BMC Urology

RESEARCH

Open Access

A highly accurate nanopore-based sequencing workflow for culture and PCR-free microbial metagenomic profiling of urogenital samples

Ben Ferneyhough[1†], Matthew Roddis[1†], Sam Millington[1†], Jess Quirk[1], Chrysolin Clements[1], Sean West[1], Rosalind Schilizzi[1], Michael D. Fischer[1] and Nick J. Parkinson[1*]

Abstract

Background The application of molecular sequencing methods for microbiome profiling of biological samples are largely restricted to research use. However, they require significant resources such as time and cost and can suffer from amplification biases that may hamper interpretation of complex systems. These issues are also a barrier to adoption as standard clinical tools in, for example, diagnosis of urogenital infections. We report a new method that utilises third generation long-read nanopore sequencing to produce fast, accurate and fully quantitated metagenomic microbiome profiles. Here, as proof of principle, we apply this methodology to reassess the healthy urogenital microbiomes of asymptomatic female and male samples.

Results We show that our method is capable of accurately and reproducibly detecting both levels and composition of a synthetic mixture of ten species comprising known amounts of hard to lyse gram-positive bacteria, gram-negative bacteria and yeast. When applied to urogenital samples, we confirm previous observations that the female asymptomatic vaginal and urinary microbiomes are predominated by *Gardnerella spp.* or one of several *Lactobacillus* species (*L. crispatus, L. gasseri, L. iners* or *L. jensenii*) that conform to previously defined community state types. We show the tight relationship between vaginal and urinary populations of the same individual at both species and strain level, provide evidence for the previously observed dynamic nature of these microbiomes over a menstrual cycle and compare biomass and complexity of male and female urobiomes.

Conclusions We set out to develop an unbiased, amplification and culture-free, fully quantitative metagenomic microbiome profiling tool. Our initial observations suggest our method represents a viable alternative to existing molecular research tools employed in the analysis of complex microbiomes.

Keywords Microbiome profiling, Urotypes, Asymptomatic, 3rd Generation Nanopore Sequencing

Background

Describing the relationship between a host and their microbiome, the community of bacteria, fungi, viruses, protozoa and archaea that co-exist with them, is critical to a more complete understanding of health and disease [1, 2].

†Ben Ferneyhough, Matthew Roddis and Sam Millington contributed equally.

*Correspondence:
Nick J. Parkinson
nickp@sbl-uk.org
[1] Systems Biology Laboratory UK, Abingdon, Oxfordshire OX14 4SA, UK

Research to monitor and understand microbial populations in both healthy and dysbiotic states increasingly looks to tools based on advanced molecular methods. Next Generation Sequencing, NGS, is commonly used to identify microbiome compositions through interrogation of specific gene targets such as 16S rRNA, 18S rRNA and chaperonin-60 (*cpn60*) [3–7]. These methods are capable of identifying multiple bacterial and yeast microbiome components simultaneously and allow a more holistic snapshot of the community to be elucidated. However, workflows are labour and time intensive, can result in organism identification lacking resolution beyond the family or genus level, are restricted to analysis of only the information contained in the short PCR amplicon captured, representing < 0.01% of the genomic material available, and produce only relative ratio-metric quantitation. Short fragment shotgun NGS metagenomic methods offer more promise but can also be cost-prohibitive, time consuming, requires specialist compute and interpretation of results and, in common to all these molecular tests, can still suffer from sample biases due to variable lysis of organisms and extensive use of PCR [8].

In stark contrast to the increasing use of molecular methods to interrogate microbiomes in research, clinical analysis techniques for the identification of dysbiotic microbiomes and infection remain reliant on culture or dipstick analysis of biomarkers such as host immune cells or bacterial metabolites [9–13]. Attempts to improve clinical tools include the use of enhanced culture protocols [13–16] and, more recently, basic molecular analyses such as semi-quantitative qPCR. Enhanced culture methods increase the breadth of identifiable species but still suffer from underlying selection biases and speed constraints. Similarly, qPCR methods are limited to identifying members of pre-defined species arrays and rely heavily on signal amplification making them prone to bias and false-positive reporting. Application of current research tools such as 16S rRNA analysis or shotgun metagenomics to the clinical setting would undoubtedly lead to more reliable pathogen identification, monitoring and treatment. However, current workflows are both cost and time prohibitive.

There remains, therefore, a clear need for an improved molecular testing system that, ideally, is capable of bridging both research and clinical settings. Such a system should be capable of fully quantitative metagenomic identification of > 99% of organisms with sufficient sensitivity and speed to become routine for both research analysis and clinical application. This tool would almost certainly lead to both a better understanding of key biological systems and ultimately improve patient outcomes [12].

To directly address this need we have designed a highly sensitive quantitative third generation metagenomic microbiome profiling system, q3MetaG. Using Oxford Nanopore Technologies long-fragment sequencing platform it allows high-confidence identification at species and sample-specific strain levels. Our technique works directly from swab or urine samples and is free of any bacterial culture or PCR amplification so is less prone to contamination, false positive identification or biasing risks. Sensitivity is underpinned by a method of internal calibration allowing reporting of fully quantitative cells/mL/species sample profiles. Our process requires less than 24 h and can be readily repeated over time to create a more complete picture of entire microbial communities.

Given the relatively mature understanding of healthy urogenital microbiomes, we report the proof-of-principle application of our new method to re-survey the asymptomatic female urinary and vaginal microbiomes and compare them to asymptomatic male urinary equivalents.

Methods and materials

Mock microbiome

To assess quantitative and qualitative behaviour, aliquots of Zymo mock community D6300 (ZymoBIOMICS) were used as an input into the standard DNA extraction method (see below).

Donors

Following local ethical review board approval, participants completed written consent for urine and vaginal swab collection for research purposes. Participants were self-reported asymptomatic female and male volunteers screened using a pre-qualifying questionnaire (see appendix). Exclusion criteria included history of recurrent urinary tract infection, UTI, or episode of UTI in the previous 6 months, antibiotic use in the past 3 months, pregnancy, multiple sclerosis, diabetes, use of a current or recent urinary catheter during previous 6 months, kidney stones or kidney/bladder surgical procedures or use of long-term steroids. All samples were pseudo-anonymised during processing and data analysis.

Sample collection

Samples were self-collected by volunteer donors using a provided home sampling kit containing 30 mL universal sodium borate urine tubes (Sterilab) and sterile hard-packed vaginal swabs (Scientific Laboratory Supplies). To minimise sample contamination donors were requested to clean around the urethra thoroughly using a sterile hygienic intimate wipe (Jeevson) before collecting their sample indirectly using a disposable sterile Pee-Canter urine collection device (MedDX Solutions) and

finally decanting into the 30 mL borate universal tube. Filled tubes and swabs were immediately returned to the laboratory in UN3373 compliant packing (AlphaLabs) at ambient temperature using a guaranteed next-day service (Royal Mail). All samples were received within 48 h and stored at 4 °C until processed.

DNA Extraction

Vaginal swabs were soaked for 5 min in 1 mL viral transport medium, VTM, (ZymoBIOMICS) and agitated to release cells. 1 mL of VTM from vaginal swabs or 5 mL aliquots of urine samples were centrifuged (14,000xg for 5 min, Prism Centrifuge, Labnet) to recover cellular material and supernatant discarded. Cell samples were then processed to lyse host cells and degrade host cell DNA (HostZero Microbial DNA Kit D4310, ZymoBIOMICS) using manufacturer's instructions. Samples were spiked with a known titre of internal calibrator species (Spike-in Control I, High Microbial Load D6320, ZymoBIOMICS), microbial cells lysed by boiling for 5 min (100 °C Accublock Digital Dry Bath, Labnet) followed by bead beating methods (FastPrep 24 5G, MP Biomedicals) and microbial DNA recovered (HostZero Microbial DNA Kit D4310, ZymoBIOMICS). Extracted DNA samples were sonicated to average 1–3 kb fragment length (Bioruptor Plus, Diagenode SA) and yields quantified (Quant-IT PicoGreen, Invitrogen).

