A Novel Two-Component Signaling System Facilitates Uropathogenic *Escherichia coli*’s Ability to Exploit Abundant Host Metabolites

Wentong Cai¹, Yvonne Wannemuehler¹, Giuseppe Dell’Anna², Bryon Nicholson¹, Nicole L. Barbieri¹,³, Subhashnie Kariyawasam⁴, Yaping Feng⁵, Catherine M. Logue¹, Lisa K. Nolan¹, Ganywu Li¹*¹

¹Department of Veterinary Microbiology and Preventive Medicine, College of Veterinary Medicine, Iowa State University, Ames, Iowa, United States of America, ²Laboratory Animal Resources, College of Veterinary Medicine, Iowa State University, Ames, Iowa, United States of America, ³Department of Veterinary and Biomedical Sciences, Pennsylvania State University, University Park, Pennsylvania, United States of America, ⁴Laurence H. Baker Center for Bioinformatics and Biological Statistics, Iowa State University, Ames, Iowa, United States of America

Abstract

Two-component signaling systems (TCSs) are major mechanisms by which bacteria adapt to environmental conditions. It follows then that TCSs would play important roles in the adaptation of pathogenic bacteria to host environments. However, no pathogen-associated TCS has been identified in uropathogenic *Escherichia coli* (UPEC). Here, we identified a novel TCS, which we termed Kgus/Kgur (Kgus: α-ketoglutarate utilization sensor; Kgur: α-ketoglutarate utilization regulator) in UPEC CFT073, a strain isolated from human pyelonephritis. Kgus/Kgur was strongly associated with UPEC but was found only rarely among other *E. coli* including commensal and intestinal pathogenic strains. An in vivo competition assay in a mouse UTI model showed that deletion of kgus/kgur in UPEC CFT073 resulted in a significant reduction in its colonization of the bladders and kidneys of mice, suggesting that Kgus/Kgur contributed to UPEC fitness in vivo. Comparative proteomics identified the target gene products of Kgus/Kgur, and sequence analysis showed that TCS Kgus/Kgur and its targeted-genes, c5032 to c5039, are encoded on a genomic island, which is not present in intestinal pathogenic *E. coli*. Expression of the target genes was induced by α-ketoglutarate (α-KG). These genes were further shown to be involved in utilization of α-KG as a sole carbon source under anaerobic conditions. Kgus/Kgur contributed to the regulation of the target genes with the direct regulation by Kgur verified using an electrophoretic mobility shift assay. In addition, oxygen deficiency positively modulated expression of kgus/kgur and its target genes. Taken altogether, this study describes the first UPEC-associated TCS that functions in controlling the utilization of α-ketoglutarate in vivo thereby facilitating UPEC adaptation to life inside the urinary tract.

Citation: Cai W, Wannemuehler Y, Dell’Anna G, Nicholson B, Barbieri NL, et al. (2013) A Novel Two-Component Signaling System Facilitates Uropathogenic *Escherichia coli*’s Ability to Exploit Abundant Host Metabolites. PLoS Pathog 9(6): e1003428. doi:10.1371/journal.ppat.1003428

Editor: Matthew A. Mulvey, University of Utah, United States of America

Received December 1, 2012; Accepted April 25, 2013; Published June 27, 2013

Copyright: © 2013 Cai et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was partly supported by the USDA NRICGP Microbial Functional Genomics program (grant no. 2008-3560418805). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: liganwu@iastate.edu

Introduction

Urinary tract infection (UTI) is one of the most common bacterial infections in humans and is a significant clinical issue worldwide. Annually, UTIs are associated with ~7 million office visits, 1 million emergency room visits, 100,000 hospitalizations, and $1.6 billion in healthcare costs [1]. In fact, ~80–90% of community-acquired UTIs are caused by uropathogenic *Escherichia coli* (UPEC) [2]. Many virulence factors are required for UPEC to cause UTIs. Typically, UTIs begin with urethral contamination with UPEC from the bowel [3]. Successful attachment to the uroepithelium requires specific adhesins including P, type 1 and other fimbriae (such as F1C, S, M, and Dr fimbriae) [4,5]. UPEC may then ascend the urethra to enter the bladder and kidneys, where several highly regulated virulence factors, including fimbriae, secreted toxins (hemolysin, Vat, Sat, and CNF) and polysaccharide capsule, may contribute to colonization and pathogenesis [6]. It is likely that UPEC’s ability to colonize the urinary tract and cause disease is affected by its adaptive responses to local environmental cues, including changes in nutrient availability.

Urine is an important host environment that UPEC encounter after transition from the intestine to urinary tract. From the standpoint view of *E. coli* nutrition, urine is a mixture of amino acids and small peptides in low concentrations, with the notable exception of urea, which is abundant [7]. UPEC adapt to growth in urine by using peptides and amino acids as their primary carbon source for fitness [8]. About 85% of UPEC strains retain a complete *dsdCXA* locus [9] and these genes are responsible for the detoxification of D-serine allowing UPEC to use D-serine as the sole carbon and nitrogen source in urine [9]. Iron is another vital nutrient for UPEC but free iron concentrations in the mammalian host are extremely low. Consequently, UPEC have evolved multiple systems and strategies to obtain needed iron during infection including ent, iro, chu, sit, iutA, and fimA [10–13]. In addition, the urinary tract is considered to be a nucleotide-...
Author Summary
Successful colonization requires bacterial pathogens to adapt their metabolism to the conditions encountered in particular infection sites. Two-component signaling systems (TCSs) enable bacterial pathogens to sense and respond to environmental cues, thus mediating their adaptation to environmental change. Though many TCSs that have been characterized in commensal E. coli have been associated with UPEC pathogenesis, no characterized TCS has been significantly associated with UPEC strains. Here, we characterized a UPEC-associated TCS that was localized to a genomic island. This novel TCS and its target genes were involved in anaerobic utilization of α-ketoglutarate, an abundant metabolite in UPEC infection site-renal proximal tubules, thus contributed to UPEC fitness in vivo. Our results also suggest that this TCS controls a variant tricarboxylic acid (TCA) cycle, the first described in E. coli, which links the oxidative and reducing branches under anaerobic conditions. Similar TCA branches have been identified in other bacterial pathogens that have adapted to oxygen-limited, obligate anaerobic conditions, and/or intracellular carbon sources. Therefore, this study provides new insight into the adaptation of bacterial pathogens to nutrient availability in vivo and makes possible the discovery of targets for antimicrobial treatment.

deficient environment. Study of gyrD and guaC deletion mutants showed that metabolism of nucleobases is required for UPEC colonization of the bladder [14,15]. Thus, UPEC metabolism appears to play a critical role in colonization and invasion of the host. Therefore, it is not surprising that acquisition of genomic islands encoding metabolic pathways is often essential for the colonization of host niches. Gene cluster ttcABC and trrRS, responsible for utilizing tetrathionate as an electron acceptor in the intestine, are encoded on Salmonella enterica subsp. enterica serotype Typhimurium pathogenicity island II and contribute to detoxification of tetrathionate and confer a competitive advantage in growth [16–18]. Gene nixA within a transferable genomic island in Helicobacter pylori aids in transport of nickel and consequently results in enhanced activity of urease, a key factor for fitness and virulence [19–22]. Besides islands encoding iron uptake systems, no other metabolic island has yet been characterized in UPEC even though their study would help understand the pathobiology and facilitate the identification of potential targets for novel therapeutic approaches to prevent this prevalent pathogen.

