Metatranscriptomics-based metabolic modeling of patient-specific urinary microbiome during infection

Running Title:

Metabolic modeling of urinary microbiome

Jonathan Josephs-Spaulding[1§], Hannah Clara Rettig[2], Johannes Zimmermann[1], Mariam Chkonia[2], Alexander Mischnik[2], Sören Franzenburg[3], Simon Graspeuntner[2,4], Jan Rupp[2,4], Christoph Kaleta[1]

[1] Research Group Medical Systems Biology, University Hospital Schleswig-Holstein Campus Kiel, 24105 Kiel University, Kiel, Schleswig-Holstein, Germany

[2] Department of Infectious Diseases and Microbiology, University of Lübeck, 23538 Lübeck, Germany

[3] Institute of Clinical Molecular Biology, University Hospital Schleswig-Holstein, Kiel University, Rosalind Franklin Strasse 12, 24105 Kiel, Germany

[4] German Center for Infection Research (DZIF), Partner Site Hamburg-Lübeck-Borstel-Riems, Lübeck, Germany

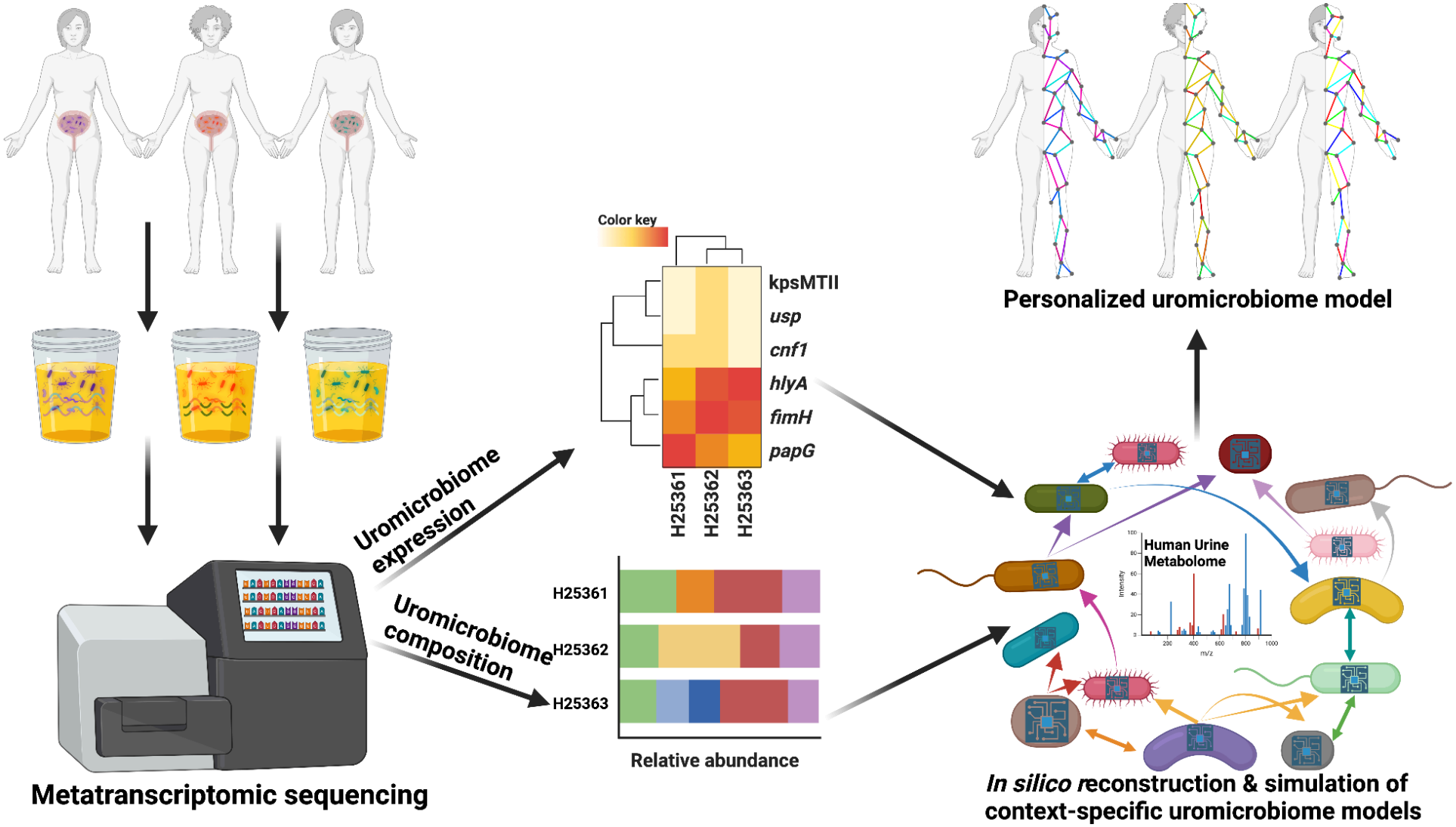
[§]Present address: The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

[*]Correspondance: c.kaleta@iem.uni-kiel.de

Abstract

Urinary tract infections (UTIs) are a major health concern which incurs significant socioeconomic costs in addition to substantial antibiotic prescriptions, thereby accelerating the emergence of antibiotic resistance. To address the challenge of antibiotic-resistant UTIs, our approach harnesses patient-specific metabolic insights to hypothesize treatment strategies. By leveraging the distinct metabolic traits of pathogens, we aim to identify metabolic dependencies of pathogens and to provide suggestions for targeted interventions. Combining patient-specific metatranscriptomic data with genome-scale metabolic modeling, we explored the metabolic aspects of UTIs from a systems biology perspective. We created tailored microbial community models to mirror the metabolic profiles of individual UTI patients' urinary microbiomes. Delving into patient-specific bacterial gene expressions and microbial interactions, we identify metabolic signatures and propose mechanisms for UTI pathology. Our research underscores the potential of integrating metatranscriptomic data using systems biological approaches, offering insights into disease metabolic mechanisms and potential phenotypic manifestations. This contribution introduces a new method that could guide treatment options for antibiotic-resistant UTIs, aiming to lessen antibiotic use by combining the pathogens' unique metabolic traits.

Keywords: Urinary tract infections, Metatranscriptomics, Systems biology, Metabolic modeling, Urogenital Microbiome


Color key
kpsMTII
usp
cnf1
hlyA
fimH
papG
H25361
H25362
H25363
Uromicrobiome expression
Uromicrobiome composition
H25361
H25362
H25363
Relative abundance
Metatranscriptomic sequencing
Personalized uromicrobiome model
Human Urine Metabolome
Intensity
m/z
In silico reconstruction & simulation of context-specific uromicrobiome models

Graphical Abstract

This study investigates the functional uromicrobiome across a female cohort. Initially, total RNA was extracted from patients' urine and sequenced to assess the metatranscriptome, providing insights into the structure and function of the uromicrobiome. Metatranscriptomic data was further utilized to construct context-specific uromicrobiome models, enabling an understanding of each patient's unique microbiome. Using metatranscriptomics and systems biology, we aimed to identify patient-specific dynamics and suggest various metabolic features that can be utilized in future studies for individualized intervention strategies.

I. Introduction

Urinary tract infections (UTIs) persist as one of the most common global health problems, with significant annual cases and healthcare costs reported worldwide (Flores-Mireles et al. 2015). Specifically, uncomplicated UTIs not associated with anatomical abnormalities frequently affect women and may result in recurrent, chronic infections (Hooton, 2012). UTIs are generally diagnosed by the presence of uropathogenic *E. coli (*UPEC*)*, a bacterium found in roughly 80% of UTI cases (Flores-Mireles et al., 2015). Even with identifying a primary causative factor, the growing issue of multi-drug resistant UTIs has compounded effective treatment; present antibiotic treatments have become increasingly ineffective and led to the elevation of UTIs from common infections to intricate medical challenges (Mediavilla et al. 2016; Shields et al. 2021). As a result, the need for a detailed understanding of microbial ecology within the human UTI microbiome has become pressing.

Scrutinizing microbial communities through shotgun metagenomics (sequencing of environmental DNA from microbial communities) is expected to investigate human diseases, but fewer studies aim to investigate metatranscriptomes (sequencing of environmental RNA from microbial communities) across human niches. A limitation of metatranscriptomics is the low resolution in distinguishing species-level organisms from one another compared to metagenomics. However, metatranscriptomics offers the unique advantage that it provides insight into active transcripts to describe microbial functions and can be used to describe how microbial communities respond to their human hosts (Ojala et al. 2023). For example, metatranscriptomics has been used to investigate the vaginal microbiome during vaginosis to study *G. vaginalis* and how this microbe circumvents stress from the antibiotic metronidazole by expressing genes for DNA repair (Deng et al. 2018). While RNA-seq alone has been utilized to investigate how uropathogenic *E. coli* responds in human urine and activates virulence mechanisms (Hagan et al. 2010), RNA-seq focuses on species-specific responses. At the same time, metatranscriptomics can investigate the gene response from entire human-associated microbial communities.

At present, a next-generation tool to investigate the interactions and phenotype of microbial communities is through the application of genome-scale metabolic models (GEMs). GEMS are mathematical reconstructions of the metabolic network of a species that provide a computational view into organismal physiology (O'Brien et al. 2015). An increasingly common and robust approach to investigating microbial response to specific conditions or contexts is the application of GEMs constrained by OMICs data (transcriptomics, proteomics, or metabolomics) to model microbial response to their environment effectively (Gopalakrishnan et al. 2023, Sen et al. 2023). While the AGORA2 database has provided a resource of 7203 gut-derived GEMS, which can be used to model drug transformation or bioremediation in a human niche accurately, context-specific GEMs have also been applied to study whole-body metabolism to

simulate the interaction between hosts and microbes (Thiele et al. 2020; Heinken et al 2023). Recently, a reproducible and flexible framework has been developed to combine metatranscriptomics with GEMs to improve the functional understanding of microbial communities (Zampieri et al. 2023).

This study employs a systems biology approach to explore UTI pathology, integrating metatranscriptomics with metabolic modeling to evaluate microbial roles from clinical isolates. We analyzed 19 UTI patient samples, assessing microbial diversity and gene expression to construct GEMs for pathogen metabolism in UTIs. Simulations used a generic urine medium based on the comprehensive Human Urine Metabolome database (Bouatra et al. 2013; Noronha et al. 2018), aiding in understanding patient-specific uromicrobiomes. Our findings offer insights into the variable microbial responses in UTIs, supporting the development of personalized treatments.

II. Material & Methods

Clinical assessment and study criteria

The samples used for the analyses herein are derived from patients and a study cohort sampled within the time frame November 2018 to August 2019 at the University Hospital Schleswig-Holstein and the University of Lübeck. The study was approved by the institutional review board ("Ethikkommission") of the University of Lübeck at the 1. February 2018 with the application number AZ18-009. The study comprises female patients with acute symptoms of uncomplicated UTIs as well as volunteer female students with cultured evidence of uropathogenic *E. coli* from self-sampled urine. All participants were 18 years or older at the time of sampling and sampling was performed before antibiotic therapy was initiated. The study excluded patients exhibiting physical malformations or comorbidities indicative of a complicated UTI. These include anatomical urogenital malformations (urethral stricture, bladder stones or tumors), previous urinary tract surgery, neurological or other diseases affecting bladder emptying, prior slight pelvis radiation, cytostatic treatment within the past year or during the study, immunosuppression, type 1 or 2 diabetes mellitus, pregnancy, recent use of vaginal probiotic therapies or vaginal spermicides, and participation in clinical studies potentially affecting urinary or renal function within 30 days before or during the study. The collected urine samples were processed as described in the following section. Acute cystitis severity was evaluated using a predetermined scoring system. Data acquired from these evaluations were used to interpret pathogen-associated factors for UTI prediction and delineate rUTI risk factors.

Metatranscriptomics sequencing & preprocessing

A complete visual summary of the workflow conducted in this study, from total RNA to the reconstruction of context-specific microbial community models, can be found in **Figure 1**.

Total RNA was extracted with MICROBEnrich and MICROBExpress mRNA enrichment kits (Ambien, Thermo Fisher Scientific) to reduce human contamination and isolate bacterial RNA. The samples underwent metatranscriptomic sequencing of 19 urine libraries at the Competence Centre for Genomic Analysis (CCGA) at Christian-Albrecht University, Kiel. The TruSeq stranded total RNA kit (Illumina) was used for RNA purification, focusing on shotgun metatranscriptomics to analyze the UTI microbiome and its functions. Initially, samples H25361-H25365 were sequenced on an Illumina MiSeq with 2x75bp paired-end reads. Subsequently, 14 additional samples were sequenced on an Illumina NovaSeq 6000 with 2x 50 bp paired-end reads as more resources became available. mRNA and rRNA were separated, post-quality control, into transcriptional units (TUs) and operational taxonomic units (OTUs), using the USEARCH algorithm. These were then clustered with PCA to detect technical bias sources **(Supplementary Figure S1**). Details on each sample and their barcoding are in **Supplementary Table S1**.

The quality control workflow processed total sequenced RNA to refine raw reads, separating mRNA and rRNA for metabolic modeling. Initially, FASTQC (Wingett et al. 2018) evaluated sequence quality, highlighting low-quality reads for removal before and after quality control. Cutadapt (Martin, 2011) and PRINSEQ (Schmieder et al. 2011) trimmed low-quality or repetitive sequences. Post-quality control, HISAT2 (Kim et al. 2019) mapped RNA to the human genome to eliminate human RNA contamination. SortMeRNA sorted the metatranscriptome, using seven indexed databases for various rRNA sequences, including the SILVA rRNA database (Yilmaz et al. 2014) and Rfam (Kalvari et al. 2021) for patient RNA indexing. Databases like Silva-Bac16s and Silva-Bac23s facilitated rRNA and tRNA extraction, focusing on bacterial 16S rRNA for taxonomy analysis. Non-mapped reads, considered mRNA, were used in functional assessment. Paired-end reads, created from combined forward and reversed samples, were filtered against RNA databases, with mapping rates detailed in **Supplementary Table S2**.