3rd Generation nanopore sequencing library preparation

Oxford Nanopore-specific barcoded sequencing libraries were synthesised using the manufacturer's standard protocols (Oxford Nanopore SQK-LSK114.24) and run on individual R10.4.1 SpotON Flongle flowcells or multiplexed and run on R10.4.1 SpotON MinION flowcells (Oxford Nanopore Technologies). Sequencing data was collected overnight, typically for 14 h.

Data processing

Resulting Fastq datasets were uploaded to our bespoke cloud-based (AWS) analysis suite (patent applied for) for processing. Individual reads were filtered for quality, length and chimerism then aligned (BLAST [17]) to a curated database of unique species-level bacterial, yeast, archaea (See appendix for list of reference species) or host references including a separately curated non-redundant plasmid and phage dataset. Alignments were secondarily filtered for various metrics using our proprietary algorithms to remove low quality mappings and noise. Per-species post-filtered mapping data was used to estimate sample input cell values by reference to mapping data for calibrator spike species. Results from individual samples were collated into 'heat-map' diagrams displaying estimated input cell numbers per species (per mL for urine input, per swab for vaginal) and displayed using a quantity-specific colourised scale. Data points resulting from less than 10 unique high-quality read mappings or equating to < 1,000 cells/mL were excluded.

Consensus assemblies were created and exported for species of particular interest where aligned read mapping exceeded 10 × coverage genome-wide, ~ 10,000 reads/species. SNP comparison tools [18] were used to compare multiple strains of the same species. Core genome phylogenies were output in Newick format and used to display cladograms [19] predicting relative sub-strain homologies or phylotypes where applicable.

Escherichia coli monoclonal cultures were grown, diluted and adjusted to approximately 1×10^6 cells/mL using OD measurements. Titres were confirmed using orthogonal measures such as cfu counts on culture plates of log series dilutions, expected vs observed yields from DNA extractions and microbial cell counter analysis (QUANTOM Tx, LOGOS Bio). These were then use as inputs into our standard DNA extraction method.

Results

Assessment of a synthetic microbiome

To assess the efficacy of our method we analysed a commercially available mock microbial community standard comprised of ten species; three gram-negative bacteria, five gram-positive bacteria and two yeasts all chosen specifically to be clinically relevant and hard to lyse. Our method was able to qualitatively identify all ten species present (Fig. 1a). Across the six replicates 0.004% of predicted cells belonged to species other than those expected in the input (Supplemental Fig. 1A data). This method, therefore, represents a very clean system for species identification with very little off-target noise. Furthermore, concordance of predicted absolute cell numbers to within ~ twofold of expected cell input values for all ten species show our method represents a highly quantitative assay (Fig. 1a; Supplemental Fig. 1A data). Standard deviation values from the six independent replicates suggest this method to be highly reproducible and, in addition, little obvious organism or gram-stain type associated deviation suggests limited bias from cell lysis and DNA isolation.

To assess method sensitivity we conducted analysis of a 10 × log dilution series of an approximate 1×10^6 cell/mL monoclonal *Escherichia coli* culture (Fig. 1b). Accuracy of measured values remained high down to ~ 1×10^3 cell inputs per species, suggesting that this assay is able to produce sensitive identification of species well below the 1×10^5 cell/mL reporting threshold employed by UK National Health Service, NHS, culture methods [20].

Highly sensitive microbiome profiling methodologies are prone to low-level background contaminants.

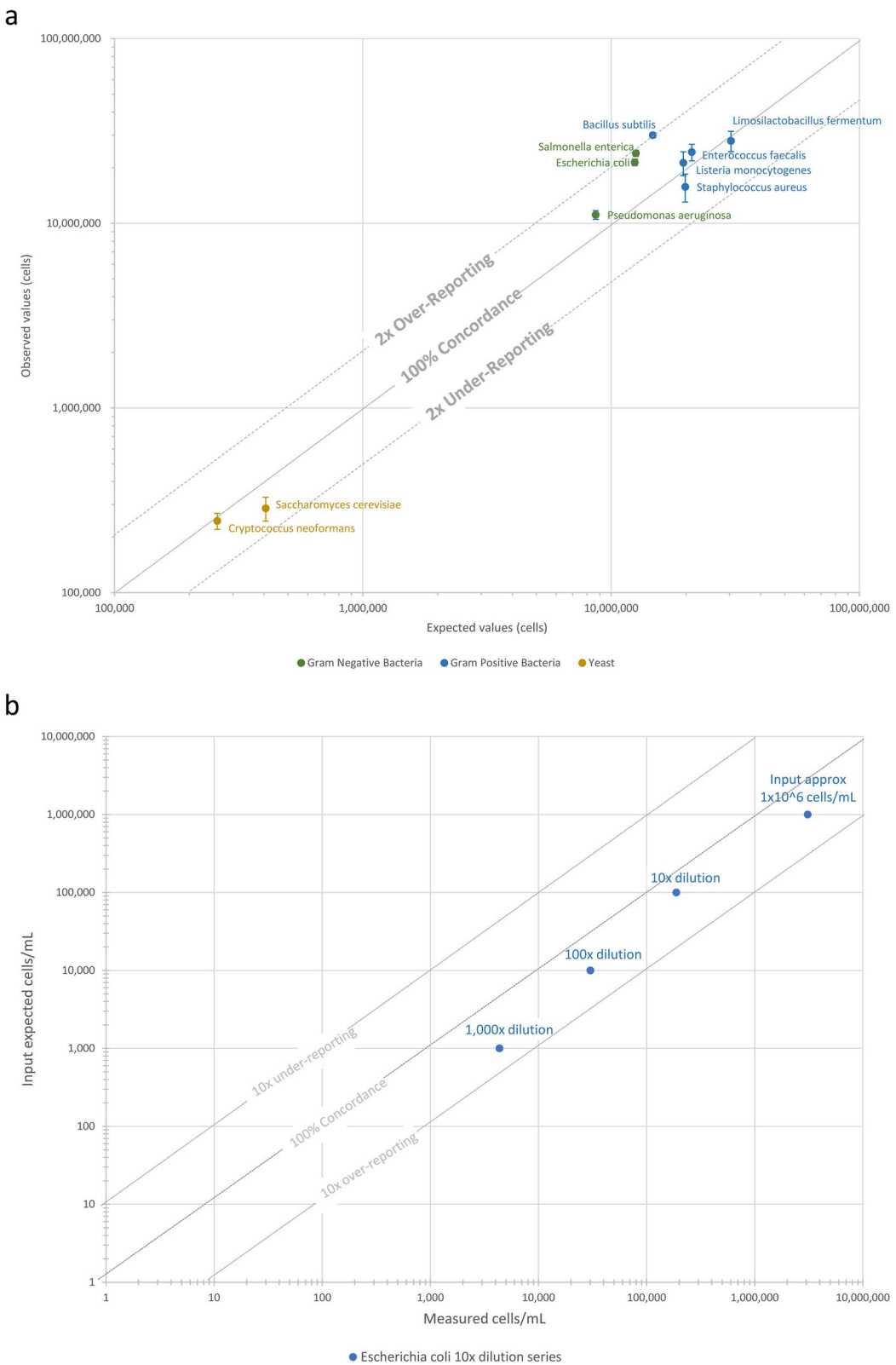


Fig. 1 **a** Accuracy and reproducibility of the meta-genomic profiling method using Zymo D6300 mock community as a known input. Average measured cell values (± 1 SD) vs predicted cell input shown for six replicate analyses of a ten species mix comprising yeasts; orange, gram-negative bacteria; green, gram-positive bacteria; blue. **b** Measured cell values vs predicted cell inputs for a 10 × log titration assay sensitivity series of *Escherichia coli*

These may be reproducibly observed for a given methodology due to intrinsic contaminants in commercial lab grade plastic-ware and reagents. To establish if a baseline microbial 'kit-ome' signal exists for our technique we processed ten independent replicates of fresh, sterile ultrapure water. Our quantitative analysis showed that 8/10 samples returned a low diversity, low biomass collection of species (>5 mapped reads/species) primarily consisting of *Sphingomonas koreensis, Cutibacterium acnes*, *Pseudomonas stutzeri* and *Pseudomonas aeruginosa* (Supplemental Fig. 1). Our analysis therefore suggests the process kit-ome is composed of readily identifiable low titre bacterial species contributing a total biomass of < 1,000 cells/mL per processed input sample.

