JB Accepted Manuscript Posted Online 16 January 2018
J. Bacteriol. doi:10.1128/JB.00698-17

Genome-Wide Identification by Transposon Insertion Sequencing of *Escherichia coli* K1 Genes Essential for *in vitro* Growth, Gastrointestinal Colonizing Capacity and Survival in Serum

Alex J. McCarthy[1*], Richard A. Stabler[2], Peter W. Taylor[1]

1 School of Pharmacy, University College London, London, WC1N 1AX, UK

2 London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK

[*]Correspondence to Peter W. Taylor, peter.taylor@ucl.ac.uk

Short title: Essential *E. coli* K1 genes

*Current address: Bacterial Infections and Immunity, Universitair Medisch Centrum Utrecht, 3584 CX Utrecht, The Netherlands

Word count: Abstract 235; text 8008

ABSTRACT *Escherichia coli* K1 strains are major causative agents of invasive disease of the new born. The age dependency of infection can be reproduced in the neonatal rat. Colonization of the small intestine following oral administration of K1 bacteria leads rapidly to invasion of the blood circulation; bacteria that avoid capture by the mesenteric lymphatic system and evade antibacterial mechanisms in the blood may disseminate to cause organ-specific infections such as meningitis. Some *E. coli* K1 surface constituents, in particular the polysialic acid capsule, are known to contribute to invasive potential but a comprehensive picture of the factors that determine the fully virulent phenotype has not so far emerged. We constructed a library and constituent sub-libraries of ~775,000 Tn*5* transposon mutants of *E. coli* K1 strain A192PP and employed transposon-directed insertion site sequencing (TraDIS) to identify genes required for fitness for infection in the two-day-old rat. Transposon insertions were lacking in 357 genes following recovery on selective agar; these genes were considered essential for growth in nutrient replete medium. Colonization of the mid-section of the small intestine was facilitated by 167 *E. coli* K1 gene products. Restricted bacterial translocation across epithelial barriers precluded TraDIS analysis of gut-to-blood and blood-to-brain transits; 97 genes were required for survival in human serum. The study revealed that a large number of bacterial genes, many not previously associated with systemic *E. coli* K1 infection, are required to realise full invasive potential.

IMPORTANCE *Escherichia coli* K1 strains cause life-threatening infections in newborn infants. They are acquired from the mother at birth and colonize the small intestine, from where they invade the blood and central nervous system. It is difficult to obtain information from acutely ill patients that shed light on physiological and bacterial factors determining invasive disease. Key aspects of naturally occurring age-dependent human infection can be

reproduced in neonatal rats. Here, we employ transposon-directed insertion site sequencing to identify genes essential for *in vitro* growth of *E. coli* K1 and genes that contribute to colonization of susceptible rats. The presence of bottlenecks to invasion of the blood and cerebrospinal compartments precluded insertion sequencing analysis but we identified genes for survival in serum.

63 **INTRODUCTION**

64 Early-onset sepsis and associated septicemia and meningitis are major causes of morbidity
65 and mortality in the first weeks of life. In the developed world, encapsulated *Escherichia coli*
66 and Group B streptococci are responsible for the large majority of these infections (1-3).
67 Over 80% of *E. coli* blood and cerebrospinal fluid isolates from infected neonates express
68 the α,2-8-linked polysialic acid (polySia) capsular K1 polysaccharide (4, 5), a polymer
69 facilitating evasion of neonatal immune defenses due to its structural similarity to the
70 polySia modulator of neuronal plasticity in the developing human embryo (6). Infections
71 arise due to colonization of the neonatal gastrointestinal (GI) tract by maternally derived *E.*
72 *coli* K1 at or soon after birth, from where the bacteria invade the systemic circulation to gain
73 entry into the central nervous system (CNS) (7, 8).

74 Essential features of the human infection can be reproduced in the neonatal rat,
75 enabling investigation of the pathogenesis of *E. coli* K1 neonatal invasive infections (9-11). In
76 susceptible two-day-old (P2) rat pups the protective mucus layer in the small intestine is
77 poorly developed but matures to full thickness over the P2-P9 period, coincident with the
78 development of resistance to invasive infection from GI-colonizing *E. coli* K1 (12). Thus, oral
79 administration of *E. coli* K1 initiates stable colonization of the small intestine in both P2 and
80 P9 pups but elicits lethal systemic infection only in younger animals (13). In the absence of
81 an effective mucus barrier at P2, the colonizing bacteria make contact with the apical
82 surface of enterocytes in the mid-region of the small intestine before translocation to the
83 submucosa by an incompletely defined transcellular pathway (12). They subsequently gain
84 access to the blood compartment by evading mesenteric lymphatic capture (10, 14). *E. coli*
85 K1 cells strongly express polySia in blood (15) and the capsule may protect the bacteria from
86 complement attack during this phase of the infection by facilitating binding of complement

regulatory factor H to surface-bound C3b to prevent activation of the alternative pathway (16, 17). Following hematogenous spread, the bacteria enter the CNS *via* the blood-cerebrospinal barrier at the choroid plexus epithelium to colonize the meninges (15). Some microorganisms that invade the CNS enter across the cerebral microvascular endothelium of the arachnoid membrane (18), although the restricted distribution of *E. coli* K1 within the CNS suggests this is not a primary route of entry for this pathogen.

Only a limited number of pathogenic bacteria have the capacity to invade the CNS from a remote colonizing site and the large majority elaborate a protective capsule that facilitates avoidance of host defenses during transit to the site of infection (19). Although the polySia capsule is clearly necessary for neonatal pathogenesis of *E. coli* K1 (11), the large majority of bacterial virulence factors that facilitate transit from GI tract to brain are unknown. A number of potential virulence factors associated with neonatal bacterial meningitis have been defined by phylogenetic analysis (20) and there is good evidence that the genotoxin colibactin and the siderophore yersiniabactin contribute to the pathogenesis of *E. coli* K1 in the experimental rat (21-23); however, a more detailed understanding of virulence mechanisms of *E. coli* K1 invasive disease will present opportunities for new modes of therapy for these devastating infections.

Transposon insertion sequencing (24, 25), a combination of traditional transposon mutagenesis and massively parallel DNA sequencing, is a powerful tool for the genome-wide enhanced genetic screening of large pools of mutants in a single experiment. It has recently been used to determine the full complement of genes required for expression of the K1 capsule by an *E. coli* uropathogenic isolate (26). The technique can be used to detect variations in genetic fitness of individual mutants undergoing selection in colonized and infected hosts. There are a number of variations of this procedure but all rely on the

creation of a pool of insertion mutants in which every locus has been disrupted at multiple sites; determination of the site of transposon insertion by sequencing of transposon junctions within chromosomal DNA before and after applying selective pressure will identify mutants attenuated under the selective condition (27). Thus, genes that confer fitness during *Klebsiella pneumoniae* (28) and *Acinetobacter baumannii* (29) lung persistence, systemic and mucosal survival of *Pseudomonas aeruginosa* (30), and spleen colonization in the mouse of uropathogenic *E. coli* (31) have been identified by this approach. In this study, we employ transposon-directed insertion site sequencing (TraDIS) (24) to interrogate a library of ~775,000 Tn*5* mutants or constituent sub-libraries of *E. coli* K1 strain A192PP for genes essential for growth *in vitro* and GI colonization, invasion and systemic survival in susceptible P2 rat pups. In addition, we identified "bottlenecks" (32) to systemic invasion that restrict population diversity and limit the potential for transposon insertion site analysis of infection in the GI-colonized neonatal rat.

RESULTS

Generation of a Tn*5* mutant library and identification of essential genes. To provide sufficient saturation density for the identification of *E. coli* K1 genes essential for growth *in vitro* and of those conferring fitness in a range of defined environments, approximately 300 individual pools, each of 1-5 x 10^3 transposon mutants of *E. coli* A192PP, were constructed and combined to form a library containing over 7.75 x 10^5 mutants. Linker PCR was performed on randomly selected mutants to confirm that Tn*5* had inserted into random genomic locations (Fig. S1). TraDIS was performed on pooled but uncultured mutants to identify Tn*5* insertion sites within the 5.52 Mbp genome of A192PP (33). Sequences of indexed amplicons were determined and 2 x 10^6 sequence reads containing Tn*5* were

mapped onto the *E. coli* K1 A192PP genome. Reads mapped to 237,860 unique Tn*5* insertion sites, and were distributed along the entire genome (Fig. 1A).

As the Tn*5* library contained a high transposon insertion density, genes with no or limited Tn*5* insertion sites are likely to be essential for growth in nutrient-replete media such as Luria-Bertani (LB) broth. We calculated insertion indices for each gene by normalizing the number of insertions in each gene by gene length. Insertion index values for two technical replicates were highly correlated (Spearman's rho = 0.9589) (Fig. 1B). A density plot of insertion indices produced a bimodal distribution, with a narrow peak representing genes with no or a limited number of Tn*5* insertions and a broad peak containing genes with a large number of Tn*5* insertions (Fig. 1C); the former comprised genes that confer lethality when mutated and the latter genes that can be mutated without affecting bacterial viability. To identify genes significantly lacking Tn*5* insertions and therefore essential for *in vitro* growth, gamma distributions from the density plot were used to determine $\log_2$ likelihood ratios. Examples of essential genes containing no or limited Tn*5* insertions are shown in Fig. 1D. A total of 357 genes were predicted to be essential for the *in vitro* growth of *E. coli* K1 A192PP and these are shown in Table S1, together with KEGG (Kyoto Encyclopedia of Genes and Genomes) descriptors for genes involved in metabolic pathways.

