Urinary biochemical ecology reveals microbiome-metabolite interactions and metabolic markers of recurrent urinary tract infection

Michael L. Neugent[1], Neha V. Hulyalkar[1], Debasish Ghosh[3], Philippe E. Zimmern[2], Vladimir Shulaev[3], Nicole J. De Nisco[1,2]*

[1]Department of Biological Sciences, The University of Texas at Dallas, Richardson, Texas
[2]Department of Urology, The University of Texas Southwestern Medical Center, Dallas, Texas
[3]Department of Biological Sciences, The University of North Texas, Denton, Texas

*Correspondence may be addressed to Nicole J. De Nisco, PhD (nicole.denisco@utdallas.edu)

Keywords: urinary microbiome, metabolomics, ecology, urinary tract infection, biomarker

Abstract

Recurrent urinary tract infections (rUTIs) are a major clinical challenge in postmenopausal women and their increasing prevalence underscores the need to define interactions between the host and the urinary microbiome that may underlie rUTI susceptibility. A body of work has identified the taxonomic profile of the female urinary microbiome associate with aging, menopause, and urinay disease. However, how this microbial community engages with the host niche, including the local biochemical environment of the urogenital tract, in health and disease is yet to be fully defined. This study directly assesses differences in the biochemical environment of the urine, or biochemical ecology, associated with recurrent urinary tract infection (UTI) and defines a microbe-metabolite association network of the female urinary microbiome. By integrating metagenomic and metabolomic data collected from a controlled cohort of women with rUTI, we find that distinct metabolites, such as methionine sulfoxide (Met-SO) and trimethylamine oxide (TMAO), are associated with differences in urinary microbiome diversity. We observe associations between microbial and biochemical beta diversity and unique metabolic networks of uropathogenic *Escherichia coli* and uroprotective *Lactobacillus* species, highlighting potential metabolite-driven ecological shifts that may influence UTI susceptibility. We identify a urinary lipid signature of active rUTI that can accurately distinguish (AUC = 0.987) cases controls. Finally, using time-to-relapse data we identify deoxycholic acid (DCA) as a new prognostic indicator for rUTI recurrence. Together these findings suggest that systemic metabolic processes may influence susceptibility, opening new avenues for therapeutic intervention and the development of more accurate diagnostic and prognostic to improve patient outcomes.

Introduction

Urinary tract infections (UTIs) are among the most common bacterial infections in adults worldwide with many UTIs progressing into recurrent UTI (rUTI) or chronic UTI.[1,2] Recurrent and chronic UTIs are often difficult to manage, and their recurrence despite antibiotic intervention suggests underlying alterations in the urogenital ecosystem, including both host and microbial factors, contribute to susceptibility.[3-5] Indeed, the decline in circulating estrogen levels associated with menopause is one of the primary risk factors for rUTI with reports of up to 50% of UTIs progressing to rUTI in this population.[1,2,5,6] Treatments for rUTI have focused on modifying the microbial environment of the urine, often through achieving urinary tract sterility via long-term and prophylactic antibiotic regimens to eliminate uropathogen populations.[4,7,8] However, the discovery of a urinary microbiome in healthy individuals has highlighted the need

to develop new treatment paradigms for rUTI focused not just on eradicating pathogens, but also on restoring a healthy urinary microbiome.[9,10]

The microbial communities of the urogenital tract of both men and women, here termed the urinary microbiome, have been extensively characterized and linked to both health and disease.[10-22] Research has shown that, in females, certain urinary microbial communities, such as those dominated by specific species of the genus *Lactobacillus*, are associated with health, while others, including those with *Escherichia coli*, are frequently linked to UTI.[5,23] Despite extensive compositional profiling of the urinary microbiome, less is known about how these microbial communities influence or are influenced by the metabolic environment of the urine. The biochemical landscape of the urinary tract, which includes a wide range of metabolites, may reflect key physiological and pathological states.[24-26] The urinary metabolome, particularly lipid and amino acid profiles, could offer insights into host responses to infection and metabolic relationships between the host and the urinary microbiota. Previous studies have suggested links between urinary metabolites like putrescine and uropathogenic *Escherichia coli*, but no comprehensive effort has been made to explore the urinary metabolome and its relationship with the taxonomic ecology of the urinary microbiome in health and disease.[27] Expanding our knowledge of both the taxonomic and biochemical ecology of the urinary microbiome can enhance our understanding of how this microbial community interacts with host health and may uncover clinically translatable biomarkers to improve the diagnosis, prognosis, and treatment of UTI, rUTI and chronic UTI.

The concept of “biochemical ecology” refers to the functional interactions between metabolites and microbial communities. Emerging evidence suggests that metabolites not only result from microbial activity but also shape the microbial ecosystem itself, potentially influencing microbiome function and susceptibility to infection.[28-30] However, little is known about the relationship between the taxonomic and biochemical ecology of the urinary tract and how these may jointly contribute to the pathophysiology of rUTI. This study aims to bridge the gap by integrating metagenomic and metabolomic data to define associations between urine biochemical ecology and taxonomic ecology in postmenopausal women with and without a history of rUTI.[5] Specifically, we investigate whether distinct biochemical signatures are linked to microbial communities and disease states using a multiplexed targeted metabolomics assay quantifying more than 600 metabolites.[31,32] Our findings suggest that the urinary metabolome reflects not just microbial activity but also broader host responses, such as oxidative stress, potentially linking local inflammation to microbial compositional diversity. In addition to identifying new relationships between the urine metabolites and urinary microbiome species, we

identified a suite of urinary lipids that have potential as highly accurate diagnostic biomarkers of active UTI. We also find that 12 metabolites, including deoxycholic acid (DCA), were prognostic for increased rUTI risk and a decrease in recurrence free days. These findings have the potential to improve our understanding of rUTI pathogenesis and open avenues for the development of metabolome-based diagnostic and therapeutic approaches.

Results

Targeted LC-MS/MS of the urinary polar metabolome and lipidome.

To quantitatively characterize the urinary metabolome associated with urinary tract infection (UTI) and the urinary microbiome, we performed targeted liquid chromatography tandem mass spectrometry (LC-MS/MS)-based metabolite profiling of more than 600 polar and non-polar metabolites in urine collected from three groups of postmenopausal women differentiated by their UTI history and current UTI status (n=75, 25 per group) (**Fig. 1A**).[5] Group 1 (No UTI History) consisted of postmenopausal women with no lifetime history of UTI and was used as a healthy comparator. Groups 2 and 3 consisted of postmenopausal women with a previous history of rUTI but differed on UTI status at time of urine donation. Group 2 (rUTI History, UTI(-)) consisted of women with a recent history of rUTI who were not experiencing a UTI at the time of urine donation, while Group 3 (rUTI History, UTI(+)) consisted of postmenopausal women with a clinically confirmed UTI at the time of urine donation.

The measured urinary metabolites included a broad range of metabolites including polar amino acids, biogenic amines, and indole derivatives, as well as less polar lipid species like ceramides, sphingolipids, and glycerophospholipids (**Fig. 1A**). We limited our analysis to metabolites that were reliably detected at levels above the calculated limit of detection in at least 7 samples across all groups. A total of 159 polar metabolites and 223 lipid species were selected for further downstream analysis (**Fig. 1B, C**). We next assessed correlation between the concentration of each metabolite and the creatinine-normalized metabolite abundance. We observed a strong correlation between the creatinine-normalized and raw micromolar metabolite concentration (R = 0.971, $P < 2.22\text{x}10^{-16}$, n = 28,350 pairwise comparisons) (**Fig. 1D**). While both values have utility for distinct questions, we chose to focus our analysis on the raw urinary concentration of each metabolite given the strong linear correlation to the creatinine-normalized abundance and our intention to associate the biochemical ecology with the local microbial ecology, a set of variables not necessarily associated with glomerular filtration and creatinine excretion rate.[3,33,34]

Urine biochemical ecology is moderately associated with the taxonomic ecology

To characterize the extent of associations between metabolic and microbial ecological signatures defining rUTI susceptibility and active UTI, we sought to test the hypothesis that there may be broad associations between urinary microbiome and biochemical ecology. We observed no difference in the biochemical alpha diversity, as measured by the inverse Simpson Hill Number, between groups even though there was a detectable difference in microbial alpha diversity between group 2 and group 3 (**Fig. 2A, B**).[5] The urinary microbiome taxonomic alpha diversity also exhibited a much larger range (7.15 to 22.05) than that of the biochemical ecology (2.26 to 7.15) (**Fig. 2B**). Additionally, we observed no association between the overall biochemical and taxonomic alpha diversity of urine (**Fig. 2C**). These data suggest that while the urinary microbiome and metabolome are components of the same ecosystem, their alpha diversities, particularly the evenness of the diversity, do not appear to be associated in the aggregate cohort studied. However, we observed significant associations between distinct urinary metabolites and taxonomic alpha diversity. Urinary concentrations of methionine sulfoxide (Met-SO), citrulline (Cit), and trimethylamine oxide (TMAO) were all significantly positively associated with both the Shannon and Simpson taxonomic alpha diversity (**Fig. 2D-F**). Met-SO is the product of methionine oxidation by reactive oxygen species (ROS) TMAO is known to be associated with cardiovascular disease and hypertension.[35]

We next hypothesized that urine samples with distinct taxonomic compositions would also have distinct urinary metabolic compositions. To test this hypothesis, we analyzed covariance of beta diversity between the taxonomic compositional data and the metabolomic data (**Fig. 2G, H, I**). For both datasets, we screened a number of commonly used methods to calculate beta diversity distance, including the Bray-Curtis, Jaccard, Euclidean, Manhattan, Canberra, Gower, and Kulczynski distances.[36-38] To measure covariance between all combinations of the distance matrices of taxonomic ecology and biochemical ecology, we used the non-parametric Spearman variant of the Mantel test.[39] We observed significant covariance ($P < 0.05$) when the biochemical beta diversity was modeled using the Canberra distance (**Fig. 2G**). However, the association detected was relatively weak (R=0.15). Since the Canberra distance calculation normalizes the feature differences before summation, it allows for equal representation of differences in high and low abundant features when calculating pairwise distance between samples. Therefore, we interpret this observation as indicative that low abundance metabolites may be driving any observed association between taxonomic and biochemical beta diversity.

We further observed that pairwise biochemical beta diversity distances were higher when comparing between groups relative to within groups, suggesting that differences in disease state are concurrently associated with altered urinary biochemical ecology (**Fig. 2H**). Between group urinary metabolic beta diversity differences were largest between Group 1 (No UTI History) and Group 3 (rUTI History, UTI (+)) ($P < 10^{-9}$). Given these results, we hypothesized that there may be a less obvious association between taxonomic and metabolic beta diversity that is not captured well in assessments of pairwise associations, such as the Mantel test. We used Procrustes analysis to assess the alignment of independent biochemical and taxonomic PCoA ordinations.[40] This analysis found significant, moderate alignment (m^2 = 0.58, P = 0.001) when modeling beta diversity distance of metabolites using the Gower distance and taxonomic composition using the Euclidean distance (**Fig. 2I**). Similarly to as observed in Fig. 2G, associations between urinary biochemical and taxonomic beta diversity were identified when modeling biochemical distances using metrics that normalize each feature to its total range across all samples. Taken together, these data suggest the covariance between urinary taxonomic and biochemical beta diversity is driven primarily by low abundance metabolites.

