Loss of type VI secretion systems in multi-drug resistant *Escherichia coli* clones

Elizabeth A. Cummins[1], Robert A. Moran[1], Ann E. Snaith[1], Rebecca J. Hall[1], Chris H. Connor[2], Steven J. Dunn[1], Alan McNally[1]

[1]Institute of Microbiology and Infection, College of Medical and Dental Sciences, University of Birmingham, Birmingham, B15 2TT, UK

Corresponding author: a.mcnally.1@bham.ac.uk

[2]Doherty Institute for Infection and Immunity, University of Melbourne, Melbourne, 3000, Australia

Keywords: type VI secretion system, *Escherichia coli*, multi-drug resistance, ST131

Abstract

The repeated emergence of pandemic multi-drug resistant (MDR) *Escherichia coli* clones is a threat to public health globally. In recent work, drug resistant *E. coli* were shown to be capable of displacing commensal *E. coli* in the human gut. We hypothesise that there are three factors potentially responsible for the apparent competitive advantage of drug resistant *E. coli* clones: enhanced metabolic capabilities; production of lytic phage or bacteriocins; or the presence of a type VI secretion system (T6SS). We employed genomic and experimental approaches to investigate. First, we searched for T6SS apparatus genes across a curated dataset of over 20,000 genomes representing the full phylogenetic diversity of *E. coli*. This revealed large, non-phylogenetic variation in the presence of T6SS genes. No association was found between T6SS gene presence and MDR clones. However, multiple MDR clones have lost essential T6SS genes. We have shown that the pandemic MDR clone ST131-H30Rx lost the essential T6SS gene *tssM* when it was interrupted by insertion of IS*Ec12*. To investigate the importance of this, a three-way competition assay was designed to assess the contact-killing capabilities of three strains - ST131-H30Rx, its drug-susceptible TssM-producing ancestor, and laboratory strain K-12 MG1655. Our results confirmed the competitive advantage of ST131-H30Rx in the absence of antibiotics, but its success was not due to the presence of phage, secreted proteins or contact-dependent killing. Our findings suggest that metabolic advantages are therefore likely a key component in the success of ST131-H30Rx.

33 # Data Summary

34 The genome sequence data generated in this study is publicly available from NCBI under
35 BioProject PRJNA943186, alongside a complete assembly in GenBank under accessions
36 CP120633-CP120634. All other sequence data used in this paper has been taken from
37 ENA with the appropriate accession numbers listed within the methods section. The *E.*
38 *coli* genome data sets used in this work are from a previous publication, the details of
39 which can be found in the corresponding supplementary data files
40 10.6084/m9.figshare.21360108 [1].

41 # Impact Statement

42 *Escherichia coli* is a globally significant pathogen that causes the majority of urinary tract
43 infections. Treatment of these infections is exacerbated by increasing levels of drug
44 resistance, with multi-drug resistant (MDR) pandemic clones such as ST131-H30Rx
45 contributing significantly to global disease burden. MDR *E. coli* clones are able to
46 colonise the human gut and displace resident commensal *E. coli.* It is important to
47 understand how this process occurs to better understand why these pathogens are so
48 successful. Type VI secretion may be one of the antagonistic systems employed by *E. coli*
49 in this process. Our findings provide the first detailed characterisation of the T6SS locus
50 in ST131 and shed light on an event in the evolutionary pathway of the prominent MDR
51 pandemic pathogen ST131-H30Rx.

52

53 # Introduction

54 *Escherichia coli* has been ranked globally as the number one causal pathogen of deaths
55 associated with bacterial antimicrobial resistance (AMR) [2] and AMR has been declared
56 by the World Health Organisation as a top ten global public health threat
57 (www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance). The
58 emergence and evolution of multi-drug resistant (MDR) *E. coli* is a pressing and relevant
59 issue for global healthcare as their extensive resistance profiles result in diminishing
60 therapeutic options for treating infections.

61 A trait increasingly ascribed to MDR *E. coli* lineages is the ability to rapidly, and
62 asymptomatically, colonise the human intestinal tract [3]. Multiple traveller studies have
63 now shown that people travelling from areas of low AMR incidence to AMR endemic
64 regions become colonised by extended spectrum beta-lactam (ESBL)-resistant or MDR
65 *E. coli* during travel [4; 5; 6]. Furthermore, genomic analysis has shown that the gain of
66 MDR *E. coli* is due to the acquisition of a new MDR strain and not the preceding
67 commensal *E. coli* becoming MDR [5]. The ability to displace and colonise may be
68 attributed to the MDR phenotype itself, but longitudinal studies from the UK and
69 Norway [7; 8] have shown that multi-drug resistance alone is not a sufficient driver for
70 epidemiological success of *E. coli*. Recent metagenomic analysis has also revealed that
71 colonisation by MDR *E. coli* does not disrupt the wider gut microbiome's composition or
72 diversity [9].

73 We hypothesise that the ability of drug-resistant *E. coli* to displace resident commensal
74 *E. coli* is a result of three possible biological scenarios. Firstly, the invading drug-resistant

75 *E. coli* has enhanced metabolic capabilities that facilitate out-competition of the resident
76 commensal *E. coli* leading to prolonged colonisation and dissemination [10; 11].
77 Secondly, the drug-resistant *E. coli* produce lytic phage or bacteriocins that can kill
78 competitors without requiring cell-to-cell contact [12; 13]. Thirdly, the drug-resistant *E.*
79 *coli* possess a type VI secretion system that allows contact-dependent killing of the
80 resident commensal *E. coli*.

81 The type VI secretion system (T6SS) is a multi-functional apparatus that some Gram-
82 negative bacteria possess to facilitate nutrient uptake [14], manipulation of host cells
83 [15], and the killing of competing bacteria [16; 17; 18; 19]. T6SS distribution varies by
84 environment [20; 21] and the presence of T6SSs varies on all taxonomic levels [20; 22;
85 23] with over-representation in Gammaproteobacteria [24]. T6SS genes can be gained
86 and lost via horizontal gene transfer (HGT) [16; 22] and T6SS presence may be influenced
87 by genomic incompatibilities, either between mobile genetic elements, donor and
88 recipient bacteria, or via assembly of multiple T6SSs [20]. T6SSs are particularly
89 prevalent in complex microbial communities that are host-associated [21; 22] and are
90 commonly found in pathogens, including *E. coli* [20; 25; 26; 27].

91 *E. coli* is a relatively understudied organism within the research field of T6SSs, with
92 studies focusing on specific pathotypes [26; 27; 28]. In *E. coli* T6SSs have been shown to
93 be directly involved in pathogenesis and antibacterial activity [25]. The existing
94 connections between T6SS and MDR in other species [29; 30] makes T6SSs in *E. coli* an
95 intriguing avenue of exploration in the context of pandemic MDR clones and how they
96 become successful. Here, we aim to assess the prevalence of T6SSs across *E. coli* and

97 identify any association with MDR lineages. Further, we will investigate, *in silico* and *in*
98 *vitro*, T6SS functionality in the successful MDR clone ST131-H30Rx.

