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

Katherine Dadswell[1], Sinead Creagh[2], Edward McCullagh[2], Mingzhi Liang[3], Ian R. Brown[3], Martin J Warren[3], Alan McNally[4], John MacSharry[1,5*], Michael B. Prentice[1,2,5,6*]

[1] School of Microbiology, University College Cork, Cork, Ireland

[2] Department of Microbiology, Cork University Hospital, Cork, Ireland

[3] School of Biosciences, University of Kent, Canterbury, UK

[4] Institute of Microbiology and Infection, University of Birmingham, Birmingham, UK

[5] APC Microbiome Ireland, University College Cork, Cork, Ireland

[6] Department of Pathology, University College Cork, Cork, Ireland

*Correspondence: Dr John MacSharry and Professor Michael B Prentice, School of Microbiology, University College Cork, Cork, Ireland

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

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

Abstract

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

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

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

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

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

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

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

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

The *eut* operon is part of the conserved *E. coli* core genome (26) having arrived in Enterobacterales by horizontal transfer (27). It contains seventeen genes including the positive transcriptional regulator *eutR*. The operon encodes enzymes required for ethanolamine metabolism and includes structural shell protein genes for the synthesis of thin porous protein shells enclosing the enzymes as bacterial microcompartments (metabolosomes) in the cytoplasm (28–30) (Fig 1A). Experiments largely conducted with *Salmonella enterica* (which contains the same operon) suggest that the enzymic breakdown of ethanolamine to ammonia (a nitrogen source) and acetaldehyde occurs within the metabolosome, with the toxic effects and evaporative loss of acetaldehyde minimised by microcompartment enclosure and onward metabolism to ethanol and acetyl-CoA (a carbon source)(30, 31). Some acetyl-CoA is further metabolised to acetyl phosphate and acetate within the metabolosome, and some is available to enter central metabolism (32). Ethanolamine in the gastrointestinal tract utilised by this pathway gives a competitive advantage to Enterohaemorrhagic *E. coli* (33) and *Salmonella enteritidis* (34). Recently it has been shown that *E. coli* ethanolamine metabolism is essential for bladder colonisation in a murine model of ascending UTI (35). The mechanism was suggested to involve resistance to

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

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

Methods

Bacterial strains and culture conditions

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

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

E. coli were routinely cultured in LB broth at 30 °C or 37 °C with aeration.To determine the ability to utilise ethanolamine, strains were cultured at 37 °C in modified M9 minimal media (33) containing 10 mM ethanolamine hydrochloride and 200 nM cobalamin with the addition of either 20mM glycerol or 20 mM ammonium chloride. Automated growth count cultures were incubated in 96-well plates in triplicate and OD_{600} measured using a Biotek Eon Microplate Spectrophotometer over 48 hours. Manual growth curves were measured in 35 ml volumes with spectrophotometric analysis of 1 ml aliquots.

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

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

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

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

Mutants

To generate deletion mutants, BW25113 knockout *E. coli* strains for the genes of interest were obtained from the Keio collection (37). Mutations were transferred to UPEC strain U1 by P1vir phage transduction (38) In brief, lysogen strains were prepared by incubating P1 lysate with the donor strain for 30 minutes at 30°C with 5µl of 1M $CaCl_2$ and the culture were plated on kanamycin selective agar. The resulting colonies were used to prepare the lysate for transduction. Lysogen colonies were grown overnight in 2ml of LB at 30ºC. The precultures were used to inoculate LB and grown until reaching an OD_{600} of 0.2. The cultures were incubated in 46ºC for 20 minutes with shaking before being moved to 37°C until complete lysis. Bacteria were centrifuged out of the culture and the supernatant was stored with chloroform to prevent bacterial growth. Overnight cultures of the recipient strain were resuspended in transduction buffer (10mM $MgSO_4$, 5mM $CaCl_2$) and 100µl of cells were incubated with lysate and incubated at 37°C for 30 minutes. Sodium citrate was added following this incubation and for a further hour. The cells were washed in LB before being

