RESEARCH ARTICLE

Comparative analysis of vaginal microbiota sampling using 16S rRNA gene analysis

Seppo Virtanen[1], Ilkka Kalliala[1,2], Pekka Nieminen[1], Anne Salonen[3]*

1 Obstetrics and Gynecology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland, **2** Institute of Reproductive and Developmental Biology, Department of Surgery and Cancer, Imperial College London, London, United Kingdom, **3** Immunobiology Research Programme, Department of Bacteriology and Immunology, University of Helsinki, Helsinki, Finland

* anne.salonen@helsinki.fi

Abstract

Background

Molecular methods such as next-generation sequencing are actively being employed to characterize the vaginal microbiota in health and disease. Previous studies have focused on characterizing the biological variation in the microbiota, and less is known about how factors related to sampling contribute to the results. Our aim was to investigate the impact of a sampling device and anatomical sampling site on the quantitative and qualitative outcomes relevant for vaginal microbiota research. We sampled 10 Finnish women representing diverse clinical characteristics with flocked swabs, the Evalyn® self-sampling device, sterile plastic spatulas and a cervical brush that were used to collect samples from fornix, vaginal wall and cervix. Samples were compared on DNA and protein yield, bacterial load, and microbiota diversity and species composition based on Illumina MiSeq sequencing of the 16S rRNA gene. We quantified the relative contributions of sampling variables versus intrinsic variables in the overall microbiota variation, and evaluated the microbiota profiles using several commonly employed metrics such as alpha and beta diversity as well as abundance of major bacterial genera and species.

Results

The total DNA yield was strongly dependent on the sampling device and to a lesser extent on the anatomical site of sampling. The sampling strategy did not affect the protein yield or the bacterial load. All tested sampling methods produced highly comparable microbiota profiles based on MiSeq sequencing. The sampling method explained only 2% (p-value = 0.89) of the overall microbiota variation, markedly surpassed by intrinsic factors such as clinical status (microscopy for bacterial vaginosis 53%, p = 0.0001), bleeding (19%, p = 0.0001), and the variation between subjects (11%, p-value 0.0001).

Conclusions

The results indicate that different sampling strategies yield comparable vaginal microbiota composition and diversity. Hence, past and future vaginal microbiota studies employing

different sampling strategies should be comparable in the absence of other technical confounders. The Evalyn® self-sampling device performed equally well compared to samples taken by a clinician, and hence offers a good-quality microbiota sample without the need for a gynecological examination. The amount of collected sample as well as the DNA and protein yield varied across the sampling techniques, which may have practical implications for study design.

Introduction

Along with the general trend in the human microbiome field, high-throughput molecular methods, predominantly next-generation sequencing (NGS), are widely used to provide untargeted identification and relative quantification of vaginal bacteria, including species that cannot be cultured. Alterations in the vaginal microbiota and host-microbe interactions have been implicated in a number of gynecological and obstetric conditions, such as bacterial vaginosis [1], cervical cancer [2,3], infertility [4], preterm birth [5] and sexually transmitted infections (STI) [6]. Apart from disease, the vaginal microbiota is affected by various intrinsic factors such as menstrual cycle [7], pregnancy [8], ethnicity [9] and menopause [10] and by external factors such as smoking [11]. Evidently, understanding the normal variation of the vaginal microbiota and the role of different host-associated factors is imperative for interpretation of the data and identifying diagnostic microbiota signatures. The contribution of technical variation introduced by specimen collection and storage, DNA extraction, and sequencing should also be known and optimally minimized to enable cross-comparison of the different studies, and to unmask the underlying biological variation.

Previously, a study comparing the performance of different sample collection methods for vaginal soluble proteins [12] found that endocervical swabs and lavage samples collected higher cytokine and antimicrobial protein concentration compared to vaginal swabs. Also, the impact of the anatomical sampling site on the vaginal microbiota composition has been previously analyzed in two culture-independent studies with contradicting results: A clone-library based shallow sequencing assessment of the microbiota from eight women [13] suggested that both the sampling site and the device had substantial impact on the resultant microbiota, and due to the high variation of the replicate samples within individual, a single swab sample is not representative of the overall vaginal microbiota. On the other hand, the Human Microbiome Project sampled 113 women from vaginal introitus, mid-vagina and posterior fornix, and concluded based on pyrosequencing that the taxonomic profiles between the sites were essentially similar [14,15].

Self-sampling of vaginal content has been validated for the detection of individual pathogens such as high-risk types of human papilloma virus (hrHPV) [16] and *Chlamydia trachomatis* [17], as well as for targeted analysis of certain commensal bacteria [18]. Self-sampling was also shown to be suitable for community-wide microbiota analysis based on comparison of self- versus physician-collected samples from mid-vagina [19]. However, there are no previous data regarding the comparison of self-collected samples to possible spatial or sampling device-specific variation of microbiota in the vagina. Other potential sources of technical variation in vaginal microbiota sampling have been studied before, e.g. sample storage conditions [20] and bacterial DNA extraction [21]. However, this is the first study to compare different vaginal sites and sampling devices with NGS.

Non-biased, reproducible and validated strategies are needed to collect vaginal samples for microbiota analysis. Therefore, we have here employed partial 16S rRNA gene deep sequencing (NGS) to test whether the bacterial diversity and composition differs between commonly used vaginal sampling devices and different vaginal sites. We also wanted to test how representative the microbiota derived from a self-sampling device is compared to samples taken by a clinician from spatially specified sites, in order to estimate the usefulness of the self-sampling device for large-scale microbiota studies.

Materials and methods

Subjects and sampling

We sampled the vaginal microbiota of 10 premenopausal women at the Department of Obstetrics and Gynecology, Helsinki University Hospital Colposcopy Clinic in Southern Finland with seven different sampling methods, summing up to total of 70 samples. All subjects were recruited and sampled in April 2014. As each woman served as their own control regarding the different sampling strategies, we did not aim for homogenous population but instead included women of varying age, colposcopy result, use of contraceptives and presence of blood in the samples (Table A in S1 Appendix). Patients were not pregnant (self-reported) and they had more than 48 hours since last intercourse, except patient number 8 who had 36 hours. A Pap smear was taken from all patients and the presence of bacterial vaginosis (BV) was evaluated from smears with microscopy, using clue-cells and Nugent score features as diagnostic criteria [22]. We also recorded smoking status, contraceptive use, last menstrual period and recent antibiotic use.

We took samples from six vaginal locations with four different sampling devices as specified in Fig 1. A Flocked swab (Floqswab regular, Copan, USA) was chosen for ease of use and offers high sample elution. A cervical brush (Cytobrush Plus, Medscand, USA) is already routinely used in colposcopy and offers accurate sampling from a specific site. The Evalyn® self-sampling device (Rovers, Netherlands) developed for hrHPV-testing was included to address the quality of self-collected microbiota samples. The Evalyn® device has a brush that is inside a protective plastic tube during insertion, and hence does not collect material from vaginal introitus or the lowest parts of vagina. After insertion, the brush comes out of the tube and rotates when a plunger at the end of the device is pushed and rotated. The tube has small wings as stoppers so that the depth of the insertion is about 2 cm (see Fig 1). We chose this device for self-sampling to minimize skin contamination and to collect microbiota from specified and reproducible depth in vagina. Scrape samples taken from the lower and upper third as well as from the entire vaginal wall with plastic spatulas (sterile PS-plastic, Aylesbury-shape) were used as a reference to estimate the sampling depth, possible skin or vulvar contamination of self-sampling, and to gather possible adherent bacteria from the vaginal wall. All samples were fresh frozen in dry 2 ml Eppendorf tubes except the left fornix flocked swabs, which were frozen in 1 ml saline to test the potential effect of storage buffer (sample 3. in Fig 1.). The Eppendorf tubes were frozen in a container filled with dry ice in the examination room and transported to -20C freezer by the end of the day. The samples were processed for analysis within 2–10 days from collection. All sampling devices were sterile except the cervical brush. The samples, except for the self-taken sample, were collected by three experienced physicians right after insertion of speculum. The speculum was lubricated with small amount of sterile saline if needed.