Asymptomatic microbiome profiles: female urine

Next, we applied our technique to profile microbiomes in urine samples from self-reported asymptomatic control female donors (Fig. 2). Initially, we conducted single time-point analysis of samples taken from 23 adult female volunteers (average age 31, median 24, range 18–53). The total estimated biomass recovered for each urine sample varied considerably (range 13,000 to 15,000,000 cells/mL, median 1,400,000 cells/mL), and represented 10× to 100× greater values than those previously estimated using 16S rRNA methods [21, 22]. Microbiome alpha diversity (Shannon diversity index, H) varied from 0 to 2.5 (median 0.76) equating to between 1 and 39 (median 7) discrete organisms per sample (Fig. 2). No obvious correlation was found between the diversity of organisms detected and total sample biomass (data not shown). 76 different species were identified across the 23-sample cohort at levels > 1,000 cells/mL. *Lactobacillus crispatus* and *Gardnerella vaginalis* were the most prevalent organisms, being present in 52% and 49% of the samples respectively. 36 species were unique to single samples.

Our data confirm the existence of discrete urotypes paralleling previously reported vaginal community state types, CST's [23–25], predominated by; *Lactobacillus crispatus (CST:I), Lactobacillus gasseri (CST:II), Lactobacillus iners (CST:III), Gardnerella spp. (CST:IV)* or *Lactobacillus jensenii (CST:V)* (Fig. 2). No obviously common pattern was seen in the remaining samples.

Donor 51 provided a sample with the highest diversity microbiome (H = 2.47) dominated by high titres of *Bifidobacterium breve* (1,200,000 cells/mL, Fig. 2). 20/37 species identified across the cohort were unique to this sample including emerging uropathogens *Oligella urethralis* [26] and *Alloscardovia omnicolens* [27], each present at > 100,000 cells/mL. Of interest, this observation matches previously reports of *Bifidobacterium breve* defining a low frequency CST subtype [23, 24, 28] as well

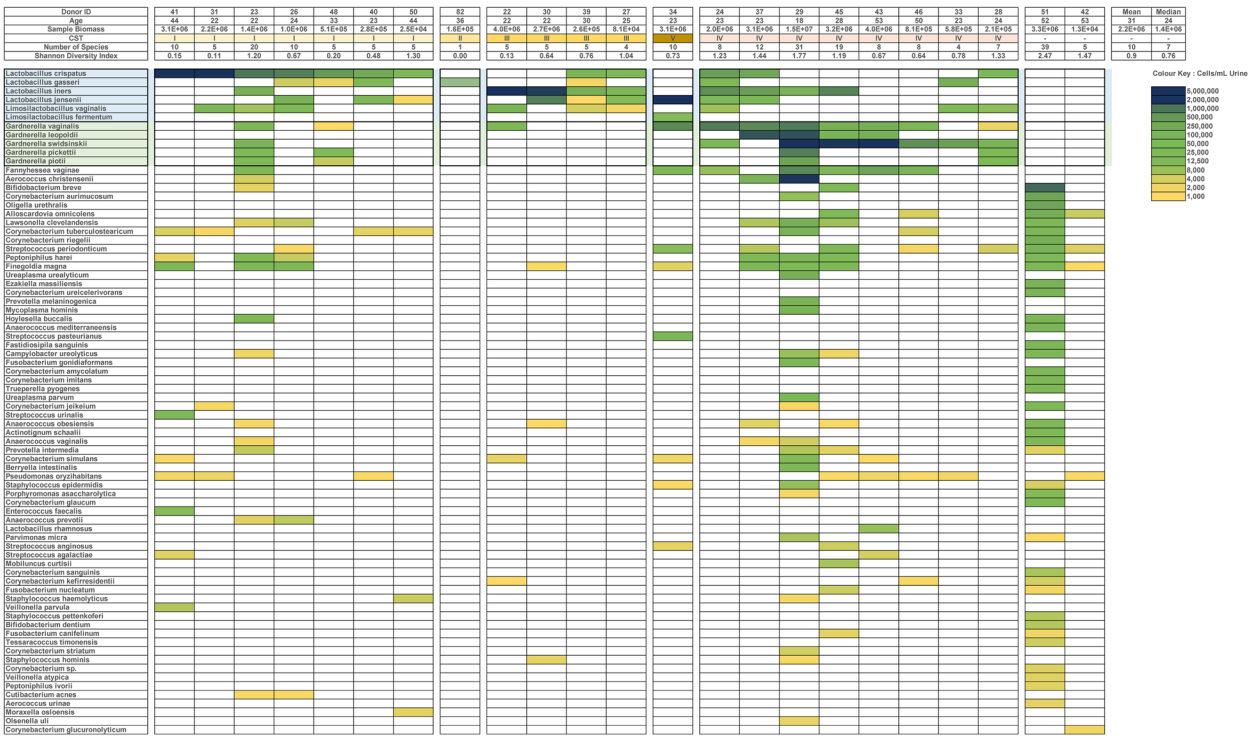


Fig. 2 Comparison of urinary profiles from 23 healthy control female donors. Values are given as predicted cells/mL urine/organism detected. Samples cluster into four major urobiomes dominated by *Gardnerella spp.*, *Lactobacillus crispatus*, *Lactobacillus iners* or *Lactobacillus jensenii*

as being a causative agent in urinary tract infections, UTI [29].

Asymptomatic microbiome profiles: vagina

Our sampling protocol captured vaginal swabs taken in parallel to each urinary test for 20 of our asymptomatic donors enabling us to co-profile asymptomatic vaginal biota (Fig. 3). Estimated cell values for each species identified at > 10,000 bacterial cells/swab are given for each sample and allow estimates of the total number of cells recovered per swab (range 64,000 to 240,000,000 cells/swab, median 14,600,000 cells/swab). As with the urine samples, the number of bacterial cells recovered for each vaginal swab varied considerably, some of this likely due to variation in individual donor swabbing technique. On average 10 discrete organisms (range 1–35, median 6) were identified per sample equating to alpha diversity measures of 0 to 2.4, median 0.8 (Fig. 3). A total of 64 different species were identified across the cohort, 50% of which were also observed in the urine samples. As in the urine samples, *Lactobacillus crispatus* was the most prevalent observed organism, found in 50% of the vaginal swab samples. Twenty-four species were unique to single donor samples.

Evidence of discrete vaginal microbiomes dominated by the same species identified in the urine analysis; *Gardnerella spp.*, *Lactobacillus iners*, *Lactobacillus gasseri* or *Lactobacillus crispatus* agree with previous community state type, CST, definitions [23–25, 28, 30] and suggest a considerable overlap between the two microbiomes of the same individual (Figs. 2 and 3). Indeed, correlation analysis between samples collected from the same donor on the same day suggests that urinary and vaginal profiles share an average 65% of identified species per individual (Fig. 4a) supporting the previously observed [25] tight relationship between these microbial communities (Fig. 4b).