Escherichia **is a metabolic hub defined by a high degree of unique metabolite associations**

We next sought to define the structure of the individual taxon-metabolite association network. To accomplish this, we performed an all-against-all Spearman correlation analysis between the urinary microbiome genus-level taxonomic profile and urinary metabolomic profile. For this and all subsequent analyses we focused solely on Groups 1 and 3 and removed the intermediate disease state Group 2 so that clear comparisons could be made between cases and controls. We used the significant ($R > |0.3|$, $P < 0.05$) correlation results between taxa-metabolite pairs to generate both positive and negative association networks with nodes representing metabolites (purple) or taxa (pink) and edges drawn only between taxa and metabolites (**Fig. 3A, B**). The positive association network was relatively sparse with a striking separation between the larger metabolite-taxa interaction network and that of a single taxon, *Escherichia* (**Fig. 3A**). Not only was *Escherichia* association with a highly distinct set of metabolites, but it was also the taxon node with highest degree, or number of first neighbors (n = 122 metabolites) (**Fig. 3A, C**). The majority of metabolites associated with *Escherichia* abundance are not strongly connected to other urinary microbiome taxa as indicated by a relatively low topological coefficient ($T(i)$ = 0.0139) (**Fig. 3C**). These data suggest that urinary *Escherichia,* which in this dataset is largely associated with active UTI, is associated with a

unique set of urinary metabolites, few of which are associated further with additional taxa.[5] Among these metabolites were putrescine (R = 0.661, P = 1.37x10^{-6}), one of the major polyamines produced by *E. coli,* and the diacylglycerol, DG(14:0, 20:0) (R = 0.638, P = 4.18x10^{-6}).[27,41] We hypothesize that this observation is evidence of a unique urinary metabolic signature associated with UTI, the majority of which in this cohort were culture-confirmed UPEC infection.[5]

When analyzed from the perspective of the metabolite nodes, the metabolite proline betaine (ProBetaine), a proline derivative also known as stachydrine, was identified as the highest degree metabolite node, being connected to more taxa than any other metabolite (n = 14 genera) (**Fig. 3D**).[42] ProBetaine was positively associated with the genera *Slackia* (R = 0.436, P = 0.003), *Schaalia* (R = 0.398, P = 0.008), *Peptococcus* (R = 0.383, P =0.011), and *Mobiluncus* (R = 0.367, P = 0.016). The negative taxa-metabolite association network was observed to be a more densely connected network with the genera, *Rothia* (n =82 metabolites), *Fusobacterium* (n = 71 metabolites), *Staphylococcus* (n = 65 metabolites), *Facklamia* (n = 62 metabolites), and *Dermabacter* (n = 58 metabolites) making up the top 5 taxa connected with the greatest number of metabolites (**Fig. 3B, E**). The hexosylceramide, HexCer(d18:1/24:0), was the metabolite node with the highest degree, being negatively associated with 22 genera (**Fig. 3F**).

The unique metabolite network of *Lactobacillus gasseri* separates it from other FUM lactobacilli.

While much of the taxa-metabolite network analysis revealed associations between metabolites and known uropathogens, we further sought to determine the *Lactobacillus*-metabolite associations to identify biochemical features associated with these beneficial microbial communities.[43] Our analysis of the species-level *Lactobacillus* positive association network revealed two non-interacting networks – one composed solely of *L. gasseri* and the other composed of *L. jensenii*, *L. crispatus*, and *L. iners* (**Fig. 3G**). *L. gasseri* positively associated with a unique and distinct set of 17 metabolites. The strongest metabolite association with *L. gasseri* was the triglyceride TG(17:2, 36:4), a triacyl glycerol (R = 0.542, P = 0.00018). The remaining *Lactobacillus* species all associated with common metabolites within a shared network neighborhood (**Fig. 3G**). Both *L. jensenii* and *L. crispatus* exhibited relatively strong associations with the fatty acid, eicosapentaenoic acid (EPA) ($R_{jensenii}$ = 0.503, $R_{crispatus}$ = 0.450), while all three taxa, *L. iners, L. jensenii,* and *L. crispatus* were correlated with the triacylglycerol, TG(16:0, 40:8) (R_{iners} = 0.447, $R_{jensenii}$ = 0.377, $R_{crispatus}$ = 0.354). The diacylglycerol, DG(16:0, 16:0), was found to be associated with *L. jensenii* (R = 0.350) and *L.*

iners (R = 0.342) and more weakly with *L. crispatus* (R = 0.301) (**Fig. 3G**). The negative association network also separated *L. gasseri* from the other FUM *Lactobacillus* species and revealed that *L. jensenii* and *L. crispatus* were highly interconnected, negatively correlating with a common set of 28 metabolites. The strongest shared anti-correlation of *L. crispatus* and *L. jensenii* was the phospholipid, PC(ae C44:3) ($R_{crispatus}$ = -0.426, $R_{jensenii}$ = -0.402), which was also negatively associated with *L. iners* (R = -0.359) (**Fig. 3H**). *L. gasseri* was negatively associated with only a single metabolite, urinary hexose (**Fig. 3H**). Given the highly similar chromatographic behavior and mass spectral features of hexoses, such as glucose, galactose, and fructose, we cannot determine which hexose(s) drive this negative metabolite association.[44] Together, our data suggests that urinary *L. gasseri* communities are both positively and negatively associated with a urinary metabolic signature that is unique among the major FUM lactobacilli.

Active UTI is associated with distinct metabolic signature

We hypothesized that the urinary metabolome between PM women who do not get UTIs and those actively presenting with a UTI would be significantly different and may harbor novel biomarkers of active UTI. To test this hypothesis, we performed a differential enrichment analysis screen of metabolites between Group 1 (No UTI History) and Group 3 (rUTI History, UTI(+)) using the Wilcoxon rank sum test (**Fig. 4A**). This analysis identified many significantly differentially abundant metabolites enriched either in the No UTI History (n=9) or rUTI History, UTI(+) (n=149) (**Fig. 4B, C**). The observation that putrescine was among the most differentially abundant metabolites enriched in urine during active rUTI (FC_{median} = 12.63, Q = 3.77x10^{-6}) is in line with expectations based on previous observations and can be attributed to largely to *Escherichia coli* as putrescine was found in be highly correlated to *Escherichia* in our taxa-metabolite network analysis in Figure 3 (**Fig. 4B, D**).[27] The remaining top 10 differentially abundant metabolites enriched during active rUTI were all lipid species, such as diacylglycerols (DG), ceramides (Cer), and glyceroceramides (Hex2Cer) (**Fig. 4B, E, F, G**).

In contrast, we observed a separate set of 9 metabolites significantly enriched in the urine of PM women in Group 1, the women with No UTI history. Differentially abundant metabolites include three fatty acids (FA 20:1, 20:3, and 18:0), methionine sulfoxide (MetSO), the amino acids isoleucine (Ile) and leucine (Leu), and a diacylglycerol (DG) (**Fig. 4C, H, I, J, K**). Interestingly, two dicarboxylic acid metabolites, DiCA(14:0) and DiCA(12:0), were the most significant differentially abundant metabolites enrich in PM women who do not get UTI ($Q < 10^{-7}$, Q=0.00074) (**Fig 4 C, H, I)**. These metabolites have previously been detected in urine and are typically linked to gamma-oxidation of long-chain fatty acids.[45] While, in many cases, we cannot

robustly decouple host-derived metabolites from microbial metabolites, these data do suggest a unique urinary metabolic signature associated with women who do not experience UTI.

Taxa associated with metabolites differentially enriched in No UTI history

Given that a unique set of metabolites were enriched in women with No UTI history that may be associated with urogenital health, we further sought to determine urinary microbiome taxonomic associations with these metabolites. We performed a Spearman correlation analysis between the quantitative abundance of the nine metabolites enriched in the No UTI History group and the urinary microbiome taxonomic profile identified by MetaPhlAn4 in the same group.[46] This analysis revealed 84 taxa-metabolite associations passing at least nominal significance at the genus and species levels (**Figure 5A, B**). Of note, we observed robust associations between methionine sulfoxide (MetSO) with 4 genera and 12 species. The genus, *Lactobacillus*, as well as the species *Lactobacillus crispatus*, were the most negatively correlated taxa with urinary MetSO concentration ($R_{Lactobacillus}$ = -0.526, $R_{L.crispatus}$ = -0.551) (**Fig. 5C, D**). Conversely, we observed a relatively strong positive association between MetSO and the dysbiotic taxon, *Streptococcus anginosus* (R = 0.550, P = 0.004) (**Fig. 5E**). The genus, *Ureaplasma*, which is considered an opportunistic pathogen in the female urogenital tract, was positively associated with DiCA (12:0) (R = 0.469, P = 0.018) (**Fig. 5A**).[47] We further observed a modest positive association between urinary fatty acid (20:1) and *Staphylococcus epidermidis* (R = 0.460, P = 0.012) (**Fig. 5F**).

Urinary lipids as biomarkers for active rUTI

Given our observation that many of the differentially abundant metabolites enriched in women with active rUTI were lipids, we hypothesized the urinary lipids may be an indicator or perhaps novel biomarker of active UTI. When summed, the total assayed lipid content detected in urine exhibited only a modest elevation in active rUTI (FC_{median} = 1.42, P = 0.00017) (**Fig. 6A**). We then hypothesized that a specific subset of urinary lipids could differentiate case (active rUTI) from control (No UTI history) groups. To test this hypothesis, we performed elastic net regularization on assayed lipids to identify an optimal subset of lipids distinguishing active rUTI case from controls. This analysis identified 7 lipids in an optimal multivariate model whose abundance was able to distinguish the active rUTI samples in the rUTI History, UTI(+) group from the control samples in the No UTI History group (**Fig. 6B**). Diacylglycerol (14:0, 20:0), which was also the most differentially abundant metabolite enriched in the UTI(+) group contributed the most to the multivariate model with ~66% of the discrimination being attributed

to this metabolite (**Fig. 6B**). To further assess the predictive power of this set of discriminating lipids, we performed logistic regression and leave one out cross validation (LOOCV) analysis (**Fig. 6C, D**). Logistic regression indicated that the seven lipid metabolites were able to strongly discriminate cases from controls (G^2 = 55.55, P = 2.03x10^{-12}) and the area under the reporter operator curve (ROC) of the 7 lipid metabolite multivariate model calculated by LOOCV was 0.987 (**Fig. 6D**). These results suggest that this discriminating lipid signature has excellent sensitivity and specificity on cross validation; however, these results should be validated in an independent cohort.

To assess the association of taxa abundant in women with active rUTI with the lipid signature of rUTI, we performed a correlation analysis with the sample-matched taxonomic genus and species level profile previously generated by whole genome metagenomics. The strongest taxa-lipid positive associations were observed between diacylglycerols (DGs) and *Escherichia* or *E. coli* (**Fig. 6E**). *E. coli* abundance was associated with DGs (14:0, 20:0) and (16:1, 18:1), as well as phosphatidyl choline (eaC44:4). *E. coli* abundance was strongly correlated with urinary DG(14:0, 20:0) concentration (R = 0.67, P = 1.3x10^{-7}). Interestingly, we did observe high concentrations of DG(14:0, 20:0) in the urine of one UTI(+) individual who did not have a UPEC infection but instead had a diverse urinary microbiome predominated by *Anaerococcus sp.* (**Fig. 6F**). From these data, we hypothesize that the increase in urinary lipids observed in individuals with active rUTI is associated with acute tissue damage (i.e. epithelial cell death and exfoliation) that occur during infection.[48] These urinary lipids may be a novel source of diagnostic biomarkers for active UTI with rapid translatability.[49]

The Urinary Metabolome Harbors Translatable Prognostic Indicators of rUTI.

We next asked if any urinary metabolites were associated with an increased hazard of recurrent UTI. Of the 25 with active UTI in Group 3, 19 had reliable time-to-next UTI metadata. We performed Cox proportional hazard analysis to identify metabolites that were significantly associated with UTI recurrence rate. The urinary concentration of 12 metabolites was significantly associated with increased hazard of UTI recurrence within the follow-up period (**Fig. 7A**). Interestingly, the secondary bile acid, Deoxycholic acid (DCA), was the metabolite most significantly associated with increased hazard of recurrent UTI (HR = 3.99, P = 0.0015). The remaining significant metabolites identified included three cholesterol esters (CE), Taurine, three Diacylglycerols (DG), the secondary bile acid Glycochendeoxycholic acid (GCDCA), Phosoptidylcholine (C30:2), Homoarginine, and Glycosylceramide (18:1, 18:0) which all had hazard ratios ranging between two and three (**Fig. 7B**).

To further characterize the association between urinary DCA concentration and the risk of UTI recurrence, we performed Kaplan-Meier analysis, dichotomizing the patients by the median urinary DCA concentration ([DCA] = 0.007μM, n_{high} = 9, n_{low} = 10) (**Fig. 7C**). Women with above median urinary concentrations of DCA exhibited a hazard ration (HR) of 2.99 (P=0.0296) for UTI recurrence compared to DCA low patients (**Fig. 7D**). Considering the differences in between continuous hazard models, such as the Cox model, and dichotomous Kaplan-Meier, we consider the hazard identified by both models to be in line with one another and supportive of DCA as a prognostic indicator for rUTI. Indeed, the median of recurrence-free days for women in high urinary DCA was 86, compared to 254.5 recurrence-free days in the DCA low group (P =0.035) (**Fig. 7E**). While these data need to be validated in a larger patient population, they suggest that the urine metabolome may harbor valuable prognostic biomarkers for improved rUTI clinical management.