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

Metabolic assays

After culture, residual urine samples were separated into cell fraction and cell free supernatant by differential centrifugation and urine supernatants were filtered with a 0.2µm membrane to remove any remaining bacteria and stored at -80°C. Urine supernatants and culture supernatants were assayed for ethanolamine, acetate and ethanol by HPLC using an Agilent 1200 HPLC system with a refractive index detector. Urines collected from CUH and bacterial culture supernatants were filter sterilised at 0.2µm to remove bacteria before being stored at -80°C until the day of experimentation. Ethanolamine was measured by gradient HPLC after derivatisation with o-phthaldialdehyde (OPA) using a method adapted from Sturms et al. (40). The mobile phase consisted of Buffer A [10% methanol (Sigma-Aldrich) - 90% 10mM Na_3PO_4 (pH7.3) (Sigma-Aldrich)], and Buffer B [80% Methanol- 20% 10mM Na_3PO_4 (pH7.3)]. Samples were prepared using an in-loop derivatization reaction where 6µl of sample were taken up followed by 6µl 10mg/ml OPA and 3-mercaptopropionic acid in 0.4M boric acid (Agilent Technologies) and incubated at room temperature for 3 minutes. The samples were injected into a 4.6 by 100mm, 2.7µm pore Infinity Lab Poroshell HPH-C18 column (Agilent Technologies) and eluted with 5ml linear gradient from 50% Buffer B to 100% Buffer B followed by 5mls of 100% Buffer B at as constant flow rate of 1ml min^{-1}. The excitation was detected at 224nm. A standard curve was created before each sequence

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

Acetate and ethanol were measured by the same HPLC system. 10µl of sample was injected into a REZEX 8µm 8% H, Organic Acid Column (Phenomenex, USA) and eluted with 15ml of 0.01M H_2SO_4 at a flow rate of 0.6ml min^{-1}. The column was maintained at 65°C for the duration of the experiments. The identification of the substrate was determined by comparison of retention time to pure compound and concentrations were quantified by comparison to known standards.

Transmission Electron Microscopy (TEM)

This was carried out as previously described (41) After growth (as indicated in the text) bacteria cells were pelleted by centrifugation, to give a pellet no larger than 100µl in volume. The bacterial pellet was fixed in 2ml of 2.5% glutaraldehyde (Fluka) diluted in 0.1M Sodium cacodylate pH 6.8 (CAB) (bioWORLD). After incubation overnight at 4°C, bacteria were washed twice with 0.1 M CAB and suspended in fresh 2ml of 2.5% glutaraldehyde diluted in CAB. The bacteria were stained for 1 hour in 1 ml of 1% osmium tetroxide (w/v) (250µl 4% osmium tetroxide; 250 µl Milli-Q H_2O; 500 µl 0.2 M CAB). The pellets were washed in 2ml Milli-Q H_2O for 10 minutes twice before the pellets were dehydrated. Pellets were dehydrated through an ethanol (EtOH) gradient as follows: 50% EtOH (v/v) x 10mins; 70% EtOH x 10min; 90% EtOH x 10mins; 100% EtOH x 10mins three times and then the bacterial pellets were washed twice in propylene oxide for 10mins. The pellets were embedded into 1.5 ml propylene oxide: LV resin at 1:1 for 30 min followed by incubation 2 × 1.5 h in 100% freshly made agar LV resin. The pellets were resuspended in 1ml of 100% LV resin and transferred to a conical bottom tube. The bacterial pellet was centrifuged at 1100xg for 5mins and was left to incubate at 60°C for 24 hours. Bacteria were sectioned to

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

DNA sequencing and sequence analysis, statistics

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

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

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

RNA and RT-PCR

RNA was extracted from bacterial pellets using the Zymo Fungal/Bacterial Mini Prep kit and from Eukaryotic cells using the Quick-RNA MiniPrep kit, following the manufacturer's instructions. After extraction genomic DNA was digested using the TURBO DNA-free (Ambion) DNase 1 treatment. The RNA was quantified using a Nanodrop 1000 spectrophotometer. cDNA was synthesised by reverse transcription carried out in nuclease free 96-well plates. RNA was diluted using molecular grade H_2O (Sigma-Aldrich) to a final concentration from of 100ng μl^{-1} in a 10µl volume. The RNA was mixed with: cDNA reaction was set up 4µl 5x Reverse transcription buffer (Roche); 3µl Random Hexamer Primer (Roche); 2µl 20mM dNTPs mix; and 1µl Reverse transcriptase/RNase Inhibitor to give a total volume of 20µl. The reaction mixture was incubated in a thermocycler in the following condition: 10mins at 25 °C; 30mins 55 °C; 5mins at 85 °C; hold at 4 °C. The cDNA was then diluted to 100µl and stored at -20 °C until use.

