Structure and regulation of the cervicovaginal microbiome in a cohort of Afro-Caribbean women

Orville St. E. Roachford¹*, Angela T. Alleyne¹, Claire Kuelbs³, Manolito G. Torralba², and Karen E. Nelson²,³

¹Department of Biological and Chemical Sciences, The University of the West Indies, Cave Hill Campus, Bridgetown, BB 11000, BARBADOS.

²J. Craig Venter Institute, 9605 Medical Center Drive Suite 150, Rockville, MD 20850, USA.

³J. Craig Venter Institute, 4120 Capricorn Lane, La Jolla, CA 92037, USA.

*Corresponding address: Department of Biological and Chemical Sciences, The University of the West Indies, Cave Hill Campus, Bridgetown, BB 11000, BARBADOS.

Telephone: +1-246-417-4808, Fax: +1-246-417-4325

E-mails: orville.roachford@mycavehill.uwi.edu; oster6413@gmail.com

ORCID ID: 0000-0003-1389-0797

Running title: Regulation of cervicovaginal microbiome
Abstract

Background

The cervicovaginal microbiome, with its well documented dynamic community state types (CSTs I-V) is known to interact with the host immune system in a complex regulatory ecosystem that guards against dysbiosis and invading pathogens. However, the regulatory mechanism of the cervicovaginal microbiome in vaginal eubiosis is not well understood. Therefore, we characterized the cervicovaginal microbiome in a cohort of Afro-Caribbean women to the species-strain taxonomic level using high-throughput targeted sequencing and shotgun metagenomics and conceptualized a hypothetical model that aids in understanding vaginal ecosystem regulation.

Results

Compared to targeted 16S rRNA V4 sequencing, whole genome shotgun (WGS) metagenomics offered greater species resolution of the cervicovaginal microbiome in the Afro-Caribbean women cohort. Community state type IV predominated in the microbiome of these women, with Prevotella (13.91%), Gardnerella (12.14%) and Lactobacillus (9.37%) being the three most abundant genera. For Prevotella and Lactobacillus the most abundant species-strains were P. timonensis DSM 22865 (5.0%) and L. iners DSM 13335 (7.0%), respectively. The less virulent strain of Gardnerella, G. vaginalis 409-05 (8.0%) was more abundant than G. vaginalis ATCC 14019 (4.0%). In the resistome, 2,753 antimicrobial resistance (AMR) genes consisting of 28 types (mostly tet and Emr; abundances 51% and 15%, respectively) which confer resistance to tetracyclines and the macrolide-lincosamide streptogramin B (MLSb phenotype) group were observed. Functional profiling showed a high abundance of biological processes (bacterial-type
flagellum-dependent cell motility, cell adhesion, response to biotic stimulus and quorum sensing) associated with biofilm activity.

Conclusions

Characterized for the first time, the cervicovaginal microbiome in Afro-Caribbean women is predominantly CST IV with the three most abundant taxa consisting of bacterial strains P. timonensis DSM 22865, G. vaginalis 409-05 and L. iners DSM 13335. Its resistome had multiple AMR genes that confer resistance to antibiotics commonly used in the treatment of sexually transmitted infections (STIs) and bacterial vaginosis (BV). In this study, the occurrence of biofilm activity within the cervicovaginal microbiome suggested a possible regulatory role. We present a conceptual immuno-munibiome model that advances a mechanistic approach for the structure and regulation of the vaginal ecosystem in the Afro-Caribbean women cohort.
Background

The external orifice (vaginal introitus) of the female reproductive tract communicates directly with another internal cavity (abdominopelvic / peritoneal cavity). Pathogens entering the lower reproductive tract (vagina) can potentially penetrate the cervical canal, enter the uterus, travel along the fallopian tubes into the abdominopelvic cavity and cause diseases such as vaginitis, cervicitis, chorioamnionitis, preterm birth, pelvic inflammatory disease, infertility, peritonitis, sepsis and even death [1, 2, 3].

Two immune defense strategies within the cervicovaginal tract have evolved along with natural physical barriers (cervical mucus plug and vaginal mucoid film) to guard against pathogens; the innate immune system (cytokines and chemokines secretion; activation of leukocytes and the complement cascade) and the adaptive immune system (antigen presentation / T-cell response) [4, 5, 6]. Presumably, vaginal microbes evolved concomitantly, entering a symbiotic interrelationship initially to survive the host’s hostile immune system and later into a unique mutualistic host-microbe relationship, giving rise to the existent cervicovaginal microbiome.

Evolutionarily, the fundamental goal of the cervicovaginal microbiome is survival and persistence; achieved by protecting its host, and itself in the presence of adverse intrinsic (e.g. pregnancy, aging, menstruation, obesity) and extrinsic (hormones, drugs, antibiotics, probiotics, coitus, douching, exogenous pathogens) factors. Two hypotheses have been proposed for the evolution of Lactobacillus-dominant cervicovaginal microbiomes involvement in host protection. The first being the “disease risk hypothesis” in which the cervicovaginal microbiome protects against sexually transmitted infections (STIs) that may negatively impact the fitness of its host (and host’s offspring) and the second the “obstetric protection hypothesis” which guards against the high risk
of pregnancy and childbirth complications due to microbial pathogens penetrating the endocervical canal [7]. This follows on the long held premise that Lactobacillus species are the principal protective organisms responsible for preventing dysmicrobism and maintaining vaginal health [8, 9]. With regards to persistence and self-protection from adverse local environmental conditions, cervicovaginal microbes perhaps have accomplished such by virtue of niche differentiation, biofilm formation and acquisition of antimicrobial resistance genes [10, 11, 12].

Contrary to past dogma, recent research has reported normal vaginal health in some women whose vaginal microbiomes consist of a low abundance (or depletion) of Lactobacillus species, high abundance of obligate anaerobes, G. vaginalis or known pathogens [13, 14, 15, 16, 17]. In addition, overgrowth of vaginal Lactobacillus is known to cause cytolytic vaginosis characterised by vaginal irritation and abnormal discharge [18].

Deep sequencing of the female cervicovaginal tract has shown that the latter harbours an extremely diverse microbiome whose composition is both temporal and individual-specific [19, 20]. However, among women of varying ethnicity and age, and regardless of the number of women under study, this diversity seems to have compositional and structural constraints [13, 16, 21, 22, 23, 24, 25]. Hence, the recognition of five distinct community state types (CSTs): I, II, III, IV and V, where CST I is predominantly of Lactobacillus crispatus, and CSTs II, III and V dominated by L. gasseri, L. iners and L. jensenii, respectively. CST IV is described as a heterogeneous group dominated by a high abundance of anaerobes such as Prevotella, Dialister, Atopobium, Gardnerella, Megasphaera, Peptoniphilus, Sneathia, Eggerthela, Aerococcus, Finegoldia and Mobiluncus [17, 19]. Therefore, regulation of the cervicovaginal microbial composition with its multitude of symbiotic interrelationships and host-microbe interactions is complex. Additionally, disease causation by pathobionts and exogenous pathogens remains poorly understood. An in-
depth assessment of the characteristics (microbial function, composition and structure) of the vaginal ecosystem, its resistome and local immune inflammatory responses are of principal importance in understanding regulation of the cervicovaginal microbiome, vaginal health and disease prevention. The cervicovaginal microbiome has been well characterized using 16S rRNA next-generation sequencing (NGS) technologies for women of the Americas (North, Central and South), Europe, Asia and Africa but not for Caribbean women to date. Consequently, the aim of this study was to fully characterize the cervicovaginal microbiome of Afro-Caribbean women using 16S rRNA V4 and whole genome shotgun metagenomics. In addition, we sought to introduce a conceptual hypothesis for the regulation and structure of the cervicovaginal microbiome.

