

1 **High-risk *Escherichia coli* clones that cause neonatal meningitis and association with**
2 **recrudescence infection**

3

4 Nguyen Thi Khanh Nhu^{1,2,3}, Minh-Duy Phan^{1,2,3}, Steven J. Hancock^{2,3#}, Kate M. Peters^{1,2,3},
5 Laura Alvarez-Fraga^{2,3^}, Brian M. Forde^{3,4}, Stacey B. Andersen⁵, Thyl Miliya⁶, Patrick N.A.
6 Harris^{4,7}, Scott A. Beatson^{2,3}, Sanmarie Schlebusch^{4,7,8}, Haakon Bergh⁷, Paul Turner^{6,9},
7 Annelie Brauner¹⁰, Benita Westerlund-Wikström¹¹, Adam D. Irwin^{3,4,12*}, and Mark A.
8 Schembri^{1,2,3*}

9

10 ¹*Institute for Molecular Bioscience (IMB), The University of Queensland, Brisbane,*
11 *Queensland, Australia*

12 ²*School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane,*
13 *Queensland, Australia*

14 ³*Australian Infectious Diseases Research Centre, The University of Queensland, Brisbane,*
15 *Queensland, Australia*

16 ⁴*University of Queensland Centre for Clinical Research, The University of Queensland,*
17 *Brisbane, Australia*

18 ⁵*Genome Innovation Hub, The University of Queensland, Brisbane, Australia*

19 ⁶*Cambodia Oxford Medical Research Unit, Angkor Hospital for Children, Siem Reap,*
20 *Cambodia.*

21 ⁷*Pathology Queensland, Health Support Queensland, Australia*

22 ⁸*Q-PHIRE Genomics and Public Health Microbiology, Forensic and Scientific Services,*
23 *Coopers Plains, Brisbane, Queensland, Australia*

24 ⁹*Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine,*
25 *University of Oxford, Oxford, UK*

26 ¹⁰ *Department of Microbiology, Tumor and Cell Biology, Division of Clinical Microbiology,*

27 *Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden*

28 ¹¹ *Molecular and Integrative Biosciences Research Programme, University of Helsinki,*

29 *Helsinki, Finland*

30 ¹² *Infection Management Prevention Service, Queensland Children's Hospital, Brisbane,*

31 *Queensland, Australia.*

32

33 #Current Address: Wellcome-Wolfson Institute for Experimental Medicine, School of

34 Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, United

35 Kingdom.

36 ^Current address: INRAE, Univ Montpellier, LBE, 102 Avenue des Etangs, Narbonne 11100,

37 France.

38

39 **Running title:** High risk neonatal meningitis *E. coli* clones

40 **Word count:** Text (3493 words), Abstract (249 words).

41

42 ***Corresponding authors:**

43 Professor Mark Schembri, Institute for Molecular Bioscience, The University of Queensland,

44 Brisbane, Queensland 4072, Australia. E-mail: m.schembri@uq.edu.au.

45 Dr Adam Irwin, Centre for Clinical Research, University of Queensland, Brisbane,

46 Queensland, Australia. Email: a.irwin@uq.edu.au.

47

49 **Abstract**

50 Neonatal meningitis is a devastating disease associated with high mortality and neurological
51 sequelae. *Escherichia coli* is the second most common cause of neonatal meningitis in full-
52 term infants (herein NMEC) and the most common cause of meningitis in preterm neonates.
53 Here we investigated the genomic relatedness of a collection of 58 NMEC isolates spanning
54 1974-2020 and isolated from seven different geographic regions. We show NMEC are
55 comprised of diverse sequence types (STs), with ST95 (34.5%) and ST1193 (15.5%) the
56 most common. No single virulence factor was conserved in all isolates; however, genes
57 encoding fimbrial adhesins, iron acquisition systems, the K1 capsule, and O antigen types
58 O18, O75 and O2 were most prevalent. Antibiotic resistance genes occurred infrequently in
59 our collection. We also monitored the infection dynamics in three patients that suffered
60 recrudescence of invasive infection caused by the original infecting isolate despite appropriate
61 antibiotic treatment based on antibiogram profile and resistance genotype. These patients
62 exhibited severe gut dysbiosis. In one patient, the causative NMEC isolate was also detected
63 in the fecal flora at the time of the second infection episode and after treatment. Thus,
64 although antibiotics are the standard of care for NMEC treatment, our data suggests that
65 failure to eliminate the causative NMEC that resides intestinally can lead to the existence of a
66 refractory reservoir that may seed recrudescence of infection.

67

68 **Key words:** *E. coli*, neonatal meningitis, genomics, recrudescence of infection, recurrent
69 infection, gut dysbiosis

70 **Introduction**

71 Neonatal meningitis (NM) is a devastating disease with a mortality rate of 10-15% and severe
72 neurological sequelae including hearing loss, reduced motor skills and impaired development
73 in 30-50% of cases [1-3]. The incidence of disease is highest in low-income countries and
74 occurs at a rate of 0.1-6.1/1000 live births [3]. *Escherichia coli* is the second most common
75 cause of NM in full-term infants (herein NMEC), after group B *Streptococcus* (GBS) [4, 5],
76 and the most common cause of meningitis in preterm neonates [5, 6]. Together, these two
77 pathogens cause ~60% of all cases, with on average one case of NMEC for every two cases
78 of GBS [7, 8]. In several countries, NM incidence caused by GBS has declined due to
79 maternal intrapartum antibiotic prophylaxis; however, NM incidence caused by *E. coli*
80 remains the same [7, 9]. Moreover, NMEC is a significant cause of relapsed infections in
81 neonates [10].

82

83 NMEC are categorised genetically based on multi-locus sequence type (ST) or by serotyping
84 based on cell-surface O antigen (O), capsule (K) and flagella (H) antigens. Analysis of
85 NMEC diversity in France revealed ~25% of isolates belong to the ST95 clonal complex
86 (STc95) [11], however, a global picture of NMEC epidemiology is lacking. NMEC possess a
87 limited diversity of serotypes, dominated by O18:K1:H7, O1:K1, O7:K1, O16:K1, O83:K1
88 and O45:K1:H7, which together account for >70% of NMEC [12-15]. Notably, ~80% of
89 NMEC express the K1 capsule, a polysaccharide comprising linear homopolymers of α 2-8-
90 linked N-acetyl neuraminic acid [12, 16]. Apart from the K1 capsule, specific NMEC
91 virulence factors are less-well defined, though studies have demonstrated a role for S
92 fimbriae [17], the outer membrane protein OmpA [18], the endothelial invasin IbeA [19] and
93 the cytotoxin necrotising factor CNF1 [20] in translocation of NMEC across the blood-brain
94 barrier and pathogenesis. A large plasmid encoding colicin V (ColV), colicin Ia bacteriocins

95 and several virulence genes including iron-chelating siderophore systems has also been
96 strongly associated with NMEC virulence [21].

97

98 Despite being the second major NM aetiology, genomic studies on NMEC are lacking, with
99 most reporting single NMEC complete genomes. Here we present the genomic analyses of a
100 collection of 58 NMEC isolates obtained from seven different geographic regions over 46
101 years to understand virulence gene content, antibiotic resistance and genomic diversity. In
102 addition, we provide a complete genome for 18 NMEC isolates representing different STs,
103 serotypes and virulence gene profiles, thus more than tripling the number of available NMEC
104 genomes that can be used as references in future studies. Three infants in our study suffered
105 recrudescence of invasive NMEC infection, and we show this was caused by the same isolate. We
106 further revealed that patients that suffered recrudescence of invasive infection had severe gut
107 dysbiosis, and detected the infecting isolate in the intestinal microflora, suggesting NMEC
108 colonisation of the gut provides a reservoir that can seed repeat infection.

109

110

111 **Results**

112 **Establishment of an NMEC collection from geographically diverse locations**

113 A collection of 52 NMEC isolates cultured from the blood or cerebrospinal fluid (CSF) of
114 neonates with meningitis was established with the addition of six completely sequenced
115 NMEC genomes available on the NCBI database. This yielded a final set of 58 NMEC
116 isolates spanning 1974 to 2020 (Supplementary Table 1). The isolates were obtained from
117 seven different geographic locations; Finland (n=17, 29.3%), Sweden (n=14, 24.1%),
118 Australia (n=15, 25.9%), Cambodia (n=7, 12.1%), USA (n=3, 5.2%), France (n=1, 1.7%) and
119 the Netherlands (n=1, 1.7%).