The pathways, conferring metabolic adaptation of bacterial pathogens to the host milieu, are usually controlled by both global and specific regulators. One mechanism used by most bacterial pathogens to sense and respond to nutrient availability is two-component signaling systems (TCSs). TCSs, composed of a membrane-bound sensor histidine kinase (HK) and a cytoplasmic response regulator (RR), have been implicated in regulating the response of bacteria to a wide array of signals and stimuli, including nutrients, quorum signals, antibiotics, etc. The recognition of physical or chemical signals by HK sensor domain typically triggers the modulation of its autophosphorylation activity. The phosphoryl group is then transferred to the RR, usually a DNA binding protein that functions by altering gene expression [23,24]. In E. coli K12, more than 30 TCSs have been identified and characterized to some extent [25–27]. Of these, several were shown to be involved in UPEC pathogenesis. TCS BarA/UvrY controlled efficient switching between glycolytic and gluconeogenic carbon sources and contributed to UPEC virulence [28]. Deletion of QscC, the kinase of a well-known TCS QscC/QscB, dysregulated nucleotide, amino acid, and carbon metabolism and resulted in attenuation of UPEC virulence [29,30]. In addition, PhoQ/PhoP [31] and the AirS system [32] have been linked to UPEC pathogenesis. To date, all TCSs linked to UPEC pathogenesis have been common to E. coli and are not pathogen-associated. Given the distinct features of environments colonized by UPEC, we hypothesize that UPEC may possess special TCSs that sense urinary tract-specific metabolic signal(s) and adapt UPEC metabolism to available nutrients.

Here, we describe the first TCS significantly associated with UPEC and its direct target genes encoded in a small genomic island. In addition, we demonstrate that the expression of these island genes is induced in response to α-KG, a nutrient found in high concentrations in renal proximal tubule cells, and that these genes are involved in the utilization of α-KG under anaerobic conditions. Our results suggest that KguS/KguR contributes to UPEC fitness in vivo by facilitating the utilization of a host abundant metabolite α-KG. This study provides a new perspective for understanding UPEC pathogenesis in vivo.

Results

Identification of a novel TCS in UPEC CFT073
TCSs enable bacteria to sense, respond, and adapt to a wide range of environments, stressors, and growth conditions [23]. Since UPEC encounter unique environmental conditions found within their host, as compared to commensal E. coli and diarrheagenic E. coli, we hypothesize that specific TCSs may exist and facilitate UPEC adaptation to certain host niches. Indeed, analysis of the UPEC CFT073 genome identified a putative TCS c5041/c5040 by the presence of specific domains in their respective predicted proteins (Fig. 1). Domains and three-dimensional structures of C5041 and C5040 were predicted by the threading method using the I-TASSER online server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) and subsequently refined using four-body contact potentials. C5041 is predicted to be a typical sensor kinase harboring a dimerization and histidine phosphorylation ATPase (DHp) domain and a histidine kinase-like ATPase domain (Fig. 1A). The C-terminal 288 amino acids of C5041 contains five short signature segments, H-, N-, G1-, F-, and G2-boxes, which are conserved across most histidine kinases [33,34]. The histidine residue within the H-box, which is very likely phosphorylated during the signaling process, was found at position 394. In the N-terminal section of C5041, two membrane helices were identified at position 15–35 and at position 297–317, respectively. Between the two transmembrane helices, a periplasmic domain, which might serve as a signal sensor, was identified according to the distribution of charged residues ahead and behind the postulated transmembrane helices. The periplasmic domain of C5041 is 25.3% identical to that of DctB, a sensor kinase protein for dicarboxylates transport in Sinorhizobium meliloti (of which crystal structure is available) [35], suggesting C5041 may sense analogs of succinate, which is sensed by DctB [36].

C5040 is 37% identical to the nitrogen regulatory protein C (NtrC) [37,38] at the amino acid sequence level. C5040 harbors the CheY-like receiver domain, phosphorylation site, AAA+ ATPase domain, and helix-turn-helix (HTH) domain (Fig. 1B). The receiver domain (position 13–123) catalyzes phosphoryl transfer from the HK to itself and regulates the output. The aspartate residue, which is predicted to be phosphorylated in the signal transduction cascade, was found at position 58. In the C-terminal part of C5040, a helix-turn-helix domain was identified, which is a DNA-binding domain that recognizes enhancer-like sequences. The CheY-like receiver domain, AAA+ ATPase...
domain, and helix-turn-helix (HTH) domains of C5040 are 32.5%, 51.7%, and 41.7% identical to that of NtrC, respectively. Such characteristics are consistent with our expectations for a TCS. Additionally, DNA sequence analyses were performed using the search engine at http://blast.ncbi.nlm.nih.gov/Blast.cgi and c5041 and c5040 did not show any homology to any characterized histidine kinase or response regulator genes, suggesting that C5041/C5040 is a novel TCS.

TCS genes c5041 and c5040 are significantly associated with UPEC and increase CFT073 fitness in vivo

Sequence analysis of all E. coli genomes available online showed that c5041/c5040 and its orthologs are absent from commensal and diarrheagenic E. coli, but present in UPEC strains. Also, genotyping of 517 E. coli clinical isolates including human diarrheagenic E. coli and UPEC (Table S3) for c5041 and c5040 revealed their differential distribution. Both genes of this TCS were amplified in 139 of 200 UPEC (69.5%) but were not detected in 12 enterotoxigenic E. coli, 66 enterohemorrhagic E. coli or 39 enteropathogenic E. coli isolates. Therefore, c5041/c5040 appears to be significantly associated with UPEC, raising the possibility that C5041/C5040 contributes to UPEC fitness or virulence in vivo.

To study the role of C5041/C5040 in UPEC fitness in vivo, we compared the UPEC CFT073 wild-type strain with its Δc5041/Δc5040 double deletion mutant strain for the ability to colonize the mouse urinary tract at 48 hours post-inoculation using an in vivo competition assay. The double deletion mutation was constructed by replacement of c5041/Δc5040 with the cat cassette conferring chloramphenicol resistance. Though Δc5041/Δc5040 mutants grew as well as the wild type (WT) in LB, the mutant was significantly outcompeted by the WT, with a nearly 10-fold reduction in the median CFU/g from both bladders and kidneys at 48 hours post-inoculation (Fig. 2A, P<0.05 for both organs). To verify that the impact on colonization is not due to a secondary mutation, in vivo complementation experiments were performed. A stable low-copy plasmid pGEN-MCS was used as it was shown to be well maintained in CFT073 up to 48 h even in the absence of antibiotic pressure [39]. Coding regions of c5041 and c5040 plus their predicted promoter region were cloned into pGEN-MCS, and the resultant construct was named p$c5041$/$c5040$. As shown in Fig. 2B and C, the Δc5041/Δc5040 mutant containing the empty vector (pGEN-MCS) demonstrated an expected colonization defect in bladder and kidneys colonization as compared to the WT (empty vector) (P<0.05) while mutants harboring the complementation plasmid (p$c5041$/$c5040$) were able to colonize the bladder and kidneys at wild-type levels at 48 hours post-inoculation. These results clearly demonstrate that c5041/c5040 has a role in UPEC colonization and fitness in vivo.