Materials and Methods

Participants and sample collection

Eighteen women (selected via simple random sampling) participated in the study between February and April 2018 following signed informed consent. Their health status was determined in a general clinic. All participants were Afro-Caribbean women between the ages of 18-74, not pregnant and without any physical disabilities, concomitant microbial infections or abnormal vaginal symptoms. For participation in this study, all women were required to refrain from vaginal douching (1 week prior to sampling), avoid antibiotic therapy (6 weeks prior to sampling), not engage in coitus and have completed menses by at least 3 days prior to sampling.

A cervicovaginal lavage procedure was carried out on each woman after taking a brief medical history and giving instructions for the sampling procedure. The lavage procedure involved the passage of a stainless-steel Cusco speculum into the vagina to visualize the cervix and related structures. Twelve millilitres (12 ml) of sterile phosphate buffer solution (PBS) with pH 7.4 was
syringed into the vagina and re-extracted after 1 minute by way of disposable sterile syringes. Approximately 10 ml of the cervicovaginal washings (CVWs) were transferred to 15 ml Falcon tubes, placed on ice and transferred directly for storage at -80°C within 1 hour of collection.

**Genomic DNA Extraction**

CVWs were thawed and 1.0 ml aliquots were placed into sterile 1.5ml Eppendorf microcentrifuge tubes stored on ice. The samples were centrifuged at room temperature for 6 min at 13, 500 rpm. Supernatant was removed, and another 1.0 ml of the CVW sample added to the pellet for re-centrifugation at 13,500 rpm for 6 min. This procedure was repeated thrice, resulting in the total use of 5.0 ml of CVW. Pellets were stored at -20°C prior to further processing. DNA was extracted from pellets using InstaGene™ Matrix (Bio-Rad Laboratories, USA). InstaGene™ Matrix (200 µl) was added to the pellets and incubated at 56°C for 30 min. The samples were vortexed at high speed for 10s and then placed in a boiling water bath for 8 min. Samples were vortexed again for 10s and centrifuged at 12,000 rpm for 5 min. The supernatant was removed and stored at -20°C prior to shipping to J. Craig Venter Institute (JCVI), California, USA for Illumina sequencing. Qualification and quantification of extracted DNA was determined by 2.0% gel electrophoresis and nanodrop spectrophotometry (NanoDrop® Technologies, Wilmington, DE, USA) at 260nm.

**16S rDNA V4 sequencing by MiSeq Illumina**

DNA extracted from the cervicovaginal samples was amplified using custom adaptor ligated primers (primers 515F 5′- GTGCCAGCMGCGGTAA-3′ and 806R 5′-GGACTACHVGGGTWTCTAAT-3′) that target the V4 hypervariable region of the 16S rRNA gene [26]. Included in the custom primers, were i5 and i7 adaptor sequences for Illumina MiSeq pyrosequencing. Unique 8 bp indices were also attached to the primers giving each sample its own
unique barcode pair. Incorporation of adaptors and index sequences onto the primers at the PCR stage minimalized the loss of sequence data in comparison to earlier methods that ligated the adaptors to each amplicon post-amplification [27]. All sequence reads generated by this method were of the same 5’-3’ orientation. Extracted DNA (approximately 100ng/µl) was used to generate amplicons using Platinum Taq polymerase (Life Technologies, CA, USA) under the following cycling conditions: 95°C for 5min for an initial denaturing step, 95°C for 30 s, 57°C for 30 s, 72°C for 30 s for 35 cycles, followed by a final elongation step at 72°C for 7 min and stored at 4°C. Amplicon purification was performed using the QIAquick PCR purification kit (Qiagen, Valencia, CA, USA), quantified using Tecan fluorometric methodology (Tecan Group, Mannedorf, Switzerland), normalized and then pooled in preparation for paired-end Illumina MiSeq sequencing using V2 chemistry dual index 2x250 bp format (Roche, Branford, CT, USA) as per the manufacturer’s protocol.

**WGS metagenomic library construction and sequencing**

Sample DNA was assessed for high quality and selected for high molecular weight in preparation for library construction using agarose gel electrophoresis. The sample DNA was then sheared using the Covaris ultrasonicator instrument (Covaris, Woburn, Massachusetts, USA). Sheared DNA was quantified, and size selected using NEBNext® Library Quant Kit for Illumina® (New England Biolabs, Ipswich, MA, USA) according to manufacturers’ specifications. Sample DNA was barcoded to allow for multiplexing of several samples per sequencing run using NexteraXT Index kit (Illumina, San Diego, CA, USA). Metagenomic libraries were sequenced using Illumina NextSeq 500 following standard manufacturer’s specifications. Approximately 1-3 Gbp of sequence data per sample was generated to obtain significant functional information for comparative analyses.
Biobinformatics analysis of 16S and WGS metagenomic dataset

Quality control checks on raw Illumina sequence data derived from WGS-metagenomic sequencing was carried out using FastQC v 0.10.1 [28]. All metagenomic sequencing data was processed and analyzed using CLC Microbial Genomics Pro Suite (Qiagen Bioinformatics, Redwood City, CA, USA) with default settings and the contigLCA algorithm within MG-RAST v4.0.3 [29]. Sequences with high Phred scores (Q >20) were demultiplexed, trimmed of primers and adaptors, as well as filtered of short reads and chimeras. Host sequences were mapped against the reference human genome (GRCh38) and removed. Sequence reads were binned into operational taxonomic units (OTUs) defined at 99% similarity and matched against 16S rRNA sequences within Greengenes v13_8 99% database. OTUs were also assigned to taxa at the genus and species level using SeqMatch version 3 in the online Ribosomal Database Project (RDP: release 11_5) [30]. Species diversity within samples and between samples were measured using alpha diversity metrics (Chao 1 bias-corrected, Simpson’s index and Shannon entropy) and beta diversity metrics (Bray-Curtis, Jaccard, Euclidean, weighted and unweighted UniFrac), respectively. Beta diversity was visualized with three-dimensional (3D) principal coordinates analysis (PCoA) within CLC Microbial Genomics Module.

Statistical, functional and antibiotic resistance analysis of the WGS metagenomic dataset

Statistical analyses were performed using PERMANOVA (permutational multivariate analysis of variance, also known as a non-parametric MANOVA) [31] in CLC Genomics Workbench v20 (https://digitalinsights.qiagen.com). Reads from whole metagenome sequencing dataset were assembled into contigs and annotated using the De Novo Assembled Metagenome and Find Prokaryote Genes tools in CLC Microbial Genomics Module (Qiagen Bioinformatics, Redwood City, CA, USA), respectively. All annotated coding sequences (CDS) with protein families (Pfam)
and Gene Ontology (GO) terms were searched against Pfam v32 and GO databases v1.2 [32, 33, 34]. Antimicrobial resistance (AMR) genes within the whole genome shotgun metagenomic dataset of the Afro-Caribbean women were identified using Find Resistance with ShortBRED (FRSB) tool, similar to ShortBRED [35]. DIAMOND v0.9.31 within the FRSB tool (CLC Microbial Genomics Module) was used to match queried sequences against the QMI-AR Peptide Marker database (released 2019-11). All default settings in the FRSB tool were retained except for the parameter “more sensitive search” in order to run DIAMOND at its highest sensitive mode. AMR genes were cross-checked against entries in Comprehensive Antibiotic Resistance Database (CARD) [36].