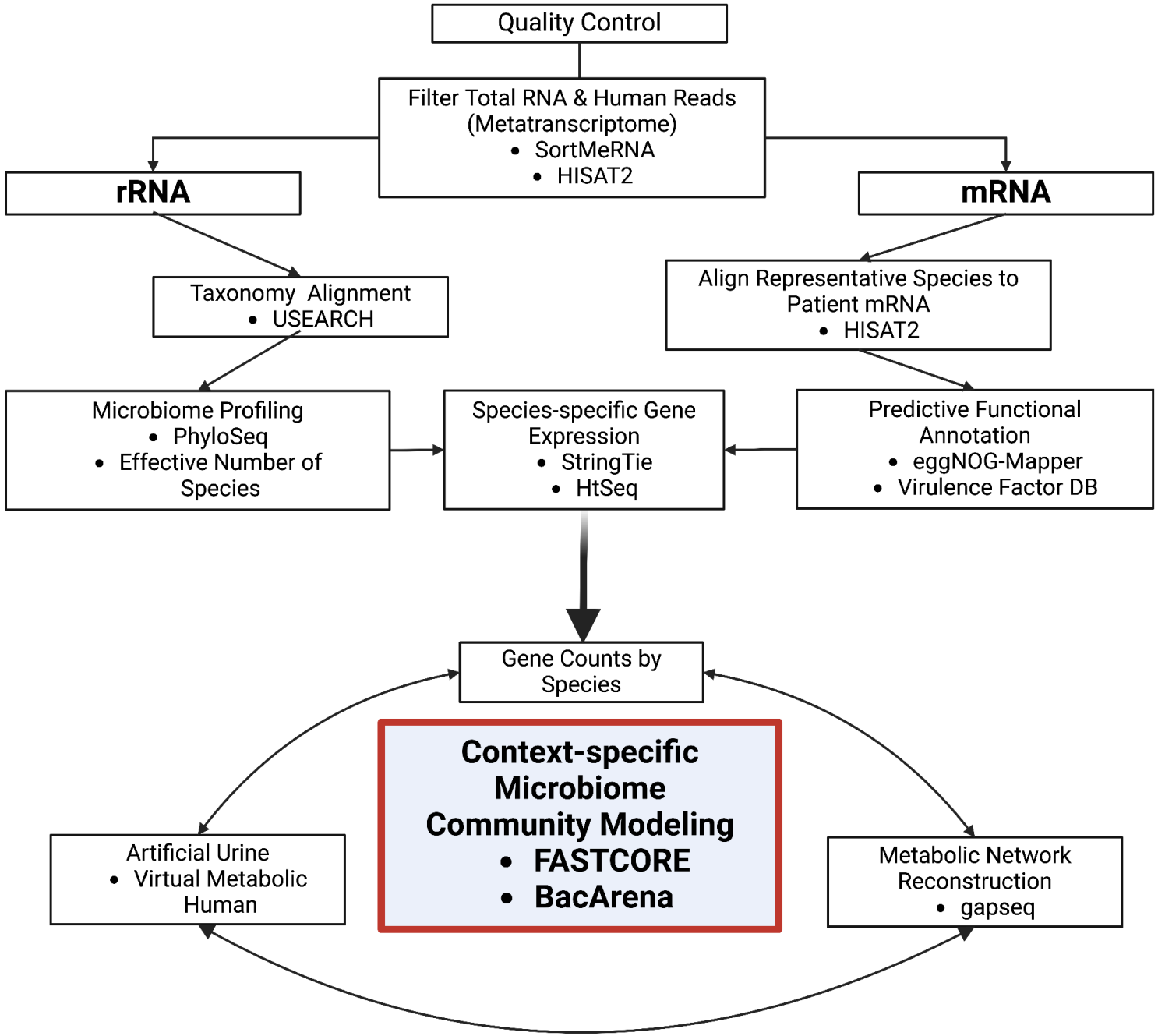

Quality Control
Filter Total RNA & Human Reads (Metatranscriptome)
• SortMeRNA
• HISAT2
rRNA
mRNA
Taxonomy Alignment
• USEARCH
Align Representative Species to Patient mRNA
• HISAT2
Microbiome Profiling
• PhyloSeq
• Effective Number of Species
Species-specific Gene Expression
• StringTie
• HtSeq
Predictive Functional Annotation
• eggNOG-Mapper
• Virulence Factor DB
Gene Counts by Species
Context-specific Microbiome Community Modeling
• FASTCORE
• BacArena
Artificial Urine
• Virtual Metabolic Human
Metabolic Network Reconstruction
• gapseq

Figure 1: Workflow of context-specific microbiome community model reconstruction based on metatranscriptomic data

Our workflow processed 19 raw metatranscriptomic datasets into context-specific microbiome community models for exploring UTI-related metabolic interactions. After splitting the metatranscriptomic data into rRNA and mRNA fractions, two separate pipelines were used to study taxonomic and functional profiles, respectively, before combining them again to reconstruct context-specific metabolic models.

rRNA taxonomy analysis

The separated rRNA was analyzed for microbial diversity and abundance. The USEARCH algorithm (Edgar, 2010) first identified and removed redundant sequences across samples, producing unique Operational Taxonomic Units (OTUs). These OTUs were denoised, clustered with a 99% bootstrap confidence, and purified from amplicons and chimeras, creating a table of zero-radius OTUs. Phyloseq (McMurdie et al., 2013) then identified each patient-specific community's taxonomy and abundance. The data was normalized by read counts, filtering out OTUs with low abundance, specifically those appearing in more than one sample and having over ten sequences. This process ensured a focus on significant microbial entities. Shannon's alpha diversity measured species diversity within communities, and the Shannon diversity values were transformed into the Effective Number of Species (ENS) to more accurately represent microbial ecosystem complexity. This method overcomes the limitations of traditional diversity indices by offering a more genuine count of equally abundant species, focusing on the most abundant species for subsequent gene-level expression counts and metabolic modeling to reconstruct context-specific models. The ENS formula is:

$ENS = e^{H}$

Where:

- ENS is the Effective Number of Species
- H represents the Shannon diversity index

Functional mRNA analysis

After SortMeRNA separation, the remaining mRNA was used to gather functional information and count species-specific gene expression for downstream community metabolic modeling. We employed HISAT2 (Kim et al. 2019), SamTools (Danecek et al. 2021), and HtSeq (Anders et al. 2015) for mRNA mapping, alignment, and counting. Reference genomes from NCBI, based on previous taxonomy analysis, were used to create species-specific genome indexes with HISAT2. We filtered out species unlikely to be part of the urinary tract microbiome, focusing on human pathogens or microbes common in human-associated environments. The selected microbes for community metabolic modeling are listed in **Supplementary Table S**3.

Each identified microbe was aligned to the mRNA from urine samples. SamTools converted aligned reads into SAM files for filtering and sorting, removing single-end reads. These were then transformed into BAM files, sorted by gene position. Stringtie (Pertea et al. 2015) with species-level gene annotations quantified gene expression, creating species-specific transcriptomes. HtSeq provided species gene-level counts, normalized to FPKM for consistent comparison across samples. Hierarchical clustering and gene expression distribution assessment revealed cohort similarities. For patient-mapped *E.coli*, gene-level FPKM counts were compared against a UPEC UTI89 reference genome (GCF_000013265.1). Virulence factors and categories were sourced from Virulence Factor DB (Liu et al. 2022) to plot gene expression associated with these categories. EggNOG Mapper (Cantalapiedra et al. 2021) extracted COG data from UPEC UTI89, identifying functional category distributions per patient UPEC.

Reconstructing microbiome community models

***In silico* human urine medium**

To enhance the accuracy of microbial community models, we created a urine metabolic medium based on data from the Urine Metabolome Database (Bouatra et al., 2013), which cataloged 445 metabolites in healthy human urine via methods like NMR and GC-MS. From these, 140 metabolites were integrated into the Virtual Metabolic Human (VMH) database as microbial exchange reactions, enabling the simulation of metabolite uptake and release (Noronha et al., 2019). We adjusted metabolite concentrations to millimolar units, basing calibration on average creatinine levels and incorporating water and acids in their conjugate base forms (Marinos et al. 2020). This procedure, including metabolite details, exchange reactions, and concentrations for simulations, is detailed in **Supplementary Table S4**. This process aids in refining microbial community models in gapseq and BacArena by offering a detailed list of metabolites compatible with VMH. We identified exchange reactions using VMH to connect microbial and human models, expressing all data in millimoles. For non-VMH metabolites, we used ModelSEED to find matching entries, facilitating the integration of extensive metabolic data into microbial community models.

Reconstruction & preprocessing of context-specific metabolic models

Our workflow employed the gapseq algorithm to automatically generate Genome-Scale Metabolic Models (GEMs) for each patient-specific microbe for simulating in silico microbial communities (Zimmermann et al. 2021). We sourced each bacterial species' genome from NCBI, and gapseq reconstructed all associated metabolic reactions, transport reactions, and pathways into GEMs. A draft metabolic model was refined by curating it with metabolic exchange fluxes from our in silico urine medium. We matched species-specific metabolic genes from the gapseq models to gene-level expression data based on chromosome position, forming a count matrix for each model. These models were then integrated into CobraPy (Ebrahim et al. 2013),

converted for MATLAB compatibility, and processed in MATLAB with patient-specific gene expression data corresponding to the gapseq model. Within MATLAB, using the COBRA Toolbox version 3.3 (Heirendt et al. 2019), gene expression data were linked to metabolic reactions via the 'mapExpressionToReactions' command. The methodology of Richelle et al. (2019a) helped identify the core set of reactions, selecting the top 25% with the highest expression levels as 'Active', designating the rest as 'Inactive'. The context-specific metabolic models, built on gapseq-derived GEMs of UTI bacteria and refined with our urine medium and microbial-specific gene expression, aimed for stability and consistency. This required all metabolic reactions to be flux-capable for microbial community simulations. Inconsistent reactions were removed using 'fastcc' from the fastcore suite (Vlassis et al. 2014), and fastcore reconstructed context-specific metabolic networks from the consistent core reactions.

Simulation of microbiome-specific community models

We created patient-specific community models using our reconstructed Genome-Scale Metabolic Models (GEMs) through the BacArena package (Bauer et al. 2017) to explore community growth, metabolic flux, and cross-feeding interactions. Two simulation types were conducted on microbial communities related to patient-specific UTIs: 1) 'non-context' simulations, involving gapseq models gapfilled with our in silico urine medium and made consistent with fastcc, and 2) 'context-specific' simulations, similar to the first type but further constrained with metatranscriptome-derived gene expression. These simulations aimed to identify whether context-specific modeling could reveal metabolic differences related to pathogenicity. For both non-context and context-specific models, GEMs representing species from each patient were formatted for R and integrated into BacArena with our artificial urine medium to simulate patient-specific microbial communities. Using a 100x100 arena based on species diversity identified by Effective Number of Species (ENS) analysis, we established species relative abundance by normalizing individual species abundance against the community's total abundance. We then simulated a community of 1000 organisms, termed 'Maximum Inoculation', calculated to mirror the relative abundance within the community for accurate representation in the simulation environment. This setup aimed for a dynamic, realistic portrayal of microbial interaction, ensuring even distribution of substrates and organisms. The community was simulated over four iterations, each representing an hour, in triplicates, allowing detailed observation of growth patterns, including metabolic outputs and cross-feeding activities among the community members.

Analysis of community simulations

We analyzed the metabolic flux distribution across patient samples using Principal Component Analysis (PCA). We assessed total flux variations stemming from different methodologies, namely non-context and context-specific models, applied to microbial community modeling. We utilized the chemical taxonomy from the Human Metabolome

Database to visualize the abundance (Z-score scaled) of metabolites across our cohort; this highlighted shifts in specific chemical classes (Wishart et al. 2022). For a deeper insight into the differences in metabolite concentrations between non-context and context-specific simulations, we conducted a Wilcoxon paired t-test. This test was chosen due to the inherent linkage between non-context and context-specific observations, ensuring more reliable differential expression detection and determining if the two sample groups had different population mean ranks, subsequently adjusting the results for the False Discovery Rate (FDR) using the Benjamini-Hochberg procedure. Our primary metric was the p-value, indicative of the likelihood that any observed change was merely by chance and plotted with a significance threshold of 0.05. Global metabolic flux, normalized by community biomass, was processed and samples were also categorized and labeled by their respective methodologies ('non-context' or 'context'). A global mean flux was computed for each patient pairing and presented as a barplot. Additionally, we formulated a presence/absence matrix for all metabolites in the cohort to delve into the nuances of context-specific metabolic models versus their non-context-specific counterparts. We also examined the log-fold changes in metabolite activity when comparing the two modeling strategies.

As a supplement to standard statistical tests to measure differences between our two groups and the impact of context-specific versus non-context-specific modeling on the metabolic profiles of microbiomes, the randomForest package was used to identify features of importance to context-specific modeling (Liaw et al. 2022). Our ultimate aim in applying random forests was to assess which metabolites are representative as the most important in distinguishing between the two modeling methodologies. The data was then partitioned into a training set (80% of the data) and a test set (20% of the data) to prepare for implementing the machine learning algorithms. We utilized two metrics, Mean Decrease Accuracy (MDA) and Mean Decrease Gini (MDG), to identify critical features observed from the random forest. MDA evaluates the impact on the model's accuracy when a feature's value is shuffled, indicating its predictive power. In contrast, MDG calculates a feature's contribution to class differentiation within the model. To underscore their relevance, we graphically presented the top 10 features, illustrating their importance. To assess the model's accuracy, the pROC package was used to make predictions on the test dataset and compare them to the actual values (Robin et al. 2011). A Receiver Operating Characteristic (ROC) curve was constructed, and the Area Under the Curve (AUC), precision, recall, and F1 score were computed to provide quantitative metrics for model evaluation.

To decipher microbial community interactions, particularly with an emphasis on metabolite cross-feeding mechanisms, we utilized the simulation outcomes from BacArena to examine microbial behaviors in a spatial context. Initially, a heatmap was constructed to provide an overarching view of species interactions based on their metabolic profiles, mainly focusing

on consumption and production of metabolites. Furthermore, we subset this data to highlight similarities and differences in metabolic cross-feeding potential in two *Lactobacillus* species (*L. inters* and *L. paucivorans*) in comparison to three gut-associated microbes (*E. coli, S. sonnei,* and *S. enterica*). Recognizing the pertinence of specific metabolites in influencing microbial survival, growth, and underlying pathology, a subsequent case study attempted to identify metabolites relevant to *E. coli*-directed UTI pathology. This involved examining metabolites such as ornithine, iron, and various sugar alcohols. This methodology facilitated a focused exploration of microbial interactions within specific patient samples, offering a detailed view of the intricate cross-feeding networks.

Statistical Testing

In our statistical analysis, we applied various tests to assess metabolic flux distribution across patient samples, comparing non-context and context-specific community models. We employed the Wilcoxon paired t-test to evaluate differences in metabolite concentrations between these models, adjusting the significance of our findings using the False Discovery Rate (FDR) through the Benjamini-Hochberg procedure. This approach enabled us to identify statistically significant variations in metabolite levels, specifically focusing on metabolites indicative of metabolic shifts during UTIs. Additionally, we visualized these differences using heatmaps and bar plots to illustrate the relative concentrations and global metabolic fluxes, providing insights into the metabolic dynamics within the UTI microbial ecosystem.

III. Results

UTI pathogenic microbiome composition and diversity vary across patients

Our analysis revealed diverse UTI microbiome compositions among patients, indicating significant biodiversity and inter-patient variation (**Figure 2**). **Figure 2A** converts absolute to relative abundance, showing a wide range of microbial species that define a unique microbial signature for each patient. We found variable abundance of microbial families, including uropathogenic genera such as *Citrobacter*, *Moraxella*, *Prevotella*, and *Salmonella*, alongside at least four *Lactobacillus* species (**Supplementary Figure S2**). Interestingly, while not typically primary UTI agents, these pathogens have roles in urinary tract colonization (Bitew et al. 2017; Gorelik et al. 2017; Sami et al. 2017; Ammitzboll et al. 2021). *Lactobacillus* species were prevalent, with one patient showing high abundance (G01). Principal Component Analysis (PCA) of species composition demonstrated the UTI microbiome's varied ecological landscape, with PC1 and PC2 explaining 37.16% and 15.23% of the variance, respectively (**Figure 2B**), underscoring the complex microbial interactions and community balance affecting UTI dynamics. The low explained variance in PC2 may be explained by our observation of outliers, especially by G01, which was observed to have a high abundance of *L. iners*. We explored the correlation between microbial community structure (PC1 scores) and Enterobacteriaceae (the

microbial family which contains *Escherichia, Klebsiella, Salmonella, Shigella,* and other human pathogens) diversity, finding a modest association (R-squared = 0.16) and a slight positive correlation (slope = 0.00891), suggesting that *Enterobacteriaceae* diversity variations are partly predictable by overall microbial community structure. Shannon's Alpha Diversity analysis showed a diversity range from 1.79 to 3.35 (**Figure 2C**), indicating variability in community diversity. We calculated the effective number of species (ENS) for each patient to inform the selection of species for downstream microbial community modeling, with 4 to 22 species used depending on the patient. Species were ranked by relative abundance, and those up to the ENS value were chosen for inoculation in community simulations (**Figure 2D**).