99 Methods

100 Genome collection and multi-drug resistant classification

101 The 20577 *E. coli* genomes used here were curated from EnteroBase [31] and were fully
102 characterised and described in a previous study (available at
103 10.6084/m9.figshare.21360108) [1]. Briefly, they were chosen to represent the
104 phylogenetic, genotypic and phenotypic diversity of the species. Encompassing 21 sequence
105 types (STs). Genomic assemblies were screened for antibiotic resistance genes using ABRicate
106 (v0.9.8) (github.com/tseemann/abricate) with the NCBI database [32]. For the purposes of this
107 study strains were classified as MDR if their genome possessed at least one extended-
108 spectrum beta-lactamase or carbapenemase gene.

109 T6SS database creation and interrogation

110 The SecReT6 [33; 34] experimentally validated database was retrieved and reformatted
111 so that sequence identifiers were a suitable format for a custom ABRicate database.
112 ABRicate (v0.9.8) was run using 25% minimum coverage and 75% minimum sequence
113 identity to obtain a list of T6SS genes that are present across our 21 *E. coli* sequence
114 types. The results were summarised and partial hits (instances where a gene hit was split
115 over multiple contigs) were accounted for and processed with a custom Python script
116 (github.com/lillycummins/InterPangenome/blob/main/process_partial_hits.py). A gene
117 was determined as present in a genome if at least 85% of the gene was covered with a

118 sequence identity above 75%. Whether T6SS genes were likely located in plasmids or in
119 chromosomes was determined using mlplasmid [35].

120 Genomic screening for insertion sequences and bacteriocins

121 Gene Construction Kit (v4.5.1) (Textco Biosoftware, Raleigh, USA) was used to annotate
122 and visualise DNA sequences. Insertion sequences (ISs) were identified using the
123 ISFinder database [36]. Target site duplications (TSDs) flanking ISs were identified
124 manually. Two junction sequences specific for a given insertion were generated by
125 taking 100 bp contiguous sequences that span the left and right ends of the insertion.
126 Each junction sequence was comprised of 50 bp of the target IS and 50 bp of its adjacent
127 sequence. To generate an ancestral sequence *in silico*, the interrupting IS and one copy
128 of its associated TSD were removed manually. A 100 bp sequence that spanned the
129 insertion point was taken to represent the naïve site. A database of all ST131 genomes
130 in the collection was generated for screening with standalone BLASTn [37]. The database
131 was queried with insertion-junction sequences and the naïve site sequence to determine
132 whether an insertion was present, with only complete and identical matches to the 100
133 bp query sequences considered positive matches.

134 Genomes were screened for bacteriocins and prophage regions using the BACTIBASE
135 [38] online database and PHASTER [39; 40], respectively.

136 Strains and culture conditions

137 Three *E. coli* isolates were selected based on their phenotypic resistance profiles to allow
138 for selective differential growth after co-culture. The strains used in this study are

139 described in Table 1. *E. coli* strains were grown in Lysogeny broth (LB) at 37°C shaking or
140 on LB agar plates. Strains were grown in the presence of 30 μg/mL kanamycin, 4 μg/mL
141 cefotaxime or 100 μg/mL rifampicin as relevant.

142 The reads for MG1655-GFP (ENA accession: ERR1957959) and UTI62 (ENA accession:
143 ERR062295) were downloaded and assembled with SPAdes (v3.12.0). No existing
144 sequence data was available for USVAST219. Therefore, DNA was extracted using the
145 Monarch Genomic DNA Extraction Kit (New England Biolabs, Massachusetts, USA)
146 before sequencing with MinION (Oxford Nanopore Technology, UK) using R9.4.1 flow
147 cells. The data was basecalled with Guppy (v6.0.1)
148 (github.com/nanoporetech/pyguppyclient) and adapters were trimmed with qcat
149 (v1.1.0) (github.com/nanoporetech/qcat). Short-read genome sequencing was provided
150 by MicrobesNG (microbesng.com) where DNA was prepared using the Nextera XT kit
151 (Illumina, San Diego, CA, USA) and sequenced on the Illumina HiSeq platform (Illumina).
152 Long and short-read data are both available within BioProject PRJNA943186. A hybrid
153 assembly using both the long and short-read sequencing data was created with Unicycler
154 (v0.4.8) [41]. To generate a rifampicin-resistant mutant, 100 μL of USVAST219 overnight
155 culture was spread on a LB plate containing 100 μg/mL rifampicin and incubated at 37
156 °C overnight. The following day, a single colony was chosen to inoculate 5 μL LB +
157 rifampicin (100 μg/mL) broth. The broth was incubated at 37 °C overnight and a glycerol
158 stock was stored at -80 °C the following day.

159

160

161

162 **Table 1** Strains used in this study.

Strain	Lineage	Isolation	Selective resistance	Ref.
K12-MG1655	ST10	-	Kanamycin	[42]
USVAST219[a]	ST131-H40	Urine sample from Veterans Affairs Medical Centre, Michigan. (2011)	Rifampicin	[43]
UTI62	ST131-H30Rx	Urine agar plate from Nottingham University Hospitals (2008-09)	Cefotaxime	[44]

163

164 [a]Rifampicin-resistant mutant.

165 ## Competition assay

166 Overnight cultures of *E. coli* strains were diluted to an OD_{600} of 1.0, mixed at a 1:1 ratio
167 on sterile filters and placed on LB agar plates. After four hours incubation at 37 °C,
168 bacteria were harvested from the filter by vigorous vortexing followed by centrifugation
169 at 12100 × g for two minutes. Harvested cells were then serially diluted and plated onto
170 LB agar plates supplemented with the appropriate antibiotic to enumerate the surviving
171 *E. coli*. All plates, including samples from time point 0, were incubated at 37 °C overnight
172 and CFUs were counted the following day. Relative fold change in growth was calculated
173 as the four-hour growth (cfu/ml) of a strain following incubation in isolation divided by
174 the four-hour growth of the strain following co-incubation. In a separate experiment,
175 the supernatant was extracted from UTI62 overnight culture with a 0.2 μm sterile syringe
176 filter. K-12 MG1655 was grown for four hours at 37 °C in the presence and absence of
177 UTI62 supernatant to test for the presence of potentially lethal phages, bacteriocins, or
178 excreted metabolites.