120

121 **ST95 and ST1193 are the two major STs of NMEC**

122 Phylogenetic analysis was performed on the 58 NMEC isolates, with an additional eight well-
123 characterised *E. coli* strains included for referencing (EC958, UTI89, MS7163, CFT073,
124 UMN026, 536, APEC01, MG1655) (Supplementary Table 1). The NMEC isolates were
125 diverse, and spanned phylogroups A, B2, C, D and F; the majority of isolates were from
126 phylogroup B2 (n=48, 82.8%). Overall, the isolates belonged to 22 STs, of which 15 STs
127 only contained one isolate. ST95 (n=20, 34.5%) and ST1193 (n=9, 15.5%) were the two most
128 common NMEC STs (Figure 1, Supplementary Table 1). ST95 isolates were obtained over
129 the entire study period, while ST1193 isolates were more recent and only obtained from
130 2013. Four isolates belonged to ST390 (6.9%), which is part of the STc95. One isolate
131 belonged to a novel ST designated ST11637, which is part of the ST14 clonal complex
132 (STc14) that also contains ST1193 (Figure 1, Supplementary Table 1). Isolates from other
133 common phylogroup B2 extra-intestinal pathogenic *E. coli* (ExPEC) lineages, ST131, ST73
134 and ST69, as well as several STs associated with environmental sources such as ST48 and
135 ST23, were detected in the collection. However, it is notable that the high incidence of NM
136 associated with ST95 and ST1193 does not reflect the broader high prevalence of major
137 ExPEC clones associated with human infections in the publicly available Enterobase database
138 [22] (Supplementary Figure 1), suggesting ST95 and ST1193 exhibit specific virulence
139 features associated with their capacity to cause NM.

140

141 Eighteen NMEC isolates were completely sequenced using complementary long-read Oxford
142 Nanopore Technology, enabling accurate comparison of NMEC genome size, genomic island
143 composition and location, and plasmid and prophage diversity (Supplementary Table 2).
144 These isolates spanned the diversity in the collection, representing 11 different STs, including

145 two ST1193 isolates (one with the dominant O75:H5 serotype and one with an unusual
146 O6:H5 serotype), five ST95 isolates with different serotypes, and one isolate from the novel
147 ST11637.

148

149 **Antibiotic resistance in NMEC**

150 Antibiotic resistance profiling revealed an overall low level of resistance in the collection.
151 The ST1193 isolates contained fluoroquinolone resistance defining mutations in *gyrA* (S83L
152 D87N) and *parC* (S80I), as previously described for this lineage [23]. In addition, 77.8% of
153 ST1193 isolates (7/9 isolates) also harboured at least one gene conferring resistance to
154 aminoglycosides (*aac(3)-IId*, *aadA5*, *aph(3'')-Ib* and *aph(6)-Id*), trimethoprim (*dftA17*) and
155 sulphonamides (*sul1* and *sul2*) (Supplementary Figure 2). Six out of the seven isolates from
156 Cambodia had more than one antibiotic resistance gene, likely reflecting increased antibiotic
157 resistance rates in this region [24]. Indeed, in addition to *gyrA* and *parC* mutations for
158 fluoroquinolone resistance, CAM-NMEC-6 contains 14 antibiotic resistance genes (including
159 resistance to third-generation cephalosporins and carbapenems) and CAM-NMEC-4 contains
160 11 antibiotic resistance genes (Supplementary Figure 2).

161

162 **Virulence factors in NMEC**

163 The isolates exhibited variable distribution of virulence genes previously linked to NMEC
164 pathogenesis. The most prevalent genes were those involved in iron uptake, including the
165 enterobactin (98%), yersiniabactin (98%), aerobactin (62%) and salmochelin (55%)
166 siderophore systems, and the heme receptors *chuA* (93%) and *hma* (62%) (Figure 1). Also
167 common were the *sitABCD* genes encoding an iron/manganese transporter (98%). The
168 presence of fimbrial and afimbrial adhesins was also diverse. The most prevalent adhesins
169 were type 1 fimbriae (100%), *mat (ecp)* fimbriae (98%) and the *fdeC* adhesin (98%). Genes

170 encoding P and S fimbriae were detected in 36% and 22% of NMEC isolates, respectively.
171 The most prevalent toxin was the uropathogenic-specific genotoxin *usp* (83%), which was
172 only found in phylogroup B2 isolates. Other toxin genes encoding the serine protease
173 autotransporters Vat (65% prevalence) and Sat (29%), hemolysin (12%) and cytotoxic
174 necrotizing factor-1 (7%) were less prevalent. Additional virulence genes included the *asIA*
175 arylsulfatase (95%), the *iss* lipoprotein (76%) and the *ibeA* invasin (33%). The ColV-plasmid
176 was present in 33% of the isolates (Figure 1, Supplementary Table 1). Direct comparison of
177 virulence factors between ST95 and ST1193, the two most dominant NMEC STs, revealed
178 that the ST95 isolates (n = 20) contained significantly more virulence factors than the
179 ST1193 isolates (n = 9); *P*-value < 0.001, Mann-Whitney two-tailed unpaired test
180 (Supplementary Table 1, Supplementary Figure 3).

181

182 **NMEC comprise a dominant K1 capsule type and a limited pool of O and H serotypes**

183 The capsule type of the NMEC isolates was determined by *in silico* typing. K1 was the
184 dominant capsule type in the collection (43/58 isolates, 74.1%) (Figure 1). Thirty-four of
185 these isolates were available for capsule testing, and we confirmed K1 expression by ELISA
186 in all but two isolates (Supplementary Figure 4). Other capsule types included K2, K5 and
187 K14 (Supplementary Table 1). A capsule type could not be resolved for 12 isolates, of which
188 eight did not possess a Group II or Group III capsule type based on the absence of the
189 conserved *kpsD* gene (Figure 1, Supplementary Table 1).

190

191 *In silico* O-antigen (O) and flagella (H) serotypes were also determined. O18 was the most
192 common O type (n=16, 27.6%), followed by O75 (n=8, 13.9%) and O2 (n=7, 12.1%). The
193 most dominant H types were H7 (n=19, 32.8%), H5 (n=13, 22.4%) and H4 (n=9, 15.5%).
194 The most common serotype was O18:H7:K1 (n=14, 24.1%); these isolates belonged to the

195 STc95 (nine ST95, four ST390 and one ST416). The second most common serotype was
196 O75:H5:K1 (n=8, 13.8%); six isolates from ST1193 possessed this serotype.

197

198 **NMEC can cause recrudescence invasive infection despite appropriate antibiotic**
199 **treatment**

200 During 2019 - 2020, three patients from which NMEC isolates were originally cultured
201 suffered recrudescence invasive infection (Figure 1; MS21522, MS21524 and MS22733),
202 providing an opportunity to compare the infecting isolates over time using whole genome
203 sequencing. In all cases, the infecting *E. coli* isolates were susceptible to the therapy, which
204 comprised cefotaxime (50mg/kg/dose 8 hourly), switched to ceftriaxone (100mg/kg/day) to
205 facilitate home parenteral antibiotic administration. Bacterial culture was performed from
206 blood, CSF, urine and/or stool during the infection period (Figure 2). These patients were
207 from different regions in Australia.

208

209 ***Patient 1***

210 Patient 1 (0-8 weeks of age) was admitted to the emergency department with fever,
211 respiratory distress and sepsis. The child was diagnosed with meningitis based on a
212 cerebrospinal fluid (CSF) pleocytosis (>2000 white blood cells; WBCs, low glucose, elevated
213 protein), positive CSF *E. coli* PCR and a positive blood culture for *E. coli* (MS21522). Two
214 weeks after completion of a 3-week course of appropriately dosed therapy with third-
215 generation cephalosporins as described above, the child developed similar symptoms of fever
216 and irritability. Lumbar puncture was performed and the CSF culture was positive for *E. coli*
217 (MS21576). Both the initial blood culture isolate and the relapse CSF isolate were non-
218 susceptible to ciprofloxacin and gentamycin, and whole genome sequencing revealed they
219 were identical (ST1193 O18:K1:H5; *fimH64*), with no single nucleotide polymorphisms

220 (SNPs) nor indels (Figure 2A). Unlike the typical ST1193 O75 serotype [23], this isolate
221 contained a unique O18 serotype. The isolate possessed mutations in *gyrA* (S83L D87N) and
222 *parC* (S80I), which explain its resistance to ciprofloxacin, as well as a multidrug resistance
223 IncF plasmid containing genes conferring resistance to aminoglycosides (*aac(3)-IId*, *aadA5*,
224 *aph(3'')-Ib* and *aph(6)-Id*), trimethoprim (*dfrA17*), sulphonamides (*sul1* and *sul2*) and
225 macrolides (*mphA*) (Supplementary Figure 2). Treatment of the relapse was extended to six
226 weeks of intravenous therapy. At follow-up, no anatomical or immunological abnormality
227 was identified and development is normal.

228

229 **Patient 2**

230 Patient 2 (0-8 weeks of age) presented to the emergency department with a febrile illness.
231 Blood and urine cultures on admission were positive for *E. coli*. CSF taken 24 hours after
232 treatment revealed pleocytosis (>300 WBCs, >95% polymorphs) but no bacteria were
233 cultured. The patient completed a 3-week course of appropriately dosed antibiotic therapy
234 with third-generation cephalosporins. In the 6-week period after discharge, the child had
235 several short admissions to hospital, but no infection was identified. At 11 weeks post initial
236 infection, the child was re-admitted to hospital with high fever. CSF cultures were negative
237 and microscopy was normal, but cultures from blood, urine and faeces were all positive for *E.*
238 *coli*. Whole genome sequencing revealed that all isolates belonged to ST537 O75:H5 (*fimH5*;
239 STc14). Pairwise comparison of the recrudescence isolates showed that the urine and fecal
240 isolates were identical to the original isolate, while the blood isolate contained one
241 nonsynonymous SNP in the *mdoH* gene encoding a glucan biosynthesis glucosyltransferase
242 (T1358G; V453G). This mutation is located in the large cytoplasmic domain of MdoH likely
243 involved in polymerisation of glucose from UDP glucose; the isolate exhibited a mucoid
244 colony morphology suggestive of increased colanic acid production. The isolates did not

245 possess plasmids nor antibiotic resistance genes. The infant experienced recurrent urinary
246 tract infections with *E. coli* and other urinary pathogens through infancy despite normal
247 urinary tract anatomy. At follow-up, no other history of invasive infection nor identified
248 immunodeficiency were noted, and the child was reported to be developing normally.

