Figures and figure supplements

Genetically diverse uropathogenic *Escherichia coli* adopt a common transcriptional program in patients with UTIs

*Anna Sintsova et al*
Clinical UPEC isolates carry a highly variable set of virulence factors. Phenotypic and genotypic information about the strains can be found in Figure 1—figure supplement 1, Figure 1—figure supplement 2, Table 1, and Table 2. (A) Clinical UPEC isolates were examined for presence of 40 virulence factors. Virulence factors were identified based on homology using BLAST searches (≥80% identity, ≥90% coverage). The heatmap shows presence (black) or absence (white) of virulence factors across 14 UPEC strains. Hierarchical clustering based on presence/absence of virulence factors shows separate clustering of B1 isolates. (B) Log2 TPM for iron acquisition genes (top panel) and adhesins (bottom panel) in urine and patient samples. Gene expression of other virulence factors is shown in Figure 1—figure supplement 3. Correlations of virulence factor expression among in vitro and patient samples is shown in Figure 1—figure supplement 4. (C) Log2 TPM of fim (top panel) and flg (bottom panel) operons across the 14 UPEC strains during in vitro urine culture and human UTI.

DOI: https://doi.org/10.7554/eLife.49748.004
Figure 1—figure supplement 1. Growth curves for 14 clinical UPEC strains cultured in LB or filter-sterilized urine.
DOI: https://doi.org/10.7554/eLife.49748.005
Figure 1—figure supplement 2. Phylogenetic tree reconstruction of 14 clinical UPEC strains isolated in this study. Antibiotic resistance profiles are indicated by filled in black circles (as determined by VITEK2 system (BioMerieux).) Patients with recurrent UTIs are indicated by filled in black square. MG1655, EC958, UTI89 and CFT073 are included for reference. DOI: https://doi.org/10.7554/eLife.49748.006
**Figure 1—figure supplement 3.** Expression of virulence factor genes in urine and patient samples.

DOI: [https://doi.org/10.7554/eLife.49748.007](https://doi.org/10.7554/eLife.49748.007)
**Figure 1—figure supplement 4.** Correlations among in vitro and patient samples measured by Pearson correlation coefficient of normalized gene expression of 40 virulence factors plotted according to hierarchical clustering of samples.

DOI: [https://doi.org/10.7554/eLife.49748.008](https://doi.org/10.7554/eLife.49748.008)
Figure 1—figure supplement 5. Treatment with MICROBEnrich does not affect measures of gene expression. Gene expression of a panel of genes was measured for HM86 (n = 3), and HM56 (n = 2) after 5 hr culture in filter-sterilized urine. After isolation RNA samples were either treated with MICROBEnrich or left untreated. Gene expression for each gene was measured by qRT-PCR. ΔCt between the gene of interest and gapA is shown.

DOI: https://doi.org/10.7554/eLife.49748.009
Figure 2. Core genome expression in patients is highly correlated. The analysis details are described in Materials and methods, and figure supplements. (A)–(B) Histogram of Pearson correlation coefficients among all samples cultured in vitro (A) or isolated from patients (B) based either on core genome or accessory genome comparisons. Accessory genome includes genes that were found in at least two but fewer than 14 of the clinical isolates. (C) Correlations among in vitro and patient samples measured by Pearson correlation coefficient of normalized gene expression plotted according to hierarchical clustering of samples. (D) Pearson correlation coefficient among all samples cultured in vitro (URINE | URINE, median = 0.92), among all samples isolated from patients (PATIENT | PATIENT, median = 0.91), between samples cultured in urine and samples isolated from patients (URINE | PATIENT, median = 0.73), and between matching urine/patient samples (ex. HM14 | URINE vs HM14 | PATIENT), (URINE | PATIENT:matched, median = 0.74). (E) Principal component analysis of normalized gene expression of 14 clinical isolates in patients and in vitro urine cultures shows distinct clustering of in vitro and patient isolates.

DOI: https://doi.org/10.7554/eLife.49748.010
Figure 2—figure supplement 1. Saturation curves. Number of mapped reads was plotted against number of expressed genes detected for each sample (in vitro samples are shown in blue; patient samples are shown in red). Vertical line shows 3 million reads cut off at which samples appear to reach saturation.