**Results**

**Read and OTU Statistics**

Following quality control and removal of human host reads, Illumina NextSeq WGS metagenome sequencing of the bacterial DNA from 18 cervicovaginal samples yielded a total of 181,820,742 sequences. Data from 3 samples were unsuitable for metagenomic analysis. The average number of high-quality sequences (Average Phred quality score 34) was 12,121,383 (ranging from 4,637,282 – 26,540,452) per sample. The average % GC content and sequence length were 44% and 151bp, respectively (Supplementary Table S1). For comparison, Illumina MiSeq sequencing of the 16S rRNA V4 hypervariable region was also used to determine the composition of 18 cervicovaginal samples of the Afro-Caribbean cohort studied. After filtration of host reads and removal of chimeric sequences, the 18 samples produced a total of 361,723 read pairs in operational taxonomic units (OTUs) with an average read length of 251bp (Supplementary Table S2).
Taxonomic allocation of OTUs

Removal of short reads from the 361,723 read pairs resulted in an input dataset of 99,322 reads which when filtered yielded 2,141 OTUs. The latter were matched (99% similarity) against Greengenes v13_8 99% (16S rRNA) database and clustered into 468 OTUs of which 46 were de novo. Each of the 468 OTUs was assigned to a given taxon, yielding 329 taxa at the genus level. These were further resolved into 236 species based on the highest similarity score (S_ab) as determined by SeqMatch version 3 in the online Ribosomal Database Project (RDP: release 11_5), shown in Supplementary Tables S3 and S4. Of the 236 species, 50% had matches with similarity scores greater than 0.970. In addition, 60% of all species were identified as part of human vagina microbiota while 25%, 10% and 5% were of gut, oral and non-human (animal, soil, water, sludge etc.) microbial communities, respectively. At the kingdom level, the 15 WGS metagenomes analysed by the online MG-RAST database were on average composed of 86.70 % bacteria, 13.24% eukaryotes, 0.04% viruses and 0.02% archaea. The contigLCA algorithm within MG-RAST identified DNA sequences from the sample that shared similarity to sequences at the genus level from humans and Asian monkeys (genus Macaca), ranging from 5.81-66.14% and 0.94-12.75%, respectively (https://www.mg-rast.org/linkin.cgi?project=mgp88579). Due to its inability to filter human and Macaca DNA sequences, MG-RAST was not used for further metagenomic analyses. In metagenomic analyses performed with Microbial Genomics Module (CLC genomics workbench 12) human host reads were completely removed and the average composition of the WGS reads (kingdom level) for bacteria and archaea were 99.88% and 0.12%, respectively. In subsequent analyses, only the bacterial component was used (Supplementary Table S5).

Comparison of 16S rRNA and WGS metagenomic analyses
WGS metagenomic analysis was successful for 15 of the 18 cervicovaginal samples compared to 16S rRNA V4 sequencing which was successful for all 18 samples. This reflects the ability of the latter sequencing technique to amplify 16S rRNA V4 within samples of limited bacterial DNA concentration. In contrast to samples sequenced using 16S rRNA V4, samples sequenced by WGS metagenomics were noted to have a high abundance of *Tetrasphaera japonica* T1-X7. The next most abundant species were of the genera *Lactobacillus, Prevotella, Gardnerella, Anaerococcus, Megasphaera, Shuttleworthia, Sneathia* and *Porphyromonas*. However, the cervicovaginal microbial communities analysed in the absence of *T. japonica* T1-X7 as determined by 16S rRNA V4 and WGS were similar (Fig. 1). Taxa identified by 16S rRNA V4 sequences were comparable to that by WGS, with discrepancies occurring mainly at species-level. WGS sequencing offered greater and more accurate resolution at the species-level. Taxonomically, WGS was better able to resolve species to strain-level by identifying 301 species (211 strains) compared to 236 (3 strains) identified by 16S rRNA V4 sequencing.

**Microbial diversity and composition of the cervicovaginal microbiomes based on 16S rRNA V4 sequencing**

The alpha diversity (which expresses the number of species represented within a given cervicovaginal microbiome) as measured by chao-1 bias-corrected, Shannon entropy and the Simpson’s index metrics showed a wide range of heterogeneity (species diversity) within the 18 cervicovaginal microbiomes. Generally, there was no distinctive clustering of the cervicovaginal microbiomes based on any of the three metrics used (Fig. 2A and Supplementary Figure S1). Similarly, when microbial diversity between the 18 cervicovaginal microbiomes were compared using five different β- diversity metrics (Bray-Curtis, Jaccard, Euclidean, weighted and unweighted UniFrac) and visualized by three-dimensional (3D) Principal Coordinates Analysis
(PCoA), no significant clustering of the cervicovaginal communities was observed (Figs. 2B, 2C and Supplementary Figure S1).

Taxa with a relative abundance as low as 5.53e-4 % of the microbial community were detected by 16S rRNA V4 sequencing. The most abundant taxa with coverage greater than 0.35 % are shown in the heatmap in Fig. 3 which also shows the diversity of these taxa in each cervicovaginal microbiome. These taxa are of the genera *Lactobacillus, Prevotella, Gardnerella, Anaerococcus, Megasphaera, Sneathia, Dialister, Shuttleworthia, Porphyromonas, Veillonella* and *Howardella.*

For the microbiome data, normalised taxon abundances were clustered by Euclidean distance. *Prevotella, Lactobacillus* and *Gardnerella* were present in all women. However, *Prevotella* was the most abundant (mean abundance 23.06%, maximum 71% when dominant), followed by *Lactobacillus* (mean abundance 22.26%, maximum 98% when dominant). For *Gardnerella,* the mean abundance and maximum abundance when dominant within a microbiome were 10% and 63%, respectively. *Shuttleworthia* was present in 94% of women and had a mean abundance of 14% with a maximum abundance of 83% when dominant. It should be noted *Shuttleworthia* shares 89% of its 16S rRNA gene V4 sequence with that of bacterial vaginosis-associated bacterium 1 (BVAB1) [37]. All other genera had mean abundances ≤5% and were present in 11% to 78% of all women. Among the 18 women, the most dominant *Lactobacillus* species was *L. iners* (19.19% mean abundance). Less abundant *Lactobacillus* species present across the cervicovaginal microbiomes were *L. crispatus, L. jensenii, L. fornicalis, L. vaginalis, L. rhamnosus, L. coleohominis, L. intestinalis and L. kitasatonis.* Also identified within the cervicovaginal microbiomes of the women were rarer (<0.05% mean abundance) *Lactobacillus species* such as *L. reuteri, L. helveticus, L. ultunensis, L. kalixensis and L. zaea* (Supplementary Tables S3 and S6).
Microbial diversity and composition of the cervicovaginal microbiomes based on shotgun whole genome (WGS) sequencing

Bray-Curtis, Jaccard and Euclidean-based Principal Coordinates Analyses (PCoA) did not reveal any distinct clustering of cervicovaginal communities with respect to the health status of the 15 women. Euclidean-based PCoA showed a cluster which consisted of cervicovaginal communities from both cohorts, however, the significance of this is not clear. In 40% of the cervicovaginal samples analyzed by WGS metagenomics, a large fraction of species remained undiscovered as reflected by the absence of plateaus in Chao-1 corrected, Shannon entropy and Simpson’s index-based alpha diversity measures (Supplementary Figure S1). However, WGS metagenomics identified a much larger number of species per sample compared to 16S rRNA V4 sequencing.