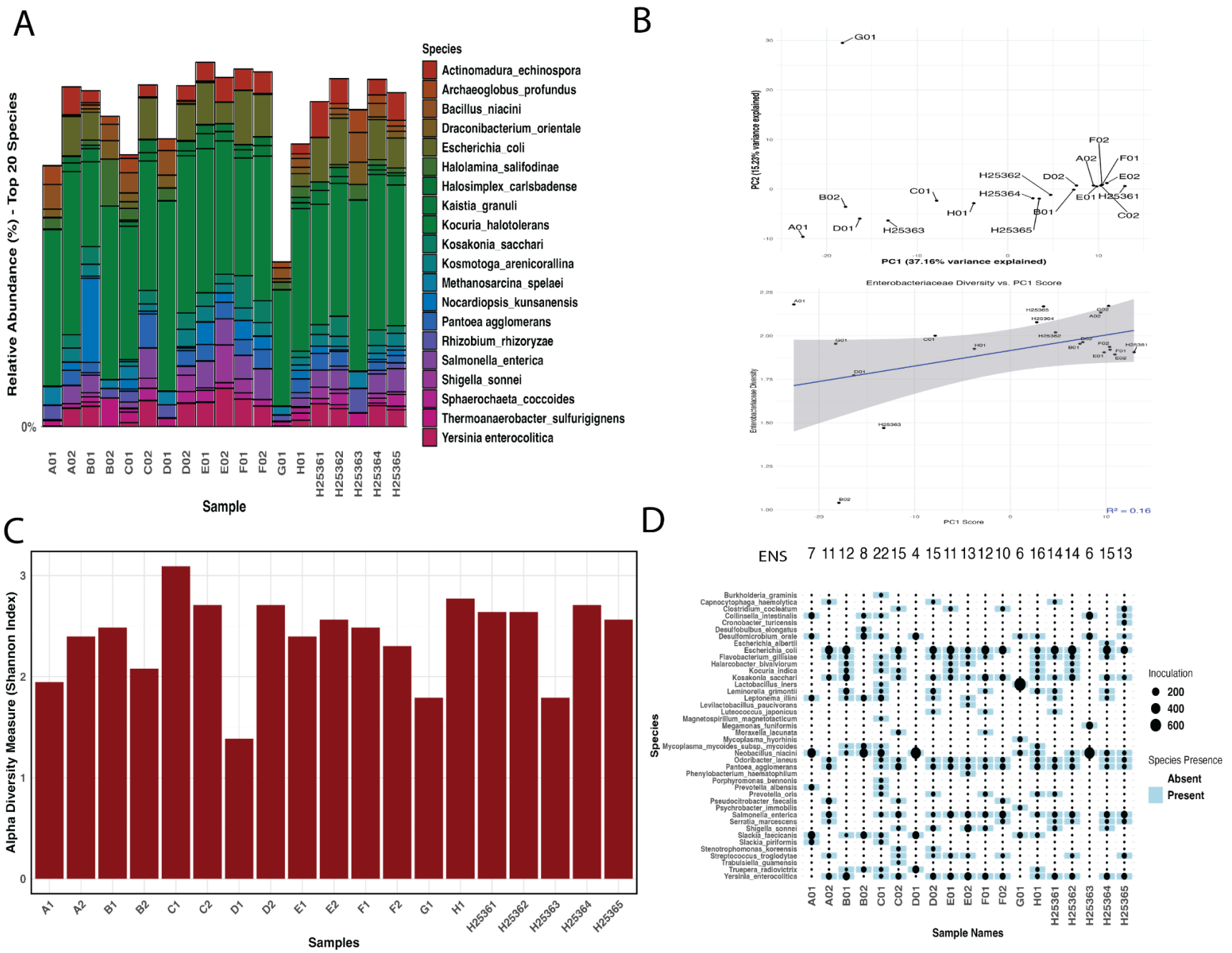

A
Relative Abundance (%) - Top 20 Species
0%
Sample
A01 A02 B01 B02 C01 C02 D01 D02 E01 E02 F01 F02 G01 H01 H25361 H25362 H25363 H25364 H25365
Species
Actinomadura_echinospora
Archaeoglobus_profundus
Bacillus_niacini
Draconibacterium_orientale
Escherichia_coli
Halolamina_salifodinae
Halosimplex_carlsbadense
Kaistia_granuli
Kocuria_halotolerans
Kosakonia_sacchari
Kosmotoga_arenicorallina
Methanosarcina_spelaei
Nocardiopsis_kunsanensis
Pantoea agglomerans
Rhizobium_rhizoryzae
Salmonella_enterica
Shigella_sonnei
Sphaerochaeta_coccoides
Thermoanaerobacter_sulfurigignens
Yersinia enterocolitica
B
PC2 (15.23% variance explained)
PC1 (37.16% variance explained)
Enterobacteriaceae Diversity vs. PC1 Score
Enterobacteriaceae Diversity
PC1 Score
R² = 0.16
C
Alpha Diversity Measure (Shannon Index)
Samples
A1 A2 B1 B2 C1 C2 D1 D2 E1 E2 F1 F2 G1 H1 H25361 H25362 H25363 H25364 H25365
D
ENS 7 11 12 8 22 15 4 15 11 13 12 10 6 16 14 14 6 15 13
Species
Burkholderia_graminis
Capnocytophaga_haemolytica
Clostridium_cocleatum
Collinsella_intestinalis
Cronobacter_turicensis
Desulfobulbus_elongatus
Desulfomicrobium_orale
Escherichia_albertii
Escherichia_coli
Flavobacterium_gillisiae
Halarcobacter_bivalviorum
Kocuria_indica
Kosakonia_sacchari
Lactobacillus_iners
Leminorella_grimontii
Leptonema_illini
Levilactobacillus_paucivorans
Luteococcus_japonicus
Magnetospirillum_magnetotacticum
Megamonas_funiformis
Moraxella_lacunata
Mycoplasma_hyorhinis
Mycoplasma_mycoides_subsp._mycoides
Neobacillus_niacini
Odoribacter_laneus
Pantoea_agglomerans
Phenylobacterium_haematophilum
Porphyromonas_bennonis
Prevotella_albensis
Prevotella_oris
Pseudocitrobacter_faecalis
Psychrobacter_immobilis
Salmonella_enterica
Serratia_marcescens
Shigella_sonnei
Slackia_faecicanis
Slackia_piriformis
Stenotrophomonas_koreensis
Streptococcus_troglodytae
Trabulsiella_guamensis
Truepera_radiovictrix
Yersinia_enterocolitica
A01 A02 B01 B02 C01 C02 D01 D02 E01 E02 F01 F02 G01 H01 H25361 H25362 H25363 H25364 H25365
Sample Names
Inoculation
200
400
600
Species Presence
Absent
Present

Figure 2: Comparative exploration of UTI pathogenic microbiome composition across UTI patients at the species level

Relative abundance across multiple microbial genera and species derived from rRNA data, as generated by PhyloSeq. This visualization highlights the diversity and distinct inter-patient variations in the composition of the UTI microbiome. (A) Top-20 species relative abundance highlights variability between the urinary microbiome during patient UTIs. (B) PCA based on species composition highlights patient-specific clustering. (C) A broad range of Alpha Diversity metrics highlights unique patient-specific differences across patient UTI microbiomes. (D) Heatmap representation of specific species' presence (light blue) or absence (white) in various samples. Overlaid points indicate the inoculation volumes, with their size reflecting the volume magnitude. Species are arranged alphabetically, and samples are displayed on the x-axis. ncorporating the Effective Number of Species (ENS) metric, derived from Shannon's alpha diversity, enhances the analysis by providing a nuanced understanding of microbial ecosystem complexity, focusing on the contribution of abundantly present species.

UPEC virulence strategies are diverse across UTI patients

In our next phase, we aligned gene expression with the UPEC UTI89 reference genome to extract FPKM values, focusing on the uropathogenic *E. coli* strain that epitomizes the primary cause of about 80% of UTIs (Flores-Mireles et al., 2015). This approach facilitated a precise examination of gene expression tied to common etiological factors of these infections. By comparing these profiles with the Virulence Factor Database (Liu et al., 2022), we identified key pathogenic mechanisms of UPEC UTI89. Virulence genes with high expression in patient samples highlighted the diversity of virulence strategies (**Figure 3**). Specifically, *fimA* and *fimI* genes, crucial for bacterial adhesion, were notably expressed in several samples, indicating their role in establishing infections (Baker et al., 2021). The ompA gene, related to invasion capabilities, was prevalent across nearly all samples, suggesting a common mechanism for UPEC strains to infiltrate host cells (Nielsen et al., 2020). Nutrient acquisition genes like *chuY*, *chuS*, *iroN*, and *fepA* were found in various samples, emphasizing the bacteria's metabolic adaptation to the host environment (Met et al., 2021; Edison et al., 2023). Additionally, the *map* gene, associated with effector protein delivery into host cells, showed high expression across most of our studied samples (Ramachandran et al., 2020). This variability in virulence gene expression underscores the complexity of UPEC pathogenicity and suggests a range of clinical manifestations across different infections.

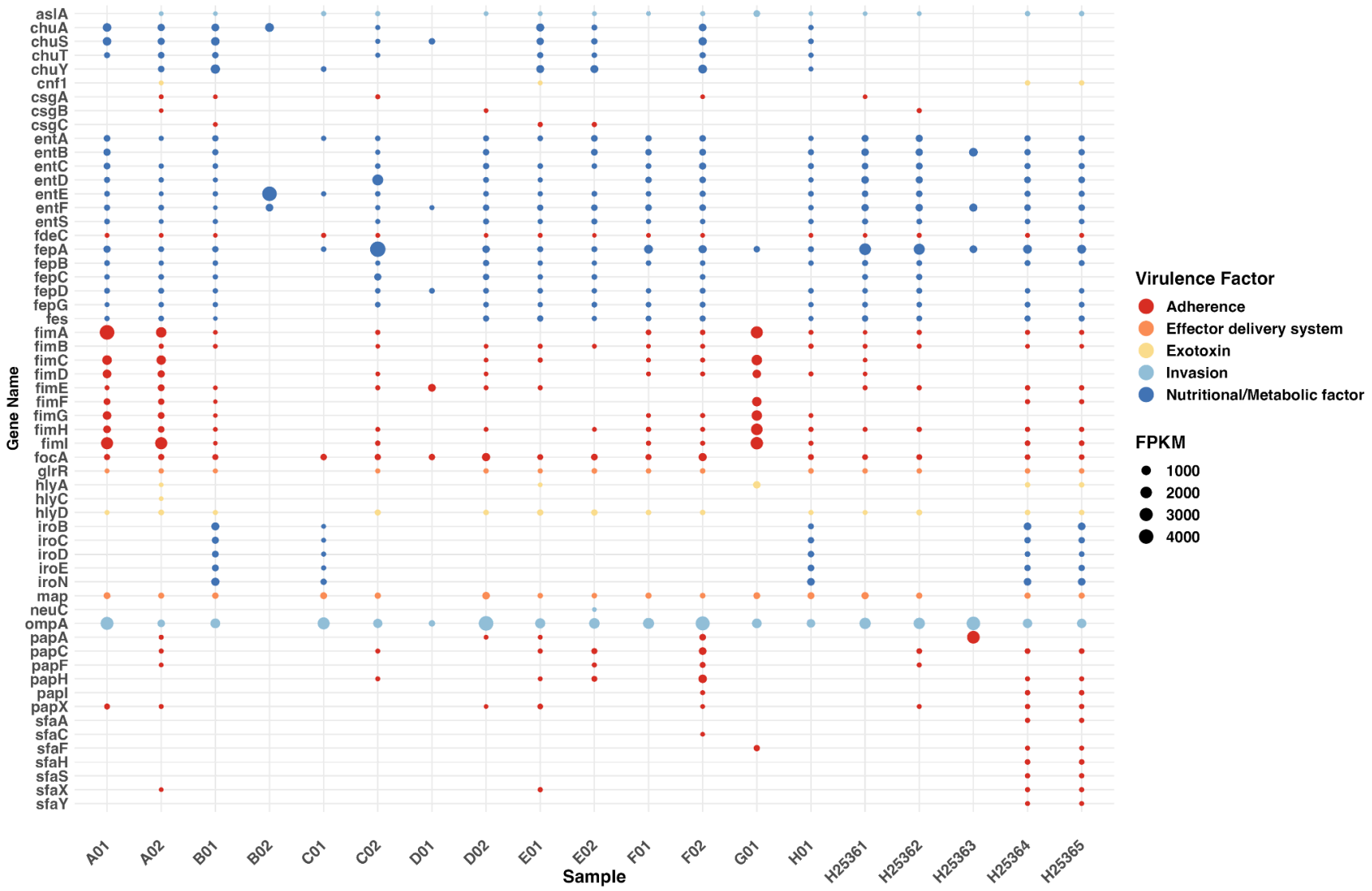

Gene Name
aslA
chuA
chuS
chuT
chuY
cnf1
csgA
csgB
csgC
entA
entB
entC
entD
entE
entF
entS
fdeC
fepA
fepB
fepC
fepD
fepG
fes
fimA
fimB
fimC
fimD
fimE
fimF
fimG
fimH
fimI
focA
glrR
hlyA
hlyC
hlyD
iroB
iroC
iroD
iroE
iroN
map
neuC
ompA
papA
papC
papF
papH
papI
papX
sfaA
sfaC
sfaF
sfaH
sfaS
sfaX
sfaY
A01
A02
B01
B02
C01
C02
D01
D02
E01
E02
F01
F02
G01
H01
H25361
H25362
H25363
H25364
H25365
Sample
Virulence Factor
Adherence
Effector delivery system
Exotoxin
Invasion
Nutritional/Metabolic factor
FPKM
1000
2000
3000
4000

Figure 3: **Mapping of Patient-specific UPEC Gene Expression to Virulence Factor Database Highlights Differing Virulence Differing Pathogenic Strategies**

Showcasing genes' normalized expression levels (FPKM), further categorized based on their association with virulence factors. Each dot represents a distinct patient-specific gene expression, color-coded by virulence factor category.

Differences in metabolic flux between non-context-specific and context-specific microbiome models

Integrating transcriptomic data into our analysis has illuminated context-specific microbial pathomechanisms during infections. Non-context-specific models simulate patient-specific scenarios without transcriptomic data, while context-specific models incorporate this data, refining metabolic flux predictions. Principal Component Analysis (PCA) showed that context-specific models resulted in a more uniform metabolic response among patients, with the first two principal components explaining 20.29% and 10.01% of variance (**Figure 4A**). This uniformity contrasts with the variability seen in non-context-specific models, which exhibited outliers indicative of extraneous metabolic reactions. Additionally, context-specific models showed reduced total metabolic flux compared to non-context-specific ones, consistent with the fastcore algorithm's focus on biologically relevant pathways (**Figure 4B**). Differences in metabolic responses or biomass accumulation between models were highlighted by Euclidean distances, with significant variations in some samples, suggesting distinct metabolic flux or biomass in context-specific modeling. Conversely, minimal differences in others suggest consistent metabolic behavior across conditions. This highlights the metabolic diversity and adaptability of microbial communities in response to environmental contexts.

