summarised in Fig. 1C, sequencing data of 29 *S. Typhimurium* ST313  
613 strains (Msefula *et al.*, 2012) were downloaded from EMBL-EBI database  
614 (<https://www.ebi.ac.uk>, accession number ERA015722). Sequence reads were assembled  
615 using Unicycler v0.4.8 (Wick *et al.*, 2017). The quality of the assemblies was assessed by  
616 Quast v5.0.2 (Gurevich *et al.*, 2013). The N50 value of all assemblies was >20kb, and the  
617 number of contigs was <600.

618 To construct the phylogenetic tree (Fig. 1C), *Salmonella* Typhimurium strains D23580,  
619 D37712, LT2 (GCA\_000006945.2), DT104 (GCA\_000493675.1), 4/74 (GCA\_000188735.1),  
620 and A130 (GCA\_902500285.1) were added as contextual genomes. Roary was used to make  
621 the core gene alignment, construct the gene presence/absence matrix and identify  
622 orthologous genes (Page *et al.*, 2015). Phylogenetic trees were constructed using  
623 Randomized Accelerated Maximum Likelihood (RAxML) (Stamatakis *et al.*, 2005), and were  
624 visualised with the interactive Tree of Life online tool (iTOL) (Letunic and Bork, 2006).

625 The assembled genome and annotation of *S. Typhimurium* ST19 representative strain 4/74  
626 (Richardson *et al.*, 2011) were obtained from GenBank (Accession number  
627 GCF\_000188735.1), while the raw sequencing data of 27 *S. Typhimurium* ST313 strains  
628 described in a previous study (Msefula *et al.*, 2012) were downloaded from EMBL-EBI  
629 database (<https://www.ebi.ac.uk>, accession number ERA015722). The raw reads were  
630 assembled using Unicycler v0.4.8 (Wick *et al.*, 2017). The quality of the assemblies was  
631 assessed by Quast v5.0.2 (Gurevich *et al.*, 2013). The N50 value of all assemblies  
632 was >20kb, and the number of contigs was <600.

633 To identify SNPs, Snippy v4.4.0 (<https://github.com/tseemann/snippy>) was used to map the  
634 raw reads against the 4/74 genome. To detect pseudogene-associated SNPs/indels in each  
635 sub-lineage, the SNPs/indels that caused nonsense or frameshifted mutations were filtered.  
636 The identifications and names of the disrupted genes were summarised, then the wild type  
637 gene sequences were extracted from the 4/74 genome. To validate the pseudogene-  
638 associated SNPs/indels, the wild type gene sequences were used to make a BLAST  
639 database with BLAST 2.9.0+ (Camacho *et al.*, 2009). The 29 genome assemblies were  
640 queried against the databases, using the BLASTn algorithm to confirm the nonsense and  
641 frameshifted mutations in all isolates.

#### 642 **Phylogenetic analysis of African *Salmonella* Typhimurium isolates dating from 1966 -** 643 **2018**

644 To examine the overall population structure of *Salmonella* Typhimurium responsible for blood  
645 infection in Malawi (Fig. 1AB and Fig. S1), the raw reads of 707 published genome  
646 sequences were downloaded (Table S7). [Trimmomatic v0.36 \(Bolger, A. M., Lohse, 2014\)](#)  
647 [was used to trim adapters and Seqtk v1.2-r94 \(https://github.com/lh3/seqtk\) was used to trim](#)  
648 [low-quality regions using the trimfq flag. Fastqc v0.11.5 \(https://www.](#)  
649 [bioinformatics.babraham.ac.uk/projects/fastqc/\) and multiqc v1.0 \(http://multiqc.info\) were](#)  
650 [used to pass sequence reads according to the following criteria: passed basic quality](#)



651 statistics, per base sequence quality, per base N content, adapter content and an average  
652 GC content of between 47% and 57%. Only high-quality reads were used in the downstream  
653 analysis. Sequence reads were aligned to the *S. Typhimurium* D23580 genome using Snippy  
654 v4.4.0- with parameter “- - mincov 5”. The recombination sites of the alignment were removed  
655 by Gubbins (Croucher *et al.*, 2015), and the phylogenetic tree was built with Raxml-ng (Kozlov  
656 et al., 2019). using GTR\_G models ad 100 bootstraps. The tree was rooted on *Salmonella*  
657 Typhi strain CT18 (GCA\_000195995.1) as the outgroup. The tree was visualised with the  
658 interactive Tree of Life online tool (iTOL) (Letunic and Bork, 2006). The sub-lineages were  
659 identified with rHierBAPS (Tonkin-Hill *et al.*, 2018). The stacked-area chart and the bar chart  
660 showing the percentage and number of isolates from each sub-lineage were made in MS  
661 Excel.

## 662 **RNA purification and growth conditions**

663 Initially, a screen of transcriptomic gene expression was performed without biological  
664 replicates. Total RNA was purified using TRIzol from *S. Typhimurium* D37712 grown in 15  
665 different conditions as described previously (Kröger *et al.*, 2013). To generate statistically-  
666 robust gene expression profiles, total RNA was subsequently purified using TRIzol from *S.*  
667 *Typhimurium* D37712 grown in four *in vitro* growth conditions (ESP, anaerobic growth,  
668 NonSPI2, InSPI2) with three biological replicates as described previously (Kröger *et al.*,  
669 2013). RNA was isolated from intra-macrophage D37712 following infection of RAW264.7  
670 murine macrophages using our published protocol (Srikumar *et al.*, 2015).

## 671 **RNA-seq of *S. Typhimurium* strain D37712 using Illumina technology**

672 For transcriptomic analyses, cDNA samples were prepared from *S. Typhimurium* RNA by  
673 Vertis Biotechnologie AG (Freising, Germany). RNA was first treated with DNase and purified  
674 using the Agencourt RNAClean XP kit (Beckman Coulter Genomics). RNA samples were  
675 sheared using ultrasound, treated with antarctic phosphatase and re-phosphorylated with T4  
676 polynucleotide kinase. RNA fragments were poly(A)-tailed using poly(A) polymerase and an  
677 RNA adapter was ligated to the 5'- phosphate of the RNA. First-strand cDNA synthesis was  
678 performed using an oligo(dT)-adapter primer and M-MLV reverse transcriptase. The resulting  
679 cDNA was PCR-amplified to about 10-20 ng/μl. The cDNA was purified using the Agencourt  
680 AMPure XP kit. The cDNA samples were pooled using equimolar amounts and size  
681 fractionated in the size range of 200-500 bp using preparative agarose gels. The cDNA pool  
682 was sequenced on an Illumina NextSeq 500 system using 75 bp read length.

683 For the biological replicates of the four growth conditions (ESP, anaerobic growth  
684 (abbreviated as NoO<sub>2</sub>), NonSPI2, and InSPI2) and the intra-macrophage RNA, cDNA  
685 samples were generated as above with some improvements in library preparation. First, after  
686 fragmentation with ultrasound, an oligonucleotide adapter was ligated to the 3' end of the  
687 RNA molecules. Second, first-strand cDNA synthesis was performed using M-MLV reverse  
688 transcriptase and the 3' adapter as primer, and, after purification, the 5' Illumina TruSeq  
689 sequencing adapter was ligated to the 3' end of the antisense cDNA. Sequencing of the

690 cDNA was performed as described above. All raw sequencing reads were deposited to the  
691 Gene Expression Omnibus (GEO) database under accession GSE161403.