Discussion

Here, we present a comprehensive analysis that integrates both metagenomics and metabolomics of the urinary microbial niche. We identify novel potential biomarkers of UTI and characterize how the urinary microbial and biochemical ecologies are interrelated. A critical challenge in metabolomic experiments that sample a host-microbiome ecological niche is distinguishing between host-derived and microbially derived metabolites.[50,51] While certain metabolites can be reasonably inferred to originate from either the host or microbiome, future research will be necessary to fully untangle the origins of metabolites and broader metabolic signatures in the host-urinary microbiome relationship. Our targeted LC-MS/MS analysis provided a broad and detailed overview of the urinary metabolome, capturing both polar and non-polar metabolites. Targeted approaches offer robust estimates of absolute analyte concentrations in contrast to the relative abundance estimates typical of untargeted metabolomics approaches.[31] However, targeted metabolomics cannot assess features not included in a predefined list of quantified targets. This limitation is significantly mitigated by the extensive range of analytes included in the Biocrates Quant500 system. Generally, we found that polar metabolites were present at higher concentrations in urine compared to the detected lipidome. This is expected due to the lower aqueous solubility of lipid species. It is important to note that, although the samples were centrifuged to remove larger particulates, no aqueous/organic partitioning was performed to separate lipids from polar metabolites. Lipid signals detected could have originated from cellular debris not pelleted by centrifugation. Given

that our aim was to identify clinically translatable metabolic markers of disease and biochemical ecology, we believe that this limited sample processing most closely reflects what would be most feasible in a clinical setting.

Little is known about the urinary microbial ecosystem beyond its taxonomic profile and how it is affected by urologic disease.[9] We aimed to characterize the urinary biochemical landscape by applying an ecological perspective to urinary metabolomic data, examining its covariance with taxonomic ecology derived from paired metagenomic data representing the same sample. Interestingly, no association was found between urinary biochemical alpha diversity and urinary microbiome alpha diversity. This is not completely unexpected; however, because in this ecosystem alpha diversity is not predictive of disease state alone as both UTI and a healthy urinary microbiome are often characterized by low alpha diversity. It is also possible that the host's dietary and lifestyle factors likely play a more significant role in shaping urinary biochemical diversity than urinary microbiome composition. However, we did identify specific metabolites linked to urinary microbiome alpha diversity. Methionine sulfoxide (Met-SO), an oxidized form of methionine, was significantly correlated with urinary microbiome alpha diversity. Met-SO is known to accumulate during oxidative stress and inflammation.[52,53] Given that a "healthy" urinary microbiome is often characterized by low alpha diversity and dominance by *Lactobacillus* species, it is possible that Met-SO serves as a marker of inflammation associated with a more diverse urinary microbiome.

We also observed an association between the beta diversity of biochemical and taxonomic ecologies of urine, finding that as urine samples became more taxonomically distinct, they also became more metabolically distinct. However, this significant association was only observed when biochemical beta diversity was calculated using the Canberra distance. We interpret this to suggest that the metabolites driving this association are likely of low abundance as the Canberra distance normalizes individual feature distances before summing to generate the aggregate distance, which in contrast with the Bray-Curtis distance, allows analysis across a wide range of feature abundances.[54] Based on these findings, we propose that the association between the urinary microbiome and metabolome is modest and primarily driven by the most abundant taxa and the least abundant metabolites. Given that host diet and lifestyle likely strongly influence urinary metabolite composition and that the urinary microbiome is a relatively low-biomass microbial community, we find this interpretation reasonable.[26]

Active UTI dramatically shifts the taxonomic and genomic landscape of the urinary microbiome, as we have previously shown in this same cohort.[5,23] We hypothesized that these large-scale taxonomic shifts would also correlate with significant changes in the urinary

metabolome. Our analysis uncovered not only novel metabolic enrichments associated with infection and health but also confirmed known metabolic enrichments observed in UPEC infection, such as putrescine.[27] Interestingly, we found a substantial number of lipids enriched in samples during active UTI. Two diacylglycerols, DG(14:0, 20:0) and DG(16:1, 18:1), were the most significantly enriched metabolites in the urine of women with active UTI. Many ceramides and ceramide derivatives were also enriched in the urine of women with active UTI, which we hypothesize are a signature of bladder epithelial tissue damage as the bladder is a particularly ceramide-rich tissue and ceramide biosynthesis is not widespread among bacteria.[55,56] The finding of a unique lipid signature associated with active UTI provides rationale for the exploration of new metabolite-based diagnostic paradigms for UTI.

The need for novel point-of-care (POC) diagnostics for UTI is significant, as current dipstick-based approaches achieve a specificity of only around 50%, and the diagnostic window for current “gold standard” culture-based methods is 24-48 hours.[34,49] We identified a multivariate model of seven lipids—including DG(14:0, 20:0), PC(ae C44:4), Cer(d18:2, 16:0), DG(16:1, 18:1), TG(16:0,38:3), lysoPC(a C20:3), and dhCer(d18:0, 16:0) that showed outstanding performance in distinguishing active UTIs from healthy controls. Although this finding needs to be validated in an independent cohort, it shows promise as rapid methods for lipid detection, such as matrix-assisted laser desorption/ionization time of (MALDI) mass spectrometry, offer opportunities to translate these biomarkers for use in for clinical diagnostics.[57,58] Furthermore, while the urinary abundance of UPEC was positively correlated with many of the lipids in this model, we also observed elevated concentrations of discriminating lipids in UTI patients without UPEC infection. Thus, this urinary lipid signature may. Have broad utility in identifying active UTI, independent of causative pathogen.

Our taxa-metabolite correlation analysis within the *No UTI History* group, focusing on metabolites enriched in healthy urine, also revealed significant associations. Notably, we found negative correlations between urinary methionine sulfoxide (Met-SO) and both the genus *Lactobacillus* and the species *L. crispatus*. This finding is in line with our additional observation the urinary Met-SO is positively associated with urinary microbiome alpha diversity and suggests a decrease in oxidative stress may be associated with the predominance of beneficial *Lactobacillus*. While we cannot determine whether the urinary Met-SO is derived from the bladder or from systemic sources, its association with increased oxidative stress may indicate local inflammation.[52,53] Interestingly, *Streptococcus anginosus*, a species linked to urogenital infections and urogenital dysbiosis, exhibited a strong positive association with urinary Met-SO.[59] Taken together, these findings suggest a plausible link between taxonomic ecological

state (healthy vs. dysbiotic) and metabolic markers of inflammation. Future work may focus on translating these findings into model systems to explore the source and implications of these associations.

There is a significant lack of molecular markers available to clinicians who aim to take a personalized approach in treating rUTI. We analyzed the association between the urinary biochemical ecology and UTI recurrence hazard to identify robust candidates for prognostic biomarkers of rUTI. Our analysis uncovered 12 candidate urinary metabolites associated with an increased hazard of UTI recurrence. The top candidate was deoxycholic acid (DCA), a microbially produced secondary bile acid, which had a hazard ratio of ~4 in a Cox Hazard Model and ~3 via Kaplan-Meier analysis when dichotomized by median DCA concentration.[60] The source of DCA is particularly noteworthy. While it is possible that local microbial communities in the urinary tract are producing DCA from host-derived cholic acid, increased urinary DCA may alternatively reflect systemic gut dysbiosis, potentially affecting gut permeability, bile acid cycling, and host-microbe interactions in the gastrointestinal tract. It is conceivable that elevated circulating DCA is secreted into the urine as a result of gut dysbiosis and inflammation, which may be linked to increased susceptibility to rUTI. This is further evidenced by our observation that an additional bile acid, glycochenodeoxycholic acid (GCDCA), was also prognostic of UTI recurrence hazard. GCDCA is a host derived secondary bile acid produced in the liver and is involved in lipid uptake.[61] This hypothesis aligns with the work of Worby et al. and Thanert et al., who contemporaneously provided evidence for gut dysbiosis, inflammation, and pathogen transmission along the gut-bladder axis in the evolution of UTI.[62,63]

Taken together our findings underscore the critical need for advancing the understanding of the urinary metabolome and its relationship with the urinary microbiome, especially in the context of rUTI and rUTI susceptibility. By identifying novel biochemical markers and elucidating metabolic associations with both microbial taxa and infection states, this work represents a meaningful step toward the development of metabolic biomarkers of UTI that can be leveraged in the development of more accurate diagnostic and prognostic paradigms. Our findings of a distinct urinary metabolome associated with microbial signatures of health and disease should be further investigated to identify metabolic interactions critical to the maintenance of a beneficial microbiota and prevention of urologic diseases like UTI. Direct interrogation of the intricate interactions between urinary biochemical and microbial ecologies interactions will be critical to inform personalized, microbiome-based approaches to UTI treatment to ultimately improve patient outcomes and reduce UTI recurrence.

Methods

Cohort Curation and definitions

Human subjects were all postmenopausal females recruited from the Urology Clinic at The University of Texas Southwestern Medical Center under IRBs STU032016-006 (University of Texas Southwestern Medical Center) and 19MR001 (University of Texas at Dallas) as previously described.[5] Written informed consent was collected from each patient. The following exclusion criteria were applied to screen patients for eligibility in the cohort: premenopausal or perimenopausal status; history of complicated recurrent urinary tract infection (rUTI); antibiotic use within four weeks prior to urine sample collection, except in cases of culture-confirmed active infection; pelvic cancer or pelvic radiation within the last three years; current chemotherapy treatment; renal insufficiency (creatinine levels >1.5 mg/dL); post-void residual volume exceeding 100 mL; stage 3 or higher pelvic organ prolapse; pelvic surgery for incontinence within six months before urine sample collection; intermittent catheter use; neurogenic bladder; any upper urinary tract abnormalities associated with UTI; and diagnosis of Type 1 or Type 2 Diabetes Mellitus. One participant (PF21) from the rUTI History, UTI(+) group had an active, culture-confirmed UTI and had received antibiotics in the four weeks prior to sample collection. All urine samples were collected via "clean-catch" midstream methods after patient education.

Targeted Urinary Metabolite Quantification

The Biocrates Quant 500 targeted metabolomics assay with the urine extension was used to quantitatively measure urinary metabolites. Assay was performed per manufacturer instructions on a Waters Xevo TQS tandem quadrupole mass spectrometer lined to an Acquity liquid chromatography system. The LC-MS system used for metabolomics passed the assay's system suitability test immediately prior to data acquisition. Briefly, urine samples were slowly thawed at 4C and a 10µl aliquot was cleared of debris by centrifugation of 4000xG. The supernatant was transferred to the Biocrates Quant 500 plate, a step which incorporates a variety of heavy isotopically labeled spike-in standards for quantitation. Samples are then subjected to a phenyl isothiocyanate derivatization of primary amines to improve chromatographic separation. Derivatized products were then subjected to LC-MS and flow injection analysis (FIA). Analytes were quantified using the MetIDQ software supplied with the assay.

Taxonomic Profiling of Metagenomic Data

Raw taxonomic profile data, generated in the original report for this cohort (PRJNA801448), was reanalyzed with the updated MetaPhLAn 4.1 database. Decontamination of potential background taxonomic signatures was performed with Decontam and analysis of the taxonomic signature identified in a water blank sample that was sequenced to observe potential environmental contaminants inevitably introduced during metagenomic DNA extraction or library preparation.[64]

Analysis of Biochemical and Taxonomic Beta Diversity

Beta diversity distances were calculated for both the Biochemical and Taxonomic profiles via the Bray-Curtis, Jaccard, Euclidean, Canberra, Kulczynski, Gower, and Manhattan methods. We then performed a screen for significant correlation between all combinations of the beta diversity distance arrays of the biochemcial and taxonomic ecology. This analysis is primarily sensitive to monotonic associations that are reflected in the pairwise distance associations.[39] We then used Procrustes of PCoA ordinations analysis over all combinations of the Beta diversity distances calculations as a more sensitive analysis to identify associations between biochemical and taxonomic ecology. For all Procrustes analyses, metabolite beta diversity was used as the target while taxonomic beta diversity was rotated and scaled. PCoA was performed over 4 dimensions for both metabolite and taxonomic ecology. All ecological analysis was performed in the R statistical coding language using the vegan package.