249

250 **Patient 3**

251 Patient 3 (0-8 weeks of age) was admitted to the paediatric intensive care unit with fever and
252 seizures. CSF and blood cultured a fully susceptible *E. coli*. Two weeks after completing a
253 four-week course of appropriate therapy with third-generation cephalosporins, the infant was
254 readmitted to hospital with fever and irritability, with further investigation identifying *E. coli*
255 in CSF, urine and blood. Three weeks after the completion of the six-week treatment course,
256 the infant experienced a second relapse, with *E. coli* isolated from both CSF and blood.
257 Whole genome sequencing revealed that all isolates were identical and belonged to ST131
258 O25b:K1:H4 (*fimH22*). These isolates contained a ColV-virulence plasmid, but did not
259 harbour acquired antibiotic resistance genes. The infant received a further 6-week course of
260 therapy. Extensive imaging studies including repeated MRI imaging of the head and spine,
261 CT imaging of the head and chest, ultrasound imaging of abdomen and pelvis, and nuclear
262 medicine imaging did not show a congenital malformation or abscess. Immunological work-
263 up did not show a known primary immunodeficiency. At follow-up, speech delay is reported
264 but no other developmental abnormality.

265

266 **The gut as a reservoir to seed recrudescence infection**

267 In all three patients that suffered NM and recrudescence invasive infection, the causative
268 isolates were susceptible to third-generation cephalosporins, suggesting the existence of a
269 persistent reservoir that could evade the cidal effect of antibiotic treatment and seed repeat

270 infection. Indeed, the fact that the causative *E. coli* isolate was detected from a fecal sample
271 at the time of the recrudescence infection in patient 2 (day 77 after initial admission), suggests
272 that NMEC could persist in the gut and cause repeat infection, an observation that has also
273 been reported for uropathogenic *E. coli* that cause recurrent urinary tract infection [25] and
274 acute pyelonephritis in infants [26]. Therefore, we retrospectively examined available stored
275 fecal samples from patient 2 at 8- and 12-week follow-up visits post recrudescence infection
276 (days 149 and 174 after initial admission) and patient 3 during treatment and at discharge
277 after the third episode (days 126 and 147 after initial admission) using short-read
278 metagenomic sequencing (Figure 3). Although no fecal samples were available for
279 comparative analysis from either patient prior to antibiotic treatment, we observed a low level
280 of diversity in the composition of the microbiome of both patients, consistent with severe
281 dysbiosis. The microbiome of patient 2 was dominated by *Enterobacter* (37.4% relative
282 abundance), *Achromobacter* (23.4% relative abundance) and *Bacteroides* (22.7% relative
283 abundance) genera at 8-weeks post recrudescence infection (day 149 after initial admission),
284 and by *Bacteroides* genera (75.8% relative abundance) at 12-weeks post recrudescence
285 infection (day 174 after initial admission). The relative abundance of *E. coli* was 2.05% and
286 4.1% in each of these samples, respectively, and further analysis using StrainGE [27] showed
287 that the isolates were most closely matched to the original causative MS21524 isolate. We
288 further employed complementary long-read metagenomic sequencing to analyse the 8-week
289 post relapse infection sample, which enabled construction of a complete *E. coli* genome that
290 was identical to the causative ST537 (*fimH5*) isolate (Figure 2, Figure 3; Supplementary
291 Table 3). In the 12-week post recrudescence infection fecal sample from patient 2, amplicon
292 sequencing targeting *fimH* identified the presence of *E. coli* with the same *fimH* type as the
293 causative isolate (*fimH5*). Thus, two independent analyses of samples taken 4 weeks apart
294 demonstrated the existence of the *E. coli* ST537 isolate in the intestinal microflora of patient

295 2. In patient 3, the microbiome was dominated by *Enterococcus* genera at both time points
296 examined (93% and 97.4% relative abundance). We were unable to detect *E. coli* by *fimH*
297 amplicon sequencing and the relative abundance of *E. coli* in these fecal samples was
298 extremely low (<0.01%) based on metagenomic sequencing (Supplementary Table 3). The
299 extensive dysbiosis revealed in this patient is likely an outcome of the three rounds of
300 antibiotic treatment.

301

302 **Discussion**

303 In this study, we present a genomic analysis of 58 NMEC isolates obtained over 46 years
304 spanning seven different geographic locations and reveal a dominance of ST95 and ST1193.
305 We also provide direct evidence to implicate the gut as a reservoir for recrudescence of invasive
306 infection in some patients despite appropriate antibiotic treatment.

307

308 The majority of the NMEC isolates in our study belonged to phylogroup B2 (82.8%), an
309 observation consistent with other reports [28, 29]. These isolates were predominantly from
310 two major STs, ST95 and ST1193. ST95 represents a major clonal lineage responsible for
311 urinary tract and bloodstream infections [30, 31], and were identified throughout the period
312 of investigation. This lineage was also previously shown to cause ~25% of NM cases in
313 France in the period 2004-2015 [11], demonstrating its enhanced capacity to cause
314 disseminated infection in newborns. ST1193, on the other hand, was first identified in 2012
315 [32], and is the second most common fluoroquinolone-resistant *E. coli* lineage after ST131
316 [23, 33]. ST1193 causing NM was first reported in the USA in 2016 [34]. Here, ST1193
317 accounted for 15.5% of NMEC isolates, all of which were obtained from 2013, and was the
318 dominant lineage since this time. This is consistent with a report in China that showed
319 ST1193 was the most common NMEC (21.4%), followed by ST95 (17.9%), between 2009-
320 2015 [35]. Concerningly, the ST1193 isolates examined here carry genes encoding several
321 aminoglycoside-modifying enzymes, generating a resistance profile that may lead to the
322 clinical failure of empiric regimens such as ampicillin and gentamicin, a therapeutic
323 combination used in many settings to treat NM and early-onset sepsis [36, 37]. This, in
324 combination with reports of co-resistance to third-generation cephalosporins for some
325 ST1193 isolates [23, 35], would limit the choice of antibiotic treatments. The dominance of
326 both ST95 and ST1193 in our collection is notable since other widespread *E. coli* phylogroup

327 B2 lineages such as ST131, ST73, ST69 and ST12 do not cause similar rates of NM disease.
328 We speculate this is due to the prevailing K1 polysialic acid capsule serotype found in ST95
329 and the newly emerged ST1193 clone [23, 38] in combination with other virulence factors
330 [15, 28, 29] (Figure 4) and the immature immune system of preterm infants. Understanding
331 the risk of these clones, as well as perinatal transmission and antibiotic resistance patterns,
332 may inform the appropriateness of interventions such as maternal screening or antimicrobial
333 prophylaxis.

334

335 Although reported rarely, recrudescence of invasive *E. coli* infection in NM patients, including
336 several infants born pre-term, has been documented in single study reports [39, 40]. In these
337 reports, infants received appropriate antibiotic treatment based on antibiogram profiling and
338 no clear clinical risk factors to explain recrudescence were identified, highlighting our limited
339 understanding of NM aetiology. Here, we tracked NMEC recurrence using whole genome
340 sequencing in three patients that suffered NM and recrudescence of invasive infection, and
341 demonstrated that the isolate causing recrudescence was the same as the original causative
342 isolate and susceptible to the initial antibiotic therapy. In one patient (patient 2), we identified
343 the causative isolate in the stool at days 77, 149 and 174 after initial detection in the
344 bloodstream, providing direct evidence of persistence in the gut, and implicating this site as a
345 reservoir to seed repeat infection. This isolate belonged to ST537 (serotype O75:H5) and is
346 from the same clonal complex as ST1193 (i.e. STc14).

347

348 This study had several limitations. First, our NMEC collection was restricted to seven
349 geographic regions, a reflection of the difficulty in acquiring isolates causing this disease.
350 Second, we did not have access to a complete set of stool samples spanning pre- and post-
351 treatment in the patients that suffered NM and recrudescence of invasive infection. This impacted

352 our capacity to monitor *E. coli* persistence and evaluate the effect of antibiotic treatment on
353 changes in the microbiome over time. Third, we did not have access to urine or stool samples
354 from the mother of the infants that suffered recrudescence, and thus cannot rule out mother-
355 to-child transmission as a mechanism of reinfection. Finally, we did not have clinical data on
356 the weeks of gestation for all patients, and thus could not compare virulence factors from
357 NMEC isolated from preterm versus term infants. Regardless, our study describes the
358 genomic diversity of NMEC, highlighting ST95 and ST1193 as the most important clonal
359 lineages associated with this devastating disease. Although antibiotics are the standard of care
360 for NMEC treatment, we show that even when appropriate antibiotics are used, in some cases
361 they do not eliminate the causative NMEC that resides intestinally. Together with associated
362 antibiotic-driven dysbiosis, this reveals a need to consider diagnostic and therapeutic
363 interventions to mitigate the risk of recrudescence infection.