### 692 **RNA-seq and dRNA-seq read processing and visualization**

693 RNA-seq data from *S. Typhimurium* 4/74 and D23580 were extracted from previously  
694 published experiments (Kröger *et al.*, 2013; Srikumar *et al.*, 2015; Canals *et al.*, 2019b; GEO  
695 dataset GSE119724). A combined reference genome was generated that contained the  
696 D23580 chromosome plus plasmids pBT1, pBT2, pBT3, pSLT-BT (from D23580) and the  
697 D37712 plasmid pCol1B9<sup>D37712</sup>. All reads were aligned and quantified using Bacpipe v0.8a  
698 (<https://github.com/apredeus/multi-bacpipe>). Briefly, basic read quality control was performed  
699 with FastQC v0.11.8. RNA-seq reads were aligned to the genome sequence using STAR  
700 v2.6.0c using “--alignIntronMin 20 --alignIntronMax 19 --outFilterMultimapNmax 20” options. A  
701 combined GFF file was generated by Bacpipe, where all features of interest were listed as a  
702 “gene”, with each gene identified by a D37712 locus tag. Subsequently, read counting was  
703 done by featureCounts v1.6.4, using options “-O -M --fraction -t gene -g ID -s 1”. For  
704 visualization, scaled gedGraph files were generated using bedtools genomecov with a scaling  
705 coefficient of 10<sup>9</sup>/(number of aligned bases), separately for sense and antisense DNA  
706 strands. Bedgraph files were converted to bigWig using bedGraphToBigWig utility  
707 ([http://hgdownload.soe.ucsc.edu/admin/exe/linux.x86\\_64/](http://hgdownload.soe.ucsc.edu/admin/exe/linux.x86_64/)). Coverage tracks, annotation, and  
708 genome sequence were visualized using JBrowse v1.16.6. Transcripts Per Million (TPM)  
709 were calculated for all samples and used as absolute expression values (Table S5). A  
710 conservative cut-off was used to distinguish between expressed (TPM >10) and not  
711 expressed (TPM ≤10), as we previously described (Kröger *et al.*, 2013). Relative expression  
712 values were calculated by dividing the TPM value for one condition in one strain by the TPM  
713 value for the same condition in a different strain. Before the calculation, all TPM values below  
714 10 were set up to 10. A conservative fold-change cut-off of 3 was used to highlight differences  
715 in expression between strains.

### 716 **Differential gene expression analysis with multiple biological replicates**

717 For differential expression analysis of *S. Typhimurium* strains 4/74, D23580, and D37712, the  
718 raw counts (Table S4) from 3-5 biological replicates in four growth conditions were used  
719 (ESP, anaerobic growth (abbreviated as NoO<sub>2</sub>), NonSPI2, and InSPI2). Differential  
720 expression analysis was done using DESeq2 v1.24.0 with default settings. A gene was  
721 considered to be differentially expressed if the absolute value of its log<sub>2</sub> fold change was at  
722 least 1 (i.e. fold change > 2), and adjusted p-value was < 0.001.

### 723 **The SalComD37712 community data resource, and the associated Jbrowse genome 724 browser**

725 SalCom provides a user-friendly Web interface that allows the visualisation and compaison of  
726 gene expression values across multiple conditions and between strains. Particular genes can  
727 be selected through pre-defined lists of interest, such as all sRNAs or all genes belonging to a

728 specific pathogenicity island. The resulting heatmap-style display highlights expression  
729 differences, and provides access to the rich, manually curated annotation of strains D37712  
730 and D23580. The actual values behind the display can be downloaded for further processing,  
731 and a link connects the current view to a genome browser interface.

732 Visualisation of all the RNA-seq and dRNA-seq (TSS) coverage tracks in JBrowse 1.16.6  
733 shows sequence reads mapped against the combined reference genome described above.  
734 Overall, the genomic distance between strains 4/74 and D23580 (approximately 1000 SNPs,  
735 or ~1 SNP per 5000 nucleotides), and between D37712 and D23580 (approximately 30  
736 SNPs, ~1 SNP per 150,000 nucleotides) allowed the alignment of RNA-seq reads to the  
737 simplified combined reference genome without significant loss of reads. The combined  
738 reference genome facilitated a direct comparison of gene coverage as well as transcriptional  
739 start sites. The unified browser is hosted at  
740 [http://hintonlab.com/jbrowse/index.html?data=Combo\\_D37/data](http://hintonlab.com/jbrowse/index.html?data=Combo_D37/data).

#### 741 **Phenotypic and mixed competitive growth experiments**

742 The swimming motility of *S. Typhimurium* strains D37712, D23580 and 4/74 was determined  
743 by a plate assay (Canals *et al.*, 2019b), which involved spotting 3  $\mu$ L overnight culture onto  
744 0.3% LB agar. Relative motility of the three strains was assessed by migration diameter after  
745 4h and 8h of incubation at 37°C.

746 Relative expression of the *ssaG* SPI2 promoter in strains D23580 and D37712 was measured  
747 at the single cell level via GFP fluorescence. Following the construction of a kanamycin-  
748 sensitive derivative of D23580 (strain JH4235), a *PssaG::gfp*<sup>+</sup> transcriptional fusion was  
749 incorporated into the chromosome of JH4235 and D37712 by inserting the *gfp*<sup>+</sup> gene  
750 downstream of the *ssaG* gene, under the control of the *PssaG* promoter. The *PssaG::gfp*<sup>+</sup>  
751 D23580 derivative (JH4692), and the *PssaG::gfp*<sup>+</sup> D37712 derivative (JH4693) are listed in  
752 Table S8.

753 The strains JH4692 and JH4693 were genome sequenced to confirm the integrity of the  
754 transcriptional fusions, and to verify that unintended nucleotide changes had not arisen.  
755 Following growth in 25 mL non-inducing NonSPI2 media in a 250 mL flask at 37°C with  
756 shaking at 220 rpm for approximately 8 hours until OD<sub>600</sub>=0.3, fluorescence was determined  
757 with a BD FACSAria Flow Cytometer. The relative fluorescence of the two strains JH4692 and  
758 JH4693, and the numbers of individual fluorescent bacteria that expressed the *PssaG::gfp*<sup>+</sup>  
759 promoter, were determined with FlowJo VX software.

760 The relative fitness of *S. Typhimurium* strains D37712 and D23580 was assessed in two  
761 independent mixed-growth experiments. First, kanamycin-resistant derivatives of each strain  
762 were constructed by inserting the *aph* kanamycin resistance gene into the chromosome at the  
763 intergenic region between the *STM4196* and *STM4197* genes, a region that we have  
764 previously shown to be transcriptionally silent (Canals *et al.*, 2019b). The strains were  
765 designated D23580::Km<sup>R</sup> JH3794 and D37712::Km<sup>R</sup>, JH4232. Mixed cultures of wild-type or

1  
2  
3 766 kanamycin-resistant derivatives of each strain were grown overnight in LB, InSPI2 and  
4 767 NonSPI2 media in a 250 mL flask at 37°C with shaking at 220 rpm. Following plating on LB  
5 768 agar or LB + kanamycin, colonies were counted and the ratio of bacterial strains was  
6 769 determined. To confirm that the insertion of kanamycin resistance at the intergenic region  
7 770 between *STM4196* and *STM4197* did not impact upon fitness, a mixed-growth experiment  
8 771 was done in both LB and NonSPI2 media (Fig. S7).  
9  
10  
11  
12 772 Second, to independently assess relative fitness, Tn7-based plasmids (Schlechter and  
13 773 Remus-Emsermann, 2019) were used to construct chromosomal sGFP2 and mScarlet  
14 774 derivatives of *S. Typhimurium* strains D23580 (sGFP2 derivative: JH4694; mScarlet  
15 775 derivative: JH4695) and D37712 (sGFP2 derivative: JH4696; mScarlet derivative: JH4697).  
16  
17 776 The gene cassettes were inserted into the *S. Typhimurium*Tn7 insertion site between the  
18 777 gene *STMMW\_38451* and *glmS*. Mixed cultures of pairs of fluorescently-labelled strains were  
19  
20 778 grown in NonSPI2 media at 37°C with shaking at 220 rpm for approximately 8 hours until  
21  
22 779  $OD_{600}=0.3$ . Levels of green and red fluorescence were determined with a BD FACSAria Flow  
23  
24 780 Cytometer.  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