Network Analysis of Metabolite-Taxon Association

All-against-all Spearman correlation analysis was performed using the metabolite concentration and taxonomic relative abundance in the 50 samples from the No UTI History and rUTI History, UTI (+) groups. Significant associations between taxa and metabolites were defined as those reaching a Spearman coefficient > |0.3| and *P*-value < 0.05. Network visualization and analysis was performed in Cytoscape (v 3.10.2)

Cross Validation Analysis of Potential Diagnostic Biomarkers

Elastic Net regularization combined with logistic regression to classify lipidomic data was performed to identify lipid features that distinguish No UTI History samples from rUTI History, UTI(+) samples. To optimize model performance, repeated cross-validated Elastic Net was used to explore a range of alpha values (from 0 to 1 in increments of 0.1) and corresponding lambda values, measuring performance with the area under the curve (AUC). The optimal model was selected based on the highest AUC, yielding the best alpha and lambda values, and the

corresponding feature coefficients were extracted to identify the most important features for classification. Leave-one-out cross-validation (LOOCV) was performed using the optimal model to predict class labels. For each observation, the model was trained on the remaining data, and probabilities of class membership were predicted for the left-out observation. The predicted probabilities were converted to class labels based on a threshold of 0.5. Model performance was assessed using confusion matrix statistics and ROC analysis, with the AUC value calculated for the LOOCV predictions.

Putative Prognostic Biomarker Analysis

To assess the association between individual metabolites and time to recurrence, a Cox proportional hazards regression analysis was performed for each metabolite in the dataset. The dataset was first loaded from a CSV file, and time to recurrence along with recurrence status were extracted as the response variables. The analysis focused on the relationship between each metabolite and the survival outcome, excluding missing values where applicable. A Cox proportional hazards model was fitted for each metabolite using the `Surv()` function, which models time to recurrence and recurrence status.

Kaplan-Meier survival analysis was performed to further explore the association between selected metabolites and time to recurrence. Significant metabolites of interest identified by the COX proportional hazard analysis were used, and for each metabolite, the data was dichotomized based on the median concentration. For each metabolite, values above the median were classified as "High" and values below or equal to the median were classified as "Low." The Kaplan-Meier estimator was then applied to assess differences in survival between the high and low groups. The survival curves were generated using the `survfit()` function, and the resulting Kaplan-Meier plots were visualized with 95% confidence intervals.

Statistical Analysis

All statistical analysis was performed using the R statistical programing language (V4.4.0). GraphPad Prism 10 was used for plotting graphs and Microsoft Excel were used for data interrogation and filtering. Pairwise hypothesis testing was performed using the non-parametric Wilcoxon Rank Sum test while the Kruskal-Wallis nonparametric ANOVA with Dunn's multiple comparison post-hoc was performed for analysis of 3 or more groups. Adjustment for multiple hypothesis testing was performed using the false discovery rate (FDR) when appropriate.

Figure Legends

Figure 1 Targeted LC-MS/MS of the urinary polar metabolome and lipidome. (A) Overview of metabolite classes detected by metabolomic assay. Y-axis refers to the proportion of the cohort where a given metabolite was detected above the limit of detection (LOD). (B) Polar metabolites retained after filtering. (C) Non-polar metabolites retained after filtering. (D) Correlation of raw and creatinine normalized metabolite abundances. *P*-value generated by permutation.

Figure 2 Urinary biochemical ecology is moderately associated with the taxonomic ecology. (A) Distribution of α diversity (Inverse Simpson Hill Number) among the three cohort groups (*n*=75 total, *n*=25 per group). P-value generate by Kruskal-Wallis nonparametric ANOVA. (B) Comparison of the distribution of biochemical (left) and taxonomic (right) α diversity (Inverse Simpson Hill Number). Violin is drawn from minimum to maximum of the distribution. Median is denoted by a line and boxes are drawn to represent the interquartile range. (C) Spearman correlation of biochemical and taxonomic α diversity. P-value generated by permutation. 95% confidence bands border the linear regression line. (D-F) Spearman correlation of taxonomic α diversity with the urinary concentration of individual metabolites, methionine sulfoxide (Met-SO) (D), citrulline (E), and trimethylamine oxide (F) by both taxonomic Shannon and Simpson metrics. *P*-value generated by permutation. 95% confidence bands border the linear regression line. (G) Mantel Test screening of the association between biochemical and taxonomic β diversity combinations. (H) Analysis of pairwise Canberra β diversity distances within and between cohort groups. Boxes draw to represent the median and interquartile range. Whiskers drawn to the minima and maxima of the data distribution. (I) Procrustes errors from analysis of the association between biochemical and taxonomic β diversity. For Procrustes analysis, PCoA ordination of the biochemical b diversity was used as the target, while taxonomic b diversity ordination was rotated and scaled. *P*-value generated by permutation Protest.

Figure 3. Taxa metabolite network analysis. (A) Network representation of significant ($R > 0.3$, $P < 0.05$) positive taxon (red nodes) metabolite (blue nodes) associations. (B) Network representation of significant ($R < -0.3$, $P < 0.05$) negative taxon (red nodes) metabolite (blue nodes) associations. Edges represent the Spearman Correlation between metabolite concentrations and mCLR transformed taxon relative abundance. Taxon node size is proportional to the number of first neighbor metabolites. (C) Network metrics, nodal degree and topological coefficient, of taxa in the positive association network. (D) Metabolite nodal degree and topological coefficient in the positive association network. (E) Taxon nodal degree and

topological coefficient in the negative association network. (F) Nodal degree and topological coefficient of metabolites in the negative association network. Network representation of significant (G) positive (R > 0.3, P < 0.05) and negative (H) species level *Lactobacillus* (red nodes) metabolite (blue nodes) positive associations. Edges represent the Spearman Correlation between metabolite concentrations and mCLR transformed taxon relative abundance. Taxon node size is proportional to the number of first neighbor metabolites.

Figure 4. Active UTI is associated with distinct urine metabolic signature. (A) Differential enrichment analysis of urinary metabolites between the No UTI History (Blue) and rUTI History, UTI(+) (Red) groups. *P*-value generated by the Wilcoxon Rank Sum test. *Q*-value generated by false discovery rate. Absolute value of the Wilcoxon Rank statistic of the top 10 metabolites differentially enriched in the (B) rUTI History, UTI (+) group and in the (C) No UTI History group. (D-G) Comparison of the distributions for individual metabolites differentially enriched in the rUTI History, UTI (+) group. (H-K) Comparison of the distributions for individual metabolites differentially enriched in No UTI History group. *P*-value generated by the Wilcoxon Rank Sum test. *Q*-value generated by false discovery rate.

Figure 5. Taxa associated with urinary metabolites differentially enriched in women with No UTI history. Spearman correlation analysis between urinary metabolites differentially enriched in the No UTI History cohort and the urinary microbiome taxonomic profile at the (A) genus and (B) species level. (C-F) Scatter plots of significant taxa-metabolite correlations. *P*-value generated by permutation.

Figure 6. Urinary lipids as biomarkers for active rUTI. (A) Comparison of the distribution of summed lipids between the No UTI History (n=25) and rUTI History, UTI(+) (n=25) groups. *P*-value generated by the Wilcoxon Rank Sum test. (B) Importance of lipid features to a discriminating model identified by elastic net regularization. (C) Logistic regression performed to assess the ability of the 7 lipids to discriminate case from control samples. *P*-value generated by log-rank test. (D) Receiver operator curve assessing the performance of the 7-lipid model in discriminating UTI from control samples. (E) Spearman correlation analysis between urinary taxa and the 7 lipids within the discriminating model. (F) Scatter plot of the correlation between diacylglycerol (14:0, 20:0) and mCLR transformed *Escherichia coli* abundance. *P*-value generated by permutation.

Figure 7. The Urinary Metabolome Harbors Translatable Prognostic Indicators of rUTI. (A) Cox proportional hazard analysis performed on metabolites on the 19 samples within the rUTI History, UTI(+) cohort that have time to recurrence metadata. (B) Hazard ratio of the 12 metabolites identified with significant association with increased hazard to recurrence. (C) DCA dichotomization scheme for Kaplan-Meier analysis. (D) Kaplan-Meier analysis of recurrence free probability and the urinary concentration of deoxycholic acid (DCA). Patients are dichotomized by the median DCA concentration. *P*-value generated by the Gehan-Breslow-Wilcoxon test. (E) Comparison of the number of Recurrence Free Survival Days between DCA low (grey) and DCA high (red) samples. *P*-value generated by Wilcoxon Rank Sum test.

Acknowledgments

We wish to acknowledge and sincerely thank the patients who participated in this study. This work was supported by grants from the Foundation for Women's Wellness (FWW), and the National Institutes of Health (1R01DK131267-01) to N.J.D., and (1F32DK128975-01) to M.L.N. This work was also supported by the Felicia and John Cain Distinguished Chair in Women's Health to P.E.Z. We would like to thank all the members of the De Nisco and Shulaev labs for their creative input throughout this study. We would further like to acknowledge and thank the Genome Center at the University of Texas at Dallas for their invaluable work and assistance in generating the metagenomic dataset used for this study.

Author Contributions

Conceptualization, M.L.N., V.Z., and N.J.D.; data curation, M.L.N., N.V.H. D.G. ; formal analysis, M.L.N; funding acquisition, V.S., P.E.Z., and N.J.D.; investigation, M.L.N., N.V.H., V.S.; methodology, M.L.N., N.V.H, P.E.Z., V.S, and N.J.D.; project administration, M.L.N., P.E.Z., and N.J.D.; resources, P.E.Z., V.S. and N.J.D.; software, M.L.N.; supervision, V.S., and N.J.D.; validation, M.L.N., N.V.H.; visualization, M.L.N., and N.J.D.; writing – original draft, M.L.N., and N.J.D.

References

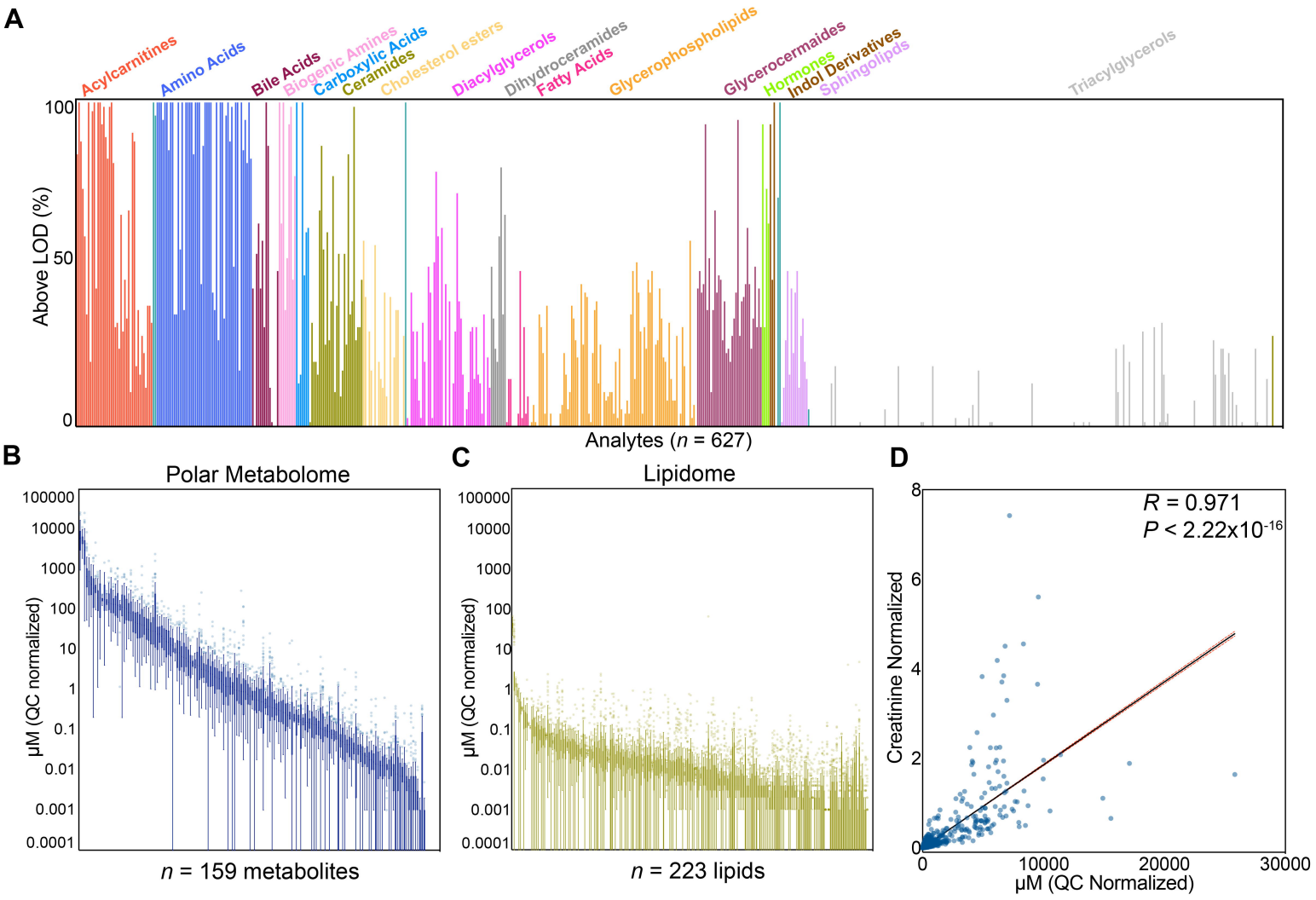
1 Gaitonde, S., Malik, R. D. & Zimmern, P. E. Financial Burden of Recurrent Urinary Tract Infections in Women: A Time-driven Activity-based Cost Analysis. *Urology* **128**, 47-54, doi:10.1016/j.urology.2019.01.031 (2019).

