The urinary tract microbiome in older women exhibits host genetics and environmental influences

AS Adebayo[1], G Ackermann[2], RC Bowyer[1], P Wells[1], G Humphreys[2], R Knight[2], TD Spector[1], CJ Steves[1]*

[1]Department of Twin Research & Genetic Epidemiology, King's College London, United Kingdom ; [2]The Knight Lab, Paediatrics, University of California San Diego, United States

*Corresponding author/Lead Contact: claire.j.steves@kcl.ac.uk

Summary

The urinary microbiome is a relatively unexplored niche despite the fact that we now know that it is not sterile. Moreover urinary microbes, especially in ageing populations, are associated with morbidity even when infection is subsequently not proven. We present the first large-scale study to explore factors defining urinary microbiome composition in community-dwelling older adult women without clinically active infection. Using 1600 twins, we estimate the contribution of genetic and environmental factors to variation in microbiome using both 16S and shotgun metagenomics. We found that the urinary microbiome is distinct from nearby sites and is unrelated to stool microbiome. Core urinary microbiome taxa were defined. The first component of weighted unifrac was heritable (18%) as were key taxa (e.g *Escherichia-Shigella* (A>0.15)). Age, menopausal status, prior UTI and host genetics were top among factors defining the urobiome. Increased composition was associated with older age, contrary to previous findings.

Keywords: microbiome, genetics, urogenital tract, ageing

Introduction

The resident microbial community (microbiome) at different human body sites, continues to generate research interest, driven by evidence of a role in human physiology. The study of the urinary microbiome (urobiome) is much less established compared to the gut microbiome; perhaps due to the previous belief that the urine was sterile in the absence of a urinary tract infection. Recently, research has shown that this is not the case and that the urinary tract is in fact, another site with a microbiome, reflective of the microbes inhabiting the bladder and closely associated organs (Wolfe et al., 2012; Siddiqui et al., 2012; Whiteside et al.; 2015). This evidence is supported

by enhanced quantitative cultures, 16S marker studies and metagenomics, in different populations (e.g Kramer et al., 2018; Adebayo et al., 2017;Wu et al., 2017).

Studies to date have identified differences in the urobiome in relation to urinary tract conditions (Sihra et al., 2018; Wolfe & Brubaker, 2019) including urinary infections (UTI). There is evidence for sex differences in the urinary microbiome which may in part be due to differences in the length of the urinary tract (Moustafa et al. 2018). Women are much more likely to develop UTI, with a lifetime risk of up to 50% (Franco, 2005). UTI is also the commonest reason for antibiotic treatment in adult women, which has implications for urinary and other microbiomes and antimicrobial resistance. Early work has indicated that the non-infected state microbiome may influence resilience to infection. Thus this paper is focused on understanding the major factors defining the urobiome in community dwelling women without active infection.

Recent studies involving urinary/bladder microbiomes have involved relatively small sample sizes (dozens or few hundreds of people) in hospital or clinic attending patients. For instance, results from our literature search (Jan 2015 to September 2018) included case-control studies on elderly/non-elderly patients (Liu et al., 2017; n=100); urinary tract infections (Moustafa et al., 2018; n=112), cancer (Wang et al., 2017; n=65), diabetes, overactive bladder (Wu et al., 2017; Fok et al., 2018,; n=55-126), chronic kidney disease (Kramer et al, 2018; n=41); surgical transplant patients (Rani et al., 2018, n=20); menopause (Curtiss et al., 2018; n=78). Reinforcing this, a recent review (covering studies up to 2016) carried out by Aragon et al. (2018) reported that the sample sizes in urinary microbiome studies varied between 8 to 60 for healthy controls and 10-197 for cases. Their report shows that many studies are commissioned on incontinence, bladder-related and gynaecologic patients. Moreover, many of the urine microbiome studies, either with 16S or shotgun metagenomes, exclude samples with non-detectable/ below detection microbiome. While the assumption maybe that the failure is completely technical, it is unknown if host factors contribute to having ‘extremely-low’ or ‘below detection’ urine microbiome.

Recently, studies of the gut microbiome, have shown a role of host genetics. While Goodrich and colleagues first reported clearly heritable components within the gut

microbiome (Goodrich et al.,2014), a finding which a few subsequent studies have also reiterated(Luca et al., 2018), Rothchilds et al reported that environmental factors may largely blur such host genetics factors (Rothchilds et al.,2018). It is unknown if genetic factors are important in the urinary microbiome.

We aimed to characterize the host influence on the urinary tract microbiome in women. Using midstream urine samples from 1600 females in the TwinsUK cohort, this study, perhaps the largest on urinary microbiome so far, reports about the urinary microbiome composition in an average female population of mainly postmenopausal women with no apparent infection. We hypothesized that, in an unselected average population, (1) the inherent core urinary bacterial community could be defined (2) that the urobiome is influenced by host-specific genetic and environmental factors, (3) that some host-specific factors may relate to undetectable microbial biomass in the urine.

Results

Urinary microbiome across studies and were distinct from proximal body sites and shared key taxa

Initially, we compared the overall composition of the urinary microbiome to similar datasets from other body sites using the same bioinformatics pipeline, using similar sized datasets of women aged >45(Supplementary Methods & Data1). Alpha diversity in the urine was, on average, reduced relative to the stool and is comparable in two urine and the vaginal datasets (Fig 1A). Stool samples in the majority ordinated separately from urine samples (Fig 1B) (Supplementary Data 1). Repeating these diversity analyses with a separate set of random 100 samples each show similar patterns and significance (SFig1A,B). In paired-sample analysis from TwinsUK (Supplementary Data1), urine microbial taxa separated from stool microbial taxa of the same individual (S1C). There was no clear correlation in the pattern of stool and urine microbiome dissimilarity for the paired samples (either obtained at same time point or not) (Mantel's $r \leq 0.02$, $p > 0.1$) and variance was not homogeneous (Levene paired test $p=0.02$) (Fig 1C-D, SData1). Thereafter, we examined the TwinsUK urinary microbiome dataset alone.

