

Kent Academic Repository

Full text document (pdf)

Citation for published version

Dadswell, Katherine and Creagh, Sinead and McCullagh, Edward and Liang, Mingzhi and Brown, Ian R. and Warren, Martin J. and McNally, Alan and MacSharry, John and Prentice, Michael B. (2019) Bacterial microcompartment-mediated ethanolamine metabolism in *E. coli* urinary tract infection. *Infection and Immunity* . ISSN 0019-9567.

DOI

<https://doi.org/10.1128/IAI.00211-19>

Link to record in KAR

<https://kar.kent.ac.uk/74189/>

Document Version

Author's Accepted Manuscript

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:

researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

1 **Bacterial microcompartment-mediated ethanolamine metabolism in *E. coli* urinary**
2 **tract infection**

3

4

5

6 Katherine Dadswell¹, Sinead Creagh², Edward McCullagh², Mingzhi Liang³, Ian R. Brown³,

7 Martin J Warren³, Alan McNally⁴, John MacSharry^{1,5*}, Michael B. Prentice^{1,2,5,6*}

8

9

10 ¹ School of Microbiology, University College Cork, Cork, Ireland

11 ² Department of Microbiology, Cork University Hospital, Cork, Ireland

12 ³ School of Biosciences, University of Kent, Canterbury, UK

13 ⁴ Institute of Microbiology and Infection, University of Birmingham, Birmingham, UK

14 ⁵ APC Microbiome Ireland, University College Cork, Cork, Ireland

15 ⁶ Department of Pathology, University College Cork, Cork, Ireland

16

17

18 *Correspondence: Dr John MacSharry and Professor Michael B Prentice, School of

19 Microbiology, University College Cork, Cork, Ireland

20 Email: J.MacSharry@ucc.ie, m.prentice@ucc.ie,

21

22 Keywords: Microcompartment, metabolosome, urinary tract infection, *E. coli*, ethanolamine

23

24

25 **Abstract**

26 Urinary tract infections (UTIs) are common, in general caused by intestinal Uropathogenic *E.*
27 *coli* (UPEC) ascending via the urethra. Microcompartment-mediated catabolism of
28 ethanolamine, a host cell breakdown product, fuels competitive overgrowth of intestinal *E.*
29 *coli*, both pathogenic enterohaemorrhagic *E. coli* and commensal strains. During UTI urease
30 negative *E. coli* thrive, despite the comparative nutrient limitation in urine. The role of
31 ethanolamine as a potential nutrient source during UTI is understudied. We evaluated the role
32 of metabolism of ethanolamine as a potential nitrogen and carbon source for UPEC in the
33 urinary tract. We analysed infected urine samples by culture, HPLC, qRT-PCR and genomic
34 sequencing. Ethanolamine concentration in urine was comparable to the most abundant
35 reported urinary amino acid D-serine. Transcription of the *eut* operon was detected in the
36 majority of urine samples screened containing *E. coli*. All sequenced UPECs had conserved
37 *eut* operons while metabolic genotypes previously associated with UTI (*dsdCXA*, *metE*) were
38 mainly limited to phylogroup B2. *In vitro* ethanolamine was found to be utilised as a sole
39 source of nitrogen by UPECs. Metabolism of ethanolamine in artificial urine medium (AUM)
40 induced metabolosome formation and provided a growth advantage at the physiological
41 levels found in urine. Interestingly, *eutE* (acetaldehyde dehydrogenase) was required for
42 UPECs to utilise ethanolamine to gain a growth advantage in AUM, suggesting ethanolamine
43 is also utilised as a carbon source. This data suggests urinary ethanolamine is a significant
44 additional carbon and nitrogen source for infecting *E. coli*.

45

46 **Introduction**

47 Urinary tract infection is a common condition with an estimated 150 million episodes
48 globally per annum (1). The most common identified cause is infection by uropathogenic
49 *Escherichia coli* (UPEC) strains (2, 3). The currently accepted paradigm for uncomplicated
50 urinary tract infection is that these *E. coli* strains residing in the gut as commensals
51 successively colonise the perineum (4), the urethra and then the bladder, where the
52 production of bacterial toxins and the host immune response lead to tissue damage and
53 symptoms such as frequency and dysuria (2). Further ascending infection to colonise the
54 kidney with more local tissue damage causing pyelonephritis and bacteraemia occurs in a
55 small percentage of cases.

56

57 Common genetic features have been noted in a variety of *E. coli* strains causing infections
58 outside the gastrointestinal tract, including UPEC, and these are collectively termed ExPEC
59 (extraintestinal pathogenic isolates of *E. coli*) (5, 6). Panels of genes whose presence is
60 associated with any *E. coli* infection outside the gastrointestinal tract (7), or specifically
61 urinary tract infection (8), have been assembled by genetic comparison of *E. coli* strains
62 isolated from the gut with those isolated from urine and other extraintestinal sites and those
63 known to be virulent in different animal models. However, the mechanism by which these
64 factors are involved in pathogenicity is obscure.

65

66 In the pathogenesis of *E. coli* urinary tract infection rapid invasion of bladder cells occurs
67 with formation of intracellular bacterial communities (IBCs) with biofilm-like properties
68 which initiate the infective process (9, 10). This bottleneck reduces diversity and has
69 prevented global searches by signature tagged mutagenesis for key genetic factors required
70 for infection (11). Assessing genome-sequenced clinical *E. coli* urinary isolates in a mouse

71 model of urinary tract infection showed that no set of genes was predictive of virulence in the
72 model (12), including genes previously specifically associated with urovirulence.

73

74 Rapid growth has been shown to be characteristic of early phase *E. coli* infection in the
75 urinary tract (13), suggesting securing nutrition in the urinary tract is a key part of *E. coli*
76 pathogenesis. *E. coli* requirements for central carbon metabolism in the urinary tract have
77 been explored by competition studies with selected mutants in murine models. Interruption of
78 gluconeogenesis (*pckA*) or the TCA cycle (*sdhB*) reduces fitness of *E. coli* to infect (14). This
79 is in contrast to the nutrient rich intestine, where glycolysis (*pgi*) or the Entner-Doudoroff
80 (*edd*) pathway are required for colonisation fitness (15).

81

82 Some metabolic loci have been linked to UPEC pathogenesis. D-serine is an abundant amino
83 acid in human urine, present at a mean concentration of 0.12 mM (16), and up to 1mM in
84 some cases (17), much higher than intestinal content levels. Some *E. coli* strains can
85 metabolise D-serine to pyruvate and ammonia (18), allowing it to be a sole carbon and
86 nitrogen source in vitro (19). This is conferred by possession of a complete D-serine
87 tolerance locus (*dsdCXA*) (20), where *dsdC* encodes a LysR-type transcriptional regulator
88 (LTTR), *dsdX* a D-serine transporter (21) and *dsdA* a D-serine dehydratase. ExPEC strains
89 usually encode a full *dsdCXA* locus, while enteric pathogenic *E. coli* frequently have a
90 truncation after *dsdC* (22). In the absence or truncation of this locus, D-serine shows
91 reversible toxicity for *E. coli* causing growth arrest at concentrations of 0.1 mM and above *in*
92 *vitro* (23).

93 A metabolic regulatory polymorphism has been associated with cobalamin-independent
94 methionine synthase (MetE) in UPEC. A promoter polymorphism (*sra* or short regulatory

95 allele) upstream of the *metE* gene in these strains is associated with increased *metE* induction
96 and enhanced ability to grow in urine *in vitro* (24).

97

98 Mutational analysis of a subset of *E. coli* genes showing a marked (>fourfold) increase in
99 transcription in infected patient urine compared to growth in urine or Luria Broth (LB) (25)
100 showed that that their knockout caused a fitness defect in the urinary bladder in a mouse
101 model of ascending urinary infection. The most marked defects were with knockout of the
102 *cus* (copper resistance) and *eut* (ethanolamine uptake and metabolism) operons.

103

104 The *eut* operon is part of the conserved *E. coli* core genome (26) having arrived in
105 Enterobacterales by horizontal transfer (27). It contains seventeen genes including the
106 positive transcriptional regulator *eutR*. The operon encodes enzymes required for
107 ethanolamine metabolism and includes structural shell protein genes for the synthesis of thin
108 porous protein shells enclosing the enzymes as bacterial microcompartments
109 (metabolosomes) in the cytoplasm (28–30) (Fig 1A). Experiments largely conducted with
110 *Salmonella enterica* (which contains the same operon) suggest that the enzymic breakdown
111 of ethanolamine to ammonia (a nitrogen source) and acetaldehyde occurs within the
112 metabolosome, with the toxic effects and evaporative loss of acetaldehyde minimised by
113 microcompartment enclosure and onward metabolism to ethanol and acetyl-CoA (a carbon
114 source)(30, 31). Some acetyl-CoA is further metabolised to acetyl phosphate and acetate
115 within the metabolosome, and some is available to enter central metabolism (32).

116 Ethanolamine in the gastrointestinal tract utilised by this pathway gives a competitive
117 advantage to Enterohaemorrhagic *E. coli* (33) and *Salmonella enteritidis* (34). Recently it has
118 been shown that *E. coli* ethanolamine metabolism is essential for bladder colonisation in a
119 murine model of ascending UTI (35). The mechanism was suggested to involve resistance to

120 innate immunity because the colonisation advantage of wild type UPEC over a $\Delta eutR$ mutant
121 was abolished in neutrophil-depleted mice. Clearance of an isogenic $\Delta eutR$ mutant *E. coli*
122 from the bladder coincided with peaking myeloperoxidase levels. However, resistance to
123 hydrogen peroxide was unchanged in the $\Delta eutR$ mutant.