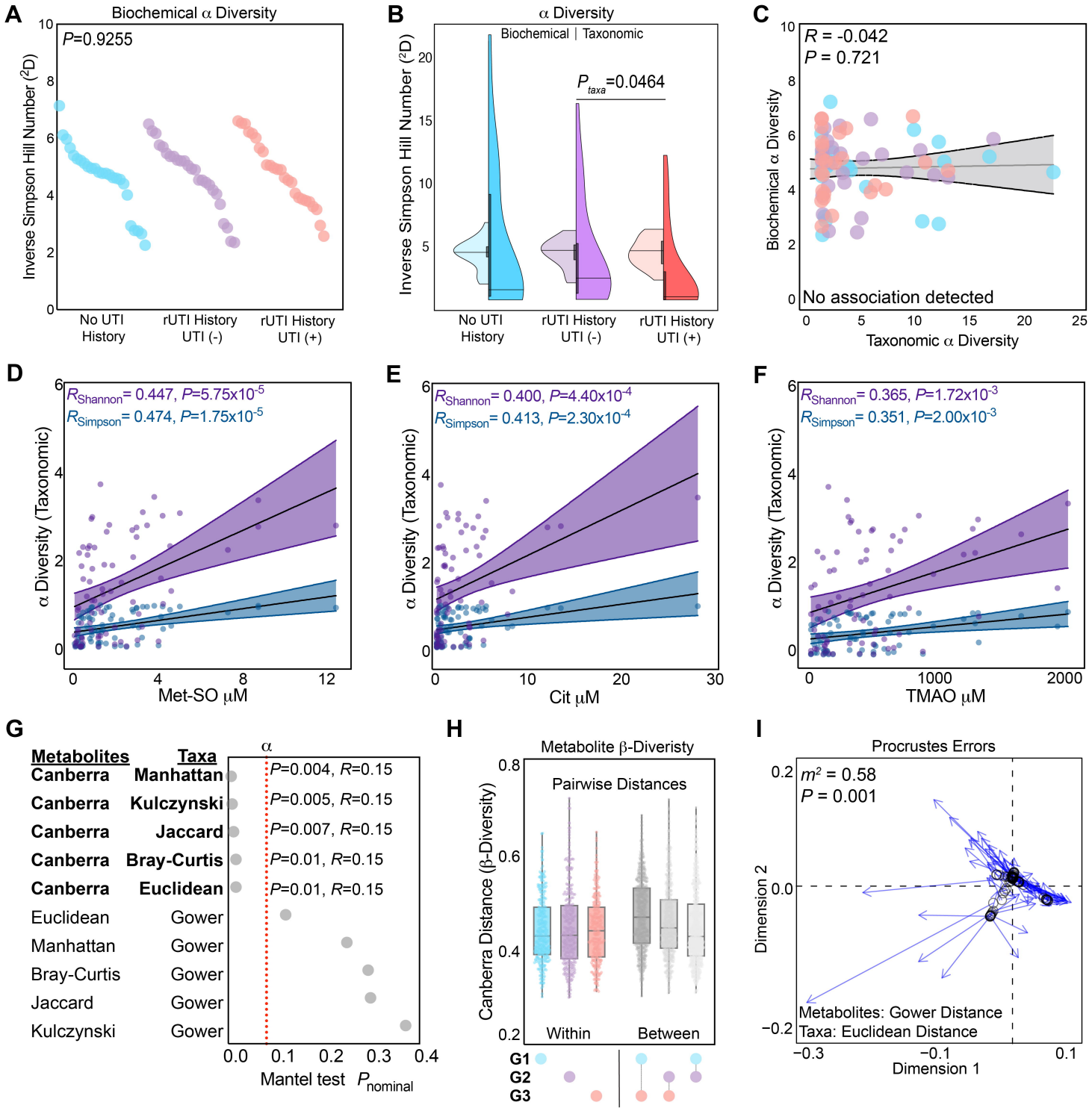
2 Jhang, J. F. & Kuo, H. C. Recent advances in recurrent urinary tract infection from pathogenesis and biomarkers to prevention. *Ci Ji Yi Xue Za Zhi* **29**, 131-137, doi:10.4103/tcmj.tcmj_53_17 (2017).
3 Neugent, M. L. *et al.* Urinary Glycosaminoglycans Are Associated with Recurrent UTI and Urobiome Ecology in Postmenopausal Women. *ACS Infect Dis*, doi:10.1021/acsinfecdis.3c00027 (2023).
4 Neugent, M. L., Hulyalkar, N. V., Nguyen, V. H., Zimmern, P. E. & De Nisco, N. J. Advances in Understanding the Human Urinary Microbiome and Its Potential Role in Urinary Tract Infection. *mBio* **11**, doi:10.1128/mBio.00218-20 (2020).
5 Neugent, M. L. *et al.* Recurrent urinary tract infection and estrogen shape the taxonomic ecology and function of the postmenopausal urogenital microbiome. *Cell Rep Med*, 100753, doi:10.1016/j.xcrm.2022.100753 (2022).
6 Thomas-White, K. *et al.* Vaginal estrogen therapy is associated with increased Lactobacillus in the urine of postmenopausal women with overactive bladder symptoms. *Am J Obstet Gynecol* **223**, 727 e721-727 e711, doi:10.1016/j.ajog.2020.08.006 (2020).
7 Flores-Mireles, A. L., Walker, J. N., Caparon, M. & Hultgren, S. J. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol* **13**, 269-284, doi:10.1038/nrmicro3432 (2015).
8 Harding, C. *et al.* Alternative to prophylactic antibiotics for the treatment of recurrent urinary tract infections in women: multicentre, open label, randomised, non-inferiority trial. *BMJ* **376**, e068229, doi:10.1136/bmj-2021-0068229 (2022).
9 Wolfe, A. J. & Brubaker, L. Urobiome updates: advances in urinary microbiome research. *Nat Rev Urol* **16**, 73-74, doi:10.1038/s41585-018-0127-5 (2019).
10 Wolfe, A. J. *et al.* Evidence of uncultivated bacteria in the adult female bladder. *J Clin Microbiol* **50**, 1376-1383, doi:10.1128/JCM.05852-11 (2012).
11 Hilt, E. E. *et al.* Urine is not sterile: use of enhanced urine culture techniques to detect resident bacterial flora in the adult female bladder. *J Clin Microbiol* **52**, 871-876, doi:10.1128/JCM.02876-13 (2014).
12 Brubaker, L. & Wolfe, A. J. The female urinary microbiota, urinary health and common urinary disorders. *Ann Transl Med* **5**, 34, doi:10.21037/atm.2016.11.62 (2017).
13 Siddiqui, H., Nederbragt, A. J., Lagesen, K., Jeansson, S. L. & Jakobsen, K. S. Assessing diversity of the female urine microbiota by high throughput sequencing of 16S rDNA amplicons. *BMC Microbiol* **11**, 244, doi:10.1186/1471-2180-11-244 (2011).
14 Price, T. K. *et al.* The Clinical Urine Culture: Enhanced Techniques Improve Detection of Clinically Relevant Microorganisms. *J Clin Microbiol* **54**, 1216-1222, doi:10.1128/JCM.00044-16 (2016).
15 Lewis, D. A. *et al.* The human urinary microbiome; bacterial DNA in voided urine of asymptomatic adults. *Front Cell Infect Microbiol* **3**, 41, doi:10.3389/fcimb.2013.00041 (2013).
16 Price, T. K. *et al.* The urobiome of continent adult women: a cross-sectional study. *BJOG*, doi:10.1111/1471-0528.15920 (2019).
17 Khasriya, R. *et al.* Spectrum of bacterial colonization associated with urothelial cells from patients with chronic lower urinary tract symptoms. *J Clin Microbiol* **51**, 2054-2062, doi:10.1128/JCM.03314-12 (2013).
18 Fouts, D. E. *et al.* Integrated next-generation sequencing of 16S rDNA and metaproteomics differentiate the healthy urine microbiome from asymptomatic bacteriuria in neuropathic bladder associated with spinal cord injury. *J Transl Med* **10**, 174, doi:10.1186/1479-5876-10-174 (2012).
19 Bucevic Popovic, V. *et al.* The urinary microbiome associated with bladder cancer. *Sci Rep* **8**, 12157, doi:10.1038/s41598-018-29054-w (2018).