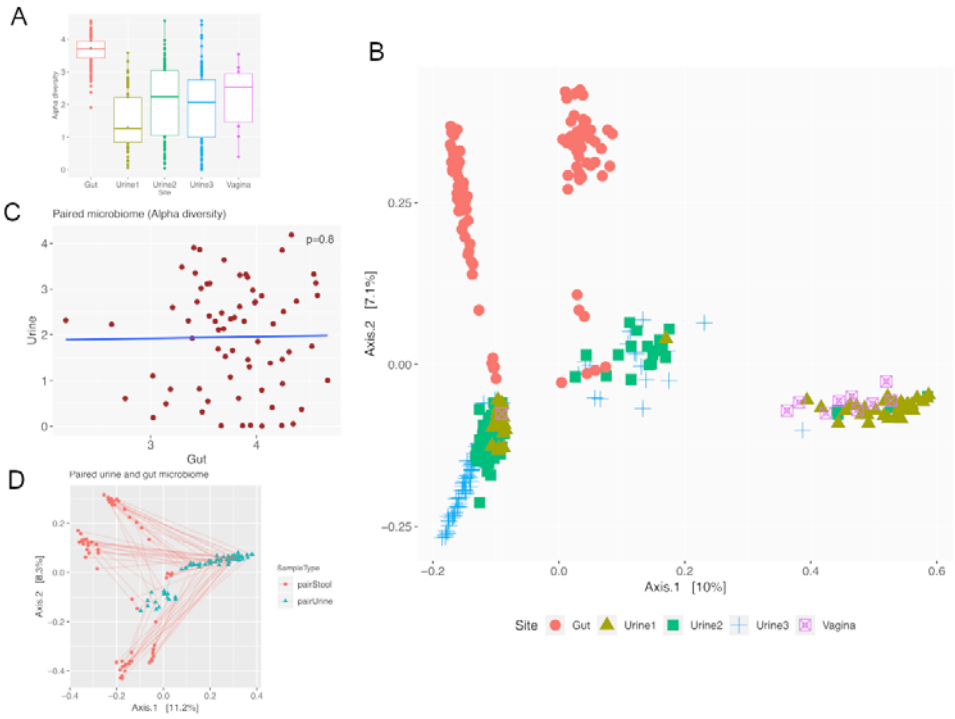

Fig 1. Urinary microbiome in older women is mostly distinct from proximal body sites and unrelated to stool microbiome. (A) Alpha diversity of urine microbiomes and other body sites. star symbol indicates significance compared to TwinsUK urobiome. **(B) Dissimilarities in urine microbiomes and other body sites.** Plots are based on unifrac distances **(C) Paired alpha diversity analysis of stool and urine collected at same time point (D) Differences in paired stool and urine microbiome from the same time point.**

<u>General description of urobiome</u>

Urine samples from 1600 mainly postmenopausal women (mean age 66.4) in the TwinsUK cohort were analysed, revealing 10955 present species-level taxa from filtered 16S data. Participant characteristics are shown in Table 1. There was high level of variability in particular species present in an individual, with only 246 (2.2%) ASVs occurring in at least 5% of samples. The use of a compositionally-sensitive analysis improved the ranking of some abundant taxa as compared to common non-compositional analysis (SFig3). To highlight intra-microbiome relationships, hierarchical balances were created, resulting in mixed-genera subclusters from 61 species-level taxa (hereafter referred to as the core urobiome). There were more Actinobacteria, Fusobacteria and Proteobacteria compared to normal gut microbiome (SFig 3B).

Having low reads (no reliably-detected microbiome (<2000 reads post-filtering)) (Supplementary Data 2) associated with younger age and lower level of health deficit; specifically, a ~20% increase in the chances of detectable microbiome for a unit increase in age (p=0.0048, OR=1.21, CI=1.07 - 1.39) and ~14% increase for a unit increase in the frailty index (OR=1.144,CI=1.01-1.30,p=0.0359). There was no association between low read status and the number of previous Urinary Tract Infections (UTIs), recent antibiotics usage, surgery episodes or number of childbirth episodes (parity); amplicon concentrations associated with parity (β=1.89,p=0.0035) alone among other demographics (Supplementary Data 2).

Host genetics' influences variation of urine microbiome

First, the quantitative twin model analysis showed considerable and significant genetic component in the first principal coordinate (PCo) of beta diversity (inter-individual) distances which capturing 57% of the variation. Heritability of this first PCo was 18% (A= 0.179, CI=0.05-0.415, p=0.003351; C=0.0049, E=0.8164, n=760 pairs) (Fig2A). Significant heritability was maintained when adjusting for other factors (Supplementary Data 3). Likewise, treating the microbiome data as Atchinson composition, the first principal component (63% of variation) on inter-sample distances showed 21% heritability (CI=0.10-0.32,C=0.00,E=0.79), and the first PC was also associated with genetic relatedness (family identity) (Kruskal-Wallis p=0.043). Some clusters showed higher heritability (Fig 2B).

In addition, the dissimilarity within relatives (twin pair) in constrained principal coordinates analysis and the average difference in Euclidean distances to the normal PCo median were both smaller for monozygotic pairs ($p \leq 0.027$) (Fig2C and Fig 3D) (Supplementary Data 3), providing further evidence of a genetic component. While the study population was majorly of British ancestry, and therefore ethnicity findings would need to be confirmed, the second PCo of the microbiome diversity differed according to the 4 major ancestry or ethnic origins present (1st PC; p=0.156; 2nd PC p=0.000143), as was the Bray-Curtis dissimilarity between the ancestry groups (Supplementary Data 3, Fig. 2D).