124

125 In this study we evaluated the role of microcompartment-mediated ethanolamine metabolism
126 in clinically infected urine samples and in laboratory cultures of *E. coli* strains isolated from
127 infected urine. The *eut* operon was induced in infected urine, and ethanolamine was present
128 in urine at a level that enhanced *E. coli* growth *in vitro*. Metabolosomes were visible on
129 TEM in a UPEC strain grown with ethanolamine. Inactivation of the *eut* operon reduced
130 growth of a UPEC strain in ethanolamine-containing nitrogen-limited minimal medium and
131 growth and competitiveness in ethanolamine-containing artificial urine medium. Selective
132 mutation of individual *eut* genes suggested that ethanolamine provided a carbon source in this
133 artificial urine medium. In summary, we have identified that microcompartment-mediated
134 metabolism of ethanolamine present in urine can give *E. coli* a growth advantage by
135 providing an additional carbon and nitrogen source.

136

137 **Methods**

138

139 **Bacterial strains and culture conditions**

140 Clinically infected urine samples received at Cork University Hospital (CUH) containing
141 visible bacteria and white cells were selected and anonymised. The protocol was approved by
142 the Clinical Research Ethics Committee of the Cork Teaching Hospitals ref ECM 4 (c)
143 12/08/14. A further 12 specimens of macroscopically clear urine with no bacteria or white
144 cells were selected as controls. Following initial culture on CLED agar pure colonies

145 subcultured on Columbia blood agar were identified by MALDI-TOF using a Microflex LT
146 mass spectrometer (Bruker Daltonik) and the MALDI Biotyper software package (version
147 3.0). Antimicrobial sensitivity was determined by the VITEK® 2.0 system (Biomérieux)
148 using EUCAST breakpoints. Strains used for gene inactivation or competitive growth assays
149 are listed in Table 1. Sixty-one *E. coli* strains were isolated, and whole genome sequences
150 obtained for 47 strains.

151

152 *E. coli* were routinely cultured in LB broth at 30 °C or 37 °C with aeration. To determine the
153 ability to utilise ethanolamine, strains were cultured at 37 °C in modified M9 minimal media
154 (33) containing 10 mM ethanolamine hydrochloride and 200 nM cobalamin with the addition
155 of either 20mM glycerol or 20 mM ammonium chloride. Automated growth count cultures
156 were incubated in 96-well plates in triplicate and OD₆₀₀ measured using a Biotek Eon
157 Microplate Spectrophotometer over 48 hours. Manual growth curves were measured in 35 ml
158 volumes with spectrophotometric analysis of 1 ml aliquots.

159

160 **Competition experiments** were carried out in a published liquid artificial urine medium
161 (AUM) (36) and with the same medium with added ethanolamine hydrochloride at 0.5 mM,
162 and 10 mM, with cell counts on LB agar. Pre-cultured *E. coli* strains were incubated in LB
163 with antibiotics where appropriate. The cultures were washed in PBS and resuspended in
164 AUM. Approximately equal concentrations of the wild type and isogenic mutant were used to
165 inoculate AUM with ethanolamine as indicated in the text to give an approximate starting
166 OD₆₀₀ of 0.1. The co-cultures were incubated at 37°C with aeration and at each time point
167 the co-culture were diluted 10-fold in PBS and plated on to LB agar. The dilutions were
168 plated onto LB agar and onto LB agar containing kanamycin to determine the concentration
169 of each strain of *E. coli*. The plates were incubated overnight at 37°C and the CFU calculated.

170 The wild type CFU was calculated by subtracting the number of CFU resistant to kanamycin
171 from the number of CFU on LB agar plates. The experiment was repeated three times and a
172 competitive index was calculated as follows:

173

$$CI = \frac{\text{eut mutant CFU recovered/Wild Type CFU recovered}}{\text{eut mutant CFU inoculum/Wild type CFU inoculum}}$$

174

175 A competitive index below 1 indicates that the wild type was outcompeting the mutant strain
176 at that time point. The CI at time zero is by definition 1.0 . Growth of *eut* operon mutants was
177 compared with wild-type strains in M9 minimal medium with 0.5 mM and 10 mM
178 ethanolamine and AUM with 10 mM ethanolamine.

179

180 **Mutants**

181 To generate deletion mutants, BW25113 knockout *E. coli* strains for the genes of interest
182 were obtained from the Keio collection (37). Mutations were transferred to UPEC strain U1
183 by P1 vir phage transduction (38) In brief, lysogen strains were prepared by incubating P1
184 lysate with the donor strain for 30 minutes at 30°C with 5µl of 1M CaCl₂ and the culture
185 were plated on kanamycin selective agar. The resulting colonies were used to prepare the
186 lysate for transduction. Lysogen colonies were grown overnight in 2ml of LB at 30°C. The
187 precultures were used to inoculate LB and grown until reaching an OD₆₀₀ of 0.2. The cultures
188 were incubated in 46°C for 20 minutes with shaking before being moved to 37°C until
189 complete lysis. Bacteria were centrifuged out of the culture and the supernatant was stored
190 with chloroform to prevent bacterial growth. Overnight cultures of the recipient strain were
191 resuspended in transduction buffer (10mM MgSO₄, 5mM CaCl₂) and 100µl of cells were
192 incubated with lysate and incubated at 37°C for 30 minutes. Sodium citrate was added
193 following this incubation and for a further hour. The cells were washed in LB before being

194 plated onto LB agar plates. Strains were selected for kanamycin resistance and transductants
195 were confirmed by genome sequencing and PCR using primers internal to the kanamycin
196 gene and upstream and downstream of the disrupted gene (Supplementary Data Table S1).
197 Complementation was with *E. coli* K-12 genes cloned in pCA24N from the ASKA library
198 (39) induced by 0.01mM IPTG.

199

200 **Metabolic assays**

201 After culture, residual urine samples were separated into cell fraction and cell free
202 supernatant by differential centrifugation and urine supernatants were filtered with a 0.2µm
203 membrane to remove any remaining bacteria and stored at -80°C. Urine supernatants and
204 culture supernatants were assayed for ethanolamine, acetate and ethanol by HPLC using an
205 Agilent 1200 HPLC system with a refractive index detector. Urines collected from CUH and
206 bacterial culture supernatants were filter sterilised at 0.2µm to remove bacteria before being
207 stored at -80°C until the day of experimentation. Ethanolamine was measured by gradient
208 HPLC after derivatisation with o-phthaldialdehyde (OPA) using a method adapted from
209 Sturms et al. (40). The mobile phase consisted of Buffer A [10% methanol (Sigma-Aldrich) -
210 90% 10mM Na₃PO₄ (pH7.3) (Sigma-Aldrich)], and Buffer B [80% Methanol- 20% 10mM
211 Na₃PO₄ (pH7.3)]. Samples were prepared using an in-loop derivatization reaction where 6µl
212 of sample were taken up followed by 6µl 10mg/ml OPA and 3-mercaptopropionic acid in
213 0.4M boric acid (Agilent Technologies) and incubated at room temperature for 3 minutes.
214 The samples were injected into a 4.6 by 100mm, 2.7µm pore Infinity Lab Poroshell HPH-
215 C18 column (Agilent Technologies) and eluted with 5ml linear gradient from 50% Buffer B
216 to 100% Buffer B followed by 5mls of 100% Buffer B at as constant flow rate of 1ml min⁻¹.
217 The excitation was detected at 224nm. A standard curve was created before each sequence

218 run. Identification of the peak and quantification was determined by comparison to retention
219 time and standard curve.

220 Acetate and ethanol were measured by the same HPLC system. 10 μ l of sample was injected
221 into a REZEX 8 μ m 8% H, Organic Acid Column (Phenomenex, USA) and eluted with 15ml
222 of 0.01M H₂SO₄ at a flow rate of 0.6ml min⁻¹. The column was maintained at 65°C for the
223 duration of the experiments. The identification of the substrate was determined by
224 comparison of retention time to pure compound and concentrations were quantified by
225 comparison to known standards.

226

227 **Transmission Electron Microscopy (TEM)**

228 This was carried out as previously described (41) After growth (as indicated in the text)
229 bacteria cells were pelleted by centrifugation, to give a pellet no larger than 100 μ l in volume.
230 The bacterial pellet was fixed in 2ml of 2.5% glutaraldehyde (Fluka) diluted in 0.1M Sodium
231 cacodylate pH 6.8 (CAB) (bioWORLD). After incubation overnight at 4°C, bacteria were
232 washed twice with 0.1 M CAB and suspended in fresh 2ml of 2.5% glutaraldehyde diluted in
233 CAB. The bacteria were stained for 1 hour in 1 ml of 1% osmium tetroxide (w/v) (250 μ l 4%
234 osmium tetroxide; 250 μ l Milli-Q H₂O; 500 μ l 0.2 M CAB). The pellets were washed in 2ml
235 Milli-Q H₂O for 10 minutes twice before the pellets were dehydrated. Pellets were
236 dehydrated through an ethanol (EtOH) gradient as follows: 50% EtOH (v/v) x 10mins; 70%
237 EtOH x 10min; 90% EtOH x 10mins; 100% EtOH x 10mins three times and then the
238 bacterial pellets were washed twice in propylene oxide for 10mins. The pellets were
239 embedded into 1.5 ml propylene oxide: LV resin at 1:1 for 30 min followed by incubation
240 2 \times 1.5 h in 100% freshly made agar LV resin. The pellets were resuspended in 1ml of 100%
241 LV resin and transferred to a conical bottom tube. The bacterial pellet was centrifuged at
242 1100xg for 5mins and was left to incubate at 60°C for 24 hours. Bacteria were sectioned to

243 60-70 nm with a diamond knife on a LEICA-EM-UC7 ultramicrotome. Sections were
244 collected on 400 mesh copper grids and stained with 4.5% (w/v) uranyl acetate in 1% acetic
245 acid (v/v) for 45mins and Reynolds lead citrate for 7 mins at room temperature. Sections
246 were then observed on a Jeol 1230 transmission electron microscope operated at an
247 accelerating voltage of 80kV and imaged with a Gatan OneView digital camera.