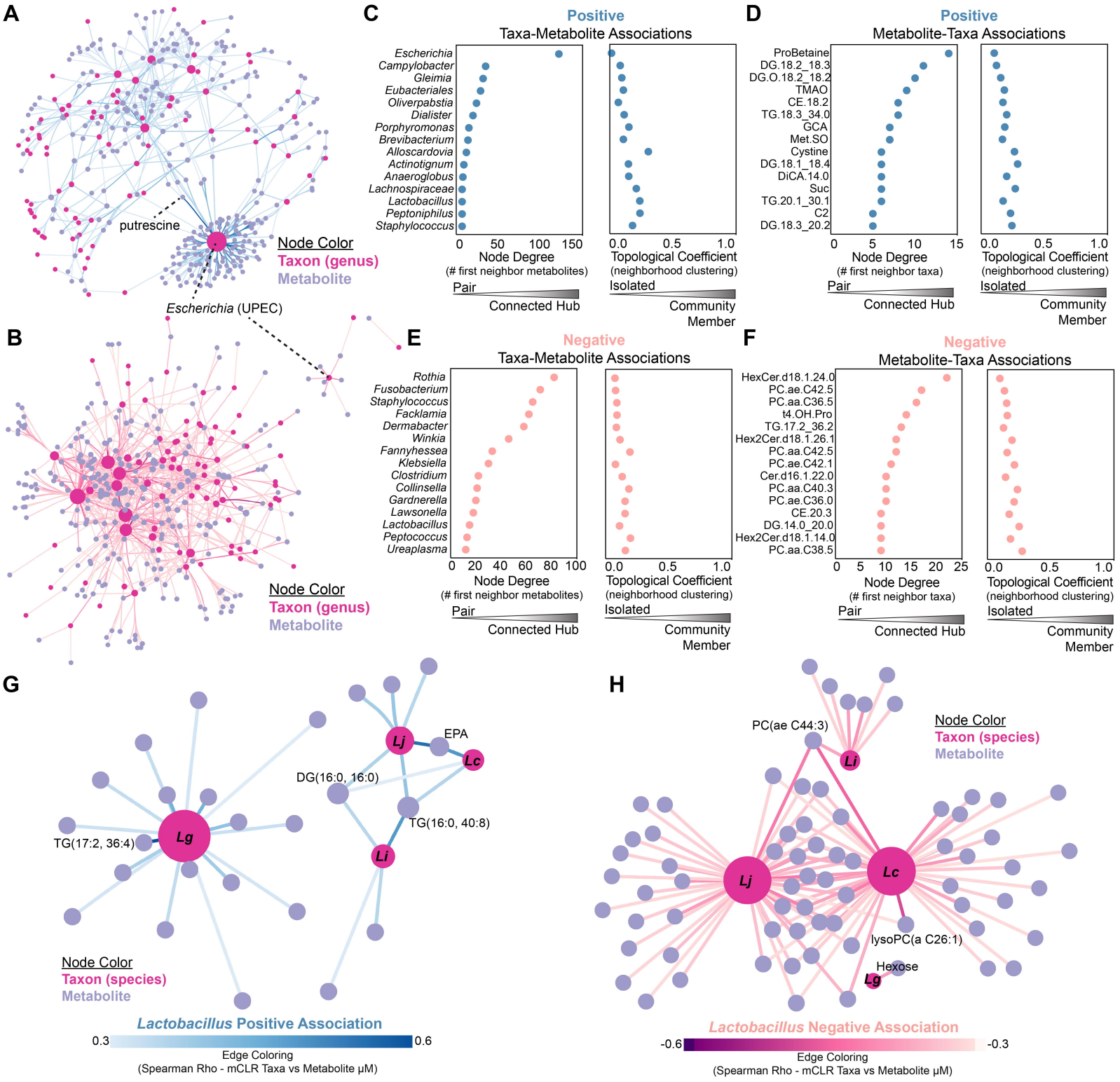
20 Karstens, L. *et al.* Does the Urinary Microbiome Play a Role in Urgency Urinary Incontinence and Its Severity? *Front Cell Infect Microbiol* **6**, 78, doi:10.3389/fcimb.2016.00078 (2016).
21 Reasoner, S. A. *et al.* Survey of the infant male urobiome and genomic analysis of Actinotignum spp. *NPJ Biofilms Microbiomes* **9**, 91, doi:10.1038/s41522-023-00457-6 (2023).
22 Kelly, M. S., Dahl, E. M., Jeries, L. M., Sysoeva, T. A. & Karstens, L. Characterization of pediatric urinary microbiome at species-level resolution indicates variation due to sex, age, and urologic history. *J Pediatr Urol* **20**, 884-893, doi:10.1016/j.jpurol.2024.05.016 (2024).
23 Vaughan, M. H. *et al.* The Urinary Microbiome in Postmenopausal Women with Recurrent Urinary Tract Infections. *J Urol* **206**, 1222-1231, doi:10.1097/JU.0000000000001940 (2021).
24 Miller, I. J. *et al.* Real-time health monitoring through urine metabolomics. *NPJ Digit Med* **2**, 109, doi:10.1038/s41746-019-0185-y (2019).
25 Thevenot, E. A., Roux, A., Xu, Y., Ezan, E. & Junot, C. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *J Proteome Res* **14**, 3322-3335, doi:10.1021/acs.jproteome.5b00354 (2015).
26 Bouatra, S. *et al.* The human urine metabolome. *PLoS One* **8**, e73076, doi:10.1371/journal.pone.0073076 (2013).
27 Puebla-Barragan, S., Renaud, J., Sumarah, M. & Reid, G. Malodorous biogenic amines in Escherichia coli-caused urinary tract infections in women-a metabolomics approach. *Sci Rep* **10**, 9703, doi:10.1038/s41598-020-66662-x (2020).
28 Milshteyn, A., Colosimo, D. A. & Brady, S. F. Accessing Bioactive Natural Products from the Human Microbiome. *Cell Host Microbe* **23**, 725-736, doi:10.1016/j.chom.2018.05.013 (2018).
29 Mayers, J. R. *et al.* A metabolomics pipeline highlights microbial metabolism in bloodstream infections. *Cell* **187**, 4095-4112 e4021, doi:10.1016/j.cell.2024.05.035 (2024).
30 Bloom, S. M. *et al.* Cysteine dependence of Lactobacillus iners is a potential therapeutic target for vaginal microbiota modulation. *Nat Microbiol* **7**, 434-450, doi:10.1038/s41564-022-01070-7 (2022).
31 Griffiths, W. J. *et al.* Targeted metabolomics for biomarker discovery. *Angew Chem Int Ed Engl* **49**, 5426-5445, doi:10.1002/anie.200905579 (2010).
32 Dudley, E., Yousef, M., Wang, Y. & Griffiths, W. J. Targeted metabolomics and mass spectrometry. *Adv Protein Chem Struct Biol* **80**, 45-83, doi:10.1016/B978-0-12-381264-3.00002-3 (2010).
33 Tang, K. W., Toh, Q. C. & Teo, B. W. Normalisation of urinary biomarkers to creatinine for clinical practice and research--when and why. *Singapore Med J* **56**, 7-10, doi:10.11622/smedj.2015003 (2015).
34 Ebrahimzadeh, T. *et al.* Urinary prostaglandin E2 as a biomarker for recurrent UTI in postmenopausal women. *Life Sci Alliance* **4**, doi:10.26508/lsa.202000948 (2021).
35 Schiattarella, G. G. *et al.* Gut microbe-generated metabolite trimethylamine-N-oxide as cardiovascular risk biomarker: a systematic review and dose-response meta-analysis. *Eur Heart J* **38**, 2948-2956, doi:10.1093/eurheartj/ehx342 (2017).
36 Bray, J. R. & Curtis, J. T. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* **27**, 325-349, doi:https://doi.org/10.2307/1942268 (1957).
37 Anderson, M. J., Ellingsen, K. E. & McArdle, B. H. Multivariate dispersion as a measure of beta diversity. *Ecol Lett* **9**, 683-693, doi:10.1111/j.1461-0248.2006.00926.x (2006).

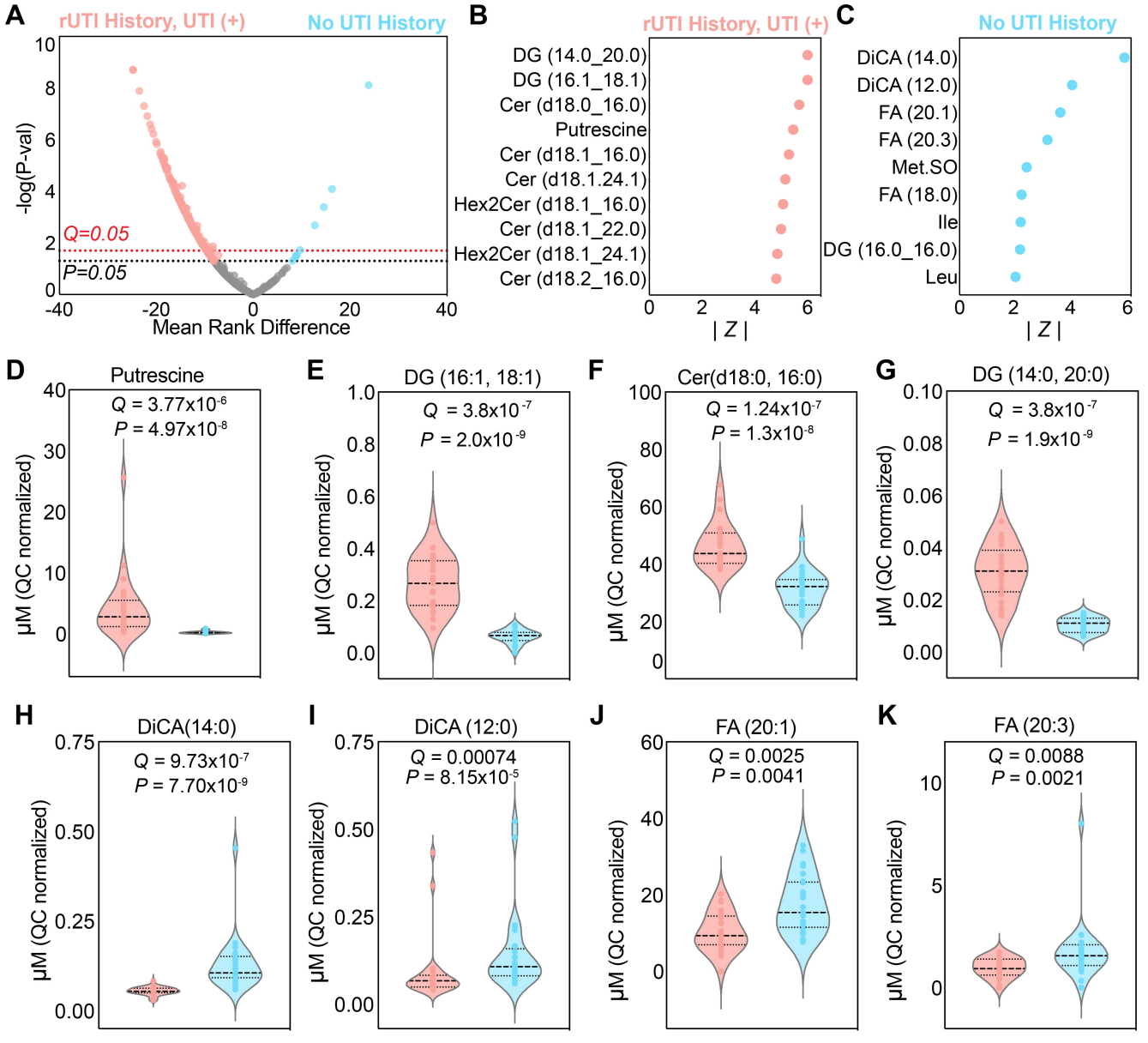
38 Legendre, P. & De Caceres, M. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecol Lett* **16**, 951-963, doi:10.1111/ele.12141 (2013).
39 Mantel, N. The detection of disease clustering and a generalized regression approach. *Cancer Res* **27**, 209-220 (1967).
40 Lisboa, F. J. *et al.* Much beyond Mantel: bringing Procrustes association metric to the plant and soil ecologist's toolbox. *PLoS One* **9**, e101238, doi:10.1371/journal.pone.0101238 (2014).
41 Schneider, B. L., Hernandez, V. J. & Reitzer, L. Putrescine catabolism is a metabolic response to several stresses in Escherichia coli. *Mol Microbiol* **88**, 537-550, doi:10.1111/mmi.12207 (2013).
42 He, Z., Li, P., Liu, P. & Xu, P. Exploring stachydrine: from natural occurrence to biological activities and metabolic pathways. *Front Plant Sci* **15**, 1442879, doi:10.3389/fpls.2024.1442879 (2024).
43 Johnson, J. A. *et al.* Commensal Urinary Lactobacilli Inhibit Major Uropathogens In Vitro With Heterogeneity at Species and Strain Level. *Front Cell Infect Microbiol* **12**, 870603, doi:10.3389/fcimb.2022.870603 (2022).
44 Nagy, G. & Pohl, N. L. Complete hexose isomer identification with mass spectrometry. *J Am Soc Mass Spectrom* **26**, 677-685, doi:10.1007/s13361-014-1072-z (2015).
45 Ranea-Robles, P. & Houten, S. M. The biochemistry and physiology of long-chain dicarboxylic acid metabolism. *Biochem J* **480**, 607-627, doi:10.1042/BCJ20230041 (2023).
46 Blanco-Miguez, A. *et al.* Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nat Biotechnol* **41**, 1633-1644, doi:10.1038/s41587-023-01688-w (2023).
47 Kasprzykowska, U., Elias, J., Elias, M., Maczynska, B. & Sobieszczanska, B. M. Colonization of the lower urogenital tract with Ureaplasma parvum can cause asymptomatic infection of the upper reproductive system in women: a preliminary study. *Arch Gynecol Obstet* **289**, 1129-1134, doi:10.1007/s00404-013-3102-7 (2014).
48 Mulvey, M. A. *et al.* Induction and evasion of host defenses by type 1-piliated uropathogenic Escherichia coli. *Science* **282**, 1494-1497, doi:10.1126/science.282.5393.1494 (1998).
49 Ebrahimzadeh, T. *et al.* Inflammatory markers for improved recurrent UTI diagnosis in postmenopausal women. *Life Sci Alliance* **7**, doi:10.26508/lsa.202302323 (2024).
50 Bauermeister, A., Mannochio-Russo, H., Costa-Lotufo, L. V., Jarmusch, A. K. & Dorrestein, P. C. Mass spectrometry-based metabolomics in microbiome investigations. *Nat Rev Microbiol* **20**, 143-160, doi:10.1038/s41579-021-00621-9 (2022).
51 Zuffa, S. *et al.* microbeMASST: a taxonomically informed mass spectrometry search tool for microbial metabolomics data. *Nat Microbiol* **9**, 336-345, doi:10.1038/s41564-023-01575-9 (2024).
52 Lee, B. C. & Gladyshev, V. N. The biological significance of methionine sulfoxide stereochemistry. *Free Radic Biol Med* **50**, 221-227, doi:10.1016/j.freeradbiomed.2010.11.008 (2011).
53 Moskovitz, J. & Smith, A. Methionine sulfoxide and the methionine sulfoxide reductase system as modulators of signal transduction pathways: a review. *Amino Acids* **53**, 1011-1020, doi:10.1007/s00726-021-03020-9 (2021).
54 Xia, Y. & Sun, J. in *Bioinformatic and Statistical Analysis of Microbiome Data: From Raw Sequences to Advanced Modeling with QIIME 2 and R* 335-395 (Springer International Publishing, 2023).
55 Watanabe, T. *et al.* The Urinary Bladder is Rich in Glycosphingolipids Composed of Phytoceramides. *J Lipid Res* **63**, 100303, doi:10.1016/j.jlr.2022.100303 (2022).