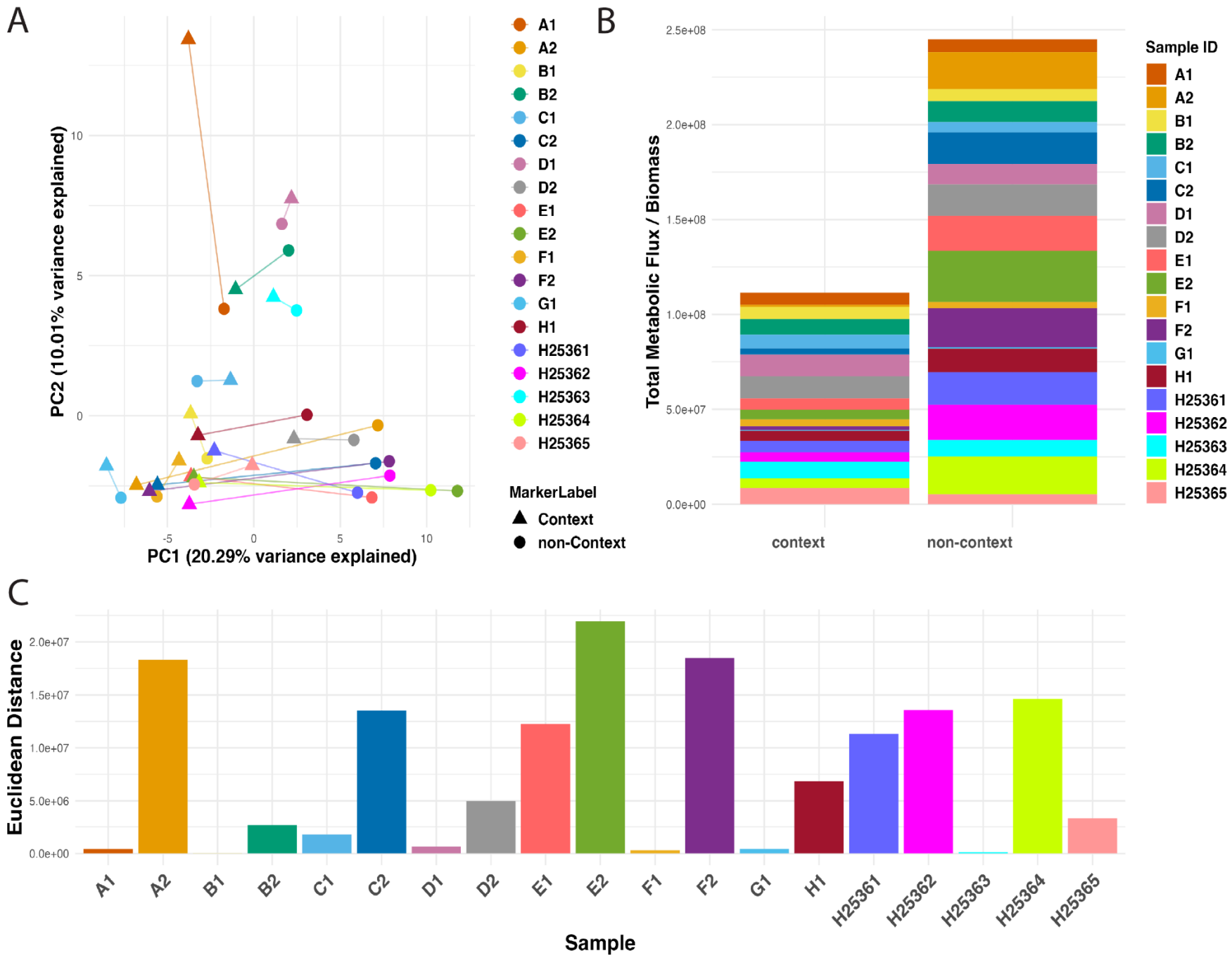

A
PC2 (10.01% variance explained)
PC1 (20.29% variance explained)
10
5
0
-5
0
5
10
A1
A2
B1
B2
C1
C2
D1
D2
E1
E2
F1
F2
G1
H1
H25361
H25362
H25363
H25364
H25365
MarkerLabel
Context
non-Context
B
Total Metabolic Flux / Biomass
2.5e+08
2.0e+08
1.5e+08
1.0e+08
5.0e+07
0.0e+00
context
non-context
Sample ID
A1
A2
B1
B2
C1
C2
D1
D2
E1
E2
F1
F2
G1
H1
H25361
H25362
H25363
H25364
H25365
C
Euclidean Distance
2.0e+07
1.5e+07
1.0e+07
5.0e+06
0.0e+00
A1
A2
B1
B2
C1
C2
D1
D2
E1
E2
F1
F2
G1
H1
H25361
H25362
H25363
H25364
H25365
Sample

Figure 4: Comparative analysis of metabolic flux across non-context-specific and context-specific microbiome models

By describing differences in metabolic flux normalized by community biomass between non-context-specific versus context-specific community modeling, changes in metabolites due to methodology can be observed. (A) A Principal Component Analysis (PCA) plot illustrates metabolic flux distribution across different patient samples and modeling methodologies. (B) Total metabolic flux between context and non-context-specific models. (C) Euclidean distances between non-context-specific and context-specific conditions for each sample, indicating the differential impact of context on metabolic flux or biomass. Differences in height represent the disparity in metabolic activity, with larger distances reflecting significant methodological effects on the microbial community's metabolic outcomes.

Elucidating metabolic profiles across patient samples and microbial community subgroup

To understand metabolic variations in UTIs, we analyzed metabolite concentrations across patient samples using a heatmap, highlighting key metabolites like riboflavin, malic acid, glycolic acid, and D-/L-lactic acid (**Figure 5**). "Homogeneous non-metal compounds" and "Organic acids and derivatives" were notably abundant, signifying their significance in microbial metabolism and the increased production or consumption of organic acids during infections. Statistical analysis showed distinct concentration differences between non-context and context community models for metabolites such as CO2, D-mannitol, D-glucose, hydrogen, and glycerol, whereas L-phenylalanine, L-arabinose, glycine, gluconic acid, and L-glutamic acid showed minimal changes, suggesting similar abundance profiles across both models (**Figure 5A**). Comparing global flux normalized by biomass between *E. coli* and *L. iners* under non-context conditions revealed significant differences (p-value: 0.0437), indicating higher mean flux in E. coli-dominated samples (**Figure 5B**). However, this difference was not significant under context-specific conditions (p-value: 0.0875), hinting that environmental factors might evenly affect microbial metabolic profiles and can potentially lead to improved niche adaptation. We identified context-specific metabolites critical to microbial metabolism and potential therapeutic targets, indicating the importance of integrating metatranscriptomic data for modeling community metabolism (**Figure 5C**). Furthermore, analysis of less significantly altered metabolites like butyric acid, D-maltose, and gamma-aminobutyric acid revealed subtle shifts in the UTI metabolic landscape, emphasizing the need for comprehensive metabolic analysis to grasp the microbial community complexities in UTIs (**Figure 5D**).

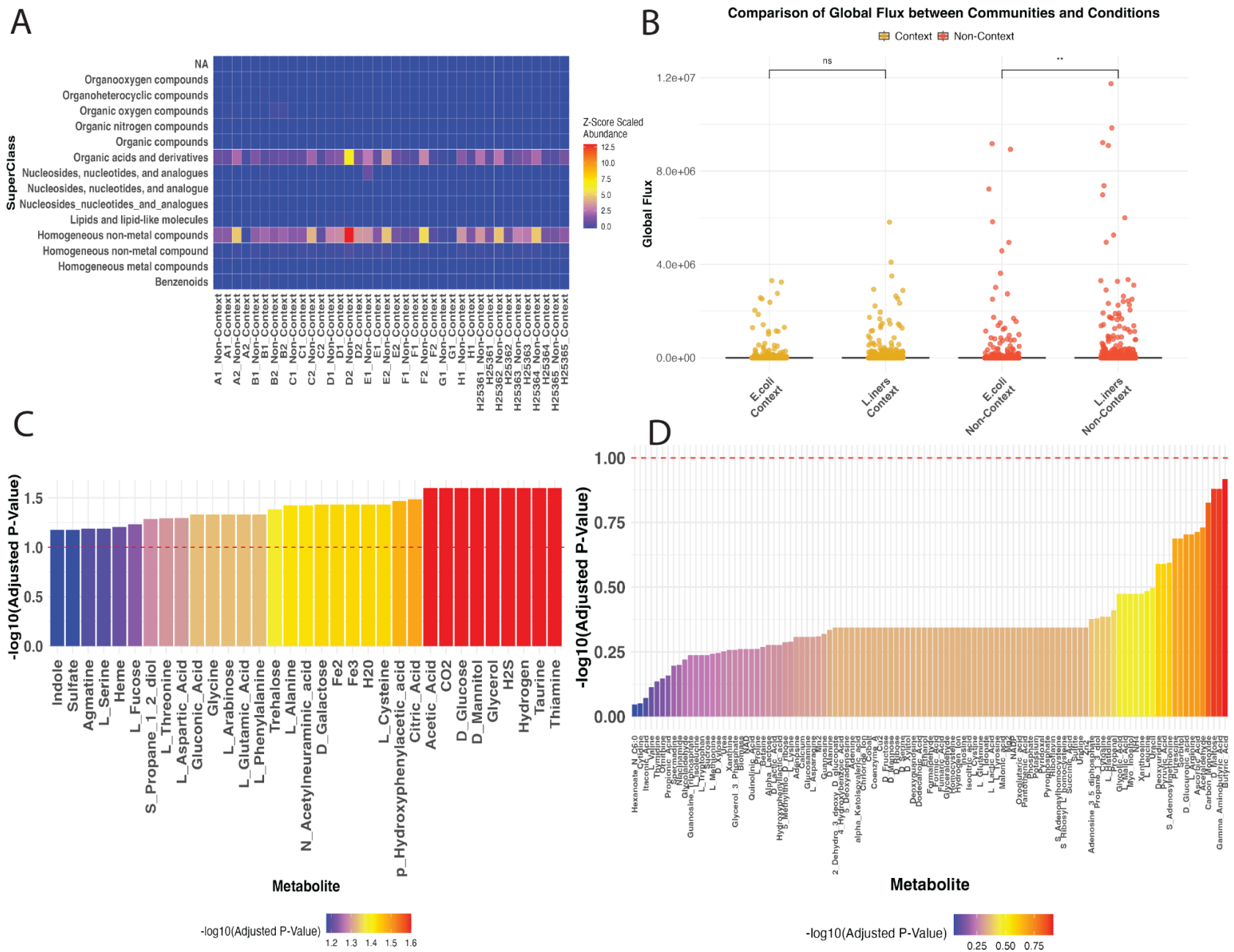

A
SuperClass
NA
Organooxygen compounds
Organoheterocyclic compounds
Organic oxygen compounds
Organic nitrogen compounds
Organic compounds
Organic acids and derivatives
Nucleosides, nucleotides, and analogues
Nucleosides, nucleotides, and analogue
Nucleosides_nucleotides_and_analogues
Lipids and lipid-like molecules
Homogeneous non-metal compounds
Homogeneous non-metal compound
Homogeneous metal compounds
Benzenoids
Z-Score Scaled Abundance
12.5
10.0
7.5
5.0
2.5
0.0
B
Comparison of Global Flux between Communities and Conditions
Context
Non-Context
ns
**
Global Flux
1.2e+07
8.0e+06
4.0e+06
0.0e+00
E.coli Context
L.iners Context
E.coli Non-Context
L.iners Non-Context
C
-log10(Adjusted P-Value)
1.5
1.0
0.5
0.0
Indole
Sulfate
Agmatine
L_Serine
Heme
L_Fucose
S_Propane_1_2_diol
L_Threonine
L_Aspartic_Acid
Gluconic_Acid
Glycine
L_Arabinose
L_Glutamic_Acid
L_Phenylalanine
Trehalose
L_Alanine
N_Acetylneuraminic_acid
D_Galactose
Fe2
Fe3
H20
L_Cysteine
p_Hydroxyphenylacetic_acid
Citric_Acid
Acetic_Acid
CO2
D_Glucose
D_Mannitol
Glycerol
H2S
Hydrogen
Taurine
Thiamine
Metabolite
-log10(Adjusted P-Value)
1.2
1.3
1.4
1.5
1.6
D
-log10(Adjusted P-Value)
1.00
0.75
0.50
0.25
0.00
Metabolite
-log10(Adjusted P-Value)
0.25
0.50
0.75

Figure 5: Characterization of Total Metabolic Fluxes Grouped by Chemical Classes

(A) The heatmap displays the scaled abundance of metabolites derived from each patient's community model, classified by chemical taxonomy derived from the Human Metabolome Database. (B) Barplot representation of total metabolic flux normalized by community biomass, categorized by E. coli or L. iners dominance and respective conditions ('non-Context' or 'Context'). Generally, 'non-Context' conditions exhibited a higher total metabolic flux, with E. coli-dominated communities consistently showing higher flux than those dominated by L. iners. (C) Differential metabolite expression between 'Context' and 'non-Context.' Paired t-test and Wilcoxon paired test were employed, with results adjusted for FDR using the Benjamini-Hochberg method. Metabolites with a FDR<=0.1 were plotted (D) All additional metabolites with a FDR<= below 1.00 are also plotted.

Evaluating metabolic impact between context-specific and non-context-specific microbiome community models

In our study, we examined the metabolic differences between context-specific and generic metabolic models within our patient cohort, focusing on metabolite activity and presence changes (**Figure 6**). **Figure 6A** presents a presence/absence matrix, revealing metabolites with unique patterns in context-specific models, demonstrating differential metabolomic responses of metabolites that are produced by some microbial communities and metabolites that are not prodiuced within other communities. **Figure 6B**, using PCA, identified distinct metabolites like N-acetylneuraminic acid and ascorbic acid that contribute significantly to variability, highlighting the precision of context-specific models in capturing biologically relevant metabolic networks. This approach revealed nuanced shifts in metabolic demands, especially in amino acids and antioxidants, under context-specific conditions. Variations along PC2 suggested changes in energy and sulfur metabolism, emphasizing the metabolic distinctions these models can unveil. **Supplementary Figure S3** shows a comprehensive metabolite presence/absence across patients. **Figure 6C** illustrates the dynamic metabolic responses within individual samples through a heatmap of the top 5 highest and lowest log fold changes (LFC) metabolites, highlighting significant metabolic transformations when employing context-specific modeling. Metabolites like S-ribosyl-L-homocysteine and formaldehyde were consistently upregulated, indicating their importance in key metabolic pathways. In contrast, homocysteine and propionic acid were downregulated, suggesting suppressed pathways in context-specific conditions. These insights underscore the critical roles of certain metabolites in microbial growth and metabolism, demonstrating the value of integrating transcriptomic data into metabolic modeling.