Sample processing

DNA extraction. We used a modification of Repeated Bead Beating (RBB) method that efficiently extracts bacterial community DNA based on comparative analysis of human fecal
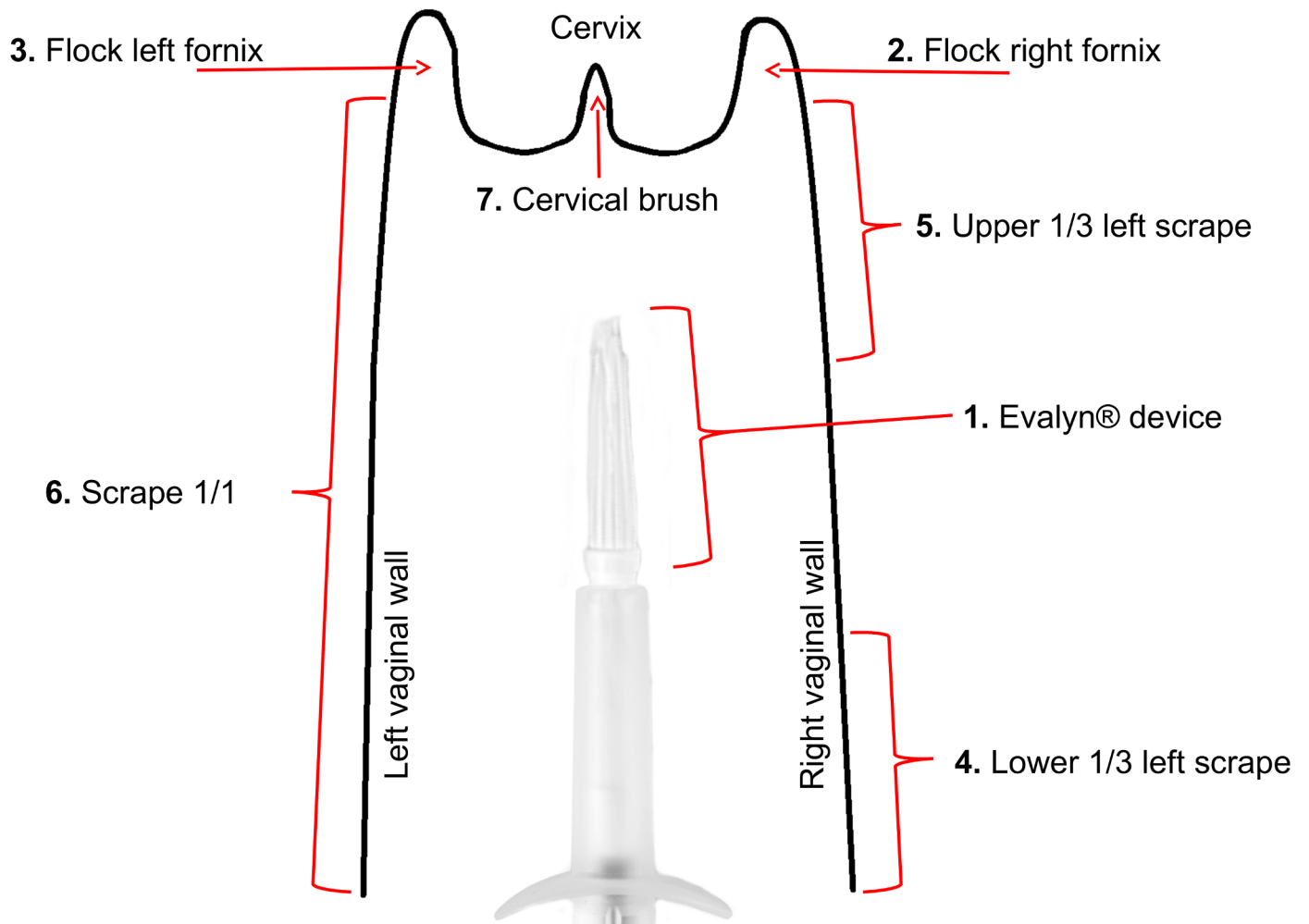
Fig 1. Sampling locations and devices. Schematic representation of vagina with arrows labeled with sampling devices and pointing to sampling locations. The locations are numbered according to the sampling order.

samples [23]. All frozen samples were thawed on ice and 1 ml of sterile ice-cold PBS was added, except for the left fornix flocked swabs that were stored in saline. Before the removal of the sampling device with sterile forceps the samples were vortexed for 30 secs to dislodge the bacteria and briefly spun down to avoid spills when opening the tube. An aliquot of the clarified supernatant was taken for the measurement of soluble protein content. The rest of the sample, including both the mucous pellet and the supernatant, was moved to 2 ml screw-cap tube and an equal volume of lysis buffer (500 mM NaCl, 50 mM Tris-HCl (pH 8.0), 50 mM EDTA, 4% SDS) was added. To break to cells, samples were bead beaten using a Precellys 24 high-throughput tissue homogenizer (Bertin Technologies, France) with 0.1 mm zirconium-silica beads (Biospec Products, Bartlesville, OK, USA) for 2 x 30 sec at 5500 rpm. DNA from the lysates was purified using QIAamp DNA Minikit (Qiagen, Hilden, Germany) without the lysis step according to the manufacturer's instructions. The samples were eluted in 200 μl of kit elution buffer.

DNA and protein quantification.   Nanodrop spectrophotometer (NanoDrop® Technologies, Wilmington, DE, USA) was used to measure DNA and protein content at 260 nm and 280 nm, respectively. The DNA and protein yields were calculated per sample as if the whole

sample was used for these analyses. 0.8% agarose gels electrophoresis was used for the quantification and qualification of the DNA extracts.

qPCR. To quantify the amount of bacterial DNA in the samples, triplicate qPCR amplifications were performed using iCycler—IQ instrument (BioRad Laboratories, USA) and 1x HOT FIREPol® EvaGreen® qPCR Mix Plus (Solis Biodyne, Estonia) in a volume of 25 μl. The thermal cycling conditions included an initial DNA denaturation step at 95˚C for 10 min followed by 40 cycles of denaturation at 95˚C for 20 s, annealing at 56˚C for 20 s, extension at 72˚C for 30 s, and an additional incubation step at 88˚C for 30 s to collect the fluorescent data. Bacterial DNA was quantified by amplifying 0.5 ng aliquots of each DNA extract with universal bacterial primers targeting the 16S rRNA gene [24]. The standard curves ranging from 10^2 to 10^7 copies were constructed using the full-length amplicons of 16S rRNA gene of *Bifidobacterium longum* to convert the threshold cycle (Ct) values into 16S rRNA gene copy numbers per ng of template DNA.

16S rRNA gene amplicon sequencing. Illumina MiSeq paired-end sequencing of the hypervariable V1-V3 regions of the 16S rRNA gene with primers pA-forward `AGAGTTTGA TCMTGGCTCAG` and pD'-reverse `GTATTACCGCGGCTGCTG` was performed at the Institute of Biotechnology, DNA Sequencing and Genomics Laboratory, University of Helsinki, Helsinki, Finland. The PCR primers contained 18-mer overhangs added to the 5' ends resulting to amplicon size of 556 nt [25]. Replicate PCR products were pooled and purified with Agencourt AMPure XP magnetic beads (Agencourt Bioscience, USA) and subjected to a second PCR round with barcoded forward primers and a reverse primer, both of which attached to the respective 18-mer overhang sequences from the primers of the first PCR amplification. Phusion polymerase (Thermo Fisher Scientific, USA) with HF buffer and 2.5% DMSO were used. Cycling conditions for both PCR reactions consisted of an initial denaturation at 98˚C for 30 s, followed by 15 cycles at 98˚C for 10 s, 65˚C for 30 s, and 72˚C for 10 s, and then a final extension for 5 min. The quality control of libraries was performed with Bioanalyzer 2100 (Agilent, USA). We excluded 10 samples that had very low DNA yields (<2.5 ng/ul) or were otherwise considered to be of insufficient quality and quantity for sequencing. Summary of the sequenced 60 samples is in Table B in S1 Appendix. The sequences have been deposited at European Nucleotide Archive, ENA accession code: PRJEB21386.