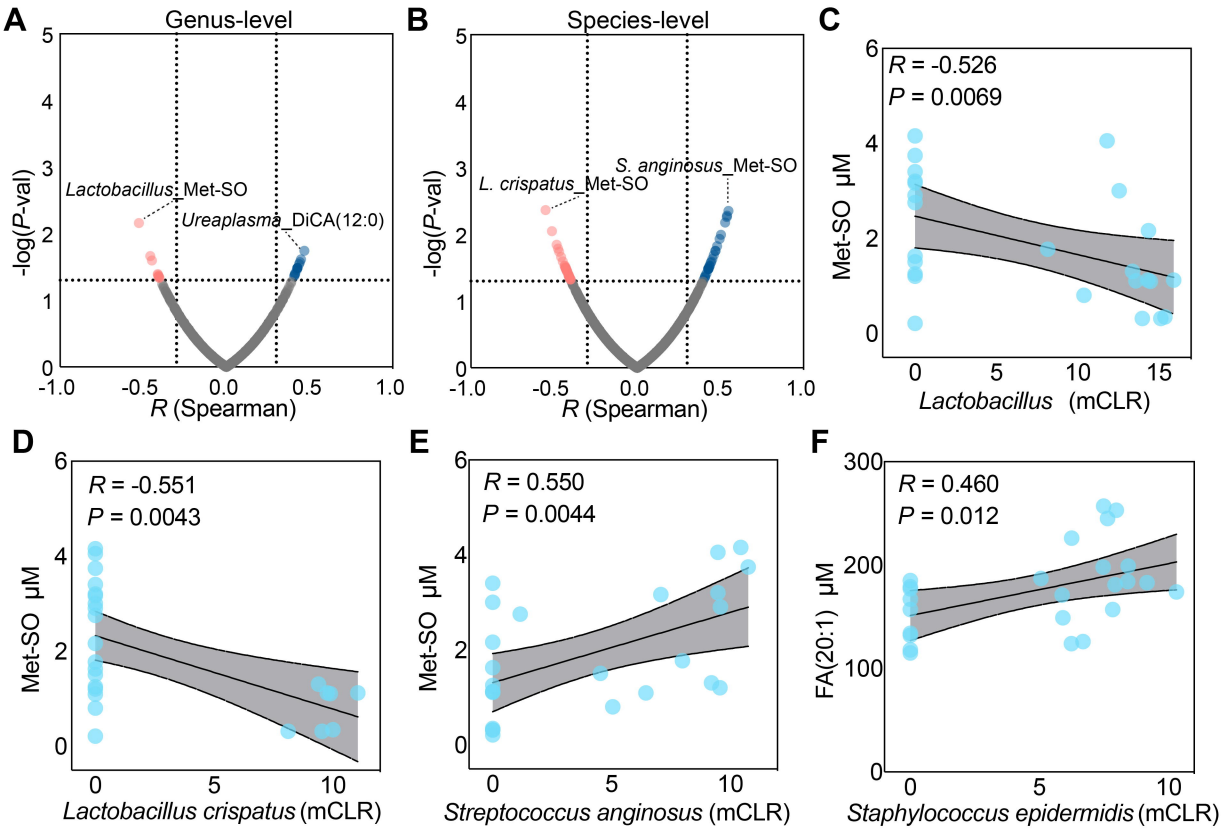
56 Stankeviciute, G. *et al.* Convergent evolution of bacterial ceramide synthesis. *Nat Chem Biol* **18**, 305-312, doi:10.1038/s41589-021-00948-7 (2022).
57 Smith, R. D. *et al.* A Novel Lipid-Based MALDI-TOF Assay for the Rapid Detection of Colistin-Resistant Enterobacter Species. *Microbiol Spectr* **10**, e0144521, doi:10.1128/spectrum.01445-21 (2022).
58 Murphy, R. C., Hankin, J. A. & Barkley, R. M. Imaging of lipid species by MALDI mass spectrometry. *J Lipid Res* **50 Suppl**, S317-322, doi:10.1194/jlr.R800051-JLR200 (2009).
59 Prasad, A. *et al.* Comparative Genomic Study of Streptococcus anginosus Reveals Distinct Group of Urinary Strains. *mSphere* **8**, e0068722, doi:10.1128/msphere.00687-22 (2023).
60 Mohanty, I. *et al.* The underappreciated diversity of bile acid modifications. *Cell* **187**, 1801-1818 e1820, doi:10.1016/j.cell.2024.02.019 (2024).
61 Chong Nguyen, C. *et al.* Circulating bile acids concentration is predictive of coronary artery disease in human. *Sci Rep* **11**, 22661, doi:10.1038/s41598-021-02144-y (2021).
62 Worby, C. J. *et al.* Longitudinal multi-omics analyses link gut microbiome dysbiosis with recurrent urinary tract infections in women. *Nat Microbiol* **7**, 630-639, doi:10.1038/s41564-022-01107-x (2022).
63 Thanert, R. *et al.* Persisting uropathogenic Escherichia coli lineages show signatures of niche-specific within-host adaptation mediated by mobile genetic elements. *Cell Host Microbe* **30**, 1034-1047 e1036, doi:10.1016/j.chom.2022.04.008 (2022).
64 Salter, S. J. *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* **12**, 87, doi:10.1186/s12915-014-0087-z (2014).

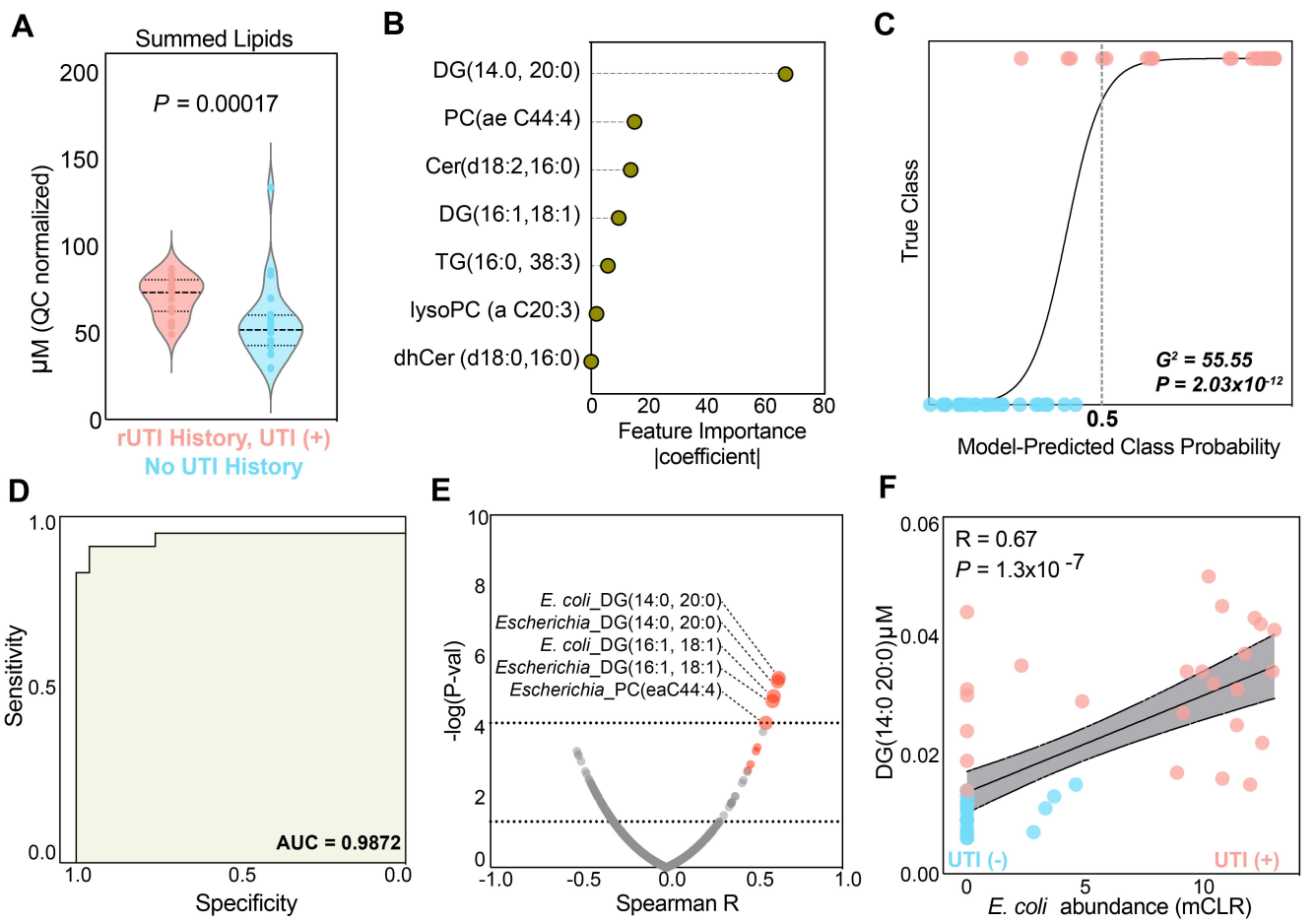

A
Acylcarnitines
Amino Acids
Bile Acids
Biogenic Amines
Carboxylic Acids
Ceramides
Cholesterol esters
Diacylglycerols
Dihydroceramides
Fatty Acids
Glycerophospholipids
Glycerocermaides
Hormones
Indol Derivatives
Sphingolipids
Triacylglycerols
Above LOD (%)
100
50
0
Analytes (n = 627)
B
Polar Metabolome
μM (QC normalized)
100000
10000
1000
100
10
1
0.1
0.01
0.001
0.0001
n = 159 metabolites
C
Lipidome
μM (QC normalized)
100000
10000
1000
100
10
1
0.1
0.01
0.001
0.0001
n = 223 lipids
D
R = 0.971
P < 2.22x10-16
Creatinine Normalized
8
6
4
2
0
0
10000
20000
30000
μM (QC Normalized)