364

365

366 **Methods**

367 **Ethics statement**

368 The study received ethical approval from the Children's Health Queensland Human Research
369 Ethics Committee (LNR/18/QCHQ/45045). Precise patient details have been removed for
370 ethics compliance; additional de-identified details are available from the corresponding
371 author upon request.

372

373 **Bacterial isolates**

374 A collection of 52 NMEC isolates obtained from 1974 to 2020 was achieved from Sweden,
375 Finland, Cambodia and Australia. Isolates were stored in glycerol at -80°C until use. All
376 isolates were cultured in Lysogeny broth. The collection comprised 42 isolates from

377 confirmed meningitis cases (29 cultured from CSF and 13 cultured from blood) and 10
378 isolates from clinically diagnosed meningitis cases (all cultured from blood) (Table S1). This
379 collection was complemented by the addition of six completely sequenced NMEC genomes
380 available on the NCBI database, namely strains IHE3034, RS218, S88, NMEC58, MCJCHV-
381 1 and CE10.

382

383 **DNA extraction, genome sequencing and analyses**

384 Genome sequencing was performed using paired-end Illumina methodology. Illumina
385 sequencing data were processed by removing adapters and low-quality reads using
386 Trimmomatic v0.36 [41], with a minimum quality score of 10 and minimum read length of
387 50. Trimmed reads were *de novo* assembled using SPAdes v3.12.0 [42] with default
388 parameters. Draft assemblies of the 52 NMEC isolates from this study, together with six
389 complete NMEC genomes and eight complete genomes from other characterised *E. coli*
390 representing different phylogroups, were subjected to phylogenetic analysis using parsnp
391 v1.5.3 [43]. A subset of 18 isolates were additionally sequenced using Oxford Nanopore
392 Technology long-read sequencing (Nanopore). Complete NMEC genomes were achieved
393 using a combination of Illumina short-read and Nanopore long-read data and analysis
394 employing the MicroPIPE tool [44].

395

396 ***In silico* and molecular analyses**

397 Virulence-associated genes, antibiotic resistance genes, plasmids and serotyping were
398 evaluated using ABRicate (<https://github.com/tseemann/abricate>) with built-in databases [45-
399 48], with the percentage nucleotide identity and coverage cut-off set at 90% and 80%,
400 respectively. Capsule typing was performed employing Kaptive [49] using an in-house *E.*
401 *coli* capsule database [38] and manually checked. Chromosomal point mutations associated

402 with antibiotic resistance were detected using PointFinder [50]. FimH amplicon sequencing
403 was performed as previously described [51, 52]; allelic variants were identified using
404 FimTyper [53].

405

406 **K1 ELISA**

407 K1 capsule expression was detected by ELISA using an anti-polysialic acid antibody single
408 chain Fv fragment [54] as the primary antibody, anti-His antibody and alkaline phosphatase
409 anti-mouse IgG as the secondary and tertiary antibodies, respectively; p-
410 nitrophenylphosphate (Sigma) was used as the substrate. Optical density was measured at 420
411 nm.

412

413 **Metagenomic sequencing and analyses**

414 Metagenomic sequencing was performed on DNA extracted from fecal samples using the
415 Illumina NovaSeq6000 platform. Adapters and low-quality reads were trimmed using
416 Trimmomatic v0.36 [41], employing a minimum quality score of 10 and minimum read
417 length of 50. Sequencing reads corresponding to human DNA were discarded by mapping the
418 trimmed reads to the human genome hg38 (accession number GCA_000001405.29) using
419 bowtie2 [55]. Taxonomical profiling was performed with Kraken2 [56] followed by Bracken
420 [57].

421

422 **Long-read metagenomic sequencing**

423 Long-read metagenomic sequencing was performed on DNA extracted from a fecal sample.
424 A HiFi gDNA library was prepared using the SMRTbell Express Template Prep Kit 2.0
425 (PacBio, 100-938-900) according to the low input protocol (PacBio, PN 101-730-400
426 Version 06 [June 2020]). As the sample DNA was already fragmented with a tight peak

427 (mode size 9.4 kb), no shearing was performed; the sample was concentrated using Ampure
428 PB beads (PacBio, PCB-100-265-900) and used directly as input into library preparation.
429 The entire quantity of purified DNA (360ng) was used to make the library as follows. The
430 DNA was treated to remove single-stranded overhangs, followed by a DNA damage repair
431 reaction and an end-repair/A-tailing reaction. Overhang barcoded adapters were ligated to the
432 A-tailed library fragments, followed by a nuclease treatment to remove damaged library
433 fragments, and then purification with AMPure PB beads. The library was size-selected to
434 remove fragments <3kb using AMPure PB beads. The final purified, size-selected library was
435 quantified on the Qubit fluorometer using the Qubit dsDNA HS assay kit (Invitrogen,
436 Q32854) to assess concentration, and run on the Agilent Femto Pulse using the 55 kb BAC
437 Analysis Kit (Agilent, FP-1003-0275) to assess fragment size distribution.

438

439 Sequencing was performed using the PacBio Sequel II (software/chemistry v10.1). The
440 library pool was prepared for sequencing according to the SMRT Link (v10.1) sample setup
441 calculator, following the standard protocol for Diffusion loading with Ampure PB bead
442 purification, using Sequencing Primer v5, Sequel II Binding Kit v2.2 and the Sequel II DNA
443 Internal Control v1. Adaptive loading was utilised, with nominated on-plate loading
444 concentration of 80pM. The polymerase-bound library was sequenced on 1 SMRT Cell with
445 a 30-hour movie time plus a 2-hour pre-extension using the Sequel II Sequencing 2.0 Kit
446 (PacBio, 101-820-200) and SMRT Cell 8M (PacBio, 101-389-001).

447

448 After sequencing, the data was processed to generate CCS reads and demultiplex samples
449 using the default settings of the CCS with Demultiplexing application in SMRT Link (v10.1).
450 The demultiplexed reads were assembled *de novo* using Hifiasm [58]. Assembled contigs
451 were subject to taxonomic profiling using kraken2 [56] and fastANI [59]. Contigs

452 taxonomically assigned as *Escherichia coli* were subjected to *in silico* sequence typing using
453 MLST (<https://github.com/tseemann/mlst>) and mlst profiles from PubMLST [60].

454

455 **Author Contributions:** Conceptualization, MAS, AB, BW-W and ADI.; Investigation,
456 NTKN, M-DP, SJH, KMP, LA-F, BMF, SBA, and TM.; Formal analysis, NTKN, M-DP,
457 ADI and MAS., Resources, PNAH, SAB, SS, HB, PT, AB, BW-W, ADI and MAS.; Writing
458 – Original Draft, NTKN, ADI and MAS.; Writing – Review & Editing, All Authors.;
459 Supervision, ADI and MAS.; Funding Acquisition, NTKN, M-DP, PNAH, SAB, SS, PT,
460 ADI, MAS.

461

462 **Declaration of interests:** The authors declare that they have no conflict of interest.

463

464 **Data sharing:** Genome sequence data have been deposited in the Sequence Read Archive
465 under the BioProjects PRJNA757133 and PRJNA893826. Sample accession numbers are
466 listed in Supplementary Table 4.

467

468 **Acknowledgments**

469 The authors would like to thank Michelle Bauer for technical expertise and the laboratories
470 contributing the isolates, Pathology Queensland and Mater Pathology. At the time of the
471 study SS was affiliated with Mater Pathology, South Brisbane, Australia. This work was
472 supported by project grants APP1181958 and APP2001431 (to MAS, MDP and NTKN) and
473 an Investigator grant GNT1197743 (to ADI) from the National Health and Medical Research
474 Council of Australia (NHMRC), a Children’s Hospital Foundation Innovator grant (50270, to
475 MAS, ADI, PNAH, SAB and SS), an Australian Infectious Diseases Research Centre grant

476 (to MAS, ADI and NTKN), a grant from the Genome Innovation Hub at the University of
477 Queensland, and a grant from the Wellcome Trust to PT (Grant number 220211). For the
478 purpose of open access, the author (PT) has applied a CC BY public copyright licence to any
479 Author Accepted Manuscript version arising from this submission.