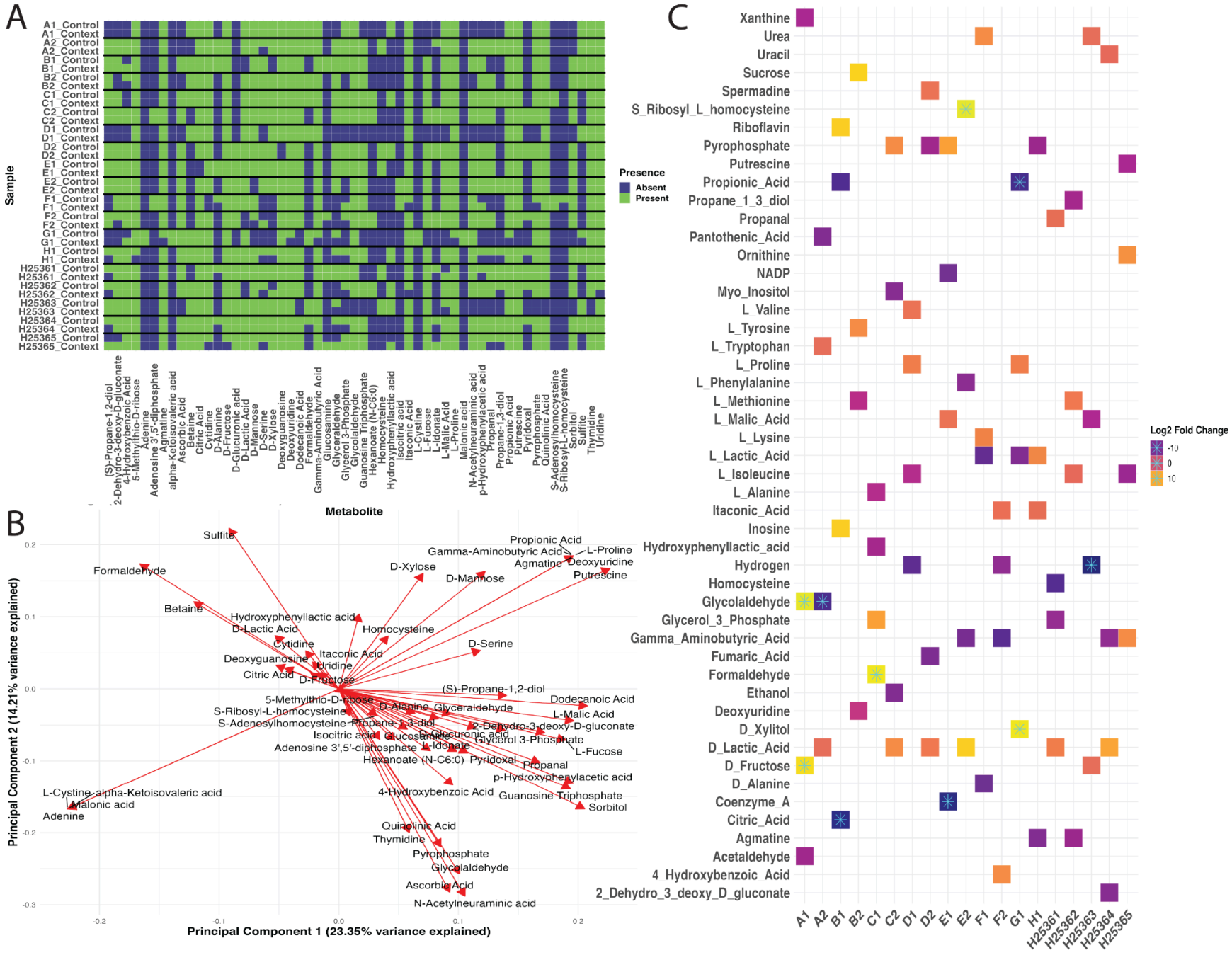

A
Sample
Metabolite
Presence
Absent
Present
B
Principal Component 2 (14.21% variance explained)
Principal Component 1 (23.35% variance explained)
C
Log2 Fold Change
-10
0
10

Figure 6: Comparative assessment of metabolic impact derived from context-specific and non-context-specific microbiome models

(A) The presence/absence matrix showcases the active metabolites across the patient cohort, highlighting gains or losses when applying our modeling methodology. (B) PCA biplot highlights the top 5 metabolites contributing to the variability. This analysis elucidates the metabolites most influential in distinguishing between context-specific and non-context-specific metabolic model changes. (C) Log-fold changes in metabolites, when contrasting the two modeling approaches, highlights the top 5 highest and lowest changes (highlighted by a cyan star) in metabolic flux for a patient when comparing results from context-specific and non-context-specific methodologies.

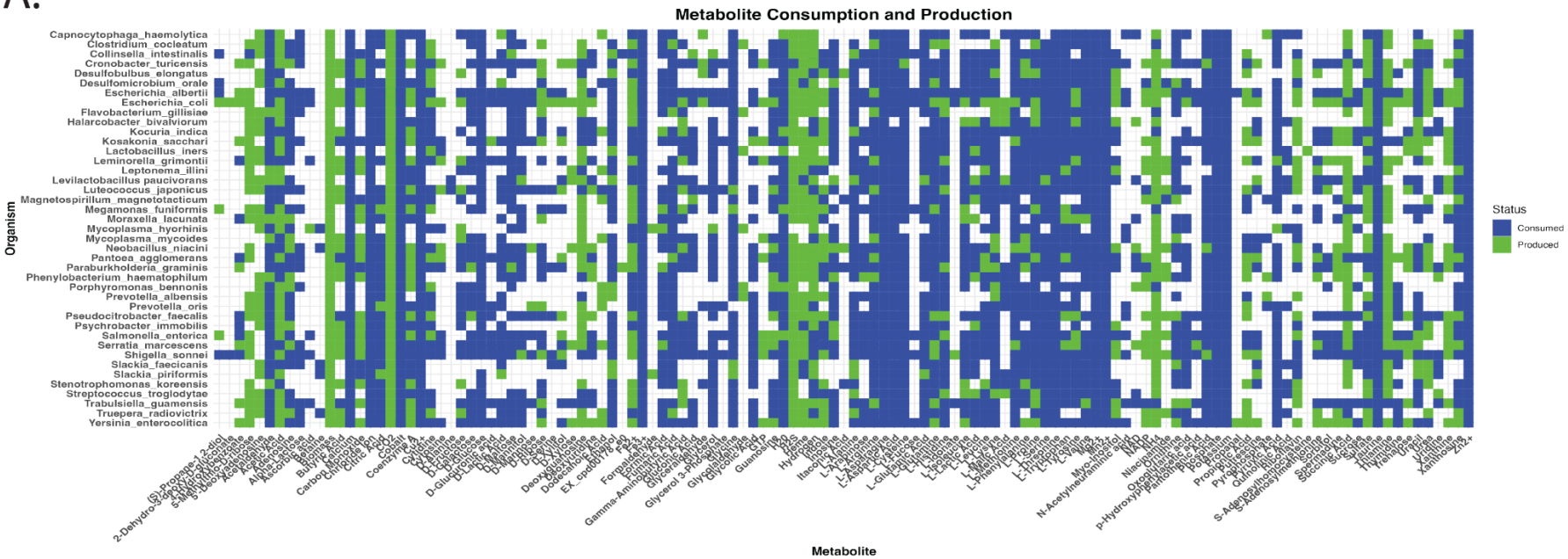
Uncovering distinguishing metabolic signatures and evaluating predictive performance via machine learning

Using a random forest algorithm complemented our statistical analyses by identifying significant metabolic features distinguishing context-specific from non-context-specific models. This approach, along with Recursive Feature Elimination (RFE) for top-k feature selection, pinpointed key metabolites defining context-specific microbial community metabolism (**Supplementary Figure S4**). From an accuracy standpoint, metabolites such as glycerol, hydrogen sulfide, thiamine, and riboflavin stood out as differentiators between modeling approaches (**Supplementary Figure S4A**). Gini Impurity further identified crucial metabolites like hydrogen sulfide, L-threonine, and glycerol (**Supplementary Figure S4B**). RFE refined this list to hydrogen sulfide, glycerol, L-threonine, and thiamine, among others, underscoring these as central to distinguishing context-specific metabolic landscapes in UTIs. Evaluating the predictive performance of our model (**Supplementary Figure S5**), the ROC curve demonstrated a moderate Area Under the Curve (AUC) of 0.66667, with precision, recall, and F1 score indicating the model's capability to classify the methodologies accurately. RFE, validated through tenfold cross-validation, compared feature subsets for classification performance, showing a high accuracy and kappa score for the 11-variable set, significantly outperforming the larger 141-variable set. However, this may be indicative of overfitting which would not allow for proper generatlization, therefore these results require further validation before application in biological contexts. Nonetheless, these findings reveal a core set of metabolites, like hydrogen sulfide and glycerol, crucial for context-specific modeling, suggesting differences in modeling approaches rather than direct implications for UTIs. The identified metabolites provide a foundation for future research to explore their roles in UTI pathogenesis and model differentiation.

Interplay of metabolite cross-feeding in UTI microbial communities reveals potential pathogenic strategies

Understanding microbial metabolic interactions is essential for deciphering the complexity of microbial ecosystems. Our comprehensive visualization of the top 150 most variable metabolites illustrates the production and consumption patterns across all simulated species, providing insights into the metabolic interplay within the microbial community (**Figure 7A/B**). **Supplementary Table S5** details the averaged metabolic flux rates, highlighting key organisms and metabolites central to the ecosystem's dynamics. Further analysis compared metabolic behaviors of gut-associated pathogens to Lactobacillus-dominated communities, identifying*g C. turicensis, B. graminis*, and *K. indica* as major consumers of metabolites like citric acid, glycine, and H2O. Conversely, *Y. enterocolitica, E. coli,* and *P. agglomerans*, known human pathogens, emerged as significant producers, with CO2, acetic acid, and hydrogen among their key products **(Figure 7C/D**). Distinct metabolic patterns were observed between gut-derived pathogens *(S. sonnei, S. enterica, & E.coli*) and *Lactobacillus* species (*L. iners & L. paucivorans*). *E. coli* showed high consumption of CO2, acetic acid, and hydrogen, contrasting with *L. iners,* which primarily consumed L-lactic acid, hydrogen, and CO2, and produced D-glucose and gluconic acid. This suggests overlapping metabolic pathways, particularly in gluconic acid production by *S. sonnei and L. iners*. Hydrogen and CO2 consumption patterns indicate a core metabolic activity shared across both groups. However, metabolite production varied among UTI pathogens, with *E. coli* focusing on citric acid and glycine, and *S. sonnei* on L-glutamine and gluconic acid, highlighting the metabolic diversity and adaptability of microbial communities.

A.


Metabolite Consumption and Production
Organism
Capnocytophaga_haemolytica
Clostridium_cocleatum
Collinsella_intestinalis
Cronobacter_turicensis
Desulfobulbus_elongatus
Desulfomicrobium_orale
Escherichia_albertii
Escherichia_coli
Flavobacterium_gillisiae
Halarcobacter_bivalviorum
Kocuria_indica
Kosakonia_sacchari
Lactobacillus_iners
Leminorella_grimontii
Leptonema_illini
Levilactobacillus paucivorans
Luteococcus_japonicus
Magnetospirillum_magnetotacticum
Megamonas_funiformis
Moraxella_lacunata
Mycoplasma_hyorhinis
Mycoplasma_mycoides
Neobacillus_niacini
Pantoea_agglomerans
Paraburkholderia_graminis
Phenylobacterium_haematophilum
Porphyromonas_bennonis
Prevotella_albensis
Prevotella_oris
Pseudocitrobacter_faecalis
Psychrobacter_immobilis
Salmonella_enterica
Serratia_marcescens
Shigella_sonnei
Slackia_faecicanis
Slackia_piriformis
Stenotrophomonas_koreensis
Streptococcus_troglodytae
Trabulsiella_guamensis
Truepera_radiovictrix
Yersinia_enterocolitica
Metabolite
Status
Consumed
Produced

B.

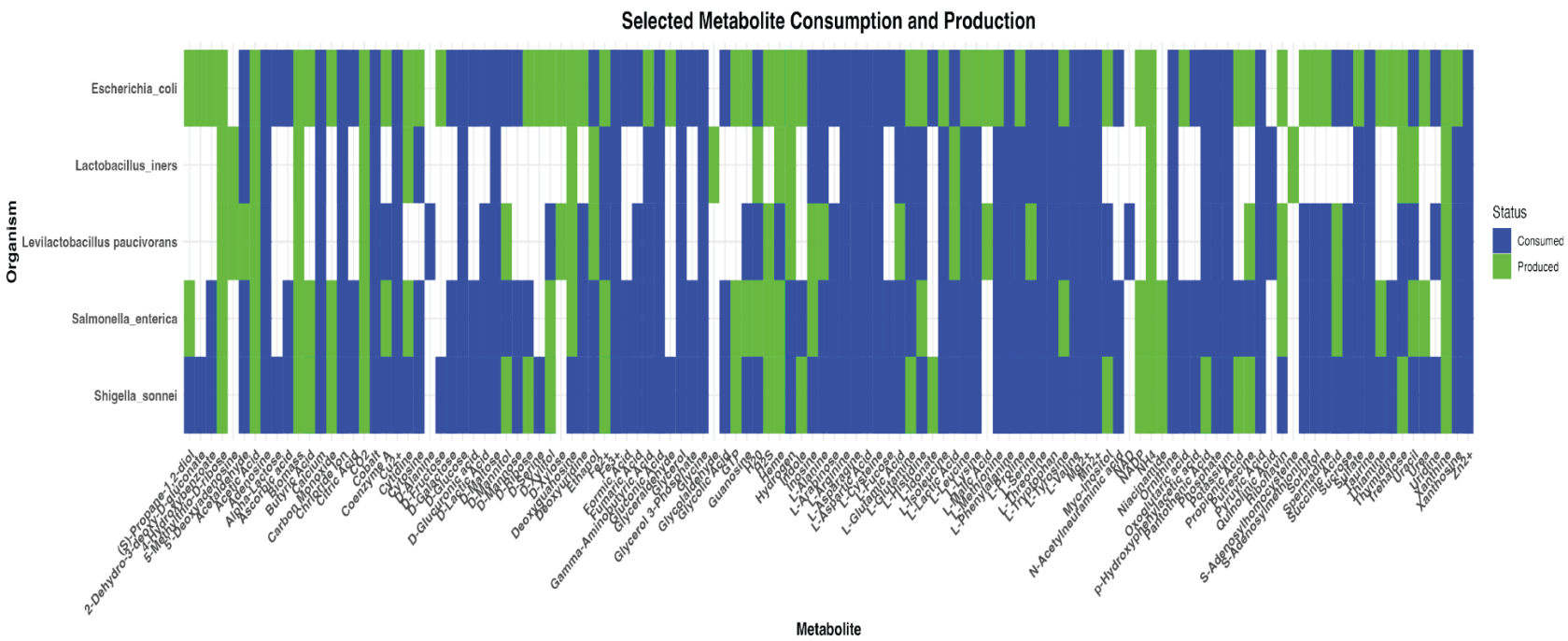

Selected Metabolite Consumption and Production
Organism
Escherichia_coli
Lactobacillus_iners
Levilactobacillus paucivorans
Salmonella_enterica
Shigella_sonnei
Metabolite
Status
Consumed
Produced

Figure 7: Metabolic interactions of microbes based on average flux values

(A) Heatmap depicting the production and consumption of the top 150 most variable compounds across all simulated microbes; rows represent microbes, while columns represent metabolites. Green represents production, and blue represents neither the observed production of metabolites. (B) A focused heatmap showcasing the metabolic interactions of Lactobacillus spp. compared to select pathogenic microbes, including E. coli. Metabolite interactions are color-coded similarly to subplot A.