Moreover, the common urobiome taxa (using balance transformations) showed heritability of 23%(95%CI=8.77 to33.7, C=1.66E-12). Almost a quarter (59 of 245)

of frequent species had heritability greater than 10%, and some of the most heritable species (e.g. *Lactobacillus iners* AB-1 *and Escherchia-Shigella sp.*) clustered together and members showed phylogenetic relatedness among themselves and with *Christenellaceae* species (SFig 1D-E). Because of the potential role of some of these heritable species in UTIs, we also tested the heritability of occurrence of prior urinary tract infections, finding prior UTI to be significantly heritable (A=0.273, 95%CI=0.178 – 0.368, p=3.073E-13, see Supplementary Data 3) possibly up to 40%.

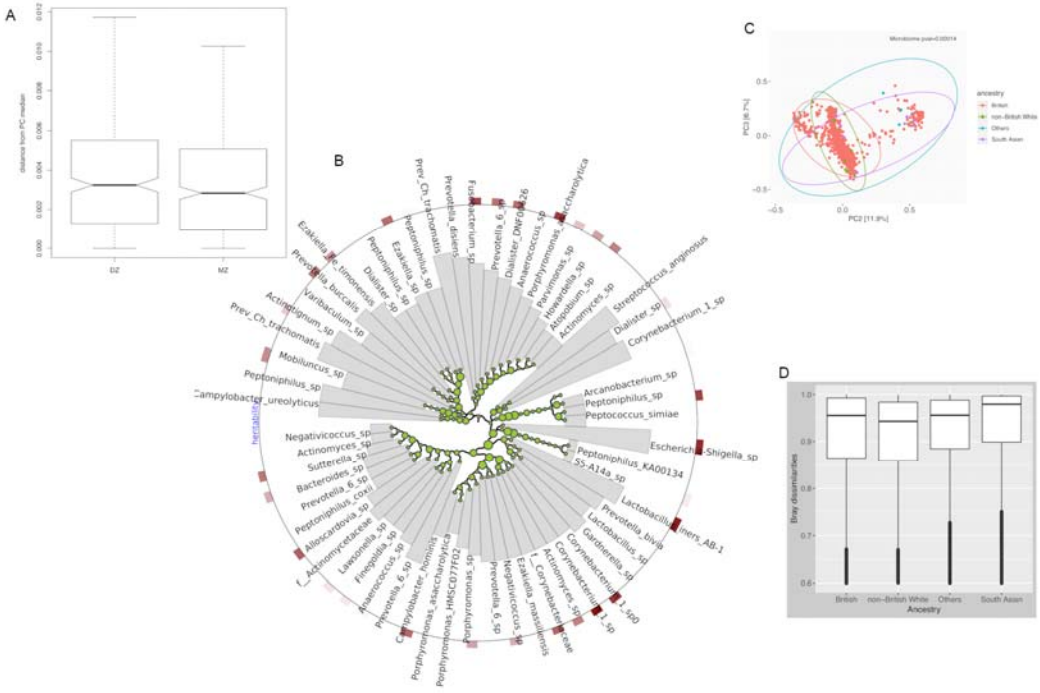

Fig 2. Host genetics considerably influences variation of urine microbiome. (A) Discordance in paired twin types for Euclidean distances to median microbiome in PC. MZ-monozygotic pair; DZ:Dizygotic pair; PC: principal coordinate (**B). Heritability and interaction of core urinary microbes.** Size of circles at each subcluster and intensity of rectangular bars at the tips represent increasing heritability of taxa. Neighbouring species in a clade show co-abundance. Taxa are annotated to indicate different species. Clusters are not phylogenetic. **(D). Microbiome principal coordinates with ancestral origin.** White British constitute>90% of individuals. P-values are derived from permutational models due to imbalanced sizes. Ellipses represent 95% confidence interval. **(E) Bray dissimilarities with the ethnic or ancestry divisions.**

175
176

177 Host-related/environmental factors in urinary microbiome, especially age, have
178 important effects

179 Age, diet, recent antibiotic usage and overall health deficit were assessed in relation to
180 the urobiome as they are known 'host-specific' influencers of gut microbiome
181 variation. Parity (previous number of births) and surgical history (had previous
182 surgery or not) were assessed as host-related "environmental" factors as they could
183 potentially alter structures in or proximal to the urinary tract. Previous history of UTI
184 was also assessed.

185 With increasing age, there is overall increase in alpha diversity (Table 1), which was
186 robust to uneven sample sizes or exclusion of small number of participants aged <50
187 ($0.10 \geq \beta \leq 0.22$, $0.00027 \leq p \geq 0.0045$). Age differed with beta diversity estimates
188 ($p<0.001$), and was a main influencer of the 3 'enterotypes' (directions) visible in the
189 PCo plot (Fig 2B). The core urobiome and one-third (22) of the subclusters, attained
190 statistical significance with age ($1.92E\text{-}30 \leq FDR \leq 0.046$).

191 The dietary index (the Healthy Eating Index), and an index of health deficit (the
192 frailty index) and antibiotics usage did not produce significant difference in alpha
193 diversity but borderline associations were found with with changes in beta diversity
194 (diet, $p=0.052$, $n=1004$; recent antibiotics usage, $p=0.041$,$n=992$; health deficit,
195 $p=0.031$, $n=1139$). Parity trended toward an association with alpha diversity reduction
196 ($p=0.058$,$n=1047$), and significantly with beta diversity ($p=0.026$,$n=1047$); surgical
197 history did not differ with beta diversity or alpha diversity ($n=540$). Occurrence of
198 UTI differed with alpha diversity ($p=0.0027$) and beta diversity ($p=0.001$). Similar
199 results were obtained using unifrac sample distances or controlling for other factors.

200 The contribution to variance that could be attributed to all factors, including host
201 genetics was then examined (Fig 3). For individuals with virtually all phenotypes
202 ($n=545$), unique contribution was obtained from R^2 decomposition on microbiome
203 beta diversity estimates, in permutational models (1000 permutations) controlling for
204 other factors. The average for each factor was used after randomly rearranging all

factors 20 times. In other scenario of measuring host genetics (Supplementary Data 3), but with a smaller sample size, the contribution of host genetics ranks higher.