The universal probe library (Roche, Indianapolis USA) was utilized to design primers for quantitative PCR. The primers used in this study are listed in Data Supplement Table S1 Amplification reactions were a mix of: 3µl of cDNA; 7µl TaqMan Probe Master buffer (Roche); 1µl 20mM primer mix (L+R primers); 0.1µl probe(Roche);and 0.9 µl molecular grade H_2O to a make a final volume of 10 µl. When the probe was not available a SYBR Green master mix was used which included: 3 µl cDNA, 5 µl 2xSYBR Green I Master buffer

(Roche); 1µl 20mM primer mix(L+R primers) and 1µl molecular grade H_2O to a final volume of 10 µl. All reactions were performed using a 384 well plate on the LightCycler®480 System (Roche) with molecular grade water included as a negative control. Thermal cycling condition were as follows: 50°C for 2 mins, 95°C for 10 mins followed by 45 cycles for 95°C for 10s, 60°C for 45 s and 72°C for 60 s. Relative gene expression was calculated using the 2-ΔΔCt (50). X-fold changes in mRNA of target gene was quantified relative to *gyrA*.

ELISA

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

Results

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

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

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

Clinically infected urine samples show stimulation of the host innate immune response . Cytokines IL-8 and IL-1β were detected in 81% of clinically infected urine samples and significantly increased in infected urines compared to non-infected urine (IL-1β $P=0.0048$, and IL-8 $P<0.001$, see Fig. S1 Supplemental material). Mean IL-6 levels were higher in infected urine than in non-infected urine but the difference was not significant (Data Supplement Fig. S1).

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

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

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

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

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

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

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

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

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

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

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

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

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

Regarding metabolic features, the *eut* operon was conserved in all 47 strains (Fig. 6B). However, strain U71 contained a novel prophage in the same site as the CPZ-55 prophage insertion between *eutA* and *eutB* characteristic of *E. coli* MG1655 (55) and other K-12 lineage strains (Fig. 1A). Genome sequencing of the knock-out strains U1 Δ*eutB* and Δ*eutE* , above, revealed the expected single gene deletions (marked by a kanamycin cassette).

A short *metE* regulatory allele was present in 30 strains and a complete D-serine tolerance locus (*dsdCXA*) was present in 29 strains (Fig. 6B). All strains contained a complete *yhaOMKJ* D-serine sensory locus. B2 strains were more likely to possess a short *metE* regulatory allele and a complete *dsdCXA* locus than non-B2 strains (2-sided P <0.0001 and 0.0022, respectively, Fisher's exact test).

Discussion

The *E. coli* strains isolated from urine in this study were phylogenetically similar to previously published urinary tract infection series, in that B2 and D2 were the commonest two phylogroups (56). We report a lower proportion of B2 strains (46%) (Fig. 6) than urosepsis and urinary tract infection studies from the USA and Spain (67%-69%) (12, 54, 57, 58), a similar proportion to Slovenia (50%) (59), and more than Denmark (34%) (60) and China (19%) (56). The PUF profile association demonstrated with phylogroup B2 (Fig. 7) is consistent with previous findings from a set of urinary tract infection isolates from the USA (12). This study found that B2 strains not associated with urinary tract infection are also enriched for these genes and that PUF profile does not correlate with virulence in animal models of UTI (12). Phylogroup B2 strains are more likely than other phylogroups to