DOI: https://doi.org/10.7554/eLife.49748.011
Figure 2—figure supplement 2. Expression ranges of core genome genes. (A) Percentage of genes in the core genome that are expressed at a given level (>1 TPM, >10 TPMs, >100 TPMs, >1000 TPMs, where TPMs are transcripts per million) is shown for patient samples that reached saturation (see Supplementary Figure 2) and corresponding in vitro samples. (B) Percentage of genes in the core genome that are expressed at a given level (>1 TPM, >10 TPMs, >100 TPMs, >1000 TPMs) is shown for patient samples that did not reach saturation and corresponding in vitro samples.

DOI: https://doi.org/10.7554/eLife.49748.012
Figure 2—figure supplement 3. Effect of phylogenetic group on core genome expression. (A) and (C) Clustering of UPEC strains cultured in filter-sterilized urine based on PCA analysis of core genome gene expression. (B) and (D) Clustering of UPEC isolated from patients based on PCA analysis of core genome gene expression. Samples in (A) and (B) are colored based on their phylogroup designation. Samples in (C) and (D) are colored based on whether the strain was isolated from a patient with recurrent UTI (Y) or without recurrent UTI (N).

DOI: https://doi.org/10.7554/eLife.49748.013
**Figure 3.** Patient-associated transcriptional signature is consistent with rapid bacterial growth. (A) The DESeq2 R package was used to compare in vitro urine cultures gene expression to that in patients. Each UPEC strain was considered an independent replicate (n = 14). Genes were considered upregulated (down-regulated) if log$_2$ fold change in expression was higher (lower) than 2 (vertical lines), and $P$ value < 0.05 (horizontal line). Using these cutoffs, we identified 149 upregulated genes, and 343 downregulated genes. GO/pathway analysis showed that a large proportion of these genes belonged to one of the four functional categories (see legend). For each category, only the genes that have met the significance cut off are shown. The sugar transporters upregulated in UTI patients are shown in figure supplement. (B) Mean normalized expression for genes belonging to differentially expressed functional categories/pathways. The number of up or down-regulated genes belonging to each category is indicated next to the category name.

DOI: https://doi.org/10.7554/eLife.49748.019
Figure 3—figure supplement 1. Gene expression of four sugar transporters upregulated in UTI patients. Heatmap shows Log2 of normalized gene expression of $ptsG$, $fruA$, $fruB$ and $gntU$ for each in vitro and patient sample. 
DOI: https://doi.org/10.7554/eLife.49748.020
Figure 4. UPEC optimize growth potential via resource reallocation during UTI. (A) Percentage of reads that aligned to the core genome (2653 genes) out of total mapped reads. (B) Percentage of core genome reads that mapped to r-proteins (ribosomal subunit proteins, 48 genes). (C) Percentage of core genome reads that mapped to catabolic genes (defined as genes regulated by Crp and present in the core genome (277 genes). (D) Percentage of core genome reads that mapped to amino acid biosynthesis genes (54 genes). The equivalent analysis of Subashchandrabose et al. (2014) dataset is shown in the figure supplement.

DOI: https://doi.org/10.7554/eLife.49748.023
Figure 4—figure supplement 1. Resource reallocation analysis of Subashchandrabose et al. (2014) dataset. Left panel. Percentage of core genome reads that mapped to r-proteins (ribosomal subunit proteins, 48 genes) in five clinical strains from Subashchandrabose et al. study. The outlier patient...
sample that has only 2% of core genome mapped to r-proteins could potentially be attributed to very low depth of sequencing for that sample (HM26, see Table 6). Right panel. Percentage of core genome reads that mapped to catabolic genes. URINE: in vitro culture in filter-sterilized urine, LB: in vitro culture in LB, PATIENT: human UTI.

DOI: https://doi.org/10.7554/eLife.49748.024
Figure 5. Increased expression of ribosomal subunit transcripts is a host specific response. (A) Growth curve for HM43 strain cultured in LB and filter-sterilized urine. (B) Percentage of HM43 core genome reads that mapped to ribosomal subunit proteins under different conditions (URINE: in vitro culture in filter-sterilized urine, LB: in vitro culture in LB, MOUSE: mice with UTI, PATIENT: human UTI. (C) Percentage of HM43 core genome reads that mapped to catabolic genes under different conditions.

DOI: https://doi.org/10.7554/eLife.49748.027
Figure 6. Differential regulon expression suggests role for multiple regulators in resource reallocation. Regulon expression for 8 out of 22 regulons enriched for genes downregulated in the patients. Expression of each gene in the regulon during in vitro culture (blue) or during UTI (red) is shown along the x-axis. Histograms show proportion of genes in the regulon expressed at any given level.

DOI: https://doi.org/10.7554/eLife.49748.030