179

180 # Results

181 ## T6SS gene presence varies between and within phylogroups and
182 sequence types

183 All 20577 *E. coli* genomes interrogated contained at least one T6SS gene from the SecReT
184 6 experimental database [33; 34]. There was variation in the average number of T6SS
185 genes present per ST and the average number of genes per ST within phylogroups (Figure
186 1). ST131 exhibited the largest range (2 to 29) of T6SS genes, highest number of T6SS
187 genes in a single genome (n = 29), and highest average number of T6SS genes (joint with
188 ST12, ST14, ST17, and ST127). The distribution of the number of T6SS genes present per
189 genome showed no phylogenetic signal. STs known to contain MDR lineages, ST131,
190 ST167, ST410, ST648 did not possess a similar range or average of T6SS genes (Figure 1),
191 which does not support a hypothesis that there is an association between the MDR
192 phenotype and T6SS carriage.
193

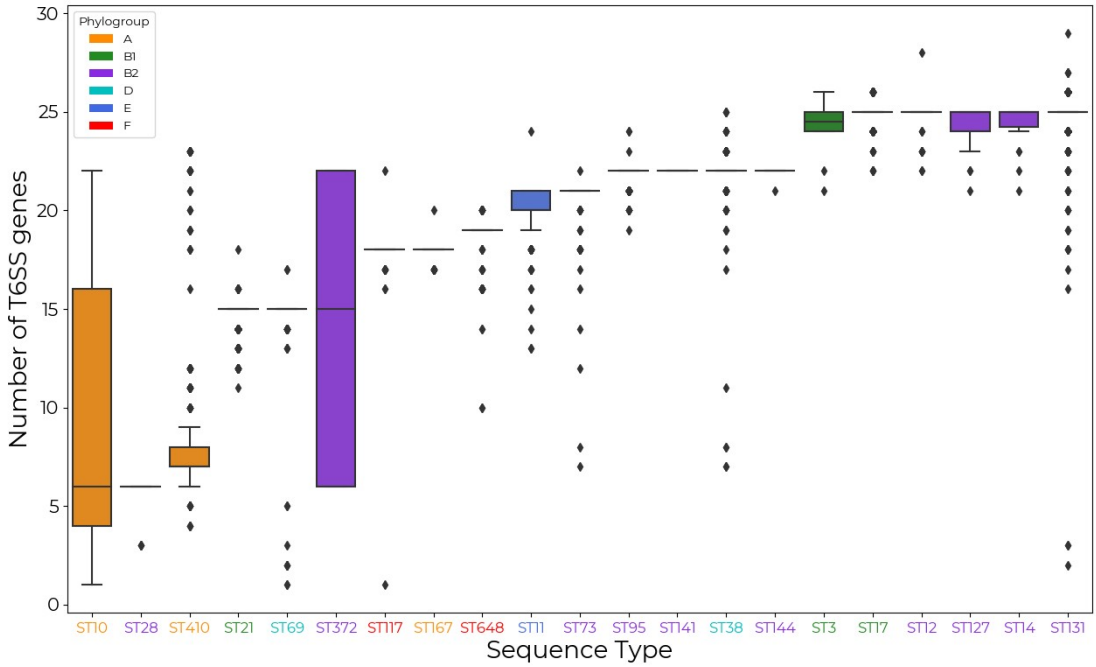
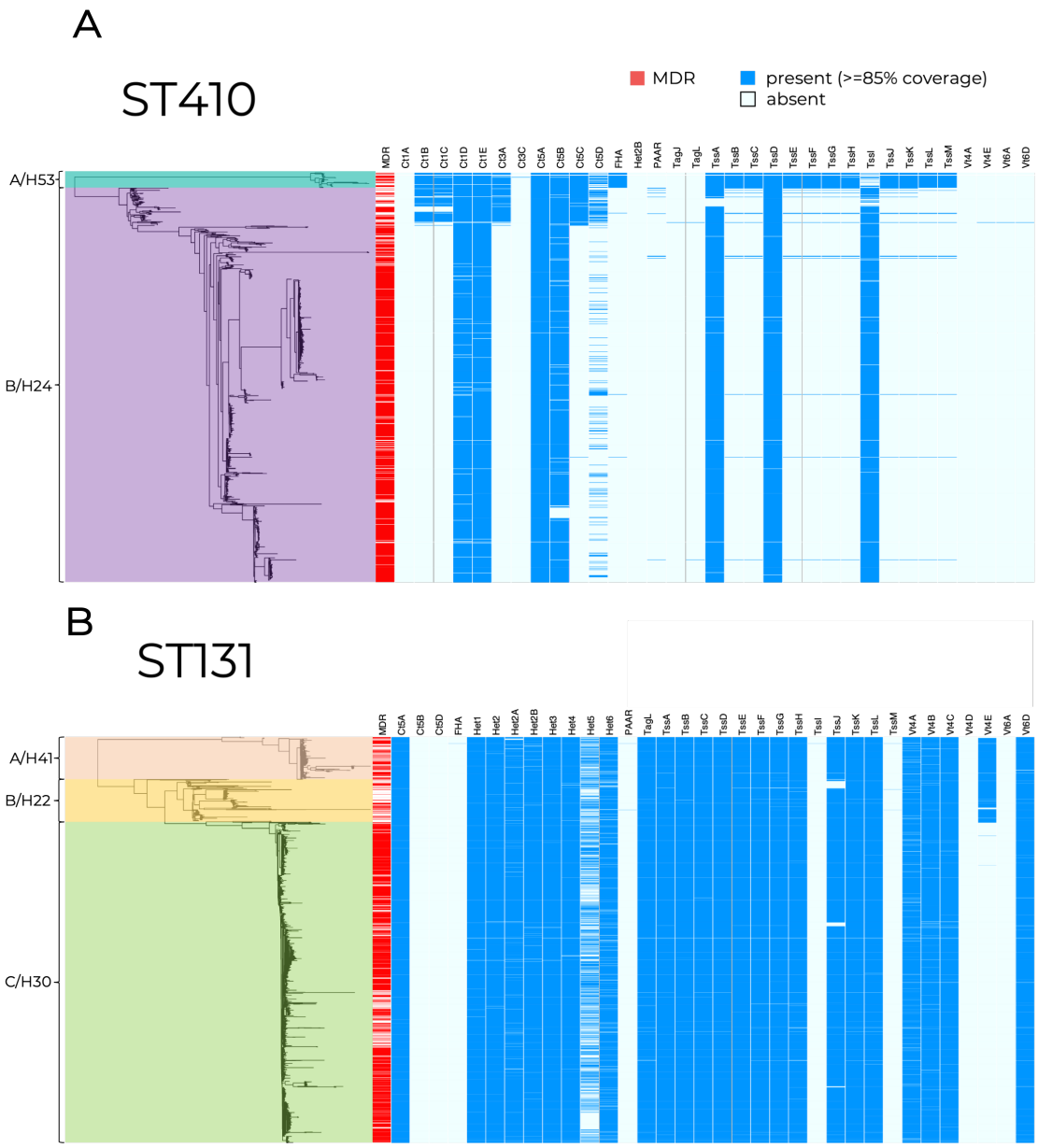


Fig. 1 Distributions of the number of type VI secretion system (T6SS) genes from the SecReT6 experimentally validated database [33; 34] across 21 sequence types of *E. coli* shows gene carriage varies by sequence type and phylogroup. Sequence types are colour coded according to their respective phylogroups: orange = A; green = B1; purple = B2; light blue = D; blue = E; red = F. A gene is determined as present if it exceeds a DNA identity of 75% and coverage of 85%.