248

249 **DNA sequencing and sequence analysis, statistics**

250 DNA was extracted from overnight cultures in LB and extracted using Qiagen DNEasy
251 Blood and Tissue (Qiagen) with RNase A treatment (Sigma). Bacterial genome sequencing
252 was carried out by MicrobesNG (see acknowledgements) using Illumina HiSeq 2500
253 2x250bp paired-end reads. Reference genomes were identified using Kraken (42) and reads
254 mapped using BWA-MEM (43). De novo read assembly was achieved using SPAdes (44),
255 with read mapping back to the resultant contigs, using BWA-MEM for quality metrics.
256 Automated annotation was performed using Prokka (45). Sequencing data are available for
257 download from the EBI European Nucleotide Archive under BioProject accessions
258 PRJEB31941, PRJEB31942, PRJEB31943, and PRJEB31944.

259

260 Phylogenetic trees were generated from contig sequences with Parsnp (Harvest tool suite
261 (46)) and edited with ITOL (47). Parsnp produces a core genome alignment and identifies
262 SNPs for tree generation by FastTree2 (48) using SH-like (Shimodaira-Hasegawa) local
263 supports for bootstrapping. Alignment with 32 reference genomes known to be representative
264 of six *E. coli* phylogroups (49) was used for phylogroup assignment. Gene presence in
265 genomes was taken as >75% identity in BLASTN search over the full reference gene
266 sequence length. Binary matrices were prepared representing sequenced genomes with PUF
267 gene presence scored as 1 and absence as zero, and phenotypic antimicrobial resistance

268 scored as 1 and sensitivity as zero. Two-dimensional cluster analysis on these matrices was
269 performed with the R software package using complete linkage clustering on the Jaccard
270 Distance. The resulting cladograms and heat maps were visualised with ITOL (47). All other
271 statistical analyses presented were generated with GraphPad Prism 7.

272

273 **RNA and RT-PCR**

274 RNA was extracted from bacterial pellets using the Zymo Fungal/Bacterial Mini Prep kit and
275 from Eukaryotic cells using the Quick-RNA MiniPrep kit, following the manufacturer's
276 instructions. After extraction genomic DNA was digested using the TURBO DNA-free
277 (Ambion) DNase 1 treatment. The RNA was quantified using a Nanodrop 1000
278 spectrophotometer. cDNA was synthesised by reverse transcription carried out in nuclease
279 free 96-well plates. RNA was diluted using molecular grade H₂O (Sigma-Aldrich) to a final
280 concentration from of 100ng μl^{-1} in a 10 μl volume. The RNA was mixed with: cDNA
281 reaction was set up 4 μl 5x Reverse transcription buffer (Roche); 3 μl Random Hexamer
282 Primer (Roche); 2 μl 20mM dNTPs mix; and 1 μl Reverse transcriptase/RNase Inhibitor to
283 give a total volume of 20 μl . The reaction mixture was incubated in a thermocycler in the
284 following condition: 10mins at 25 °C; 30mins 55 °C; 5mins at 85 °C; hold at 4 °C. The
285 cDNA was then diluted to 100 μl and stored at -20 °C until use.

286

287 The universal probe library (Roche, Indianapolis USA) was utilized to design primers for
288 quantitative PCR. The primers used in this study are listed in Data Supplement Table S1
289 Amplification reactions were a mix of: 3 μl of cDNA; 7 μl TaqMan Probe Master buffer
290 (Roche); 1 μl 20mM primer mix (L+R primers); 0.1 μl probe(Roche);and 0.9 μl molecular
291 grade H₂O to a make a final volume of 10 μl . When the probe was not available a SYBR
292 Green master mix was used which included: 3 μl cDNA, 5 μl 2xSYBR Green I Master buffer

293 (Roche); 1 μ l 20mM primer mix(L+R primers) and 1 μ l molecular grade H₂O to a final
294 volume of 10 μ l. All reactions were performed using a 384 well plate on the
295 LightCycler®480 System (Roche) with molecular grade water included as a negative control.
296 Thermal cycling condition were as follows: 50°C for 2 mins, 95°C for 10 mins followed by
297 45 cycles for 95°C for 10s, 60°C for 45 s and 72°C for 60 s. Relative gene expression was
298 calculated using the 2- $\Delta\Delta$ Ct (50). X-fold changes in mRNA of target gene was quantified
299 relative to *gyrA*.

300

301 **ELISA**

302 Frozen urines were analysed using Meso Scale Discovery (MSD) V-PLEX proinflammatory
303 panel I and Cytokine Panel II (MSD, Rockville, MD) enzyme-linked immunosorbent assays
304 (ELISAs). Assays were performed according to the manufacturer's instructions and measured
305 using MESO QuickPlex SQ120. Calibrators were run in duplicate with the urines and used to
306 form a standard curve. The concentration of cytokines in the urine were extrapolated from the
307 standard curve. Values which fell below the limits of detection were excluded from statistical
308 analysis.

309

310 **Results**

311 **Ethanolamine is present in urine and infecting *E. coli* strains show *eut* operon induction**

312 One hundred and three clinically infected urine samples were selected from which 61 *E. coli*
313 strains were isolated, 47 of which were sequenced and used for *in vitro* metabolic analysis.
314 The mean concentration of ethanolamine in 54 clinically infected urine samples was 0.55
315 mM (mean \pm 0.076) and 0.66 mM (mean \pm 0.155) in 12 control urine samples which were not
316 clinically infected (contained no white cells or bacteria on microscopy) (Figure 1B). The
317 difference between infected and control urines was not significant. In 24 *E. coli* infected

318 urine samples from which RNA was extracted, transcription of *eut* operon genes was detected
319 in the majority of cases for *eutB* (88%) *eutS* (68%) and *eutR* (63%). Expression of *eutB*
320 significantly correlated with the ethanolamine concentration in urine (Fig. 1C). Because of
321 anonymisation individual patient details are not available. Audit of diagnostic urine
322 specimens in our laboratory shows that 75% come from general practice, 25 % from hospital
323 sources, and 75% overall from women.

324

325 **Clinically infected urine samples show stimulation of the host innate immune response .**

326 Cytokines IL-8 and IL-1 β were detected in 81% of clinically infected urine samples and
327 significantly increased in infected urines compared to non-infected urine (IL-1 β P=0.0048,
328 and IL-8 P<0.001, see Fig. S1 Supplemental material). Mean IL-6 levels were higher in
329 infected urine than in non-infected urine but the difference was not significant (Data
330 Supplement Fig. S1).

331

332 **Uropathogenic *E. coli* strains utilise ethanolamine *in vitro* resulting in enhanced growth,
333 formation of bacterial microcompartments, and production of acetate and ethanol**

334 Forty-five out of 47 (96%) *E. coli* strains isolated from urine showed increased overnight
335 growth with 10 mM ethanolamine as the sole nitrogen source in M9 minimal medium (Data
336 Supplement Fig. S2). No increased growth was detected with 10 mM ethanolamine as a sole
337 carbon source in M9 for four strains shown to actively metabolise ethanolamine as a nitrogen
338 source (Data Supplement Fig.S2, S3). For these selected strains (U1, U13, U17, U38) growth
339 in M9 medium with ethanolamine containing glycerol as a carbon source commenced after
340 10 hours (Fig. 2A) with ethanolamine consumption from around eight hours (Fig. 2C).
341 Addition of 10 mM ethanolamine to artificial urine medium (AUM) also increased growth of
342 these strains (Fig. 2B) with consumption of ethanolamine from around four hours incubation

343 onwards (Fig. 2D). Acetate and ethanol were produced by *E. coli* U1 growth in both M9 and
344 AUM media when ethanolamine was added (Fig. 2C,D) and corresponded with induction of
345 the *eut* operon at 4 and 8 hours of incubation with ethanolamine in AUM (Data Supplement
346 Fig. S4). TEM of *E. coli* U1 grown in AUM with added ethanolamine showed 100-130 nm
347 cytoplasmic inclusions with straight edges (Fig. 3A) in the majority of cells visualised (43/69
348 = 62%). These structures are typical of bacterial microcompartments. They were not
349 observed in cells grown in the absence of ethanolamine (Fig. 3B) and were seen in a minority
350 of cells grown in minimal medium with ethanolamine (Supplementary data Fig. S5). The
351 difference in TEM appearances between M9 and AUM medium may be growth phase-
352 related. Cells were collected for TEM at 8 hours incubation which is approximately the
353 starting time for ethanolamine consumption in M9 minimal medium, but the time of most
354 rapid consumption in AUM (Fig. 2). Acetate was detected in nearly all infected urine samples
355 tested (Supplementary Data Fig. S8).