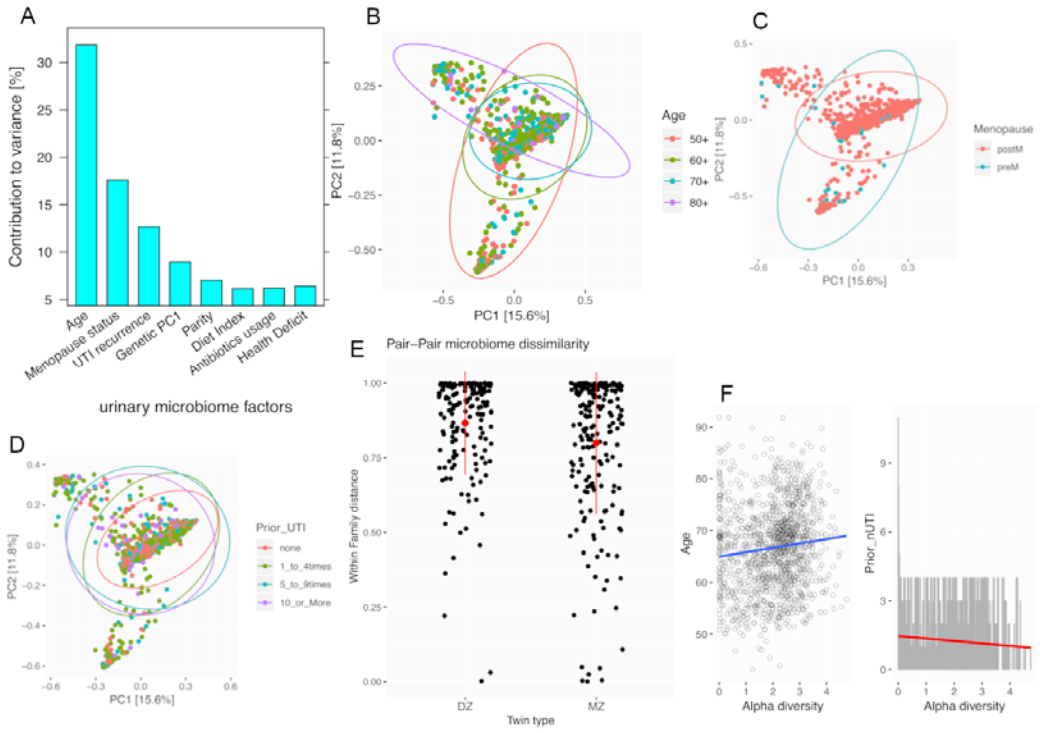

Fig 3. Top contributors to urinary microbiome variation. (A) Relative contributions to urinary microbiome. Bars represents average R^2 for each variable, controlled for the presence of other factors. Microbial variation was measured using Bray-Curtis dissimilarities. Genetic contribution shown was derived from principal components of genetic kinship calculated from whole genome data. **(B-E) Microbiome dissimilarities with (B) age (C) menopause (D) prior number of UTI. (E) within family of twin pairs (F) Trends in intra-individual Shannon diversity with age and prior number of UTI.**

Metagenomes confirm overall 16S microbiome data variation

Using shotgun metagenome data for a subset of 178 individuals, we also examined how closely the overall patterns of the 16S data are replicated in the metagenome data. The classified metagenome reads were 99.64% Bacteria (Supplementary Data 5) and a greater number of urine metagenomes (total and per individual) were obtained than earlier reported in literature. Sample-sample variation or inter-sample distances in the microbiome data were highly correlated from metagenome and 16S data (for

Atchinson compositions with Euclidean distance, Mantel's r=0.859, p=0.002; and for Bray dissimilarities, Mantel's r=0.799,p=0.001). Sixteen of the top 20 abundant taxa are also within the top 20 of the metagenome data. The core microbiome found in 16S data was largely recapitulated in the metagenomics analysis; 27 of the 31 genera (87%) forming the core urobiome using 16S data were also replicated in the metagenome data. From this core, the total number of species identifiable approximately doubled (125 vs 61 in total, 94 vs 53 in the replicated genera) most likely to due to better species assignment.

Discussion

In this study, we utilised new approaches in (urinary) microbiome analysis - using amplicon sequence variants rather than OTUs, creating microbial balances from highly frequent taxa, compositional analysis, and eliminating common batch environment effect in twin-pairs - to explore host factors in an relatively large, unselected community-based study population of women. These approaches strengthen deductions made from factors in urinary microbiome variation, for instance, increased diversity with age contrary to previous studies (e.g. Curtiss et al., 2018; Kramer et al., 2018; Liu et al., 2017;Wang et al., 2017).

Urine and other body sites

The ordination patterns of the microbiomes support current thinking that the urobiome is a distinct site, similar to the observations that most bladder microbiome (urine obtained directly by catheter) differ from vaginal or stool microbiome (Wolfe & Brubaker, 2019). The more divergent of the urine studies (Urine1 cohort) involved patients with incontinence and collection was wholly catheterized. In a very small minority of individuals where urine microbiome taxa appear closer to stool, this is most likely due to phylogenetic or genome similarity in species (as no such closeness occur with non-phylogenetic measures) rather than common demographics (SFig2). In all, the current study show clear dissimilarities in stool and urine for the average population.

Host-related factors and host genetics' contribution in urinary microbiome

Parity (childbirth episodes), previous UTI occurrence, recent antibiotics usage and diet showed changes with urine microbiome diversity. Using heritability analysis, the current study showed a considerable genetic influence in the microbiome of ageing women, reaching 15% in 57% of urine taxa variation. The remainder of contribution was largely due to variance unique to individuals. Some clinically important, "uropathogenic" genera such as *Escherichia* had variants with high heritability estimates, In addition, *Lactobacillus. iners*, a commonly found vaginal microbe which is phylogenetically close to the heritable gut microbe Christenellaceae was found to be heritable in urine.