COG (Clusters of Orthologous Groups) was used to identify the functional category of each gene essential for growth *in vitro* from the A192PP whole genome sequence (accession number PRJEB9141). Genes involved in ribosomal structure (11% of total number of essential genes) and protein biosynthesis (15%) featured prominently and were significantly enriched in relation to their representation within the whole genome as were genes encoding proteins for DNA replication (3%), cell wall (peptidoglycan, lipopolysaccharide)

158 biosynthesis (6.25%) and membrane biogenesis (3%) (Fig. 2). Genes for protein secretion
159 and export as well as ABC transporter genes were also well represented; the remaining
160 essential genes were involved in a wide variety of cellular catabolic and anabolic functions.
161 The list features 254 genes that were found by TraDIS (34) to be essential for growth in Luria
162 broth of an *E. coli* ST131 multi-drug resistant urinary tract isolate (from a total of 315
163 essential genes). In similar fashion, 253 genes determined as essential for growth of *E. coli*
164 K12 MG1655 in LB broth were also identified as essential in the current study (Table S1); the
165 K12 study employed a comprehensive set of precisely defined, in-frame single-gene deletion
166 mutants (35), not transposon insertion sequencing.

167 **Maintaining Tn*5* library diversity.** The polySia capsule is a major determinant of virulence in
168 *E. coli* K1 and is central to the capacity of K1 clones to cause neonatal systemic infection (11,
169 36). PolySia biosynthesis imposes a substantial metabolic burden on producer strains (37).
170 As TraDIS and other transposon insertion sequencing procedures generally employ growth
171 in liquid medium for recovery and expansion of the output pool (38), we investigated the
172 impact of batch culture on the expression of the K1 capsule within the Tn*5* library. The
173 complete Tn*5* library was inoculated into LB broth, incubated for 8 h at 37^{O}C and the
174 proportion of encapsulated and non-encapsulated A192PP bacteria determined by
175 susceptibility to the *E. coli* K1-specific bacteriophage K1E within the population. Non-
176 encapsulated mutants initially comprised 4.66% of the bacterial population but by the end
177 of the incubation period this had risen to 98.24% (Fig. 3A). Growth rates in LB broth of *E. coli*
178 A192PP and a non-encapsulated mutant of A192PP randomly selected from the Tn*5* library
179 did not differ significantly (Fig. S2A).

The cultured Tn*5* library was avirulent as determined by administration to P2 neonatal rat pups whereas GI colonization with 2-6 x 10^6 CFU *E. coli* A192PP and the uncultured Tn*5* library were lethal. A similar colonizing inoculum of the cultured (8 h; 37^OC) *E. coli* A192PP-Tn*5* library had no impact on survival and all pups remained healthy over the seven-day observation period (Fig. 3B), even though all animals remained heavily colonized with K1 bacteria throughout the experiment (data not shown). Thus, culture of the library prior to challenge resulted in loss of phenotypic diversity and virulence. The complete Tn*5* library contained 2.81 x 10^5 unique Tn*5* insertions, of which 750 (2.66% of the bacterial population) possessed transposon insertions in genes determining capsule biosynthesis (data not shown). The probability that cultured sub-libraries of more than 5 x 10^3 mutants contained a non-capsulated mutant was calculated to be ≥0.98 but only 0.55 for sub-libraries of 1 x 10^3. Low complexity libraries of 10^3 mutants maintained virulence in P2 neonatal rat pups after culture whereas more complex libraries did not (Fig. S3), due to the absence of mutants lacking the capacity to express the polySia capsule within the inoculum. To minimise bias, in all subsequent experiments libraries of sufficient complexity to contain multiple numbers of non-encapsulated mutants were used; for experiments utilizing neonatal rats the period between colonization initiation and tissue harvesting was kept to a minimum and tissue homogenates were cultured directly on to selective agar plates with no intervening liquid culture step.

Genes required for GI colonization. *E. coli* A192PP colonize the small intestine of neonatal rats following oral administration of the bacterial bolus, with 10^7-10^8 K1 bacteria/g intestinal tissue persisting for at least one week (12, 13). Translocation of the neonatal pathogen to the blood compartment *via* the mesenteric lymphatic system occurs predominantly, and in

all likelihood exclusively, across the epithelium of the mid-section of the small intestine (MSI), even though the density of colonizing bacteria in this region of the GI tract is no greater than that within neighbouring proximal (PSI) or distal (DSI) locations (12).

Few attempts have been made to determine the genes or gene products required by *E. coli* K1 for colonization of the GI tract (39). To prevent loss of diversity of the *E. coli* K1 A192PP-Tn*5* library, we minimized the period of colonization before sampling the *E. coli* K1 population of the MSI. The colonizing *E. coli* K1 population in proximal, middle and distal regions of the small intestine did not expand beyond 4 h after initiation of colonization (Fig. 4A); GI tissues were therefore excised at this time point. To identify mutants with decreased capacity to colonize the MSI, P2 rats were fed 1 x 10^9 CFU of an *E. coli* K1 A192PP-Tn*5* library containing 2 x 10^5 mutants, the pups sacrificed after 4 h and *E. coli* K1 bacteria in the MSI enumerated. The bacterial load of rats colonized with the Tn*5* library was comparable to that of rats colonized with the wildtype strain (data not shown). MSI tissues from four rats were pooled, homogenized and cultured on LB agar containing kanamycin to ensure that mutant frequency was not overestimated by inclusion of measurements of DNA from dead bacteria; kan^R colonies were then pooled, DNA extracted and the fitness of each mutant determined by TraDIS. Input and output pools each comprised 2 x 10^5 CFU and the ratio of input:MSI read counts were expressed as $\log_2$ fold change. A wide distribution of fitness scores (40) were detected (Fig. 4B). The majority of transposon insertions did not have a strong negative or positive effect on colonisation of the MSI. A total of 387 transposon insertions, within 167 genes, had significantly decreased in normalized read counts between input and output pools (negative $\log_2$ fold change and $P< 0.05$; Table S2). Of the 387 insertion sites, 180 were not detectable in the output pool, demonstrating complete loss in the output pool. Many of these transposon insertion sites occurred within the same gene

(Table S2). For example, within the *neuC* gene, 70 unique transposon insertion sites were identified as lost during colonization. Transposon-interrupted genes were identified as important for colonization of the MSI and were grouped into seven arbitrary categories: (i) genes encoding surface structures, including pili, (ii) genes encoding secretory components, (iii) genes involved in intermediary metabolism, (iv) stress response genes, (v) cytoplasmic membrane (CM)-located genes, (vi) genes for iron acquisition and (vii) others and hypothetical genes.

A high proportion of mutations associated with decreased MSI colonizing capacity were located in genes affecting the biosynthesis of surface structures (Table S2). A few genes were involved in lipopolysaccharide (LPS) biosynthesis (*yrbH*, *yiaH*) and OM proteins (*ompG*, *ycbS*) but the majority affected the polySia capsule, with genes of the *neu* operon (41), accounting for 194 of the 387 colonization-attenuated mutants. There is some evidence that capsular polysaccharides may promote adhesion to biological and non-biological surfaces during biofilm formation (37) but there has been little or no consideration of a role for capsules as mediators of GI colonization.

A limited number of genes associated with type II and IV secretion were identified as required for colonisation of the MSI; these multiprotein complexes translocate a wide range of proteins and protein complexes across host membranes (42, 43) and are implicated in adherence and intestinal colonization of enterohemorrhagic *E. coli* in farm animals (44). Genes for assembly of pilus proteins, including some encoded by the *tra* locus, likely to be located on plasmids, that initiate conjugation, were also linked to colonization; pili are virulence factors that may mediate attachment to and infection of host cells (45). Colonization by both commensals and pathogens is dependent on nutrient scavenging, sensing chemical signals and regulation of gene expression as the bacteria adapt to a new

and potentially hostile environment that in the case of *E. coli* K1 appears to rely on stress response genes such has as *yhiM* (encodes a protein aiding survival at low pH) and the heat shock protein genes *clpB* and *yrfH*, as well as DNA repair genes. A large number encoded enzymes involved in the metabolism of sugars (e.g., *gcd*, *rpiR*, *glgC*), amino acids (*dadX*, *metB*, *tdcB*), fatty acids (*yafH*, *fixA*), growth factors (*bisC*, *yigB*, *thiF*) and other secondary metabolites (*yicP*). Transporters and permeases involved in central intermediary metabolism also featured prominently: these included permeases of the major facilitator superfamily (YjiZ), the hexose phosphate transport protein UhpT, the carnitine transporter CaiT and a range of CM-located sugar transporters. Of note was the impact of mutation of the *fucR* L-fucose operon activator on colonization; fucose is abundant in the GI tract and the fucose-sensing system in enterohemorrhagic *E. coli* regulates colonization and controls expression of virulence and metabolic genes (46). Availability of free iron is severely limited in the GI tract and ingestion of iron predisposes to infection (47); the importance of iron acquisition for *E. coli* K1 during GI colonization is reflected in the requirement for a number of genes related to iron uptake (e.g., *feoB*, *fepA*).

GI colonizing capacity and virulence of single gene mutants. To investigate the contribution of the polySia capsule to colonization of the neonatal rat GI tract, we disrupted the *neuC* gene of *E. coli* A192PP genes using bacteriophage λ Red recombinase to produce a capsule free mutant as judged by resistance to *E. coli* K1-specific phage K1E. We also produced other single gene mutants in genes identified by the TraDIS GI screen: *vasL* (encoding a type IV secretion system protein), *yfeC* (predicted to form part of a toxin/anti-toxin locus) and two genes with unknown function, *yaeQ* and *A192PP_3010* (the latter is present in genomes of other extra-intestinal *E. coli* pathogens, including IHE3034, UTI89, RS218, PMV-1 and S88).

Growth rates of these mutants, in particular the capsule-negative *neuC* mutant (Fig. S2B), were indistinguishable from that of the *E. coli* A192PP parent in LB medium. All were examined for their capacity to colonize the GI tract and cause lethal infection in P2 rat pups (Fig. 4C and 4D).