Proteomic determination of gene products targeted by C5041/C5040

To determine how C5041/C5040 affects UPEC fitness in vivo, 2D-DIGE was used to identify differentially expressed gene products in the proteomes of the WT and single deletion mutants of c5041 and c5040 grown aerobically in human urine. Human urine was used as the growth medium to simulate the urinary tract environment [40–42]. c5041 and c5040 single mutants were...
constructed by replacement with Chl\(^+\) gene, and the resultant mutants were named LMP100Chl (CFT073 \(\Delta c5041:\)Chl\(^+\)) and LMP101Chl (CFT073 \(\Delta c5040:\)Chl\(^+\)). Using a cutoff of 1.5-fold change, seventeen proteins in the mutants were differentially expressed, as compared to the WT (Fig. S1). Of these, five were induced and five were repressed in \(\Delta c5041:\)Chl\(^+\) mutant; while six were induced and five were repressed in the \(\Delta c5040:\)Chl\(^+\) mutant. To determine the identities of the differentially expressed proteins, protein spots in the gel were excised and subjected to enzymatic digestion with trypsin followed by tandem mass spectroscopy. Functions of these proteins primarily fell into three categories: metabolism, cell envelope constituents, and translational machinery (Table 1). C5035, which showed a 24.6-fold induction, is a putative 2-oxoglutarate dehydrogenase presumably involved in tricarboxylic acid (TCA) cycle. XylA, a xylose isomerase, was increased 8.7-fold. It catalyzes the interconversion between xylose and xylulose. SitA, induced 4-fold, is a putative periplasmic iron-binding protein. Several outer membrane proteins including OmpA, OmpC, OmpF and OmpX were uniformly induced in \(\Delta c5040:\)Chl\(^+\) mutant. 508 ribosomal proteins, such as RplL, RplQ, and RplF, were also differentially expressed in the mutants as compared to the WT. These results indicate that C5041/5040 regulate multiple genes, suggesting that they may serve as a pleiotropic two-component system.

Genomic island genes targeted by C5041/C5040 contribute to UPEC fitness in vivo

The most differentially expressed protein identified by proteomic analysis was C5035, which was induced nearly 25-fold in the \(\Delta c5041:\)Chl\(^+\) mutant. qRT-PCR results confirmed that c5035 was up-regulated in \(\Delta c5041:\)Chl\(^+\) mutant, but not in \(\Delta c5040:\)Chl\(^+\) mutant, as compared to the WT when grown aerobically in human urine (Fig. 3A). To rule out the possibility of polar effects, the Chl\(^+\) gene was removed from both \(\Delta c5041:\)Chl\(^+\) and \(\Delta c5040:\)Chl\(^+\) mutants to generate mutants LMP100 (CFT073 \(\Delta c5041\)) and LMP101 (CFT073 \(\Delta c5040\)), respectively. Interestingly, the expression of c5035 was detected at extremely low levels not only in the mutant strains \(\Delta c5041\) and \(\Delta c5040\), but also in the WT. No significant differences in the c5035 expression were detected among the WT, \(\Delta c5041\) and \(\Delta c5040\) mutants when grown in human urine under aerobic conditions (Fig. 3A). Note that in the \(\Delta c5041:\)Chl\(^+\) mutant, c5041 was replaced by the Chl\(^+\) gene, which transcribes in the same orientation as its downstream gene, c5040. We further compared the expression levels of RR c5040 in the WT, \(\Delta c5041\), and \(\Delta c5041:\)Chl\(^+\) mutants. The results showed that expression of c5040 was extremely low in the WT and \(\Delta c5041\) mutant, but it was significantly upregulated in the \(\Delta c5041:\)Chl\(^+\) mutant when grown in human urine under aerobic conditions (Fig. 3B). Therefore, we suspect that the upregulation of c5035 in \(\Delta c5041:\)Chl\(^+\) mutant was due to the constitutive expression of c5040 caused by the upstream Chl\(^+\) gene. If so, overexpression of c5040 in the \(\Delta c5041\) mutant should induce the expression of c5035. Thus, c5040 was cloned into plasmid pMAL-MCS under the control of promoter P\(_{ara}\), and the resultant plasmid transformed into the \(\Delta c5041\) mutant. Indeed, induction of c5040 significantly upregulated the expression of c5035 in the \(\Delta c5041\) mutant (Fig. 3C). We also tested such regulation under different culturing condition and the results revealed similar regulatory pattern. In summary, constitutive expression of c5040 induced the expression of c5035 regardless of growth conditions.

Further analysis of genomic localization of c5035 in CFT073 revealed that genes c5032-5039 including c5035, together with TCS genes c5041 and c5040, are encoded by a genomic island inserted between E. coli K12 genes tyrB and aphA1 (Fig. 4A). Not surprisingly, island genes c5032-5039 were strongly associated with UPEC, and about 91.4% of UPEC strains tested containing c5040/c5041 loci also possess c5032-5039 genes (data not shown).
Island genes \(c5032-5039\) contribute to anaerobic \(\alpha\)-KG utilization and \(\alpha\)-KG induces their expression under anaerobic conditions

The proteins encoded by \(c5032\) and \(c5034\), and \(c5035\) are homologous to the E1, E2 and E3 components of \(\alpha\)-KG dehydrogenase in \(E.\ coli\), with similarity of 64\%, 69\%, and 55\%, respectively; whereas, the proteins encoded by \(c5036\) and \(c5037\) are 70\% and 83\% similar to the beta- and alpha-subunits of succinyl-CoA synthetase from \(E.\ coli\). \(\alpha\)-KG dehydrogenase converts \(\alpha\)-KG to succinyl-CoA, which can be further transformed to succinate by succinyl-CoA synthetase [43]. \(c5038\) is predicted to encode a putative dicarboxylate transporter with 13 transmembrane alpha-helices (TMHMM program [44]), showing 49\% similarity to succinyl-CoA synthetase from \(E.\ coli\). The proteins encoded by \(c5038-5039\) are 70\% and 83\% similar to the beta- and alpha-subunits of succinyl-CoA synthetase from \(E.\ coli\). We further tested their growth in anaerobic conditions, no significant growth differences were observed among these strains (Fig. S5). However, when tested under anaerobic conditions, deletion of \(c5032-5039\) completely abolished the growth, and the \(\Delta c5032-5037\) and \(\Delta c5038-5039\) mutants also displayed statistically significant growth defects, as compared to the WT (Fig. S5A). We further tested their growth in M9 minimal medium with \(\alpha\)-KG as the sole carbon source. Under aerobic conditions, no significant growth differences were observed among these strains (Fig. S5). However, when tested under anaerobic conditions, deletion of \(c5032-5039\) completely abolished the growth, and the \(\Delta c5032-5037\) and \(\Delta c5038-5039\) mutants also displayed statistically significant growth defects, as compared to the WT (Fig. S5A). We further tested their growth in M9 minimal medium with \(\alpha\)-KG as the sole carbon source.

\[\text{Table 1. Differentially expressed proteins identified in mutant strains of UPEC CFT073 as compared to the wild-type when cultured in human urine.}\]

| Order | Name or ORF | Function | MW* \(/pI^b\) | No. of peptides matched | Total ion score | Total ion C.I.\(^c\) \% | Fold-Change\(^d\) |
|-------|-------------|----------|---------------|------------------------|----------------|------------------------|------------------|
| Differentially expressed proteins in \(\Delta c5041-\text{Chl}\) mutant | | | | | | | |
| 1 | C5035 | putative 2-oxoglutarate dehydrogenase | 49832/6.1 | 16 | 172 | 100 | 24.6 |
| 2 | XylA | xylose isomerase | 49634/5.63 | 16 | 63 | 100 | 8.7 |
| 3 | SitA | putative periplasmic iron-binding protein | 31544/8.7 | 12 | 284 | 100 | 4.0 |
| 4 | TufB | EF-Tu | 43427/5.3 | 7 | 556 | 95 | 2.2 |
| 5 | GlmA | glutamine synthetase | 52099/5.26 | 2 | 76 | 95 | 2.0 |
| 6 | YbiS | hypothetical protein | 33059/5.99 | 10 | 232 | 100 | 2.1 |
| 7 | SpeE | spermidine synthase | 32286/5.33 | 7 | 228 | 100 | 2.2 |
| 8 | RplF | 50S ribosomal protein L6 | 17763/9.49 | 9 | 208 | 100 | 2.2 |
| 9 | RplQ | 50S ribosomal protein L17 | 14355/11.05 | 6 | 115 | 100 | 3.2 |
| 10 | RplD | 50S ribosomal protein L17 | 15759/11.05 | 14 | 474 | 100 | 4.0 |

*Molecular weight.