Preprocessing of reads. MiSeq reads were processed in R [26] with package mare (Microbiota Analysis in R Easily) [27] as shown in Fig 2. The default parameters of mare, such as quality filtering based on read abundance rather than quality scores, use of only forward reads instead of the merged reads, and the truncation of sequences to 150 nt to include only the region of highest quality, were applied as they maximized the similarity between the expected
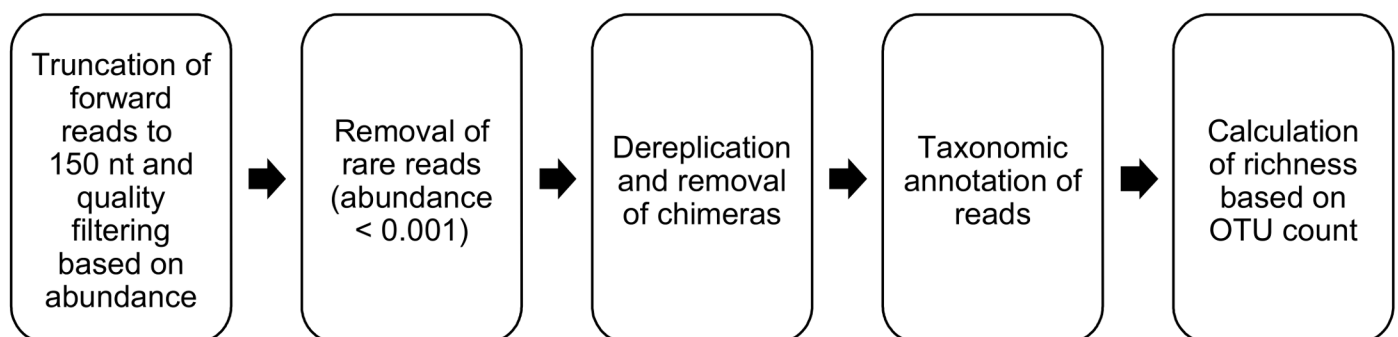

Fig 2. Pre-processing pipeline. MiSeq sequence pre-processing workflow with R-package mare [27].

and observed microbial richness and composition in mock communities [28]. The composition of mock communities used in parameter optimization has been described elsewhere [29]. The forward reads were truncated to length of 150 bases with mare's "ProcessReads" command. We used default settings for minimum quality score (2) and maximum expected errors (1). Reads with abundance below 0.0001 were removed. Chimera removal and dereplication of the reads was done using USEARCH8 [30]. Truncated, filtered and dereplicated reads were then annotated using curated vaginal 16S V1-3 reference database from Fettweis et al. [31] to get species level classification with "TaxonomicTable" function of the mare package. Richness and diversity (inverse Simpson [32]) were estimated after clustering the reads to operational taxonomic units (OTUs) using USEARCH8.

Statistical data analysis

The MiSeq data analysis was done without rarefaction or transformations utilizing statistical and visualization tools included in the mare-package [27]. To control for the effect of the variable read count between the samples, the number of reads was used as an offset in all statistical models. Basic statistical inference in mare is based on generalized linear models with negative binomial distribution utilizing glm.nb function of the R package MASS [33]. Community dissimilarity was estimated with principal coordinates analysis (PCoA) using Bray-Curtis dissimilarity as the distance measure. Bray-Curtis dissimilarity is 0 for identical samples and 1 for communities that do not share any species [34]. PCoA was calculated with the capscale function of R package vegan and Bray-Curtis dissimilarities with function vegdist of the same package [35]. Continuous variables in the PCoA-space were interpolated with gstat function of the gstat package in R [36]. To calculate the relative contribution of different variables in the microbiota variation, permutational multivariate analysis of variance using distance matrices was performed with adonis function in the vegan package. Taxonomic alpha-diversity was estimated as the number of observed OTUs and by the inverse Simpson's diversity index. For univariate data (DNA yield, protein content, qPCR results, Bray-Curtis dissimilarity, richness, diversity) the differences between the samples was determined by repeated measures ANOVA in combination with Tukey's post-hoc tests. We compared the relative abundances of the bacterial species collected with different methods with generalized linear mixed models within "GroupTest" function of the mare package using subject as the random factor. This function uses the glmmADMB package (generalized linear mixed models built on AD Model Builder) of R software on background and assumes negative binomial distribution of abundance (for the command and source code see S2 Appendix) [37,38]. P-values for taxon-specific differences were collected for false discovery rate (FDR; Benjamini–Hochberg [39]) with function p. adjust of the stats package in R [26].

Results

We sampled 10 women at six different vaginal sites with four different devices to optimize and validate vaginal microbiota sampling for molecular studies. The comparison was based on quantitative and qualitative assessment, focusing on the microbiota profiles generated by sequencing of the hypervariable V1-V3 region of the 16S rRNA gene.

Protein content

Protein content was measured for the flocked swabs, the self-taken sample and the cervical brush, because in contrast to the scrape-samples, those were thought to collect sufficiently material for the analysis of soluble immunomarkers and other proteins. All the tested devices collected roughly 2 mg of protein based on the Nanodrop measurement. There were no
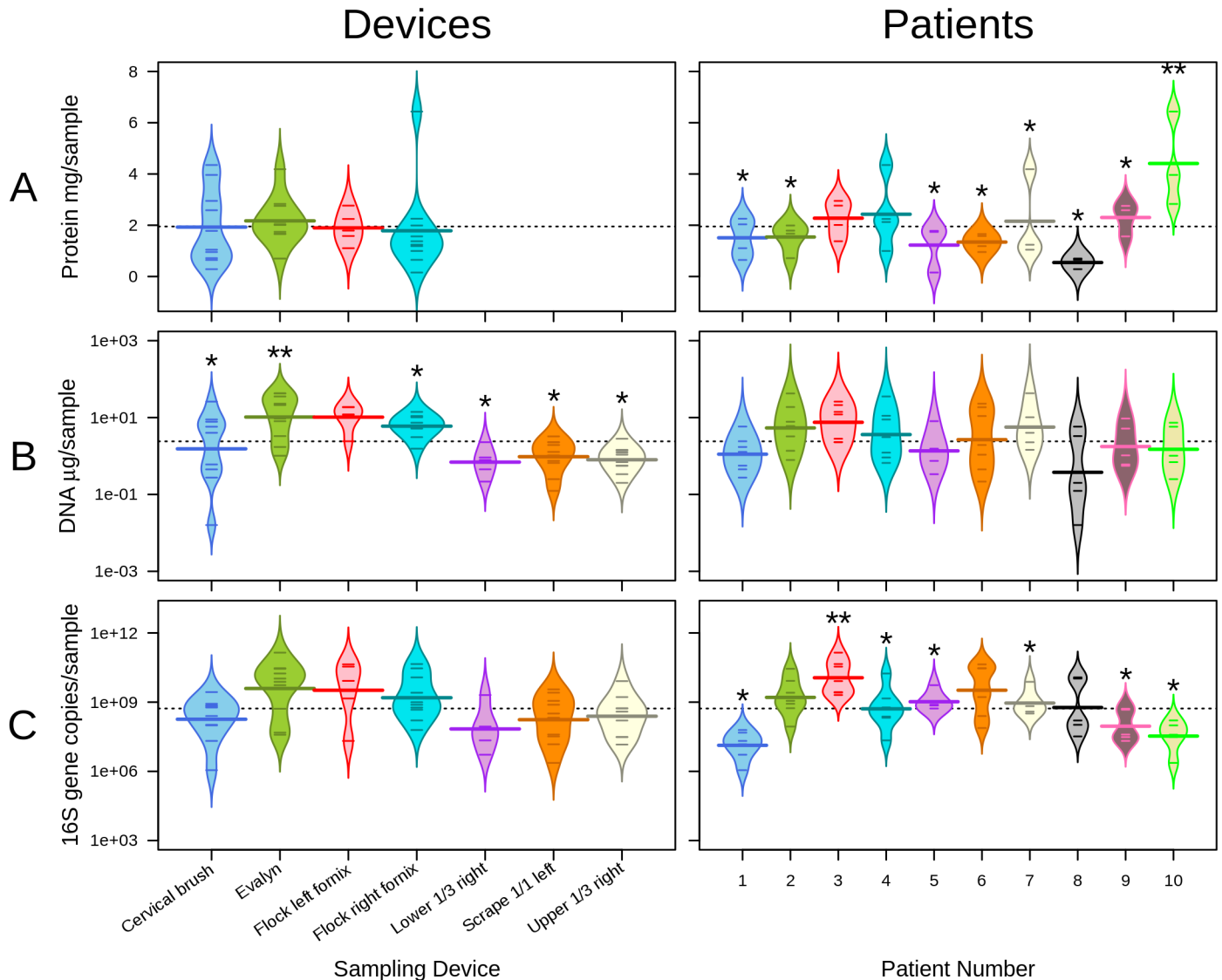
Fig 3. Protein and DNA results for different devices and patients. Protein yield (row A), DNA yield (row B) and total bacteria measured by qPCR (row C). The individual lines depict mean and the overall dotted line the group mean. The left columns show values according to the sampling methods (technical variation) and the and right according to the patients (biological variation). Protein yield was not measured from scrape samples due to small sample volume. Mean values not sharing the same number of asterisks signify samples with statistically significant ($p < 0.05$) pair-wise differences.

https://doi.org/10.1371/journal.pone.0181477.g003

statistically significant differences in protein yield between the devices or vaginal sites, but there were significant, up to 8-fold, differences in protein yield between the patients (Fig 3A). Bleeding or the clinical status including cytological diagnosis and presence or absence of bacterial vaginosis did not have significant effect on the protein yield (Fig A-A in S1 Appendix).