The *E. coli* A192PP parent strain or single gene mutants (2-6x10^6) were administered orally to P2 rats; all members of a litter of 12 pups received the same strain. Pups were sacrificed 24 h after initiation of colonization and *E. coli* K1 bacteria in the small intestine (PSI, MSI and DSI) and colon enumerated. The capacity of all mutants to transit the upper portion of the alimentary canal, pass through the stomach and colonize the small intestine was markedly inferior to the wildtype strain (Fig. 4C). Reductions in colonization of the PSI, MSI and DSI by the mutants, including *E. coli* A192PPΔ*neuC*::*kan*, were significant, the only exception being colonization of the DSI by A192PPΔ*yfeC*::*kan*, with no significant difference between parent and mutant. Interestingly, no increases in the numbers of viable A192PPΔ*neuC*::*kan*, A192PPΔ*vasL*::*kan*, A192PPΔ*3010*::*kan* and A192PPΔ*yaeQ*::*kan* recovered from the colon were noted to compensate for reductions in colonization of the small intestine. There was a significant increase in the colonic burden of viable A192PPΔ*yfeC*::*kan* bacteria compared to the parent strain. We have established (12) that *E. coli* A192PP transits to the blood circulation *via* the mesenteric lymphatic system by exploiting a vesicular pathway through the GI epithelium only at the MSI. As mutant numbers colonizing this region of the small intestine were much reduced compared to the parent strain, we determined the capacity of the single gene mutants to elicit lethal systemic infection following GI colonization by oral administration of 2-6x10^6 bacteria at P2 (Fig. 4D). Four of the five mutants (A192PPΔ*neuC*::*kan*, A192PPΔ*vasL*::*kan*, A192PPΔ*3010*::*kan* and A192PPΔ*yfeC*::*kan*) displayed significantly reduced lethal potential

298 compared to the A192PP parent. Loss of capsule (A192PPΔ*neuC*::*kan*) resulted in complete
299 loss of lethality over the seven day observation period. Administration of A192PPΔ*vasL*::*kan*
300 elicited a lethal response in 41.6% of pups; 33.3% and 25% survived after receiving,
301 respectively, A192PPΔ*3010*::*kan* and A192PPΔ*yfeC*::*kan* at P2. For A192PPΔ*yaeQ*::*kan*, 75%
302 of pups succumbed to lethal infection but did not reach levels of significance when
303 compared to the 100% lethality engendered by the A192PP parent (P >0.05). Overall, these
304 data indicate that the TraDIS screen efficiently identified genes important for MSI
305 colonization that impact on pathogenic potential.

306 **A bottleneck to infection in the neonatal rat prevents identification of genes for**
307 **translocation across the gastrointestinal epithelium.** Our initial intention was to exploit the
308 high degree of susceptibility of the P2 neonatal rat to systemic infection, sepsis and
309 meningitis following oral administration of an effective dose of *E. coli* A192PP in order to
310 determine all genes required to enable the neonatal pathogen to overcome previously
311 defined (12-15) physical and immunological barriers to invasion of the blood circulation and
312 dissemination to the meninges. However, earlier studies indicate that relatively few *E. coli*
313 K1 bacteria migrate from colonized sites within the GI tract to the blood (10), constraining
314 the genetic diversity of the translocated bacterial population and eliminating genotypes
315 from the translocated gene pool in a stochastic manner that does not reflect the fitness of
316 individual genes to contribute to genotypes with invasive potential (32). We therefore
317 determined if bottlenecks existed which would compromise the identification of mutants
318 with attenuated capacity to translocate from the GI tract to the blood compartment; if any
319 experimental bottlenecks are narrower than the complexity of the *E. coli* A192PP Tn*5*
320 library, many relevant transposon insertion mutants will be lost entirely by chance (38).

Further, the existence of a restrictive bottleneck would limit the complexity of the library that could be used for TraDIS evaluation of populations colonizing the MSI (input pool) and reaching the blood (output pool).

We constructed an *E. coli* A192PPΔ*lacZ* mutant by bacteriophage λ Red recombineering and confirmed that there was no significant difference in lethal potential between *E. coli* A192PP and the *lacZ* mutant (Fig. 5A). We then used mixtures of parent and mutant to investigate the existence of bottlenecks that restrict translocation to the blood compartment. A 1:1 mixture (total 2-4 x10^6 CFU) of *E. coli* A192PP and A192PPΔ*lacZ* was administered orally to P2 rat pups, the animals sacrificed after 24 h and GI tissue homogenates plated for quantification of each strain. The competitive index (CI), the ratio of input A192PP: A192PPΔ*lacZ* to output A192PP: A192PPΔ*lacZ*, was calculated for excised PSI, MSI, DSI, colon and mesenteric lymphatic tissue and for blood. CI values in the PSI, MSI, DSI and colon were not significantly different from 1 (one-sample *t*-test), indicating that the composition of the colonizing inoculum was maintained in each rat pup (Fig. 5B). However, there was more heterogeneity in CI values of bacterial populations from the blood and in five pups only one strain could be recovered from the blood (four animals parent strain only, one animal A192PPΔ*lacZ* only). The highly restrictive bottleneck between GI epithelial transport and entry into the blood circulation supports the argument that reduced virulence of the complete, cultured library in comparison to less complex sub-libraries (Fig. S3) is at least in part due to a reduced likelihood that a fully virulent mutant would randomly escape capture by the mesenteric lymphatic system. The presence of significant bottlenecks between the GI tract, blood circulation and brain was confirmed by determination of the complexity of recovered Tn-*5* library populations from these sources (Fig. 5C).

344 **Identification of *E. coli* K1 A192PP genes required for survival in human serum.** Systemic
345 infection in the neonatal rat is likely to be maintained only if *E. coli* A192PP bacteria survive
346 in the blood circulation. Due to limited exposure to antigens *in utero* coupled with deficits in
347 adaptive immunity, neonates depend on innate immunity for protection against infection.
348 The complement system provides front line innate defense against Gram-negative bacterial
349 infection and the polySia capsule in turn enables *E. coli* K1 to avoid successful complement-
350 mediated attack by host immune mechanisms. To obtain insights into *E. coli* K1
351 pathogenesis during the invasive phase of the infection, and in light of restrictions placed on
352 the neonatal rat model with regard to the use of TraDIS by the gut-to-blood bottleneck, we
353 used the *E. coli* A192PP Tn*5* library to investigate genes essential for A192PP fitness in
354 pooled normal human serum, a reliable and plentiful source of all soluble components of
355 the three complement pathways (48).

356 *E. coli* A192PP is resistant to the bactericidal action of human serum (Fig. 6A). A
357 portion of the A192PP-Tn*5* library containing 2 x 10^4 mutants (1 x 10^9 CFU) was incubated in
358 either 30% human serum or 30% heat-inactivated serum (final volume 375 µl) at 37^{o}C for 3
359 h. Kan^R bacteria in the input and output pools (each 2 x 10^5) were collected, DNA extracted
360 from each pool and transposon insertion sites sequenced. A wide distribution of fitness
361 scores were detected (Fig. 6B). Mutation of 97 genes (negative $\log_2$-fold change and $P< 0.05$)
362 resulted in decreased survival in normal serum, but not in heat-inactivated serum (Fig. 6C &
363 Table S3).

364 A high proportion of genes identified in the TraDIS screen as contributing to
365 resistance encoded cell surface constituents. It is well established that the polySia capsule
366 protects *E. coli* K1 from complement attack (16, 17) and three mutations in the *kps* capsule
367 gene cluster compromised serum survival. The central region of the cluster contains the *neu*

genes that direct the biosynthesis, activation and polymerization of the *N*-acetylneuraminic acid building block of polySia. *neuC* encodes the UDP *N*-acetylglucosamine 2-epimerase that catalyzes the formation of *N*-acetylmannosamine (49) and the *O*-acetyltranferase *neuD* acetylates monomeric neuraminic acid at carbon position 7 or 9 (50). KpsM is a component of the multimeric ATP-binding cassette transporter involved in the translocation of the polySia capsule through a transmembrane corridor to the cell surface (41, 51). Disruption of the genes encoding these proteins will prevent polySia expression (41); interruption of *rfaH*, identified in the TraDIS screen, will also prevent capsule expression but its loss will have a more profound effect on the surface topography of *E. coli* A192PP, as this transcriptional anti-terminator is required for the expression of operons that direct the synthesis, assembly and export of LPS core components, pili and toxins in addition to the capsule (52, 53). Indeed, survival in serum is dependent on anti-termination control by RfaH (54). Another gene identified that impacts on capsule formation was *bipA*; BipA is a tyrosine-phosphorylated GTPase that regulates through the ribosome a variety of cell processes, including some associated with virulence (55, 56). Other genes involved in LPS biosynthesis and pilus formation were also identified: *waaW* is a UDP-galactose:(galactosyl) LPS alpha1,2-galactosyltransferase involved in the synthesis of the R1 and R4 LPS core oligosaccharides (57) and *wzzE* encodes a polysaccharide copolymerase that catalyzes the polymerization of LPS O-antigen oligosaccharide repeat units into a mature polymer within the periplasmic space in readiness for export to the cell surface (58). Both mutations will prevent attachment of LPS O-side chains to the core oligosaccharide of LPS. The 16 genes that specify pilus synthesis that were identified in the screen included the majority of genes of the *tra* locus.

391 The TraDIS screen identified a range of proteins that are embedded in the OM (Fig.
392 6C), none of which had been previously implicated in complement resistance, and which
393 could influence the topography of the bacterial surface. Of the remaining genes with
394 assigned function, the majority were involved with cell metabolism and the stress response;
395 it is well established that metabolic processes are intimately associated with the
396 complement-mediated bacterial killing process (59, 60). To verify the screen, we
397 constructed four single gene mutants of *E. coli* A192PP by bacteriophage λ Red
398 recombineering. Genes with roles in LPS synthesis (*rfaH* and *waaW*), capsule synthesis
399 (*neuC*) and pilus assembly (*traL*) were mutated; none showed any reduction in growth rate
400 in LB broth. All displayed significant reductions in complement resistance following
401 incubation in pooled human serum (Fig. 6D). *E. coli* A192PPΔ*rfaH* was exquisitely susceptible
402 with no colonies detected after 30 min. The viability of A192PPΔ*neuC* was also
403 compromised with a threefold log reduction in viability over the 3 h incubation period.
404 Killing of A192PPΔ*traL* and A192PPΔ*waaW* was less marked but these mutations
405 significantly reduced viability. Complementation of the mutants with the functional gene
406 introduced on a pUC19 vector completely restored resistance in all cases (Fig. 6D). These
407 genes also contributed to lethality in the P2 neonatal rat (Fig. 6E). The lethal capacity of
408 A192PPΔ*neuC*, A192PPΔ*rfaH* and A192PPΔ*waaW* was completely attenuated in comparison
409 to *E. coli* A192PP; 42% of pups administered A192PPΔ*traL* succumbed to systemic infection
410 (all $P < 0.01$).