781 **Figure Legends**

782 **Fig. 1. Emergence of *S. Typhimurium* ST313 sublineages L2.2 and L2.3 in Malawi. (A)**  
 783 Evolutionary dynamics of *S. Typhimurium* lineages in Blantyre, Malawi from 1996 to 2018. **A**  
 784 [maximum likelihood tree constructed with 1000 bootstraps using the GTRGAMMA model in](#)  
 785 [RaxML rooted on ST19, LT2.](#) The genomes of 549 *S. Typhimurium* ST313 isolates from  
 786 bacteraemic patients at the Queen Elizabeth Hospital in Blantyre, Malawi were used for this  
 787 analysis. The proportions of the five lineages/sublineages are shown. **(B)** The total number of  
 788 isolates of each lineage/sublineage per year. **(C)** Phylogenetic comparison between  
 789 representative strains of *S. Typhimurium* ST19 and four ST313 lineages/sublineages (L1, L2.0,  
 790 L2.2, L2.3) showing the presence and absence of plasmids, prophages and the *spvD*  
 791 pseudogene. The complete phylogenetic analysis of 707 *S. Typhimurium* genomes is shown in  
 792 Fig.S1.

793 **Fig. 2. Key genetic similarities and differences between the chromosome and plasmid**  
 794 **profiles of D23580 (lineage 2) and D37712 (L2.2). (A)** A comparison of the D23580 (L2.0)  
 795 and D37712 (L2.2) chromosomes. The dots around the chromosome are different kinds of  
 796 SNPs identified. Phages and *Salmonella* pathogenicity islands are shown in blue and red  
 797 respectively. **(B)** Plasmid profile of D37712 versus D23580. The pSLT-BT virulence plasmid is  
 798 present in both D37712 and D23580, and carries the Tn-21 transposable element; **(C)** pCol1B9  
 799 is present in D37712 and absent from D23580 (D) pBT3 is present in both D37712 and D23580.  
 800 **(E)** Absence of *ssel* gene and the STM1050 coding sequence in L2.2 (D37712), as compared  
 801 to *S. Typhimurium* ST19 4/74 and *S. Typhimurium* ST313 L2.0 (D23580). **(F)** List of  
 802 pseudogenes in D37712 and D23580, with reference to 4/74. The colour blue means  
 803 pseudogene/disrupted gene while grey indicates functional genes. *macB* is a pseudogene in  
 804 D23580 (L2.0) but not in L2.2, while *spvD* is a pseudogene in L2.2 but not in L2.0. All L2.2  
 805 strains share similar pseudogenes.

806 **Fig. 3. General comparison of expression profiles of strains 4/74, D23580, and D37712**  
 807 **under 17 different *in vitro* conditions. (A)** Principal component analysis (PCA) plot of the  
 808 individual RNA-seq samples, indicating the overall similarity in gene expression between the  
 809 three strains. The 17 growth conditions have been defined previously (Kröger *et al.*, 2013).  
 810 **(B)** Visualization of SPI-2 pathogenicity island expression with the Jbrowse genomic browser,  
 811 [underat](#) mid-exponential phase (MEP), InSPI2, and NonSPI2 *in vitro* conditions-, [which can](#)  
 812 [be accessed here.](#) **(C)** Boxplot visualization of SPI-2 gene expression [underat](#) mid-  
 813 exponential phase (MEP), InSPI2, and NonSPI2 *in vitro* conditions. [The y-axis shows the](#)  
 814 [combined log TPM values for 45 genes located in the SPI2 pathogenicity island, namely](#)  
 815 [ssaU, ssaT, ssaS, ssaR, ssaQ, ssaP, ssaO, ssaN, ssaV, ssaM, ssaL, ssaK, STnc1220,](#)  
 816 [STM1410, ssaJ, ssal, ssaH, ssaG, sseG, sseF, sscB, sseE, sseD, sseC, sscA, sseB, sseA,](#)  
 817 [ssaE, ssaD, ssaC, ssaB, ssrA, ssrB, orf242, orf319, orf70, ttrR, ttrS, ttrC, ttrB, ttrA, orf408,](#)  
 818 [orf245, orf32, and orf48.](#) The elevated expression of SPI-2 genes in strain D37712 cultured  
 819 [underin](#) NonSPI2 [conditionsmedia](#) is highlighted in a red box.



820 **Fig. 4. Differential gene expression of *S. Typhimurium* 4/74, D37712, and D23580 under**  
 821 **4 in vitro conditions. (A)** Boxplots indicating the number of differentially-expressed genes  
 822 identified in the following *in vitro* growth conditions: early stationary phase, ESP; anaerobic  
 823 growth, NoO<sub>2</sub>; SPI-2 inducing medium, InSPI2; SPI-2 non-inducing minimal medium, NonSPI2.  
 824 Multiple (3 to 5) biological replicates were used for comparison. DESeq2 was used for  
 825 differential analysis; only genes with  $|\log_2FC| \geq 1$  and with adjusted  $p$ -value  $\leq 0.001$  were  
 826 retained. **(B)** Heatmap of the genes differentially expressed between D23580 and D37712.  
 827 Functional groups and operons of interest are highlighted on the right of Panel B.

828 **Fig. 5. Phenotypes that distinguish ST313 L2.2 from ST313 L2.0. (A)** Swimming motility  
 829 assay of strains D23589, D37712 and 4/74, with a representative plate shown on the left.  
 830 Average migration diameters were measured after 4 and 8 hours. Each bar represents the  
 831 mean of three biological replicates, with *error bars representing* standard deviation.  
 832 Significant difference (\*\*\*) indicates  $P$  value ( $t$  test)  $< 0.001$ . In Panels B & C, comparison of  
 833 *ssaG* expression by flow cytometry using D23580 and D37712 derivatives containing a  
 834 chromosomal *ssaG*-GFP<sup>+</sup> transcriptional fusion, strains SZS008 and SZS032, respectively.  
 835 Cells were collected at 8 hours after inoculation in NonSPI2 media. Ten thousand events were  
 836 acquired for each sample. **(B)** Mean fluorescent intensity signal of *ssaG*-GFP<sup>+</sup> for D23580  
 837 (SZS008, dark grey) and D37712 (SZS032, grey) grown in NonSPI2 media. Significant  
 838 difference (\*\*\*) indicates  $P$  value ( $t$  test)  $< 0.001$ . **(C)** The proportions of bacterial cells that  
 839 expressed *ssaG*-GFP<sup>+</sup> during growth in NonSPI2 media was determined. Percentage of  
 840 positive GFP-expressing (green) and negative non-fluorescent cells (white) for *ssaG* expression  
 841 in each sample D23580 (SZS008) and D37712 (SZS032) is shown. Each bar represents the  
 842 mean of three biological replicates, error bars show standard deviation. Significant difference  
 843 (\*\*\*) indicates  $P$  value ( $t$  test)  $< 0.001$ . **(D)** Relative fitness of wild-type D23580 and D37712  
 844 and their kanamycin-resistant derivatives. Bacterial numbers were determined by after  
 845 overnight culture of a 1:1 mixture (wild-type versus Km<sup>R</sup>) in NonSPI2 (red LB (left), InSPI2  
 846 (blue middle) and LB (black NonSPI2 (right) media. Each bar dot represents the log-transformed  
 847 mean competitive index of three biological replicates with *error bars* representing 95%  
 848 confidence interval from standard error.  $P$  values were determined by  $t$  test (\*\*\*:  $P < 0.001$ ; \*\*:  
 849  $P < 0.01$ ; \*:  $P < 0.05$ ; ns: no significance) deviation. A competitive index of 1 indicates the equal  
 850 fitness of two strains, while a log number higher than 10 reflects the increased fitness of  
 851 kanamycin-resistant derivatives.  $P$  values were determined by  $t$  test (\*\*\*:  $P < 0.001$ ; \*\*:  $P <$   
 852  $0.01$ ; \*:  $P < 0.05$ ; ns: not significant).