DNA yield and purity

The DNA yield varied with different sampling sites and devices. The mean DNA yields per sampling device varied up to 31-fold, the actual yields ranging from 0.16 ng to 42 μg per sample (Fig 3B). The mean DNA yield per device thus ranged from 596 ng to 19 μg. The highest DNA yield was obtained with self-sampling brush, followed by flock swabs, cervical brush and

scrapes. The DNA yield of the self-taken sample was significantly higher than that of scrapes, cervical brush samples or right fornix flocked swab samples (all p-values < 0.05). Pair-wise comparison between the other devices did not yield significant differences. Bleeding or BV-status did not have effect on DNA yield (Fig A-B in S1 Appendix).

To assess the purity of different DNA preps we compared their A260/A280 and A260/A230 ratios from the Nanodrop measurement. The median of A260/A280 ratio was good (1.8–2.0), except for the scrape samples it varied from 2.4 to 4.8. As all samples were treated with RNAse during the DNA extractions, the higher A260/A280 ratio of scrape samples probably reflects the limitations of the Nanodrop performance with very low concentration samples (median 2.2–5.8 ng/μl DNA) or vulnerability to extraction artifacts/errors for the same reason. The A260/A230 ratio was also good (median 1.3–1.8) in all samples other than the scrapes ($\leq$ 0.7), assumingly for the same reasons as above.

Quantification of bacteria

To specifically quantify bacterial DNA in the different extracts, we used bacteria-specific qPCR assay and calculated the results as the number of 16S rRNA genes per ng of template DNA (ng^{-1} DNA) in order to normalize for the high device-driven variation in the absolute DNA yields. We then multiplied the number of 16S rRNA copies per ng with the DNA yield (corrected for the protein aliquot) to get the number of copies per sample. The mean amount of bacterial DNA varied 74 -fold between the methods and 648-fold between the patients. There were no statistical differences in the amount of bacteria collected across the tested devices and locations (Fig 3C), and no correlation between the DNA yield and bacterial load (Spearman r = 0.24, p = 0.07). Bleeding did not affect the bacterial content, but BV was associated with significantly higher bacterial counts compared to other clinical groups based on microscopy (Fig A-C in S1 Appendix).

Composition and diversity analysis of the vaginal microbiota

Overview of the vaginal microbiota composition. Illumina MiSeq sequencing was used to determine the impact of sampling devices and anatomical sites on the vaginal bacterial community composition. After preprocessing, we had an average of 79 500 high-quality sequences per sample (range 43 000–107 000). Due to the excess number of reads, there was no association between the read count and microbiota richness (Fig C in S1 Appendix). In the studied 10 Finnish asymptomatic women, we detected altogether 45 bacterial species. Most abundant bacteria were *Lactobacillus iners*, *Lactobacillus crispatus*, *Megasphera paucivorans/sueciensis*, *Prevotella bivia* and *Atopobium vaginae*. The variation of microbial composition was high between the individuals, six out of 10 women having *Lactobacillus*-dominated microbiota and the rest had *Enterobacter*, *Lachnospiraceae* BVAB1, *Prevotella* or *Megasphera* dominated community (Fig 4). Four patients had *L. iners* dominated (Community State Type III; CST III) and two *L. crispatus* dominated bacteria (CST I) [40]. We did not detect *Gardnerella vaginalis* in any of the samples, despite the fact that nine patients had abnormal bacterial findings suggestive of bacterial vaginosis in the Pap smear. The lack of *Gardnerella* sequences is likely reflecting PCR bias as the used V1 forward primer does not capture *Bifidobacteriales* including *G. vaginalis* [41].

Comparison of bacterial communities across sampling devices and sites. Our focus was to determine the impact of sampling devices and sites on the overall community composition and diversity of the vaginal microbiota. Principal coordinates analysis (PCoA) based on Bray-Curtis dissimilarity of all the samples revealed that the vaginal microbial communities clustered according to the patient (Fig 5.). Permutational multivariate ANOVA, with Bray-
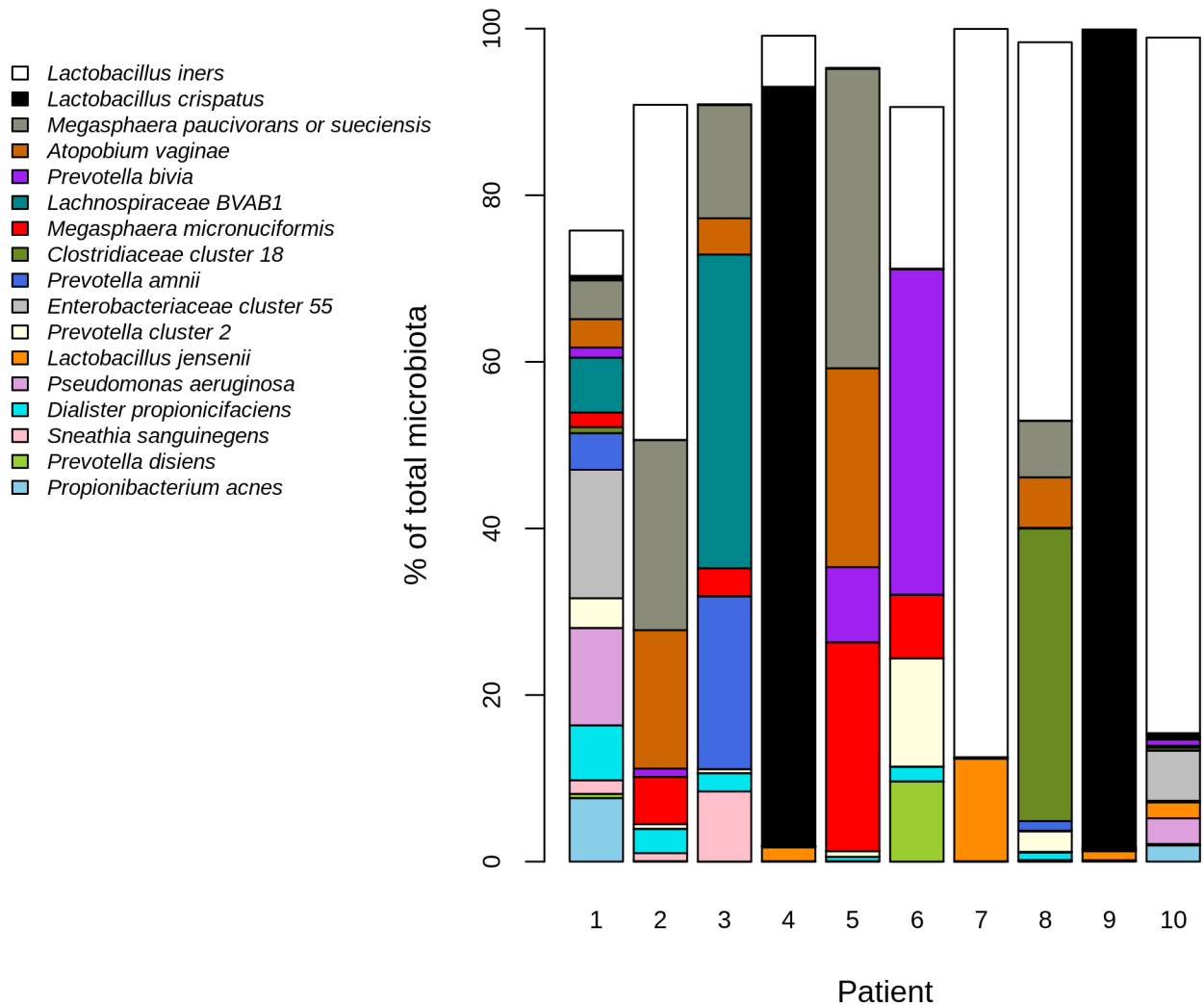
Legend:
- ☐ *Lactobacillus iners*
- ■ *Lactobacillus crispatus*
- ■ *Megasphaera paucivorans or sueciensis*
- ■ *Atopobium vaginae*
- ■ *Prevotella bivia*
- ■ *Lachnospiraceae BVAB1*
- ■ *Megasphaera micronuciformis*
- ■ *Clostridiaceae cluster 18*
- ■ *Prevotella amnii*
- ■ *Enterobacteriaceae cluster 55*
- ☐ *Prevotella cluster 2*
- ■ *Lactobacillus jensenii*
- ■ *Pseudomonas aeruginosa*
- ■ *Dialister propionicifaciens*
- ■ *Sneathia sanguinegens*
- ■ *Prevotella disiens*
- ■ *Propionibacterium acnes*