411 **DISCUSSION**

412 Systemic infection with meningeal involvement arises spontaneously after GI colonization of
413 neonatal rats with a high proportion of *E. coli* K1 isolates and the pathway to infection

414 mirrors to a large extent that of natural infections in the human host. In contrast to models
415 of bacterial infection that create an artificial pathogenesis bypassing some or all of the
416 barriers to infection by injection of a bacterial bolus directly into the blood circulation, the
417 neonatal rat model provides an opportunity to investigate in stepwise fashion the progress
418 of the pathogen as it transits from gut to blood to brain. TraDIS and other transposon
419 sequencing methods enable simultaneous and rapid determination of the fitness
420 contribution of every gene for a given condition and therefore have the potential to enable
421 the identification of genes that are essential for, or significantly contribute to, each step of
422 the infection process. However, stochastic loss will become evident if each mutant in the
423 input pool does not have an equal chance to overcome the physical, physiological and
424 immunological barriers presented by the host (61). This was clearly the case with epithelial
425 transit of *E. coli* A192PP, with evidence that on occasion systemic infection arose due to only
426 one viable bacterial cell entering the blood circulation (Fig. 5B), and complements other
427 studies showing single or low-cell-number bottlenecks in models of severe infection (62-64).
428 As translocation from MSI colonizing sites to the blood was not amenable to analysis by
429 TraDIS we determined genes essential for survival in the presence of complement, a major
430 component of the innate immune system that protects against extracellular systemic
431 pathogens (17).

432 The high density of transposon insertion into random genomic positions along the
433 entire *E. coli* A192PP chromosome, with minimal insertional bias (Fig. 1A), enabled the
434 identification of genes essential for growth in nutrient replete LB medium. Of the 357 *E. coli*
435 A192PP genes considered essential, orthologues of 254 (from 315) had been previously
436 identified using TraDIS in a multi-drug-resistant uropathogenic strain of *E. coli* ST131 grown
437 in LB (34) and 253 in an *E. coli* K12 strain (35), confirming the existence of a core set of

essential genes in *E. coli*. As anticipated, a high proportion of these genes encoded enzymes involved in a range of key metabolic functions such as carbohydrate, protein and nucleobase metabolism, and the remainder were associated with essential functions such as transport, cell organisation and biogenesis.

During characterization of the *E. coli* A192PP mutant library we examined the impact of culture in liquid medium on the expression of the polySia capsule, which places large demands on cell energy expenditure, as lengthy incubation times before marker selection may lower library complexity (38). Unexpectedly, we found that prolonged culture of the library enriched the proportion of non-encapsulated mutants (Fig. 3A). We anticipated that loss of capsule would enable the non-encapsulated mutants to grow at a faster rate than capsule-replete mutants and wildtype, and out-compete capsule-bearing library members. However, growth of a non-encapsulated mutant selected at random from the library was virtually identical to, and not significantly different from, the *E. coli* A192PP parent strain (Fig. S2A). There was also no difference in the climax populations of the strains at the end of the logarithmic phase of growth. In similar fashion, the growth curve for a *neuC* single gene mutant was identical to *E. coli* A192PP (Fig. S2B). *neuC* is involved in the synthesis of the *N*-acetylneuraminic acid monomeric unit of polySia, and as a consequence is unable to elaborate the capsule. It is clearly impractical to evaluate the growth kinetics of every distinct non-encapsulated mutant in the Tn5 library but it currently appears that differences in growth rate of individual library members cannot explain the highly reproducible enrichment that we observed. Indeed, use of transposon insertion libraries is predicated on the assumption that there are no significant differences in the growth rate of individual mutants. At present, the basis of the loss of mutants expressing capsule in TraDIS library cultures cannot be readily explained.

462 A sub-library of 2 x 10^5 mutants was used to establish genes involved in GI
463 colonization. To minimize bias due to any outgrowth of non-encapsulated mutants on the GI
464 epithelium we harvested *E. coli* K1 from the MSI after 4 h, by which time maximal CFU had
465 been achieved; bacteria were plated directly on to solid medium to further avoid outgrowth.
466 Bias due to this restricted timeline is likely to be low as the majority of genes involved in
467 adhesion and complement resistance are expressed constitutively. TraDIS identified the
468 polySia capsule as a major determinant of GI colonization associated with *E. coli* K1. There is
469 little or no evidence from the literature that capsules of Gram-negative bacteria enhance GI
470 colonization; indeed, it has been reported that they interfere with adhesive interactions by
471 obstructing binding of underlying surface molecules to mucosal surfaces (65, 66). The single
472 gene mutant *E. coli* A192PPΔ*neuC*::*kan* displayed a reduced capacity to colonize the MSI
473 (Fig. 4E), although it should be borne in mind that passage through the upper alimentary
474 canal and stomach may impact on the number of mutant bacteria gaining access to the
475 small intestine. In this context it should be noted that capsular exopolysaccharide protects
476 *E. coli* from the environmental stress of stomach acid (67).
477 Other cell surface structures that are likely to have an impact on adhesion and
478 colonization of the mucosal layer associated with the MSI were identified by TraDIS. Pili are
479 established mediators of adhesion of *E. coli* to the host epithelium, although a large
480 proportion of the evidence comes from enterotoxigenic and enteropathogenic strains (68,
481 69). LPS and OM protein encoding genes were also implicated, as were genes involved in the
482 stress response, reflecting ongoing adaptation to a new and hostile environment. The
483 involvement of genes encoding metabolic enzymes, including some for anaerobic
484 respiration, equates to increases in bacterial cell numbers in the anaerobic environment of
485 the small intestine and for iron acquisition genes this reflects the low availability of

intestinal luminal iron (47, 70). Genes encoding some components of type II and type IV secretion systems were found with decreased frequency in the output pool. Members of these gene categories were also identified by Martindale et al (39) as necessary for GI colonization of *E. coli* K1 faecal isolate RS228 using signature-tagged mutagenesis; no genes found in this study were identified in the current study, in spite of the close genetic relatedness of the strains employed.

The intestinal lumen represents a potentially important portal of entry for pathogens into the host through adhesion, invasion or disruption of the epithelial barrier (71). In neonatal rats, *E. coli* K1 induces no detectable disruption of barrier integrity but exploits an intracellular pathway to access the submucosa (12). Only small numbers of bacteria breech the mesenteric lymphatic barrier in apparently random fashion (Fig. 5) and this precludes analysis by TraDIS. To accumulate data on genes and gene products facilitating invasion and survival/replication in the blood circulation, we examined essentiality for avoiding complement-mediated bactericidal effects. Although not all *E. coli* K1 isolates from cases of systemic infection are resistant to complement, resistance amongst K1 and K5 capsular types is more frequently encountered that for other K types (72); *E. coli* O18:K1 strains (such as A192) are in turn more often resistant than other O:K serotype combinations (73) due to the capacity of the polySia capsule to prevent complement activation. It is assumed, but not established, that the polySia capsule surrounding susceptible strains does not completely mask either OM-located activators of complement or lipid domains on the outer surface of the cell that are targets for OM intercalation of the C5b-9 membrane attack complex, the entity responsible for bacterial killing (59). In addition, long and numerous LPS O-side chains are necessary but not sufficient to enable the target cell to avoid complement killing (74) and they are able to bind C1 inhibitor to arrest classical or lectin pathway activation at the

early C1 stage (75). The importance of these structures for the complement resistance of *E. coli* K1 is supported by the decreased frequency of key LPS and capsule genes in the output pool along with a large number of OM-embedded proteins.

A small number of OM proteins, such as TraT and Iss, have been implicated as determinants of complement resistance (74) but they have been introduced into low-resistance backgrounds in high copy number; their role in the intrinsic resistance of clinical isolates is unclear and no mechanisms have been invoked to account for increases in resistance. The insertion of large numbers of protein molecules into the OM may fortuitously alter the biophysical properties of the bilayer, reducing the surface area and fluidity of lipid patches that are essential for binding and assembly of the C5b-9 membrane attack complex. The identification by TraDIS of a range of OM proteins as putative complement resistance determinants creates an opportunity to systematically investigate their precise function through generation of single gene mutants and we intend to pursue this line of investigation. We suggest that the architecture of the external surface of the OM, together with other more external macromolecular structures such as polysaccharide capsules, influences the capacity of the pore-generating C5b-9 complex to perturb the integrity of the OM. Thus, the surface of susceptible strains contains a sufficient number of exposed lipid domains to facilitate C5b-9 generation and penetration whereas the spatial and temporal organization of the OM of resistant bacteria is dominated by supramolecular protein assemblages to a degree where insufficient hydrophobic domains are available to act as C5b-9 assembly and binding sites, and this state persists throughout the growth cycle. The data we have generated in this study is compatible with this hypothesis. An array of metabolic genes emerged as essential for maintenance of the complement resistant phenotype (Fig. 6D) and may be indicative of repair processes invoked due to complement

attack. Exposure of resistant *E. coli* to complement results in minor perturbation of membrane integrity and metabolic homeostasis (76, 77) and C5b-9 intercalation into the OM has profound effects on cellular metabolic parameters (60).