Co-profiled urine and vaginal microbiomes: temporal series

In-depth analysis of weekly urine and vaginal samples from three volunteers produced a temporal series spanning a full menstrual cycle (Fig. 5a-c). High levels of concordance are apparent between urinary and vaginal microbiomes of the same individual across the majority of this testing period. However, transitory changes were observed in vaginal and urinary microbiome composition of some individuals. Of interest, the nature of these changes seem specific to each location; the *Lactobacillus*

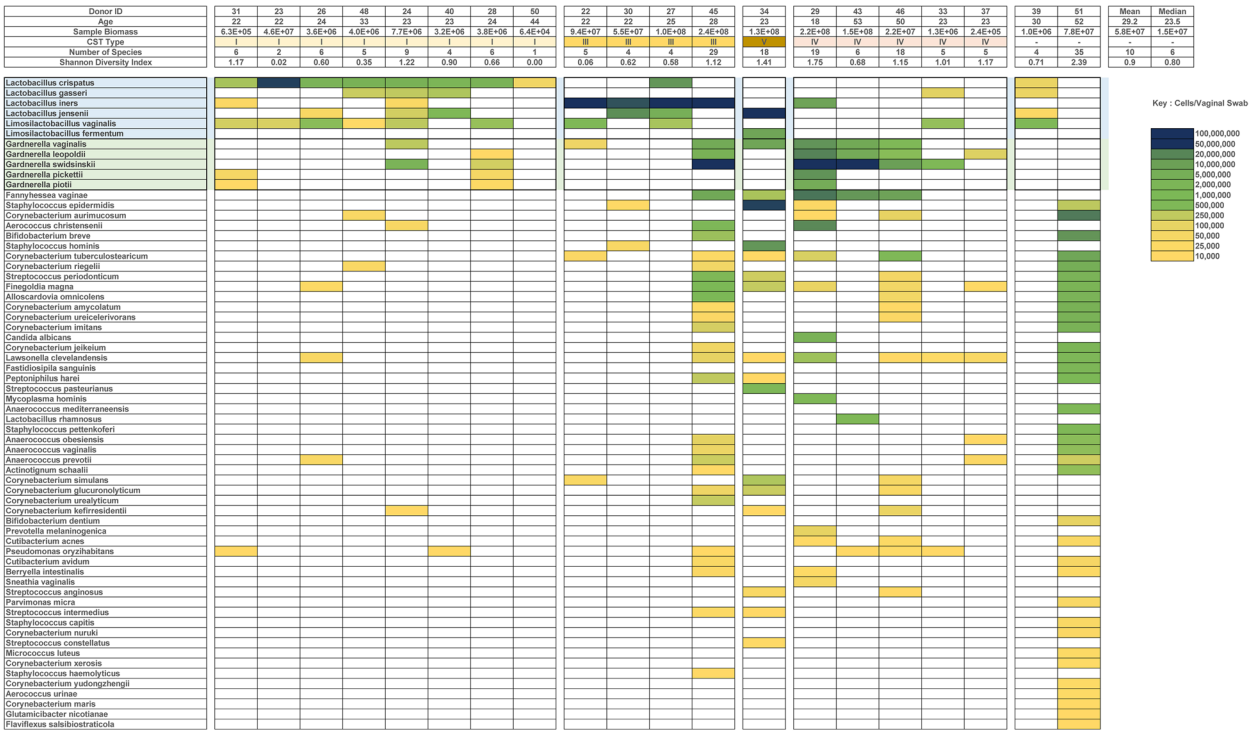


Fig. 3 Comparison of vaginal swab profiles from 20 healthy control female donors. Values are given as predicted cells/swab/organism detected. Akin to urine samples (Fig. 2.) samples cluster into four major microbiomes dominated by *Gardnerella spp.*, *Lactobacillus crispatus, Lactobacillus iners* or *Lactobacillus jensenii*

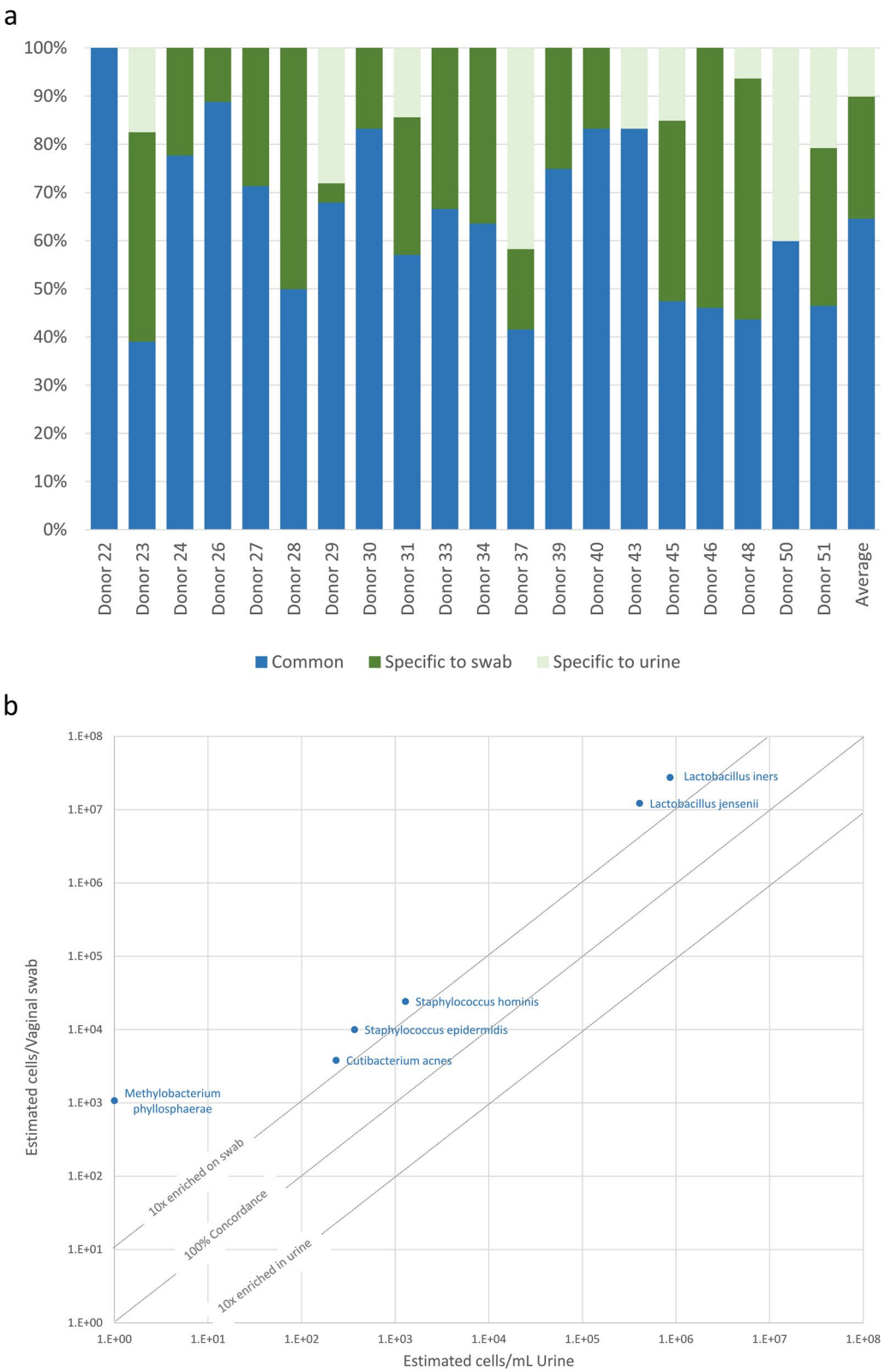


Fig. 4 **a** Comparison of urinary and vaginal microbiome species concordance from the same donor. Histogram shows the percentage of species in each donor that were present in both vaginal and urinary microbiomes (dark blue), specific to vaginal samples (dark green) or specific to urine samples (light green). Where species were found at > 1,000 cells/swab any detected level in urine for the same species was scored as a match. **b** Example of relationship between measured species (> 1,000 cells) in vaginal and urinary microbiomes in samples from donor 30