853

#### 854 Supporting information

855 **Fig. S1, Maximum-likelihood phylogeny of 707 African *S. Typhimurium* isolates.** All  
 856 genome sequences have been published (Msefula *et al.*, 2012, Pulford *et al.*, 2021, Canals *et*  
 857 *al.*, 2019b). Raw sequence reads were aligned to the *S. Typhimurium* D23580 genome  
 858 (FN424405) using Snippy. The recombination sites of the alignment were removed by Gubbins,

859 and the phylogenetic tree was built with Raxml-ng. The tree is rooted on *Salmonella* Typhi strain  
 860 CT18 as the outgroup. The MLST sequence types, HierBAPS level 1 and level 2 clusters are  
 861 shown in coloured concentric rings as indicated. The *S. Typhimurium* ST313 isolates are  
 862 categorised as Lineage 1, Lineage 2 or Lineage 3 according to HierBAPS level 1 clustering.  
 863 ST313 Lineage 2 was then sub- divided into 3 sub-lineages according to HierBAPS level 2  
 864 clustering: ST313 L2.0, ST313 L2.2 and ST313 L2.3. The metadata and lineage designations  
 865 of all the *S. Typhimurium* isolates are in Table S7.

866 **Fig. S2.** PCR-based confirmation of the deletion of the *sseI* gene from *S. Typhimurium* L2.2  
 867 D37712. Arrows from left to right show the forward strand while the left strand is shown by  
 868 arrows from right to left. However, *sseI* gene in D23580 is a pseudogene with a SNP  
 869 indicated as a red line.

870

871 **Fig. S3.** Genomic comparison of plasmids pCol1B9<sup>4/74</sup> and pCol1B9<sup>D37712</sup> using Artemis  
 872 Comparison Tool (ACT). Bottom panel details the differences observed in the most divergent  
 873 regions, including colicin toxin-antitoxin system (in pCol1B9) and *impC-umuC-umuD* operon (in  
 874 pCol1B9).

875 **Fig. S4. RDAR Phenotypes of 4/74, D23580, D37712 and BKQZM9.** The top panel shows  
 876 the RDAR morphology assay and the bottom panel shows a complementary experiment that  
 877 involves the induction of biofilm formation on 1% tryptone agar (MacKenzie *et al.*, 2019).  
 878 Strain 4/74 was used as a RDAR-positive control, which has concentric rings and a wrinkled  
 879 appearance (Pulford *et al.*, 2021). The *S. Typhimurium* ST313 L3 strain BKQZM9 is shown for  
 880 comparative purposes.

881 **Fig. S5. Competitive index analysis of D23580 and D37712 using fluorescently-tagged *S.***  
 882 ***Typhimurium* strains (A)** Km<sup>R</sup>-sGFP2 and Gm<sup>R</sup>-mScarlet were inserted into the transposon  
 883 Tn7 site of D23580 or D37712. Bent arrows represent promoters and directional arrows  
 884 represent genes. **(B)** A 1:1 mix of Km<sup>R</sup>-sGFP2 and Gm<sup>R</sup>-mScarlet marked strain was inoculated  
 885 in NonSPI2 media, followed by an overnight incubation in 37°C. **Percentage** The percentage of  
 886 sGFP2 (green) and mScarlet (Red) -marked cells was **measureddetermined** by flow cytometry.  
 887 Raw data are shown in Figure S7, 10,000 events were acquired for each sample. **(C)**  
 888 Competitive index analysis of Km<sup>R</sup>-sGFP2 and Gm<sup>R</sup>-mScarlet marked strain. Bacterial numbers  
 889 were determined by counting CFU for overnight culture of a 1:1 mixture in NonSPI2 media.  
 890 Each dot represents a single biological replicate and the lane represents mean value. A  
 891 competitive index of 1 indicates the equal fitness of two strains, while a number higher than 1  
 892 reflects an increased fitness of D37712.

893 **Fig. S6. Raw flow cytometric data related to Fig. S5B. (A)** JH4695 + JH4698 and **(B)**  
 894 JH4696 + JH4697.- A 1:1 mix of the Km<sup>R</sup>-sGFP2 and Gm<sup>R</sup>-mScarlet marked strains were  
 895 inoculated in NonSPI2 media, followed by growth at 37°C until OD<sub>600</sub> = 0.3. The X-axis  
 896 (labelled FITC) shows the GFP level and the Y-axis (labelled PE Yell-Grn) indicates the

897 mScarlet level. Quadrant gates were used to separate four populations, and the black  
898 numbers indicate the percentage of events in each quadrant. In total, 10,000 events were  
899 acquired for each sample.

900 **Fig. S7. The insertion of GFP-Km or RFP-Gm did not impact on fitness.** A 1:1 mix of  
901 Km<sup>R</sup>-sGFP2 and Gm<sup>R</sup>-mScarlet marked strains were inoculated in LB or NonSPI2 media,  
902 followed by overnight incubation in 37°C. The competitive index (CI) was calculated using the  
903 formula  $(CFU_{Gm})/(CFU_{Km})$ . Each dot represents the CI from a single replicate and the  
904 horizontal bars indicate the mean of each dataset.

#### 905 **Supplementary data**

906 **Table S1:** SNP and indel variants that differentiate L2.2 (strain D37712) and L2.3 (strain  
907 D49679).

908 **Table S2:** SNP and indel variants that differentiate L2.2 (strain D37712) and L2.0 (strain  
909 D23580).

910 **Table S3:** Pseudogenes carried by ST19 and ST313 L2.0 and L2.2 (strains 4/74, D23580 and  
911 D37712).

912 **Table S4:** Raw read counts for all processed RNA-seq samples shown in Figures 3 and 4  
913 (strains 4/74, D23580, and D37712).

914 **Table S5:** TPM values for all processed RNA-seq samples shown in Figures 3 and 4 (strains  
915 4/74, D23580, and D37712).

916 **Table S6:** ~~Differential DESeq2-based differential gene~~ expression analysis ~~using DESeq2~~ for  
917 strains D23580 vs D37712 grown in four *in vitro* conditions.

918 **Table S7:** Metadata and lineage designations of the 708 *S. Typhimurium* isolates used to  
919 generate the maximum likelihood phylogeny (Fig. S1).

920 **Table S8:** Bacterial strains used in this study.