Pattern of T6SS loss in MDR clones

We further focused on STs containing MDR clones in greater detail. Figure 2 displays the presence of T6SS genes within the major MDR lineages ST410 (Figure 2a) and ST131 (Figure 2b). Both STs demonstrate a correlation between multidrug resistance and the absence of T6SS genes.



207

208 **Fig. 2** Presence of MDR genotype and experimentally validated type VI secretion genes

209 taken from the SecReT6 database [33; 34] across (A) 1006 ST410 genomes and (B) 3186

210 ST131 genomes. Phylogenetic clades are shaded and labelled clade name/*fimH* allele.

211 Within ST410, a *tss* operon was present in 100% (n = 35) of clade A/H53 (Figure 2a) and

212 1.44% (n = 14) of clade B/H24 genomes. While isolates from clade A/H53 are largely drug

213 susceptible, clade B/H24 is a globally distributed multidrug resistant lineage [45;46]. The

214 conserved T6SS operon of ST410-A/H53 was confirmed as being chromosomal using
215 mlplasmid [35]. In the 14 rare instances where a full *tss* operon was present within the
216 B/H24 clade, nine genomes contained T6SS regions most closely related to plasmid
217 borne T6SS regions. Of these nine instances, five were most closely related to the F-type
218 plasmid pSTEC299_1 (GenBank accession: CP022280) and four to another F-type
219 plasmid, pRHB33C08_3 (GenBank accession: CP057209). The two plasmid-associated *tss*
220 regions were clearly distinct, sharing just 82.75% sequence identity (12289 bp
221 compared). The sporadically present ST410-B/H24 T6SS genes also shared low amino
222 acid sequence identity with the conserved T6SS genes of ST410-A/H53 (33 - 99%
223 identity). These findings together suggest multiple acquisitions of T6SS regions have
224 occurred within this clade.

225 Within ST131, the ancestral and less drug-resistant clades A and B possess a gene
226 labelled Vt4E, which is not present in the MDR clone of clade C (Figure 2b). To deduce
227 the function of Vt4E, we used BLASTn to search [37] the GenBank nonredundant
228 nucleotide database, which revealed that Vt4E shared 100% sequence identity with
229 *vasK*. The *vasK* gene is a homologue of *tssM* [47], which is a component of the inner
230 membrane complex of T6SS [25]. The absence of a functional VasK protein has been
231 shown to render T6SS non-functional in *Vibrio cholerae* [47].

232 Together, the observed patterns of presence/absence of T6SS genes in ST410 and ST131
233 could conceivably point towards a common evolutionary trajectory, where the transition
234 from a largely drug-susceptible to a multi-drug resistant phenotype coincides with the
235 loss of T6SS genes.

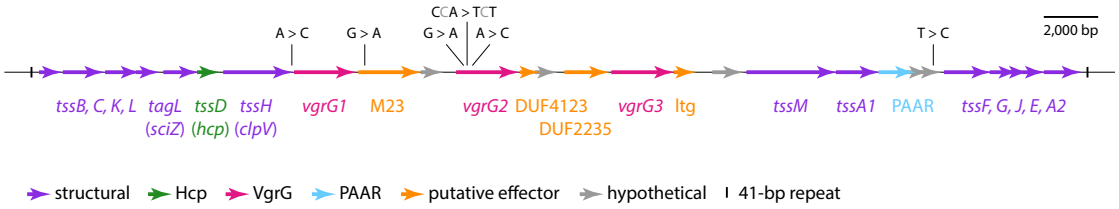
236 ## Characterisation of the ST131 type VI secretion locus

237 To facilitate comparative analyses, we sought to characterise the T6SS-determining
238 region of the ancestral ST131 lineage, ST131-H41 (clade A). First, we generated a
239 complete genome sequence for the ST131-H41 strain USVAST219 (PRJNA943186). We
240 found a complete and uninterrupted sequence of the T6SS locus in a 39211 bp region of
241 the USVAST219 chromosome, flanked by copies of a perfect 41 bp repeat sequences
242 (Figure 3). The region contained 27 open reading frames (ORFs), 22 of which could be
243 assigned functions. Among these were representatives of all 13 *tss* genes required for
244 synthesis of a type VI secretion apparatus (*tssA* to *tssM*), along with the structural gene
245 *tagL* and a second, different *tssA* gene. Genes for Hcp and PAAR-domain proteins, along
246 with three different *vgrG* (also called *tssI*) genes, encode the core and spike of the
247 secretion apparatus. Downstream of each *vgrG*, we found ORFs that encode putative
248 T6SS effector proteins; a M23-family peptidase, a lytic transglycosylase and proteins
249 with DUF4123 and DUF2235 domains, which have been associated with T6SS effectors
250 [48]. We conclude that this region of the ST131 chromosome includes all determinants
251 sufficient for production of a functional T6SS.

252 Reference strain EC958 (GenBank accession: HG941718) [49] was used to represent
253 ST131-H30Rx (clade C) in comparisons with USVAST219. The T6SS regions of USVAST219
254 and EC958 were almost identical, differing by just seven single nucleotide
255 polymorphisms (SNPs) (shown in Figure 3). The only SNP located in a structural gene
256 (*tssH*) was silent. We therefore conclude that this chromosomal region encodes the
257 native T6SS of ST131, and is conserved in clades A, B, and C, consistent with the

258 presence/absence data shown in Figure 2. However, the EC958 T6SS region does not

259 encode TssM.



260

261 **Fig. 3** The ST131 chromosomal T6SS locus. The extents and orientations of open reading

262 frames are indicated by labelled arrows, with colours corresponding to encoded protein

263 types as outlined in the key below. SNPs that differentiate EC598 from USVAST219 are

264 shown in black text above. M23: M23-family peptidase, ltg: lytic transglycosylase. Drawn

265 to scale from GenBank accession CP120633.

266 *tssM* is interrupted by IS*Ec12* in ST131-H30Rx

267 To determine the cause of *tssM* loss in ST131-H30Rx, we examined the complete

268 genome of reference strain EC958 [49] in further detail. We found that *tssM* had been

269 interrupted by a copy of IS*Ec12*, which was flanked by the 5 bp target site duplication

270 ACTGC (Figure 4). The *tssM* reading frame was split approximately in half by the

271 insertion, with 1,570 bp located to the left of IS*Ec12*, and 1,768 bp to the right. To deduce

272 whether other *tssM* homologues were present in EC958, and could encode

273 replacements for the interrupted *tssM*, we used tBLASTn to query the complete EC958