\(^b\)Isoelectric point.

\(^c\)Confidence interval (Ion Confidence Interval % greater than 95 were considered significant).

\(^d\)Average from two biological repeats.

\(\text{doi:10.1371/journal.ppat.1003428.t001}\)
among the strains tested (data not shown). These results indicated that genomic island genes c5032-5039 specifically contribute to α-KG utilization under anaerobic conditions but play no or a very limited role in α-KG utilization under aerobic conditions.

The ability of α-KG to induce expression of c5032-5039 was also examined. Chromosomal c5032-lacZ and c5038-lacZ transcriptional reporter fusions in CFT073 were constructed in order to study the expression of operon c5032-5037 and operon c5038-5039, and β-galactosidase activity in M9 medium using α-KG (40 mM) as the sole carbon source was also determined. Under anaerobic conditions, α-KG significantly induced the expression of c5032 and c5038, as compared to the rich medium LB or M9 medium containing glycerol or glucose as the sole carbon source (Fig. 5B); while the analogs of α-KG such as 2-HO-glutarate, glutamate, glutarate, fumarate, or succinate could not. To verify that other target genes on the genomic island are also induced by α-KG, qRT-PCR was performed. RNA was isolated from bacteria grown in M9 medium (glycerol and TMAO) with or without inducer α-KG. As shown in Fig. S4, all of the genes tested were greatly induced in the presence of α-KG. We then examined the effect of α-KG concentration on the stimulation of c5038 expression. Bacteria were grown in M9 medium plus glycerol and TMAO in the presence of various amounts of α-KG. We found that the induction of c5038 by the presence of α-KG was dose-dependent with concentrations as low as 200 μM able to induce their expression under anaerobic conditions (Fig. 5C). These results indicate that island genes c5032-5039 are induced by α-KG and involved in anaerobic utilization of α-KG.

C5040 directly regulates the expression of c5032-5039 and contributes to α-KG utilization under anaerobic conditions

To determine the regulatory role of C5040 in α-KG induction of target genes, chromosomal c5032-lacZ and c5038-lacZ transcriptional reporter fusions in the Δc5040 mutant were constructed. Strains were grown in M9 medium containing glycerol and TMAO in the presence or absence of inducer α-KG under anaerobic conditions. As shown in Fig. 6A, in wild-type genetic background, c5032 and c5038 expression was considerably induced in response to α-KG; in contrast, deletion of c5040 abolished the expression of c5032 and c5038 to the background level under anaerobic conditions. Introduction of a plasmid construct p5041 carrying c5041 and c5040 genes with their native promoter into c5040 mutants (LMP206 and LMP209) markedly increased expression of c5032 and c5038; whereas, the empty plasmid vector did not affect the expression of c5032 and c5038. In addition, the regulation of other target genes (c5034, c5035, c5036, c5037, and c5039) in the genomic island in response to α-KG by C5040 was confirmed using qRT-PCR (Fig. S4). Together, these results indicate that α-KG induction of c5032-5039 is C5040-dependent (Fig. 6A).

To determine if C5040 directly regulates c5032 and c5038 expression, an electrophoretic mobility shift assay (EMSA) was performed. The promoter regions of c5032 and c5038 were predicted by BProm program (http://linux1.softberry.com) [45]. The potential binding sites for c5032 and c5038 were identified and found to be highly similar, with TGTGTG-N13-CGCGCA predicted by BProm program (http://linux1.softberry.com) [45]. DNA fragments containing the potential binding sites were then PCR amplified for use as probes (305 nucleotides in size for c5032, starting from −272 to +32 relative to translational start codon; 196 nucleotides in size for c5038, starting from −163 to +32 relative to translational start codon). Fragments amplified from the coding region of c5036 were used as negative controls. Gene c5040 encoding the RR was cloned into the expression vector pMal-c2x and fused with MBP-His6. MBP-C5040-His6 fusion protein retained its regulatory biological function since introduction of
the plasmid construct expressing MBP-C5040-His6 fusion protein into Δc5040 mutant activated the expression of c5038; whereas, the empty plasmid did not affect expression of c5038 (Fig. S5). As shown in Fig. 6B, the purified MBP-C5040-His6 fusion protein was able to shift the promoter fragments of both c5032 and c5038, but not the control fragment. Meanwhile, use of the purified MBP-His6 fusion protein without C5040 was not associated with a detectable protein-DNA complex (Fig. 6B). These results demonstrate that C5040 directly binds to the promoters of the two operons.

To determine if C5040 is involved in the utilization of α-KG as the sole carbon source under anaerobic conditions, the growth properties of the WT, Δc5040 mutants, and the complemented strains were compared. When grown in M9 medium with α-KG as the sole carbon source under anaerobic conditions, the Δc5040 mutant strains showed a growth defect as compared to the WT. The c5040-complemented strain (Δc5040 carrying pC5040/c5040) significantly increased growth in this medium; while the empty plasmid vector (pGEN-MCS) did not improve growth of the mutant (Δc5040) (Fig. 6C).

The role of C5041 in the expression of island genes c5032-5039

The regulatory role of HK C5041 was also determined. Chromosomal c5032-lacZ and c5038-lacZ transcriptional reporter fusions in the Δc5041 mutant (LMP106) were obtained to create LMP205 and LMP208, respectively. Strains were grown in M9 medium (supplemented with glycerol and TMAO) with or without inducer α-KG under anaerobic conditions. Deletion of c5041 abolished the expression of c5032 and c5038 to background levels in the presence of α-KG. Re-introduction of the plasmid pC5041 carrying the c5041 gene with its native promoter into LMP205 and LMP208 mutants significantly increased their expression (P<0.05), indicating that C5041 positively affects expression of the target genes (Fig. 7A). This suggests that phosphorylation is very important for the regulatory function of RR C5040. In
addition, the expression of other genes (c5034, c5035, c5036, c5037, and c5039) in the genomic island affected by C5041 was also confirmed using qRT-PCR (Fig S4). The growth phenotypes of the c5041 mutant and its complementation strain in difference medium were then tested. Interestingly, the growth of the WT and its derivative mutants (c5041 mutant and complementation strain) in all media tested was statistically indistinguishable (Fig 7B).

Oxygen tension modulates expression of c5032-5037 and c5038-5039 directly via the TCS C5041/C5040

Island genes c5032-5039 contribute to anaerobic utilization of α-KG but play a limited role under aerobic conditions causing us to suspect that oxygen might modulate expression of these genes. Indeed, the c5032-5037 operon encoding a putative α-KG dehydrogenase and a putative succinyl-CoA synthetase was upregulated >5-fold in a reduced oxygen environment. Similarly, the expression of the c5038-5039 operon, encoding a putative transporter of α-KG and a protein with unknown function, was upregulated 2-fold under anaerobic conditions, as compared to aerobic conditions (Fig 8A).

To determine if oxygen regulates c5032-5039 expression via C5041/C5040 but not through other independent pathways, the effects of c5041/c5040 double deletion on the expression of c5032-5039 were examined. Deletion of c5041/c5040 completely abolished expression of the two operons under both aerobic and anaerobic conditions, supporting the hypothesis that oxygen regulates these two operons via TCS C5041/C5040 (Fig. 8A).