#### 921 **Acknowledgements**

922 The authors thank Brian Coombes and Rob Kingsley for their constructive comments during  
923 the peer review process. We are grateful to present and former members of the Hinton  
924 laboratory for helpful discussions, and to Paul Loughnane for his expert technical assistance.  
925

926  
927 This work was supported by a Wellcome Trust Investigator award [grant numbers  
928 106914/Z/15/Z and 222528/Z/21/Z] to J.C.D.H., and by the Malawi-Liverpool-Wellcome  
929 Research Centre Director's Fund. B.K. was funded by an AESA-RISE fellowship from the  
930 African Academy of Sciences [Grant Number: RPDF-18-04]. For the purpose of open access,  
931 the authors have applied a CC BY public copyright licence to any Author Accepted  
932 Manuscript version arising from this submission.  
933

#### 934 **Author contributions**

935

*Kumwenda et. al.*

27

1  
2  
3 936 **Conceptualization:** B.K., R.H., R.H., M.A.G., C.L.M. and J.C.D.H.  
4 937  
5 938 **Data curation:** B.K., R.C., A.V.P., C.V.P., P.A.  
6 939  
7 940 **Formal analysis:** B.K., R.C., C.V.P., A.V.P., X.Z., C.K., S.V.O., Y.L., P.A., A.D. and J.C.D.  
8 941  
9 942 **Funding acquisition:** B.K., R.C., A.V.P., X.Z. and J.C.D.H.  
10 943  
11 944 **Investigation:** B.K., R.C.A., A.V.P., X.Z. and J.C.D.H.  
12 945  
13 946 **Methodology:** B.K., R.H., M.A.G., C.L.M. and J.C.D.H.  
14 947  
15 948 **Project administration:** B.K. and J.C.D.H.  
16 949  
17 950 **Resources:** B.K., R.H., M.A.G, C.L.M. and J.C.D.H.  
18 951  
19 952 **Software:** B.K. and A.V.P.  
20 953  
21 954 **Supervision:** M.A.G., C.L.G., C.L.M. and J.C.D.H.  
22 955  
23 956 **Validation:** B.K., R.C., A.V.P., X.Z. and J.C.D.H.  
24 957  
25 958 **Visualization:** B.K., R.C., Y.L., C.V.P., A.V.P. and J.C.D.H.  
26 959  
27 960 **Writing original draft:** B.K., R.C. and J.C.D.H  
28 961  
29 962 **Writing reviews and editing:** B.K., R.C., A.V.P., X.Z., C.K., S.V.O., A.D., R.H., M.G and  
30 963 J.C.D.H.  
31 964  
32 965 **Equal contribution:** Authors B.K., R.C. and A.V.P. made equal contributions to this work.  
33 966  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## References

- 967  
968
- 969 Ackermann M, Stecher B, Freed N, Songhet P, Hardt W, and Doebeli M (2008)  
970 Self-destructive cooperation mediated by phenotypic noise. *454:987–990*.
- 971 Andrews S (2010) FastQC: a quality control tool for high throughput sequence  
972 data. Available online at:  
973 <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- 974 Aulicino A, Antanaviciute A, Frost J, Sousa Geros A, Mellado E, Attar M,  
975 Jagielowicz M, Hublitz P, Sinz J, Preciado-Llanes L, Napolitani G, Bowden R,  
976 Koohy H, Drakesmith H, and Simmons A (2022) Dual RNA sequencing reveals  
977 dendritic cell reprogramming in response to typhoidal *Salmonella* invasion.  
978 *Commun Biol* 5:111.
- 979 Bogomolnaya LM, Santiviago CA, Yang H-J, Baumler AJ, and Andrews-  
980 Polymenis HL (2008) 'Form variation' of the O12 antigen is critical for  
981 persistence of *Salmonella* Typhimurium in the murine intestine. *Mol Microbiol*  
982 70:1105–1119.
- 983 Bolger AM, Lohse M, and Usadel B (2014) Trimmomatic: a flexible trimmer for  
984 Illumina sequence data. *Bioinformatics* 30:2114–2120.
- 985 Branchu P, Bawn M, and Kingsley RA (2018) Genome Variation and Molecular  
986 Epidemiology of *Salmonella enterica* Serovar Typhimurium Pathovariants.  
987 *Infect Immun* 86:e00079-18.
- 988 Brink T, Leiss V, Siegert P, Jehle D, Ebner JK, Schwan C, Shymanets A, Wiese  
989 S, Nürnberg B, Hensel M, Aktories K, and Orth JHC (2018) *Salmonella*  
990 Typhimurium effector Ssel inhibits chemotaxis and increases host cell survival  
991 by deamidation of heterotrimeric Gi proteins. *PLOS Pathog* 14:e1007248.
- 992 Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, and  
993 Madden TL (2009) BLAST+: architecture and applications. *BMC Bioinformatics*  
994 10:421.
- 995 Canals R, Chaudhuri RR, Steiner RE, Owen SV, Quinones-Olvera N, Gordon  
996 MA, Baym M, Ibba M, and Hinton JCD (2019) [\(20192019a\)](#) The fitness landscape of  
997 the African *Salmonella* Typhimurium ST313 strain D23580 reveals unique  
998 properties of the pBT1 plasmid. *PLOS Pathog* 15:e1007948.
- 999 Canals R, Hammarlöf DL, Kröger C, Owen SV, Fong WY, Lacharme-Lora L,  
1000 Zhu X, Wenner N, Carden SE, Honeycutt J, Monack DM, Kingsley RA,  
1001 Brownridge P, Chaudhuri RR, Rowe WPM, Predeus AV, Hokamp K, Gordon  
1002 MA, and Hinton JCD (2019) [\(20192019b\)](#) Adding function to the genome of African  
1003 *Salmonella* Typhimurium ST313 strain D23580. *PLOS Biol* 17:e3000059.
- 1004 Carden SE, Walker GT, Honeycutt J, Lugo K, Pham T, Jacobson A, Bouley D,  
1005 Idoyaga J, Tsolis RM, and Monack D (2017) Pseudogenization of the Secreted



- 1  
2  
3 1006 Effector Gene *ssel* Confers Rapid Systemic Dissemination of *S. Typhimurium*  
4 1007 ST313 within Migratory Dendritic Cells. *Cell Host Microbe* 21:182–194.  
5  
6 1008 Chirwa E, Dale, H, Gordon, MA, and Ashton, PM (2023) What is the Source of  
7 1009 Infections Causing Invasive Nontyphoidal Salmonella Disease? *Open Forum*  
8 1010 *Infect Diseases* February 2023.  
9  
10  
11 1011 Crump JA, Sjölund-Karlsson M, Gordon MA, and Parry CM (2015)  
12 1012 Epidemiology, Clinical Presentation, Laboratory Diagnosis, Antimicrobial  
13 1013 Resistance, and Antimicrobial Management of Invasive Salmonella Infections.  
14 1014 *Clin Microbiol Rev* 28:901–937.  
15  
16 1015 Ewels P, Magnusson M, Lundin S, and Käller M (2016) MultiQC: summarize  
17 1016 analysis results for multiple tools and samples in a single report. *Bioinformatics*  
18 1017 32:3047–3048.  
19  
20  
21 1018 Fass E, and Groisman E (2009) Control of Salmonella pathogenicity island-2  
22 1019 gene expression. *Curr Opin Microbiol* 12:199–204.  
23  
24 1020 Feasey NA, Dougan G, Kingsley RA, Heyderman RS, and Gordon MA (2012)  
25 1021 Invasive non-typhoidal salmonella disease: an emerging and neglected tropical  
26 1022 disease in Africa. *The Lancet* 379:2489–2499.  
27  
28  
29 1023 Feasey NA, Hadfield J, Keddy KH, Dallman TJ, Jacobs J, Deng X, Wigley P,  
30 1024 Barquist L, Langridge GC, Feltwell T, Harris SR, Mather AE, Fookes M, Aslett  
31 1025 M, Msefula C, Kariuki S, Maclennan CA, Onsare RS, Weill F-X, Le Hello S,  
32 1026 Smith AM, McClelland M, Desai P, Parry CM, Cheesbrough J, French N,  
33 1027 Campos J, Chabalgoity JA, Betancor L, Hopkins KL, Nair S, Humphrey TJ,  
34 1028 Lunguya O, Cogan TA, Tapia MD, Sow SO, Tennant SM, Bornstein K, Levine  
35 1029 MM, Lacharme-Lora L, Everett DB, Kingsley RA, Parkhill J, Heyderman RS,  
36 1030 Dougan G, Gordon MA, and Thomson NR (2016) Distinct Salmonella Enteritidis  
37 1031 lineages associated with enterocolitis in high-income settings and invasive  
38 1032 disease in low-income settings. *Nat Genet* 48:1211–1217.  
39  
40  
41 1033 Gilchrist J, and MacLennan C (2019) Invasive Nontyphoidal Salmonella  
42 1034 Disease in Africa. *Cell Mol Biol E Coli Salmonella Enterobact*, doi: doi:10.1128/  
43 1035 ecosalplus.ESP-0007-2018.  
44  
45  
46 1036 Grabe GJ, Zhang Y, Przydacz M, Rolhion N, Yang Y, Pruneda JN, Komander  
47 1037 D, Holden DW, and Hare SA (2016) The Salmonella Effector SpvD Is a  
48 1038 Cysteine Hydrolase with a Serovar-specific Polymorphism Influencing Catalytic  
49 1039 Activity, Suppression of Immune Responses, and Bacterial Virulence. *J Biol*  
50 1040 *Chem* 291:25853–25863.  
51  
52  
53 1041 Gurevich A, Saveliev V, Vyahhi N, and Tesler G (2013) QUASt: quality  
54 1042 assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075.  
55  
56 1043 Hammarlöf DL, Kröger C, Owen SV, Canals R, Lacharme-Lora L, Wenner N,  
57 1044 Schager AE, Wells TJ, Henderson IR, Wigley P, Hokamp K, Feasey NA,  
58 1045 Gordon MA, and Hinton JCD (2018) Role of a single noncoding nucleotide in  
59  
60