356

357 **The effect of inactivation of individual enzyme-encoding genes in the *eut* operon**
358 **suggests ethanolamine growth stimulation in artificial urine medium is due to provision**
359 **of an additional carbon source**

360

361 Mutation of the *eut* operon genes *eutB* and *eutE* was achieved in strain U1 (Table 1). *eutB*
362 encodes the heavy chain component of ethanolamine ammonia lyase required to liberate
363 ammonia from ethanolamine, and *eutE* encodes a reversible acetaldehyde dehydrogenase,
364 acting after *eutBC* in the ethanolamine catabolism pathway (see schematic, Fig. 1A). EutE is
365 required to generate acetyl-CoA, which is the route for carbon assimilation from
366 ethanolamine (Fig. 1A).

367 Growth stimulation in nitrogen-limited minimal (M9) medium by addition of ethanolamine
368 (0.5 mM or 10 mM) was abolished by deletion of *eutB* in U1 and retained after deletion of
369 *eutE* (Fig. 4A, 4B, Data Supplement Fig. S6, Table 2). RT-PCR showed that ethanolamine
370 induced *eutE* transcription in the *eutB* mutant and vice versa, demonstrating that these were
371 not polar mutations (Supplementary Data Fig. S4). Ammonia generation from ethanolamine
372 alone is therefore sufficient to stimulate *E. coli* U1 growth in nitrogen-limited minimal (M9)
373 medium. Complementation of the *eutB* mutant restored the wild-type phenotype in
374 ethanolamine-containing minimal medium (Fig. 4A).

375 In contrast to this phenotype in nitrogen-limited minimal (M9) medium, in AUM medium
376 which contains 25 mM ammonium chloride and no glycerol as carbon source, growth
377 stimulation by ethanolamine was absent in U1 Δ *eutE*, although ethanolamine was still
378 metabolised by this strain (Fig. 4D,E, Table 2). Growth enhancement by ethanolamine in
379 AUM was restored by *eutE* complementation. Therefore, in AUM, unlike nitrogen limited
380 M9, the growth stimulation conferred by ethanolamine metabolism is not due to ammonia
381 generation, but appears to be caused by the provision of an additional carbon source from
382 acetyl-CoA.

383 **A functional *eut* operon is essential for competitive growth of a UPEC strain in the**
384 **presence of ethanolamine *in vitro***

385 Competitive growth assays in AUM containing 10 mM ethanolamine between wild type *E.*
386 *coli* strain U1 and Δ *eutB* and Δ *eutE* mutants showed a significant advantage for the wild-type
387 after 32 hours (incorporating a 24-hour subculture) for both mutants (Fig. 5). The Δ *eutE*
388 mutant showed a significant disadvantage from 12 hours onwards. The competitive index
389 (CI) of both mutants at all time intervals from 12 hours onwards was less than 0.8 (Data
390 Supplement Table S2) No significant difference was found in competitive growth between

391 wild type and mutants in AUM with 0.5 mM ethanolamine (Data Supplement Fig. S7) or in
392 the absence of ethanolamine (data not shown).

393

394 **The *eut* operon is conserved in all UPEC strains sequenced while putative urovirulence**
395 **factors and metabolic polymorphisms previously associated with UPEC are**
396 **phylogroup-related**

397 A SNP-based tree from a core genome alignment of the 47 urine *E. coli* isolates and 32
398 representative reference strains by Parsnp (46) assigned all urine strains to phylogroups (Fig.
399 6). The largest single grouping of urine *E. coli* isolates was formed by 22 phylogroup B2
400 strains (46%) (Fig.6), followed by 11 phylogroup D2 (23%), 7 A 15%), 4 B1(9%), 2 D1(4%)
401 and one phylogroup E (2%). The tree shown used U7 from this study as the reference strain
402 for SNPs and the core 79-genome alignment (47 from this study plus 32 phylogroup
403 representatives) included 53% of the U7 genome. The same phylogroup assignments were
404 found in trees generated with finished closed GenBank genome sequence strains from each
405 phylogroup as the SNP reference strain, as expected (46).

406

407 The presence of a set of 31 previously described (12) putative virulence factors (PUFs)
408 determined by BLASTN searching was used to score each of the 47 *E. coli* genomes. These
409 represented a compilation of genes previously found to be enriched in UTI *E. coli* strains
410 compared to other *E. coli* (51–54). All 31 PUFs were found in the set of genomes and the
411 median PUF count was 13 (range 2-25). Phylogroup B2 *E. coli* urine isolates had higher PUF
412 counts than non-B2 strains ($P < 0.001$, Mann-Whitney U test) (Figure 7A). Hierarchical
413 clustering of PUF carriage profiles showed PUF profile patterns related to B2 clade
414 membership, while clustering of antimicrobial resistance phenotypic profiles showed no
415 obvious phylogenetic relationship (Figure 7B).

416

417 Regarding metabolic features, the *eut* operon was conserved in all 47 strains (Fig. 6B).

418 However, strain U71 contained a novel prophage in the same site as the CPZ-55 prophage

419 insertion between *eutA* and *eutB* characteristic of *E. coli* MG1655 (55) and other K-12

420 lineage strains (Fig. 1A). Genome sequencing of the knock-out strains U1 Δ *eutB* and Δ *eutE*,

421 above, revealed the expected single gene deletions (marked by a kanamycin cassette).

422

423 A short *metE* regulatory allele was present in 30 strains and a complete D-serine tolerance

424 locus (*dsdCXA*) was present in 29 strains (Fig. 6B). All strains contained a complete

425 *yhaOMKJ* D-serine sensory locus. B2 strains were more likely to possess a short *metE*

426 regulatory allele and a complete *dsdCXA* locus than non-B2 strains (2-sided P <0.0001 and

427 0.0022, respectively, Fisher's exact test).

428

429

430 Discussion

431 The *E. coli* strains isolated from urine in this study were phylogenetically similar to

432 previously published urinary tract infection series, in that B2 and D2 were the commonest

433 two phylogroups (56). We report a lower proportion of B2 strains (46%) (Fig. 6) than

434 urosepsis and urinary tract infection studies from the USA and Spain (67%-69%) (12, 54, 57,

435 58), a similar proportion to Slovenia (50%) (59), and more than Denmark (34%) (60) and

436 China (19%) (56). The PUF profile association demonstrated with phylogroup B2 (Fig. 7) is

437 consistent with previous findings from a set of urinary tract infection isolates from the

438 USA (12). This study found that B2 strains not associated with urinary tract infection are also

439 enriched for these genes and that PUF profile does not correlate with virulence in animal

440 models of UTI (12). Phylogroup B2 strains are more likely than other phylogroups to

441 colonise the gut (61, 62) and these putative urovirulence factors may in fact be more
442 important in the gut. Similarly, we found that the metabolic loci proposed to be helpful for
443 growth in urine such as D-serine tolerance and short *metE* allele were also associated with
444 phylogroup B2 (Fig. 6).

445

446 In contrast the *eut* operon was conserved in all isolates (Fig.6) and the ability to utilise
447 ethanolamine *in vitro* was observed in 96% of strains (Supplementary Data Fig. S2). This is
448 not surprising, because the *E. coli* core genome includes the *eut* operon (26). Therefore, the
449 presence of ethanolamine accessible in urine is potentially a significant nutritional resource
450 for all phylogroups of UPEC.

451

452 We found similar concentrations of ethanolamine in infected urine from patients 0.55 mM
453 (mean \pm 0.076) and non-infected urine controls 0.66 mM (mean \pm 0.155) (Fig. 1B). The levels
454 are consistent with previous reports on smaller numbers of samples from healthy controls
455 using different methodology such as NMR (0.38 mM) (63) and LC/MS (0.47 mM) (64). The
456 NMR study found ethanolamine in all 22 urine specimens processed (63). The lack of
457 ethanolamine in a minority of our infected specimens (9/54, Fig.1B) may reflect limitations
458 of the HPLC assay. The maximal ethanolamine concentration in bovine intestinal content
459 (BIC,the filtered contents of jejunum and ileum), where enterohaemorrhagic *E. coli* has been
460 shown to gain an *in vitro* competitive advantage by ethanolamine utilisation, is 2.2 mM (33).
461 For comparison, D-serine is regarded as an abundant substrate for *E. coli* metabolism in
462 human urine (65) where it has been reported at a mean concentration of 0.12 mM out of a
463 total mean urine serine concentration of 0.33 mM (16).

464

465 We found evidence that ethanolamine in infected urine was sensed by *E. coli* with induction
466 of the *eut* operon regulator *eutR*, and was being metabolised, with induction of the
467 ethanolamine deaminase component *eutB* correlating with measured ethanolamine levels in
468 urine (Fig. 1C,D). *In vitro*, UPEC strains produced acetate and ethanol when metabolising
469 ethanolamine in both minimal medium and artificial urine medium (Fig. 2), as expected (Fig.
470 1A) (66). Acetate was also detected in infected urine (Data Supplement Fig S8), as
471 previously reported for infected urine samples with a variety of different bacterial causes
472 (67). Acetogenic growth of *E. coli in vivo* is hypothesized to be an essential property in
473 urinary tract infection (68, 69) and has been ascribed to metabolism of D-serine via pyruvate
474 to acetyl-CoA and acetyl phosphate (68, 70). We propose the consistent presence of host-
475 derived ethanolamine in urine at higher concentrations than D-serine also contributes to this
476 phenotype. Acetate is an important regulator of *E. coli* gene expression (70) and the host
477 immune response (71) and may contribute to the previously reported (35) phenotype linking
478 the *eut* operon to resistance to innate immunity.