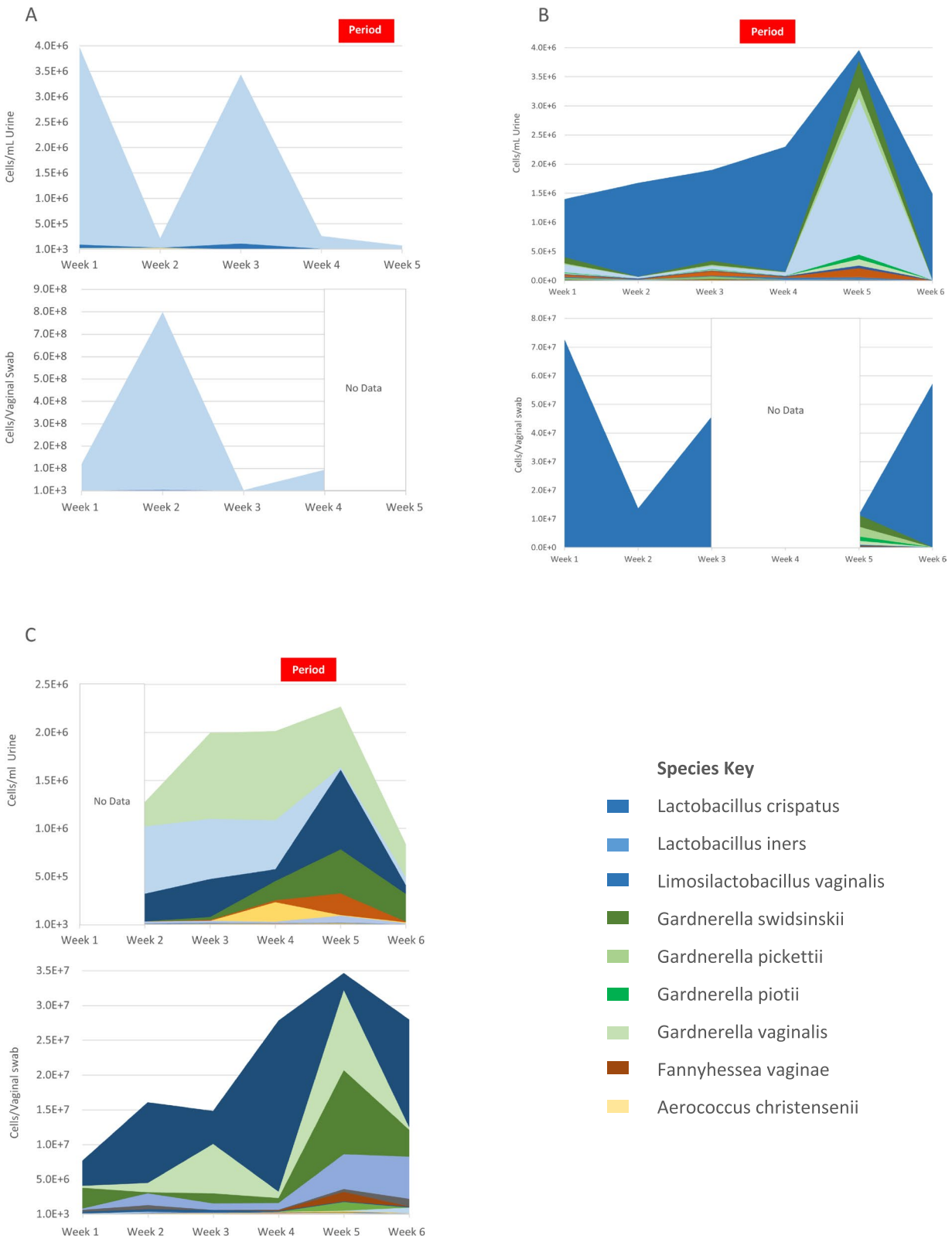


Fig. 5 **a**-**c** Temporal series of urine (top) and vaginal swab (bottom) microbiome profiles for three asymptomatic control female donors; Donor 22, Donor 23, Donor 24 respectively. Five separate sequential weekly assays spanning a full menstrual cycle are shown for each. Estimated cells/mL/organism of urine or cells/swab/organism are plotted for each species at each time-point

crispatus:*Lactobacillus iners* ratio is transiently inverted in donor 23 & 24 urine samples whilst in the vagina in the same donors the *Gardnerella vaginalis & Gardnerella swidsinskii*:*Lactobacillus crispatus* ratios display a similar phenomenon over the same time frame. These findings are compatible with previous studies describing similar phenomena in both urinary [22] and vaginal microbiomes [3, 30].

Lactobacillus crispatus strains in urine vs vaginal samples

Our method provides consensus metagenome references for any bacterial species present with greater than tenfold aligned sequence coverage. Extracted references can be compared using genome-wide SNP comparison tools to produce cladograms highlighting relationships to key published reference strains of the same species. Six of our asymptomatic control donors displayed *Lactobacillus crispatus* levels with sufficient coverage to construct independent consensus references for both their urine and vaginal swab samples. Comparison of these references against previously published vaginal and gut strains [31] shows that, in all six cases, the closest homologies were between urine and vaginal *L. crispatus* references computed from the same donor (Fig. 6). This pairing of donor specific references strongly suggests that *L. crispatus* strains in both the urine and vaginal samples are sufficiently unique to an individual donor and shared across urogenital sites. Of note, all consensus references show highest homology to published vaginal rather than gut strains, further supporting the hypothesised anatomical niche-specific strain adaptation of this genetically diverse species [31].

Asymptomatic microbiome profiles: male urine

To investigate the sex difference of urinary microbiomes in asymptomatic individuals we profiled urine donated by 18 asymptomatic male volunteers (Fig. 7). In general, lower numbers of species were returned for each sample with 9/18 failing to return any organisms present at > 1,000 cells/mL, similar to that previously reported [32]. On average male urine samples returned ~ 300 × lower total biomass/mL compared to female urine samples (range 320 to 160,000 cell/mL, median = 4,700 cells/mL) with between 1 and 18, median 3, discrete organisms equating to alpha diversity (H) measures of 0 to 2.6, median 0.68 (Fig. 7). Of note, bacteria associated with the 'kit-ome', *Cutibacterium acnes* and *Pseudomonas stutzeri* were seen to be present across the majority of samples at < 1,000 cells/species (*C. acnes* median value 480 cells/mL, *P stutzeri* median value 200 cells/ml) suggesting that this should be a consideration