In addition, we wished to determine if oxygen deficiency regulates expression of the TCS itself. Under anaerobic conditions, the expression of c5041 was about 3-fold greater than that observed under aerobic conditions (Fig. 8B), suggesting that c5041 expression is significantly induced in response to anaerobiosis. Similarly, under anaerobic conditions, c5040 expression is dramatically increased as compared to that under aerobic conditions (Fig 8B). These results suggested that oxygen deficiency upregulated the expression of c5041 and c5040 and raised the possibility that oxygen modulates the expression of c5032-5037 and c5038-5039 operons via controlling c5041/c5040 expression. However, these results did not rule out the possibility that anaerobiosis stimulates the typical TCS phosphotransfer reactions between C5041 and C5040, leading to increased DNA binding and up-regulation of target gene expression.

Discussion

One of the greatest challenges confronted by all microorganisms is adapting to rapid changes of nutrient availability in different habitats. In the course of evolution, bacteria have developed several mechanisms to sense and utilize available nutrient sources associated with particular niches or to favor the most efficiently metabolizable nutrient sources when exposed to a range of
choices. TCSs are major mechanisms enabling bacteria to couple environmental stimuli to adaptive responses [46]. The TCSs in *E. coli* K12 have been extensively studied [25–27]. Many TCSs that are common to both commensal and pathogenic *E. coli*, such as PhoQ/PhoP [31], QseC/QseB [29,30], BarA/UvrY [28], and the AirS system [32] have been shown to contribute to virulence by mediating bacterial adaption to the host environment. However, no pathogen-associated TCS has yet been characterized in *E. coli*.

Driven by the hypothesis that UPEC-associated TCSs exist to sense and respond to host environment signals distinct from those of *E. coli*.

Figure 6. The role of c5040. (A) Effects of c5040 on the expression of target genes in response to α-KG. Bacteria were anaerobically grown in M9 medium containing glycerol as sole carbon source and TMAO as electron acceptor in the absence or presence of α-KG. The expression of c5032-lacZ or c5038-lacZ in WT, Δc5040 (LMP107), and the complemented strains were represented by the β-galactosidase activity. (B) Non-radioactive EMSA studying the binding of MBP-C5040-His6 to the promoter regions. PCR products of c5032 promoter region (left panel) and c5038 promoter region (right panel) were used as probes. Purified MBP-C5040-His6 fusion protein was added in different concentration in each reaction mixture as indicated and MBP-His6 was used as negative control at 300 ng per each reaction. For the lane on the right on each panel, a negative control DNA fragment amplified from coding region of c5036 was used. DNA fragments were stained with SYBR green. (C) In vitro growth of Δc5040 mutants. Optical density of the CFT073 WT, Δc5040 mutant, and the complemented strain during growth in M9 medium containing α-KG as the sole carbon source under anaerobic conditions was measured. Growth bars represent the average measurement at each time point from triplicate experiments. doi:10.1371/journal.ppat.1003428.g006
of intestinal *E. coli*, we identified a novel TCS C5041/C5040 significantly associated with UPEC strains, but not with EHEC, EPEC or ETEC strains. This TCS activated the expression of a genomic island involved in transport and metabolism of α-KG under anaerobic conditions. In view of these findings, C5041 was renamed KguS (α-ketoglutarate utilization sensor) and C5040 renamed KguR (α-ketoglutarate utilization regulator). These results indicated that UPEC might actively import α-KG and that α-KG could be an important carbon source for UPEC in *vivo*. Consistent with this observation, deletion of the TCS or its target island genes resulted in a significant reduction in a UPEC’s colonization of the murine urinary tract. To our knowledge, this is the first report of a pathogen-associated TCS in *E. coli* that contributes to UPEC pathogenesis.

α-KG is an intermediate in the citric acid/tricarboxylic acid (TCA) cycle, which is a metabolic pivot for both catabolic and anabolic processes that supply key metabolic intermediates and energy in most eubacteria [47]. Many intermediates in the TCA cycle can be sensed by TCSs, which elicit an adaptive response in *E. coli*. Four-carbon (C4) dicarboxylate-sensing DcuS/DcuR was shown to control fumarate transportation (*dcuB*) and respiration (*frdABCD*) under anaerobic conditions, and under aerobic condi-

![Figure 7. The role of c5041.](image)

(A) Effects of c5041 on the expression of the target genes in response to α-KG. The expression of c5032-lacZ or c5038-lacZ in the WT, Δc5041 (LMP106), and the complemented strains were represented by the β-galactosidase activity. (B) *In vitro* growth of Δc5041 mutants. Optical density of CFT073 WT, Δc5041 mutant, and the complemented strain during growth in M9 medium containing α-KG as the sole carbon source under anaerobic conditions was measured. Growth bars represent the average measurement at each time point from triplicate experiments.

doi:10.1371/journal.ppat.1003428.g007
Figure 8. Oxygen tension modulates the expression of two operons in genomic island via TCS C5041/C5040. (A) Target genes' expression in C5041ΔC5040 mutants under aerobic and anaerobic conditions. The expression of c5032-lacZ or c5038-lacZ in WT and ΔC5041/c5040 mutant (LMP108) under aerobic and anaerobic conditions were compared. (B) Induction of c5041/c5040 expression by oxygen deficiency. c5041-lacZ and c5040-lacZ transcriptional fusion strains were grown in M9 medium containing α-KG as the sole carbon source under aerobic and anaerobic conditions, respectively. β-galactosidase activity was measured and the values shown are means plus standard deviations of triplicate samples.

doi:10.1371/journal.ppat.1003428.g008
colonization in the bladders and kidneys of mice. However, deletion of kgsS/KguR or its target genomic island genes did not affect growth in human urine in vitro under both aerobic and anaerobic conditions (data not shown). We reasoned that human urine might only be able to partially imitate the in vitro conditions of UPEC infection, whereas the contributions of KgsS/KguR to UPEC pathogenesis may be restricted to some particular site of infection or pathogenic process, which cannot be represented by urine.

Interestingly, Chlamydia trachomatis, an obligate intracellular bacterial pathogen of urogenital infections [60], possesses an incomplete TCA cycle. Because of their small genome, chlamydiae lack many metabolic pathways retaining only the functions for performing key steps and interconversions of metabolites from the host cells. The fact that chlamydiae lack the first three enzymes of the TCA pathway and obtains α-KG from their host cell via a membrane transporter [56,61] suggests that utilization of intracellular α-KG is important for the survival of this urogenital intracellular pathogen. It could be inferred from Chlamydia and hypothesized that α-KG might be one of the preferred carbon and energy sources for UPEC during infection. Previous studies supporting this hypothesis showed that the oxygen tension (PO2) in UPEC-infected tissue dropped to 0 mmHg within 3.5–4 h [62], suggesting UPEC encountered anaerobic conditions during infection. It was also reported that the concentration of α-KG in human cells and blood is about 10 μM, a concentration that was able to induce the expression of KguS/KguR-controlled pathway in vitro under anaerobic conditions in this study. The KguS/KguR-controlled transporter (C5035) differs from the constitutively expressed Kgp encoded by all E. coli [63,64] and was induced under anaerobic conditions. During infection, this transporter allows UPEC to import α-KG from the infected tissue into the cell, and α-KG dehydrogenase and succinyl-CoA synthetase allow UPEC to take advantage of α-KG under anaerobic conditions, thus contributing to UPEC’s in vivo fitness. More interestingly, proximal tubule cells in the kidney are able to accumulate α-KG. Proximal tubules are the primary site where organic anions (OA) of physiological, pharmacological, and toxicological importance are cleared from the body. Their clearance involves uptake of OA via OA-dicarboxylate exchange driven by α-KG gradients. Thus, intracellular concentration of α-KG in the proximal tubule cells is ~100–400 μM, which is 10 to 40-fold higher than that found in any other cells [63–67]. UPEC undergo an intracellular lifestyle where they form biofilm-like communities [68], and UPEC are able to infect proximal tubule cells [62,69–71]. However, very little is known about the metabolism of UPEC within these host cells. It is very likely that KguS/KguR-controlled anaerobic utilization pathway of α-KG would increase UPEC fitness during infections of renal proximal tubule cells. Future studies in our laboratory should help clarify these questions.