Our analysis of context-specific metabolic cross-feeding interactions within microbial communities revealed significant variability in metabolite distribution and frequency, highlighting the diverse metabolic capacities of context-specific microbial ecosystems (**Figure 8**). Hydrogen, notably prevalent in sample C1 with 506 instances of microbial interaction, underscores its role in syntrophic relationships and electron transfer within the uromicrobiome. Acetic acid's prominence indicates its contribution to fermentation or as a substrate for acetogenic bacteria, while the abundance of Fe^{2+} emphasizes iron's critical role in microbial growth and electron transfer processes. Amino acid metabolism diversity was evident, with L-arginine and L-glutamine exchanges pointing to varied protein synthesis, nitrogen assimilation, or degradation pathways within uropathogenic communities. The substantial variability across contexts suggests distinct metabolic states or microbial compositions, reflecting the dynamic nature of microbial ecosystems in response to environmental changes. Particularly in context G1, dominated by *L. iners*, a unique metabolic signature emerged, despite a lower frequency of exchanged metabolites. The presence and ratios of acetic acid, L-glutamine, and Fe2+ in G1 provide insights into *L. iners*' metabolic activities and ecological roles. The balanced profile of amino acids, energy substrates, and micronutrients in G1 may highlight *L. iners'* unique metabolic capabilities and its contributions to the ecosystem's resilience and functionality. Overall, our metabolic profiling underscores the complexity of microbial metabolisms and the pivotal roles of specific metabolites and species, like *L. iners* in G1, in defining ecological interactions and functionalities within diverse microbial communities. **Supplementary Table S6** offers an exhaustive overview of the bidirectional metabolic fluxes between strains in each patient-specific sample, further illustrating the intricate metabolic interplays at play.

Our study explores metabolic cross-feeding within UTI microbiome communities, highlighting the complexity of interactions between microbial species in two patient samples, H1 and H25364, as detailed in **Figure 8B/C**. These case studies, while unique, reflect broader patterns observable across various patients. In patient, H1, *E. coli* emerges as a critical player in the microbial network (**Figure 8A)**. It acquires Fe^{2+} from the community, essential for its growth and possibly serving in electron transfer processes during pathogenesis. Concurrently, *E. coli* supplies ornithine, crucial for bacterial proliferation, demonstrating a mutualistic interaction that could influence UTI progression. Additionally, *E. coli's* role extends to distributing sorbitol,

indicating its multifaceted influence on the UTI microbiome's metabolic dynamics. Conversely, patient H25364 showcases distinct interactions, with a notable three-way exchange among *E. coli, S. sonnei, and S. enterica*, focusing on the transfer of ornithine and iron. This interaction underscores the significance of these pathogens in the community, potentially identifying them as targets for therapeutic interventions. Furthermore, *E. coli's* provision of mannitol and *Y. enterocolitica's* contribution of sugar alcohols like xylitol and sorbitol reveal a complex network of metabolite exchange. These insights into metabolic cross-feeding illuminate the intricate microbial interactions within UTI communities, offering potential pathways for targeted treatments and highlighting the need for further investigation into these complex ecosystems.

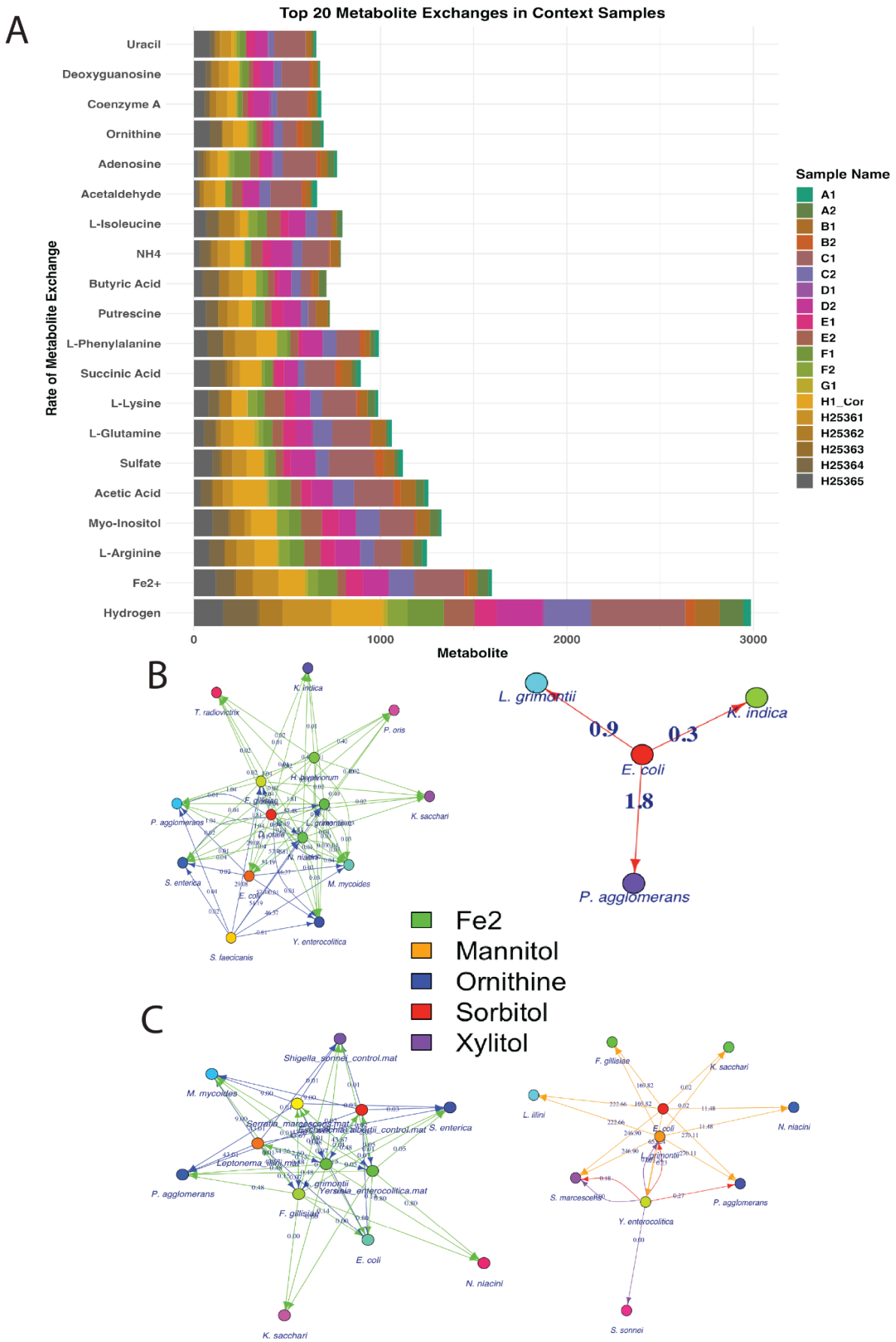

A
Top 20 Metabolite Exchanges in Context Samples
Uracil
Deoxyguanosine
Coenzyme A
Ornithine
Adenosine
Acetaldehyde
L-Isoleucine
NH4
Butyric Acid
Putrescine
L-Phenylalanine
Succinic Acid
L-Lysine
L-Glutamine
Sulfate
Acetic Acid
Myo-Inositol
L-Arginine
Fe2+
Hydrogen
Rate of Metabolite Exchange
0
1000
2000
3000
Metabolite
Sample Name
A1
A2
B1
B2
C1
C2
D1
D2
E1
E2
F1
F2
G1
H1_Cor
H25361
H25362
H25363
H25364
H25365
B
K. indica
T. radiovictrix
P. oris
H. hyalriorum
F. gillisiae
K. sacchari
P. agglomerans
L. grimontii
D. otala
N. niacini
S. enterica
M. mycoides
E. coli
Y. enterocolitica
S. faecicanis
L. grimontii
K. indica
0.9
0.3
E. coli
1.8
P. agglomerans
Fe2
Mannitol
Ornithine
Sorbitol
Xylitol
C
Shigella_sonnei_control.mat
M. mycoides
Serratia_marcescens.mat
Escherichia_albertii_control.mat
S. enterica
Leptonema_illini.mat
L. grimontii
Yersinia_enterocolitica.mat
P. agglomerans
F. gillisiae
E. coli
N. niacini
K. sacchari
F. gillisiae
K. sacchari
L. illini
N. niacini
E. coli
L. grimontii
S. marcescens
P. agglomerans
Y. enterocolitica
S. sonnei

Figure 8: Metabolic flux exchange and cross-feeding interactions in UTI microbiome communities

(A) This figure illustrates the top 20 metabolic interactions among context-specific microbial communities. This visual summary captures the dynamic metabolic exchanges within microbial ecosystems, showcasing key metabolites and the impact of microbial diversity on metabolic pathways. Examples of predicted cross-feeding interactions within UTI microbiome communities from two context-specific patient samples: H1 and H25364. Both panels are split into two parts to illustrate different cross-feeding dynamics, with the left side (B) illustrates the metabolic interactions within the entire microbial community, while (C) zooms in on E. coli, highlighting examples of metabolic cross-talk that suggest potential pathogenic strategies within the urinary tract infection (UTI) ecosystem. In these visual representations, nodes symbolize the different species within the community, and edges depict the metabolites exchanged between them, directed from the metabolite-secreting species to the consuming ones. This directional flow of metabolites, dictated by the metabolic flux across the population, underscores the complex network of metabolite cross-feeding. By mapping out these interactions, the panels not only shed light on the dynamics of UTIs but also propose metabolic avenues to curb the influence of primary community pathogens, offering insights for targeted therapeutic strategies.

IV. Discussion

The outdated belief that the urinary tract is sterile has been recently refuted by previous authors, acknowledging that a unique microbiome exists within the urinary tract, even in the absence of infections (Ackerman & Toby, 2019; Pohl et al. 2020). Our study leverages metatranscriptomics to delve into the urinary microbiome of women experiencing UTIs, revealing a complex network of microbial compositions, functions, and interactions. Contrary to the notion of UTIs being caused by single pathogens, the findings presented in this manuscript and observed by others suggest a communal and synergistic bacterial interplay (de Vos et al. 2017; Zandbergen et al. 2021). Through the integration of metatranscriptomics, Genome-Scale Metabolic Models (GEMs), and insights from the Human Urine Metabolome database, we crafted personalized microbiome models for patients (Bouatra et al. 2013). This approach enhances our comprehension of the metabolic dynamics within the UTI microbiome, highlighting the microbial diversity and complexity in the urinary tracts of patients with UTIs. Recognizing these intricate microbial interactions is pivotal for devising more effective, patient-specific treatments.

The diversity in the uromicrobiome among UTI patients highlights the need for personalized treatment approaches, challenging the one-size-fits-all nature of standard therapies. This diversity, encompassing a range of uropathogenic and probiotic genera, suggests that UTIs may emerge from complex microbial interactions rather than from a single pathogen.

Such variability can influence immune responses, infection susceptibility, and treatment outcomes, with factors like genetics, diet, age, and antibiotic use playing significant roles (Zheng et al. 2020; Schembri et al. 2022). The *Lactobacillus* genus, in particular, may provide protective benefits by maintaining urinary tract health and preventing pathogen colonization (Singhal et al. 2020). The Shannon index in combination with ENS, was useful to indicate and model patient-specific diversity within patient uromicrobiomes, an approach that can be readily applicable to other microbiomes across the body (Huttenhower et al. 2012; Saheb et al. 2022). However, traditional diversity metrics may overlook the impact of low-abundance species within the uromicrobiome on community dynamics and disease progression. These minor constituents, crucial for understanding urinary health and disease, necessitate more nuanced analytical methods to fully appreciate their roles (Feehily et al. 2020; Pust et al. 2022). Our findings also point to the presence of gut-associated microbes like *E. coli and S. sonnei* in the urinary tract, underscoring a link between gastrointestinal and urogenital microbiomes in UTI development; this observation supports the hypothesis of microbial migration between these sites, though it necessitates cautious interpretation without comparative fecal sample analysis (Paalanne et al. 2018; Magruder et al. 2019; Schembri et al. 2022; Worby et al. 2022). Furthermore, we encountered discrepancies in sequencing data from Illumina MiSeq and NovaSeq platforms, due to varying sequencing depths and biases, which highlights a challenge in cross-platform microbial analysis (Singer et al. 2019).

UPEC strains display diverse virulence tactics, with variable expression of virulence genes among UTI patients, highlighting their adaptability and posing challenges for infection management due to different pathogenic mechanisms (Hagan et al. 2010; Frick-Cheng et al. 2020). Specific adherence genes, *fimA* and *fimI*, show pronounced expression in certain samples, suggesting these strains have enhanced capabilities to attach to host cells, contributing to the persistence of infections. This supports findings on the critical role of adherence in bacterial colonization and infection establishment (Schuroff et al. 2021). The *ompA* gene, essential for bacterial invasion of host cells, appears nearly universally, indicating its key role in UPEC's ability to evade the immune response and potentially cause recurrent infections (Nielsen et al. 2020). Genes like *chuY* and *chuS*, involved in heme utilization, and iroN and fepA, related to iron acquisition, underline UPEC's strategies to secure essential nutrients in the iron-scarce urinary tract, crucial for their virulence and survival (Mey et al. 2021; Edison et al. 2023). Additionally, the presence of effector genes like *map* suggests mechanisms by which UPEC modulates host cell functions to facilitate invasion. This multifaceted virulence approach, combining adherence, invasion, and nutritional strategies, may contribute to the severity and treatment resistance of infections (Ambite et al. 2021). However, the influence of host factors on gene expression and the functional activity of expressed virulence factors remain important considerations in understanding UPEC pathogenicity.

Our research adeptly leverages transcriptomic data to elucidate microbial physiology, bridging gene expression with metabolic activities for an in-depth analysis of microbial dynamics across health and disease spectra (Angione et al., 2019). Through the comparative analysis of metabolic flux patterns in both context-specific and non-context-specific microbiome models, our study underscores the criticality of incorporating biological context into metabolic simulations. The observed consistency in metabolic responses among patient samples in context-specific modeling likely mirrors the actual metabolic states of microbial communities within UTI environments, advocating for the customization of models to specific biological scenarios (Richelle et al., 2019b; Joshi et al., 2020; Viera et al., 2022). This approach contrasts with the variability noted in non-context-specific models, which may stem from the inclusion of irrelevant metabolic reactions, suggesting a potential pitfall of generic modeling approaches. Employing Random Forest machine learning, our study discerns metabolic distinctions between these modeling approaches, pinpointing metabolites such as H2S, glycerol, L-threonine, thiamine, inosine, and riboflavin as pivotal in context-specific models. This aligns with previous research underscoring the significance of these metabolites in UTI dynamics (Gupta et al., 2006; Debarbouille et al., 2009; Campagna et al., 2020; Zhao et al., 2021), thus aiding in the prediction of metabolite behavior across various contexts. However, the translation of gene expression data into actual metabolic flux or protein activity remains a challenge, highlighting a limitation in the current modeling approach. Despite these limitations, the methodological framework presented offers profound insights into microbial functionality and interactions, significantly enriching our understanding of the metabolic intricacies within patient-specific microbial ecosystems. This study lays a foundation for future research, which could be augmented by integrating patient-specific metabolomics data for enhanced validation and accuracy, ultimately fostering a deeper comprehension of microbial community dynamics in UTIs.