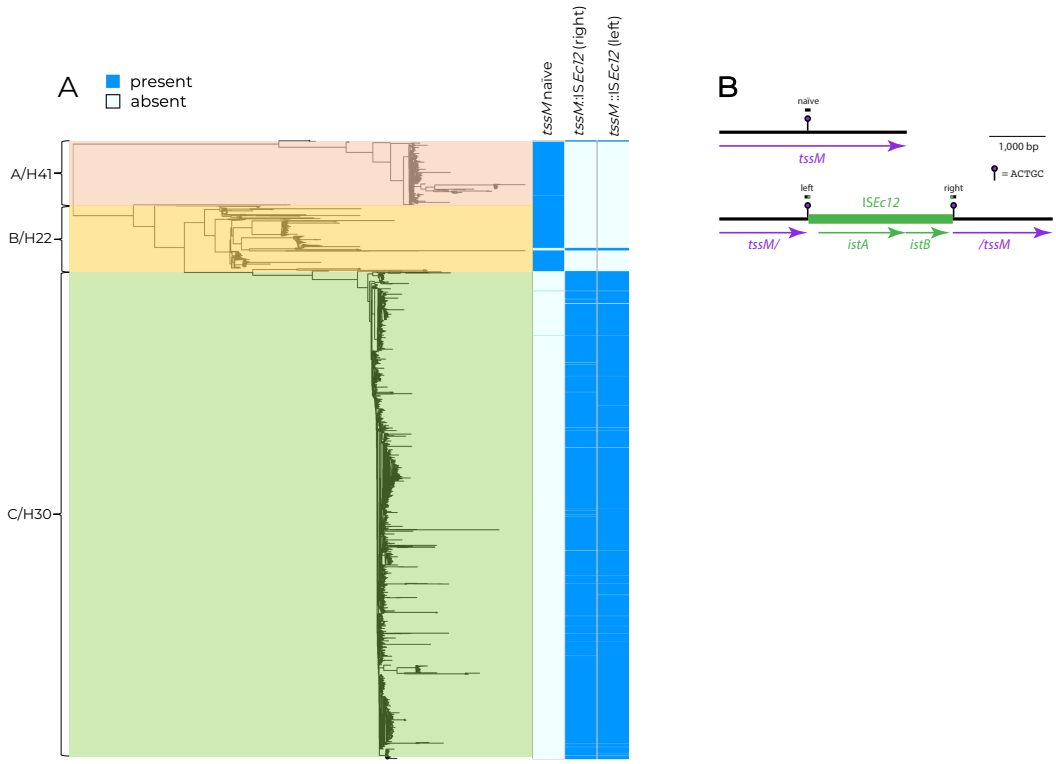
274 chromosome with the TssM sequence encoded by USVAST219. This query returned no
275 hits, apart from the fragmented sequences produced by the IS*Ec12*-interrupted *tssM*.
276 The IS*Ec12* insertion has rendered *tssM* non-functional in EC958, and in the absence of
277 *tssM* redundancy, it appears most likely that the T6SS itself is also non-functional.

278 From the EC958 genome, we generated 100 bp sequences that span the left and right
279 ends of the IS*Ec12* insertion (marked left and right in Figure 4b), and from the
280 USVAST219 genome a 100 bp sequence that represents an ancestral *tssM* uninterrupted
281 by this insertion (marked naïve in Figure 4b). To determine whether the same insertion
282 event was responsible for the absence of *tssM* in all ST131-H30Rx genomes, we screened
283 the ST131 collection for the three 100 bp sequences. The naïve sequence was absent in
284 just 2.97% (n = 21) of clade A and B genomes, confirming that almost all of them
285 contained uninterrupted *tssM* (Figure 4a). 2462 out of the 2478 (99.48%) clade C
286 genomes lacked the naïve sequence but contained one or both of the left and right
287 IS*Ec12* junction sequences (Figure 3). Of the 16 remaining clade C genomes, three
288 genomes contained the naïve sequence, and 13 genomes did not include either the naïve
289 sequence or either of the IS*Ec12* junction sequences (Figure 4a). Manual inspection of
290 the 13 genomes that lacked indicator sequences revealed that in six the IS*Ec12* insertion
291 was present, but low assembly quality had resulted in truncated contigs that were too
292 short for recognition with the strict threshold used for junction sequence detection. The
293 remaining seven genomes were excluded from further analysis, as low coverage appears
294 to have resulted in highly fragmented assemblies such that the structure of the T6SS
295 region was impossible to determine accurately. Thus, while *tssM* is uninterrupted in
296 97.03% (n = 687) of ST131 clade A and B genomes, it has been interrupted by IS*Ec12* at
297 exactly the same position in 99.78% (n = 2465) of clade C genomes. We conclude that

298 the loss of *tssM* in ST131 clade C was the result of a single IS*Ec12* insertion event that

299 occurred in an ancestral strain.

300



301

302 **Fig. 4** (A) ST131 phylogeny. Genomes are shaded according to their clade and *fimH* allele.

303 The presence of 100 bp IS*Ec12* insertion event indicator sequences is dictated by blue

304 shading to the right of the phylogeny. (B) Interruption of *tssM* by IS*Ec12*. Chromosomal

305 sequence is shown as a thin black line, and IS*Ec12* as a thicker blue line, with open

306 reading frames indicated by arrows beneath. The positions of target site duplication

307 sequences and extents of indicator sequences are shown above. Drawn to scale using

308 sequence obtained from GenBank accessions: CP120633 and HG94718.

309 **ST131-H30Rx out-competes without a complete T6SS**

310 To determine whether USVAST219 (ST131-H41, T6SS+) and UTI62 (ST131-H30Rx,
311 T6SS-) exhibit differential T6SS-mediated antibacterial activity, we used a kanamycin-
312 resistant K-12 MG1655 strain as prey in bacterial killing assays. Growth during co-
313 incubation was compared to growth in monoculture for each strain in all experiments. A
314 clear reduction in growth of MG1655 was observed after a four-hour co-incubation with
315 UTI62 compared to monoculture (Figure 5a). Coincubation of MG1655 with USVAST219
316 did not show a clear reduction in growth when compared to monoculture of MG1655
317 (Figure 5a). Next, we tested the growth levels of the two ST131 strains, UTI62 and
318 USVAST219, after a four-hour co-incubation period. Neither ST131 strain exhibited a
319 reduction in growth following co-incubation compared to monoculture (Figures 5b, c).
320 Our results show that the growth of ST131-H30Rx strain UTI62 (T6SS-) is not diminished
321 by the presence of other *E. coli*, and that it out-competes MG1655 (T6SS-) *in vitro*.
322