479

480 TEM revealed that cells metabolising *E. coli in vitro* in AUM produced numerous plane-
481 edged cytoplasmic inclusions typical of bacterial microcompartments (Fig. 3) in the majority
482 of cells imaged. Although Eut microcompartments have been extensively imaged from
483 *Salmonella enteritidis*, we are not aware of previous publications showing these from
484 uropathogenic *E. coli*.

485

486 Ethanolamine is not synthesized by mammals (72) and is obtained from the diet, with the
487 ultimate source being plant and animal cell membranes. It is incorporated in
488 phosphatidylethanolamine (PE), an aminophospholipid that is an essential constituent of cell
489 membranes, particularly those of mitochondria and the endoplasmic reticulum (72). The

490 source for ethanolamine detected in urine has not been established. Cell lines *in vitro* release
491 ethanolamine into culture medium from cell membrane turnover (73). Within the
492 gastrointestinal tract available ethanolamine is assumed to derive from the breakdown of
493 phospholipid from the turnover of the epithelium and dietary phospholipid (74). There is a
494 constant supply of ethanolamine in urine in both health and infection (Fig. 1B) (63, 64), and
495 the source in health seems unlikely to be cell turnover in the urinary tract, because this occurs
496 at a relatively slow rate compared to the gastrointestinal tract. The cell membranes of
497 neutrophils and bladder epithelial cells are additional potential sources in infected urine.

498

499 There is some evidence to regard *E. coli* as relatively nitrogen limited in the urinary tract
500 because it lacks urease to metabolise the most abundant nitrogen source in urine. Induction of
501 the high-ammonium affinity glutamine synthase and glutamate oxo-glutarate
502 aminotransferase pathway (GS/GOCAT) for nitrogen assimilation occurs in *E. coli* infected
503 urine (69, 75).

504

505 Because ethanolamine metabolism yields ammonia and acetate (Fig. 1A), in theory it should
506 promote *E. coli* growth as either a sole carbon or nitrogen source. *E. coli* utilisation of
507 ethanolamine as a sole nitrogen source in minimal media has been reported at concentrations
508 of 30 mM (33). We found that 96 % of clinical UPEC strains showed utilisation of 10 mM
509 ethanolamine as a sole nitrogen source (Fig. 2A, Data Supplement Fig.S3A). Contradicting
510 the assertion that concentrations of ethanolamine below 1 mM (76) do not support growth of
511 *E. coli*, we found that 0.5 mM, the level present in urine, could sustain small amounts of *E.*
512 *coli* growth in nitrogen limited media (Supplementary Data Fig. S2, S6A). Utilisation of
513 ethanolamine by *E. coli* strains as a sole carbon source *in vitro* is reported to require a high
514 ethanolamine concentration (1 g l^{-1} , 82 mM) (77). Even at this concentration some strains

515 showing active ethanolamine metabolism, for example the O157:H7 EHEC strain EDL933,
516 have been reported as unable to use ethanolamine as a sole carbon source (33). Likewise, we
517 found no *in vitro* growth promotion of known-ethanolamine metabolising UPEC strains by
518 10 mM ethanolamine in carbon limited minimal media (Data Supplement Fig S3B).
519 However, in artificial urine medium (AUM) where the nitrogen sources are urea and
520 ammonia, and the carbon sources are amino acids, lactate and citrate (36), ethanolamine at 10
521 mM and 0.5 mM (Fig 2B, 4B, 4C, Data Supplement Fig. S6B) promoted additional growth of
522 *E. coli*.

523

524 In M9 nitrogen limited media the phenotype of *eutE* mutants showed that the ammonia
525 liberated by the first reaction in ethanolamine metabolism catalysed by *eutBC* (Fig.1A) was
526 sufficient for growth (Fig. 4B, Supplementary Data Fig. S6A). However, this was not
527 sufficient for growth stimulation by ethanolamine in AUM where *eutE* was also required
528 (Fig. 4D), suggesting generation of acetyl-CoA as an additional carbon source was
529 responsible for additional growth in this medium. A second pathway for ethanolamine
530 conversion to acetyl CoA has been predicted (but not defined) in *Salmonella enterica* from
531 the ability of *eutBC* mutants to grow on ethanolamine as a carbon source in the presence of
532 concentrations of carbon dioxide sufficient to change intracellular pH (31), but no carbon
533 dioxide was provided in our experiments.

534

535 The observation that ethanolamine at 10 mM confers a competitive growth advantage on a
536 wild type UPEC strain co-cultured with $\Delta eutE$ and $\Delta eutB$ mutants in artificial urine media
537 (Fig. 5) also supports a role for acetyl-CoA generation in growth enhancement, because
538 extracellular acetate or ammonia deriving from wild-type cells metabolising ethanolamine is
539 apparently insufficient to confer growth enhancement on mutants in this medium. In contrast,

540 *E. coli* strains engineered for enhanced takeup of amino acids to grow faster on amino acids
541 than a wild type strain when cultured in isolation, lose any growth advantage in co-culture
542 with the wild type (78). This is because of extracellular ammonia leak from enhanced amino
543 acid metabolism in the engineered strains providing nitrogen to the wild-type strain (78).

544

545 Although we did not demonstrate a competitive advantage of wild-type *E. coli* over *eut*
546 operon mutants in co-culture in a physiological ethanolamine concentration of 0.5 mM (Data
547 Supplement Fig. S7), this may well be due to methodological limitations. Following a 4 hour
548 lag period, ethanolamine is removed from AUM medium by *E. coli* at a rate of approximately
549 0.75 mM per hour (Fig. 2D), so any selective advantage due to 0.5 mM ethanolamine must be
550 necessarily brief and difficult to detect in a competition assay based on batch culture.

551 However, *in vivo*, host-derived ethanolamine would be continuously passing into urine at the
552 same time as bacterial ethanolamine catabolism. The level of ethanolamine seen in non-
553 infected urine is maintained in infected urine (Fig. 1B) containing large numbers of *E. coli*
554 with induced *eut* operons (Fig. 1C,D), suggesting it is an equilibrium level. The assertion that
555 concentrations of ethanolamine below 1 mM (76), the level present in urine, do not support
556 growth of *E. coli* is contradicted by our *in vitro* data in both minimal medium where
557 ethanolamine functions as the sole nitrogen source (Fig. S6A) and the complex AUM where
558 it appears to function as a carbon source additional to amino acids (Fig 6B). Ethanolamine in
559 urine is an important nutritional resource that infecting uropathogenic *E. coli* can access to
560 augment growth by microcompartment-mediated metabolism. These conserved metabolic
561 pathways and structures distinct from the host offer opportunities for detection and treatment
562 of infection.

563

564 **Acknowledgements**

565 This manuscript has emanated from research supported in part by Science Foundation
566 Ireland (SFI) Grant Number SFI/12/RC/2273 to APC Microbiome Ireland. It was also
567 supported by the British Biotechnology and Biological Sciences Research Council (BBSRC)
568 grants BB/M002969 and BB/H013180. Bacterial genome sequencing was performed by
569 MicrobesNG (<http://www.microbesng.uk>), which is supported by the BBSRC (grant number
570 BB/L024209/1). We thank Dr Brendan Palmer for assistance with R and Mr Daniel Walsh
571 for technical assistance with HPLC. Ethical approval for this project was obtained from the
572 Clinical Research Ethics Committee of the Cork Teaching Hospitals (reference ECM 4 (c)
573 12/08/14).

574

575

576

577

578

579 **Tables**

580

581 **Table 1 Plasmids and strains in this study**

	Genotype/designation	Source
Plasmids		
pCA24N	High copy number expression vector, <i>cat</i>	NBRP <i>E.coli</i> , Japan (39)
pCA24N:: <i>eutB</i>	ASKA clone JW2434	“
pCA24N:: <i>eutE</i>	ASKA clone JW2439	“
Strains		
<i>E. coli</i> U1	<i>E.coli</i> Phylogroup A urine isolate	This study
<i>E.coli</i> JW2434-1	BW25113 Δ <i>eutB</i>	Keio collection, Japan (37)
<i>E.coli</i> JW2439-1	BW25113 Δ <i>eutE</i>	“
<i>E. coli</i> U1 Δ <i>eutB</i>	Δ <i>eutB</i> :: <i>kan</i>	This study
<i>E. coli</i> U1 Δ <i>eutE</i>	Δ <i>eutE</i> :: <i>kan</i>	This study
U2-U79 (46 strains)	<i>E. coli</i> urine isolates.	This study

582

583

584

585

586

587

588 **Table 2** *In vitro* growth phenotype of wild type U1 and *eut* operon mutants with
 589 additional ethanolamine (Eth)
 590

Genotype	M9	M9	AUM	AUM
	10mM Eth	0.5mM Eth	10mM Eth	0.5mM Eth
U1 wild type	+	+	+	+
U1 Δ <i>eutB</i>	—	—	—	—
U1 Δ <i>eutE</i>	+	+	—	—
U1 Δ <i>eutB</i> / pCA24N:: <i>eutB</i>	+	+	+	ND
U1 Δ <i>eutE</i> / pCA24N:: <i>eutE</i>	+	+	+	ND

591

592

593 + growth enhancement compared to growth without ethanolamine. — no growth
 594 enhancement compared to growth without ethanolamine. *Growth enhancement only
 595 observed after 30 hours. ND no data. M9 minimal medium, AUM Artificial Urine Medium.