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

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

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

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

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

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

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

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

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

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

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

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

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

Although we did not demonstrate a competitive advantage of wild-type *E. coli* over *eut* operon mutants in co-culture in a physiological ethanolamine concentration of 0.5 mM (Data Supplement Fig. S7), this may well be due to methodological limitations. Following a 4 hour lag period, ethanolamine is removed from AUM medium by *E. coli* at a rate of approximately 0.75 mM per hour (Fig. 2D), so any selective advantage due to 0.5 mM ethanolamine must be necessarily brief and difficult to detect in a competition assay based on batch culture. However, *in vivo*, host-derived ethanolamine would be continuously passing into urine at the same time as bacterial ethanolamine catabolism. The level of ethanolamine seen in non-infected urine is maintained in infected urine (Fig. 1B) containing large numbers of *E. coli* with induced *eut* operons (Fig. 1C,D), suggesting it is an equilibrium level. The assertion that concentrations of ethanolamine below 1 mM (76), the level present in urine, do not support growth of *E. coli* is contradicted by our *in vitro* data in both minimal medium where ethanolamine functions as the sole nitrogen source (Fig. S6A) and the complex AUM where it appears to function as a carbon source additional to amino acids (Fig 6B). Ethanolamine in urine is an important nutritional resource that infecting uropathogenic *E. coli* can access to augment growth by microcompartment-mediated metabolism. These conserved metabolic pathways and structures distinct from the host offer opportunities for detection and treatment of infection.

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

Tables

Table 1 Plasmids and strains in this study

	Genotype/designation	Source
Plasmids		
pCA24N	High copy number expression vector, *cat*	NBRP *E.coli*, Japan (39)
pCA24N::*eutB*	ASKA clone JW2434	"
pCA24N::*eutE*	ASKA clone JW2439	"
Strains		
E. coli U1	*E.coli* Phylogroup A urine isolate	This study
E.coli JW2434-1	BW25113Δ*eutB*	Keio collection, Japan (37)
E.coli JW2439-1	BW25113Δ*eutE*	"
E. coli U1Δ*eutB*	Δ*eutB::kan*	This study
E. coli U1Δ*eutE*	Δ*eutE::kan*	This study
U2-U79 (46 strains)	*E. coli* urine isolates.	This study

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

Genotype	M9 10mM Eth	M9 0.5mM Eth	AUM 10mM Eth	AUM 0.5mM Eth
U1 wild type	+	+	+	+
U1Δ*eutB*	—	—	—	—
U1Δ*eutE*	+	+	—	—
U1Δ*eutB*/ pCA24N::*eutB*	+	+	+	ND
U1Δ*eutE*/ pCA24N::*eutE*	+	+	+	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

References

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

9. **Mulvey, MA, Schilling, JD, Martinez, JJ, Hultgren, SJ**. 2000. Bad bugs and beleaguered bladders: interplay between uropathogenic Escherichia coli and innate host defenses. Proc Natl Acad Sci U S A **97**:8829–8835.

10. **Rosen, DA, Hooton, TM, Stamm, WE, Humphrey, PA, Hultgren, SJ**. 2007. Detection of intracellular bacterial communities in human urinary tract infection. PLoS Med **4**:e329.

11. **Conover, MS, Hadjifrangiskou, M, Palermo, JJ, Hibbing, ME, Dodson, KW, Hultgren, SJ**. 2016. Metabolic Requirements of Escherichia coli in Intracellular Bacterial Communities during Urinary Tract Infection Pathogenesis. MBio **7**:e00104–16.

12. **Schreiber, HL, Conover, MS, Chou, W-C, Hibbing, ME, Manson, AL, Dodson, KW, Hannan, TJ, Roberts, PL, Stapleton, AE, Hooton, TM, Livny, J, EarL, AM, Hultgren, SJ**. 2017. Bacterial virulence phenotypes of Escherichia coli and host susceptibility determine risk for urinary tract infections. Science Translational Medicine **9**:eaaf1283.

13. **Forsyth, VS, Armbruster, CE, Smith, SN, Pirani, A, Springman, AC, Walters, MS, Nielubowicz, GR, Himpsl, SD, Snitkin, ES, Mobley, HLT**. 2018. Rapid Growth of Uropathogenic Escherichia coli during Human Urinary Tract Infection. mBio **9**:e00186–18.

14. **Alteri, CJ, Smith, SN, Mobley, HLT**. 2009. Fitness of Escherichia coli during Urinary Tract Infection Requires Gluconeogenesis and the TCA Cycle. PLoS Pathog **5**:e1000448.

15. **Conway, T, Cohen, PS**. 2015. Commensal and pathogenic escherichia coli metabolism in the gut. Microbiology spectrum **3**:MBP–0006.