*T. japonica* T1-X7, a polyphosphate accumulating Gram-positive cocci, was the predominant organism found in the 15 cervicovaginal samples following WGS metagenomic analysis but this is believed to be a contaminant, possibly related to the previous run of the Illumina sequencer on sludge samples. *Tetrasphaera* are normally isolated from activated sludge in wastewater treatment systems with enhanced biological phosphorus removal (EBPR) [38, 39] though a species, *Tetrasphaera remsis*, has reportedly been isolated from air (40). Not considering *T. japonica* T1-X7, and classifying percentage relative abundance in accordance with Ranjan et al. [41], the most abundant taxa (i.e. taxa with abundances > 1.00%) were *Prevotella* spp. (13.91%), *Gardnerella* spp. (12.14%), *Lactobacillus* spp. (9.37%), *Mobiluncus* spp. (7.15%), *Sneathia* spp. (5.59%), *Megasphaera* genomosp (3.00%), *Atopobium* spp. (2.31%), *Bifidobacterium breve* (2.00%), *Mageeibacillus indolicus* (2.00%) and *Porphyromonas asaccharolytica* (2.00%). Other taxa with low (0.01 – 0.49%), moderate (0.05 – 0.49%) and high (>0.50–0.99%) abundances are featured in Table 1 (Supplementary Tables S5 and S6). A greater number of species/strains were identified
for \textit{Prevotella} (36/34) than for \textit{Lactobacillus} (22/15) or any other genera. The most common \textit{Prevotella} species/strains were \textit{P. timonensis} 4401737, \textit{P. amnii} DSM 23384, \textit{P. bivia} DSM 20514 and \textit{P. ihumii} with mean abundances of 5.00\%, 3.00\%, 2.00\% and 1.00\%, respectively. The four most abundant \textit{Lactobacillus} species/strains identified were \textit{L. iners} DSM 13335, \textit{L. crispatus} ST1, \textit{L. gasseri} and \textit{L. jensenii} with mean abundances of 7.00\%, 2.00\%, 0.14\% and 0.08\%, respectively (supplementary Table S5). Of note, was the identification of \textit{Mycoplasma} species-strains; \textit{M. hominis} ATCC 23144, \textit{M. gallisepticum} str. R, \textit{M. flocculare} ATCC 27399 and \textit{M. alligatoris} A21JP2. \textit{M. hominis} ATCC 23144, the only human associated \textit{Mycoplasma} identified, had a mean abundance of 0.12\%. Based on the metadata (Supplementary Table S7), 47\% of the participants in the study were noted to have systemic lupus erythematosus (SLE). Hence, we compared the cervicovaginal microbial compositions of women with and without SLE by performing a PERMANOVA analysis [31]. The analysis for the 3 most abundant species/strains (\textit{Prevotella}, \textit{Gardnerella}, \textit{Lactobacillus}) and \textit{Mycoplasma} is detailed in Supplementary Table S8. Of \textit{Prevotella} species, only \textit{P. brevis} ATCC 19188, \textit{P. oryzae} DSM 17970 and \textit{P. paludivivens} DSM 17968 were more significantly abundant in the SLE group with FDR (false discovery rate) p-values 0.02, 0.02, 0.01 and Log$_2$ fold changes of 7.11, 7.11, 7.41, respectively. \textit{L. crispatus} ST1 (representing CST I) was significantly more abundant in the cervicovaginal microbiomes of women without SLE having FDR p-value 0.02 and Log$_2$ fold change of 6.72. While \textit{L. iners} DSM 13335 (CST III) was more abundant in the SLE group with FDR p-value 0.02 and a Log$_2$ fold change of 4.09. There were no significant differences in the abundancies of \textit{G. vaginalis} 409-05, \textit{G. vaginalis} ATCC 23114 and \textit{M. hominis} ATCC 23114 (FDR p-values 0.11, 0.82 and 0.17, respectively). Whether there is a meaningful correlation between \textit{Prevotella} species, \textit{L. iners} DSM
13 and SLE would require a large high-powered study due to the age disparity of the two groups, sampling size and associated confounding factors (such as parity and co-morbidities).

Function of the cervicovaginal microbiome

Functional analysis (genome annotation) of the cervicovaginal microbiome whole metagenome shotgun sequencing dataset using Pfam (Protein families) and Gene Ontology (GO) databases generated a set of 98 functional profiles for the 15 cervicovaginal samples with the following ontologies: biological processes (38), cellular components (25) and molecular functions (35). The biological functions which were visualized in the form of a heatmap (Supplementary Figure S2 and Table S9) highlighted the differential gene expressions and associated biological functions, in terms of relative abundance, within and amongst the cervicovaginal microbiomes of the Afro-Caribbean women. Of the biological processes, there was upregulation of cell adhesion, bacterial-type flagellum-dependent cell motility, sporulation, quorum sensing, response to biotic stimulus, carbohydrate binding, viral processes and host cell viral entry. Among the more abundant molecular functions were peroxidase and antioxidant activities, electron transfer activity, and amino acid and iron-sulphur cluster binding. Cell well, outer membrane and bacterial-type flagellum were some of the more abundant cellular components.

Antimicrobial resistance genes of the cervicovaginal microbiome

Among the cervicovaginal shotgun metagenomes of the 15 Afro-Caribbean women, 2,753 antimicrobial resistance (AMR) genes consisting of 28 types (ileS, IsaC, mel, tet32, tetM, tetO, tetQ, tetS, tetW, rpoB, CfxA2, catl, ErmA, ErmB, ErmF, ErmT, ErmX, emrB, emrK, yegN, MexW, patB, TolC, YojI, AcrS, cpxA, gadX and hns) were detected (Supplementary Table S10). This meta-resistome conferred resistance to a number of drug classes (tetracyclines, macrolides, lincosamides, penicillins, phenicols, aminoglycosides, quinolones, rifamycins and mupirocin).
The relative abundances of the 14 most abundant AMR genes are visualized in the area chart (Fig. 4). The most abundant drug class was tetracyclines (51%), followed by the macrolide-lincosamide streptogramin B (MLSb phenotype) group with a relative abundancy of 15%. This group consisted of common antibiotics such as azithromycin, clindamycin and nitroimidazoles. Antibiotic resistance genes tetM, tetQ and tetW were noted to be associated with conjugate and non-conjugate transposons. The tetO genes were associated with conjugative plasmids, genetic elements that are involved in horizontal gene transfer (HGT) between microbial species. Also, identified within the meta-resistome were CfxA2, a class A beta-lactamase AMR gene, found in Prevotella intermedia, and antimicrobial resistance genes ileS and rpoB found in the lactic acid producing Bifidobacterium species. Several mechanisms of antibiotic resistance were identified. The most common resistance mechanism found in this study was antibiotic target protection mediated by tetracycline-resistant ribosomal protection proteins (tet32, tetM, tetQ, tetO and tetW). Resistance genes tet32, tetM, tetO and tetW have been found in the Firmicutes (consisting of Gram-positive genera such as Lactobacillus, Streptococcus and Megasphaera. Whereas, the tetQ gene is found in Gram-negative bacteria such as Prevotella and Porphyromonas of the Bacteroidetes phylum [42, 43]. Two other AMR mechanisms were protein synthesis inhibition by mutational alterations in the erm genes (ErmA, ErmB, ErmF, ErmT and ErmX ) that encode Erm 23S ribosomal RNA methyltransferase and confer antibiotic resistance to the macrolide-lincosamide streptogramin B group, and the efflux pump mechanism in which antibiotics are transported out of the bacterial cells. For example, TolC an outer membrane efflux protein subunit of multidrug efflux complexes in Gram negative bacteria pumps aminoglycosides, penicillins, cephalosporins, fluoroquinolones and nitroimidazoles out of the cells. Other resistance genes within the cervicovaginal microbiomes
of the Afro-Caribbean women involving antibiotic efflux were *emrB, emrK, yegN, MexW, patB* and *YojI*.