596

597 **References**

- 598 1. **Stamm, WE, Norrby, SR.** 2001. Urinary tract infections: disease panorama and
599 challenges. *J Infect Dis* **183 Suppl 1**:S1–4.
- 600 2. **Flores-Mireles, AL, Walker, JN, Caparon, M, Hultgren, SJ.** 2015. Urinary tract
601 infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev*
602 *Microbiol* **13**:269–284.
- 603 3. **Foxman, B.** 2014. Urinary tract infection syndromes: occurrence, recurrence,
604 bacteriology, risk factors, and disease burden. *Infect Dis Clin North Am* **28**:1–13.
- 605 4. **Czaja, CA, Stamm, WE, Stapleton, AE, Roberts, PL, Hawn, TR, Scholes, D,**
606 **Samadpour, M, Hultgren, SJ, Hooton, TM.** 2009. Prospective cohort study of
607 microbial and inflammatory events immediately preceding *Escherichia coli* recurrent
608 urinary tract infection in women. *J Infect Dis* **200**:528–536.
- 609 5. **Russo, TA, Johnson, JR.** 2000. Proposal for a new inclusive designation for
610 extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC. *The Journal of infectious*
611 *diseases* **181**:1753–1754.
- 612 6. **Dale, AP, Woodford, N.** 2015. Extra-intestinal pathogenic *Escherichia coli* (ExPEC):
613 Disease, carriage and clones. *J Infect* **71**:615–626.
- 614 7. **Johnson, JR, Murray, AC, Gajewski, A, Sullivan, M, Snippes, P, Kuskowski, MA,**
615 **Smith, KE.** 2003. Isolation and molecular characterization of nalidixic acid-resistant
616 extraintestinal pathogenic *Escherichia coli* from retail chicken products. *Antimicrob*
617 *Agents Chemother* **47**:2161–2168.
- 618 8. **Spurbeck, RR, Dinh, PC, Walk, ST, Stapleton, AE, Hooton, TM, Nolan, LK, Kim,**
619 **KS, Johnson, JR, Mobley, HLT.** 2012. *Escherichia coli* isolates that carry *vat*, *fyuA*,
620 *chuA*, and *yfcV* efficiently colonize the urinary tract. *Infection and immunity* **80**:4115–
621 4122.

- 622 9. **Mulvey, MA, Schilling, JD, Martinez, JJ, Hultgren, SJ.** 2000. Bad bugs and
623 beleaguered bladders: interplay between uropathogenic *Escherichia coli* and innate host
624 defenses. *Proc Natl Acad Sci U S A* **97**:8829–8835.
- 625 10. **Rosen, DA, Hooton, TM, Stamm, WE, Humphrey, PA, Hultgren, SJ.** 2007.
626 Detection of intracellular bacterial communities in human urinary tract infection. *PLoS*
627 *Med* **4**:e329.
- 628 11. **Conover, MS, Hadjifrangiskou, M, Palermo, JJ, Hibbing, ME, Dodson, KW,**
629 **Hultgren, SJ.** 2016. Metabolic Requirements of *Escherichia coli* in Intracellular
630 Bacterial Communities during Urinary Tract Infection Pathogenesis. *MBio* **7**:e00104–
631 16.
- 632 12. **Schreiber, HL, Conover, MS, Chou, W-C, Hibbing, ME, Manson, AL, Dodson,**
633 **KW, Hannan, TJ, Roberts, PL, Stapleton, AE, Hooton, TM, Livny, J, EarL, AM,**
634 **Hultgren, SJ.** 2017. Bacterial virulence phenotypes of *Escherichia coli* and host
635 susceptibility determine risk for urinary tract infections. *Science Translational Medicine*
636 **9**:eaaf1283.
- 637 13. **Forsyth, VS, Armbruster, CE, Smith, SN, Pirani, A, Springman, AC, Walters, MS,**
638 **Nielubowicz, GR, Himpfl, SD, Snitkin, ES, Mobley, HLT.** 2018. Rapid Growth of
639 Uropathogenic *Escherichia coli* during Human Urinary Tract Infection. *mBio*
640 **9**:e00186–18.
- 641 14. **Alteri, CJ, Smith, SN, Mobley, HLT.** 2009. Fitness of *Escherichia coli* during Urinary
642 Tract Infection Requires Gluconeogenesis and the TCA Cycle. *PLoS Pathog*
643 **5**:e1000448.
- 644 15. **Conway, T, Cohen, PS.** 2015. Commensal and pathogenic *Escherichia coli* metabolism
645 in the gut. *Microbiology spectrum* **3**:MBP–0006.

- 646 16. **Huang, Y, Nishikawa, T, Satoh, K, Iwata, T, Fukushima, T, Santa, T, Homma, H,**
647 **Imai, K.** 1998. Urinary excretion of D-serine in human: comparison of different ages
648 and species. *Biol Pharm Bull* **21**:156–162.
- 649 17. **Anfora, AT, Haugen, BJ, Roesch, P, Redford, P, Welch, RA.** 2007. Roles of serine
650 accumulation and catabolism in the colonization of the murine urinary tract by
651 *Escherichia coli* CFT073. *Infect Immun* **75**:5298–5304.
- 652 18. **Metzler, DE, Snell, EE.** 1952. Deamination of serine. II. D-Serine dehydrase, a
653 vitamin B6 enzyme from *Escherichia coli*. *J Biol Chem* **198**:363–373.
- 654 19. **Maas, WK, Maas, R, McFall, E.** 1995. D-serine deaminase is a stringent selective
655 marker in genetic crosses. *Journal of bacteriology* **172**:459–461.
- 656 20. **Nørregaard-Madsen, M, McFall, E, Valentin-Hansen, P.** 1995. Organization and
657 transcriptional regulation of the *Escherichia coli* K-12 D-serine tolerance locus. *J*
658 *Bacteriol* **177**:6456–6461.
- 659 21. **Anfora, AT, Welch, RA.** 2006. DsdX is the second D-serine transporter in
660 uropathogenic *Escherichia coli* clinical isolate CFT073. *J Bacteriol* **188**:6622–6628.
- 661 22. **Moritz, RL, Welch, RA.** 2006. The *Escherichia coli* argW-dsdCXA genetic island is
662 highly variable, and *E. coli* K1 strains commonly possess two copies of dsdCXA. *J Clin*
663 *Microbiol* **44**:4038–4048.
- 664 23. **Cosloy, SD, McFall, E.** 1973. Metabolism of D-serine in *Escherichia coli* K-12:
665 mechanism of growth inhibition. *J Bacteriol* **114**:685–694.
- 666 24. **Oren, Y, Smith, MB, Johns, NI, Kaplan Zeevi, M, Biran, D, Ron, EZ, Corander, J,**
667 **Wang, HH, Alm, EJ, Pupko, T.** 2014. Transfer of noncoding DNA drives regulatory
668 rewiring in bacteria. *Proc Natl Acad Sci U S A* **111**:16112–16117.
- 669 25. **Subashchandrabose, S, Hazen, TH, Brumbaugh, AR, Himpsl, SD, Smith, SN,**
670 **Ernst, RD, Rasko, DA, Mobley, HL.** 2014. Host-specific induction of *Escherichia coli*

- 671 fitness genes during human urinary tract infection. *Proc Natl Acad Sci U S A*
672 **111**:18327–18332.
- 673 26. **Ding, W, Baumdicker, F, Neher, RA.** 2016. panX: pan-genome analysis and
674 exploration. *bioRxiv* 072082.
- 675 27. **Tsoy, O, Ravcheev, D, Mushegian, A.** 2009. Comparative genomics of ethanolamine
676 utilization. *J. Bacteriol.* **JB.00838–09.**
- 677 28. **Kerfeld, CA, Aussignargues, C, Zarzycki, J, Cai, F, Sutter, M.** 2018. Bacterial
678 microcompartments. *Nat Rev Microbiol* **16**:277–290.
- 679 29. **Kaval, KG, Garsin, DA.** 2018. Ethanolamine Utilization in Bacteria. *mBio* **9**:e00066–
680 18.
- 681 30. **Brinsmade, SR, Paldon, T, Escalante-Semerena, JC.** 2005. Minimal functions and
682 physiological conditions required for growth of salmonella enterica on ethanolamine in
683 the absence of the metabolosome. *J Bacteriol* **187**:8039–8046.
- 684 31. **Penrod, JT, Roth, JR.** 2006. Conserving a Volatile Metabolite: a Role for
685 Carboxysome-Like Organelles in Salmonella enterica. *J. Bacteriol.* **188**:2865–2874.
- 686 32. **Moore, TC, Escalante-Semerena, JC.** 2015. The EutQ and EutP Proteins are Novel
687 Acetate Kinases Involved in Ethanolamine Catabolism: Physiological Implications for
688 the Function of the Ethanolamine Metabolosome in Salmonella enterica. *Mol Microbiol*
689 **99**:497–511.
- 690 33. **Bertin, Y, Girardeau, JP, Chaucheyras-Durand, F, Lyan, B, Pujos-Guillot, E,**
691 **Harel, J, Martin, C.** 2011. Enterohaemorrhagic Escherichia coli gains a competitive
692 advantage by using ethanolamine as a nitrogen source in the bovine intestinal content.
693 *Environmental Microbiology* **13**:365–377.
- 694 34. **Thiennimitr, P, Winter, SE, Winter, MG, Xavier, MN, Tolstikov, V, Huseby, DL,**
695 **Sterzenbach, T, Tsolis, RM, Roth, JR, Bäumlner, AJ.** 2011. Intestinal inflammation