Constraint-based metabolic modeling offers a dynamic view of a biological system's physiological state, capturing specific metabolite clusters such as organoxygen compounds—encompassing carbohydrate conjugates, alcohols, and amino acids—essential for various cellular functions, from energy production to signaling (Damiani et al. 2020; Robinson et al. 2019). The abundance of "Homogeneous non-metal compounds" and "Organic acids and derivatives" during UTIs underscores their potential importance in disease processes. Shifts in metabolites like D-mannitol, D-glucose, and thiamine, serving as primary metabolic intermediates, provide insights into metabolic transitions between health and disease states in UTIs (Lupo et al. 2021; Paudel et al. 2023). A comparative analysis revealed that *E. coli* exhibits a higher mean flux than *L. iners*, suggesting a more metabolically active or aggressive stance by *E. coli* in UTI conditions, highlighting microbial competition and dynamics in the urinary tract. However, the study acknowledges limitations, emphasizing that metabolite abundance does not directly correlate with functional significance without considering the broader biological context and the interplay of factors (Joshi et al. 2020). Furthermore, variations in mean flux across

conditions may reflect inherent biases of the context-specific modeling method, which, while informative, may also influence the interpretation of data (Zampieri et al. 2023).

Our research explores metabolic interactions within UTI microbial communities, uncovering complex cross-feeding mechanisms that may underpin both protective and pathogenic behaviors. We identified a syntrophic relationship among microbes such as *C. turicensis, B. graminis, and K. indica*, which utilize metabolites like citric acid and glycine, in contrast to *Y. enterocolitica, E. coli, and P. agglomerans*, producers of CO2 and acetic acid. Particularly intriguing is the consistent presence of Lactobacillus species in UTI patients, with *L. iners* noted for metabolizing substances into D-glucose and gluconic acid. This activity hints at *Lactobacillus's* capacity to foster a less pathogenic urinary ecosystem, possibly by acidifying the environment or generating metabolites detrimental to pathogen proliferation. The study also emphasizes the significance of hydrogen, acetic acid, and Fe^{2+} as central metabolites, highlighting their roles in microbial survival and interaction. For instance, *E. coli's* dual role in consuming iron and producing metabolites associated with UTI pathogenesis reflects its intricate involvement in shaping the metabolic framework of the community, potentially aiding pathogenic strategies. This detailed examination of metabolite cross-feeding enriches our understanding of microbial community dynamics within UTIs, suggesting avenues for metabolic-based interventions in UTI management. Despite these insights, the complexity of these interactions calls for further investigation to fully elucidate the metabolic network driving UTI pathogenicity, including the protective potential of lactobacilli in creating a less hospitable environment for pathogens.

V. Conclusion

Our data-driven modeling of the UTI microbiome through metatranscriptomics has provided valuable metabolic insights into metabolic pathways and cross-feeding interactions relevant to UTIs. This paves the way for the design of specific intervention strategies that interrupt those interactions e.g. via dietary supplements to suppress infection (Yang et al. 2020). However, it's crucial to recognize the limitations of our current model. The unique phenotypes of different strains can be influenced by individual factors and micro-environmental conditions—such as diet, pre- or coexisting health conditions, local microbiota, pH, and nutrient presence—which our current framework does not entirely encapsulate. As we delve deeper into defining the functional uromicrobiome, future models will factor in patient-specific dietary and metabolic information, refining future simulations and revolutionizing how clinicians perceive and manage UTIs. Our study highlights the remarkable capability of metatranscriptomics in revealing the dynamic interplay of the uromicrobiome and suggests a roadmap for future research. As we decipher these intricate relationships between the host and its microbiome in UTIs, the vision of tailoring therapeutic strategies based on individual microbiome compositions emerges, heralding a new era of enhanced care for UTI patients.

Acknowledgments

The authors are thankful to the Medical Systems Biology team for useful discussions and support in this project. The authors thank Rebecca Höög for creating the graphical abstract, which was created using Biorender. The authors declare that they have no conflict of interest.

Funding

This research was funded through the DFG Excellence Cluster Precision Medicine in Chronic Inflammation (code EXC2167). Additionally funding for metatranscriptomics sequencing was awarded to J.J-S through a Kiel University ZMB Young Scientist Grant.

DATA/CODE AVAILABILITY

All in-house generated sequences were deposited in the NCBI-Sequence Read Archive database (PRJNA1091563). The accession number of the deposited reads is provided in **Supplementary Table S1**. Code/data/software required to replicate the figures and findings in this manuscript are all available on GitHub (https://github.com/PoeticPremium6/UTI-Microbiome-Model).

References

Adebayo, A. S., G. Ackermann, R. C. E. Bowyer, P. M. Wells, G. Humphreys, R. Knight, T. D. Spector, and C. J. Steves. "The urinary tract microbiome in older women exhibits host genetic and environmental influences." *Cell Host & Microbe* 28, no. 2 (2020): 298-305.

Ackerman, A. Lenore, and Toby C. Chai. "The bladder is not sterile: an update on the urinary microbiome." *Current bladder dysfunction reports* 14 (2019): 331-341.

Ambite, Ines, Daniel Butler, Murphy Lam Yim Wan, Therese Rosenblad, Thi Hien Tran, Sing Ming Chao, and Catharina Svanborg. "Molecular determinants of disease severity in urinary tract infection." *Nature Reviews Urology* 18, no. 8 (2021): 468-486.

Ammitzbøll, Nadia, Benedikt Paul Josef Bau, Caspar Bundgaard-Nielsen, Annemarie Brusen Villadsen, Ann-Maria Jensen, Peter Derek Christian Leutscher, Karin Glavind, Søren Hagstrøm, Louise Thomsen Schmidt Arenholt, and Suzette Sørensen. "Pre-and postmenopausal women have different core urinary microbiota." *Scientific reports* 11, no. 1 (2021): 2212.

Anders, Simon, Paul Theodor Pyl, and Wolfgang Huber. "HTSeq—a Python framework to work with high-throughput sequencing data." *bioinformatics* 31, no. 2 (2015): 166-169.

Angione, Claudio. "Human systems biology and metabolic modelling: a review—from disease metabolism to precision medicine." *BioMed research international* 2019 (2019).

Baker, Joseph L., Tobias Dahlberg, Esther Bullitt, and Magnus Andersson. "Impact of an alpha helix and a cysteine–cysteine disulfide bond on the resistance of bacterial adhesion pili to stress." *Proceedings of the National Academy of Sciences* 118, no. 21 (2021): e2023595118.

Bauer, Eugen, Johannes Zimmermann, Federico Baldini, Ines Thiele, and Christoph Kaleta. "BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities." *PLoS computational biology* 13, no. 5 (2017): e1005544.

Bitew, Adane, Tamirat Molalign, and Meseret Chanie. "Species distribution and antibiotic susceptibility profile of bacterial uropathogens among patients complaining urinary tract infections." *BMC infectious diseases* 17, no. 1 (2017): 1-8.

Bouatra, Souhaila, Farid Aziat, Rupasri Mandal, An Chi Guo, Michael R. Wilson, Craig Knox, Trent C. Bjorndahl et al. "The human urine metabolome." *PloS one* 8, no. 9 (2013): e73076.

Campagna, Giovanni, Andre Espaillat, Thomas Pfeiffer, Jacquelyn M. Powers, Meghan McClure, and Lauren M. Hess. "A Case of Sulfhemoglobinemia Secondary to a Urinary Tract Infection." *Journal of Pediatric Hematology/Oncology* 42, no. 8 (2020): e765-e767.

Cantalapiedra, Carlos P., Ana Hernández-Plaza, Ivica Letunic, Peer Bork, and Jaime Huerta-Cepas. "eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale." *Molecular biology and evolution* 38, no. 12 (2021): 5825-5829.

Damiani, Chiara, Daniela Gaglio, Elena Sacco, Lilia Alberghina, and Marco Vanoni. "Systems metabolomics: From metabolomic snapshots to design principles." *Current opinion in biotechnology* 63 (2020): 190-199.

Danecek, Petr, James K. Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O. Pollard, Andrew Whitwham et al. "Twelve years of SAMtools and BCFtools." *Gigascience* 10, no. 2 (2021): giab008.

Debarbouille, Michel, Shaynoor Dramsi, Olivier Dussurget, Marie-Anne Nahori, Elisabeth Vaganay, Grégory Jouvion, Alain Cozzone, Tarek Msadek, and Bertrand Duclos. "Characterization of a serine/threonine kinase involved in virulence of Staphylococcus aureus." *Journal of bacteriology* 191, no. 13 (2009): 4070-4081.

Deng, Zhi-Luo, Cornelia Gottschick, Sabin Bhuju, Clarissa Masur, Christoph Abels, and Irene Wagner-Döbler. "Metatranscriptome analysis of the vaginal microbiota reveals potential mechanisms for protection against metronidazole in bacterial vaginosis." *Msphere* 3, no. 3 (2018): 10-1128.

de Vos, Marjon GJ, Marcin Zagorski, Alan McNally, and Tobias Bollenbach. "Interaction networks, ecological stability, and collective antibiotic tolerance in polymicrobial infections." *Proceedings of the National Academy of Sciences* 114, no. 40 (2017): 10666-10671.

Durand, Jérôme MB, and Glenn R. Bjork. "Metabolic control through ornithine and uracil of epithelial cell invasion by Shigella flexneri." *Microbiology* 155, no. 8 (2009): 2498-2508.

Ebrahim, A., J. A. Lerman, B. O. Palsson, and D. R. Hyduke. "COBRApy: COnstraints-Based Reconstruction and Analysis for Python. BMC Syst Biol." *COBRApy: COnstraints-Based Reconstruction and Analysis for Python. BMC Syst Biol* 7 (2013).

Edgar, Robert. *Usearch*. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2010.

Edison, Lekshmi K., Indira T. Kudva, and Subhashinie Kariyawasam. "Comparative Transcriptome Analysis of Shiga Toxin-Producing Escherichia coli O157: H7 on Bovine Rectoanal Junction Cells and Human Colonic Epithelial Cells during Initial Adherence." *Microorganisms* 11, no. 10 (2023): 2562.

Feehily, Conor, David Crosby, Calum J. Walsh, Elaine M. Lawton, Shane Higgins, Fionnuala M. McAuliffe, and Paul D. Cotter. "Shotgun sequencing of the vaginal microbiome reveals both a species and functional potential signature of preterm birth." *NPJ biofilms and microbiomes* 6, no. 1 (2020): 50.

Flores-Mireles, Ana L., Jennifer N. Walker, Michael Caparon, and Scott J. Hultgren. "Urinary tract infections: epidemiology, mechanisms of infection and treatment options." *Nature reviews microbiology* 13, no. 5 (2015): 269-284.

Frick-Cheng, Arwen E., Anna Sintsova, Sara N. Smith, Michael Krauthammer, Kathryn A. Eaton, and Harry LT Mobley. "The gene expression profile of uropathogenic Escherichia coli in women with uncomplicated urinary tract infections is recapitulated in the mouse model." *MBio* 11, no. 4 (2020): 10-1128.

Garfoot, Andrew L., Gary M. Wilson, Joshua J. Coon, and Laura J. Knoll. "Proteomic and transcriptomic analyses of early and late-chronic Toxoplasma gondii infection shows novel and stage specific transcripts." *BMC genomics* 20 (2019): 1-11.

Gopalakrishnan, Saratram, Chintan J. Joshi, Miguel Á. Valderrama-Gómez, Elcin Icten, Pablo Rolandi, William Johnson, Cleo Kontoravdi, and Nathan E. Lewis. "Guidelines for extracting biologically relevant context-specific metabolic models using gene expression data." *Metabolic Engineering* 75 (2023): 181-191.

Gorelik, Yuri, Mical Paul, Yuval Geffen, and Mogher Khamaisi. "Urinary tract infections due to nontyphoidal Salmonella." *The American Journal of the Medical Sciences* 353, no. 6 (2017): 529-532.

Gupta, Ashish, Mayank Dwivedi, GA Nagana Gowda, A. A. Mahdi, Amita Jain, Archana Ayyagari, Raja Roy, M. Bhandari, and C. L. Khetrapal. "1H NMR spectroscopy in the diagnosis of Klebsiella pneumoniae-induced urinary tract infection." *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In vivo* 19, no. 8 (2006): 1055-1061.

Hagan, Erin C., Amanda L. Lloyd, David A. Rasko, Gary J. Faerber, and Harry LT Mobley. "Escherichia coli global gene expression in urine from women with urinary tract infection." *PLoS pathogens* 6, no. 11 (2010): e1001187.

Heirendt, Laurent, Sylvain Arreckx, Thomas Pfau, Sebastián N. Mendoza, Anne Richelle, Almut Heinken, Hulda S. Haraldsdóttir et al. "Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v. 3.0." *Nature protocols* 14, no. 3 (2019): 639-702.

Heinken, Almut, Johannes Hertel, Geeta Acharya, Dmitry A. Ravcheev, Malgorzata Nyga, Onyedika Emmanuel Okpala, Marcus Hogan et al. "Genome-scale metabolic reconstruction of 7,302 human microorganisms for personalized medicine." *Nature Biotechnology* (2023): 1-12.

Ho, Sung Yang, Limsoon Wong, and Wilson Wen Bin Goh. "Avoid oversimplifications in machine learning: going beyond the class-prediction accuracy." *Patterns* 1, no. 2 (2020).

Hooton, Thomas M. "Uncomplicated urinary tract infection." *New England Journal of Medicine* 366, no. 11 (2012): 1028-1037.

Huttenhower, Curtis, Dirk Gevers, Rob Knight, Sahar Abubucker, Jonathan H. Badger, Asif T. Chinwalla, Heather H. Creasy et al. "Structure, function and diversity of the healthy human microbiome." *Nature* 486, no. 7402 (2012): 207-214.

Joshi, Chintan J., Song-Min Schinn, Anne Richelle, Isaac Shamie, Eyleen J. O'Rourke, and Nathan E. Lewis. "StanDep: Capturing transcriptomic variability improves context-specific metabolic models." *PLoS computational biology* 16, no. 5 (2020): e1007764.

Jost, Lou. "Entropy and diversity." *Oikos* 113, no. 2 (2006): 363-375.

Jost, Lou. "Partitioning diversity into independent alpha and beta components." *Ecology* 88, no. 10 (2007): 2427-2439.

Kalvari, Ioanna, Eric P. Nawrocki, Nancy Ontiveros-Palacios, Joanna Argasinska, Kevin Lamkiewicz, Manja Marz, Sam Griffiths-Jones et al. "Rfam 14: expanded coverage of metagenomic, viral and microRNA families." *Nucleic Acids Research* 49, no. D1 (2021): D192-D200.

Kim, Daehwan, Joseph M. Paggi, Chanhee Park, Christopher Bennett, and Steven L. Salzberg. "Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype." *Nature biotechnology* 37, no. 8 (2019): 907-915.