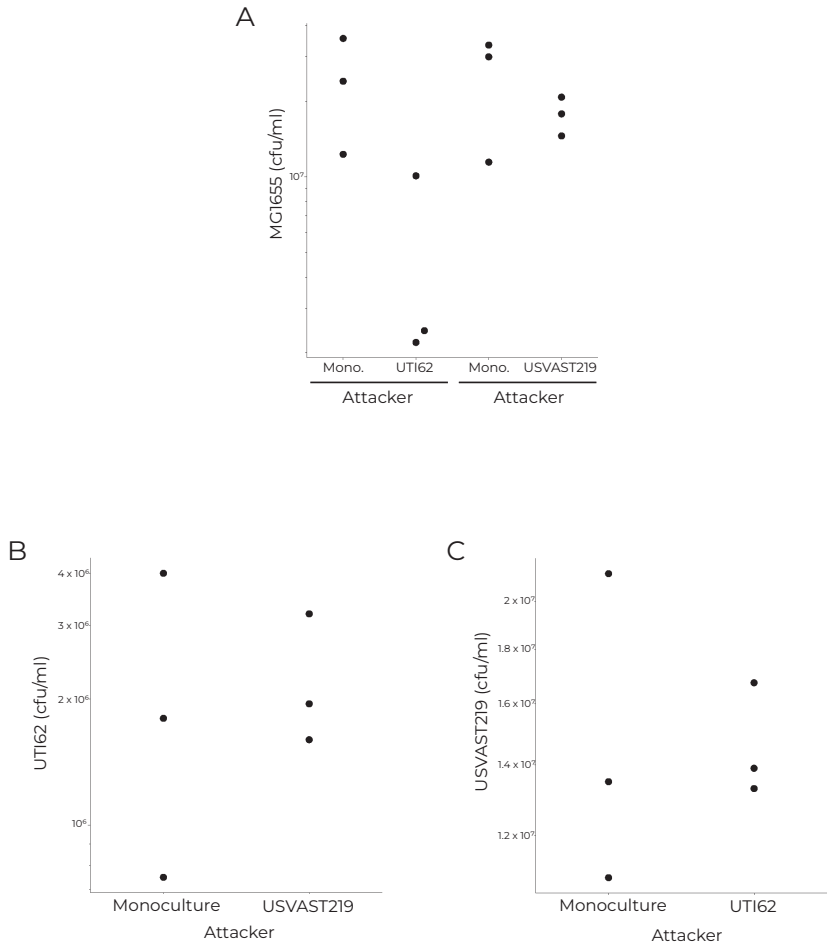


Fig. 5 Pairwise competition assays. Measurements taken after 4 hours incubation. (A) Growth of MG1655 in monoculture (mono.) and in the presence of UTI62 (T6SS-) and USVAST219 (T6SS+). (B) Growth of UTI62 in monoculture and in the presence of USVAST219. (C) Growth of USVAST219 in monoculture and in the presence of UTI62.

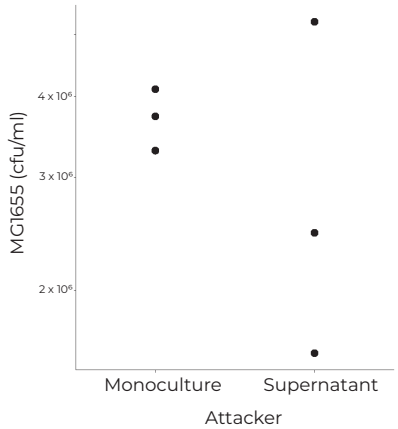
UTI62 does not produce lytic phage or bacteriocins in lab growth conditions

To further deduce the biological mechanism allowing UTI62 to out-compete MG1655, we grew MG1655 to mid-exponential phase and then challenged with the filter sterilised

331 supernatant from a UTI62 liquid culture to test for the production of lytic phages or
332 bacteriocins by UTI62 that could be responsible for the reduction in growth of MG1655.
333 Growth during co-incubation was compared to growth during incubation of MG1655 in
334 monoculture. Co-incubation of MG1655 with the supernatant of UTI62 did not result in as
335 large a reduction of growth for MG1655 in comparison to co-incubation with UTI62 itself
336 (Figure 6). This contrasts with our previous result, where the growth of MG1655 was
337 reduced in the presence of UTI62 (Figure 5a). This finding suggests the production of lytic
338 phages or bacteriocins is not responsible for UTI62 out-competing MG1655 *in vitro*.
339 However, phages or toxins might still be produced under different growth conditions. To
340 investigate this possibility, the genome of UTI62 was screened for colicin/bacteriocin genes
341 using the BACTIBASE database [38], and for phages with PHASTER [39; 40]. No bacteriocins
342 were detected and none of 12 prophage regions detected by PHASTER contained lysis genes.
343 These results suggest that diffusible elements were not playing a role in the observed
344 reduction of MG1655 growth in the presence of UTI62.

345

346



347

348 **Fig. 6** Competition assay. Measurements taken after 4 hours incubation. Growth of
349 MG1655 in monoculture compared to growth in the presence of liquid supernatant of
350 UTI62.

351 Discussion

352 Extensive work has been done to characterise the T6SS in terms of its structure, function,
353 variability and prevalence in various genera. However, this is the first study, to the best
354 of our knowledge, to examine the prevalence of T6SSs across a broad range of *E. coli*
355 phylogroups, sequence types and pathotypes. We determined the prevalence of the *tss*
356 operon within lineages known to contain pandemic MDR clones, completed the first
357 detailed characterisation of the T6SS locus in ST131, and identified an evolutionary event
358 within ST131 that led to the loss of the key structural T6SS component TssM. We also
359 tested the competitive capabilities of ST131 strains with and without a complete *tss*
360 operon experimentally. Our study's combination of wet and dry lab approaches has

361 yielded new insights on T6SS functionality in the MDR *E. coli* lineage ST131 and provided
362 a broader overview of T6SS presence in other lineages.

363 The similarities we found between ST410 and ST131 suggests we may have uncovered a
364 common evolutionary trajectory in two MDR lineages, where a functional T6SS was lost
365 before the expansion of an MDR clone. We have previously described the important role
366 of potentiating mutations involved in colonisation factors, anaerobic metabolism genes,
367 and intergenic regions as being common evolutionary trajectories in MDR clone
368 formation [11]. The loss or degradation of a T6SS could perhaps integrate into the
369 stepwise evolutionary course of ST410 and ST131 pandemic clones.

370 We found that the T6SS in ST131-H30 was inactivated by an insertion event. The
371 interruption of *tssM* by IS*Ec12* occurring prior to or during the formation of ST131-H30
372 (clade C) implies that this event may have contributed to the successful emergence and
373 expansion of the clade. The rare instances where *tssM* is interrupted in clade A/B, and
374 uninterrupted in clade C (Figure 4a), might be explained by inter-lineage horizontal gene
375 transfer and homologous recombination, but the T6SS region is too well-conserved
376 (>99.99% ID between lineages) to accurately detect recombinant sequences. We
377 speculate that the lack of a functional T6SS allowed clade C to become more receptive
378 to cell-to-cell contact, and therefore plasmid transfer, which may have enabled the
379 acquisition of MDR plasmids that have previously been described as a vital step in the
380 evolution of pandemic clones [11].

381 A caveat of this work is that contact-dependent killing capabilities were only tested *in*
382 *vitro*. An *in vivo* experiment would be better suited to capture the complex host-
383 associated microbial community of a mammalian gut, as numerous external factors

384 cannot be applied under standard laboratory conditions. Interactions with other species
385 within the gut microbiome could not be replicated here, although recent research has
386 shown that the displacement of commensal *E. coli* with MDR lineages does not disrupt
387 the wider gut microbiome [9]. *In vivo* work may also more accurately replicate the
388 diversity of nutrients available within the gut and any potential interactions with host
389 cells. These factors would need to be considered to draw further conclusions on the
390 superior displacement and colonisation capabilities that could be associated with the
391 loss of T6SSs in MDR lineages.