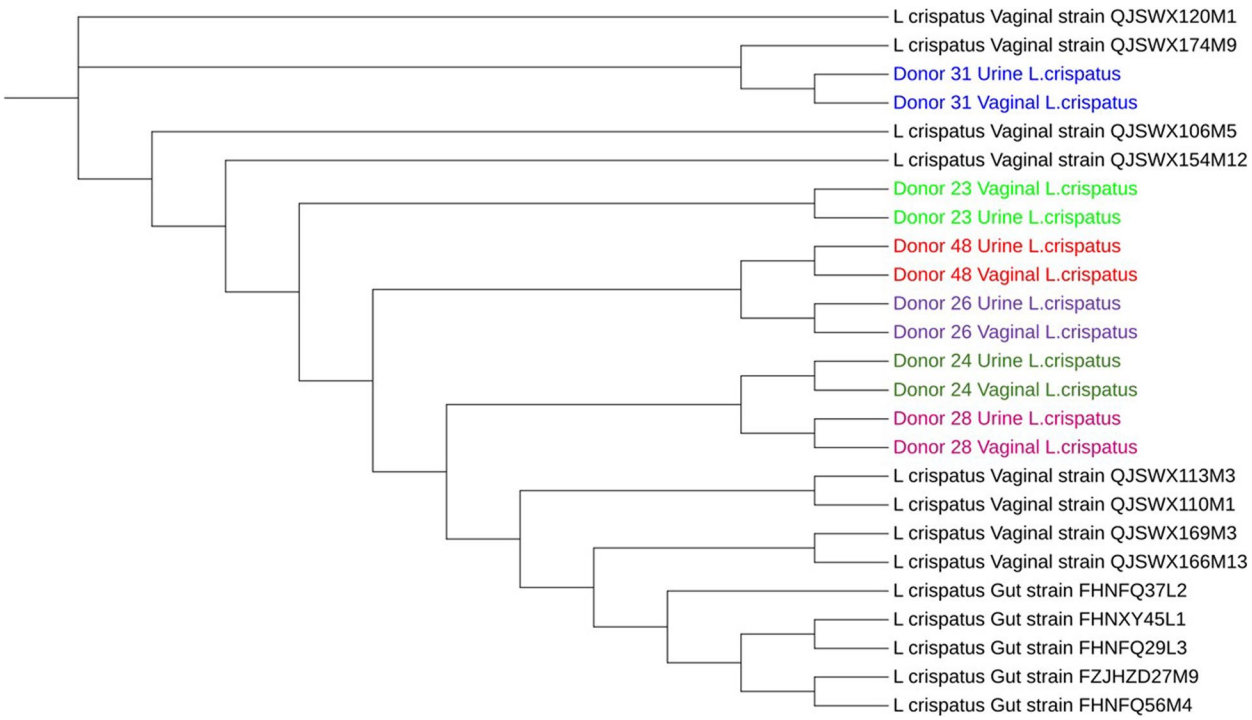
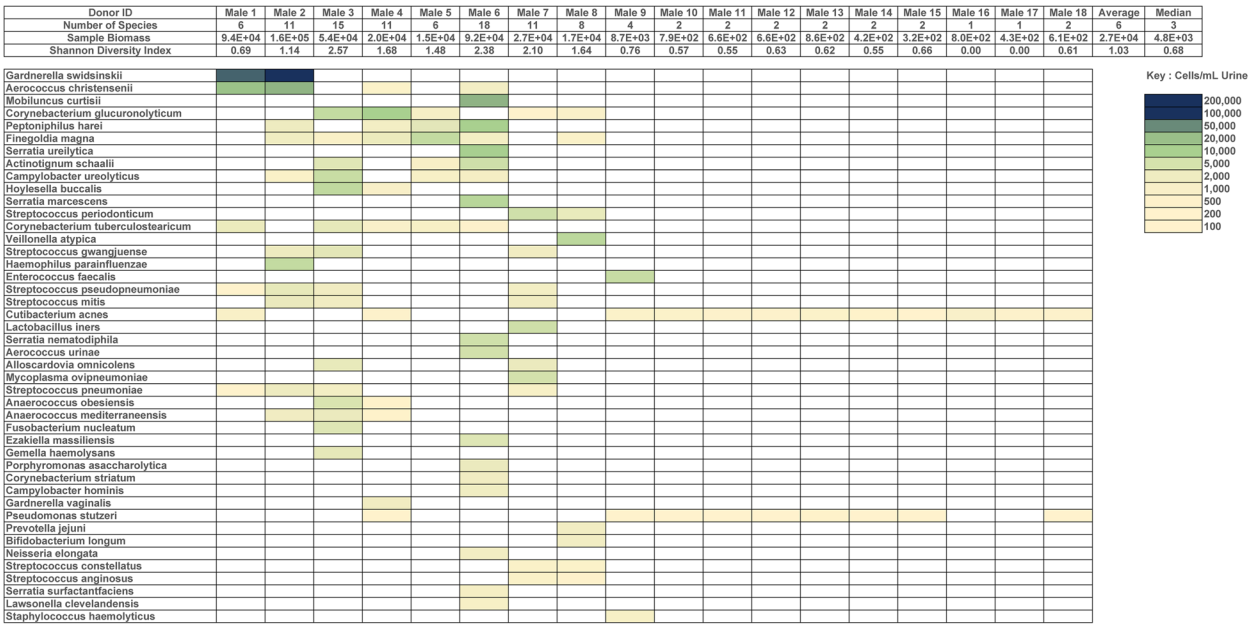


Fig. 6 Cladogram displaying SNP based relationship between consensus builds of *Lactobacillus crispatus* from urine and vaginal swab samples from six donors with comparison to eight 'Vaginal strains' and five 'Gut' strain references taken from [31]. Vaginal and urine consensus references from the same individual show highest homology to each other and are highlighted in the same colour



Donor ID	Male 1	Male 2	Male 3	Male 4	Male 5	Male 6	Male 7	Male 8	Male 9	Male 10	Male 11	Male 12	Male 13	Male 14	Male 15	Male 16	Male 17	Male 18	Average	Median
Number of Species	6	11	15	11	6	18	11	8	4	2	2	2	2	2	2	1	1	2	6	3
Sample Biomass	9.4E+04	1.6E+05	5.4E+04	2.0E+04	1.5E+04	9.2E+04	2.7E+04	1.7E+04	8.7E+03	7.9E+02	6.6E+02	6.6E+02	8.6E+02	4.2E+02	3.2E+02	8.0E+02	4.3E+02	6.1E+02	2.7E+04	4.8E+03
Shannon Diversity Index	0.69	1.14	2.57	1.68	1.48	2.38	2.10	1.64	0.76	0.57	0.55	0.63	0.62	0.55	0.66	0.00	0.00	0.61	1.03	0.68
Gardnerella swidsinskii																				
Aerococcus christensenii																				
Mobiluncus curtisii																				
Corynebacterium glucuronolyticum																				
Peptoniphilus harei																				
Finegoldia magna																				
Serratia ureilytica																				
Actinotignum schaalii																				
Campylobacter ureolyticus																				
Hoylesella buccalis																				
Serratia marcescens																				
Streptococcus periodonticum																				
Corynebacterium tuberculostearicum																				
Veillonella atypica																				
Streptococcus gwangjuense																				
Haemophilus parainfluenzae																				
Enterococcus faecalis																				
Streptococcus pseudopneumoniae																				
Streptococcus mitis																				
Cutibacterium acnes																				
Lactobacillus iners																				
Serratia nematodiphila																				
Aerococcus urinae																				
Alloscardovia omnicolens																				
Mycoplasma ovipneumoniae																				
Streptococcus pneumoniae																				
Anaerococcus obesiensis																				
Anaerococcus mediterraneensis																				
Fusobacterium nucleatum																				
Ezakiella massiliensis																				
Gemella haemolysans																				
Porphyromonas asaccharolytica																				
Corynebacterium striatum																				
Campylobacter hominis																				
Gardnerella vaginalis																				
Pseudomonas stutzeri																				
Prevotella jejuni																				
Bifidobacterium longum																				
Neisseria elongata																				
Streptococcus constellatus																				
Streptococcus anginosus																				
Serratia surfactantfaciens																				
Lawsonella clevelandensis																				
Staphylococcus haemolyticus																				

Fig. 7 Comparison of urinary profiles from 18 healthy control male donors. Values are cells/ml Urine. 9/18 samples failed to record any species above this threshold

with interpreting datasets that approach the lowest practical detection limit of our method.

Previous reports suggest enrichment of *Corynebacterium* and *Streptococcus* species in male urine [4, 32, 33]. We confirm the presence of five *Streptococcus* species in more than one sample; *S. gwangjuense*, *S. mitis*, *S. periodonticum*, *S. pneumonia* and *S. pseudopneumoniae* as well as two species of *Corynebacterium*; *C. glucuronolyticum* and *C. tuberculostearicum* (Fig. 7).

Discussion

Assessing the accuracy and reproducibility of a new method for quantitative co-profiling of microbial communities is challenging. Our initial investigation therefore focused on profiling a commercially available synthetic mixture of species of known titres. We were able to show that our new tool can accurately identify both relative and absolute quantities of a mixture of multiple gram-negative, gram-positive and yeast species with high reproducibility and very little off-target signal. Furthermore, we demonstrated that our method has the sensitivity to quantitatively detect bacterial species to 1×10^3 cells/ml and has a defined and reproducible process associated 'kit-ome' typically equivalent below 1,000 total cells/sample.

Applying this new tool to re-investigate asymptomatic female urine, vaginal and male urinary microbiomes has reproduced many of the key observations made using 16S rRNA and short fragment metagenomic studies. These include confirmation that *Lactobacilli* spp. and *Gardnerella spp.* are highly enriched in female urogenital samples [4, 21, 25, 33, 34] and can be defined into community state types; that a high degree of commonality between urinary and vaginal microbial populations exits [21, 35–37]; that microbiome relationships can be dynamic across a menstrual cycle [22] and that *L. crispatus* and *G. vaginalis* [22, 38, 39] display a largely exclusive relationship that generally holds true in vivo. Confirmation of these key findings supports the validity of our new tool as a method for profiling complex microbial communities.