**Discussion**

This research presents the first characterization and comparative metagenomic analysis of the cervicovaginal microbiomes of a group of Afro-Caribbean women based on high-throughput next generation sequencing (targeted 16S rRNA V4 sequencing and whole genome shotgun metagenomics).

**16S rRNA V4 sequencing vs whole genome shotgun metagenomics**

Based on our results, WGS metagenomic sequencing resulted in a greater depth of sequencing and produced greater phylogenetic resolution at the species/strain levels as seen in other studies [44, 45]. This is not surprising as the extent of microbial diversity determined by 16S rRNA V4 sequencing is limited by the availability of species-strain directed primers and the short sequence area (V4) covered [41, 46, 47]. In both cases, there was some misidentification of species. This is due to the occurrence of base calling errors, point errors in reads, artificial replicates, over-represented reads, GC skewing as well as inherent errors embedded in bioinformatic algorithms within databases such as Greengenes and RDP that were used for taxonomic classification. However, WGS was the better technique for studying the vaginal microbiome for three reasons. Firstly, it allowed for the detection of rare and unknown microbes that might play a significant role in vaginal dysbiosis and associated diseases. Secondly, it can offer more consistent phylogenetic resolution compared to partial 16S rRNA sequences, as the various hypervariable regions of the 16S rRNA gene have different evolutionary rates and can produce varied phylogenetic resolutions [48]. Thirdly, WGS better differentiated species to the strain level. This
is particularly important in understanding the cervicovaginal microbial interrelationships as species-specific strains display different microbial inhibitory effects. For example, L. gasseri strain KS123.1 has no inhibitory effect on G. vaginalis and P. bivia, whereas L. gasseri strain KS120.1 does [49].

Cervicovaginal microbiome of Afro-Caribbean women

The composition of the cervicovaginal microbiomes of the group of Afro-Caribbean women was found to be individual-specific regardless of the sequencing technique utilized (16S and WGS). This specificity is in keeping with research findings from other studies involving large numbers of women [16, 24]. The cervicovaginal microbiomes showed a wide range of microbial richness within samples (Alpha diversity) and the level of diversity was comparable among the women (Beta diversity). Species diversity was more pronounced with WGS metagenomic sequencing. About 25% of the bacteria identified in the cervicovaginal samples were associated with the human gut. Previous studies have documented microbes that are common to both microbiomes, with the gut being a likely source for vaginal colonization [16, 50]. Using 16S rRNA V4 sequencing, 4 out of the 5 CSTs previously defined by Ravel et al. [17], were recognised: CST’s I (L. crispatus), III (L. iners), IV (heterogenous microbes) and V (L. jensenii). Whereas for WGS metagenomic sequencing, only CSTs I, III and IV were observed (Heatmap, Fig. 5). CST II which is defined by a predominance of L. gasseri was not identified by either sequencing technique. This might be due to the small sampling size. L. fornicalis was identified in CST IV in low abundance. CST IV was the most common microbial structure for Afro-Caribbean women, and was dominated by Prevotella species (P. timenosis, P. amnii, P. histolitica, P. bivia, P. ihumii and P. bergensis). Other genera of lower abundance within this community state type were Gardnerella, Sneathia, Megasphaera and Anaerococcus (Figs. 2 and 3). This finding is in accordance with previous
studies that show CST IV to be diverse and more common in African Americans [17, 20, 51, 52]. In general, based on both sequencing techniques used, the relative abundances of *Prevotella*, *Lactobacillus* and *Gardnerella* dominating the cervicovaginal microbiomes of Afro-Caribbean women were 23.06%, 22.26%, 10% and 13.91%, 9.37%, 12.14% for 16Sr RNA V4 and WGS, respectively. The high abundance of *Prevotella* found in the Afro-Caribbean cohort has also been reported for African American women though not as the dominant genus [17]. In contrast, native West African women have been noted to have very low abundances (0.08-1.2%) of *Prevotella* [16]. Given that both Afro-Caribbean and African American women are descendants of West Africans, one might have expected these ethnic groups to have similar abundances of *Prevotella*. However, this was not the case among this Afro-Caribbean cohort. Therefore, in addition to genetic factors [53], it is likely that cervicovaginal microbial communities are also influenced by geophysical environments and sub-culture factors acting upon host gene expression (analogous to the impact of diet on foetal/neonatal gut microbiome development) [54, 55]. Based on our results, the highly diverse cervicovaginal microbiome of the Afro-Caribbean women is a reservoir of antibiotic resistance genes. Discovery of AMR genes to antibiotics such as azithromycin (marketed as Zithromax™), Clindamycin (Cleocin™) and nitroimidazoles (metronidazole/Flagyl™) in this resistome is significant as these drugs are the mainstay of treatment of many common STIs and BV. As is documented in literature, black women with predominantly CST IV cervicovaginal bacterial communities have a greater prevalence of BV which is associated with increased risk of STIs, HIV and HPV acquisition [13, 37, 56].

**Immuno-Munibiome Model: a conceptual hypothesis of cervicovaginal microbiome structure and regulation**
The regulatory mechanism involving CST and vaginal eubiosis is not well understood. The absence or dominance of specific planktonic prokaryotes such as *Lactobacilli* or a core microbiome do not completely account for vaginal health. Our findings, along with the analysis of referenced vaginal microbiome studies, suggest that emphasis should perhaps be placed more on microbial functions rather than on the abundance of a specific bacterial genus comprising vaginal microbiomes. Thus a “core microbiome” should fundamentally be thought of as a “munibiome” (termed by authors, Latin; *muni* “functions”, *biome* “bacterial community”) where the function(s) of a group of organisms assumes greater relevance.

Here, based in part on the findings of this study, we introduce a hypothetical concept, the immunomunibiome model, for the regulation of the vaginal ecosystem in the Afro-Caribbean cohort. We posit that vaginal eubiosis is maintained by two opposing microbial states. The first being a facultative anaerobic state (FAS) governed by lactic acid producing microbes (such as *Lactobacillus* and *Bifidobacterium* species) and the second an obligate anaerobic state (OAS) governed by species (such as *Prevotella* and *Gardnerella*) that produce biogenic amines (Figs. 6 and 7). We suggest that one anaerobic state is more dominant than the other. Based on our research findings, the cervicovaginal microbiome CST IV (with *Prevotella* having the highest relative abundance) was the dominant community state type for our Afro-Caribbean cohort. Therefore, for this group we considered the OAS to be the dominant state; one that is likely heritable. It has been suggested that host genetics may in fact be a determining factor in defining the community state types among females of different ethnicities [57]. In addition, heritability estimates determined by variance component methods using Sequential Oligogenic Linkage Analysis Routines (SOLAR) in a recent study involving 542 Korean women showed *Prevotella* to be the most heritable (72.2%) taxa within the vaginal microbiota of women with BV. For the *Lactobacillus* species, *L. crispatus*
and L. iners had heritable estimates of 36.9% and 41.2%, respectively [51]. Interestingly, CST IV has been described in the literature as asymptomatic BV, albeit an ill-defined purported disease [58].