Fig 4. Abundant microbiota in the study subjects. The plot shows the most abundant bacteria (mean relative abundance > 0.01) per patient, representing mean abundance derived from all the sampling methods. Patients 3,5,6,8 had bacterial vaginosis according to Pap smear at the time of sampling. Patient 1 gave birth 2 months before sampling and post-partum atrophy was seen in Pap smear. Area not covered by the bars represents taxa below the abundance filter.

https://doi.org/10.1371/journal.pone.0181477.g004

Curtis dissimilarities, was used to assess the proportion of variation in the microbiota composition attributable to different biological and technical factors, such as the sample donor, sampling site and device as well as microscopic findings (clinical phenotype). Of the tested variables, microscopic findings (BV-status) explained most of the microbiota variation (53%, p-value 0.0001) followed by bleeding status (19%, p-value 0.0001), the sample donor (11%, p-value 0.0001) and the read count (6%, p-value 0.005). As the read count was significantly associated with the microbiota variation, it was used as offsets in all statistical models. The sampling method explained only 2% (p-value 0.89) of the total variance and hence was practically not associated with the overall microbiota variation. The use of storage buffer (flocked swabs stored dry versus in sterile saline) did not affect the microbiota composition (1%, p-value 0.90).

Despite the high overall similarity of the microbiota profiles between the sampling methods, we quantified their relative distance based on Bray-Curtis dissimilarity using flocked swab, the
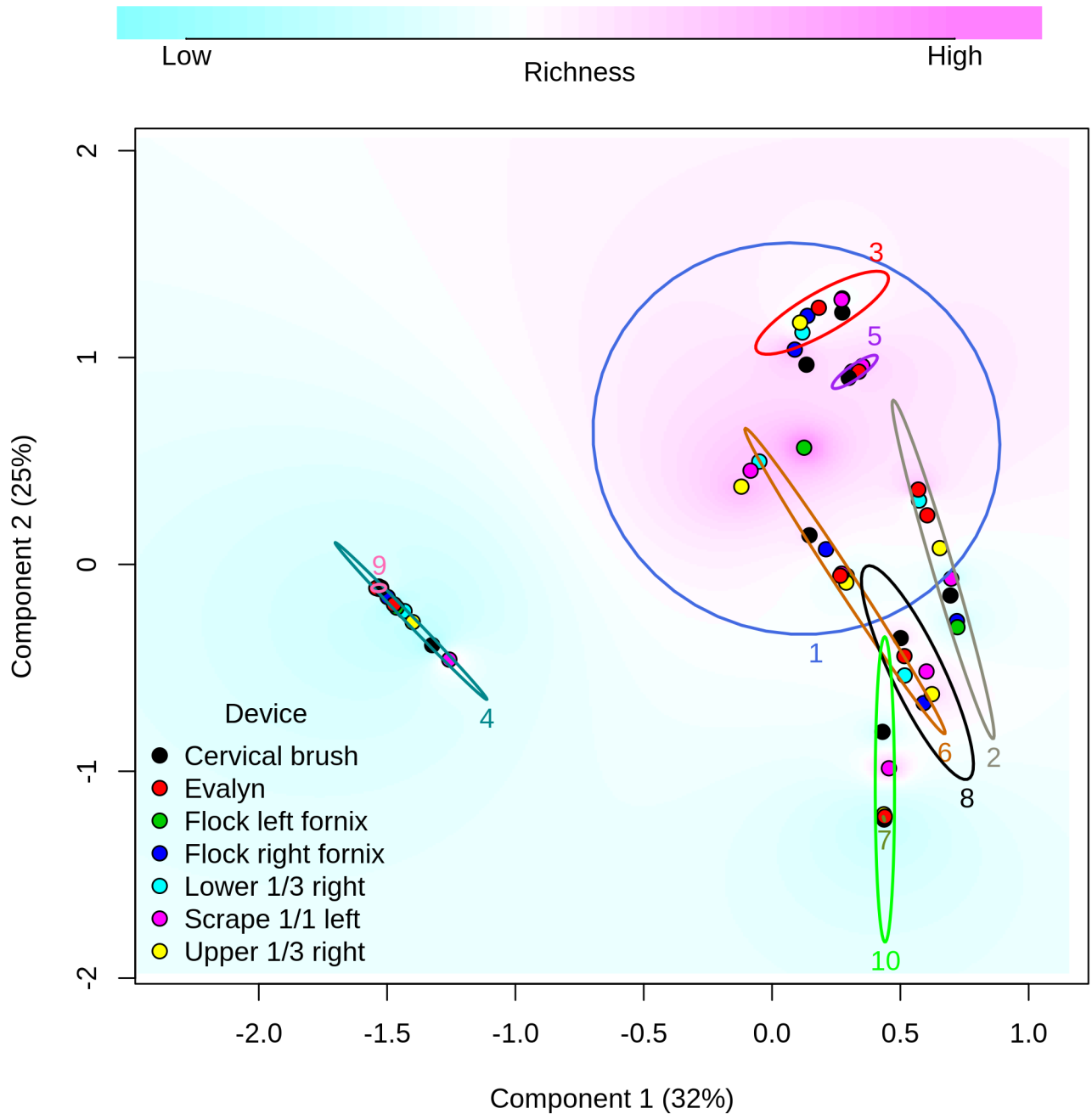
Fig 5. Principal coordinates analysis (PCoA) plot based on Bray-Curtis dissimilarity. Samples are colored by the sampling method, which explained only 2% of the microbiota variation (p = 0.89). Ellipses are drawn around every patients' samples and numbered accordingly (patient; explaining 11%, p-value 0.0001). Component 1 explains 32% of variation and is dominated by *L. crispatus*, + *L. iners* and *Ureaplasma*. Component 2 explains 25% and is dominated by *L. iners*, + *Prevotella* and -*Eggerthella*.

https://doi.org/10.1371/journal.pone.0181477.g005

most commonly employed sampling strategy in the current vaginal microbiota studies [2,40,42,43], as the reference device (Fig C in S1 Appendix). There were no significant differences, however the scape from the lower vaginal 1/3 and cervical brush showed highest dissimilarity to the reference.

The observed microbiota richness ranged from 13 to 47 (median 27) and the diversity from 1.0 to 15.5 (median 2.8) among the samples, being highly variable between the patients ($p < 0.05$), but essentially invariable between the sampling methods ($p > 0.05$ for both measures). Because of the small sample size and heterogeneity of the clinical phenotypes (Table A in S1 Appendix), we did not attempt to couple the microbiota composition to subject characteristics.