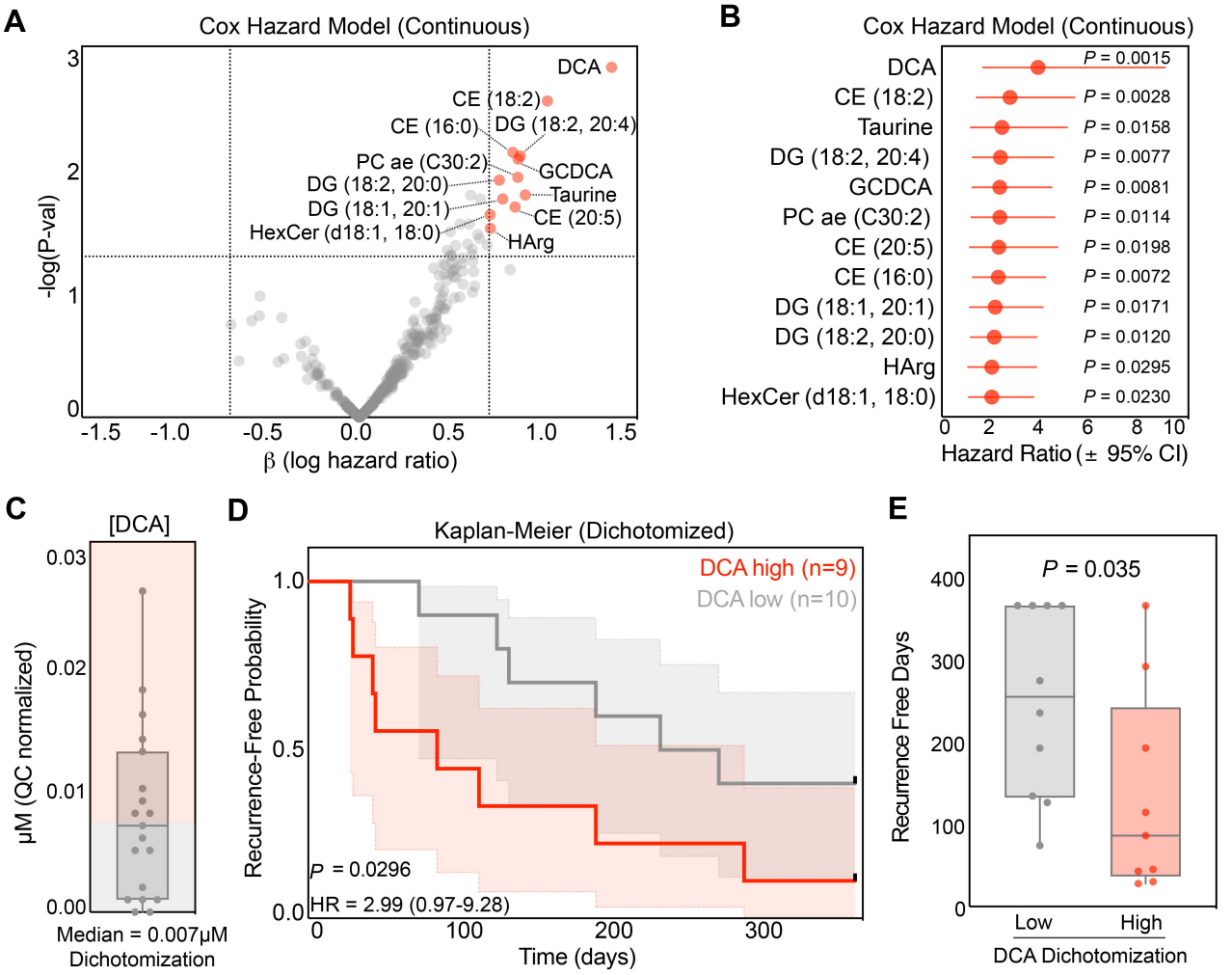

A
Biochemical α Diversity
P=0.9255
Inverse Simpson Hill Number (²D)
No UTI History
rUTI History UTI (-)
rUTI History UTI (+)
B
α Diversity
Biochemical | Taxonomic
P taxa=0.0464
Inverse Simpson Hill Number (²D)
No UTI History
rUTI History UTI (-)
rUTI History UTI (+)
C
R = -0.042
P = 0.721
Biochemical α Diversity
No association detected
Taxonomic α Diversity
D
RShannon= 0.447, P=5.75x10-5
RSimpson= 0.474, P=1.75x10-5
α Diversity (Taxonomic)
Met-SO μM
E
RShannon= 0.400, P=4.40x10-4
RSimpson= 0.413, P=2.30x10-4
α Diversity (Taxonomic)
Cit μM
F
RShannon= 0.365, P=1.72x10-3
RSimpson= 0.351, P=2.00x10-3
α Diversity (Taxonomic)
TMAO μM
G
Metabolites
Taxa
α
Canberra Manhattan P=0.004, R=0.15
Canberra Kulczynski P=0.005, R=0.15
Canberra Jaccard P=0.007, R=0.15
Canberra Bray-Curtis P=0.01, R=0.15
Canberra Euclidean P=0.01, R=0.15
Euclidean Gower
Manhattan Gower
Bray-Curtis Gower
Jaccard Gower
Kulczynski Gower
Mantel test Pnominal
H
Metabolite β-Diveristy
Pairwise Distances
Canberra Distance (β-Diversity)
Within
Between
G1
G2
G3
I
Procrustes Errors
m² = 0.58
P = 0.001
Dimension 2
Metabolites: Gower Distance
Taxa: Euclidean Distance
Dimension 1


A
putrescine
Escherichia (UPEC)
Node Color
Taxon (genus)
Metabolite
B
Node Color
Taxon (genus)
Metabolite
C
Positive
Taxa-Metabolite Associations
Escherichia
Campylobacter
Gleimia
Eubacteriales
Oliverpabstia
Dialister
Porphyromonas
Brevibacterium
Alloscardovia
Actinotignum
Anaeroglobus
Lachnospiraceae
Lactobacillus
Peptoniphilus
Staphylococcus
0 50 100 150
Node Degree
(# first neighbor metabolites)
0.0 0.5 1.0
Topological Coefficient
(neighborhood clustering)
Pair
Connected Hub
Isolated
Community Member
D
Positive
Metabolite-Taxa Associations
ProBetaine
DG.18.2_18.3
DG.O.18.2_18.2
TMAO
CE.18.2
TG.18.3_34.0
GCA
Met.SO
Cystine
DG.18.1_18.4
DiCA.14.0
Suc
TG.20.1_30.1
C2
DG.18.3_20.2
0 5 10 15
Node Degree
(# first neighbor taxa)
0.0 0.5 1.0
Topological Coefficient
(neighborhood clustering)
Pair
Connected Hub
Isolated
Community Member
E
Negative
Taxa-Metabolite Associations
Rothia
Fusobacterium
Staphylococcus
Facklamia
Dermabacter
Winkia
Fannyhessea
Klebsiella
Clostridium
Collinsella
Gardnerella
Lawsonella
Lactobacillus
Peptococcus
Ureaplasma
0 20 40 60 80 100
Node Degree
(# first neighbor metabolites)
0.0 0.5 1.0
Topological Coefficient
(neighborhood clustering)
Pair
Connected Hub
Isolated
Community Member
F
Negative
Metabolite-Taxa Associations
HexCer.d18.1.24.0
PC.ae.C42.5
PC.aa.C36.5
t4.OH.Pro
TG.17.2_36.2
Hex2Cer.d18.1.26.1
PC.aa.C42.5
PC.ae.C42.1
Cer.d16.1.22.0
PC.aa.C40.3
PC.ae.C36.0
CE.20.3
DG.14.0_20.0
Hex2Cer.d18.1.14.0
PC.aa.C38.5
0 5 10 15 20 25
Node Degree
(# first neighbor taxa)
0.0 0.5 1.0
Topological Coefficient
(neighborhood clustering)
Pair
Connected Hub
Isolated
Community Member
G
Lg
Lj
Lc
Li
EPA
DG(16:0, 16:0)
TG(16:0, 40:8)
TG(17:2, 36:4)
Node Color
Taxon (species)
Metabolite
Lactobacillus Positive Association
0.3
0.6
Edge Coloring
(Spearman Rho - mCLR Taxa vs Metabolite μM)
H
PC(ae C44:3)
Li
Node Color
Taxon (species)
Metabolite
Lj
Lc
lysoPC(a C26:1)
Hexose
Lg
Lactobacillus Negative Association
-0.6
-0.3
Edge Coloring
(Spearman Rho - mCLR Taxa vs Metabolite μM)


A
rUTI History, UTI (+)
No UTI History
-log(P-val)
Q=0.05
P=0.05
Mean Rank Difference
B
rUTI History, UTI (+)
DG (14.0_20.0)
DG (16.1_18.1)
Cer (d18.0_16.0)
Putrescine
Cer (d18.1_16.0)
Cer (d18.1.24.1)
Hex2Cer (d18.1_16.0)
Cer (d18.1_22.0)
Hex2Cer (d18.1_24.1)
Cer (d18.2_16.0)
| Z |
C
No UTI History
DiCA (14.0)
DiCA (12.0)
FA (20.1)
FA (20.3)
Met.SO
FA (18.0)
Ile
DG (16.0_16.0)
Leu
| Z |
D
Putrescine
Q = 3.77x10-6
P = 4.97x10-8
μM (QC normalized)
E
DG (16:1, 18:1)
Q = 3.8x10-7
P = 2.0x10-9
μM (QC normalized)
F
Cer(d18:0, 16:0)
Q = 1.24x10-7
P = 1.3x10-8
μM (QC normalized)
G
DG (14:0, 20:0)
Q = 3.8x10-7
P = 1.9x10-9
μM (QC normalized)
H
DiCA(14:0)
Q = 9.73x10-7
P = 7.70x10-9
μM (QC normalized)
I
DiCA (12:0)
Q = 0.00074
P = 8.15x10-5
μM (QC normalized)
J
FA (20:1)
Q = 0.0025
P = 0.0041
μM (QC normalized)
K
FA (20:3)
Q = 0.0088
P = 0.0021
μM (QC normalized)


A
Genus-level
Lactobacillus_Met-SO
Ureaplasma_DiCA(12:0)
-log(P-val)
R (Spearman)
B
Species-level
L. crispatus_Met-SO
S. anginosus_Met-SO
-log(P-val)
R (Spearman)
C
R = -0.526
P = 0.0069
Met-SO μM
Lactobacillus (mCLR)
D
R = -0.551
P = 0.0043
Met-SO μM
Lactobacillus crispatus (mCLR)
E
R = 0.550
P = 0.0044
Met-SO μM
Streptococcus anginosus (mCLR)
F
R = 0.460
P = 0.012
FA(20:1) μM
Staphylococcus epidermidis (mCLR)


A
Summed Lipids
P = 0.00017
μM (QC normalized)
200
150
100
50
0
rUTI History, UTI (+)
No UTI History
B
DG(14.0, 20:0)
PC(ae C44:4)
Cer(d18:2,16:0)
DG(16:1,18:1)
TG(16:0, 38:3)
lysoPC (a C20:3)
dhCer (d18:0,16:0)
0
20
40
60
80
Feature Importance
|coefficient|
C
True Class
G² = 55.55
P = 2.03x10⁻¹²
0.5
Model-Predicted Class Probability
D
Sensitivity
1.0
0.5
0.0
AUC = 0.9872
1.0
0.5
0.0
Specificity
E
-log(P-val)
10
8
6
4
2
0
E. coli_DG(14:0, 20:0)
Escherichia_DG(14:0, 20:0)
E. coli_DG(16:1, 18:1)
Escherichia_DG(16:1, 18:1)
Escherichia_PC(eaC44:4)
-1.0
-0.5
0.0
0.5
1.0
Spearman R
F
R = 0.67
P = 1.3x10 -7
DG(14:0 20:0)μM
0.06
0.04
0.02
0.00
UTI (-)
UTI (+)
0
5
10
E. coli abundance (mCLR)


A
Cox Hazard Model (Continuous)
DCA
CE (18:2)
CE (16:0)
DG (18:2, 20:4)
PC ae (C30:2)
GCDCA
DG (18:2, 20:0)
Taurine
DG (18:1, 20:1)
CE (20:5)
HexCer (d18:1, 18:0)
HArg
-log(P-val)
β (log hazard ratio)
B
Cox Hazard Model (Continuous)
DCA P = 0.0015
CE (18:2) P = 0.0028
Taurine P = 0.0158
DG (18:2, 20:4) P = 0.0077
GCDCA P = 0.0081
PC ae (C30:2) P = 0.0114
CE (20:5) P = 0.0198
CE (16:0) P = 0.0072
DG (18:1, 20:1) P = 0.0171
DG (18:2, 20:0) P = 0.0120
HArg P = 0.0295
HexCer (d18:1, 18:0) P = 0.0230
Hazard Ratio (± 95% CI)
C
[DCA]
μM (QC normalized)
Median = 0.007μM
Dichotomization
D
Kaplan-Meier (Dichotomized)
DCA high (n=9)
DCA low (n=10)
Recurrence-Free Probability
P = 0.0296
HR = 2.99 (0.97-9.28)
Time (days)
E
P = 0.035
Recurrence Free Days
Low
High
DCA Dichotomization