In this study, there were increased biological processes related to bacterial-type flagellum-dependent cell motility, cell adhesion, response to biotic stimulus and quorum sensing. These processes are known to facilitate biofilm formation [11, 59, 60]. Curiously, sporulation was also noted to be upregulated. Sporulation is probably involved in vaginal biofilm development since mutations in spoOA genes that encode major early sporulation transcription factor required for sporulation have been found to inhibit biofilm formation in Bacillus subtilis [61]. Additionally, fragile wall-less planktonic mycoplasmas like M. hominis ATCC 23114 and U. urealyticum identified in this research are not expected to survive the hostile vaginal environment and so it is inferred that these and other pathobionts are protected by their integration into vaginal biofilms: dynamic, highly organized adherent bacterial communities self-encased in an extracellular polysaccharide (EPS) that forms a matrix. Within the matrix, microbial interactions and controlled differential gene expression to extracellular stimuli and other vaginal bacteria are facilitated by quorum sensing and signalling. Switching between the planktonic and complex multicellular modes is also regulated by chemical signalling [11, 12, 62].

It is known that G. vaginalis can strive in the vagina in the presence of high concentrations of lactic acid, hydrogen peroxide, bactericidins and antibiotics due to the formation of protective biofilms [63]. G. vaginalis biofilms, enhanced by Prevotella, harbour a plethora of obligate anaerobes (such as Mycoplasma, Sneathia, Gemella, Mobiluncus, Atopobium and Megasphaera amongst other anaerobes) [64, 65]. Some species or strains of Gardnerella do not cause disease [11, 14, 66]. In this research, G. vaginalis 409-05 was more abundant than G. vaginalis ATCC
14019 among this group of asymptomatic women (Fig. 5 and Supplementary Table S5).

Intriguingly, the genome of *G. vaginalis* 409-05 lacks several mucin degrading and antibiotic resistance genes (important virulence factors) compared to *G. vaginalis* ATCC 14019 [67]. Like *Prevotella* and *Gardnerella* species, *Lactobacillus jensenii* has been shown to form vaginal biofilms in vivo [12, 68, 69]. Therefore, we further propose that the two opposing anaerobic states (FAS and OAS) exist as dynamic vaginal biofilms. We envision that metabolic activities within these anaerobic states induce local vaginal anti-inflammatory (Iₐ) and pro-inflammatory (Iₚ) responses that once in equilibrium maintain vaginal health.

Biogenic amines (e.g. putrescine, cadaverine, phenethylamine, tyramine) synthesized by obligate anaerobes (e.g. *Prevotella* and *Gardnerella*), Gram-negative microbial lipopolysaccharides (LPS) and other bacterial products (lipoproteins/lipopeptides) within OAS induce pro-inflammatory responses via TLR-4 and NF-κB resulting in the release of cytokines and chemokines: an increase in IL-1α, IL-1β, IL-6, IL-8, TNF-α, IFN-γ, RANTES and a decrease in Lipocalin-2 (Neutrophil gelatinase-associated lipocalin, NGAL) secreted by neutrophils [13, 70, 71, 72, 73, 74]. In this hypothesis, the pro-inflammatory response is mitigated by an anti-inflammatory immune response triggered within the FAS by lactic acid produced by *Lactobacillus* and other LABs (e.g. *Bifidobacterium*, *Corynebacterium*, *Aerococcus*, *Enterococcus* and *Atopobium vaginae*), in which IL-6, IL-8, TNF-α, CCL5 (a chemotactic cytokine) and macrophage inflammatory protein, MIP-3α are inhibited [75, 76, 77]. Vaginal epithelial cells secrete less than 15% of L-lactic acid via anaerobic metabolism of glycogen. *Lactobacillus* species produce and regulate the majority of D and L enantiomers of lactic acid within the vaginal mucus [8]. This lactic acid is generated from the hydrolysis of maltose, maltotriose, maltotetraose and α-dextrins which are products of cell-free vaginal glycogen catabolized by α-amylase. The latter is secreted by endocervical and fallopian
tube epithelial cells [78]. The quantity of vaginal cell-free glycogen is not significantly correlated with oestrogen levels [79]. However, high levels of cell-free glycogen positively correlate with a high abundance of *L. crispatus* and *L. jensenii*, but not *L. iners* [80, 81]. Free glycogen is released from vaginal epithelial cells via cell lysis and epithelial cell membrane perforation by the action of matrix metalloproteinase-8 (MMP-8) hyaluronidase-1 and cytolysin generated by *Lactobacillus* [9]. At pH 3.86, 50% of the lactic acid is in equilibrium with lactate and hydrogen ions. Uptake of hydrogen ions by obligate anaerobes are utilized in decarboxylation of amino acids within the bacterial cytoplasm to produce biogenic amines. The release of the biogenic amines elevates the pH (>4.5) within the vaginal ecosystem inhibiting *Lactobacillus* growth [71].

A pro-inflammatory response is associated with a high pH which is unfavourable for facultative anaerobes. Under these circumstances, perhaps through quorum sensing, lactate dehydrogenase is upregulated and *L. iners* (presumably along with other non-lactobacillus LABs) which is better adapted (than *L. crispatus*, *L. gasseri* and *L. jensenii*) to this environment proliferates to become dominant (CST III) [82]. The resulting increase in L-lactic acid production effectively reduces the pH, increases release of intracellular glycogen via pore-forming cholesterol-dependent cytolysin, inerolysin [83] shifting the IA-IP equilibrium leftwards. We postulate that with this equilibrium shift (lower pH and higher cell-free glycogen concentrations) that either *L. crispatus*, *L. gasseri* or *L. jensenii* which are more efficient at producing the more protective D-lactic acid [84] out compete other species, with the most adaptable species (and hence CSTs I, II or V) becoming dominant. Similar to the OAS biofilm, various *Lactobacilli* species are expected to co-exist in LAB biofilm of the FAS. In essence, we believe that the cervicovaginal microbiome cycles through various *Lactobacillus*-dominant CSTs depending on the local vaginal environment and external influences (e.g. contraceptive use, antibiotic therapy, coitus, douching, menstruation). We,
therefore, speculate that CSTs I, II, III and V are snapshots of an active facultative anaerobic state.

In fact, the temporal dynamics of these transient CSTs at the individual level have been well documented [20, 45, 85].

In this model, only the FAS (dominant in Caucasians) is influenced by oestrogen levels. Therefore, any factors (such as age, obesity/ body mass index (BMI) and ethnicity) that influence oestrogen levels should not have any significant impact on the OAS. Hence, as predicted, in contrast to white women no correlations have been demonstrated to date between oestrogen levels, body mass index (BMI) and the cervicovaginal microbiome in black women [86, 87] even though they possess higher levels of oestrogen compared to Caucasians [88].