392 We observed that the T6SS-negative UTI62 (ST131-H30Rx) out-competed other *E. coli*
393 and confirmed through genomic analysis that this was not due to UTI62 secreting toxins,
394 or the interference of lytic phages. Therefore, the data presented here does not support
395 the hypothesis that the ability of drug-resistant *E. coli* to displace resident commensal *E.*
396 *coli* is due to the production of a T6SS, lytic phage or bacteriocins, so another explanation
397 is needed. This work strengthens our hypothesis [10, 11] that metabolism is the key
398 factor in MDR *E. coli* displacing commensal *E. coli* from the gut.

399 Author Statements

400 Author Contributions

401 E.A.C.: data curation, methodology, formal analysis, visualisation, writing – original draft
402 preparation and review and editing. R.A.M.: methodology, formal analysis, writing –
403 original draft preparation and review and editing. A.E.S: methodology. R.J.H.:

methodology, writing - review and editing. C.H.C.: data curation. S.J.D.: data curation. A.M.: conceptualisation, supervision, writing – review and editing.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Funding Information

EAC was funded by the Wellcome Antimicrobial and Antimicrobial Resistance (AAMR) DTP (108876B15Z).

Acknowledgements

We would like to thank Professor Jose Bengoechea for helpful discussion and all members of the McNally group for their support.

References

[1] Cummins EA, Hall RJ, Connor C, McInerney JO, McNally A. Distinct evolutionary trajectories in the *Escherichia coli* pangenome occur within sequence types. Microbial Genomics. 2022;8(11).

[2] Murray CJ, Ikuta KS, Sharara F, Swetschinski L, Robles Aguilar G, Gray A, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. The Lancet. 2022 2;399(10325):629-55.

421 [3] Connor CH, Zucoloto AZ, Yu IL, Corander J, McDonald B, McNally A, et al. Multi-drug
422 resistant E. coli displace commensal E. coli from the intestinal tract, a trait
423 associated with elevated levels of genetic diversity in carbohydrate metabolism
424 genes. bioRxiv. 2022.
425
426 [4] Arcilla MS, van Hattem JM, Haverkate MR, Bootsma MCJ, van Genderen PJJ,
427 Goorhuis A, et al. Import and spread of extended-spectrum *β*-lactamase producing
428 Enterobacteriaceae by international travellers (COMBAT study): a prospective,
429 multicentre cohort study. The Lancet Infectious Diseases. 2017 1;17(1):78-85.

430 [5] Bevan ER, McNally A, Thomas CM, Piddock LJV, Hawkey PM. Acquisition and loss
431 of CTX-M-producing and non-producing *Escherichia coli* in the fecal microbiome of
432 travelers to South Asia. mBio. 2018 11;9(6).

433 [6] Kantele A, Kuenzli E, Dunn SJ, Dance DAB, Newton PN, Davong V, et al. Dynamics
434 of intestinal multidrug-resistant bacteria colonisation contracted by visitors to a
435 high-endemic setting: a prospective, daily, real-time sampling study. The Lancet
436 Microbe. 2021;2(4):e151-8.
437
438 [7] Kallonen T, Brodrick HJ, Harris SR, Corander J, Brown NM, Martin V, et al.
439 Systematic longitudinal survey of invasive Escherichia coli in England demonstrates
440 a stable population structure only transiently disturbed by the emergence of
441 ST131. Genome Research. 2017;27(8):1437-49.

442 [8] Gladstone RA, McNally A, Pöntinen AK, Tonkin-hill G, Lees JA, Skytén K, et al.
443 Emergence and dissemination of antimicrobial resistance in Escherichia coli

444 causing bloodstream infections in Norway in 2002 – 17 : a nationwide , longitudinal
445 , microbial population genomic study. The Lancet Microbe.
446 2021;5247(21):1-11.

447 [9] Davies M, Galazzo G, van Hattem JM, Arcilla MS, Melles DC, de Jong MD, et al.
448 Enterobacteriaceae and Bacteroidaceae provide resistance to travel-associated
449 intestinal colonization by multi-drug resistant Escherichia coli.
450 Gut Microbes. 2022;14(1).

451 [10] Dunn SJ, Connor C, McNally A. The evolution and transmission of multidrug
452 resistant Escherichia coli and Klebsiella pneumoniae: the complexity of clones and
453 plasmids. Current Opinion in Microbiology. 2019;51:51-6.

454 [11] Cummins EA, Snaith AE, McNally A, Hall RJ. The role of potentiating mutations in
455 the evolution of pandemic Escherichia coli clones. European Journal of Clinical
456 Microbiology & Infectious Diseases. 2021.

457 [12] Riley MA, Gordon DM. The ecological role of bacteriocins in bacterial competition.
458 1999.

459 [13] Ellis EL, Delbruck M. THE GROWTH OF BACTERIOPHAGE. The Journal of General
460 Physiology. 1938.

461 [14] Mariano G, Trunk K, Williams DJ, Monlezun L, Strahl H, Pitt SJ, et al. A family of Type
462 VI secretion system effector proteins that form ion-selective pores. Nature
463 Communications. 2019 12;10(1).

464 [15] Pukatzki S, Ma AT, Revel AT, Sturtevant D, Mekalanos JJ. Type VI secretion system
465 translocates a phage tail spike-like protein into target cells where it cross-links
466 actin. Proceedings of the National Academy of Sciences of the United States of
467 America. 2007 9;104(39):15508-13.

468 [16] Salomon D, Klimko JA, Trudgian DC, Kinch LN, Grishin NV, Mirzaei H, et al. Type VI
469 Secretion System Toxins Horizontally Shared between Marine Bacteria. PLoS
470 Pathogens. 2015 8;11(8).

471 [17] Bernal P, Allsopp LP, Filloux A, Llamas MA. The Pseudomonas putida T6SS is a plant
472 warden against phytopathogens. ISME Journal. 2017 4;11(4):972-
473 87.

474 [18] Troselj V, Treuner-Lange A, Søgaard-Andersen L, Wall D. Physiological
475 Heterogeneity Triggers Sibling Conflict Mediated by the Type VI Secretion System
476 in an Aggregative Multicellular Bacterium. mBio. 2018;9(1).

477 [19] Tang L, Yue S, Li GY, Li J, Wang XR, Li SF, et al. Expression, secretion and bactericidal
478 activity of type VI secretion system in Vibrio anguillarum. Archives of Microbiology.
479 2016 10;198(8):751-60.

480 [20] Unni R, Pintor K, Diepold A, Unterweger D. Presence and absence of type VI
481 secretion systems in bacteria. Microbiology. 2022;in print:1-13.