Previously, Rothschild and colleagues (2018) reported that environmental factors such as sharing household may blur genetic influence in gut microbiome composition, while Goodrich and colleagues (2014) showed host genetics played roles in gut microbiome patterns of twin-pairs. The current study, indicates significant contributions of genetics to the pattern of urine microbial composition; and controlling for cohabitation (participants asked if they live together or close with their sibling) and other known factors in urine microbial variation, did not alter the estimated the significant contributions to the pattern. Other parameters from this study bolster the observation on genetic influence: (1) samples of a member in a twin-pair were not extracted or sequenced in the same batch as the other member,(2) adding genetic relatedness statistically explained much more in the pattern of constrained ordination, (3) there was lower intra-twin difference distance to centroid among monozygotic pairs, and (4) there were differences along the lines of ethnic ancestry though the proportion of white British was dominant. Thus we conclude that host genetics influenced variation in urinary microbiome composition in this population of women.

Relative to other factors, only age, menopause status and prior history of current UTI were greater than the influence of genetics. Incidental to our main purpose, we also report here for the first time in humans that history urinary tract infection itself has a significant heritability as suggested in other species (Norris et al., 2000).

Heritable urinary pathogens

While Corynebacterium species were frequent among top core urobiome taxa with high heritability, the patterns detected for *Lactobacillus iners/jensenii and Escherichia* variants deserve mention. The *Escherichia-Shigella* taxon, renamed as

such to reflect the extreme sequence similarity of *Escherichia coli* and *Shigella*, is apparently ubiquitous in the normal urine microbiota from this data. The current study shows that presence of this taxon is influenced by (1) host genetic make up (its proportions had one of the highest heritability estimates (A=0.17,CI=0.11-0.29) of all frequent urine microbial species); and (2) age (its coefficient in age, 0.43, is more than double that of UTI history, 0.20). The relatively high heritability of these taxa were also replicated in the subset with metagenomics data and in all, the findings may have implications in the mixed success of *E. coli* vaccine trials (Huttner et al., 2017).

The current study has limitations. Questionnaire data, which is subject to accurate recall and self-report by participants, was part of measures used in deriving variables such as UTI, diet and frailty. Another limitation may be the use of a single midstream urine sample set from an individual, and as such, prior microbiome stability information is unknown. Clearly, further research is needed to confirm if the findings also relate to the male urinary microbiome.

To conclude, this is the first 'large-scale' human study to identify the factors influencing composition of the female urinary microbiome. The urinary microbiome was distinct and apparently unrelated to stool microbiome. It shows a significant contribution of host genetics. Key species known to have pathogenic potential were among the most heritable microbes. Age and menopausal status were the factors with greatest influence on the urinary microbiome in women.

Acknowledgement

We thank Dr Alan Wolfe and Roberto Limeira of Health Sciences Division, Loyola University, Chicago, United States for providing access to raw sequence data from two urine studies; the phenotype data team at TwinsUK; laboratory team at TwinsUK for sample handling; and Rachel Horsfall, Marina Mora Ortiz, Mary NiLochlainn for discussions and comments on the manuscript. CS received research funding through the Chronic Disease Research Foundation which receives funds from the Denise Coates Foundation. We also thank all participants in TwinsUK (www.twinsuk.ac.uk).

Author Contributions

Conceptualization: C.J.S, T.S. and A.S.A; Investigation: C.J.S., G.H., G.A., R.B., P.W. and R.K.; Methodology: C.J.S. and A.S.A.; Formal Analysis: C.J.S. and A.S.A; Writing: A.S.A, C.J.S, G.H., T.S. and R.K.; Funding Acquisition: C.J.S. and T.S; Supervision: C.J.S., T.S. and R.K.

Declaration of Interests

The authors declare no competing interests

Figure Legends

Fig 1. Urinary microbiome in older women is mostly distinct from proximal body sites and unrelated to stool microbiome. (A) Alpha diversity of urine microbiomes and other body sites. star symbol indicates significance compared to TwinsUK urobiome. **(B) Dissimilarities in urine microbiomes and other body sites.** Plots are based on unifrac distances **(C) Paired alpha diversity analysis of stool and urine collected at same time point (D) Differences in paired stool and urine microbiome from the same time point.**

Fig 2. Host genetics considerably influences variation of urine microbiome. (A) Discordance in paired twin types for Euclidean distances to median microbiome in PC. MZ-monozygotic pair; DZ:Dizygotic pair; PC: principal coordinate (**B). Heritability estimates in species and clusters of highly frequent urinary microbes in paired twins.** Cb represents cluster names, Size of circles at each subcluster and intensity of rectangular bars at the tips represent increasing heritability of taxa. Taxa are annotated to indicate different species. Only species in at least 20% of population form clusters. Clusters are hierarchical but not phylogenetic. **(D). Microbiome principal coordinates with ancestral origin.** White British constitute>90% of individuals. P-values are derived from permutational models due to imbalanced sizes. Ellipses represent 95% confidence interval. **(E) Bray dissimilarities with the ethnic or ancestry divisions.**

Fig 3. Top contributors to urinary microbiome variation. (A) Relative contributions to urinary microbiome. Bars represents average R^2 for each variable, controlled for the presence of other factors. Microbial variation was measured using

Bray-Curtis dissimilarities. Genetic contribution shown was derived from principal components of genetic kinship calculated from whole genome data. **(B-E) Microbiome dissimilarities with (B) age (C) menopause (D) prior number of UTI. (E) within family of twin pairs (F) Trends in intra-individual Shannon diversity with age and prior number of UTI.**