- 696 allows Salmonella to use ethanolamine to compete with the microbiota. Proceedings of
697 the National Academy of Sciences **108**:17480–17485.
- 698 35. **Sintsova, A, Smith, S, Subashchandrabose, S, Mobley, HL.** 2018. Role of
699 ethanolamine utilization genes in host colonization during urinary tract infection. *Infect*
700 *Immun* **86**:e00542–17.
- 701 36. **Brooks, T, Keevil, CW.** 1997. A simple artificial urine for the growth of urinary
702 pathogens. *Lett Appl Microbiol* **24**:203–206.
- 703 37. **Baba, T, Ara, T, Hasegawa, M, Takai, Y, Okumura, Y, Baba, M, Datsenko, KA,**
704 **Tomita, M, Wanner, BL, Mori, H.** 2006. Construction of Escherichia coli K-12 in-
705 frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**:2006.0008.
- 706 38. **Thomason, LC, Costantino, N, Court, DL.** 2007. *E. coli* genome manipulation by P1
707 transduction. *Current protocols in molecular biology* **79**:1.17.1–1.17.8.
- 708 39. **Kitagawa, M, Ara, T, Arifuzzaman, M, Ioka-Nakamichi, T, Inamoto, E, Toyonaga,**
709 **H, Mori, H.** 2005. Complete set of ORF clones of Escherichia coli ASKA library (a
710 complete set of E. coli K-12 ORF archive): unique resources for biological research.
711 *DNA Res* **12**:291–299.
- 712 40. **Sturms, R, Streauslin, NA, Cheng, S, Bobik, TA.** 2015. In Salmonella enterica,
713 Ethanolamine Utilization Is Repressed by 1,2-Propanediol To Prevent Detrimental
714 Mixing of Components of Two Different Bacterial Microcompartments. *J Bacteriol*
715 **197**:2412–2421.
- 716 41. **Parsons, JP, Dinesh, SD, Deery, E, Leech, HK, Brindley, AA, Heldt, D, Frank, S,**
717 **Smales, CM, Lunsdorf, H, Rambach, A, Gass, MH, Bleloch, A, McClean, KJ,**
718 **Munro, AW, Rigby, SEJ, Warren, MJ, Prentice, MB.** 2008. Biochemical and
719 structural insights into bacterial organelle form and biogenesis. *J. Biol. Chem.*
720 **283**:14366–14375.

- 721 42. **Wood, DE, Salzberg, SL.** 2014. Kraken: ultrafast metagenomic sequence classification
722 using exact alignments. *Genome biology* **15**:R46.
- 723 43. **Li, H, Durbin, R.** 2010. Fast and accurate long-read alignment with Burrows-Wheeler
724 transform. *Bioinformatics* **26**:589–595.
- 725 44. **Bankevich, A, Nurk, S, Antipov, D, Gurevich, AA, Dvorkin, M, Kulikov, AS,**
726 **Lesin, VM, Nikolenko, SI, Pham, S, Prjibelski, AD, Pyshkin, AV, Sirotkin, AV,**
727 **Vyahhi, N, Tesler, G, Alekseyev, MA, Pevzner, PA.** 2012. SPAdes: a new genome
728 assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*
729 **19**:455–477.
- 730 45. **Seemann, T.** 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*
731 **30**:2068–2069.
- 732 46. **Treangen, TJ, Ondov, BD, Koren, S, Phillippy, AM.** 2014. The Harvest suite for
733 rapid core-genome alignment and visualization of thousands of intraspecific microbial
734 genomes. *Genome biology* **15**:1–15.
- 735 47. **Letunic, I, Bork, P.** 2016. Interactive tree of life (iTOL) v3: an online tool for the
736 display and annotation of phylogenetic and other trees. *Nucleic Acids Res* **44**:W242–5.
- 737 48. **Price, MN, Dehal, PS, Arkin, AP.** 2010. FastTree 2--approximately maximum-
738 likelihood trees for large alignments. *PLoS One* **5**:e9490.
- 739 49. **Dixit, PD, Pang, TY, Studier, FW, Maslov, S.** 2015. Recombinant transfer in the
740 basic genome of *Escherichia coli*. *Proc Natl Acad Sci U S A* **112**:9070–9075.
- 741 50. **Schmittgen, TD, Livak, KJ.** 2008. Analyzing real-time PCR data by the comparative
742 CT method. *Nature protocols*
- 743 51. **Johnson, JR, Stell, AL.** 2000. Extended virulence genotypes of *Escherichia coli* strains
744 from patients with urosepsis in relation to phylogeny and host compromise. *The Journal*
745 *of infectious diseases* **181**:261–272.

- 746 52. **Johnson, JR, O'Bryan, TT, Delavari, P, Kuskowski, M, Stapleton, A, Carlino, U,**
747 **Russo, TA.** 2001. Clonal relationships and extended virulence genotypes among
748 *Escherichia coli* isolates from women with a first or recurrent episode of cystitis. *The*
749 *Journal of infectious diseases* **183**:1508–1517.
- 750 53. **Luo, Y, Ma, Y, Zhao, Q, Wang, L, Guo, L, Ye, L, Zhang, Y, Yang, J.** 2012.
751 Similarity and divergence of phylogenies, antimicrobial susceptibilities, and virulence
752 factor profiles of *Escherichia coli* isolates causing recurrent urinary tract infections that
753 persist or result from reinfection. *J Clin Microbiol* **50**:4002–4007.
- 754 54. **Johnson, JR, Porter, S, Johnston, B, Kuskowski, MA, Spurbeck, RR, Mobley, HL,**
755 **Williamson, DA.** 2015. Host Characteristics and Bacterial Traits Predict Experimental
756 Virulence for *Escherichia coli* Bloodstream Isolates From Patients With Urosepsis.
757 *Open Forum Infect Dis* **2**:ofv083.
- 758 55. **Hayashi, K, Morooka, N, Yamamoto, Y, Fujita, K, Isono, K, Choi, S, Ohtsubo, E,**
759 **Baba, T, Wanner, BL, Mori, H, Horiuchi, T.** 2006. Highly accurate genome
760 sequences of *Escherichia coli* K-12 strains MG1655 and W3110. *Mol Syst Biol* **2**:2006
761 0007.
- 762 56. **Cao, X, Cavaco, LM, Lv, Y, Li, Y, Zheng, B, Wang, P, Hasman, H, Liu, Y,**
763 **Aarestrup, FM.** 2011. Molecular characterization and antimicrobial susceptibility
764 testing of *Escherichia coli* isolates from patients with urinary tract infections in 20
765 Chinese hospitals. *J Clin Microbiol* **49**:2496–2501.
- 766 57. **Zhang, L, Foxman, B, Marrs, C.** 2002. Both urinary and rectal *Escherichia coli*
767 isolates are dominated by strains of phylogenetic group B2. *J Clin Microbiol* **40**:3951–
768 3955.

- 769 58. **Moreno, E, Andreu, A, Pigrau, C, Kuskowski, MA, Johnson, JR, Prats, G.** 2008.
770 Relationship between *Escherichia coli* strains causing acute cystitis in women and the
771 fecal *E. coli* population of the host. *J Clin Microbiol* **46**:2529–2534.
- 772 59. **Rijavec, M, Starcic Erjavec, M, Ambrozic Avgustin, J, Reissbrodt, R, Fruth, A,**
773 **Krizan-Hergouth, V, Zgur-Bertok, D.** 2006. High prevalence of multidrug resistance
774 and random distribution of mobile genetic elements among uropathogenic *Escherichia*
775 *coli* (UPEC) of the four major phylogenetic groups. *Curr Microbiol* **53**:158–162.
- 776 60. **Skjøt-Rasmussen, L, Hammerum, AM, Jakobsen, L, Lester, CH, Larsen, P,**
777 **Frimodt-Møller, N.** 2011. Persisting clones of *Escherichia coli* isolates from recurrent
778 urinary tract infection in men and women. *J Med Microbiol* **60**:550–554.
- 779 61. **Nowrouzian, FL, Adlerberth, I, Wold, AE.** 2006. Enhanced persistence in the colonic
780 microbiota of *Escherichia coli* strains belonging to phylogenetic group B2: role of
781 virulence factors and adherence to colonic cells. *Microbes and Infection* **8**:834–840.
- 782 62. **Nowrouzian, FL, Wold, AE, Adlerberth, I.** 2005. *Escherichia coli* strains belonging
783 to phylogenetic group B2 have superior capacity to persist in the intestinal microflora of
784 infants. *The Journal of infectious diseases* **191**:1078–1083.
- 785 63. **Bouatra, S, Aziat, F, Mandal, R, Guo, AC, Wilson, MR, Knox, C, Bjorndahl, TC,**
786 **Krishnamurthy, R, Saleem, F, Liu, P, Dame, ZT, Poelzer, J, Huynh, J, Yallou, FS,**
787 **Psychogios, N, Dong, E, Bogumil, R, Roehring, C, Wishart, DS.** 2013. The human
788 urine metabolome. *PLoS One* **8**:e73076.
- 789 64. **Guo, K, Li, L.** 2009. Differential ¹²C-/¹³C-isotope dansylation labeling and fast liquid
790 chromatography/mass spectrometry for absolute and relative quantification of the
791 metabolome. *Anal Chem* **81**:3919–3932.