In addition the quantitative nature of our method allows us to demonstrate that average female urobiomes have two orders of magnitude more biomass/mL than males and that, of the 27 species identified to be common to urobiomes from both sexes, most were present at comparable cells/ml titres. Given the marked difference in total biomass between the sexes this highlights the majority contribution made by a subset of species that are heavily enriched or unique to female urogenital microbiomes including *Gardnerella* spp. *Fanyhessea vaginae* and six species of the genus *Lactobacillus*. In comparison, *Lactobacillus* species were almost completely absent in male urine samples that, in agreement with previous observations [32], are populated by low biomass of *Streptococci*; *S. gwangjuense*, *S. mitis*, *S. pneumonia* and *S. pseudopneumoniae* and members of the *Serratia* genus; *S. urielytica, S. marcescens* and *S. nematodiphila*. In line with previous reports [32], half of our

input asymptomatic male urine samples returned no detectable organisms other than those compatible with contribution from the process kit-ome. Hence, they may be considered practically sterile as assayed by our technique. The high frequency of this finding in our male samples raises the possibility that, in stark contrast to females, this may represent the normal status of asymptomatic male urine. It also displays the dynamic range of our method that is able to report across 6 log units between the highest (1.3×10^8 cells/swab, *Lactobacillus iners*, Female vaginal swab 54, Fig. 3) and lowest (1×10^2 cells/mL, *Pseudomonas stutzeri*, Male sample 14, Fig. 7) recorded individual values.

The use of a full metagenomic workflow allows our method to provide automated full consensus genomes computed directly from sample sequencing data. Using this facility were are able to show that *Lactobacillus crispatus* strains have donor-specific SNP patterns. Furthermore, these patterns are most closely shared by vaginal and urinary strains of the same individual underlining the previously described tight relationship between female urogenital microbiomes [25]. Automatic differentiation of bacterial strains at genome wide SNP level directly from urine or swab samples will prove useful as both a research and clinical tool in, for example, identifying pathogenic sub-strains of bacterial species and monitoring the epidemiology of infection.

Choice of analytical technique is central to the design of a successful microbiome profiling study. PCR-based methods including 16S rRNA and short fragment shotgun sequencing profiling have high innate sensitivity but can be prone to report false-positives in low biomass samples [40]. In response, our system has been designed to be completely PCR and culture free to minimise biasing issues whilst maintaining a level of sensitivity that allows practical measurement of organisms down to 1,000 cells/mL. This arbitrary and deliberately conservative level is defined by sequencing depth and bioinformatic confidence thresholds both of which can be modulated to increase sensitivity further but with an associated risk of false positive prediction. Our method's innate sensitivity limitations remain an important consideration in experimental design and choice of assay.

Our method is restricted in its scope to make molecular identifications through it's database of species level genomic references. However, unlike 16S RNA or ITS based methods our technique is able to identify any organism with a DNA based genome so is able to provide holistic single sample co-profiles that include entities as diverse as Bacteria, Yeast, Viruses, Plasmids and Phages. The breadth and quality of organism genome references are both publically available and constantly improving. Thus, periodic updates to the database used by our method will continue to expand it's ability to identify a wider selection of organisms and allow the reanalysis of historic datasets as new species of interest become available.

Conclusions

We set out to develop a low bias, fully-quantitative metagenomic profiling tool for analysis of microbiomes directly from biological samples and clinical isolates. The asymptomatic urogenital profiling data presented here reproduces many of the key findings previously observed using more established molecular technologies. This is our first step towards validating our new method for use in research applications. Ultimately our aim is to apply this method to leverage the power of molecular analysis to improve pathogen identification in cases of clinical infection such as chronic UTI and vaginoses.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12894-025-01723-9.

Supplementary Material 1.
Supplementary Material 2.
Supplementary Material 3.
Supplementary Material 4.
Supplementary Material 5.
Supplementary Material 6.

Acknowledgements

The authors would like to acknowledge Prof. A Wolfe (Loyola University) and Melissa Kramer (Live UTI Free) for useful discussion in the preparation of this manuscript.

Authors' contributions

The study was devised by MDF and NJP and project managed by NJP. Experiments were performed by BF, MJR, RS and JQ. Analysis and sample tracking code was written by SM, CC, SW. Data analysis was undertaken by SM and NJP. The manuscript was written by NJP and MDF. All authors reviewed the manuscript.

Funding

BF, MJR, RS, JQ, SM, CC, SW, MDF & NJP were funded by the Fischer Family Foundation, FFF.

Data availability

The sequencing reads datasets generated during and/or analysed during the current study are available in the ncbi biosample repository, http://www.ncbi.nlm.nih.gov/bioproject/1034292: bioproject id: prjna1034292, biosample accessions samn38056575-samn38056714 (140 files).

Declarations

Ethics approval and consent to participate

The collection and use of anonymised human urine and vaginal swab sample from healthy controls for this study was approved via Systems Biology Laboratory's local ethical review board. Informed consent was given by volunteers prior to sample collection and analysis.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 23 July 2024 Accepted: 21 February 2025
Published online: 28 February 2025