482 [21] Kempnich MW, Sison-Mangus MP. Presence and abundance of bacteria with the
483 Type VI secretion system in a coastal environment and in the global oceans. PLoS
484 ONE. 2020 12;15(12 December 2020).

485 [22] García-Bayona L, Coyne MJ, Comstock LE. Mobile Type VI secretion system loci of
486 the gut Bacteroidales display extensive intra-ecosystem transfer, multi-species
487 spread and geographical clustering. PLoS Genetics. 2021
488 4;17(4).

489 [23] Abby SS, Cury J, Guglielmini J, Néron B, Touchon M, Rocha EPC. Identification of
490 protein secretion systems in bacterial genomes. Scientific Reports.
491 2016 3;6.

492 [24] Bingle LE, Bailey CM, Pallen MJ. Type VI secretion: a beginner's guide.
493 Current Opinion in Microbiology. 2008 2;11(1):3-8.

494 [25] Journet L, Cascales E. The Type VI Secretion System in Escherichia coli and Related
495 Species. EcoSal Plus. 2016;7(1).

496 [26] Ma J, Bao Y, Sun M, Dong W, Pan Z, Zhang W, et al. Two functional type VI secretion
497 systems in avian pathogenic Escherichia coli are involved in different pathogenic
498 pathways. Infection and Immunity. 2014;82(9):3867-79.

499 [27] De Pace F, Nakazato G, Pacheco A, De Paiva JB, Sperandio V, Da Silveira WD. The
500 type VI secretion system plays a role in type 1 fimbria expression and pathogenesis
501 of an avian pathogenic Escherichia coli strain. Infection and Immunity. 2010
502 12;78(12):4990-8.

503 [28] Brunet YR, Espinosa L, Harchouni S, Mignot T, Cascales E. Imaging Type VI
504 Secretion-Mediated Bacterial Killing. Cell Reports. 2013;3(1):36-41.

505 [29] Weber BS, Ly PM, Irwin JN, Pukatzki S, Feldman MF. A multidrug resistance plasmid
506 contains the molecular switch for type VI secretion in Acinetobacter baumannii.
507 Proceedings of the National Academy of Sciences of the United States of America.
508 2015;112(30):9442-7.

509 [30] Venanzio GD, Moon KH, Weber BS, Lopez J, Ly PM, Potter RF, et al. Multidrug-
510 resistant plasmids repress chromosomally encoded T6SS to enable their
511 dissemination. Proceedings of the National Academy of Sciences of the United
512 States of America. 2019;116(4):1378-83.

513 [31] Zhou Z, Alikhan NF, Mohamed K, Fan Y, Achtman M. The EnteroBase user's guide,
514 with case studies on Salmonella transmissions, Yersinia pestis phylogeny, and
515 Escherichia core genomic diversity. Genome Research. 2020;30(1):138-52.

516 [32] Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, et al. Validating
517 the AMRFINder tool and resistance gene database by using antimicrobial resistance
518 genotype-phenotype correlations in a collection of isolates.
519 Antimicrobial Agents and Chemotherapy. 2019;63(11).

520 [33] Li J, Yao Y, Xu HH, Hao L, Deng Z, Rajakumar K, et al. SecReT6: A web-based resource
521 for type VI secretion systems found in bacteria. Environmental Microbiology. 2015
522 7;17(7):2196-202.

523 [34] Zhang J, Guan J, Wang M, Li G, Djordjevic M, Tai C, et al. SecReT6 update:
524 a comprehensive resource of bacterial Type VI Secretion Systems. Science China
525 Life Sciences. 2022 11.

526 [35] Arredondo-Alonso S, Rogers MRC, Braat JC, Verschuuren TD, Top J, Corander J, et
527 al. Mlplasmids: a User-Friendly Tool To Predict Plasmid- and Chromosome-Derived
528 Sequences for Single Species. Microbial genomics. 2018;4(11).
529
530 [36] Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. ISfinder: the reference
531 centre for bacterial insertion sequences. Nucleic acids research.
532 2006;34 (Database issue).
533 [37] Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al.
534 BLAST+: Architecture and applications. BMC Bioinformatics. 2009 12;10.
535 [38] Hammami R, Zouhir A, Ben Hamida J, Fliss I. BACTIBASE: A new web-accessible
536 database for bacteriocin characterization. BMC Microbiology. 2007;7.
537
538 [39] Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: A Fast Phage Search Tool.
539 Nucleic Acids Research. 2011 7;39(SUPPL. 2).
540 [40] Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster
541 version of the PHAST phage search tool. Nucleic Acids Research.
542 2016 7;44(W1):W16-21.
543 [41] Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome
544 assemblies from short and long sequencing reads. PLoS Computational Biology.
545 2017 6;13(6).
546 [42] Bottery MJ, Wood AJ, Brockhurst MA. Adaptive modulation of antibiotic resistance
547 through intragenomic coevolution. Nature Ecology and Evolution.

548 2017 9;1(9):1364-9.

549 [43] Colpan A, Johnston B, Porter S, Clabots C, Anway R, Thao L, et al. Escherichia coli
550 sequence type 131 (ST131) subclone h30 as an emergent multidrug-resistant
551 pathogen among US Veterans. Clinical Infectious Diseases. 2013 11;57(9):1256-65.

552 [44] Croxall G, Weston V, Joseph S, Manning G, Cheetham P, McNally A. Increased
553 human pathogenic potential of Escherichia coli from polymicrobial urinary tract
554 infections in comparison to isolates from monomicrobial culture samples. Journal
555 of Medical Microbiology. 2011 1;60(1):102-9.

556 [45] Feng Y, Liu L, Lin J, Ma K, Long H, Wei L, et al. Key evolutionary events in the
557 emergence of a globally disseminated, carbapenem resistant clone in the
558 Escherichia coli ST410 lineage. Communications Biology. 2019;2(1):1-13.

559 [46] Roer L, Overballe-Petersen S, Hansen F, Schønning K, Wang M, Røder BL, et al.
560 Escherichia coli Sequence Type 410 Is Causing New International HighRisk Clones.
561 mSphere. 2018;3(4):00337-18.

562 [47] Kostiuk B, Santoriello FJ, Diaz-Satizabal L, Bisaro F, Lee KJ, Dhody AN, et al. Type VI
563 secretion system mutations reduced competitive fitness of classical Vibrio cholerae
564 biotype. Nature Communications. 2021;12(1):1-11.

565 [48] Salomon D, Kinch LN, Trudgian DC, Guo X, Klimko JA, Grishin NV, et al. Marker for
566 type VI secretion system effectors. Proceedings of the National Academy of
567 Sciences of the United States of America. 2014 6;111(25):9271-6.

568 [49] Forde BM, Ben Zakour NL, Stanton-Cook M, Phan MD, Totsika M, Peters KM, et al.
569 The complete genome sequence of *Escherichia coli* EC958: A high quality reference
570 sequence for the globally disseminated multidrug resistant E. coli O25b:H4-ST131
571 clone. PLoS ONE. 2014 8;9(8).