In summary, our findings suggested a model that describes a novel regulatory and metabolic pathway in UPEC. For optimal growth during infection, HK KguS may sense α-KG, an abundant metabolite in UPEC infection site-renal proximal tubules, and then phosphorylate the RR KguR, which finally activates the import and utilization of α-KG (Fig. 9). These findings provide compelling evidence that this first UPEC-associated TCS enables E. coli to sense infection niche-specific stimuli and adapt to local nutrient availability, thus increasing its in vivo fitness. Hopefully, this and future studies on this regulatory system and metabolism pathway could lead to a better understanding of correlations between bacterial metabolism and virulence, but also the molecular pathogenesis of UPEC.

Materials and Methods

Ethics statement

All animal procedures were conducted in accordance with NIH guidelines, the Animal Welfare Act and US federal law. The experimental protocol (Protocol number 4-11-7111-Z) for handling animals was approved by Institutional Animal Care and Use Committee at Iowa State University (IACUC). All surgery was performed under isoflurane anesthesia, and all efforts were made to minimize suffering.

The study using human urine was approved by Institutional Review Board (IRB ID: 04-171). Written informed consent was obtained from human participants and/or their legal guardians for urine collection.

Bacterial strains and culture conditions

Strains and plasmids used in this study are listed in Table S1. Aerobic growth was achieved by shaking in air at 180 rpm and anaerobic growth by incubating in a Bactron chamber (Sheldon Manufacturing, Inc., OR) filled with gas mixture (N2, 90%; CO2, 5%; H2, 5%). For genetic manipulations, all E. coli strains were grown routinely in lysogenic broth (LB) medium. For growth and gene expression studies, bacteria were generally grown aerobically or anaerobically in M9 minimal salts with certain carbon sources indicated, supplemented with 2 mM MgSO4, 0.1 mM CaCl2, and 1 mg/ml vitamin B1. When used, trimethylamine N oxide (TMAO) and di- or tri-carboxylates were present at 40 mM. Glucose (0.5% v/v) or glycerol (0.25% v/v) was added as energy substrates, as indicated. Fresh mid-stream human urine was collected from six male and female consenting donors, pooled, and filter sterilized. Selective antibiotics and IPTG were added when necessary at the following concentrations: ampicillin (Amp), 100 μg ml⁻¹; kanamycin (Kan), 50 μg ml⁻¹; chloramphenicol (Chl), 25 μg ml⁻¹; IPTG, 1 mM.

Recombinant DNA techniques

Polymerase chain reaction (PCR), DNA ligation, electroporation and DNA gel electrophoresis were performed according to Sambrook and Russel [72] unless otherwise indicated. All oligonucleotide primers were purchased from Integrated DNA Technologies (Iowa) and are listed in Table S2. All restriction and DNA-modifying enzymes were purchased from New England Biolabs and used based on the suppliers’ recommendations. Recombinant plasmids, PCR products, and restriction fragments were purified using QIAquick PCR purification kit or MinElute gel extraction kit (Qiagen, CA) as recommended by the supplier. DNA sequencing was performed at the DNA facility, Iowa State University. DNA and amino acid sequence analyses were performed using CloneManager software (Scientific & Educational Software, NC) and the search engine (http://blast.ncbi.nlm.nih.gov/Blast.cgi.) was used to identify conserved domain structures of the two-component system.

Deletion mutants were constructed using the lambda red recombinase system described by Datsenko and Wanner [73]. Chromosomal transcriptional lacZ fusion was constructed by homologous recombination of the suicidal plasmid pVIK112 carrying a fragment of complete 5′-region, 3′-region, or internal fragment of the target gene [74]. Briefly, PCR fragments of target genes were cloned into pVIK112 using EcoRI and XbaI sites. The resultant pVIK112 derivatives were introduced into CFT073ΔlacZ2E::Chl by conjugation and the integration was allowed to occur. Conjugants were selected and confirmed by PCR. Chl was removed by transforming pCP20 plasmid carrying flippase [75]. For complementation, the selectable sequences of genes plus their putative promoter regions were amplified from the CFT073 genome and independently cloned into pGEN-MGS [39] using EcoRI and SalI restriction sites.
To construct the plasmid overproducing MBP-KguR-His6 fusion protein, a 1.4-kb fragment containing the coding region of $kguR$ was obtained by PCR from genomic DNA using MalE/c5040Histag-F and MalE/c5040Histag-R carrying codons for His and subsequently cloned into pMAL-c2x (New England Biolabs) using BamHI and HindIII sites. The resultant plasmid contains MBP-KguR-His6 under the control of the Ptac promoter.

The construction of control vector pMAL-MCS was achieved by replacing the MalE coding sequence with the MCS fragment from pEGFP plasmid using NdeI and EcoRI sites.

**PCR genotyping**

The nucleotide sequences publicly available of $c5041/c5040$ genes were aligned by using the ClustalW2 program. The primers were selected from a relatively conserved region and on the basis of G/C content, annealing temperature and size of the amplicon. Multiplex PCR was carried out according to Johnson et al. [76].

**β-Galactosidase assays**

Overnight LB cultures of *E. coli* containing the gene of interest-lacZ fusions were washed with PBS once and then were diluted 1:100 in LB or M9 medium with the carbon sources indicated and grown at 37°C to log phase. These cultures were diluted 1:10 in Z buffer and assayed for β-galactosidase activity using ortho-Nitrophenyl-β-D-galactoside (ONPG) as a substrate as described previously [77].

**2D-DIGE**

To prepare bacterial protein samples, isolated colonies were used to inoculate LB and cultures grown overnight at 37°C, the culture was further diluted 1:50 to fresh LB medium and re-incubated; after OD 600 reached 1.0, 2 ml of culture (approximately 10⁸ CFU) was pelleted and the supernatant removed. Bacterial pellets were re-suspended in 20 ml of human urine and grown aerobically for 4 hours at 37°C. Bacteria grown in human urine were then collected by centrifugation and washed 3 times with sterile PBS. The resultant bacterial pellets were snap-frozen at −80°C and submitted to Applied Biomics, Inc. in California for Fluorescence difference in gel electrophoresis (2D-DIGE) [78].

**Reverse transcription PCR and quantitative real-time RT-PCR**

RNA from *E. coli* CFT073 and its derivatives was stabilized by RNAProtect Bacterial Reagent (QIAGEN) and extracted using an RNeasy Mini Kit (QIAGEN) with a one-hour in-tube DNase digestion (QIAGEN) to remove possible DNA contamination according to the manufacturer's instructions. Three biological replicates of the each sample were prepared. The concentration of RNA was determined using a Spectrophotometer (ND-1000) (NanoDrop).

For the co-transcription test, one microgram of total RNA was reverse transcribed in triplicate using random hexamers and ImProm-II reverse transcriptase (Promega). Reactions without reverse transcriptase were used as a DNA contamination control. cDNA was then used as template for subsequent PCR reactions. For quantitative real-time RT-PCR, melting curve analyses were performed after each reaction to ensure amplification specificity. Differences (n-fold) in transcripts were calculated using the relative comparison method, and amplification efficacies of each primer set were verified as described by Schmittgen et al [79]. RNA levels were normalized using the housekeeping gene *tus*.