480

481

482 **References**

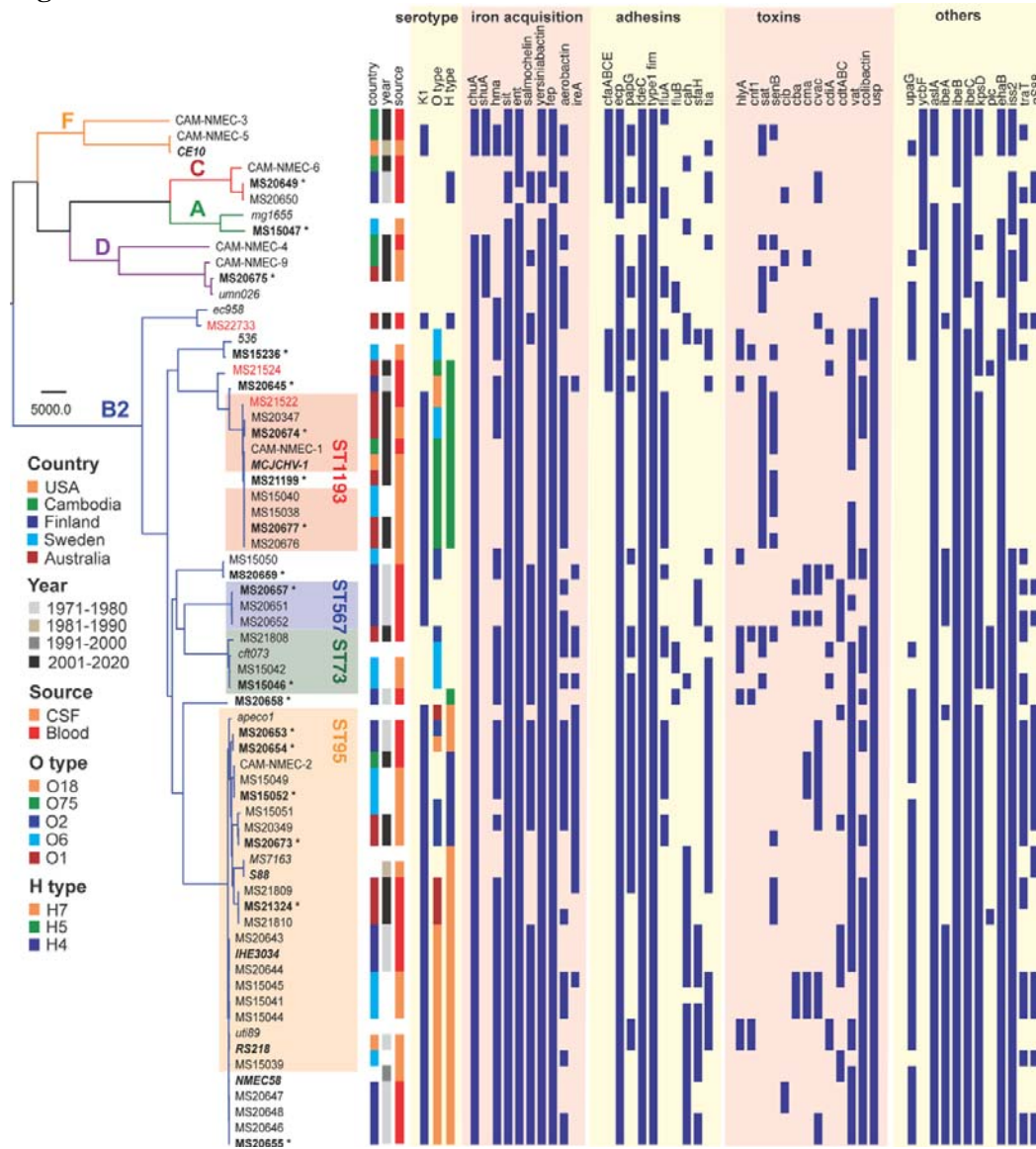
- 483 1. Doctor BA, Newman N, Minich NM, Taylor HG, Fanaroff AA, Hack M. Clinical
484 outcomes of neonatal meningitis in very-low birth-weight infants. *Clin Pediatr (Phila)*
485 **2001**; 40(9): 473-80.
- 486 2. Stevens JP, Eames M, Kent A, Halket S, Holt D, Harvey D. Long term outcome of
487 neonatal meningitis. *Arch Dis Child Fetal Neonatal Ed* **2003**; 88(3): F179-84.
- 488 3. Harvey D, Holt DE, Bedford H. Bacterial meningitis in the newborn: a prospective
489 study of mortality and morbidity. *Semin Perinatol* **1999**; 23(3): 218-25.
- 490 4. Ouchenir L, Renaud C, Khan S, et al. The Epidemiology, Management, and Outcomes
491 of Bacterial Meningitis in Infants. *Pediatrics* **2017**; 140(1).
- 492 5. Gaschignard J, Levy C, Romain O, et al. Neonatal Bacterial Meningitis: 444 Cases in 7
493 Years. *Pediatr Infect Dis J* **2011**; 30(3): 212-7.
- 494 6. Basmaci R, Bonacorsi S, Bidet P, et al. Escherichia Coli Meningitis Features in 325
495 Children From 2001 to 2013 in France. *Clin Infect Dis* **2015**; 61(5): 779-86.
- 496 7. May M, Daley AJ, Donath S, Isaacs D, Australasian Study Group for Neonatal I. Early
497 onset neonatal meningitis in Australia and New Zealand, 1992-2002. *Arch Dis Child*
498 *Fetal Neonatal Ed* **2005**; 90(4): F324-7.
- 499 8. Holt DE, Halket S, de Louvois J, Harvey D. Neonatal meningitis in England and Wales:
500 10 years on. *Arch Dis Child Fetal Neonatal Ed* **2001**; 84(2): F85-9.
- 501 9. van der Flier M. Neonatal meningitis: small babies, big problem. *Lancet Child Adolesc*
502 *Health* **2021**; 5(6): 386-7.
- 503 10. Anderson SG, Gilbert GL. Neonatal gram negative meningitis: a 10-year review, with
504 reference to outcome and relapse of infection. *J Paediatr Child Health* **1990**; 26(4):
505 212-6.
- 506 11. Geslain G, Birgy A, Adiba S, et al. Genome sequencing of strains of the most
507 prevalent clonal group of O1:K1:H7 Escherichia coli that causes neonatal meningitis
508 in France. *BMC Microbiol* **2019**; 19(1): 17.
- 509 12. Sarff LD, McCracken GH, Schiffer MS, et al. Epidemiology of Escherichia coli K1 in
510 healthy and diseased newborns. *Lancet* **1975**; 1(7916): 1099-104.
- 511 13. Plainvert C, Bidet P, Peigne C, et al. A new O-antigen gene cluster has a key role in
512 the virulence of the Escherichia coli meningitis clone O45:K1:H7. *J Bacteriol* **2007**;
513 189.
- 514 14. Bidet P, Mahjoub-Messai F, Blanco J, et al. Combined multilocus sequence typing and
515 O serogrouping distinguishes Escherichia coli subtypes associated with infant
516 urosepsis and/or meningitis. *J Infect Dis* **2007**; 196(2): 297-303.
- 517 15. Johnson JR, Oswald E, O'Bryan TT, Kuskowski MA, Spanjaard L. Phylogenetic
518 distribution of virulence-associated genes among Escherichia coli isolates associated
519 with neonatal bacterial meningitis in the Netherlands. *J Infect Dis* **2002**; 185(6): 774-
520 84.
- 521 16. Robbins JB, McCracken GH, Jr., Gotschlich EC, Orskov F, Orskov I, Hanson LA.
522 Escherichia coli K1 capsular polysaccharide associated with neonatal meningitis. *N*
523 *Engl J Med* **1974**; 290(22): 1216-20.
- 524 17. Prasadarao NV, Wass CA, Hacker J, Jann K, Kim KS. Adhesion of S-fimbriated
525 Escherichia coli to brain glycolipids mediated by sfaA gene-encoded protein of S-
526 fimbriae. *J Biol Chem* **1993**; 268(14): 10356-63.