TraDIS has also been employed by Schembri and coworkers to define the serum resistome of a globally disseminated, multidrug resistant clone of *E. coli* ST131 (34). They identified, and in most cases validated, 56 genes that contributed to the high level of complement resistance displayed by this pathogen. In similar fashion to our study, genes involved in the synthesis and expression of cell surface components were prominent. A number of genes contributing to LPS biosynthesis such as those of the *waa* operon, the *wzz* locus and *rfaH*, were common to both studies, as was the gene encoding the intermembrane protein AcrA. Genes of the plasmid-encoded *tra* locus, which we determined to be components of the *E. coli* A192PP serum resistome, were not present in *E. coli* ST131 (34) but other OM-located proteins may fulfill a similar role in reducing the fluidic properties of the bilayer. In contrast to the well-established role of the *E. coli* K1 polysialyl polymer in prevention of complement activation, no capsule genes were identified as components of the serum resistome of *E. coli* ST131, but different ST131 isolates express different capsule types due to extensive mosaicism at the capsule locus (78) and these uronic acid-containing polymers are unlikely to prevent complement activation (75). Thus, the different strategies employed by the two strains to prevent successful complement attack, together with differences in the bacterial surface composition and topography, probably explain variations in the serum resistomes of these related pathogens.

In summary, we identified *E. coli* K1 genes required for growth in standard laboratory liquid medium and for colonization of the GI tract of P2 neonatal rat pups. Both data sets provide insights into the biology of K1 neuropathogens and could provide the basis

for drug discovery programs for identification of selective antibacterial or colonization-inhibiting agents. In our rodent model, the stochastic nature of invasion of blood and probably brain prevented TraDIS analysis of gene essentiality for crossing gut epithelial and choroid plexus borders but some indication of genes necessary for survival in blood were obtained from output pool analyses after incubation of *E. coli* A192PP in human serum, a potent source of complement.

MATERIALS AND METHODS

Ethics statement. Animal experiments were approved by the Ethical Committee of the UCL School of Pharmacy and the United Kingdom Home Office and were conducted in accordance with national legislation.

Bacteria and culture conditions. *E. coli* strain A192PP was obtained by serial passage in P2 neonatal rats of *E. coli* A192 (serotype O18:K1) isolated from a patient with septicemia (79), as described earlier (11). Carriage of the polysialyl K1 capsule was determined with phage K1E (80): colonies were streaked onto MH agar, 10 µl of ~10^9 PFU/ml phage suspension dropped on each streak, the plates incubated overnight at 37^{o}C and the proportion of encapsulated bacteria within cultures quantified by comparing the ratio of phage-susceptible and phage-resistant colonies. *E. coli* A192PP single gene mutants (Table 1) were constructed using bacteriophage λ Red recombination (81); the oligonucleotides employed for construction of targeted mutants, for confirmation of targeted mutants and for construction of complemented mutants are shown in Tables S3-S5. All were cultured in Luria-Bertani (LB) and on LB agar at 37^{o}C; media were supplemented with either 100 µg/ml ampicillin or 50 µg/ml kanamycin as required.

Tn*5* library construction. The EZ-Tn*5* <KAN-2> Tnp transposome (Epicentre Biotechnologies) was introduced into *E. coli* A192PP by electroporation. Transformants were selected by overnight growth on LB plates containing 50 µg/ml kanamycin. Pools of 1-5 x 10^3 colonies were collected and frozen at -80°C in PBS containing 20% glycerol. Aliquots of individual pools were combined to create larger populations of mutants of up to 7.75 x 10^5. Genomic DNA was extracted from 1 ml cultures using the PurElute Bacterial Genomic Kit (Edge Biosystems) following standard protocol.

Linker PCR of Tn*5* insertion sites. Linker PCR was used to test individual transformant colonies and to confirm individual random-insertion events. DNA (2.5 µg) was digested with AluI restriction enzyme (Promega) and purified using MinElute PCR purification kit (QIAgen). A linker, formed by annealing of oligonucleotides 254 ($^{5'}$CGACTGGACCTGGA$^{3'}$) and 256 ($^{5'}$GATAAGCAGGGATCGGAACCTCCAGGTCCAGTCG$^{3'}$), was ligated to purified fragments (50 ng) with Quick Ligation kit (NEB). Linker PCR was performed with linker- and transposon-specific oligonucleotides (258 $^{5'}$GATAAGCAGGGATCGGAACC$^{3'}$ and $^{5'}$GCAATGTAACATCAGAGATTTTGAG$^{3'}$ respectively) using HotStart Taq Mastermix kit (QIAgen) and thermocycling conditions of 95°C for 5 min, 35 cycles of 94°C for 45 s, 56°C for 1 min and 72°C for 1 min, and 72°C for 10 min. Resulting amplicons were separated on 1.5% agarose gels at 100 V for 60 min.

Illumina sequencing. For sequencing of Tn*5* insertion sites, approximately 2 µg of genomic DNA was degraded to ~ 500 bp fragments by ultrasonication using a Covaris instrument. Fragments were end-repaired and A-tailed using the NEBNext DNA library preparation reagent kit for Illumina sequencing (NEB). Adapters Ind_Ad_T (ACACTCTTTCCCTACACGACGCTCTTCCGATC*T; where * indicates phosphorothionate) and

603 Ind_Ad_B (pGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCTC) were annealed and
604 ligated to DNA fragments. PCR was performed with transposon- and adapter- specific
605 primers Tn-FO
606 $^{5'}$TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCGGGGATCCTCTAGAGTCGACCTGC$^{3'}$ and
607 Adapt-RO
608 $^{5'}$GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACACTCTTTCCCTACACGACGCTCTTCCGATC$^{3'}$. Tn-FO and Adapt-RO contain a forward overhang and reverse overhang for indexing of
610 amplicons by Nextera index primers (Illumina). PCR was performed using HotStart Taq
611 Mastermix kit (QIAgen) and thermocycling conditions of 95°C for 5 min, 22 cycles of 94°C for
612 45 s, 56°C for 1 min and 72°C for 1 min, and 72°C for 10 min. Resulting amplicons were
613 separated on 1.5% agarose gels at 70 V for 90 min, and those between 150 and 700 bp
614 selected and purified using QIAquick Gel Extraction kit (QIAgen). Samples were indexed with
615 oligonucleotides from Nextera XT Index Kit (Illumina) using HotStart ReadyMix (Kapa
616 Biosystems) and thermocycling conditions of 95°C for 3 min, 8 cycles of 95°C for 30 s, 55°C
617 for 30 s and 72°C for 30 s, and 72°C for 5 min applied. Indexed amplicons were purified
618 using the AMPure XP system (Agencourt). The final concentration of samples was confirmed
619 using Qubit dsDNA BR assays (Thermofisher Scientific). Indexed amplicons were sequenced
620 on an Illumina Mi-Seq platform as 151-bp paired-end reads following manufacturer's
621 protocol (Illumina).

622 **Bioinformatic and statistical analysis.** Raw sequence reads that passed Trimmomatic quality
623 control filters (82) and contained the Tn*5* transposon were mapped to the *E. coli* K1 A192PP
624 reference genome (14) using Bowtie (83), permitting zero mismatches and excluding reads
625 that did not map to a single site. The reference genome assembly contains ORFs located on

626 contigs that were mapped to the IHE3034 chromosome and ORFs located on other contigs
627 that are likely to map to plasmids and other mobile genetic elements. An in-house pipeline
628 based on the SAMtools (http://samtools.sourceforge.net) and BCFtools toolkits was utilised
629 on the alignment files to determine insertion sites and coverage. To identify essential and
630 non-essential genes, the insertion index was calculated for each gene by dividing the
631 number of unique insertions in the gene by gene length. Observed insertion index values
632 were fitted to a bimodal distribution with a gamma distribution (or an exponential
633 distribution for genes with no observed insertion sites) corresponding to essential and non-
634 essential genes. A $\log_2$ likelihood, and corresponding *P* values, of each gene belonging to
635 essential or non-essential sets was calculated using R software. To compare the fitness of
636 individual mutants in input and output populations, reads were normalised and tested for
637 differential base means by calculating $\log_2$-fold changes and corresponding *P* values at a
638 false discovery rate of 0.1 using DESeq with R software. Raw read data for all transposon
639 insertions have been deposited in the European Nucleotide Archive (ENA); accession
640 numbers are as follows: ERR2235345 and ERR2235346 for identification of essential genes
641 for replicates 1 and 2; ERR2235567 for input population; ERR2235568 for output population
642 of rat MSI genes; ERR2235569 for output population of serum-exposed *E. coli* A192PP;
643 ERR2235570 for output population of bacteria exposed to heat-inactivated serum.

644 **Colonisation and infection of neonatal rats.** Timed-birth Wistar rat pup litters (usually *n* =
645 12) were purchased from Harlan UK, delivered at P2 and colonized on the same day. Pups
646 were retained throughout each experiment with the natural mothers in a single dedicated
647 cage under optimal conditions (19-21^{o}C, 45-55% humidity, 15-20 changes of air/h, 12 h
648 light/dark cycle) and were returned to the mother immediately after colonization. Mothers

had unrestricted access to standard rat chow and water. The procedure has been described in detail (84). In brief, all members of P2 rat pup litters were fed 20µl of mid-logarithmic-phase *E. coli* (2-6 x 10^6 CFU unless otherwise stated) from an Eppendorf micropipette. GI colonization was confirmed by culture of perianal swabs on MacConkey agar and bacteremia detected by MacConkey agar culture of blood taken *post mortem*. Disease progression was monitored by daily evaluation of symptoms of systemic infection and neonates culled by decapitation and recorded as dead once a threshold had been reached: pups were regularly examined for skin color, agility, agitation after abdominal pressure, presence of a milk line, temperature, weight and behaviour in relation to the mother. Neonates were culled immediately when abnormalities for three of these criteria were evident. After sacrifice, GI tissues were excised aseptically without washing, colon separated and the SI segmented into 2 cm portions representing proximal, middle and distal small intestinal tissue. Tissues were then transferred to ice-cold phosphate-buffered saline, and homogenized. Bacteria were quantified by serial dilution culture on MacConkey agar supplemented with 25 µg/ml kanamycin as appropriate. The presence of *E. coli* K1 was confirmed with phage K1E: 20 lactose-fermenting colonies were streaked onto MH agar, 10 µl of ~10^9 PFU/ml phage suspension dropped on each streak and the plates incubated overnight. *E. coli* K1 bacteria were quantified by multiplying total CFU by the proportion of K1E susceptible colonies. In all cases at least 19 colonies were susceptible to the K1 phage; *E. coli* K1 was never found in samples from non-colonized pups.