Conclusions

This is the first study to characterize the cervicovaginal microbiome in a cohort of Afro-Caribbean women. Functional profiling of the cervicovaginal microbial communities suggested the occurrence of biofilm activity in this cohort. High abundances of Prevotella species and the less virulent G. vaginalis 409-05 strain in the cervicovaginal microbiomes of these women (that were clinically asymptomatic of vaginal infection), along with published discoveries of vaginal Lactobacillus biofilms point towards a possible protective role of (facultative and obligate anaerobe) biofilms in vaginal health. The conceptualized immuno-munibiome model (based on inflammatory responses controlled by metabolic activity within biofilms of opposing obligate and facultative anaerobic states) accounts for the CSTs and regulation of this vaginal ecosystem. In addition, expression of macrolide-lincosamide streptogramin B and nitroimidazole AMR genes within the cervicovaginal microbiome in this ethnic group may reduce the effectiveness of azithromycin, clindamycin and metronidazole currently used to treat STIs and BV in Afro-
Caribbean women. Future studies focusing on microbial ecophysiology, biofilm activity and bacterial-induced inflammatory responses at the species-strain level could further contribute to our understanding of the structure and regulation of the cervicovaginal microbiome.

**Declarations**

**Ethics approval and consent to participate**

The research protocol of this study was approved (IRB# 170710-B) by the Research Ethics Committee (REC) and Institutional Review Board (IRB) of The University of the West Indies, Cave Hill Campus, Barbados. Written informed consent was obtained from all participating volunteers.

**Consent for publication**

Not applicable

**Availability of data and materials**

The nucleotide sequence data for 16S rRNA V4 and WGS metagenomics for this study were deposited in the European Nucleotide Archive (ENA) and MGnify at EMBL-EBI under accession numbers PRJEB34744 (MGYS00005125) and PRJEB34967 (MGYS00005124), respectively. Data are also available on MG-RAST server: mgp89180 and mgp88579.

**Competing Interests**

The authors declare no competing interests.

**Funding**
This research was financially supported by the UWI Graduate Student Research Award Committee.

Author Contributions

In this study OR conceptualized, designed, performed laboratory work, bioinformatic analyses and wrote the manuscript. CK prepared the libraries and MT performed Illumina next-generation sequencing on samples. AA, KN and MT critically reviewed the manuscript.

Acknowledgments

The authors are grateful for necessary facilities provided by the Head, Department of Biological and Chemical Sciences, University of the West Indies, Cave Hill Campus. We thank Dr. Bidyut Mohapatra for his expert advice and all volunteers who participated in this research.

Supporting Information

Table S1. Sequence statistics of 15 cervicovaginal samples from Afro-Caribbean women (Whole genome shotgun metagenomics). (MS)

Table S2. OTU statistics of 18 cervicovaginal samples from Afro-Caribbean women (16S rRNA V4 sequencing). (MS)

Table S3. Relative abundance values and operational taxonomic units (OTUs) derived from partial 16S rRNA V4 sequences for cervicovaginal microbiota in Afro-Caribbean women (sheet 1A). The identified taxa based on OTUs sequences and raw abundances are listed in sheet 1B. Color codes are used to represent the species niche or conflicts: Red= conflicts, non-bacteria, non-human; Blue= gut and oral microbiota; Green= Soil or plant microbiota. (XLSX)
Table S4. Results of taxa obtained from blasting OTU sequences (from 16S rRNA V4 sequencing) using SeqMatch version 3 in the online Ribosomal Database Project (RDP: release 11_5) (Cole et al., 2014). (PDF). Available at https://data.mendeley.com/datasets/gn5jgvgd4s/draft?a=0e91656e-f307-4574-94de-b448741b14b2

Table S5. Relative abundance values for all taxa in cervicovaginal samples from Afro-Caribbean women based on WGS sequencing (Sheet 1A). Relative abundance values for top 61 taxa in cervicovaginal samples from Afro-Caribbean women (Sheet 1B). The mean combined abundances for top 30 genera and species/strain are shown for all Afro-Caribbean women (Sheet 1C). (XLSX)

Table S6. Most abundant taxa for cervicovaginal microbiota in Afro-Caribbean women based on 16S rRNA V4 sequencing and whole genome shotgun metagenomics. Results for the most abundant taxa based on WGS sequencing are summarized in Table 1. (XLSX)

Table S7. Metadata for 18 Afro-Caribbean women

Table S8. Statistical analysis of differential abundance of three most abundant taxa and Mycoplasma. (PDF)

Table S9. Antibiotic resistance genes in the cervicovaginal microbiomes of Afro-Caribbean women. (XLSX)

Figure S1. Alpha and Beta diversity metrics used for 16S rRNA V4 sequencing and whole genome shotgun metagenomics. Metrics for alpha diversity: Chao-1 bias-corrected, Shannon entropy and Simpson’s index. Metrics for beta diversity: Bray-Curtis, Jaccard, Euclidean, weighted UniFrac and unweighted UniFrac. Beta diversity was visualized using 3D principal coordinates analysis (PCoA). (MS)
Figure S2. Heatmap showing the relative abundance of a set of 98 functional profiles for the cervicovaginal microbiome of Afro-Caribbean women based on whole genome shotgun metagenomic sequence analysis.

References


7. Miller EA, Beasley DAE, Dunn RR, Archie EA. Lactobacilli dominance and vaginal pH:


16. Okoli AC, Agbakoba NR, Ezeanya CC, Oguejiofor CB, Anukam KC. Comparative abundance and functional biomarkers of the vaginal and gut microbiome of Nigerian


Herbst FA, Dueholm MS, Wimmer R, Nielsen PH. The proteome of Tetrasphaera elongata is adapted to changing conditions in wastewater treatment plants. Proteomes. 2019;7(2).


48. Vilo C, Dong Q. Evaluation of the RDP Classifier Accuracy Using 16S rRNA Gene


Marconi C, Donders GG, Parada CMGL, Giraldo PC, Da Silva MG. Do Atopobium
vaginae, Megasphaera sp. and Leptotrichia sp. change the local innate immune response

73. Roachford OSE, Nelson KE, Mohapatra BR. Comparative genomics of four *Mycoplasma*
species of the human urogenital tract: Analysis of their core genomes and virulence genes.

74. Roachford O, Nelson KE, Mohapatra BR. Virulence and molecular adaptation of human

75. Hearps AC, Tyssen D, Sribnovski D, Bayigga L, Diaz DJD, Aldunate M, et al. Vaginal
lactic acid elicits an anti-inflammatory response from human cervicovaginal epithelial cells
and inhibits production of pro-inflammatory mediators associated with HIV acquisition.

HY7042 ameliorates *Gardnerella vaginalis*-induced vaginosis by killing *Gardnerella*

77. Rizzo A, Fiorentino M, Buommino E, Donnarumma G, Losacco A, Bevilacqua N.
*Lactobacillus crispatus* mediates anti-inflammatory cytokine interleukin-10 induction in

α-amylase present in lower-genital-tract mucosal fluid processes glycogen to support


86. Brookheart RT, Lewis WG, Peipert JF, Lewis AL, Allsworth JE. Association between


FIGURE LEGENDS

**Fig. 1.** Relative abundance of selected taxa at the species level. Microbial community structures determined by 16S rRNA V4 and WGS metagenomics are similar. (A) Relative abundance of taxa as determined by targeted 16S rRNA V4 sequencing (B) Relative abundance of taxa as determined by whole genome shotgun (relative abundances for *T. japonica* not included) (C) Relative abundance of taxa as determined by whole genome shotgun (relative abundances for *T. japonica* included). Cervicovaginal samples are labeled L100 to H210.

**Fig. 2.** (A) 16S rRNA V4 sequencing data analyzed using Chao-1 bias-corrected $\alpha$-diversity metric. (B) 16S rRNA V4 sequencing data visualized with 3D principle coordinates analysis (PCoA) for using Bray-Curtis $\beta$-diversity metrics. (C) Whole genome shotgun metagenomics data visualized with 3D principal coordinates analysis (PCoA) Bray-Curtis using $\beta$-diversity metrics. Refer to supplementary Figure S1 for additional analyses.