Tables

Table 1. Summary of participants in TwinsUK urinary microbiome study

Phenotype category	Subcategory	α-D index (mean±SD)	Ave. no of unique taxa(mean±SD)	No. of samples	Age (mean±SD)
Participants		2.01±1.05	65.7±48.8	1600	66.7±8.3
Previous UTI occurrence[$]	0 times	2.14±1.0	66.1±43.1	393	67.6±8.2[$]
	1-4 times	2.02±1.04	67.5±51.0	719	65.9±7.8
	5-9 times	1.98±1.03	65.4±45.2	208	66.3±8.3
	10times >	1.79±1.17	60.0±53.9	201	65.7±8.3
Age[$]	<50-54	1.56±0.76	45.9±32.2	117	-
	55-59	1.86±1.13	61.7±49.7	210	-
	60-64	2.00±0.98	63.5±44.8	327	-
	65-69	2.04±1.03	66.0±49.8	409	-
	70-74	2.16±0.97	71.5±50.6	276	-
	75-79	2.26±1.12	74.5±50.1	170	-
	80-84	2.02±1.12	63.7±41.9	68	-
	85-	1.73±1.42	71.7±62.3	23	-
RecentAntibiotic usage:3mths	Yes	1.97±1.20[ns]	70.0±53.0[ns]	47	68.3±8.0[ns]
	No	2.03±1.06	66.0±49.0	945	66.6±8.3
Frailty	<0.15	2.05±1.01[ns]	67.0+49.0[ns]	511	65.9±7.5[$]
	0.15-0.29	1.99±1.05	64.8±49.0	834	66.1±8.0
	0.3-0.44	2.04±1.15	67.5±48.0	227	68.4±8.9
	>0.45	1.75±1.17	62.0±47.0	28	68.5±8.2

Legend. α-D: Shannon H index of alpha diversity; No. of taxa refers to number of unique sequence variant per sample i.e. no of potential species. Diversity measures were calculated after subsampling to 2000. S/NS indicates statistical significance or not for tests of a phenotype as a continuous variable. Post-hoc pairwise comparisons showed no difference in alpha diversity after 75years.

STAR Methods

2.1 Cohort and Phenotypes

The TwinsUK cohort has been described elsewhere (Verdi et al. 2019). Participants in the cohort are community dwelling twin pairs, recruited without any specific clinical phenotype. Various demographics were examined. Medical history questionnaires were used to define age (from birth date), history of urinary tract infections (UTIs), cohabitation, antibiotic usage, previous hysterectomy, previous oophorectomy, caesarian section and menopause status. The frailty index, calculated from clinical, physiological and mental domains (Livshits et al., 2017) was used as a measure of health deficit, and the Healthy Eating Index (Bowyer et al. 2018) based on food frequency questionnaires used to assess diet.

2.2 16S Microbiome Sequencing and Analysis

Twin-pair samples were separated for processing. Extraction and Sequencing was performed at the Knight Lab, University of California San Diego. A low biomass pipeline designed to extract optimal yields of DNA was used with 16S V4 marker-based paired-end sequencing on IlluminaMiSeq platform. Multilevel quality filtering procedures and data analysis were applied to remove potential contaminants (Suppl Methods). In summary, amplicon sequence variants (ASVs), were filtered, and analysed as individual taxa and as clusters based on highly frequent variants, with subsequent compositional balance transformations (Morton et al.,2017) (Supplementary Methods). The current data was also compared to those of previous microbiome studies with similar age-range of participants after re-analysis of such data to produce ASVs (Supplementary Methods). Diversity analysis was carried out with Shannon index, Unifrac, Bray and Atchinson distances, and permutational analysis of variance was used to test inter-sample differences. Taxa counts were centred-log ratio transformed after adding a pseudocount of 1, and independent taxa associations were pruned for presence in at least 5% of samples.

2.3 Metagenome Analysis

Shotgun metagenomic sequencing was carried out for 178 of the participants using newer approaches (Hillman et al., 2018), with additional 14 blanks for quality control. This subset of participants were chosen to include equal numbers of dizygotic pairs and monozygotic twin pairs, as well as equal numbers of twin pairs showing discordance and concordance in 16S microbial diversity. After quality control filtering, and mapped human reads removal (based on hg19) one sample was

410 excluded, and the final analysis included 177 samples, comprising 43 pairs of
411 dizygotic twins and 45 pairs of monozygotic twins (n=176). Potential contaminant
412 species were also removed (Supplementary Methods).

413

414 2.4 Host genetics analyses

415 Heritability was calculated using an ACE model in which the component of
416 phenotypes explained by genetics in twin pairs was estimated. Samples from co-twin
417 were separated into different batches for sample preparation and sequencing to
418 remove the shared technical environment related to batching. This further solidified
419 the deductions made from the analysis of the genetic effects. Where constrained
420 principal coordinates analysis was used, microbiome data was ordinated with the
421 family ID tested as a predictor, then the dissimilarity within a family was then
422 extracted to compare twin types. Discordance analysis was based on quantitative
423 difference in pairs of monozygotic and dizygotic twins. Analysis on ethnic origin of
424 participants based on information obtained from questionnaires. To represent host
425 genetic variation, first principal component from SNP-based kinship data, raw whole
426 genomic sequence data (available for a separate subset of unrelated participants)
427 which were part of a previous study (Long et al. 2017), and zygosity:family nested
428 model variance (only for twin-pairs) were obtained. Each of these were analysed
429 separately as a measure of genetic relatedness.
430

431 Throughout analysis, technical covariates, including extraction kit lots, mastermix kit
432 lot, batch, extraction and sequencing processors, and depth/library sizes (sequence
433 reads post-QC filtering) were controlled for. Raw sequence data is available from
434 qiita, phenotype data is available on request TwinsUK data access committee at
435 http://twinsuk.ac.uk/resources-for-re- searchers/access-our-data.html. Scripts and
436 codes used are available at github.com/urobiome-host-genetics
437

438 **Supplemental Information**

439

440 **SFig1A-B. Replicate diversity analysis to compare urinary microbiome from**
441 **various body sites. (C). Plots showing the ordination of paired stool and urine**

samples (D) Top heritable species. Species displayed in line bars have more than 15% heritability and star symbol indicate species detected in at least 20%.