Susceptibility to human serum. Serum was obtained from healthy volunteers and used immediately. Bacteria were grown to late logarithmic phase in LB broth in an orbital incubator (minimum 200 orbits/min), 500 µl culture removed, washed twice with gelatin-

Veronal buffered saline plus magnesium and calcium ions (pH 7.35) (GVB^{++}), and suspended in an equal volume of GVB^{++}. Fresh human serum was diluted 1:3 in GVB^{++} and pre-warmed to 37°C. Bacterial suspensions and serum solutions were mixed 1:2 to give a final concentration of ~10^7 CFU/ml and incubated at 37°C for 3 h in a total volume of 125 µl containing 22% serum. Surviving *E. coli* were quantified by serial dilution and overnight incubation on LB agar. Pre-warmed, heat-inactivated (56°C, 30 min) serum served as control.

ACKNOWLEDGMENTS

This work was supported by research grant MR/K018396/1 from the Medical Research Council. The National Institute for Health Research University College London Hospitals Biomedical Research Centre provided infrastructural support.

695 **REFERENCES**

696 1. Simonsen KA, Anderson-Berry AL, Delair SF, Davies HD. 2014. Early-onset neonatal
697 sepsis. Clin Microb Rev 27:21-47.
698 2. Bonacorsi S, Bingen E. 2005. Molecular epidemiology of *Escherichia coli* causing
699 neonatal meningitis. Int J Med Microbiol 295:373-381.
700 3. Tsai MH, Lee CW, Chu SM, Lee IT, Lien R, Huang HR, Chiang MC, Fu RH, Hsu JF, Huang
701 YC. 2016. Infectious complications and morbidities after neonatal bloodstream
702 infections: an observational cohort study. Medicine (Baltimore) 95:e3078.
703 4. Robbins JB, McCracken GH, Gotschlich EC, Ørskov F, Ørskov I, Hanson LA. 1974.
704 *Escherichia coli* K1 capsular polysaccharide associated with neonatal meningitis. N
705 Eng J Med 290:1216-1220.
706 5. Korhonen TK, Valtonen MV, Parkkinen J, Väisänen-Rhen V, Finne J, Ørskov F, Ørskov
707 I, Svenson SB, Mäkelä PH. 1985. Serotypes, hemolysin production, and receptor
708 recognition of *Escherichia coli* strains associated with neonatal sepsis and meningitis.
709 Infect Immun 48:486-491.
710 6. Rutishauser U. 2008. Polysialic acid in the plasticity of the developing and adult
711 vertebrate nervous system. Nat Rev Neurosci 9:26-35.
712 7. Sarff LD, McCracken GH, Schiffer MS, Glode MP, Robbins JB, Ørskov I, Ørskov F.
713 1975. Epidemiology of *Escherichia coli* K1 in healthy and diseased newborns. Lancet
714 i:1099-1104.
715 8. Tunkel AR, Scheld WM. 1993. Pathogenesis and pathophysiology of bacterial
716 meningitis. Clin Microbiol Rev 6:118-136.

9. Glode MP, Sutton A, Moxon ER, Robbins JB. 1977. Pathogenesis of neonatal *Escherichia coli* meningitis: induction of bacteremia and meningitis in infant rats fed *E. coli* K1. Infect Immun 16:75-80.
10. Pluschke G, Mercer A, Kusećek B, Pohl A, Achtman M. 1983. Induction of bacteremia in newborn rats by *Escherichia coli* K1 is correlated with only certain O (lipopolysaccharide) antigen types. Infect Immun 39:599-608.
11. Mushtaq N, Redpath MB, Luzio JP, Taylor PW. 2004. Prevention and cure of systemic *Escherichia coli* K1 infection by modification of the bacterial phenotype. Antimicrob Agents Chemother 48:1503-1508.
12. Birchenough GMH, Dalgakiran F, Witcomb LA, Johansson MEV, McCarthy AJ, Hansson GC, Taylor PW. 2017. Postnatal development of the small intestinal mucosa drives age-dependent regio-selective susceptibility to *Escherichia coli* K1 infection. Sci Rep 7:83.
13. Birchenough GMH, Johansson MEV, Stabler RA, Dalgakiran F, Hansson GC, Wren BW, Luzio JP, Taylor PW. 2013. Altered innate defenses in the neonatal gastrointestinal tract in response to colonization by neuropathogenic *Escherichia coli*. Infect Immun 81:3264-3275.
14. Witcomb LA, Collins JW, McCarthy AJ, Frankel G, Taylor PW. 2015. Bioluminescent imaging reveals novel patterns of colonization and invasion in systemic *Escherichia coli* K1 experimental infection in the neonatal rat. Infect Immun 83:4528-4540.
15. Zelmer A, Bowen M, Jokilammi A, Finne J, Luzio JP, Taylor PW. 2008. Differential expression of the polysialyl capsule during blood-to-brain transit of neuropathogenic *Escherichia coli* K1. Microbiology 154:2522-2532.

16. Meri S, Pangburn MK. 1990. Discrimination between activators and nonactivators of the alternative pathway of complement: regulation via a sialic acid/polyanion binding site on factor H. Proc Natl Acad Sci USA 87:3982-3986.
17. Taylor PW. 1993. Non-immunoglobulin activators of the complement system, p 37-68. *In* Sim RB (ed), Activators and Inhibitors of Complement. Kluwer Academic Publishers, Dordrecht.
18. Nassif X, Bourdoulous S, Eugène E, Couraud PE. 2002. How do extracellular pathogens cross the blood-brain barrier? Trends Microbiol 10:227-232.
19. Brouwer MC, Tunkel AR, van de Beek D. 2010. Epidemiology, diagnosis, and antimicrobial treatment of acute bacterial meningitis. Clin Microbiol Rev 23:467-492.
20. Johnson JR, Oswald E, O'Bryan TT, Kuskowski MA, Spanjaard L. 2002. Phylogenetic distribution of virulence-associated genes among *Escherichia coli* isolates associated with neonatal bacterial meningitis in the Netherlands. J Infect Dis 185:774-784.
21. Payros D, Secher T, Boury M, Brehin C, Ménard S, Salvador-Cartier C, Cuevas-Ramos G, Watrin C, Marcq I, Nougayrède JP, Dubois D, Bedu A, Garnier F, Clermont O, Denamur E, Plaisancié P, Theodorou V, Fioramonti J, Olier M, Oswald E. 2014. Maternally acquired genotoxic *Escherichia coli* alters offspring's intestinal homeostasis. Gut Microbes 5:313-325.
22. McCarthy AJ, Martin P, Cloup E, Stabler RA, Oswald E, Taylor PW. 2015. The genotoxin colibactin is a determinant of virulence in *Escherichia coli* K1 experimental neonatal systemic infection. Infect Immun 83:3704-3711.
23. Garcie C, Tronnet S, Garénaux A, McCarthy AJ, Brachmann AO, Pénary M, Houle S, Nougayrède JP, Piel J, Taylor PW, Dozois CM, Genevaux P, Oswald E, Martin P. 2016. The bacterial stress-responsive Hsp90 chaperone is required for the production of

764 the genotoxin colibactin and the siderophore yersiniabactin by *Escherichia coli*. J
765 Infect Dis 214:916-924.
766 24. Langridge GC, Phan MD, Turner DJ, Perkins TT, Parts L, Haase J, Charles I, Maskell DJ,
767 Peters SE, Dougan G, Wain J, Parkhill J, Turner AK. 2009. Simultaneous assay of every
768 *Salmonella* Typhi gene using one million transposon mutants. Genome Res 19:2308-
769 2316.
770 25. van Opijnen T, Bodi KL, Camilli A. 2009. Tn-seq: high-throughput parallel sequencing
771 for fitness and genetic interaction studies in microorganisms. Nat Methods 6:767-
772 772.
773 26. Goh KGK, Phan MD, Forde BM, Chong TM, Yin WF, Chan KG, Ulett GC, Sweet MJ,
774 Beatson SA, Schembri MA. 2017. Genome-wide discovery of genes required for
775 capsule production by uropathogenic *Escherichia coli*. mBio 8:e01558-17.
776 27. van Opijnen T, Camilli A. 2013. Transposon insertion sequencing: a new tool for
777 systems-level analysis of microorganisms. Nat Rev Microbiol 11:435-442.
778 28. Bachman MA, Breen P, Deornellas V, Mu Q, Zhao L, Wu W, Cavalcoli JD, Mobley HL.
779 2015. Genome-wide identification of *Klebsiella pneumoniae* fitness genes during lung
780 infection. mBio 6:e00775.
781 29. Wang N, Ozer EA, Mandel MJ, Hauser AR. 2014. Genome-wide identification of
782 *Acinetobacter baumannii* genes necessary for persistence in the lung. mBio
783 5:e01163-14.
784 30. Skurnik D, Roux D, Aschard H, Cattoir V, Yoder-Himes D, Lory S, Pier GB. 2013. A
785 comprehensive analysis of *in vitro* and *in vivo* genetic fitness of *Pseudomonas*
786 *aeruginosa* using high-throughput sequencing of transposon libraries. PLoS Pathog
787 9:e1003582.

31. Subashchandrabose S, Smith SN, Spurbeck RR, Kole MM, Mobley HL. 2013. Genome-wide detection of fitness genes in uropathogenic *Escherichia coli* during systemic infection. PLoS Pathog 9:e1003788.

32. Abel S, Abel zur Wiesch P, Davis BM, Waldor MK. 2015. Analysis of bottlenecks in experimental models of infection. PLoS Pathog 11:e1004823.