Infection and Immunity

16. **Huang, Y, Nishikawa, T, Satoh, K, Iwata, T, Fukushima, T, Santa, T, Homma, H, Imai, K**. 1998. Urinary excretion of D-serine in human: comparison of different ages and species. Biol Pharm Bull **21**:156–162.

17. **Anfora, AT, Haugen, BJ, Roesch, P, Redford, P, Welch, RA**. 2007. Roles of serine accumulation and catabolism in the colonization of the murine urinary tract by Escherichia coli CFT073. Infect Immun **75**:5298–5304.

18. **Metzler, DE, Snell, EE**. 1952. Deamination of serine. II. D-Serine dehydrase, a vitamin B6 enzyme from Escherichia coli. J Biol Chem **198**:363–373.

19. **Maas, WK, Maas, R, McFall, E**. 1995. D-serine deaminase is a stringent selective marker in genetic crosses. Journal of bacteriology **172**:459–461.

20. **Nørregaard-Madsen, M, McFall, E, Valentin-Hansen, P**. 1995. Organization and transcriptional regulation of the Escherichia coli K-12 D-serine tolerance locus. J Bacteriol **177**:6456–6461.

21. **Anfora, AT, Welch, RA**. 2006. DsdX is the second D-serine transporter in uropathogenic Escherichia coli clinical isolate CFT073. J Bacteriol **188**:6622–6628.

22. **Moritz, RL, Welch, RA**. 2006. The Escherichia coli argW-dsdCXA genetic island is highly variable, and E. coli K1 strains commonly possess two copies of dsdCXA. J Clin Microbiol **44**:4038–4048.

23. **Cosloy, SD, McFall, E**. 1973. Metabolism of D-serine in Escherichia coli K-12: mechanism of growth inhibition. J Bacteriol **114**:685–694.

24. **Oren, Y, Smith, MB, Johns, NI, Kaplan Zeevi, M, Biran, D, Ron, EZ, Corander, J, Wang, HH, Alm, EJ, Pupko, T**. 2014. Transfer of noncoding DNA drives regulatory rewiring in bacteria. Proc Natl Acad Sci U S A **111**:16112–16117.

25. **Subashchandrabose, S, Hazen, TH, Brumbaugh, AR, Himpsl, SD, Smith, SN, Ernst, RD, Rasko, DA, Mobley, HL**. 2014. Host-specific induction of Escherichia coli

fitness genes during human urinary tract infection. Proc Natl Acad Sci U S A **111**:18327–18332.

26. **Ding, W, Baumdicker, F, Neher, RA**. 2016. panX: pan-genome analysis and exploration. bioRxiv 072082.

27. **Tsoy, O, Ravcheev, D, Mushegian, A**. 2009. Comparative genomics of ethanolamine utilization. J. Bacteriol. JB.00838–09.

28. **Kerfeld, CA, Aussignargues, C, Zarzycki, J, Cai, F, Sutter, M**. 2018. Bacterial microcompartments. Nat Rev Microbiol **16**:277–290.

29. **Kaval, KG, Garsin, DA**. 2018. Ethanolamine Utilization in Bacteria. mBio **9**:e00066–18.

30. **Brinsmade, SR, Paldon, T, Escalante-Semerena, JC**. 2005. Minimal functions and physiological conditions required for growth of salmonella enterica on ethanolamine in the absence of the metabolosome. J Bacteriol **187**:8039–8046.

31. **Penrod, JT, Roth, JR**. 2006. Conserving a Volatile Metabolite: a Role for Carboxysome-Like Organelles in Salmonella enterica. J. Bacteriol. **188**:2865–2874.

32. **Moore, TC, Escalante-Semerena, JC**. 2015. The EutQ and EutP Proteins are Novel Acetate Kinases Involved in Ethanolamine Catabolism: Physiological Implications for the Function of the Ethanolamine Metabolosome in Salmonella enterica. Mol Microbiol **99**:497–511.

33. **Bertin, Y, Girardeau, JP, Chaucheyras-Durand, F, Lyan, B, Pujos-Guillot, E, Harel, J, Martin, C**. 2011. Enterohaemorrhagic Escherichia coli gains a competitive advantage by using ethanolamine as a nitrogen source in the bovine intestinal content. Environmental Microbiology **13**:365–377.