(E). Phylogenetic tree of frequent species in urinary microbiome of older women and their heritability. Tree edges and branch length are coloured by increasing heritability estimates (from green to red). Species displayed in tree were detected at least 5% of study population.

SFig2. Comparison of demographics for individuals with closer urine and gut microbiome

SFig3. Comparison of top abundant urinary microbiome taxa using various approaches

SFig4 Additional variation explained from relatedness in twin pairs. A without relatedness B. with relatedness

References

Adebayo, A.S., Survayanshi, M., Bhute, S., Agunloye, A.M., Isokpehi, R.D., Anumudu, C.I., and Shouche, Y.S. (2017). The microbiome in urogenital schistosomiasis and induced bladder pathologies. PLoS Negl. Trop. Dis. *11*.

Alanee, S., El-Zawahry, A., Dynda, D., McVary, K., Karr, M., and Braundmeier-Fleming, A. (2019). Prospective examination of the changes in the urinary microbiome induced by transrectal biopsy of the prostate using 16S rRNA gene analysis. Prostate Cancer Prostatic Dis.

Aragón, I.M., Herrera-Imbroda, B., Queipo-Ortuño, M.I., Castillo, E., Del Moral, J.S.G., Gómez-Millán, J., Yucel, G., and Lara, M.F. (2018). The Urinary Tract Microbiome in Health and Disease. Eur. Urol. Focus *4*, 128–138.

Bowyer, R.C.E., Jackson, M.A., Pallister, T., Skinner, J., Spector, T.D., Welch, A.A., and Steves, C.J. (2018). Use of dietary indices to control for diet in human gut microbiota studies. Microbiome *6*, 77

Brown, K., Church, D., Lynch, T., and Gregson, D. (2014). Bloodstream infections due to Peptoniphilus spp.: Report of 15 cases. Clin. Microbiol. Infect. *20*, O857–O860.

Curtiss, N., Balachandran, A., Krska, L., Peppiatt-Wildman, C., Wildman, S., and Duckett, J. (2018). Age, menopausal status and the bladder microbiome. Eur. J. Obstet. Gynecol. Reprod. Biol. *228*, 126–129.

Eline Slagboom, P., van den Berg, N., and Deelen, J. (2018). Phenome and genome based studies into human ageing and longevity: An overview. Biochim. Biophys. Acta - Mol. Basis Dis. *1864*, 2742–2751.

Felice, V.D., and O'Mahony, S.M. (2017). The microbiome and disorders of the central nervous system. Pharmacol. Biochem. Behav. *160*, 1–13.

Fok, C.S., Gao, X., Lin, H., Thomas-White, K.J., Mueller, E.R., Wolfe, A.J., Dong, Q., and Brubaker, L. (2018). Urinary symptoms are associated with certain urinary microbes in urogynecologic surgical patients. Int. Urogynecol. J. *29*, 1765–1771.

Franco, A.V. (2005). Recurrent urinary tract infections. Best Practice & Research Clinical Obstetrics and Gynaecology. *19*, 861–873.

Goodman, B., and Gardner, H. (2018). The microbiome and cancer. J. Pathol. *244*, 667–676.

Goodrich, J.K., Waters, J.L., Poole, A.C., Sutter, J.L., Koren, O., Blekhman, R., Beaumont, M., Van

485 Treuren, W., Knight, R., Bell, J.T., et al. (2014). Human genetics shape the gut microbiome. Cell *159*,
486 789–799.

487 Karstens, L., Asquith, M., Caruso, V., Rosenbaum, J.T., Fair, D.A., Braun, J., Gregory, W.T., Nardos,
488 R., and McWeeney, S.K. (2018). Community profiling of the urinary microbiota: considerations for
489 low-biomass samples. Nat. Rev. Urol. *15*, 735–749.

490 Kline, K.A., and Lewis, A.L. (2016). Gram-Positive Uropathogens, Polymicrobial Urinary Tract
491 Infection, and the Emerging Microbiota of the Urinary Tract. Microbiol. Spectr. *4*.

492 Könönen, E., and Wade, W.G. (2015). Actinomyces and related organisms in human infections. Clin.
493 Microbiol. Rev. *28*, 419–442.

494 Kramer, H., Kuffel, G., Thomas-White, K., Wolfe, A.J., Vellanki, K., Leehey, D.J., Bansal, V.K.,
495 Brubaker, L., Flanigan, R., Koval, J., et al. (2018). Diversity of the midstream urine microbiome in
496 adults with chronic kidney disease. Int. Urol. Nephrol. *50*, 1123–1130.

497 Livshits, G., Lochlainn, M.N., Malkin, I., Bowyer, R., Verdi, S., Steves, C.J., and Williams, F.M.K.
498 (2018). Shared genetic influence on frailty and chronic widespread pain: A study from TwinsUK. Age
499 Ageing *47*, 119–125.

500 Luca F., Kupfer S.S., Knights D, Khoruts A and Blekhman R. (2018). Functional Genomics of Host–
501 Microbiome Interactions in Humans. Trends in Genetics *34*,30-40.

502 Ma, Z. (Sam), Li, L., and Gotelli, N.J. (2019). Diversity-disease relationships and shared species
503 analyses for human microbiome-associated diseases. ISME J. 1.

504 Moayyeri, A., Hammond, C.J., Hart, D.J., and Spector, T.D. (2013). The UK adult twin registry
505 (twinsUK resource). Twin Res. Hum. Genet. *16*, 144–149.

506 Morton, J.T., Sanders, J., Quinn, R.A., McDonald, D., Gonzalez, A., Vázquez-Baeza, Y., Navas-
507 Molina, J.A., Song, S.J., Metcalf, J.L., Hyde, E.R., et al. (2017). Balance Trees Reveal Microbial Niche
508 Differentiation. MSystems *2*, 1–11.