- 1  
2  
3 1046 the evolution of an epidemic African clade of *Salmonella*. *Proc Natl Acad Sci*  
4 1047 115.  
5  
6 1048 Hautefort I, Proença MJ, and Hinton JCD (2003) Single-Copy Green  
7 1049 Fluorescent Protein Gene Fusions Allow Accurate Measurement of *Salmonella*  
8 1050 Gene Expression In Vitro and during Infection of Mammalian Cells. *Appl Environ*  
9 1051 *Microbiol* 69:7480–7491.  
10  
11  
12 1052 Honeycutt JD, Wenner N, Li Y, Brewer SM, Massis LM, Brubaker SW,  
13 1053 Chairatana P, Owen SV, Canals R, Hinton JCD, and Monack DM (2020)  
14 1054 Genetic variation in the MacAB-TolC efflux pump influences pathogenesis of  
15 1055 invasive *Salmonella* isolates from Africa. *PLOS Pathog* 16:e1008763.  
16  
17  
18 1056 Jennings E, Thurston TLM, and Holden DW (2017) *Salmonella* SPI-2 Type III  
19 1057 Secretion System Effectors: Molecular Mechanisms And Physiological  
20 1058 Consequences. *Cell Host Microbe* 22:217–231.  
21  
22 1059 Johnson R, and Simon M (1985) Hin-mediated site-specific recombination  
23 1060 requires two 26 bp recombination sites and a 60 bp recombinational enhancer.  
24 1061 *Cell* 41:781–791.  
25  
26  
27 1062 Kariuki S, Revathi G, Kariuki N, Kiiru J, Mwituria J, Muyodi J, Githinji JW,  
28 1063 Kagendo D, Munyalo A, and Hart CA (2006) Invasive multidrug-resistant non-  
29 1064 typhoidal *Salmonella* infections in Africa: zoonotic or anthroponotic  
30 1065 transmission? *J Med Microbiol* 55:585–591.  
31  
32  
33 1066 Kasumba IN, Pulford CV, Perez-Sepulveda BM, Sen S, Sayed N, Permala-  
34 1067 Booth J, Livio S, Heavens D, Low R, Hall N, Roose A, Powell H, Farag T,  
35 1068 Panchalingham S, Berkeley L, Nasrin D, Blackwelder WC, Wu Y, Tamboura B,  
36 1069 Sanogo D, Onwuchekwa U, Sow SO, Ochieng JB, Omore R, Oundo JO,  
37 1070 Breiman RF, Mintz ED, O'Reilly CE, Antonio M, Saha D, Hossain MJ,  
38 1071 Mandomando I, Bassat Q, Alonso PL, Ramamurthy T, Sur D, Qureshi S, Zaidi  
39 1072 AKM, Hossain A, Faruque ASG, Nataro JP, Kotloff KL, Levine MM, Hinton JCD,  
40 1073 and Tennant SM (2021) Characteristics of *Salmonella* Recovered From Stools  
41 1074 of Children Enrolled in the Global Enteric Multicenter Study. *Clin Infect Dis*  
42 1075 73:631–641.  
43  
44  
45 1076 Kingsley RA, Msefula CL, Thomson NR, Kariuki S, Holt KE, Gordon MA, Harris  
46 1077 D, Clarke L, Whitehead S, Sangal V, Marsh K, Achtman M, Molyneux ME,  
47 1078 Cormican M, Parkhill J, MacLennan CA, Heyderman RS, and Dougan G (2009)  
48 1079 Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive  
49 1080 disease in sub-Saharan Africa have a distinct genotype. *Genome Res* 19:2279–  
50 1081 2287.  
51  
52  
53 1082 Kirk MD, Pires SM, Black RE, Caipo M, Crump JA, Devleeschauwer B, Döpfer  
54 1083 D, Fazil A, Fischer-Walker CL, Hald T, Hall AJ, Keddy KH, Lake RJ, Lanata CF,  
55 1084 Torgerson PR, Havelaar AH, and Angulo FJ (2015) World Health Organization  
56 1085 Estimates of the Global and Regional Disease Burden of 22 Foodborne  
57 1086 Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. *PLOS Med*  
58 1087 12:e1001921.  
59  
60