33. McCarthy AJ, Negus D, Martin P, Pechincha C, Oswald E, Stabler RA, Taylor PW. 2016. Pathoadaptive mutations of *Escherichia coli* K1 in experimental neonatal systemic infection. PLoS One 11:e0166793.

34. Phan MD, Peters KM, Sarkar S, Lukowski SW, Allsopp LP, Gomes Moriel D, Achard ME, Totsika M, Marshall VM, Upton M, Beatson SA, Schembri MA. 2013. The serum resistome of a globally disseminated multidrug resistant uropathogenic *Escherichia coli* clone. PLoS Genet 9:e1003834.

35. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2:2006.0008.

36. Gonzalez MD, Lichtensteiger CA, Vimr ER. 2001. Adaptation of signature-tagged mutagenesis to *Escherichia coli* K1 and the infant-rat model of invasive disease. FEMS Microbiol Lett 198:125-128.

37. Roberts IS. 1996. The biochemistry and genetics of capsular polysaccharide production in bacteria. Ann Rev Microbiol 50:285-315.

38. Chao MC, Abel S, Davis BM, Waldor MK. 2016. The design and analysis of transposon insertion sequencing experiments. Nat Rev Microbiol 14:119-128.

39. Martindale J, Stroud D, Moxon ER, Tang CM. 2000. Genetic analysis of *Escherichia coli* K1 gastrointestinal colonization. Mol Microbiol 37:1293-1305.

812 40. Anders S, Huber W. 2010. Differential expression analysis for sequence count data.
813 Genome Biol 11:R106.
814 41. Vimr ER, Steenbergen SM. 2006. Mobile contingency locus controlling *Escherichia*
815 *coli* K1 polysialic acid capsule acetylation. Mol Microbiol 60:828-837.
816 42. Fronzes R, Christie PJ, Waksman G. 2009. The structural biology of type IV secretion
817 systems. Nat Rev Microbiol 7:703-714.
818 43. Patrick M, Gray MD, Sandkvist M, Johnson TL. 2010. Type II Secretion in *Escherichia*
819 *coli*. EcoSal Plus 4: doi 10.1128/ecosalplus.4.3.4.
820 44. Ho TD, Davis BM, Ritchie JM, Waldor MK. 2008. Type 2 secretion promotes
821 enterohemorrhagic *Escherichia coli* adherence and intestinal colonization. Infect
822 Immun 76:1858-1865.
823 45. Kline KA, Dodson KW, Caparon MG, Hultgren SJ. 2010. A tale of two pili: assembly
824 and function of pili in bacteria. Trends Microbiol 18:224-232.
825 46. Pacheco AR, Curtis MM, Ritchie JM, Munera D, Waldor MK, Moreira CG, Sperandio V.
826 2012. Fucose sensing regulates bacterial intestinal colonization. Nature 492:113-117.
827 47. Weinberg ED. 2009. Iron availability and infection. Biochim Biophys Acta 1790:600-
828 605.
829 48. Serruto D, Rappuoli R, Scarselli M, Gros P, van Strijp JA. 2010. Molecular mechanisms
830 of complement evasion: learning from staphylococci and meningococci. Nat Rev
831 Microbiol 8:393-399.
832 49. Vann WF, Daines DA, Murkin AS, Tanner ME, Chaffin DO, Rubens CE, Vionnet J, Silver
833 RP. 2004.The NeuC protein of *Escherichia coli* K1 is a UDP *N*-acetylglucosamine 2-
834 epimerase. J Bacteriol 186:706-712.

50. Steenbergen SM, Lee YC, Vann WF, Vionnet J, Wright LF, Vimr ER. 2006. Separate pathways for O acetylation of polymeric and monomeric sialic acids and identification of sialyl *O*-acetyl esterase in *Escherichia coli* K1. J Bacteriol 188:6195-6206.

51. Pigeon RP, Silver RP. 1997. Analysis of the G93E mutant allele of KpsM, the membrane component of an ABC transporter involved in polysialic acid translocation in *Escherichia coli* K1. FEMS Microbiol Lett 156:217-22.

52. Bailey MJ, Hughes C, Koronakis V. 1997. RfaH and the *ops* element, components of a novel system controlling bacterial transcription elongation. Mol Microbiol 26:845-851.

53. Hu K, Artsimovitch I. 2017. A screen for *rfaH* suppressors reveals a key role for a connector region of termination factor Rho. mBio 8:e00753-17.

54. Garrett SB, Garrison-Schilling KL, Cooke JT, Pettis GS. 2016. Capsular polysaccharide production and serum survival of *Vibrio vulnificus* are dependent on antitermination control by RfaH. FEBS Lett 590:4564-4572.

55. Rowe S, Hodson N, Griffiths G, Roberts IS. 2000. Regulation of the *Escherichia coli* K5 capsule gene cluster: evidence for the roles of H-NS, BipA, and integration host factor in regulation of group 2 capsule gene clusters in pathogenic *E. coli*. J Bacteriol 182:2741-2745.

56. Krishnan K, Flower AM. 2008. Suppression of Δ*bipA* phenotypes in *Escherichia coli* by abolishment of pseudouridylation at specific sites on the 23S rRNA. J Bacteriol 190:7675-7683.

857 57. Heinrichs DE, Yethon JA, Whitfield C. 1998. Molecular basis for structural diversity in
858 the core regions of the lipopolysaccharides of *Escherichia coli* and *Salmonella*
859 *enterica*. Mol Microbiol 30:221-232.
860 58. Kalynych S, Cherney M, Bostina M, Rouiller I, Cygler M. 2015. Quaternary structure
861 of WzzB and WzzE polysaccharide copolymerases. Protein Sci 24:58-69.
862 59. Taylor PW, Kroll HP. 1985. Effect of lethal doses of complement on the functional
863 integrity of target enterobacteria. Curr Top Microbiol Immunol 121:135-158.
864 60. Dankert JR, Esser AF. 1987. Bacterial killing by complement. C9-mediated killing in
865 the absence of C5b-8. Biochem J 244:393-399.
866 61. Eckert SE, Dziva F, Chaudhuri RR, Langridge GC, Turner DJ, Pickard DJ, Maskell DJ,
867 Thomson NR, Stevens MP. 2011. Retrospective application of transposon-directed
868 insertion site sequencing to a library of signature-tagged mini-Tn*5*Km2 mutants of
869 *Escherichia coli* O157:H7 screened in cattle. J Bacteriol 193:1771-1776.
870 62. Moxon ER, Murphy PA. 1978. *Haemophilus influenzae* bacteremia and meningitis
871 resulting from survival of a single organism. Proc Natl Acad Sci USA 75:1534-1536.
872 63. Barnes PD, Bergman MA, Mecsas J, Isberg RR. 2006. *Yersinia pseudotuberculosis*
873 disseminates directly from a replicating bacterial pool in the intestine. J Exp Med
874 203:1591-1601.
875 64. Kono M, Zafar MA, Zuniga M, Roche AM, Hamaguchi S, Weiser JN. 2016. Single cell
876 bottlenecks in the pathogenesis of *Streptococcus pneumoniae*. PLoS Pathog
877 12:e1005887.
878 65. Favre-Bonte S, Joly B, Forestier C. 1999. Consequences of reduction of *Klebsiella*
879 *pneumoniae* capsule expression on interactions of this bacterium with epithelial
880 cells. Infect Immun 67:554-561.

881 66. Taylor CM, Roberts IS. 2005. Capsular polysaccharides and their role in virulence.
882 Contrib Microbiol 12:55-66.
883 67. Mao Y, Doyle MP, Chen J. 2001. Insertion mutagenesis of *wca* reduces acid and heat
884 tolerance of enterohemorrhagic *Escherichia coli* O157:H7. J Bacteriol 183:3811-3815.
885 68. Mu XQ, Savarino SJ, Bullitt E. 2008. The three-dimensional structure of CFA/I
886 adhesion pili: traveler's diarrhea bacteria hang on by a spring. J Mol Biol 376:614-
887 620.
888 69. Cleary J, Lai LC, Shaw RK, Straatman-Iwanowska A, Donnenberg MS, Frankel G,
889 Knutton S. 2004. Enteropathogenic *Escherichia coli* (EPEC) adhesion to intestinal
890 epithelial cells: role of bundle-forming pili (BFP), EspA filaments and intimin.
891 Microbiology 150:527-538.
892 70. Cherayil BJ, Ellenbogen S, Shanmugam NN. 2011. Iron and intestinal immunity. Curr
893 Opin Gastroenterol 27:523-528.
894 71. Doran KS, Banerjee A, Disson O, Lecuit M. 2013. Concepts and mechanisms: crossing
895 host barriers. Cold Spring Harb Perspect Med 3:a10090.
896 72. Falkenhagen U, Zingler G, Naumann G. 1991. Serum resistance in different serotypes
897 of *Escherichia coli*. Zentralbl Bakteriol 275:216-222.
898 73. Pluschke G, Mayden J, Achtman M, Levine RP. 1983. Role of the capsule and the O
899 antigen in resistance of O18:K1 *Escherichia coli* to complement-mediated killing.
900 Infect Immun 42:907-913.
901 74. Taylor PW. 1995. Resistance of bacteria to complement, p 49-64. *In* Roth JA, Bolin
902 CA, Brogden KA, Minion FC, Wannemuehler MJ (ed), Virulence Mechanisms of
903 Bacterial Pathogens 2nd ed. ASM Press, Washington DC.

75. Rautemaa R, Meri S. 1999. Complement-resistance mechanisms of bacteria. Microbes Infect 1:785-794.

76. Kroll HP, Bhakdi S, Taylor PW. 1983. Membrane changes induced by exposure of *Escherichia coli* to human serum. Infect Immun 42:1055-1066.

77. Taylor PW, Kroll HP. 1984. Interaction of human complement proteins with serum-sensitive and serum-resistant strains of *Escherichia coli*. Mol Immunol 21:609-620.

78. Alqasim A, Scheutz F, Zong Z, McNally A. 2014. Comparative genome analysis identifies few traits unique to the *Escherichia coli* ST131 H30Rx clade and extensive mosaicism at the capsule locus. BMC Genomics.15:830.