- 792 65. **Connolly, JP, Roe, AJ.** 2016. When and where? Pathogenic *Escherichia coli*
793 differentially sense host D-serine using a universal transporter system to monitor their
794 environment. *Microb Cell* **3**:181–184.
- 795 66. **Scarlett, FA, Turner, JM.** 1976. Microbial metabolism of amino alcohols.
796 Ethanolamine catabolism mediated by coenzyme B12-dependent ethanolamine
797 ammonia-lyase in *Escherichia coli* and *Klebsiella aerogenes*. *Journal of general*
798 *microbiology* **95**:173–176.
- 799 67. **Gupta, A, Dwivedi, M, Mahdi, AA, Khetrapal, CL, Bhandari, M.** 2012. Broad
800 identification of bacterial type in urinary tract infection using (1)h NMR spectroscopy. *J*
801 *Proteome Res* **11**:1844–1854.
- 802 68. **Anfora, AT, Halladin, DK, Haugen, BJ, Welch, RA.** 2008. Uropathogenic
803 *Escherichia coli* CFT073 is adapted to acetatogenic growth but does not require acetate
804 during murine urinary tract infection. *Infection and Immunity* **76**:5760–5767.
- 805 69. **Hagan, EC, Lloyd, AL, Rasko, DA, Faerber, GJ, Mobley, HLT.** 2010. *Escherichia*
806 *coli* global gene expression in urine from women with urinary tract infection. *PLoS*
807 *Pathogens* **6**:e1001187.
- 808 70. **Wolfe, AJ.** 2005. The acetate switch. *Microbiol Mol Biol Rev* **69**:12–50.
- 809 71. **Mariño, E, Richards, JL, McLeod, KH, Stanley, D, Yap, YA, Knight, J, McKenzie,**
810 **C, Kranich, J, Oliveira, AC, Rossello, FJ, Krishnamurthy, B, Nefzger, CM, Macia,**
811 **L, Thorburn, A, Baxter, AG, Morahan, G, Wong, LH, Polo, JM, Moore, RJ,**
812 **Lockett, TJ, Clarke, JM, Topping, DL, Harrison, LC, Mackay, CR.** 2017. Gut
813 microbial metabolites limit the frequency of autoimmune T cells and protect against
814 type 1 diabetes. *Nat Immunol* **18**:552–562.
- 815 72. **Vance, JE, Tasseva, G.** 2013. Formation and function of phosphatidylserine and
816 phosphatidylethanolamine in mammalian cells. *Biochim Biophys Acta* **1831**:543–554.

- 817 73. **Schmitt, J, Noble, A, Otsuka, M, Berry, P, Maitland, NJ, Rumsby, MG.** 2014.
818 Phorbol ester stimulates ethanolamine release from the metastatic basal prostate cancer
819 cell line PC3 but not from prostate epithelial cell lines LNCaP and P4E6. *British journal*
820 *of cancer* **111**:1646.
- 821 74. **Cotton, PB.** 1972. Non-dietary lipid in the intestinal lumen. *Gut* **13**:675–681.
- 822 75. **Alteri, CJ, Himpls, SD, Mobley, HL.** 2015. Preferential use of central metabolism in
823 vivo reveals a nutritional basis for polymicrobial infection. *PLoS Pathog* **11**:e1004601.
- 824 76. **Kendall, MM, Gruber, CC, Parker, CT, Sperandio, V.** 2012. Ethanolamine Controls
825 Expression of Genes Encoding Components Involved in Interkingdom Signaling and
826 Virulence in Enterohemorrhagic *Escherichia coli* O157:H7. *mBio* **3**
- 827 77. **Jones, PW, Turner, JM.** 1984. Interrelationships Between the Enzymes of
828 Ethanolamine Metabolism in *Escherichia coli*. *Journal of General Microbiology*
829 **130**:299–308.
- 830 78. **Wang, J, Yan, D, Dixon, R, Wang, YP.** 2016. Deciphering the Principles of Bacterial
831 Nitrogen Dietary Preferences: a Strategy for Nutrient Containment. *MBio* **7**
832
833

834 **Figure Legends Dadswell et al**

835

836 **Figure 1. Ethanolamine is present in urine and urinary ethanolamine concentration**
837 **correlates with expression of *eut* operon genes in *E. coli* infected urine**

838 A. Microcompartment-mediated ethanolamine metabolic pathway and *eut* operon: black
839 arrows metabolite translocation or reaction, dotted arrows metabolite translocation impeded
840 by microcompartment, blue hexagon microcompartment, microcompartment associated
841 enzymes in blue, cytoplasmic enzymes in black. Yellow arrows below show the *eut* operon
842 (red arrow at prophage insertion hot spot). B. Ethanolamine concentration in urine. There is
843 no significant difference in ethanolamine concentration between clinically infected urine
844 samples and control non-infected samples (Mann-Whitney U-test). C. Correlation between
845 ethanolamine concentration in infected urine and expression of *eutB* (relative to *gyrA*),
846 Spearman's rank correlation coefficient $r=0.815$, $***p < 0.001$. D. Correlation between
847 ethanolamine concentration in infected urine and expression of *eutR* (relative to *gyrA*),
848 Spearman's rank correlation coefficient $r=0.423$.

849

850 **Figure 2. Ethanolamine metabolism promotes UPEC growth in nitrogen-limited**
851 **minimal medium and Artificial Urine Medium (AUM).**

852 Aerobic growth of selected UPECs at 37 °C in: (A) ammonia-free modified M9 media with
853 glycerol (20mM) (B) AUM. Hollow data points are without ethanolamine, solid data points
854 with additional 10mM Ethanolamine. Concentration of ethanolamine (Eth) (green), acetate
855 (red) and ethanol (blue) over time during U1 growth in (C) ammonia-free M9 media with
856 glycerol (20mM) and (D) AUM, both supplemented with an initial 10mM ethanolamine.
857 Values are Mean \pm SEM. $N \geq 3$.

858

859 **Figure 3. Growth of UPEC strain U1 in Artificial Urine Medium with ethanolamine**
860 **promotes formation of bacterial microcompartments.**

861 Transmission electron microscopy following culture for eight hours. A. in AUM with 10 mM
862 ethanolamine. B. In AUM alone. White arrows indicate microcompartments.

863

864 **Figure 4. *eutE* inactivation in UPEC strain U1 abolishes ethanolamine growth**
865 **stimulation in AUM medium despite preserved ethanolamine catabolism**

866 A. Growth of U1, U1 Δ *eutB* mutant and complement in modified M9 plus 10 mM
867 ethanolamine. B. Growth of U1, U1 Δ *eutE* mutant and complement in modified M9 with 10
868 mM ethanolamine. C. Growth of U1, U1 Δ *eutB* mutant and complement in AUM plus 10
869 mM ethanolamine. D. Growth of U1, U1 Δ *eutE* and complement in AUM with 10 mM
870 ethanolamine. In A-D growth of U1 in control medium without ethanolamine is shown as
871 open circles. E. Percentage change in ethanolamine concentration measured by HPLC over
872 24 hours of U1, U1 Δ *eutB* and U1 Δ *eutE*, and their complements in AUM with initial 10mM
873 ethanolamine. Significant difference with wild-type*** $P < 0.001$, 1-way ANOVA. All
874 Values are Mean \pm SEM. N=3

875

876 **Figure 5. Inactivation of *eut* operon genes reduces competitiveness of *E. coli* UPEC**
877 **strain U1 in artificial urine medium containing 10 mM ethanolamine**

878 Competition of U1 vs A: U1 Δ *eutB* with 10mM ethanolamine, B: U1 Δ *eutE* with 10mM
879 ethanolamine. Mann-Whitney U test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Values are Mean \pm
880 SEM. N=3.

881

882 **Figure 6. Phylogenetic distribution of *E. coli* urine isolates from this study and**
883 **conservation of metabolic operons**

884 A. The phylogeny of 47 strains (taxon labels in red) isolated from infected urine analysed by
885 core genome alignment using Parsnp with 32 reference strains representative of six *E.coli*
886 phylogroups (taxon labels in black). Bootstrap values for all internal nodes were 1.0 apart
887 from the node (0.25) between the reference strains APECO1 and IHE3034, which constitute
888 the least diverged core genome pair in the reference set. Clade assignments shown in the
889 vertical bar on the right. B. Parsnp alignment of the 47 strains alone, B2 phylogroup coloured
890 blue. Vertical bars/circles indicate presence of a complete *eut* operon (red), a complete
891 *dsdCXA* locus (green), and a short regulatory *metE* allele (grey) in each strain

892

893 **Figure 7 Carriage of PUFs (putative virulence factors) but not antimicrobial resistance**
894 **is associated with clade B2 *E. coli* urine isolates**

895 A. PUF scores differ between B2 and non-B2 groups. Mann-Whitney U test *** $p < 0.0001$

896 B. Antimicrobial resistance scores (number of different antimicrobials to which the strain is
897 resistant) does not differ between B2 and non-B2 groups

898 C. Genome sequences of clinical urine *E.coli* isolates were screened for the presence of 31
899 previously-described PUFs (y-axis labels) using BLASTN. Presence (black squares) or
900 absence (grey squares) is shown for each PUF in relation to each isolate. Two dimensional
901 hierarchical clustering shows PUF co-occurrence by strain (upper y-axis dendrogram) and
902 PUF association with phylogeny (x-axis dendrogram). Clade B2 strains are indicated by
903 white names on a black background (x-axis labels). Lower diagram shows hierarchical
904 clustering of resistance (dark grey squares) and sensitivity (pale grey squares) to nine
905 different antimicrobials (lower y-axis dendrogram) by strain phylogeny. Abbreviations as
906 follows: Gent, Gentamicin; Nitro, Nitrofurantoin; Cipro, Ciprofloxacin; Levo, Levofloxacin;
907 Tetra, Tetracycline; Cotrim, Cotrimoxazole; PipTaz, Piperacillin/tazobactam; Amp,
908 Ampicillin; Coamox, Amoxicillin/clavulanic acid.

909
910