- 1088 Koolman L, Prakash R, Diness Y, Msefula C, Nyirenda TS, Olgemoeller F,  
1089 Perez-Sepulveda B, Hinton JCD, Owen SV, Feasey NA, Ashton PM, and  
1090 Gordon MA (2022) *Case-control investigation of invasive Salmonella disease*  
1091 *in Africa – comparison of human, animal and household environmental isolates*  
1092 *find no evidence of environmental or animal reservoirs of invasive strains,*  
1093 *Infectious Diseases (except HIV/AIDS).*
- 1094 Kröger C, Colgan A, Srikumar S, Händler K, Sivasankaran SK, Hammarlöf DL,  
1095 Canals R, Grissom JE, Conway T, Hokamp K, and Hinton JCD (2013) *An*  
1096 *Infection-Relevant Transcriptomic Compendium for Salmonella enterica*  
1097 *Serovar Typhimurium. Cell Host Microbe* 14:683–695.
- 1098 Kröger C, Dillon SC, Cameron ADS, Papenfort K, Sivasankaran SK, Hokamp  
1099 K, Chao Y, Sittka A, Hébrard M, Händler K, Colgan A, Leekitcharoenphon P,  
1100 Langridge GC, Lohan AJ, Loftus B, Lucchini S, Ussery DW, Dorman CJ,  
1101 Thomson NR, Vogel J, and Hinton JCD (2012) *The transcriptional landscape*  
1102 *and small RNAs of Salmonella enterica serovar Typhimurium. Proc Natl Acad*  
1103 *Sci* 109.
- 1104 Lai MA, Quarles EK, López-Yglesias AH, Zhao X, Hajjar AM, and Smith KD  
1105 (2013) *Innate Immune Detection of Flagellin Positively and Negatively*  
1106 *Regulates Salmonella Infection. PLoS ONE* 8:e72047.
- 1107 Letunic I, and Bork P (2006) *Interactive Tree Of Life (iTOL): an online tool for*  
1108 *phylogenetic tree display and annotation. Bioinformatics* 23:1283–1287.
- 1109 Lian ZJ, Phan M, Hancock SJ, Nhu NTK, Paterson DL, and Schembri MA  
1110 (2023) *Genetic basis of I-complex plasmid stability and conjugation. PLOS*  
1111 *Genet* 19:e1010773.
- 1112 Löber S, Jäckel D, Kaiser N, and Hensel M (2006) *Regulation of Salmonella*  
1113 *pathogenicity island 2 genes by independent environmental signals. Int J Med*  
1114 *Microbiol* 296:435–447.
- 1115 Lucchini S, Rowley G, Goldberg MD, Hurd D, Harrison M, and Hinton JCD  
1116 (2006) *H-NS Mediates the Silencing of Laterally Acquired Genes in Bacteria.*  
1117 *PLoS Pathog* 2:e81.
- 1118 Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, Jones TF,  
1119 Fazil A, and Hoekstra RM (2010) *The Global Burden of Nontyphoidal*  
1120 *Salmonella Gastroenteritis. Clin Infect Dis* 50:882–889.
- 1121 Marchello C, Dale, A. P., Pisharody, S., Rubach, M. P., and Crump, J.A. (2019)  
1122 *A Systematic Review and Meta-analysis of the Prevalence of Community-*  
1123 *Onset Bloodstream Infections among Hospitalized Patients in Africa and Asia.*  
1124 *Antimicrobial Agents and Chemotherapy.* 64.
- 1125 Marchello CS, Dale AP, Pisharody S, Rubach MP, and Crump JA (2019) *A*  
1126 *Systematic Review and Meta-analysis of the Prevalence of Community-Onset*  
1127 *Bloodstream Infections among Hospitalized Patients in Africa and Asia.*  
1128 *Antimicrob Agents Chemother* 64:e01974-19.

- 1  
2  
3 1129 Msefula CL, Kingsley RA, Gordon MA, Molyneux E, Molyneux ME, MacLennan  
4 1130 CA, Dougan G, and Heyderman RS (2012) Genotypic Homogeneity of  
5 1131 Multidrug Resistant *S. Typhimurium* Infecting Distinct Adult and Childhood  
6 1132 Susceptibility Groups in Blantyre, Malawi. *PLoS ONE* 7:e42085.
- 8  
9 1133 Musicha P, Cornick JE, Bar-Zeev N, French N, Masesa C, Denis B, Kennedy  
10 1134 N, Mallewa J, Gordon MA, Msefula CL, Heyderman RS, Everett DB, and  
11 1135 Feasey NA (2017) Trends in antimicrobial resistance in bloodstream infection  
12 1136 isolates at a large urban hospital in Malawi (1998–2016): a surveillance study.  
13 1137 *Lancet Infect Dis* 17:1042–1052.
- 15  
16 1138 Nedialkova LP, Denzler R, Koepfel MB, Diehl M, Ring D, Wille T, Gerlach RG,  
17 1139 and Stecher B (2014) Inflammation Fuels Colicin Ib-Dependent Competition of  
18 1140 *Salmonella* Serovar Typhimurium and *E. coli* in Enterobacterial Blooms. *PLoS*  
19 1141 *Pathog* 10:e1003844.
- 21  
22 1142 Okoro CK, Barquist L, Connor TR, Harris SR, Clare S, Stevens MP, Arends MJ,  
23 1143 Hale C, Kane L, Pickard DJ, Hill J, Harcourt K, Parkhill J, Dougan G, and  
24 1144 Kingsley RA (2015) Signatures of Adaptation in Human Invasive *Salmonella*  
25 1145 Typhimurium ST313 Populations from Sub-Saharan Africa. *PLoS Negl Trop Dis*  
26 1146 9:e0003611.
- 28  
29 1147 Okoro CK, Kingsley RA, Connor TR, Harris SR, Parry CM, Al-Mashhadani MN,  
30 1148 Kariuki S, Msefula CL, Gordon MA, de Pinna E, Wain J, Heyderman RS, Obaro  
31 1149 S, Alonso PL, Mandomando I, MacLennan CA, Tapia MD, Levine MM, Tennant  
32 1150 SM, Parkhill J, and Dougan G (2012) Intracontinental spread of human invasive  
33 1151 *Salmonella* Typhimurium pathovariants in sub-Saharan Africa. *Nat Genet*  
34 1152 44:1215–1221.
- 36  
37 1153 Owen SV, Wenner N, Canals R, Makumi A, Hammarlöf DL, Gordon MA,  
38 1154 Aertsen A, Feasey NA, and Hinton JCD (2017) Characterization of the  
39 1155 Prophage Repertoire of African *Salmonella* Typhimurium ST313 Reveals High  
40 1156 Levels of Spontaneous Induction of Novel Phage BTP1. *Front Microbiol* 8.
- 41  
42 1157 Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M,  
43 1158 Falush D, Keane JA, and Parkhill J (2015) Roary: rapid large-scale prokaryote  
44 1159 pan genome analysis. *Bioinformatics* 31:3691–3693.
- 46  
47 1160 Perez-Sepulveda BM, and Hinton JCD (2018) Functional Transcriptomics for  
48 1161 Bacterial Gene Detectives. *Microbiol Spectr* 6:6.5.06.
- 49  
50 1162 Piccini G, and Montomoli E (2020) Pathogenic signature of invasive non-  
51 1163 typhoidal *Salmonella* in Africa: implications for vaccine development. *Hum*  
52 1164 *Vaccines Immunother* 16:2056–2071.
- 53  
54 1165 Post AS, Diallo SN, Guiraud I, Lompo P, Tahita MC, Maltha J, Van Puyvelde S,  
55 1166 Mattheus W, Ley B, Thriemer K, Rouamba E, Derra K, Deborggraeve S, Tinto  
56 1167 H, and Jacobs J (2019) Supporting evidence for a human reservoir of invasive  
57 1168 non-Typhoidal *Salmonella* from household samples in Burkina Faso. *PLoS*  
58 1169 *Negl Trop Dis* 13:e0007782.
- 59  
60