Identification of taxon-specific differences between sampling devices and sites. To identify the individual taxa that potentially differed systemically in abundance between the sampling methods, we compared the bacterial abundance across the sequenced samples. The results indicate that the flocked swab collected less *Atopobium vaginae* than cervical brush and less *Prevotella* cluster 2 than scrape 1/1 (p-values 0.048 and 0.027, fold change 1.03 and 1.28, respectively). Evalyn collected less *Prevotella* cluster 2 than scrape 1/1 (p-value 0.026, fold change 1.40). The cervical brush collected less *A. vaginae* and more *Megasphaera paucivorans/sueciensis* than the upper scrape (p-values 0.049 and 0.049, fold change 1.13 and 0.82, respectively) and less *Prevotella* cluster 2 than the scrape 1/1 (p-value 0.028 and fold change 1.33). In all cases the fold changes between the mean abundances of the devices were negligible, maximally 1.4-fold, suggesting that although statistically significant, the biological relevance of the detected species-specific differences is small.

Discussion

This study provides the first next generation sequencing–based analysis addressing in parallel the impact of different sampling devices and anatomical sites on the vaginal microbiota composition and species diversity. Our results, based on excessively high 16S rRNA gene amplicon sequence coverage of one self-sampled and six different physician-collected samples from 10 women, indicate that both the sampling device and anatomical site have little effect on the representation of the cervico-vaginal bacterial community membership and abundance. In contrast, the amount of the captured total genomic DNA, but not the proportion of bacterial DNA, is strongly dependent on the sampling device and to a lesser amount on the anatomical site of the sampling.

Among the tested sampling strategies, the choice of a sampling tool had the largest impact on the DNA yield from vaginal samples. The self-sampling brush collected a mean of 19 μg of genomic DNA per sample, vastly exceeding the DNA yields obtained with the other tools (mean ranging from 0.6 μg to 9.1 μg), making it therefore ideal for applications requiring a large amount of cells or DNA. We further observed a non-significant, descending trend in the DNA yield from fornix and cervix to vaginal walls (Fig 3B). All scrapes collected invariably low amounts of DNA (Fig 3B). Based on the visual inspection of the cell pellets during DNA extraction, the large surface area of the self-sampling brush collected several times more vaginal content than the other devices, especially the scrapes, and this evidently contributed to the large difference in the DNA yields between the devices. We did not attempt to normalize the DNA concentrations for sample weight but rather calculated the total yield of genomic DNA per each device to reflect the real-life situation. We are not aware of other studies that have compared genomic DNA yield from different types of vaginal samples. Self-collected vaginal Dacron swabs have been reported to yield 2.5 to 5 μg of total DNA [7], which is ca. 2-fold lower than our DNA yields from physician-collected, spatially constrained flocked swabs.

In order to specifically quantify the amount of bacterial DNA in the samples, we quantified the 16S rRNA gene in each sample using qPCR. To normalize for the high device-driven variation in the ab-solute DNA yields, we calculated the results per fixed amount of template DNA (1 ng). While the sampling device had the largest impact on the DNA yield, the amount of

bacterial DNA in the samples was driven strongly by inter-individual variation with minor, non-significant differences arising from the sampling methods (Fig 3C). This indicates that each woman had a characteristic bacterial load in the vagina that was robustly collected with all the tested methods. Hence, although self-sampling brush collected significantly more DNA in absolute terms, these samples had similar bacterial content to the scrapes and other samples, highlighting that the majority of the DNA collected by the self-sample is of non-bacterial origin.

We observed a trend towards a lower bacterial content in the cervix compared to the vagina, in line with a previous study [44] that collected vaginal and cervical samples from 100 women and observed higher bacterial load in the vagina measured by qPCR. We also observed some-what lower amount of bacteria in the lower third compared to the rest of the vagina. However, these differences were minor compared to the biological variation independent of the sampling strategy. In line with an earlier study [45], the BV samples had ca. 10-fold higher total bacterial load compared to other samples. In summary, our results of the quantitative analysis indicate that the vaginal DNA yield depends on the sampling method, but the bacterial counts in the vagina are regulated by biological factors, being insensitive to the choice of the sampling strategy. This also applied to the microbiota diversity and richness that were significantly affected by the sample donor, bleeding and BV-status, but not by the sampling site or device.

Our data indicate that the vaginal microbiota composition assessed by NGS is robust to variation in sampling devices and anatomical sampling sites, as the microbiota profiles and diversity were highly similar between all the tested seven sampling approaches, and sampling strategy explained only 2% of the detected microbiota variation. This study confirms and extents the previous molecular studies reporting the similarity of microbiota profiles derived from self- vs. physician-collected swabs from the mid-vagina [19], and from the physician-collected swabs from the vaginal introitus, the posterior fornix, and the midpoint of the vagina [15]. However, our results disagree with the only previous study that has addressed the impact of vaginal microbiota sampling device and anatomical sampling site in parallel in the same women [46]. The authors reported that both the vaginal sites and sampling methods introduce substantial variation to the detected microbiota profiles. Comparison of the swabs collected from the cervix, fornix and lower 1/3 of the vaginal canal of eight women showed marked, up to phylum-level microbiota differences that however were inconsistent across the study subjects.

The study by Kim et al [46] was the first molecular study to address the niche variation and potential impact of sampling method on the vaginal microbiota. Although comprehensive regarding the number of the tested sampling variables, the work was based on Sanger sequencing of clone libraries and can therefore be considered being shallow according today's standards for sequencing. Hence, the detected microbiota variability between the samples may at least partly reflect stochastic variation due low sampling depth. Our MiSeq data was saturated in terms of the read counts, and in that sense our results, showing negligible effect of the sampling device on the microbiota profile, can be considered more reliable. Therefore the single swab that has been used to sample the vaginal microbiota in majority of the large-scale studies to date [2,40,42,43] seems to be representative.

When compared to self-sampling methods used in previous studies [18,19,40,47], the Evalyn self-sampling device has two major differences: the external tube to minimize contact with vulva and to keep the sampling depth constant, and the actual sample collecting part being a silicone brush instead of a swab. We did not do any direct comparison to other self-sampling methods, since we did not want to risk losing possible spatial variation by mixing the bacteria across vaginal sites with multiple self-sampling. There are other devices for self-sampling on the market, but we chose Evalyn because of the protective tube, the Finnish-language patient

manual available and the fact that we happened to have few extra devices left from another study [48].

Based on deep Illumina MiSeq sequencing and validated bioinformatic pipeline for sequence processing, we detected altogether 45 bacterial species in the studied 10 Finnish asymptomatic women. The variation of microbial composition was high between the individuals, six out of 10 women having Lactobacillus-dominated microbiota and the rest had *Enterobacter*, *Lachnospiraceae* BVAB1, *Prevotella* or *Megasphera* dominated community (Fig 4). Four patients had *L. iners* dominated (Community State Type III; CST III) and two *L. crispatus* dominated bacteria (CST I) [15]. Our small cohort represents the first study to measure the vaginal microbiota with culture-independent methods from Finnish women. Species composition was similar to what has been reported in studies from our neighboring countries Estonia and Sweden [49,50] and in the most recent large meta-analysis of vaginal microbiota [51]. This suggests that our results on the robustness of the microbiota profiles to sampling strategy can be generalized also to healthy women even though many of the 10 women sampled in this study happened to have bacterial vaginosis and/or abnormal cells in Pap smear. The limitation of our study is the relatively small number of study subjects. Hence, we cannot prove that the sampling methods or locations would always produce equal results.

In accordance with the community-wide comparison methods, we detected very few taxon-specific abundance differences between the sampling strategies. The three bacteria (*Atopobium vaginae*, *Megasphaera paucivorans/sueciensis*, *Prevotella* cluster 2) reaching statistical significance differed between cervical brush versus flocked swab, and cervical brush versus upper 1/3 scrape. All the differences had very low fold change in relative abundance (maximally 1.4-fold) and hence the biological relevance of these differences seems virtual. With the targeted 10 patients, we did not identify bacteria that would have been unique e.g. to cervix, or to the adherent material collected from the vaginal wall by the scrapes. Also previously, the microbiota composition between cervical and vaginal samples has been reported to be similar [44]. *G. vaginalis*, an important species in the vaginal microbiota, was not covered in our analysis due to the mismatch of the forward primer 27F. Consequently, our study cannot evaluate performance of these sampling methods for *G. vaginalis* or any other bacteria belonging to order *Bifidobacteriales* [41]. The parameters used in the preprocessing pipeline filtered low abundant bacteria and thus this study does not address possible variation in the very low abundant species between sampling sites and devices.