34. **Thiennimitr, P, Winter, SE, Winter, MG, Xavier, MN, Tolstikov, V, Huseby, DL, Sterzenbach, T, Tsolis, RM, Roth, JR, Bäumler, AJ**. 2011. Intestinal inflammation

allows Salmonella to use ethanolamine to compete with the microbiota. Proceedings of the National Academy of Sciences **108**:17480–17485.

35. **Sintsova, A, Smith, S, Subashchandrabose, S, Mobley, HL**. 2018. Role of ethanolamine utilization genes in host colonization during urinary tract infection. Infect Immun **86**:e00542–17.

36. **Brooks, T, Keevil, CW**. 1997. A simple artificial urine for the growth of urinary pathogens. Lett Appl Microbiol **24**:203–206.

37. **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.

38. **Thomason, LC, Costantino, N, Court, DL**. 2007. *E. coli* genome manipulation by P1 transduction. Current protocols in molecular biology **79**:1.17.1–1.17.8.

39. **Kitagawa, M, Ara, T, Arifuzzaman, M, Ioka-Nakamichi, T, Inamoto, E, Toyonaga, H, Mori, H**. 2005. Complete set of ORF clones of Escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research. DNA Res **12**:291–299.

40. **Sturms, R, Streauslin, NA, Cheng, S, Bobik, TA**. 2015. In Salmonella enterica, Ethanolamine Utilization Is Repressed by 1,2-Propanediol To Prevent Detrimental Mixing of Components of Two Different Bacterial Microcompartments. J Bacteriol **197**:2412–2421.

41. **Parsons, JP, Dinesh, SD, Deery, E, Leech, HK, Brindley, AA, Heldt, D, Frank, S, Smales, CM, Lunsdorf, H, Rambach, A, Gass, MH, Bleloch, A, McClean, KJ, Munro, AW, Rigby, SEJ, Warren, MJ, Prentice, MB**. 2008. Biochemical and structural insights into bacterial organelle form and biogenesis. J. Biol. Chem. **283**:14366–14375.

42. **Wood, DE, Salzberg, SL**. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome biology **15**:R46.

43. **Li, H, Durbin, R**. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics **26**:589–595.

44. **Bankevich, A, Nurk, S, Antipov, D, Gurevich, AA, Dvorkin, M, Kulikov, AS, Lesin, VM, Nikolenko, SI, Pham, S, Prjibelski, AD, Pyshkin, AV, Sirotkin, AV, Vyahhi, N, Tesler, G, Alekseyev, MA, Pevzner, PA**. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol **19**:455–477.

45. **Seemann, T**. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics **30**:2068–2069.

46. **Treangen, TJ, Ondov, BD, Koren, S, Phillippy, AM**. 2014. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome biology **15**:1–15.

47. **Letunic, I, Bork, P**. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res **44**:W242–5.

48. **Price, MN, Dehal, PS, Arkin, AP**. 2010. FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One **5**:e9490.

49. **Dixit, PD, Pang, TY, Studier, FW, Maslov, S**. 2015. Recombinant transfer in the basic genome of Escherichia coli. Proc Natl Acad Sci U S A **112**:9070–9075.

50. **Schmittgen, TD, Livak, KJ**. 2008. Analyzing real-time PCR data by the comparative CT method. Nature protocols

51. **Johnson, JR, Stell, AL**. 2000. Extended virulence genotypes of Escherichia coli strains from patients with urosepsis in relation to phylogeny and host compromise. The Journal of infectious diseases **181**:261–272.

52. **Johnson, JR, O'Bryan, TT, Delavari, P, Kuskowski, M, Stapleton, A, Carlino, U, Russo, TA**. 2001. Clonal relationships and extended virulence genotypes among Escherichia coli isolates from women with a first or recurrent episode of cystitis. The Journal of infectious diseases **183**:1508–1517.

53. **Luo, Y, Ma, Y, Zhao, Q, Wang, L, Guo, L, Ye, L, Zhang, Y, Yang, J**. 2012. Similarity and divergence of phylogenies, antimicrobial susceptibilities, and virulence factor profiles of Escherichia coli isolates causing recurrent urinary tract infections that persist or result from reinfection. J Clin Microbiol **50**:4002–4007.