79. Achtman M, Mercer A, Kusecek B, Pohl A, Heuzenroeder M, Aaronson W, Sutton A, Silver RP. 1983. Six widespread bacterial clones among *Escherichia coli* K1 isolates. Infect Immun 39:315-335.

80. R J Gross, T Cheasty, B Rowe. 1977. Isolation of bacteriophages specific for the K1 polysaccharide antigen of *Escherichia coli*. J Clin Microbiol 6:548–550.

81. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci USA 97:6640-6645.

82. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120.

83. Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25.

84. Dalgakiran F, Witcomb L, McCarthy A, Birchenough GMH, Taylor PW. 2014. Non-invasive model of neuropathogenic *Escherichia coli* infection in the neonatal rat. J Vis Exp 92:e52018.

928 **LEGENDS**

929 **FIG 1.** A high-density transposon library for identification of genes essential for *in vitro*
930 growth of *E. coli* K1 A192PP. (A) Distribution of Tn*5* insertions along the *E. coli* K1 A192PP
931 genome. The number of sequence reads mapped to each single genomic location are
932 plotted to show representation of the entire genome. (B) Insertion index values for two
933 biological replicates are strongly correlated. (C) Density plot showing the frequency of
934 insertion index values for all genes. A biomodal distribution is evident, with the left peak
935 representing "essential" genes in which Tn5 insertion is lethal for growth on selective Luria-
936 Bertani agar; the left peak represents "non-essential" genes into which Tn*5* inserted without
937 induction of lethality. Green and red lines indicate gamma distributions used to estimate
938 likelihood ratios and *P* values. (D) Tn*5* insertion site reads plotted to a 9 Kb region of the *E.*
939 *coli* A192PP genome. The height of each line on the y axis indicates the number of reads at
940 each Tn*5* insertion site. The genes *lytB* and *dapB* possess no insertion sites, indicating they
941 are putative essential genes.

942 **FIG 2.** Essential *E. coli* A192PP genes in each selected KEGG (Kyoto Encyclopedia of Genes
943 and Genomes) functional orthologs (KO). Gene frequencies (light grey; expressed as % of
944 essential coding DNA sequences [CDSs] for each category) are compared to their frequency
945 within the whole genome (dark grey). KO (x/y) where y is the number of CDSs in the whole
946 genome and x is the number of identified essential genes.

947 **FIG 3**. Culture of the *E. coli* K1 A192PP-Tn*5* library results in loss of population diversity and
948 enrichment of non-encapsulated mutants. (A) Changes in the proportion of *E. coli*
949 encapsulated and non-capsulated A192PP bacteria during culture of the *E. coli* A192PP-Tn*5*
950 library in LB media at 37^{o}C (200 orbits/min) (n = 3; ±1SD; Student's t, $p < 0.01$). CFU of

951 encapsulated and non-encapsulated bacteria were determined from the proportion of
952 bacteria susceptible to the K1E bacteriophage. (B) Survival of P2 rats colonized with *E.*
953 *coli* K1 A192PP, the uncultured *E. coli* A192PP-Tn*5* library and the cultured (LB broth; 8 h;
954 37^{o}C) *E. coli* A192PP-Tn*5* library. Pups (*n* = 12 for each group) were colonized with 2-4 x 10^{6}
955 CFU by the oral route. Log-rank [Mantel-Cox] to compare survival of the cultured library
956 with wildtype strain and the uncultured library: ns, non-significant, * $P < 0.05$, ** $P < 0.01$.

957 **FIG 4.** Identification using a high-density transposon library of genes promoting GI
958 colonization of *E. coli* A192PP in the neonatal rat. (A) Colonization of PSI, MSI, DSI and colon
959 after oral administration of 2-6x10^{6} CFU *E. coli* K1 A192PP to P2 pups. (B) Log_2-fold change
960 and average Tn*5* insertion site read abundance of each gene after MSI colonization of P2 rat
961 pups (*n* = 4) over a 4 h stabilization period expressed as MA-plot. An inoculum containing 2 x
962 10^{4} unique *E. coli* K1 A192PP-Tn*5* mutants was prepared and 1 x 10^{9} CFU administered
963 orally. *E. coli* colonies (2 x 10^{5}) were recovered from the inoculum (input pool) and from MSI
964 homogenates (output pool) by culture on to LB agar containing 50 µg/ml kanamycin. Red
965 data points represent Tn*5* insertion sites determined as differentially expressed in the
966 output pool compared to the input pool using a negative binomial test with a false discovery
967 rate of 0.1. (C) Mutations in 167 genes significantly decreased fitness for colonization of the
968 MSI and encoded proteins with a range of functions. (D) Colonization of P2 rat intestine by
969 *E. coli* K1 A192PP and single gene mutants. Bacteria (2-4 x 10^{6}) were administered orally to
970 P2 rats (*n* = 12/group). Pups were sacrificed and the *E. coli* K1 burden in intestinal sections
971 enumerated 24 h after initiation of colonization. Parent and mutant strain CFU values were
972 compared using Student's *t*-test: * $P < 0.05$, ** $P < 0.01$. (E) Survival of P2 rats colonized with
973 *E. coli* K1 A192PP and single gene mutants. Bacteria (2-4 x10^{6}) were administered orally to

974 P2 rats (*n* = 12/group). Log-rank [Mantel-Cox] test: ns, non-significant, * $P < 0.05$, ** $P <$
975 0.01.

976 **FIG 5.** Bottleneck to systemic infection in the neonatal rat. (A) Survival of rats colonized at
977 P2 by oral administration of *E. coli* K1 A192PP or A192PPΔ*lacZ*::*kan*. *n* = 12 pups for both
978 groups. Log-rank [Mantel-Cox] test: ns, non-significant, * $P < 0.05$, ** $P < 0.01$. (B)
979 Competitive indices of intestinal colonization and gut-to-blood transit of *E. coli* K1 A192PP
980 and A192PPΔ*lacZ*::*kan*. A 1:1 mixture of *E. coli* K1 A192PP and A192PPΔ*lacZ*::*kan* (total 2-4
981 x10^6 CFU) was administered orally to P2 pups. After 24 h, ratios of A192PP and
982 A192PPΔ*lacZ*::*kan* were enumerated in segmented GI tissues and in the blood as indicated
983 using selective media. Animals in which only *E. coli* K1 A192PP or A192PPΔ*lacZ*::*kan* were
984 detected in the blood are coloured red and blue respectively, indicating the existence of a
985 bottleneck to infection. (C) Loss of diversity of *E. coli* K1 A192PP-Tn*5* populations recovered
986 from the blood (red) and brain (blue) following translocation from the GI tract (black).

987 **FIG 6.** Identification using a high-density transposon library of genes contributing to the
988 complement resistance of *E. coli* A192PP. (A) Survival of *E. coli* A192PP and *E. coli* K12 strain
989 MG1655 in 22% pooled human serum. The latter was used a serum susceptible control; *n* =
990 3, error bars represent range of values. (B) Log_2-fold change and average Tn*5* insertion site
991 read abundance of each gene after incubation of 1 x 10^6 CFU containing 2 x 10^4 unique *E.*
992 *coli* K1 A192PP-Tn*5* mutants in 22% pooled human serum for 3 h at 37°C. Colonies (2 x 10^5)
993 were obtained by culture of diluted aliquots on LB agar containing 50 µg/ml kanamycin. The
994 inoculum served as the input pool. Red data points represent Tn*5* insertion sites determined
995 as differentially expressed in the output pool compared to the input pool using a negative
996 binomial test with a false discovery rate of 0.1. (C) Survival of 1 x10^6 *E. coli* K1 A192PP and

single gene mutants in 22% normal human and heat-inactivated (56^{o}C; 30 min) serum. Final volume of the reaction mixture was 1.5 ml; *n* = 3, error bars represent range of values. Complementation with the functional gene restored resistance in all cases. (D) Survival of P2 rats colonized with *E. coli* K1 A192PP and single gene mutants. Bacteria (2-4 x10^{6}) were administered orally to P2 rats (*n* = 12/group). Log-rank [Mantel-Cox] test: ns, non-significant, * $P < 0.05$, ** $P < 0.01$.

1020 **TABLE 1.** Strains used in this study

Strain	**Description**
E. coli K1	
A192PP	018:K1:H7; virulent in neonatal rat model of infection
A192PPΔ*lacZ*::*kan*	*lacZ* mutant of A192PP; Kan^r
A192PPΔ*neuC*::*kan*	*neuC* mutant of A192PP; Kan^r
A192PPΔ*neuC*::*kan* + pUC19.*neuC*	Complemented A192PPΔ*neuC*::*kan*; Kan^r Amp^r
A192PPΔ*rfaH*::*kan*	*rfaH* mutant of A192PP; Kan^r
A192PPΔ*rfaH*::*kan* + pUC19.*rfaH*	Complemented A192PPΔ*rfaH*::*kan*; Kan^r Amp^r
A192PPΔ*traL*::*kan*	*traL* mutant of A192PP; Kan^r
A192PPΔ*traL*::*kan* + pUC19.*traL*	Complemented A192PPΔ*traL*::*kan*; Kan^r Amp^r
A192PPΔ*vasL*::*kan*	*vasL* mutant of A192PP; Kan^r
A192PPΔ*waaW*::*kan*	*waaW* mutant of A192PP; Kan^r
A192PPΔ*waaW*::*kan* + pUC19.*waaW*	Complemented A192PPΔ*waaW*::*kan*; Kan^r Amp^r
A192PPΔ*yaeQ*:*kan*	*yaeQ* mutant of A192PP; Kan^r
A192PPΔ0678::*kan*	0678 mutant of A192PP; Kan^r
A192PPΔ3010::*kan*	3010 mutant of A192PP; Kan^r
E. coli K12	
MG1655	F- lambda- *ilvG- rfb*-50 *rph*-1
MG1655Δ*lacZ*::*kan*	*lacZ* mutant of MG1655; Kan^r

1021