- 1  
2  
3 1170 Preciado-Llanes L, Aulicino A, Canals R, Moynihan PJ, Zhu X, Jambo N,  
4 1171 Nyirenda TS, Kadwala I, Sousa Gerós A, Owen SV, Jambo KC, Kumwenda B,  
5 1172 Veerapen N, Besra GS, Gordon MA, Hinton JCD, Napolitani G, Salio M, and  
6 1173 Simmons A (2020) Evasion of MAIT cell recognition by the African *Salmonella*  
7 1174 Typhimurium ST313 pathovar that causes invasive disease. *Proc Natl Acad Sci*  
8 1175 117:20717–20728.
- 9  
10  
11 1176 Pulford CV, Perez-Sepulveda BM, Canals R, Bevington JA, Bengtsson RJ,  
12 1177 Wenner N, Rodwell EV, Kumwenda B, Zhu X, Bennett RJ, Stenhouse GE,  
13 1178 Malaka De Silva P, Webster HJ, Bengoechea JA, Dumigan A, Tran-Dien A,  
14 1179 Prakash R, Banda HC, Alufandika L, Mautanga MP, Bowers-Barnard A,  
15 1180 Beliavskaia AY, Predeus AV, Rowe WPM, Darby AC, Hall N, Weill F-X, Gordon  
16 1181 MA, Feasey NA, Baker KS, and Hinton JCD (2021) Stepwise evolution of  
17 1182 *Salmonella* Typhimurium ST313 causing bloodstream infection in Africa. *Nat*  
18 1183 *Microbiol* 6:327–338.
- 19  
20  
21 1184 Rankin, J.D. and Taylor R.J. (1966) The estimation of doses of *Salmonella*  
22 1185 typhimurium suitable for the experimental production of disease in calves. *Vec*  
23 1186 *Rec* 78:706–7.
- 24  
25  
26 1187 Richardson EJ, Limaye B, Inamdar H, Datta A, Manjari KS, Pullinger GD,  
27 1188 Thomson NR, Joshi RR, Watson M, and Stevens MP (2011) Genome  
28 1189 Sequences of *Salmonella enterica* Serovar Typhimurium, Choleraesuis, Dublin,  
29 1190 and Gallinarum Strains of Well- Defined Virulence in Food-Producing Animals.  
30 1191 *J Bacteriol* 193:3162–3163.
- 31  
32  
33 1192 Schlechter R, and Remus-Emsermann M (2019) Delivering &quot;Chromatic  
34 1193 Bacteria&quot; Fluorescent Protein Tags to Proteobacteria Using Conjugation.  
35 1194 *BIO-Protoc* 9.
- 36  
37 1195 Sikand A, Jaszczur M, Bloom LB, Woodgate R, Cox MM, and Goodman MF  
38 1196 (2021) The SOS Error-Prone DNA Polymerase V Mutasome and  $\beta$ -Sliding  
39 1197 Clamp Acting in Concert on Undamaged DNA and during Translesion  
40 1198 Synthesis. *Cells* 10:1083.
- 41  
42  
43 1199 [Skidmore PD, Canals R, and Ramasamy MN \(2023\). The iNTS-GMMA vaccine:  
44 1200 a promising step in non-typhoidal \*Salmonella\* vaccine development. \*Expert Rev\*  
45 1201 \*Vaccines\* \(In Press\) doi: 10.1080/14760584.2023.2270596.](#)  
46 1202
- 47 1203 Srikumar S, Kröger C, Hébrard M, Colgan A, Owen SV, Sivasankaran SK,  
48 1204 Cameron ADS, Hokamp K, and Hinton JCD (2015) RNA-seq Brings New  
49 1205 Insights to the Intra-Macrophage Transcriptome of *Salmonella* Typhimurium.  
50 1206 *PLOS Pathog* 11:e1005262.
- 51  
52  
53 1207 Stamatakis A, Ludwig T, and Meier H (2005) RAxML-III: a fast program for  
54 1208 maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics*  
55 1209 21:456–463.
- 56  
57 1210 Stanaway JD, Parisi A, Sarkar K, Blacker BF, Reiner RC, Hay SI, Nixon MR,  
58 1211 Dolecek C, James SL, Mokdad AH, Abebe G, Ahmadian E, Alahdab F,  
59 1212 Alemnew BTT, Alipour V, Allah Bakeshei F, Animut MD, Ansari F, Arabloo J,



- 1213 Asfaw ET, Bagherzadeh M, Bassat Q, Belayneh YMM, Carvalho F, Daryani A,  
1214 Demeke FM, Demis ABB, Dubey M, Duken EE, Dunachie SJ, Eftekhari A,  
1215 Fernandes E, Fouladi Fard R, Gedefaw GA, Geta B, Gibney KB, Hasanzadeh  
1216 A, Hoang CL, Kasaeian A, Khater A, Kidanemariam ZT, Lakew AM,  
1217 Malekzadeh R, Melese A, Mengistu DT, Mestrovic T, Miazgowski B,  
1218 Mohammad KA, Mohammadian M, Mohammadian-Hafshejani A, Nguyen CT,  
1219 Nguyen LH, Nguyen SH, Nirayo YL, Olagunju AT, Olagunju TO, Pourjafar H,  
1220 Qorbani M, Rabiee M, Rabiee N, Rafay A, Rezapour A, Samy AM, Sepanlou  
1221 SG, Shaikh MA, Sharif M, Shigematsu M, Tessema B, Tran BX, Ullah I, Yimer  
1222 EM, Zaidi Z, Murray CJL, and Crump JA (2019) The global burden of non-  
1223 typhoidal salmonella invasive disease: a systematic analysis for the Global  
1224 Burden of Disease Study 2017. *Lancet Infect Dis* 19:1312–1324.
- 1225 Tack B, Phoba M-F, Barbé B, Kalonji LM, Hardy L, Van Puyvelde S, Ingelbeen  
1226 B, Falay D, Ngonda D, van der Sande MAB, Deborggraeve S, Jacobs J, and  
1227 Lunguya O (2020) Non-typhoidal Salmonella bloodstream infections in Kisantu,  
1228 DR Congo: Emergence of O5-negative Salmonella Typhimurium and extensive  
1229 drug resistance. *PLoS Negl Trop Dis* 14:e0008121.
- 1230 Van Puyvelde S, Pickard D, Vandelannoote K, Heinz E, Barbé B, de Block T,  
1231 Clare S, Coomber EL, Harcourt K, Sridhar S, Lees EA, Wheeler NE, Klemm  
1232 EJ, Kuijpers L, Mbuyi Kalonji L, Phoba M-F, Falay D, Ngbonda D, Lunguya O,  
1233 Jacobs J, Dougan G, and Deborggraeve S (2019) An African Salmonella  
1234 Typhimurium ST313 sublineage with extensive drug-resistance and signatures  
1235 of host adaptation. *Nat Commun* 10:4280.
- 1236 [Van Puyvelde S, de Block T, Sridhar S, Bawn M, Kingsley RA, Ingelbeen B,](#)  
1237 [Beale MA, Barbé B, Jeon HJ, Mbuyi-Kalonji L, Phoba MF, Falay D, Martiny D,](#)  
1238 [Vandenberg O, Affolabi D, Rutanga JP, Ceysens PJ, Mattheus W, Cuypers](#)  
1239 [WL, van der Sande MAB, Park SE, Kariuki S, Otieno K, Lusingu JPA, Mbwana](#)  
1240 [JR, Adjei S, Sarfo A, Agyei SO, Asante KP, Otieno W, Otieno L, Tahita MC,](#)  
1241 [Lompo P, Hoffman IF, Mvalo T, Msefula C, Hassan-Hanga F, Obaro S,](#)  
1242 [Mackenzie G, Deborggraeve S, Feasey N, Marks F, MacLennan CA, Thomson](#)  
1243 [NR, Jacobs J, Dougan G, Kariuki S, Lunguya O. \(2023\) A genomic appraisal of](#)  
1244 [invasive Salmonella Typhimurium and associated antibiotic resistance in sub-](#)  
1245 [Saharan Africa. \*Nat Commun.\* 14:6392.](#)
- 1246 Wick RR, Judd LM, Gorrie CL, and Holt KE (2017) Unicycler: Resolving  
1247 bacterial genome assemblies from short and long sequencing reads. *PLOS*  
1248 *Comput Biol* 13:e1005595.
- 1249 Zhang S, Kingsley RA, Santos RL, Andrews-Polymenis H, Raffatellu M,  
1250 Figueiredo J, Nunes J, Tsolis RM, Adams LG, and Bäumlér AJ (2003)  
1251 Molecular Pathogenesis of *Salmonella enterica* Serotype Typhimurium-  
1252 Induced Diarrhea. *Infect Immun* 71:1–12.
- 1253