- 527 18. Prasadarao NV, Wass CA, Weiser JN, Stins MF, Huang SH, Kim KS. Outer membrane
528 protein A of *Escherichia coli* contributes to invasion of brain microvascular
529 endothelial cells. *Infect Immun* **1996**; 64(1): 146-53.
- 530 19. Huang SH, Chen YH, Kong G, et al. A novel genetic island of meningitic *Escherichia*
531 *coli* K1 containing the *ibeA* invasion gene (*GimA*): functional annotation and carbon-
532 source-regulated invasion of human brain microvascular endothelial cells. *Funct*
533 *Integr Genomics* **2001**; 1(5): 312-22.
- 534 20. Wang MH, Kim KS. Cytotoxic necrotizing factor 1 contributes to *Escherichia coli*
535 meningitis. *Toxins (Basel)* **2013**; 5(11): 2270-80.
- 536 21. Peigne C, Bidet P, Mahjoub-Messai F, et al. The plasmid of *Escherichia coli* strain S88
537 (O45:K1:H7) that causes neonatal meningitis is closely related to avian pathogenic *E.*
538 *coli* plasmids and is associated with high-level bacteremia in a neonatal rat
539 meningitis model. *Infect Immun* **2009**; 77(6): 2272-84.
- 540 22. Zhou Z, Alikhan NF, Mohamed K, Fan Y, Agama Study G, Achtman M. The Enterobase
541 user's guide, with case studies on *Salmonella* transmissions, *Yersinia pestis*
542 phylogeny, and *Escherichia* core genomic diversity. *Genome Res* **2020**; 30(1): 138-52.
- 543 23. Johnson TJ, Elnekave E, Miller EA, et al. Phylogenomic Analysis of Extraintestinal
544 Pathogenic *Escherichia coli* Sequence Type 1193, an Emerging Multidrug-Resistant
545 Clonal Group. *Antimicrob Agents Chemother* **2019**; 63(1).
- 546 24. Reed TAN, Krang S, Miliya T, et al. Antimicrobial resistance in Cambodia: a review. *Int*
547 *J Infect Dis* **2019**; 85: 98-107.
- 548 25. Forde BM, Roberts LW, Phan MD, et al. Population dynamics of an *Escherichia coli*
549 ST131 lineage during recurrent urinary tract infection. *Nat Commun* **2019**; 10(1):
550 3643.
- 551 26. Tullus K, Horlin K, Svenson SB, Kallenius G. Epidemic outbreaks of acute
552 pyelonephritis caused by nosocomial spread of P fimbriated *Escherichia coli* in
553 children. *J Infect Dis* **1984**; 150(5): 728-36.
- 554 27. van Dijk LR, Walker BJ, Straub TJ, et al. StrainGE: a toolkit to track and characterize
555 low-abundance strains in complex microbial communities. *Genome Biol* **2022**; 23(1):
556 74.
- 557 28. Wijetunge DS, Gongati S, DebRoy C, et al. Characterizing the pathotype of neonatal
558 meningitis causing *Escherichia coli* (NMEC). *BMC Microbiol* **2015**; 15: 211.
- 559 29. Bidet P, Mahjoub-Messai F, Blanco J, et al. Combined Multilocus Sequence Typing
560 and O Serogrouping Distinguishes *Escherichia coli* Subtypes Associated with Infant
561 Urosepsis and/or Meningitis. *J Infect Dis* **2007**; 196.
- 562 30. Manges AR, Geum HM, Guo A, Edens TJ, Fibke CD, Pitout JDD. Global Extraintestinal
563 Pathogenic *Escherichia coli* (ExPEC) Lineages. *Clin Microbiol Rev* **2019**; 32(3).
- 564 31. Kallonen T, Brodrick HJ, Harris SR, et al. Systematic longitudinal survey of invasive
565 *Escherichia coli* in England demonstrates a stable population structure only
566 transiently disturbed by the emergence of ST131. *Genome Res* **2017**.
- 567 32. Platell JL, Trott DJ, Johnson JR, et al. Prominence of an O75 clonal group (clonal
568 complex 14) among non-ST131 fluoroquinolone-resistant *Escherichia coli* causing
569 extraintestinal infections in humans and dogs in Australia. *Antimicrob Agents*
570 *Chemother* **2012**; 56(7): 3898-904.
- 571 33. Tchesnokova VL, Rechkina E, Larson L, et al. Rapid and Extensive Expansion in the
572 United States of a New Multidrug-resistant *Escherichia coli* Clonal Group, Sequence
573 Type 1193. *Clin Infect Dis* **2019**; 68(2): 334-7.

- 574 34. Nielsen DW, Ricker N, Barbieri NL, et al. Complete Genome Sequence of the
575 Multidrug-Resistant Neonatal Meningitis Escherichia coli Serotype O75:H5:K1 Strain
576 mcjchv-1 (NMEC-O75). *Microbiol Resour Announc* **2018**; 7(10).
- 577 35. Ding Y, Zhang J, Yao K, Gao W, Wang Y. Molecular characteristics of the new
578 emerging global clone ST1193 among clinical isolates of Escherichia coli from
579 neonatal invasive infections in China. *Eur J Clin Microbiol Infect Dis* **2021**; 40(4): 833-
580 40.
- 581 36. Fleiss N, Schwabenbauer K, Randis TM, Polin RA. What's new in the management of
582 neonatal early-onset sepsis? *Arch Dis Child Fetal Neonatal Ed* **2023**; 108(1): 10-4.
- 583 37. Fuchs A, Bielicki J, Mathur S, Sharland M, Van Den Anker JN. Reviewing the WHO
584 guidelines for antibiotic use for sepsis in neonates and children. *Paediatr Int Child*
585 *Health* **2018**; 38(sup1): S3-S15.
- 586 38. Goh KGK, Phan MD, Forde BM, et al. Genome-Wide Discovery of Genes Required for
587 Capsule Production by Uropathogenic Escherichia coli. *mBio* **2017**; 8(5).
- 588 39. Vissing NH, Monster MB, Nordly S, et al. Relapse of Neonatal Escherichia coli
589 Meningitis: Did We Miss Something at First? *Children (Basel)* **2021**; 8(2).
- 590 40. Bingen E, Cave H, Aujard Y, et al. Molecular analysis of multiply recurrent meningitis
591 due to Escherichia coli K1 in an infant. *Clin Infect Dis* **1993**; 16(1): 82-5.
- 592 41. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina
593 sequence data. *Bioinformatics* **2014**; 30(15): 2114-20.
- 594 42. Bankevich A, Nurk S, Antipov D, et al. SPAdes: a new genome assembly algorithm
595 and its applications to single-cell sequencing. *J Comput Biol* **2012**; 19(5): 455-77.
- 596 43. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-
597 genome alignment and visualization of thousands of intraspecific microbial genomes.
598 *Genome Biol* **2014**; 15(11): 524.
- 599 44. Murigneux V, Roberts LW, Forde BM, et al. MicroPIPE: validating an end-to-end
600 workflow for high-quality complete bacterial genome construction. *BMC Genomics*
601 **2021**; 22(1): 474.
- 602 45. Chen L, Zheng D, Liu B, Yang J, Jin Q. VFDB 2016: hierarchical and refined dataset for
603 big data analysis--10 years on. *Nucleic Acids Res* **2016**; 44(D1): D694-7.
- 604 46. Feldgarden M, Brover V, Haft DH, et al. Using the NCBI AMRFinder Tool to Determine
605 Antimicrobial Resistance Genotype-Phenotype Correlations Within a Collection of
606 NARMS Isolates. **2019**: 550707.
- 607 47. Carattoli A, Zankari E, Garcia-Fernandez A, et al. In silico detection and typing of
608 plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob*
609 *Agents Chemother* **2014**; 58(7): 3895-903.
- 610 48. Ingle DJ, Valcanis M, Kuzevski A, et al. In silico serotyping of E. coli from short read
611 data identifies limited novel O-loci but extensive diversity of O:H serotype
612 combinations within and between pathogenic lineages. *Microb Genom* **2016**; 2(7):
613 e000064.
- 614 49. Wyres KL, Wick RR, Gorrie C, et al. Identification of Klebsiella capsule synthesis loci
615 from whole genome data. *Microb Genom* **2016**; 2(12): e000102.
- 616 50. Zankari E, Allesoe R, Joensen KG, Cavaco LM, Lund O, Aarestrup FM. PointFinder: a
617 novel web tool for WGS-based detection of antimicrobial resistance associated with
618 chromosomal point mutations in bacterial pathogens. *J Antimicrob Chemother* **2017**;
619 72(10): 2764-8.

- 620 51. Willner D, Low S, Steen JA, et al. Single clinical isolates from acute uncomplicated
621 urinary tract infections are representative of dominant in situ populations. *mBio*
622 **2014**; 5(2): e01064-13.
- 623 52. Chen Z, Phan MD, Bates LJ, et al. The urinary microbiome in patients with refractory
624 urge incontinence and recurrent urinary tract infection. *Int Urogynecol J* **2018**;
625 29(12): 1775-82.
- 626 53. Roer L, Tchesnokova V, Allesoe R, et al. Development of a Web Tool for *Escherichia*
627 *coli* Subtyping Based on fimH Alleles. *J Clin Microbiol* **2017**; 55(8): 2538-43.
- 628 54. Nagae M, Ikeda A, Hane M, et al. Crystal structure of anti-polysialic acid antibody
629 single chain Fv fragment complexed with octasialic acid: insight into the binding
630 preference for polysialic acid. *J Biol Chem* **2013**; 288(47): 33784-96.
- 631 55. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*
632 **2012**; 9(4): 357-9.
- 633 56. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2.
634 *Genome Biol* **2019**; 20(1): 257.
- 635 57. Jennifer Lu FPB, Peter Thielen, Steven L. Salzberg. Bracken: estimating species
636 abundance in metagenomics data. *PeerJ Computer Science* **2017**; 3(e104).
- 637 58. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo
638 assembly using phased assembly graphs with hifiasm. *Nat Methods* **2021**; 18(2): 170-
639 5.
- 640 59. Jain C, Rodriguez RL, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI
641 analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun*
642 **2018**; 9(1): 5114.
- 643 60. Jolley KA, Maiden MC. BIGSdb: Scalable analysis of bacterial genome variation at the
644 population level. *BMC Bioinformatics* **2010**; 11: 595.
- 645
646