**Figure 9. A schematic model for the regulatory mechanism of C5041/C5040 and its controlled target genes.** In certain host environments of UPEC, such as the epithelial cells of renal proximal tubule, where oxygen tension is low and α-KG is abundant, KguS/KguR can sense and respond to these stimuli by activating the expression of $c5038$ encoding dicarboxylate transporter and $c5032-5037$ genes encoding metabolic enzymes, involved in the utilization of α-KG as carbon and energy source. Possession and functionality of this system may meet special metabolic needs for UPEC during urinary tract infections. Question mark indicates function is unknown. Blue dots, α-KG (α-ketoglutarate); Succ-CoA, succinyl-CoA; Succ, succinate.

doi:10.1371/journal.ppat.1003428.g009
encoding DNA replication terminus site-binding protein as an endogenous control [80]. Quantitative real-time RT-PCR (qRT-PCR) was performed with a Bio-Rad IQ5 iCycler detection system using iScript one-step RT-PCR kit with SYBR Green (Bio-Rad) according to the manufacturer’s instruction [81].

Electrophoretic mobility shift assays

To study the binding of KguR to the DNA probe, electrophoretic mobility shift assays (EMSAs) were performed using the commercialized EMSA kit (Invitrogen, California) [82]. MBP-KguR-His6 fusion protein was purified to homogeneity using Ni-NTA Spin Columns and dialyzed against binding buffer. DNA probes were PCR amplified using specific primers and gel purified using a QIAGEN MinElute gel extraction kit. EMSAs were performed by adding increasing amounts of purified MBP-KguR-His6 fusion protein (0 to 300 ng) to the DNA probe (60 ng) in binding buffer (10 mM Tris (pH 7.5), 1 mM EDTA, 1 mM dithiothreitol, 50 mM KCl, 50 mM MgCl₂, 1 μg ml⁻¹ bovine serum albumin (NEB)) for 30 min at room temperature. Reaction mixtures were then subjected to electrophoresis on a 6% polyacrylamide gel in 0.5×TBE buffer (44.5 mM Tris, 44.5 mM boric acid, 1 mM EDTA, pH 8.0) at 200 V for 45 min. The gel was stained in 0.5×TBE buffer containing 1×SYBR Gold nucleic acid staining solution for 30 min.

Experimental UTIs

Mouse infection studies were performed according to the methods of Johnson et al [83]. Female CBA/J mice (six to ten weeks of age) were anesthetized and inoculated via transurethral catheterization with a 20 μl (2×10⁶ CFU) challenge inocula per mouse. Overnight LB cultures for CFT073 and the mutant strain were pelleted and resuspended in sterile PBS, mixed in equal number and adjusted to make challenge inocula. To determine the initial CFU/mL, dilutions of each inoculum were plated onto LB plates. Plates were enumerated. The numbers of colonies on selective plates were subtracted from those on LB plates to obtain the number of wild-type bacteria. In the case of in vivo complementation studies, recovered bacteria were plated on LB plates with ampicillin and LB with ampicillin and chloramphenicol. After 48 h, the mice were euthanized and the bladder and kidneys were aseptically removed, weighed, and homogenized in tubes containing PBS. Dilutions of the homogenized tissue were then plated onto duplicate LB plates with and without chloramphenicol. After overnight incubation, distinct colonies on plates were enumerated. The numbers of colonies on selective plates were subtracted from those on LB plates to obtain the number of wild-type bacteria. In the case of in vivo complementation studies, recovered bacteria were plated on LB plates with ampicillin and LB with ampicillin and chloramphenicol. Similarly, the numbers of colonies on ampicillin and chloramphenicol plates were subtracted from those on LB plates with ampicillin to obtain the number of wild-type bacteria. A group of 10 mice for each dual-strains challenge were used to determine alterations in fitness. The WT/Δc5041/Δc5040 competition assay was performed twice while other assays were performed once. For statistical analysis, a two-tailed Wilcoxon matched pairs test was used (Prism software, CA) and the threshold for statistical significance was a P value <0.05.

Supporting Information

Figure S1 Comparison of the proteomes from the WT, Δc5041 and Δc5040 mutants during growth in human urine. Whole cell proteins of WT, Δc5041::Chï, and Δc5040::Chï mutants cultured in urine were labeled with cy2, cy3, and cy5 respectively and analyzed by 2D-DIGE in a single gel. Red spots represent proteins induced in the mutant as compared to the WT; green spots indicate proteins repressed in the mutant as compared to the WT. Differentially expressed proteins were circled and numbered. (A) Comparison between the WT and Δc5041::Chï mutants. Ten proteins were identified to be differentially expressed. (B) Comparison between the WT and Δc5040::Chï mutants. Eleven proteins were identified to be differentially expressed.

Figure S2 Operon formation represented by reverse transcription (RT)-PCR results using cDNA and a negative control without reverse transcriptase. Primers were designed to span ORFs c5032 & c5034, c5034 & c5035, c5035 & c5036, c5036 & c5037, c5037 & c5038, c5038 & c5039, and c5039 & c5040. UPEC CFT073 were cultured in M9 medium containing glycerol as the sole carbon source. RNA was purified and reverse transcribed to cDNA. The RNA that was not reverse transcribed served as a negative control.

Figure S3 In vitro growth of island gene mutants in M9 medium containing α-KG as the sole carbon and energy source under aerobic conditions. Optical density of CFT073 WT and island gene mutants during growth in M9 medium containing α-KG as the sole carbon source under aerobic conditions was determined. Growth curves represent the average measurement at each time point from triplicate experiments.

Figure S4 Effects of C5041 and C5040 on the induction of target genes on genomic island by α-KG. Wild-type or mutant bacteria were anaerobically grown in M9 medium containing glycerol and TMAO in the absence or presence of α-KG. qRT-PCR expression values in the presence of α-KG are presented as relative values, as compared to that in the absence of α-KG. Error bars represent the standard deviations for 3 independent experiments.

Figure S5 The induction of c5038 expression by over-produced MalE/C5040 fusion protein. The expression of c5038-lacZ in the WT transformed with the plasmid vector control (pMAL), pMAL-c5040, or pMAL-malE/c5040 were represented by the β-Galactosidase activity. IPTG was present at 1 mM for induction of Pmal promoter.

Table S1 Strains and plasmids. The genotypes of all strains of E. coli utilized or constructed in this study and information about the plasmids used in this study.

Table S2 Oligonucleotides. Oligonucleotide sequences used as PCR primers.

Table S3 Prevalence of c5041/c5040 locus among E. coli isolates.

Acknowledgments

The authors gratefully acknowledge Drs. Michael J. Wannemuehler and Glenn J. Songer for providing access to an anaerobic chamber and Dr. Qijing Zhang for use of the BioRad iQ5 iCycler. We thank Ting Meng and Capricia Thompson for the help in performing transurethral catheterization of mouse. Also, we appreciate the kind gift of pGEN plasmids from Dr. Harry L. T. Mobley and intestinal pathogenic E. coli strains from Dr. Nancy Cornick.
Author Contributions

Conceived and designed the experiments: WC GL. Performed the experiments: WC YW GD BNL SK. Analyzed the data: WC CML