In summary, our results indicate that different sampling strategies collect comparable vaginal microbiota composition and diversity. Hence, biological variability of the vaginal microbiota exceeds the extent of technical variation rising from the sampling collection technique (this study) and DNA extraction method [21].Therefore past and future vaginal microbiota studies employing different sampling strategies should be comparable in the absence of other technical confounders, such as sample storage, sequencing strategy or bioinformatic parameters. The Evalyn® self-sampling device performed equally well compared to samples taken by a clinician, and hence offer a good-quality microbiota sample without the need for a gynecological examination. The self-sampling option enables higher participant rates e.g. in large longitudinal epidemiological studies. Our results emphasize the fact that the DNA yield from vaginal samples is merely a qualitative readout, which cannot be used as an indication of bacterial load or microbiota diversity in the samples. However, it might have practical implications for study design, e.g. when selecting a method for other down-stream applications than NGS that require larger input sample volumes, or when aliquoting the samples for multiple assays.

Conclusion

Vaginal microbiota can be sampled many ways and this study shows that four different devices and six vaginal locations can produce comparable results with almost no differences in the microbiota composition. Ultimately the choice for optimal sampling device depends on the sampling conditions, e.g. if gynecological examination is feasible or not, and on the sampling location of interest. Combined protein and microbiota sampling from one sample should also be possible with a brush or a flocked swab, if multiple sampling is not possible, but a single sample for multiple purposes is very vulnerable and might also lead to multiple freeze-thaw cycles in the laboratory compromising the sample quality.

Ethical statement

The study was approved by the Ethical Committee of Helsinki University Central Hospital (HUCH, 21/13/03/03/2014) and all participants provided a written informed consent.

Supporting information

S1 Appendix. Additional tables and figures.
(PDF)

S2 Appendix. R-code. R commands and source code for statistical analyses.
(R)

Acknowledgments

The authors want to thank physicians and nurses who made the sample collection possible, patients who volunteered to this study, Prof. Willem de Vos for support in setting up the MiSeq analysis, Dr. Jarmo Ritari for support in the initial steps of sequence preprocessing and special thanks to Dr. Katri Korpela for developing the mare R package and continuing support in data analysis.

Author Contributions

Conceptualization: Ilkka Kalliala, Pekka Nieminen, Anne Salonen.

Data curation: Seppo Virtanen, Anne Salonen.

Formal analysis: Seppo Virtanen, Anne Salonen.

Funding acquisition: Pekka Nieminen.

Methodology: Seppo Virtanen, Anne Salonen.

Resources: Anne Salonen.

Supervision: Ilkka Kalliala, Pekka Nieminen, Anne Salonen.

Visualization: Seppo Virtanen.

Writing – original draft: Seppo Virtanen.

Writing – review & editing: Seppo Virtanen, Ilkka Kalliala, Pekka Nieminen, Anne Salonen.

References

1. Srinivasan S, Hoffman NG, Morgan MT, Matsen F a, Fiedler TL, Hall RW, et al. Bacterial communities in women with bacterial vaginosis: high resolution phylogenetic analyses reveal relationships of

microbiota to clinical criteria. PLoS One. 2012; 7: e37818. https://doi.org/10.1371/journal.pone.0037818 PMID: 22719852

2. Mitra A, MacIntyre DA, Lee YS, Smith A, Marchesi JR, Lehne B, et al. Cervical intraepithelial neoplasia disease progression is associated with increased vaginal microbiome diversity. Sci Rep. England; 2015; 5: 16865. https://doi.org/10.1038/srep16865 PMID: 26574055

3. Kyrgiou M, Mitra A, Moscicki A-B. Does the vaginal microbiota play a role in the development of cervical cancer? Transl Res. Elsevier Inc.; 2016; https://doi.org/10.1016/j.trsl.2016.07.004 PMID: 27477083

4. Hyman RW, Herndon CN, Jiang H, Palm C, Fukushima M, Bernstein D, et al. The dynamics of the vaginal microbiome during infertility therapy with in vitro fertilization-embryo transfer. J Assist Reprod Genet. 2012; 29: 105–115. https://doi.org/10.1007/s10815-011-9694-6 PMID: 22222853

5. Hyman RW, Fukushima M, Jiang H, Fung E, Rand L, Johnson B, et al. Diversity of the Vaginal Microbiome Correlates With Preterm Birth. Reprod Sci. United States; 2014; 21: 32–40. https://doi.org/10.1177/1933719113488838 PMID: 23715799

6. Borgdorff H, Tsivtsivadze E, Verhelst R, Marzorati M, Jurriaans S, Ndayisaba GF, et al. Lactobacillus-dominated cervicovaginal microbiota associated with reduced HIV/STI prevalence and genital HIV viral load in African women. ISME J. Nature Publishing Group; 2014; 8: 1781–93. https://doi.org/10.1038/ismej.2014.26 PMID: 24599071

7. Gajer P, Brotman RM, Bai G, Sakamoto J, Schutte UME, Zhong X, et al. Temporal Dynamics of the Human Vaginal Microbiota. Sci Transl Med. 2012; 4: 132ra52–132ra52. https://doi.org/10.1126/scitranslmed.3003605 PMID: 22553250

8. Aagaard K, Riehle K, Ma J, Segata N, Mistretta T-A, Coarfa C, et al. A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. PLoS One. 2012; 7: e36466. https://doi.org/10.1371/journal.pone.0036466 PMID: 22719832

9. Fettweis JM, Paul Brooks J, Serrano MG, Sheth NU, Girerd PH, Edwards DJ, et al. Differences in vaginal microbiome in African American women versus women of European ancestry. Microbiol (United Kingdom). 2014; 160: 2272–2282. https://doi.org/10.1099/mic.0.081034–0

10. Brotman RM, Shardell MD, Gajer P, Fadrosh D, Chang K, Silver MI, et al. Association between the vaginal microbiota, menopause status, and signs of vulvovaginal atrophy. Menopause. United States; 2014; 21: 450–458. https://doi.org/10.1097/GME.0b013e3182a4690b PMID: 24080849

11. Brotman RM, He X, Gajer P, Fadrosh D, Sharma E, Mongodin EF, et al. Association between cigarette smoking and the vaginal microbiota: a pilot study. BMC Infect Dis. England; 2014; 14: 471. https://doi.org/10.1186/1471-2334-14-471 PMID: 25169082

12. Dezzutti CS, Hendrix CW, Marrazzo JM, Pan Z, Wang L, Louissaint N, et al. Performance of swabs, lavage, and diluents to quantify biomarkers of female genital tract soluble mucosal mediators. PLoS One. 2011; 6: e23136. https://doi.org/10.1371/journal.pone.00231366 PMID: 21858008

13. Kim TK, Thomas SM, Ho M, Sharma S, Reich CI, Frank J a, et al. Heterogeneity of vaginal microbial communities within individuals. J Clin Microbiol. 2009; 47: 1181–9. https://doi.org/10.1128/JCM.00854-08 PMID: 19158255

14. Huse SM, Ye Y, Zhou Y, Fodor AA. A core human microbiome as viewed through 16S rRNA sequence clusters. PLoS One. 2012; 7: 1–12. https://doi.org/10.1371/journal.pone.0034242 PMID: 22719824

15. Human Microbiome Project Consortium and others. Structure, function and diversity of the healthy human microbiome. Nature. Nature Publishing Group; 2012; 486: 207–214. https://doi.org/10.1038/nature11234 PMID: 22699609

16. Arbyn M, Verdoodt F, Snijders PJF, Verhoef VMJ, Suonio E, Dillner L, et al. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: A meta-analysis. Lancet Oncol. Elsevier Ltd; 2014; 15: 172–183. https://doi.org/10.1016/S1470-2045(13)70570-9 PMID: 24433684

17. Lunny C, Taylor D, Hoang L, Wong T, Gilbert M, Lester R, et al. Self-collected versus clinician-collected sampling for chlamydia and gonorrhea screening: A systemic review and meta-analysis. PLoS One. 2015; 10: 1–23. https://doi.org/10.1371/journal.pone.0132776 PMID: 26168051