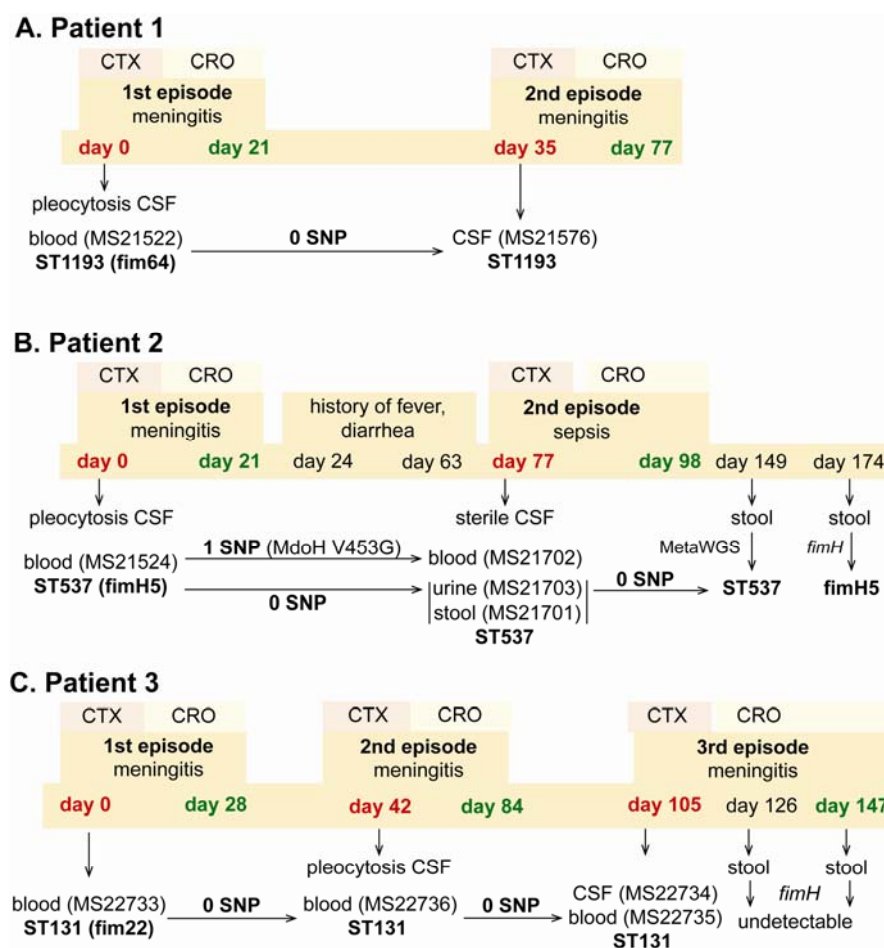
647 **Figures**



648
649

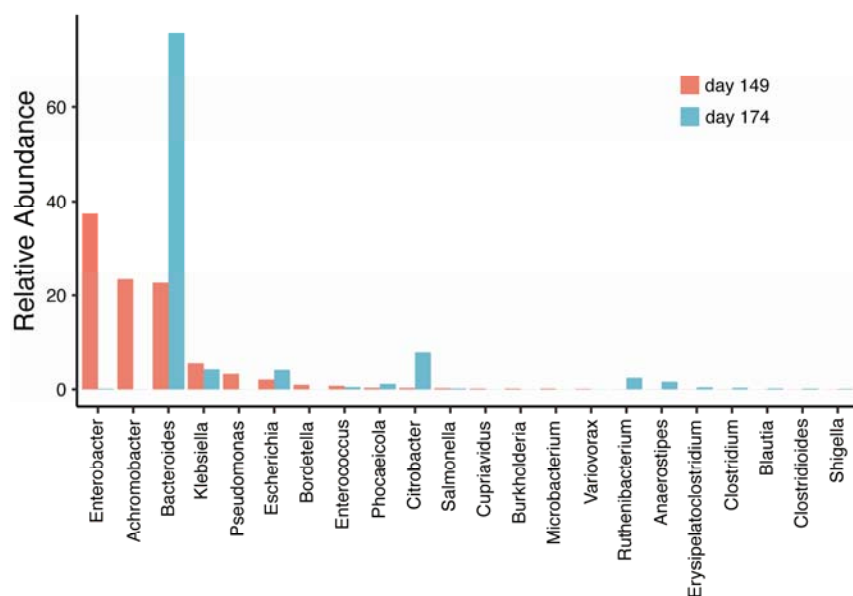
650 **Figure 1.** Maximum likelihood phylogram displaying the relationship of the NMEC isolates with
 651 their associated serotype and virulence factor profile. Non-NMEC isolates used in the analysis for
 652 referencing are italicized. The phylogram was built and recombination regions removed employing
 653 Parsnp, using 185,911 core single nucleotide polymorphisms (SNPs) and NMEC strain IHE3034 as
 654 the reference. The scale bar indicates branch lengths in numbers of SNPs. NMEC isolates with
 655 available complete genomes are bold-italicized, while NMEC isolates that were completely sequenced
 656 in this study are indicated in bold and marked with an asterisk. The NMEC isolates that caused
 657 recrudescence in this study are indicated in red. Branches are colored according to
 658 phylogroups: orange, phylogroup F; red, phylogroup C; green, phylogroup A; violet, phylogroup D
 659 and blue, phylogroup B2. The presence of specific virulence factors is indicated in dark blue. The
 660 phylogeny can be viewed interactively at <http://microreact.org/project/oNfA4v16h3tQbqREoYtCXj->

661 [high-risk-escherichia-coli-clones-that-cause-neonatal-meningitis-and-association-with-recrudescence-](#)
 662 [infection.](#)
 663

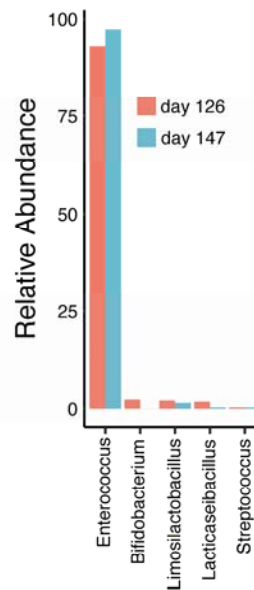


664
 665 **Figure 2.** Infection and treatment profile of patients suffering from NM and recrudescence of invasive
 666 infection. Indicated is the hospital admission history of patients, together with the timeline of sample
 667 collection, identified *E. coli* isolates and their infection source, and isolate identification based on
 668 whole genome sequencing, metagenomic sequencing (MetaWGS) or *fimH* amplicon sequencing.
 669 Genomic relatedness is indicated based on the number of single nucleotide polymorphisms (SNPs).
 670 The time of admission for the initial episode is indicated as Day 0, with subsequent timepoints
 671 indicated as days post initial admission. Admission and discharge days are indicated in red and green,
 672 respectively.
 673

A. Patient 2



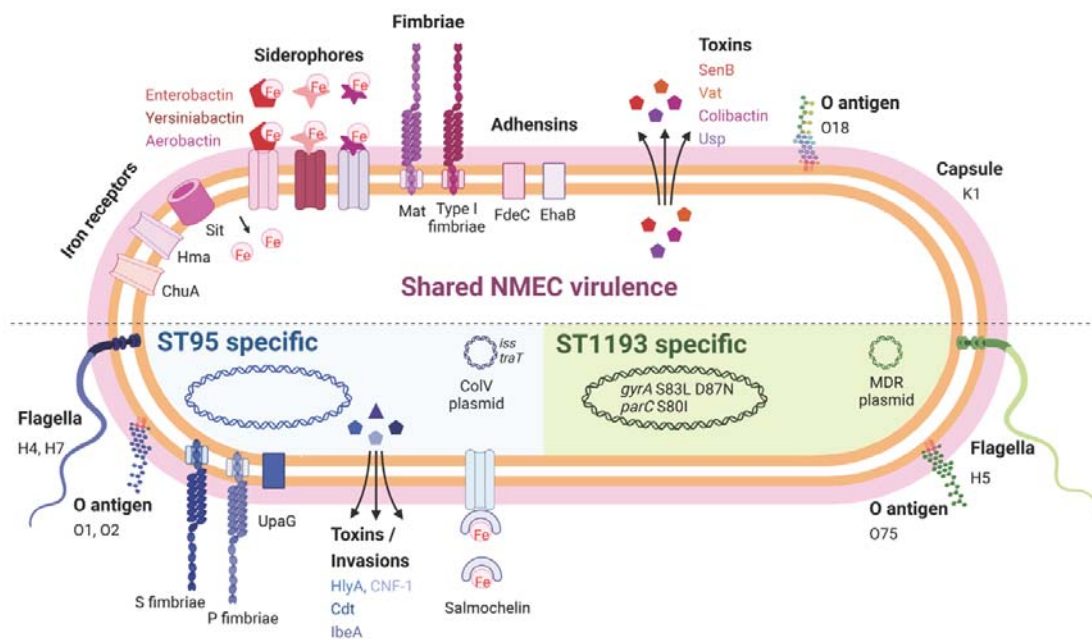
B. Patient 3



674

675 **Figure 3.** Relative abundance of bacterial genera ($\geq 0.01\%$) in the gut microbiome of patient 2 at 8-
 676 and 12-week follow-up post relapsed infection (days 149 and 174 after initial admission) (A) and
 677 patient 3 during treatment and at discharge after the third episode (days 126 and 147 after initial
 678 admission) (B).

679



680

681 **Figure 4.** Summary of key NMEC virulence genes based on genome profiling performed in this
 682 study. Shown are shared virulence genes common to most NMEC, as well as ST95- and ST1193-
 683 specific NMEC virulence genes. Figure created with BioRender.com.

684

685 **Supplementary Figure Legends**

686

687 **Figure S1.** Number of human-derived *E. coli* strains from ST95, ST1193, ST38, ST131, ST73, ST10
688 and ST69 available in the Enterobase database. Strains were stratified based on their year of isolation,
689 spanning the periods before 2000, 2001-2005, 2006-2010, 2011-2015 and 2016-2022.