509 Moustafa, A., Li, W., Singh, H., Moncera, K.J., Torralba, M.G., Yu, Y., Manuel, O., Biggs, W.,
510 Venter, J.C., Nelson, K.E., et al. (2018). Microbial metagenome of urinary tract infection. Sci. Rep. *8*,
511 1–12.

512 Pearce, M.M., Hilt, E.E., Rosenfeld, A.B., Zilliox, M.J., Thomas-White, K., Fok, C., Kliethermes, S.,
513 Schreckenberger, P.C., Brubaker, L., Gai, X., et al. (2014). The female urinary microbiome: A
514 comparison of women with and without urgency urinary incontinence. MBio *5*.

515 Rothschild, D., Weissbrod, O., Barkan, E., Kurilshikov, A., Korem, T., Zeevi, D., Costea, P.I.,
516 Godneva, A., Kalka, I.N., Bar, N., et al. (2018). Environment dominates over host genetics in shaping
517 human gut microbiota. Nature *555*, 210–215.

518 Shrestha, E., White, J.R., Yu, S.H., Kulac, I., Ertunc, O., De Marzo, A.M., Yegnasubramanian, S.,
519 Mangold, L.A., Partin, A.W., and Sfanos, K.S. (2018). Profiling the Urinary Microbiome in Men with
520 Positive versus Negative Biopsies for Prostate Cancer. J. Urol. *199*, 161–171.

521 Sihra, N., Goodman, A., Zakri, R., Sahai, A., and Malde, S. (2018). Nonantibiotic prevention and
522 management of recurrent urinary tract infection. Nat. Rev. Urol. *15*, 750–776.

523 Thomas-White, K.J., Hilt, E.E., Fok, C., Pearce, M.M., Mueller, E.R., Kliethermes, S., Jacobs, K.,
524 Zilliox, M.J., Brincat, C., Price, T.K., et al. (2016). Incontinence medication response relates to the
525 female urinary microbiota. Int. Urogynecol. J. *27*, 723–733.

526 Thomas-White, K.J., Kliethermes, S., Rickey, L., Lukacz, E.S., Richter, H.E., Moalli, P., Zimmern, P.,
527 Norton, P., Kusek, J.W., Wolfe, A.J., et al. (2017). Evaluation of the urinary microbiota of women with
528 uncomplicated stress urinary incontinence. Am. J. Obstet. Gynecol. *216*, 55.e1-55.e16.

529 Verma, R., Morrad, S., and Wirtz, J.J. (2017). Peptoniphilus asaccharolyticus -associated septic
530 arthritis and osteomyelitis in a woman with osteoarthritis and diabetes mellitus. BMJ Case Rep. *2017*.

531 Whiteside, S.A., Razvi, H., Dave, S., Reid, G., and Burton, J.P. (2015). The microbiome of the urinary
532 tract - A role beyond infection. Nat. Rev. Urol. *12*, 81–90.

533 Wolfe, A.J., and Brubaker, L. (2019). Urobiome updates: advances in urinary microbiome research.
534 Nat. Rev. Urol. *16*, 73–74.

535 Wolfe, A.J., Toh, E., Shibata, N., Rong, R., Kenton, K., FitzGerald, M.P., Mueller, E.R.,
536 Schreckenberger, P., Dong, Q., Nelson, D.E., et al. (2012). Evidence of uncultivated bacteria in the

adult female bladder. J. Clin. Microbiol. *50*, 1376–1383.

Wu, P., Chen, Y., Zhao, J., Zhang, G., Chen, J., Wang, J., and Zhang, H. (2017). Urinary microbiome and psychological factors in women with Overactive bladder. Front. Cell. Infect. Microbiol. *7*.

El-Zawahry, A., Dynda, D., McVary, K., Karr, M., and Braundmeier-Fleming, A. (2019). Prospective examination of the changes in the urinary microbiome induced by transrectal biopsy of the prostate using 16S rRNA gene analysis. Prostate Cancer Prostatic Dis.

Hillmann, B., Al-ghalith, G.A., Shields-cutler, R.R., Zhu, Q., Gohl, D.M., Beckman, K.B., Knight, R., and Knights, D. (2018). crossm Metagenomics. MSystems *3*, 1–12.

Huttner, A. et al. (2017). Safety, immunogenicity, and preliminary clinical efficacy of a vaccine against extraintestinal pathogenic Escherichia coli in women with a history of recurrent urinary tract infection: a randomised, single-blind, placebo-controlled phase 1b trial. Lancet Infect. Dis. *17*, 528–537.

Liu, F., Ling, Z., Xiao, Y., Yang, Q., Zheng, L., Jiang, P., Li, L., and Wang, W. (2017). Characterization of the urinary microbiota of elderly women and the effects of type 2 diabetes and urinary tract infections on the microbiota. Oncotarget *8*, 100678–100690.

Rani, A., Ranjan, R., McGee, H.S., Andropolis, K.E., Panchal, D. V., Hajjiri, Z., Brennan, D.C., Finn, P.W., and Perkins, D.L. (2017). Urinary microbiome of kidney transplant patients reveals dysbiosis with potential for antibiotic resistance. Transl. Res. *181*, 59–70.

Verdi, Serena, Golboo Abbasian, Ruth C. E Bowyer, Genevieve Lachance, Darioush Yarand, Paraskevi Christofidou, Massimo Mangino, Cristina Menni, Jordana T. Bell, Mario Falchi, Kerrin S. Small, Frances M. K Williams, Christopher J. Hammond, Deborah J. Hart, C.J.S. (2019). TwinsUK: The UK Adult Twin Registry. Twin Res. Hum. Genet. *Accepted*.

Wang, H., Altemus, J., Niazi, F., Green, H., Calhoun, B.C., Sturgis, C., Grobmyer, S.R., and Eng, C. (2017). Breast tissue, oral and urinary microbiomes in breast cancer. Oncotarget *8*, 88122–88138.