54. **Johnson, JR, Porter, S, Johnston, B, Kuskowski, MA, Spurbeck, RR, Mobley, HL, Williamson, DA**. 2015. Host Characteristics and Bacterial Traits Predict Experimental Virulence for Escherichia coli Bloodstream Isolates From Patients With Urosepsis. Open Forum Infect Dis **2**:ofv083.

55. **Hayashi, K, Morooka, N, Yamamoto, Y, Fujita, K, Isono, K, Choi, S, Ohtsubo, E, Baba, T, Wanner, BL, Mori, H, Horiuchi, T**. 2006. Highly accurate genome sequences of Escherichia coli K-12 strains MG1655 and W3110. Mol Syst Biol **2**:2006 0007.

56. **Cao, X, Cavaco, LM, Lv, Y, Li, Y, Zheng, B, Wang, P, Hasman, H, Liu, Y, Aarestrup, FM**. 2011. Molecular characterization and antimicrobial susceptibility testing of Escherichia coli isolates from patients with urinary tract infections in 20 Chinese hospitals. J Clin Microbiol **49**:2496–2501.

57. **Zhang, L, Foxman, B, Marrs, C**. 2002. Both urinary and rectal Escherichia coli isolates are dominated by strains of phylogenetic group B2. J Clin Microbiol **40**:3951–3955.

58. **Moreno, E, Andreu, A, Pigrau, C, Kuskowski, MA, Johnson, JR, Prats, G**. 2008. Relationship between Escherichia coli strains causing acute cystitis in women and the fecal E. coli population of the host. J Clin Microbiol **46**:2529–2534.

59. **Rijavec, M, Starcic Erjavec, M, Ambrozic Avgustin, J, Reissbrodt, R, Fruth, A, Krizan-Hergouth, V, Zgur-Bertok, D**. 2006. High prevalence of multidrug resistance and random distribution of mobile genetic elements among uropathogenic Escherichia coli (UPEC) of the four major phylogenetic groups. Curr Microbiol **53**:158–162.

60. **Skjøt-Rasmussen, L, Hammerum, AM, Jakobsen, L, Lester, CH, Larsen, P, Frimodt-Møller, N**. 2011. Persisting clones of Escherichia coli isolates from recurrent urinary tract infection in men and women. J Med Microbiol **60**:550–554.

61. **Nowrouzian, FL, Adlerberth, I, Wold, AE**. 2006. Enhanced persistence in the colonic microbiota of Escherichia coli strains belonging to phylogenetic group B2: role of virulence factors and adherence to colonic cells. Microbes and Infection **8**:834–840.

62. **Nowrouzian, FL, Wold, AE, Adlerberth, I**. 2005. Escherichia coli strains belonging to phylogenetic group B2 have superior capacity to persist in the intestinal microflora of infants. The Journal of infectious diseases **191**:1078–1083.

63. **Bouatra, S, Aziat, F, Mandal, R, Guo, AC, Wilson, MR, Knox, C, Bjorndahl, TC, Krishnamurthy, R, Saleem, F, Liu, P, Dame, ZT, Poelzer, J, Huynh, J, Yallou, FS, Psychogios, N, Dong, E, Bogumil, R, Roehring, C, Wishart, DS**. 2013. The human urine metabolome. PLoS One **8**:e73076.

64. **Guo, K, Li, L**. 2009. Differential 12C-/13C-isotope dansylation labeling and fast liquid chromatography/mass spectrometry for absolute and relative quantification of the metabolome. Anal Chem **81**:3919–3932.

65. **Connolly, JP, Roe, AJ**. 2016. When and where? Pathogenic Escherichia coli differentially sense host D-serine using a universal transporter system to monitor their environment. Microb Cell **3**:181–184.

66. **Scarlett, FA, Turner, JM**. 1976. Microbial metabolism of amino alcohols. Ethanolamine catabolism mediated by coenzyme B12-dependent ethanolamine ammonia-lyase in Escherichia coli and Klebsiella aerogenes. Journal of general microbiology **95**:173–176.