690

691 **Figure S2.** Antibiotic resistance gene profile of NMEC strains in the collection. The presence of each
692 resistance gene is denoted by black shading.

693

694 **Figure S3.** ST95 NMEC strains contain more virulence factors than ST1193 NMEC strains. (A) The
695 number of virulence genes (grouped as in Fig.1) for each strain within each ST. (B) The number of
696 virulence genes grouped by their functions in ST95 versus ST1193 strains. P-value were calculated
697 using Mann-Whitney two-tailed unpaired test.

698

699 **Figure S4.** K1 capsule production in NMEC. K1 capsule production was detected by ELISA using a
700 monoclonal antibody specific for polysialic acid. Strains with an $OD_{420} > 0.133$ (mean + 3 standard
701 deviations of a negative control *kpsD* mutant; dashed line) were considered positive for K1 capsule
702 production. Data points represent independent biological replicates with horizontal lines as the mean.

703

704 **Supplementary Tables**

705

706 **Supplementary Table 1.** Isolates used in this study.

707

708 **Supplementary Table 2.** Completely sequenced NMEC isolates.

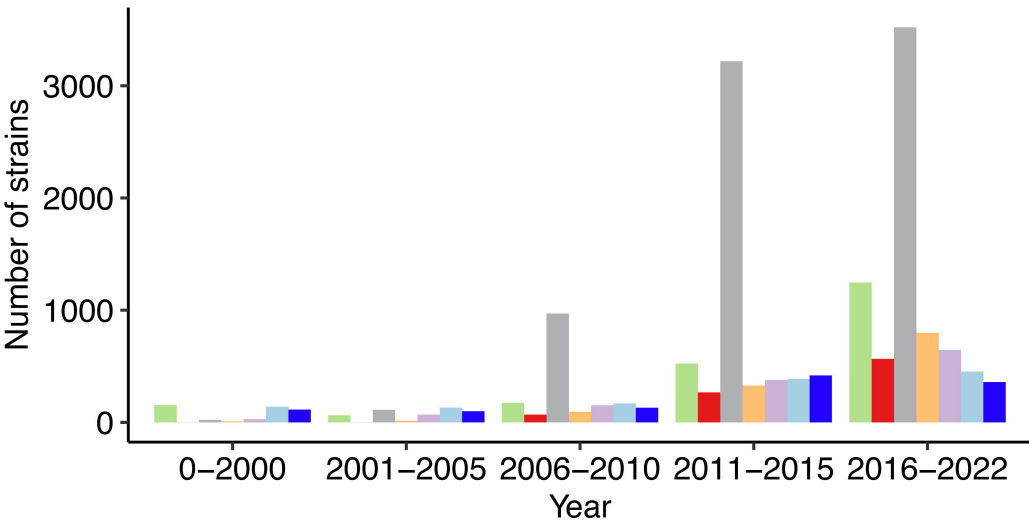
709

710 **Supplementary Table 3.** Metagenomic sequence analysis.

711

712 **Supplementary Table 4.** Accession numbers of strains sequenced in the study

ST 10 1193 131 38 69 73 95



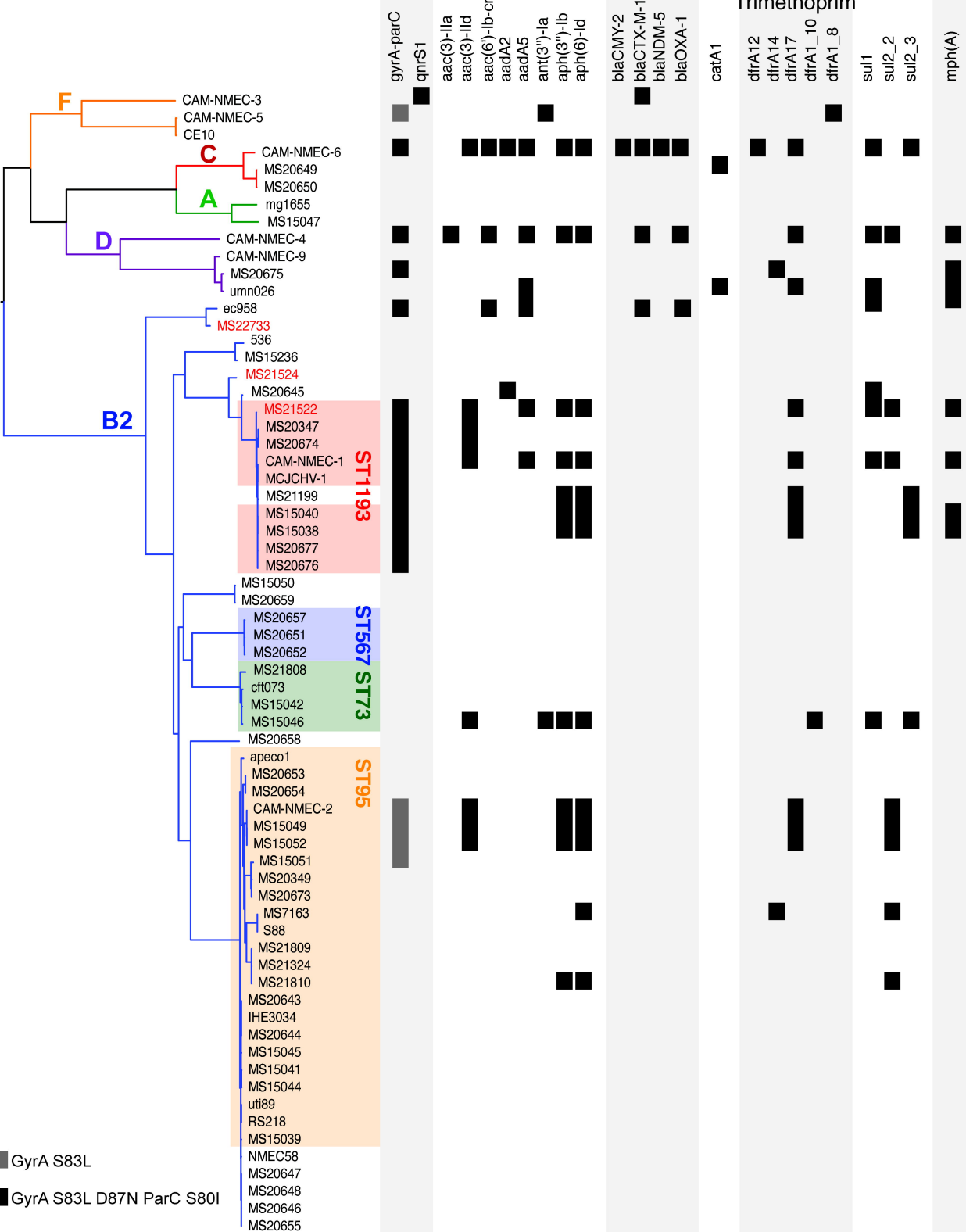
(fluor)quinolones

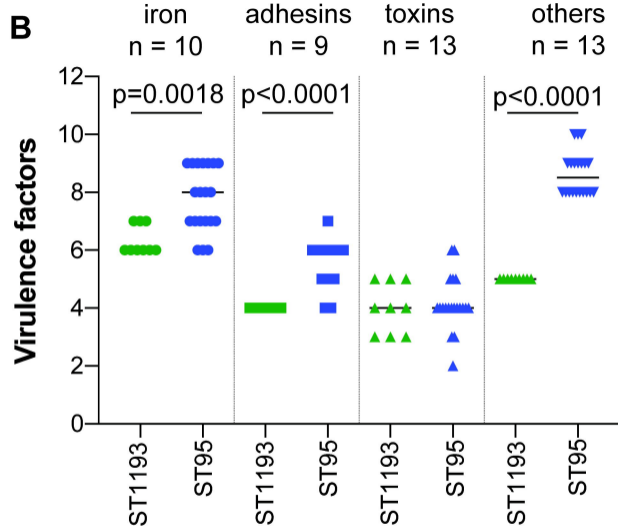
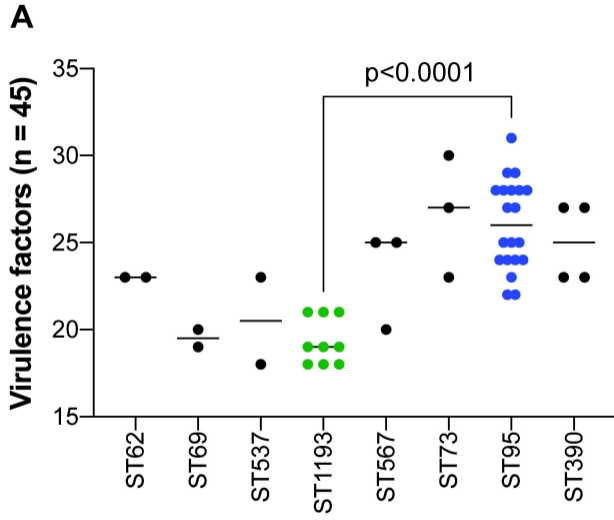
Aminoglycosides

beta-lactam

Trimethoprim

Sulfonamides





OD420