18. Srinivasan S, Liu C, Mitchell CM, Fiedler TL, Thomas KK, Agnew KJ, et al. Temporal variability of human vaginal bacteria and relationship with bacterial vaginosis. PLoS One. 2010; 5. https://doi.org/10.1371/journal.pone.0010197 PMID: 20419168

19. Forney LJ, Gajer P, Williams CJ, Schneider GM, Koenig SSK, McCulle SL, et al. Comparison of self-collected and physician-collected vaginal swabs for microbiome analysis. J Clin Microbiol. 2010; 48: 1741–8. https://doi.org/10.1128/JCM.01710-09 PMID: 20200290

20. Bai G, Gajer P, Nandy M, Ma B, Yang H, Sakamoto J, et al. Comparison of storage conditions for human vaginal microbiome studies. PLoS One. 2012; 7: e36934. https://doi.org/10.1371/journal.pone.0036934 PMID: 22655031

21. Gill C, Wijgert JHHM Van De, Blow F, Darby AC. Evaluation of Lysis Methods for the Extraction of Bacterial DNA for Analysis of the Vaginal Microbiota. PLoS One. 2016; 11: 1–16. https://doi.org/10.1371/journal.pone.0163148 PMID: 27643503

22. Nugent RP, Krohn MA, Hillier SL. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. J Clin Microbiol. 1991; 29: 297–301. PMID: 1706728

23. Salonen A, Nikkilä J, Jalanka-Tuovinen J, Immonen O, Rajilić-Stojanović M, Kekkonen R a, et al. Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. J Microbiol Methods. 2010; 81: 127–34. https://doi.org/10.1016/j.mimet.2010.02.007 PMID: 20171997

24. Nadkarni M, Martin FE, Jacques NA, Hunter N. Determination of bacterial load by real-time PCR using a broad range (universal) probe and primer set. Microbiology. 2002; 148: 257–266. https://doi.org/10.1099/00221287-148-1-257 PMID: 11782518

25. Edwards U, Rogall T, Blöcker H, Emde M, Böttger EC. Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. Nucleic Acids Res. 1989; 17: 7843–7853. https://doi.org/10.1093/nar/17.19.7843 PMID: 2798131

26. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria; 2016. Available: https://www.r-project.org/

27. Korpela K. mare: Microbiota Analysis in R Easily. R package version 1.0. [Internet]. 2016 [cited 23 Oct 2016]. 10.5281/zenodo.50310

28. Korpela K. mare: Background Reading. In: GitHub [Internet]. 2016 [cited 23 Oct 2016]. Available: https://github.com/katrikorpela/mare/blob/master/inst/mareBackground.pdf

29. Ramiro-Garcia J, Hermes GDA, Giatsis C, Sipkema D, Zoetendal EG, Schaap PJ, et al. NG-Tax, a highly accurate and validated pipeline for analysis of 16S rRNA amplicons from complex biomes. F1000Research. 2016; 5: 1791. https://doi.org/10.12688/f1000research.9227.1

30. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010; 26: 2460–2461. https://doi.org/10.1093/bioinformatics/btq461 PMID: 20709691

31. Fettweis JM, Serrano MG, Sheth NU, Mayer CM, Glascock AL, Brooks JP, et al. Species-level classification of the vaginal microbiome. BMC Genomics. 2012; 13 Suppl 8: S17. https://doi.org/10.1186/1471-2164-13-S8-S17 PMID: 23282177

32. SIMPSON EH. Measurement of Diversity. Nature. 1949; 163: 688–688. https://doi.org/10.1038/163688a0

33. Venables WN, Ripley BD. Modern Applied Statistics with S [Internet]. Fourth. New York: Springer; 2002. Available: http://www.stats.ox.ac.uk/pub/MASS4

34. Bray JR, Curtis JT. An Ordination of the upland forest community of southern Wisconsin.pdf. Ecology Monographs. 1957. pp. 325–349. https://doi.org/10.2307/1942268

35. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. vegan: Community Ecology Package [Internet]. 2016. Available: https://cran.r-project.org/package=vegan

36. Pebesma EJ. Multivariable geostatistics in S: the gstat package. Comput Geosci. 2004; 30: 683–691.

37. Fournier DA, Skaug HJ, Ancheta J, Ianelli J, Magnusson A, Maunder MN, et al. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim Methods Softw. 2012; 27: 233–249. https://doi.org/10.1080/10556788.2011.597854

38. Skaug H, Fournier D, Bolker B, Magnusson A, Nielsen A. Generalized Linear Mixed Models using "AD Model Builder" [Internet]. Available: http://glmmadmb.r-forge.r-project.org/

39. Benjamini Y and Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing [Internet]. Journal of the Royal Statistical Society. 1995. pp. 289–300. https://doi.org/10.2307/2346101

40. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, et al. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci. 2011; 108: 4680–4687. https://doi.org/10.1073/pnas.1002611107 PMID: 20534435

41. Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl Environ Microbiol. 2008; 74: 2461–2470. https://doi.org/10.1128/AEM.02272-07 PMID: 18296538

42. Fettweis JM, Brooks JP, Serrano MG, Sheth NU, Girerd PH, Edwards DJ, et al. Differences in vaginal microbiome in African American women versus women of European ancestry. Microbiology. England; 2014; 160: 2272–2282. https://doi.org/10.1099/mic.0.081034-0 PMID: 25073854

43. Lee JE, Lee S, Lee H, Song Y-M, Lee K, Han MJ, et al. Association of the Vaginal Microbiota with Human Papillomavirus Infection in a Korean Twin Cohort. Medeiros R, editor. PLoS One. United States; 2013; 8: e63514. https://doi.org/10.1371/journal.pone.0063514 PMID: 23717441

44. Ling Z, Liu X, Chen X, Zhu H, Nelson KE, Xia Y, et al. Diversity of Cervicovaginal Microbiota Associated with Female Lower Genital Tract Infections. Microb Ecol. 2011; 61: 704–714. https://doi.org/10.1007/s00248-011-9813-z PMID: 21287345

45. Zozaya-Hinchliffe M, Lillis R, Martin DH, Ferris MJ. Quantitative PCR assessments of bacterial species in women with and without bacterial vaginosis. J Clin Microbiol. 2010; 48: 1812–1819. https://doi.org/10.1128/JCM.00851-09 PMID: 20305015

46. Kim TK, Thomas SM, Ho M, Sharma S, Reich CI, Frank JA, et al. Heterogeneity of vaginal microbial communities within individuals. J Clin Microbiol. 2009; 47: 1181–1189. https://doi.org/10.1128/JCM.00854-08 PMID: 19158255

47. Santiago GLDS, Cools P, Verstraelen H, Trog M, Missine G, El Aila N, et al. Longitudinal study of the dynamics of vaginal microflora during two consecutive menstrual cycles. PLoS One. 2011; 6: e28180. https://doi.org/10.1371/journal.pone.0028180 PMID: 22140538

48. Karjalainen L, Anttila A, Nieminen P, Luostarinen T, Virtanen A. Self-sampling in cervical cancer screening: comparison of a brush-based and a lavage-based cervicovaginal self-sampling device. BMC Cancer. BMC Cancer; 2016; 16: 221. https://doi.org/10.1186/s12885-016-2246-9 PMID: 26979237

49. Drell T, Lillsaar T, Tummeleht L, Simm J, Aaspõllu A, Väin E, et al. Characterization of the Vaginal Micro- and Mycobiome in Asymptomatic Reproductive-Age Estonian Women. Fredricks DN, editor. PLoS One. United States; 2013; 8: e54379. https://doi.org/10.1371/journal.pone.0054379 PMID: 23372716

50. Shipitsyna E, Roos A, Datcu R, Hallén A, Fredlund H, Jensen JS, et al. Composition of the vaginal microbiota in women of reproductive age—sensitive and specific molecular diagnosis of bacterial vaginosis is possible? Coenye T, editor. PLoS One. United States; 2013; 8: e60670. https://doi.org/10.1371/journal.pone.0060670 PMID: 23585843

51. van de Wijgert JHHM, Borgdorff H, Verhelst R, Crucitti T, Francis S, Verstraelen H, et al. The vaginal microbiota: what have we learned after a decade of molecular characterization? PLoS One. 2014; 9: e105998. https://doi.org/10.1371/journal.pone.0105998 PMID: 25148517