References

1. Russo TA, Johnson JR (2003) Medical and economic impact of extraintestinal infections due to Escherichia coli: focus on an increasingly important endemic problem. Microbes Infect 5: 449-456.
2. Mobley HLT, Warren JW (1996) Urinary Tract Infections: Molecular Pathogenesis and Clinical Management. Washington D.C.: American Society for Microbiology.
3. Kaper JB, Nataro JP, Mobley HL (2004) Pathogenic Escherichia coli. Nat Rev Microbiol 2: 123-140.
4. Panek R, Zalewski B, Koski O, Ferens M, Nowicki R et al. (2003) Molecular aspects of biogenesis of Escherichia coli Dfr: expression of phd and dfrA genes in B. cereus and E. coli. J Bacteriol 185: 55-67.
5. Johnson JR (1991) Virulence factors in Escherichia coli urinary tract infections. Clin Microbiol Rev 4: 80-108.
6. Johnson JR, Stefl AL (2000) Extended virulence genotypes of Escherichia coli strains from patients with urosepsis in relation to phylogeny and host compromise. J Infect Dis 181: 261-272.
7. Haugen JR, Pellet-S, Redford P, Hamilton H, Roesch PL, et al. (2007) In vivo gene expression analysis identifies genes required for enhanced colonization of the mouse urinary tract by uropathogenic Escherichia coli strain CFT073 dhaA. Infect Immun 75: 278-289.
8. Al-Ali CJ, Smith SN, Mobley HLT, (2009) Identification of a new iron-regulated virulence gene, iraA, in an extraintestinal pathogenic isolate of Escherichia coli. Infect Immun 77: 6179-6216.
9. Russo TA, Carlini UB, Johnson JR (2001) Identification of a novel type I fimbriae in a urinary tract infection patient. Clin Infect Dis 34: 1669-1671.
10. Bahrani-Mougeot FK, Buckles EL, Lockatell CV, Hebel JR, Johnson DE, et al. (2004) Uropathogenic Escherichia coli use d-serine deaminase to modulate infection of the murine urinary tract. Mol Microbiol 49: 55-67.
11. Opal SM, Cross AS, Genski P, Lyhte LW (1990) Aerobactin and alpha-hemolysin as virulence determinants in Escherichia coli isolated from human blood, urine, and stool. J Infect Dis 161: 794-796.
12. Russo TA, Carlini UB, Johnson JR (2001) Determination of a new iron-regulated virulence gene, iraA, in an extraintestinal pathogenic isolate of Escherichia coli. Infect Immun 70: 7156-7160.
13. Torres AG, Redford P, Batchelet S, Moritz RL, Pellet-S, et al. (2003) Uropathogenic Escherichia coli use d-serine deaminase to modulate infection of the murine urinary tract. Mol Microbiol 49: 55-67.
14. Fischer W, Windhager L, Rohrer S, Zeiller M, Karnholz A, et al. (2010) Strain-specific genes of Helicobacter pylori: genome evolution driven by a novel type IV secretion system and genomic island formation. Nucleic Acids Res 38: 6089-6101.
15. Hadjifrangiskou M, Kostaki M, Chen SL, Henderson JP, Greene SE, et al. (2011) A central metabolic circuit controlled by QseC in pathogenic Escherichia coli. Mol Microbiol 80: 1516-1529.
16. Oshima T, Aiba H, Masuda Y, Kanaya S, Sugiura M, et al. (2002) Novel TCS in UPEC pathogenesis. Mol Microbiol 46: 281-291.
17. Yamamoto K, Hiroa K, Oshima T, Aiba H, Usumi R, et al. (2005) Functional characterization in vitro of all two-component signal transduction systems from Escherichia coli. J Biol Chem 280: 1448-1456.
18. Tomenson H, Fernestig AK, Jonas K, Georgelin D, Molloy R, et al. (2006) The Escherichia coli BarA/UVRV two-component system is a virulence determinant in the urinary tract. BMC Microbiol 6: 27.
19. Hensel M (2000) Salmonella pathogenicity island 2. Mol Microbiol 36: 1015–1032.
20. Bahrani-Mougeot FK, Lockatell CV, Hebel JR, Johnson DE, et al. (2004) Identification of a new iron-regulated virulence gene, iraA, in an extraintestinal pathogenic isolate of Escherichia coli. Infect Immun 72: 7156-7160.
21. Lin C, Nau B, Nau J, Ma Q, Panjikar S, et al. (2008) C4-dicarboxylates sensing mechanism revealed by the crystal structures of DctB sensor domain. J Mol Biol 383: 49-61.
22. Hensel M (2000) Salmonella pathogenicity island 2. Mol Microbiol 46: 1015-1023.
23. Follmer C (2010) Ureases as a target for the treatment of gastric and urinary tract infections. J Clin Pathol 63: 424–430.
24. Hoch JA (2000) Two-component and phosphorelay signal transduction. Curr Opin Microbiol 3: 165-170.
25. Mizuno T (1997) Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of Escherichia coli. DNA Res 4: 161-168.
26. Oshima T, Aiba H, Masuda Y, Kanaya S, Sugimura M, et al. (2002) Transcriptome analysis of all two-component regulatory system mutations of Escherichia coli K-12. Mol Microbiol 46: 281-291.
55. Guest JR (1995) The Leeuwenhoek Lecture, 1995: Adaptation to Life without Oxygen. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences 350: 199–202.

56. Cordwell SJ (1999) Microbial genomes and “missing” enzymes: redefining biochemical pathways. Arch Microbiol 172: 269–279.

57. Piton SM, Mendz GL, Srinivasan S, Hazell SL (1999) The tricarboxylic acid cycle of Helicobacter pylori. Eur J Biochem 266: 256–267.

58. Tian J, Bryk R, Itoh M, Suematsu M, Nathan C (2005) Variant tricarboxylic acid cycle in Mycobacterium tuberculosis: identification of alpha-ketoglutarate decarboxylase. Proc Natl Acad Sci U S A 102: 10670–10675.

59. Lloyd AK, Rashko DA, Mobley HL (2007) Defining genomic islands and uropathogen-specific genes in uropathogenic Escherichia coli. J Bacteriol 189: 3332–3346.

60. Stephens RS, Kalman S, Lammel C, Fan J, Marathe R, et al. (1998) Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis. Science 282: 754–759.

61. Kubo A, Stephens RS (2001) Substrate-specific diffusion of select dicarboxylates through Chlamydia trachomatis PorB. Microbiology 147: 3133–3140.

62. Mielean K, Boekel J, Messon LE, Sandowal RM, Tanner GA, et al. (2008) Bacterial infection-mediated mucosal signalling induces local renal ischaemia as a defence against sepsis. Cell Microbiol 10: 1987–1998.

63. Seol W, Shatkin AJ (1992) Escherichia coli kgtP encodes an alpha-ketoglutarate permease. Proc Natl Acad Sci U S A 89: 3902–3906.

64. Seol W, Shatkin AJ (1991) Escherichia coli kgtP encodes an alpha-ketoglutarate transporter. J Biol Chem 266: 6409–6413.

65. Pritchard JB (1995) Intracellular alpha-ketoglutarate controls the efficacy of growth of Helicobacter pylori. J Bacteriol 177: 5586–5591.

66. Boyd TA, Goldstein L (1979) Kidney metabolite levels and ammonia production in acute acid-base alterations in the rat. Am J Physiol 236: E289–295.

67. Martin M, Ferrier B, Baverel G (1989) Transport and utilization of alpha-ketoglutarate by the rat kidney in vivo. Pflugers Arch 413: 217–224.

68. Anderson GG, Palermo JJ, Schilling JD, Roth R, Hessler J, et al. (2003) Intracellular bacterial biofilm-like pods in urinary tract infections. Science 301: 105–107.

69. Chassin C, Vimont S, Cluzeaud F, Bens M, Goujon JM, et al. (2008) TLR4 cycles of Helicobacter pylori. Eur J Biochem 266: 256–267.

70. Pichon C, Hechard C, du Merle L, Chaudray C, Boune I, et al. (2009) Uropathogenic Escherichia coli AL511 requires flagellum to enter renal collecting duct cells. Cell Microbiol 11: 616–628.