References

1. Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. Nat Rev Microbiol. 2015;13(5):269–84.
2. Reynoso-Garcia J, Miranda-Santiago AE, Melendez-Vazquez NM, Acosta-Pagan K, Sanchez-Rosado M, Diaz-Rivera J, Rosado-Quinones AM, Acevedo-Marquez L, Cruz-Roldan L, Tosado-Rodriguez EL et al: A complete guide to human microbiomes: Body niches, transmission, development, dysbiosis, and restoration. Front Syst Biol. 2022;2. https://doi.org/10.3389/fsysb.2022.951403.
3. Albert AY, Chaban B, Wagner EC, Schellenberg JJ, Links MG, van Schalkwyk J, Reid G, Hemmingsen SM, Hill JE, Money D, et al. A Study of the Vaginal Microbiome in Healthy Canadian Women Utilizing cpn60-Based Molecular Profiling Reveals Distinct Gardnerella Subgroup Community State Types. PLoS ONE. 2015;10(8): e0135620.
4. Fouts DE, Pieper R, Szpakowski S, Pohl H, Knoblach S, Suh MJ, Huang ST, Ljungberg I, Sprague BM, Lucas SK, 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. 2012;10:174.
5. Kullar R, Chisari E, Snyder J, Cooper C, Parvizi J, Sniffen J. Next-Generation Sequencing Supports Targeted Antibiotic Treatment for Culture Negative Orthopedic Infections. Clin Infect Dis. 2023;76(2):359–64.
6. Perez-Carrasco V, Soriano-Lerma A, Soriano M, Gutierrez-Fernandez J, Garcia-Salcedo JA. Urinary Microbiome: Yin and Yang of the Urinary Tract. Front Cell Infect Microbiol. 2021;11: 617002.
7. Wensel CR, Pluznick JL, Salzberg SL, Sears CL: Next-generation sequencing: insights to advance clinical investigations of the microbiome. J Clin Invest. 2022;132(7). https://doi.org/10.1172/JCI154944.
8. Greninger AL. The challenge of diagnostic metagenomics. Expert Rev Mol Diagn. 2018;18(7):605–15.
9. Heytens S, De Sutter A, Coorevits L, Cools P, Boelens J, Van Simaey L, Christiaens T, Vaneechoutte M, Claeys G. Women with symptoms of a urinary tract infection but a negative urine culture: PCR-based quantification of Escherichia coli suggests infection in most cases. Clin Microbiol Infect. 2017;23(9):647–52.
10. Mambatta AK, Jayarajan J, Rashme VL, Harini S, Menon S, Kuppusamy J. Reliability of dipstick assay in predicting urinary tract infection. J Family Med Prim Care. 2015;4(2):265–8.
11. Sathiananthamoorthy S, Malone-Lee J, Gill K, Tymon A, Nguyen TK, Gurung S, Collins L, Kupelian AS, Swamy S, Khasriya R et al: Reassessment of Routine Midstream Culture in Diagnosis of Urinary Tract Infection. J Clin Microbiol. 2019;57(3). https://doi.org/10.1128/jcm.01452-18.
12. Brubaker L, Chai TC, Horsley H, Khasriya R, Moreland RB, Wolfe AJ: Tarnished gold—the "standard" urine culture: reassessing the characteristics of a criterion standard for detecting urinary microbes. Frontiers in Urology. 2023;3. https://doi.org/10.3389/fruro.2023.1206046.
13. Price TK, Hilt EE, Dune TJ, Mueller ER, Wolfe AJ, Brubaker L. Urine trouble: should we think differently about UTI? Int Urogynecol J. 2018;29(2):205–10.
14. Hilt EE, McKinley K, Pearce MM, Rosenfeld AB, Zilliox MJ, Mueller ER, Brubaker L, Gai X, Wolfe AJ, Schreckenberger PC. Urine is not sterile: use of enhanced urine culture techniques to detect resident bacterial flora in the adult female bladder. J Clin Microbiol. 2014;52(3):871–6.
15. Coorevits L, Heytens S, Boelens J, Claeys G. The resident microflora of voided midstream urine of healthy controls: standard versus expanded urine culture protocols. Eur J Clin Microbiol Infect Dis. 2017;36(4):635–9.
16. Khasriya R, Sathiananthamoorthy S, Ismail S, Kelsey M, Wilson M, Rohn JL, Malone-Lee J. Spectrum of bacterial colonization associated with urothelial cells from patients with chronic lower urinary tract symptoms. J Clin Microbiol. 2013;51(7):2054–62.
17. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–10.
18. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol. 2014;15(11):524.
19. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49(W1):W293–6.
20. PHE: SMI B41: investigation of urine. Information on UK standards for microbiology investigations of urine. Public Health England; 2018. https://www.rcpath.org/static/de4a6639-b118-46ea-9d3b8a0be4014944/UK-SMI-B-41i87-January-2019-Investigationof-urine.pdf.
21. Pearce MM, Hilt EE, Rosenfeld AB, Zilliox MJ, Thomas-White K, Fok C, Kliethermes S, Schreckenberger PC, Brubaker L, Gai X, et al. The female urinary microbiome: a comparison of women with and without urgency urinary incontinence. mBio. 2014;5(4):e01283-01214.
22. Price TK, Wolff B, Halverson T, Limeira R, Brubaker L, Dong Q, Mueller ER, Wolfe AJ: Temporal Dynamics of the Adult Female Lower Urinary Tract Microbiota. mBio. 2020;11(2). https://doi.org/10.1128/mbio.00475-20.
23. Feehily C, Crosby D, Walsh CJ, Lawton EM, Higgins S, McAuliffe FM, Cotter PD. Shotgun sequencing of the vaginal microbiome reveals both a species and functional potential signature of preterm birth. NPJ Biofilms Microbiomes. 2020;6(1):50.
24. France MT, Ma B, Gajer P, Brown S, Humphrys MS, Holm JB, Waetjen LE, Brotman RM, Ravel J. VALENCIA: a nearest centroid classification method for vaginal microbial communities based on composition. Microbiome. 2020;8(1):166.
25. Thomas-White K, Forster SC, Kumar N, Van Kuiken M, Putonti C, Stares MD, Hilt EE, Price TK, Wolfe AJ, Lawley TD. Culturing of female bladder bacteria reveals an interconnected urogenital microbiota. Nat Commun. 2018;9(1):1557.
26. Farfour E, Vasse M, Vallee A. Oligella spp.: A systematic review on an uncommon urinary pathogen. Eur J Clin Microbiol Infect Dis. 2024;43(6):1037–50.
27. Mahlen SD, Clarridge JE 3rd. Site and clinical significance of Alloscardovia omnicolens and Bifidobacterium species isolated in the clinical laboratory. J Clin Microbiol. 2009;47(10):3289–93.
28. Brooks JP, Buck GA, Chen G, Diao L, Edwards DJ, Fettweis JM, Huzurbazar S, Rakitin A, Satten GA, Smirnova E, et al. Changes in vaginal community state types reflect major shifts in the microbiome. Microb Ecol Health Dis. 2017;28(1):1303265.
29. Pathak P, Trilligan C, Rapose A: Bifidobacterium--friend or foe? A case of urinary tract infection with Bifidobacterium species. BMJ Case Rep. 2014;2014. https://doi.org/10.1136/bcr-2014-205122.
30. Chaban B, Links MG, Jayaprakash TP, Wagner EC, Bourque DK, Lohn Z, Albert AY, van Schalkwyk J, Reid G, Hemmingsen SM, et al. Characterization of the vaginal microbiota of healthy Canadian women through the menstrual cycle. Microbiome. 2014;2:23.
31. Zhang Q, Zhang L, Ross P, Zhao J, Zhang H, Chen W: Comparative Genomics of Lactobacillus crispatus from the Gut and Vagina Reveals Genetic Diversity and Lifestyle Adaptation. Genes (Basel). 2020;11(4). https://doi.org/10.3390/genes11040360.
32. Bajic P, Van Kuiken ME, Burge BK, Kirshenbaum EJ, Joyce CJ, Wolfe AJ, Branch JD, Bresler L, Farooq AV. Male Bladder Microbiome Relates to Lower Urinary Tract Symptoms. Eur Urol Focus. 2020;6(2):376–82.
33. Modena BD, Milam R, Harrison F, Cheeseman JA, Abecassis MM, Friedewald JJ, Kirk AD, Salomon DR. Changes in Urinary Microbiome Populations Correlate in Kidney Transplants With Interstitial Fibrosis and Tubular Atrophy Documented in Early Surveillance Biopsies. Am J Transplant. 2017;17(3):712–23.
34. Pearce MM, Zilliox MJ, Rosenfeld AB, Thomas-White KJ, Richter HE, Nager CW, Visco AG, Nygaard IE, Barber MD, Schaffer J, et al. The female urinary microbiome in urgency urinary incontinence. Am J Obstet Gynecol. 2015;213(3):347 e341-311.
35. Gottschick C, Deng ZL, Vital M, Masur C, Abels C, Pieper DH, Wagner-Dobler I. The urinary microbiota of men and women and its changes in women during bacterial vaginosis and antibiotic treatment. Microbiome. 2017;5(1):99.

36. Pohl HG, Groah SL, Perez-Losada M, Ljungberg I, Sprague BM, Chandal N, Caldovic L, Hsieh M. The Urine Microbiome of Healthy Men and Women Differs by Urine Collection Method. Int Neurourol J. 2020;24(1):41–51.
37. Wolfe AJ, Toh E, Shibata N, Rong R, Kenton K, Fitzgerald M, Mueller ER, Schreckenberger P, Dong Q, Nelson DE, et al. Evidence of uncultivated bacteria in the adult female bladder. J Clin Microbiol. 2012;50(4):1376–83.
38. Castro J, Henriques A, Machado A, Henriques M, Jefferson KK, Cerca N. Reciprocal interference between Lactobacillus spp. and Gardnerella vaginalis on initial adherence to epithelial cells. Int J Med Sci. 2013;10(9):1193–8.
39. Ojala T, Kankainen M, Castro J, Cerca N, Edelman S, Westerlund-Wikstrom B, Paulin L, Holm L, Auvinen P. Comparative genomics of Lactobacillus crispatus suggests novel mechanisms for the competitive exclusion of Gardnerella vaginalis. BMC Genomics. 2014;15:1070.
40. Kennedy KM, de Goffau MC, Perez-Munoz ME, Arrieta MC, Backhed F, Bork P, Braun T, Bushman FD, Dore J, de Vos WM, et al. Questioning the fetal microbiome illustrates pitfalls of low-biomass microbial studies. Nature. 2023;613(7945):639–49.

Publisher's Note