67. **Gupta, A, Dwivedi, M, Mahdi, AA, Khetrapal, CL, Bhandari, M**. 2012. Broad identification of bacterial type in urinary tract infection using (1)h NMR spectroscopy. J Proteome Res **11**:1844–1854.

68. **Anfora, AT, Halladin, DK, Haugen, BJ, Welch, RA**. 2008. Uropathogenic Escherichia coli CFT073 is adapted to acetatogenic growth but does not require acetate during murine urinary tract infection. Infection and Immunity **76**:5760–5767.

69. **Hagan, EC, Lloyd, AL, Rasko, DA, Faerber, GJ, Mobley, HLT**. 2010. Escherichia coli global gene expression in urine from women with urinary tract infection. PLoS Pathogens **6**:e1001187.

70. **Wolfe, AJ**. 2005. The acetate switch. Microbiol Mol Biol Rev **69**:12–50.

71. **Mariño, E, Richards, JL, McLeod, KH, Stanley, D, Yap, YA, Knight, J, McKenzie, C, Kranich, J, Oliveira, AC, Rossello, FJ, Krishnamurthy, B, Nefzger, CM, Macia, L, Thorburn, A, Baxter, AG, Morahan, G, Wong, LH, Polo, JM, Moore, RJ, Lockett, TJ, Clarke, JM, Topping, DL, Harrison, LC, Mackay, CR**. 2017. Gut microbial metabolites limit the frequency of autoimmune T cells and protect against type 1 diabetes. Nat Immunol **18**:552–562.

72. **Vance, JE, Tasseva, G**. 2013. Formation and function of phosphatidylserine and phosphatidylethanolamine in mammalian cells. Biochim Biophys Acta **1831**:543–554.

73. **Schmitt, J, Noble, A, Otsuka, M, Berry, P, Maitland, NJ, Rumsby, MG**. 2014. Phorbol ester stimulates ethanolamine release from the metastatic basal prostate cancer cell line PC3 but not from prostate epithelial cell lines LNCaP and P4E6. British journal of cancer **111**:1646.

74. **Cotton, PB**. 1972. Non-dietary lipid in the intestinal lumen. Gut **13**:675–681.

75. **Alteri, CJ, Himpsl, SD, Mobley, HL**. 2015. Preferential use of central metabolism in vivo reveals a nutritional basis for polymicrobial infection. PLoS Pathog **11**:e1004601.

76. **Kendall, MM, Gruber, CC, Parker, CT, Sperandio, V**. 2012. Ethanolamine Controls Expression of Genes Encoding Components Involved in Interkingdom Signaling and Virulence in Enterohemorrhagic Escherichia coli O157:H7. mBio **3**

77. **Jones, PW, Turner, JM**. 1984. Interrelationships Between the Enzymes of Ethanolamine Metabolism in Escherichia coli. Journal of General Microbiology **130**:299–308.

78. **Wang, J, Yan, D, Dixon, R, Wang, YP**. 2016. Deciphering the Principles of Bacterial Nitrogen Dietary Preferences: a Strategy for Nutrient Containment. MBio **7**

Figure Legends Dadswell et al

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

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

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

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

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

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

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

A. Growth of U1, U1 Δ*eutB* mutant and complement in modified M9 plus 10 mM ethanolamine. B. Growth of U1, U1Δ*eutE* mutant and complement in modified M9 with 10 mM ethanolamine. C. Growth of U1, U1 Δ*eutB* mutant and complement in AUM plus 10 mM ethanolamine. D. Growth of U1, U1Δ*eutE* and complement in AUM with 10 mM ethanolamine. In A-D growth of U1 in control medium without ethanolamine is shown as open circles. E. Percentage change in ethanolamine concentration measured by HPLC over 24 hours of U1, U1Δ*eutB* and U1Δ*eutE*, and their complements in AUM with initial 10mM ethanolamine. Significant difference with wild-type*** $P<0.001$, 1-way ANOVA. All Values are Mean ± SEM. N=3

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

Competition of U1 vs A: U1Δ*eutB* with 10mM ethanolamine, B: U1Δ*eutE* with 10mM ethanolamine. Mann-Whitney U test . *p<0.05, **p<0.01, ***p<0.001. Values are Mean ± SEM. N=3.

Figure 6. Phylogenetic distribution of *E. coli* urine isolates from this study and 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