Lara, Elena G., Isabelle van der Windt, Douwe Molenaar, Marjon GJ de Vos, and Chrats Melkonian. "Using functional annotations to study pairwise interactions in urinary tract infection communities." *Genes* 12, no. 8 (2021): 1221.

Liaw, Andy, and Matthew Wiener. "Classification and regression by randomForest." *R news* 2, no. 3 (2002): 18-22.

Lieven, Christian, Moritz E. Beber, Brett G. Olivier, Frank T. Bergmann, Meric Ataman, Parizad Babaei, Jennifer A. Bartell et al. "MEMOTE for standardized genome-scale metabolic model testing." *Nature biotechnology* 38, no. 3 (2020): 272-276.

Lindner, Scott E., Kristian E. Swearingen, Melanie J. Shears, Michael P. Walker, Erin N. Vrana, Kevin J. Hart, Allen M. Minns, Photini Sinnis, Robert L. Moritz, and Stefan HI Kappe. "Transcriptomics and proteomics reveal two waves of translational repression during the maturation of malaria parasite sporozoites." *Nature communications* 10, no. 1 (2019): 4964.

Lupo, Federico, Molly A. Ingersoll, and Miguel A. Pineda. "The glycobiology of uropathogenic E. coli infection: the sweet and bitter role of sugars in urinary tract immunity." *Immunology* 164, no. 1 (2021): 3-14.

Kopylova, Evguenia, Laurent Noé, and Hélène Touzet. "SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data." *Bioinformatics* 28, no. 24 (2012): 3211-3217.

Kulkarni, Ashwinikumar, Ashley G. Anderson, Devin P. Merullo, and Genevieve Konopka. "Beyond bulk: a review of single cell transcriptomics methodologies and applications." *Current opinion in biotechnology* 58 (2019): 129-136.

Liu, Bo, Dandan Zheng, Siyu Zhou, Lihong Chen, and Jian Yang. "VFDB 2022: a general classification scheme for bacterial virulence factors." *Nucleic acids research* 50, no. D1 (2022): D912-D917.

Magruder, Matthew, Adam N. Sholi, Catherine Gong, Lisa Zhang, Emmanuel Edusei, Jennifer Huang, Shady Albakry et al. "Gut uropathogen abundance is a risk factor for development of bacteriuria and urinary tract infection." *Nature communications* 10, no. 1 (2019): 5521.

Marinos, Georgios, Christoph Kaleta, and Silvio Waschina. "Defining the nutritional input for genome-scale metabolic models: A roadmap." *PloS one* 15, no. 8 (2020): e0236890.

Martin, Marcel. "Cutadapt removes adapter sequences from high-throughput sequencing reads." *EMBnet. journal* 17, no. 1 (2011): 10-12.

McMurdie, Paul J., and Susan Holmes. "phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data." *PloS one* 8, no. 4 (2013): e61217.

Mediavilla, José R., Amee Patrawalla, Liang Chen, Kalyan D. Chavda, Barun Mathema, Christopher Vinnard, Lisa L. Dever, and Barry N. Kreiswirth. "Colistin-and carbapenem-resistant Escherichia coli harboring mcr-1 and bla NDM-5, causing a complicated urinary tract infection in a patient from the United States." *MBio* 7, no. 4 (2016): 10-1128.

Mey, Alexandra R., Camilo Gómez-Garzón, and Shelley M. Payne. "Iron transport and metabolism in Escherichia, Shigella, and Salmonella." *EcoSal plus* 9, no. 2 (2021): eESP-0034.

Neugent, Michael L., Ashwani Kumar, Neha V. Hulyalkar, Kevin C. Lutz, Vivian H. Nguyen, Jorge L. Fuentes, Cong Zhang et al. "Recurrent urinary tract infection and estrogen shape the taxonomic ecology and function of the postmenopausal urogenital microbiome." *Cell Reports Medicine* 3, no. 10 (2022).

Nielsen, Daniel W., Nicole Ricker, Nicolle L. Barbieri, Heather K. Allen, Lisa K. Nolan, and Catherine M. Logue. "Outer membrane protein A (OmpA) of extraintestinal pathogenic Escherichia coli." *BMC research notes* 13, no. 1 (2020): 1-7.

Noronha, Alberto, Jennifer Modamio, Yohan Jarosz, Elisabeth Guerard, Nicolas Sompairac, German Preciat, Anna Dröfn Daníelsdóttir et al. "The Virtual Metabolic Human database: integrating

human and gut microbiome metabolism with nutrition and disease." *Nucleic acids research* 47, no. D1 (2019): D614-D624.

O'Brien, Edward J., Jonathan M. Monk, and Bernhard O. Palsson. "Using genome-scale models to predict biological capabilities." *Cell* 161, no. 5 (2015): 971-987.

Ojala, Teija, Esko Kankuri, and Matti Kankainen. "Understanding human health through metatranscriptomics." *Trends in Molecular Medicine* 29, no. 5 (2023): 376-389.

Paalanne, Niko, Aleksi Husso, Jarmo Salo, Oskari Pieviläinen, Mysore V. Tejesvi, Pirjo Koivusaari, Anna Maria Pirttilä et al. "Intestinal microbiome as a risk factor for urinary tract infections in children." *European Journal of Clinical Microbiology & Infectious Diseases* 37 (2018): 1881-1891.

Paudel, Santosh, Sarah Guedry, Chloe LP Obernuefemann, Scott J. Hultgren, Jennifer N. Walker, and Ritwij Kulkarni. "Defining the Roles of Pyruvate Oxidation, TCA Cycle, and Mannitol Metabolism in Methicillin-Resistant Staphylococcus aureus Catheter-Associated Urinary Tract Infection." *Microbiology spectrum* 11, no. 4 (2023): e05365-22.

Pertea, Mihaela, Geo M. Pertea, Corina M. Antonescu, Tsung-Cheng Chang, Joshua T. Mendell, and Steven L. Salzberg. "StringTie enables improved reconstruction of a transcriptome from RNA-seq reads." *Nature biotechnology* 33, no. 3 (2015): 290-295.

Pohl, Hans G., Suzanne L. Groah, Marcos Pérez-Losada, Inger Ljungberg, Bruce M. Sprague, Neel Chandal, Ljubica Caldovic, and Michael Hsieh. "The urine microbiome of healthy men and women differs by urine collection method." *International Neurourology Journal* 24, no. 1 (2020): 41.

Puebla-Barragan, Scarlett, Justin Renaud, Mark Sumarah, and Gregor Reid. "Malodorous biogenic amines in Escherichia coli-caused urinary tract infections in women—a metabolomics approach." *Scientific Reports* 10, no. 1 (2020): 9703.

Pust, Marie-Madlen, and Burkhard Tümmler. "Bacterial low-abundant taxa are key determinants of a healthy airway metagenome in the early years of human life." *Computational and Structural Biotechnology Journal* 20 (2022): 175-186.

Qi, Yanjun. "Random forest for bioinformatics." *Ensemble machine learning: Methods and applications* (2012): 307-323.

Richelle, Anne, Chintan Joshi, and Nathan E. Lewis. "Assessing key decisions for transcriptomic data integration in biochemical networks." *PLoS computational biology* 15, no. 7 (2019a): e1007185.

Richelle, Anne, Austin WT Chiang, Chih-Chung Kuo, and Nathan E. Lewis. "Increasing consensus of context-specific metabolic models by integrating data-inferred cell functions." *PLoS computational biology* 15, no. 4 (2019b): e1006867.

Ramachandran, Rachana Pattani, Chaya Spiegel, Yael Keren, Tsafi Danieli, Naomi Melamed-Book, Ritesh Ranjan Pal, Efrat Zlotkin-Rivkin, Ilan Rosenshine, and Benjamin Aroeti. "Mitochondrial targeting of the enteropathogenic Escherichia coli map triggers calcium mobilization, ADAM10-MAP kinase signaling, and host cell apoptosis." *MBio* 11, no. 5 (2020): 10-1128.

Robin, Xavier, Natacha Turck, Alexandre Hainard, Natalia Tiberti, Frédérique Lisacek, Jean-Charles Sanchez, and Markus Müller. "pROC: an open-source package for R and S+ to analyze and compare ROC curves." *BMC bioinformatics* 12, no. 1 (2011): 1-8.

Robinson, John I., William H. Weir, Jan R. Crowley, Tiffany Hink, Kimberly A. Reske, Jennie H. Kwon, Carey-Ann D. Burnham, Erik R. Dubberke, Peter J. Mucha, and Jeffrey P. Henderson.

"Metabolomic networks connect host-microbiome processes to human Clostridioides difficile infections." *The Journal of Clinical Investigation* 129, no. 9 (2019): 3792-3806.

Rosales-Colunga, Luis Manuel, and Antonio De León Rodríguez. "Escherichia coli and its application to biohydrogen production." *Reviews in Environmental Science and Bio/Technology* 14 (2015): 123-135.

Saheb Kashaf, Sara, Diana M. Proctor, Clay Deming, Paul Saary, Martin Hölzer, Monica E. Taylor, Heidi H. Kong, Julia A. Segre, Alexandre Almeida, and Robert D. Finn. "Integrating cultivation and metagenomics for a multi-kingdom view of skin microbiome diversity and functions." *Nature microbiology* 7, no. 1 (2022): 169-179.

Sami, Hiba, Asfia Sultan, Meher Rizvi, Fatima Khan, Shariq Ahmad, Indu Shukla, and Haris M. Khan. "Citrobacter as a uropathogen, its prevalence and antibiotics susceptibility pattern." *Chrismed: Journal of Health & Research* 4, no. 1 (2017).

Schmieder, Robert, and Robert Edwards. "Quality control and preprocessing of metagenomic datasets." *Bioinformatics* 27, no. 6 (2011): 863-864.

Schembri, Mark A., Nguyen Thi Khanh Nhu, and Minh-Duy Phan. "Gut–bladder axis in recurrent UTI." *Nature Microbiology* 7, no. 5 (2022): 601-602.

Schüroff, Paulo A., Fábia A. Salvador, Cecilia M. Abe, Haleluya T. Wami, Eneas Carvalho, Rodrigo T. Hernandes, Ulrich Dobrindt, Tânia AT Gomes, and Waldir P. Elias. "The aggregate-forming pili (AFP) mediates the aggregative adherence of a hybrid-pathogenic Escherichia coli (UPEC/EAEC) isolated from a urinary tract infection." *Virulence* 12, no. 1 (2021): 3073-3093.

Sen, Partho, and Matej Orešič. "Integrating Omics Data in Genome-Scale Metabolic Modeling: A Methodological Perspective for Precision Medicine." *Metabolites* 13, no. 7 (2023): 855.

Shields, Ryan K., Yun Zhou, Hemanth Kanakamedala, and Bin Cai. "Burden of illness in US hospitals due to carbapenem-resistant Gram-negative urinary tract infections in patients with or without bacteraemia." *BMC infectious diseases* 21, no. 1 (2021): 1-12.

Singhal, Lipika, Varsha Gupta, Menal Gupta, Poonam Goel, and Jagdish Chander. "Identification and Sensitivity of Vaginal and Probiotic Lactobacillus species to Urinary Antibiotics." *Journal of Laboratory Physicians* 12, no. 02 (2020): 111-114.

Singer, G. A. C., Nicole A. Fahner, J. G. Barnes, Avery McCarthy, and Mehrdad Hajibabaei. "Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater." *Scientific reports* 9, no. 1 (2019): 5991.

Song, Seung Hoon, and Claire Vieille. "Recent advances in the biological production of mannitol." *Applied microbiology and biotechnology* 84 (2009): 55-62.

Thiele, Ines, Swagatika Sahoo, Almut Heinken, Johannes Hertel, Laurent Heirendt, Maike K. Aurich, and Ronan MT Fleming. "Personalized whole-body models integrate metabolism, physiology, and the gut microbiome." *Molecular systems biology* 16, no. 5 (2020): e8982.

Vieira, Vítor, Jorge Ferreira, and Miguel Rocha. "A pipeline for the reconstruction and evaluation of context-specific human metabolic models at a large-scale." *PLoS computational biology* 18, no. 6 (2022): e1009294.

Vlassis, Nikos, Maria Pires Pacheco, and Thomas Sauter. "Fast reconstruction of compact context-specific metabolic network models." *PLoS computational biology* 10, no. 1 (2014): e1003424.

Werber, Dirk, Martina Bielaszewska, Christina Frank, Klaus Stark, and Helge Karch. "Watch out for the even eviler cousin—sorbitol-fermenting E coli O157." *The Lancet* 377, no. 9762 (2011): 298-299.

Wingett, Steven W., and Simon Andrews. "FastQ Screen: A tool for multi-genome mapping and quality control." *F1000Research* 7 (2018).

Wishart, David S., AnChi Guo, Eponine Oler, Fei Wang, Afia Anjum, Harrison Peters, Raynard Dizon et al. "HMDB 5.0: the human metabolome database for 2022." *Nucleic acids research* 50, no. D1 (2022): D622-D631.

Worby, Colin J., Henry L. Schreiber IV, Timothy J. Straub, Lucas R. van Dijk, Ryan A. Bronson, Benjamin S. Olson, Jerome S. Pinkner et al. "Longitudinal multi-omics analyses link gut microbiome dysbiosis with recurrent urinary tract infections in women." *Nature microbiology* 7, no. 5 (2022): 630-639.

Yang, Huan, Yiran Sun, Rui Cai, Ying Chen, and Bing Gu. "The impact of dietary fiber and probiotics in infectious diseases." *Microbial pathogenesis* 140 (2020): 103931.

Yilmaz, Pelin, Laura Wegener Parfrey, Pablo Yarza, Jan Gerken, Elmar Pruesse, Christian Quast, Timmy Schweer, Jörg Peplies, Wolfgang Ludwig, and Frank Oliver Glöckner. "The SILVA and “all-species living tree project (LTP)” taxonomic frameworks." *Nucleic acids research* 42, no. D1 (2014): D643-D648.

Zampieri, Guido, Supreeta Vijayakumar, Elisabeth Yaneske, and Claudio Angione. "Machine and deep learning meet genome-scale metabolic modeling." *PLoS computational biology* 15, no. 7 (2019): e1007084.

Zampieri, Guido, Stefano Campanaro, Claudio Angione, and Laura Treu. "Metatranscriptomics-guided genome-scale metabolic modeling of microbial communities." *Cell Reports Methods* 3, no. 1 (2023).

Zandbergen, Laurens E., Thomas Halverson, Jolanda K. Brons, Alan J. Wolfe, and Marjon GJ De Vos. "The good and the bad: ecological interaction measurements between the urinary microbiota and uropathogens." *Frontiers in Microbiology* 12 (2021): 659450.

Zhao, Xian-liang, Zhuang-gui Chen, Tian-ci Yang, Ming Jiang, Jie Wang, Zhi-xue Cheng, Man-jun Yang et al. "Glutamine promotes antibiotic uptake to kill multidrug-resistant uropathogenic bacteria." *Science Translational Medicine* 13, no. 625 (2021): eabj0716.

Zheng, Danping, Timur Liwinski, and Eran Elinav. "Interaction between microbiota and immunity in health and disease." *Cell research* 30, no. 6 (2020): 492-506.

Zimmermann, Johannes, Christoph Kaleta, and Silvio Waschina. "gapseq: Informed prediction of bacterial metabolic pathways and reconstruction of accurate metabolic models." *Genome biology* 22, no. 1 (2021): 1-35.