**Fig. 3.** Heatmap indicating relative abundance of bacteria in the cervicovaginal microbiome of Afro-Caribbean women based on 16S rRNA V4 sequence analysis.

**Fig. 4.** Area chart showing relative abundance of antimicrobial (AMR) resistance genes in the cervicovaginal microbiome of Afro-Caribbean women.

**Fig. 5.** Heatmap indicating relative abundance of bacteria in the cervicovaginal microbiome of Afro-Caribbean women based on whole genome shotgun metagenomic sequence analysis.

**Fig. 6.** Rooted circular cladogram of taxa of cervicovaginal microbiome of Afro-Caribbean women based on 16S rRNA V4 sequence analysis. Dominant facultative anaerobes (*Lactobacillus* species) and dominant obligate anaerobes (*Prevotella* species) are indicated by red and green boundaries, respectively.
Fig. 7. Schematic of the immune-munibiome model depicting the dynamic regulation of the cervicovaginal ecosystem (pink area indicates lower vaginal tract). The left border of the diagram represents the obligate anaerobe state (OAS) biofilm. On the right border is the biofilm of the facultative anaerobic state. A low pH (high lactic acid concentration) generated by lactic acid bacilli (LAB) bacteria is mitigated by the consumption of protons and production of biogenic amines by obligate anaerobes. Lactic acid induces an anti-inflammatory immune response, while lipopolysaccharides (LPS) and lipoproteins trigger a proinflammatory immune response. The net result is equilibrium of inflammatory responses and pH changes that maintains vaginal eubiosis.
Fig. 1.

(A)

(B)

(C)
Fig. 2.

(A) Chao 1 bias-corrected

(B) PCo 2 (15%)

(C) PCo 1 (25%)
Fig. 3.

Clustering: Sample and feature clustering
Data: Normalized abundances
Dissimilarity: Euclidean distance
Linkage: Complete linkage

Taxon Abundance

- Prevotella amnii
- Prevotella sp.
- Prevotella sp.
- Sneathia amnii
- Megasphaera sp.
- Sneathia sanguinegens
- Prevotella sp. S7-1-8
- Dialister sp.
- Prevotella amnii
- Prevotella timonensis
- Prevotella histicola
- Anaerococcus jeddahensis
- Veillonella montpellieriensis
- Howardella sp.
- Shuttleworthia satelles
- Porphryromonas asaccharolytica
- Prevotella timonensis
- Gardnerella vaginalis
- Lactobacillus fomicalis
- Lactobacillus jensenii
- Lactobacillus iners
- Lactobacillus iners
- Lactobacillus iners
- Lactobacillus crispatus
- Lactobacillus crispatus
AMR = Antimicrobial resistance

**Fig. 4.**

Relative abundance of AMR genes (%)
Fig. 5.

Clustering: Sample and feature clustering
Data: Normalized abundances
Dissimilarity: Euclidean distance
Linkage: Complete linkage
Vaginal epithelial cells can be lysed by inerolysin and mucin degraded by sialidase, leading to extragenital disease.

Cell adhesion
Co-aggregation
Matrix: extracellular polymeric substance (EPS)
Quorum sensing
Antibiotic resistance

Obligate anaerobic state (OAS)
- Obligate anaerobes
  - α-amylase
  - Glycogen
  - LAB e.g. Lactobacilli, Bifidobacterium
  - Biogenic Amines → Amino acids + H^+

Facultative anaerobic state (FAS)
- Facultative anaerobes
  - LPS/lipoproteins
  - TLR / NF-κB
  - Pro-inflammatory
    - ↑IL-1α, IL-1β, IL-6, IL-8, TNF-α
    - ↑IFN-γ, RANTES, ↓Lipocalin-2
  - Anti-inflammatory
    - ↑IL-10
    - ↓IL-6, IL-8, TNF-α

LAB biofilm
- More dominant than Lactobacilli biofilm

Biofilm alternates between planktonic and multicellular modes.

Fig. 797

Affected by:
1. BMI
2. Ethnicity
3. Age
4. Drugs
Table 1. Relative abundance for taxa in the cervicovaginal microbiome of Afro-Caribbean women based on 1012 WGS and 16S rRNA V4 sequencing.

<table>
<thead>
<tr>
<th>WGS sequencing (Overall n=15)</th>
<th>16S rRNA V4 sequencing (Overall n=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Taxa</strong></td>
<td><strong>Abundance (%)</strong></td>
</tr>
<tr>
<td><strong>Prevotella spp.</strong></td>
<td>13.91</td>
</tr>
<tr>
<td><strong>Gardnerella spp.</strong></td>
<td>12.14</td>
</tr>
<tr>
<td><strong>Lactobacillus spp.</strong></td>
<td>9.37</td>
</tr>
<tr>
<td><strong>Mobiluncus spp.</strong></td>
<td>7.15</td>
</tr>
<tr>
<td><strong>Sneathia spp.</strong></td>
<td>5.59</td>
</tr>
<tr>
<td><strong>Megasphaera genomosp. type_1 str. 28L</strong></td>
<td>3.00</td>
</tr>
<tr>
<td><strong>Atopobium spp.</strong></td>
<td>2.31</td>
</tr>
<tr>
<td><strong>Bifidobacterium breve DSM 20213</strong></td>
<td>2.00</td>
</tr>
<tr>
<td><strong>Mageeibacillus indolicus UPII9-5</strong></td>
<td>2.00</td>
</tr>
<tr>
<td><strong>Porphyromonas asaccharolytica DSM 20707</strong></td>
<td>2.00</td>
</tr>
<tr>
<td><strong>Dialister micraerophilus DSM 19965</strong></td>
<td>0.66</td>
</tr>
<tr>
<td><strong>Aerococcus christensenii</strong></td>
<td>0.31</td>
</tr>
<tr>
<td><strong>Peptoniphilus lacrimalis DSM 7455</strong></td>
<td>0.30</td>
</tr>
<tr>
<td><strong>Eubacterium sulci ATCC 35585</strong></td>
<td>0.23</td>
</tr>
<tr>
<td><strong>Propionibacterium acnes KPA171202</strong></td>
<td>0.21</td>
</tr>
<tr>
<td><strong>Staphylococcus aureus subsp. aureus NCTC 8325</strong></td>
<td>0.19</td>
</tr>
<tr>
<td><strong>Alloprevotella tannerae ATCC 51259</strong></td>
<td>0.19</td>
</tr>
<tr>
<td><strong>Anaerococcus</strong></td>
<td>0.17</td>
</tr>
<tr>
<td><strong>Streptococcus pseudoporcinos LQ 940-04</strong></td>
<td>0.17</td>
</tr>
<tr>
<td><strong>Thermogemmatispora carboxidivorans</strong></td>
<td>0.14</td>
</tr>
<tr>
<td><strong>Lachnospiraceae</strong></td>
<td>0.13</td>
</tr>
<tr>
<td><strong>Mycoplasma hominis ATCC 23114</strong></td>
<td>0.12</td>
</tr>
<tr>
<td><strong>Gemella asaccharolytica</strong></td>
<td>0.11</td>
</tr>
<tr>
<td><strong>Prevotellamassilia timonensis</strong></td>
<td>0.10</td>
</tr>
</tbody>
</table>