RESEARCH ARTICLE

Association of key species of vaginal bacteria of recurrent bacterial vaginosis patients before and after oral metronidazole therapy with short- and long-term clinical outcomes

Ashomathi Mollin[1], Mounika Katta[1], Jack D. Sobel[2], Robert A. Akins[1]*

1 Department of Biochemistry, Microbiology, & Immunology, Wayne State University School of Medicine, Detroit, Michigan, United States of America, **2** Division of Infectious Diseases, Wayne State University School of Medicine, Detroit, Michigan, United States of America

* rakins@med.wayne.edu

Abstract

Bacterial vaginosis (BV) is associated with a state of vaginal dysbiosis typically involving depletion of otherwise dominant populations of *Lactobacillus*. The causes of this microbial succession are not known; there may be multiple causes. Standard treatment includes oral metronidazole, which typically restores *Lactobacillus* species to dominance. However, recurrence rates are high; recurrent BV patients recur 3–4 times annually and are often refractory to treatment. Our previous qPCR-based study of recurrent BV patients pointed to putatively more virulent species of *Gardnerella* that were associated with refractory responses to oral metronidazole, and less robust recovery of *Lactobacillus* species associated with recurrence after an initial period of remission. However, these associations did not account for outcomes in all patients, suggesting that other bacterial species were involved. In this follow-up study, we sequenced the V4 domain of 16S rRNA sequences of 41of these same patients pre- and posttreatment. Overall compositions among pretreatment clinical outcome groups were not different, although alpha diversity significantly decreased: refractory > recurrent > remission. Combinations of key species were associated with and prognostic for outcome. Higher pretreatment abundance of *Megasphaera lornae* together with lower abundance of *Gardnerella Gsp07* and *Finegoldia magna* predicted long term remission after oral metronidazole. Furthermore, a subset of refractory patients that did not have high levels of *Gardnerella Gsp07*, instead had elevated levels of alternative species including *Atopobium vaginae*, *Mageeibacillus indolicus (BVAB3)*, and *Prevotella timonensis*. Patients who recurred after transient remission had elevated abundance of species including *Atopobium vaginae*, *Gardnerella*, and *Aerococcus christensenii*, compared to long-term remission patients. Core bacterial species among refractory patients did not change in abundance after metronidazole, suggesting resistance or tolerance, in contrast to the loss in abundance of the same species among recurrent or remission patients. These findings have potential prognostic and therapeutic implications.

Introduction

Bacterial vaginosis (BV) is a state of dysbiosis of vaginal microbiota in which typically dominant species of *Lactobacillus* are displaced by an array of BV-associated species [1, 2]. There is no known cause for BV or for the displacement of lactobacilli, although some believe that some species of *Gardnerella* play a perhaps necessary but not sufficient role in initiating the succession [3–5]. There is some uncertainty because subpopulations of healthy (asymptomatic) women host non-*Lactobacillus* dominant vaginal microbiota [6–8], and at least some among these do not seem to be in transition to or from symptomatic BV [9]. The issue is further complicated in that recurrent BV, defined as 3 or more episodes per year, may have different causes than sporadic or incident BV [10, 11]. Nevertheless, women who achieve remission from BV upon treatment, typically with oral metronidazole, almost always concomitantly acquire a *Lactobacillus*-dominant vaginal microbiome [12–16].

For decades, *Gardnerella vaginalis* was the only species classified under the genus *Gardnerella*. As this species was found in 87% of healthy women and therefore did not satisfy Koch's postulates as the cause of BV, heterogeneity among the species was suspected [17]. Numerous cultivation–dependent and cultivation—independent studies have established that *G. vaginalis* is genetically and phenotypically heterogeneous [18–36]. Polymorphisms in its cpn60 gene cluster *G. vaginalis* strains into two groups (A & B) and corresponding clades (Clades 1–4) [23, 24, 37]. More recently, whole genome sequencing and next generation sequencing studies have identified 9–13 distinct genomospecies under the genus *Gardnerella* [25, 26]. Associations of individual *Gardnerella* clades or species with BV pathogenesis have been inconsistent across different studies. However, many have observed that the presence of multiple clades is associated with BV [15, 24, 28, 38].

Metagenomic studies in parallel with diagnostic tools have shown that while healthy and asymptomatic Caucasian and Asian women tend to have vaginal compositions that are *Lactobacillus* dominant, a subset of healthy, asymptomatic African American and Hispanic women tend to have vaginal compositions that resemble BV. Together, these studies have also indicated that a given microbial profile is not a definitive indicator for BV; species composition is not enough to distinguish BV-positive from BV-negative patients, even as defined by Amsel criteria [6, 8, 9, 39–41]. Anaerobic species present in high prevalence in BV positive women can also be present at lower prevalence in healthy women. Furthermore, species that are detected at low prevalence in healthy women are poor markers for BV diagnosis or prognosis because they are not sufficiently prevalent in symptomatic BV patients [41]. Finally, changes by orders of magnitude occur daily in vaginal bacterial compositions [42–46]. These are some of the factors that complicate the identification of bacterial species that are diagnostic of BV or prognostic of recurrence.

Individual studies vary as to what bacteria are considered important in the diagnosis or prognosis of BV. Broadly, NGS and qPCR studies have proposed that species under the genera *Atopobium*, *Dialister*, *Gardnerella*, *Lactobacillus* (*L. iners*), and *Prevotella* are be associated with incidental BV; *Corynebacterium*, *Lachnospiraceae_1(BVAB1)*, and *Ruminococcaceae_2 (BVAB2)* are associated with recurrent BV; and *Megasphaera lornae* is associated with both incidental and persistent BV [8, 41, 47–55]. These variations may be attributed to different demographic factors such as geographical location, incidence of sex, ethnicity, and personal hygiene, but they also suggest that there may be many paths to BV. Commercial BV qPCR tests are available and have varying targets. For example, LabCorp offers a research-lab semi-quantitative test for *Atopobium vaginae*, *Ruminococcaceae_2(BVAB2)*, and *Megasphaera lornae*, Diagnostic Laboratory of Oklahoma tests for *Atopbium vaginae*, *Gardnerella vaginalis*, *Mycoplasma genitalium*, *Mycoplasma hominis*, and *Ureaplasma urealyticum*, and the BD Max

system Vaginal Panel is a qPCR test for *Lactobacillus spp. (L. crispatus* and *L. jensenii), Gardnerella vaginalis, Atopobium vaginae, Ruminococcaceae_2(BVAB2),* and *Megasphaera lornae.*

The most widely used therapy for BV in the United States is oral or vaginal metronidazole, followed by oral clindamycin. Yet, the recurrence rate after metronidazole is as high as 80 percent [10, 56–58]. Several factors may contribute to this. One may be reinfection, although this may be more relevant in incident BV rather than recurrent BV [57, 59–62]. Recurrence may occur due to metronidazole resistance; several species or strains within species are resistant in vitro [11, 49, 57, 63–66]. Another recurrence mechanism may be inactivation or sequestration of metronidazole by one or several species [67] or by sequestering bacterial populations within biofilms, which are likely to have initiated with *Gardnerella* species [27, 28, 35, 36, 68, 69]. A cultivation–dependent study proposed that the clinical response post–metronidazole treatment is highly dependent on the pretreatment abundance of *Lactobacillus spp.* versus BV–associated species, particularly *Gardnerella*. Employing the concept of bacterial competition, this study suggested that elevated pretreatment abundance of *Lactobacillus spp.* relative to BV–associated species lowered the efficacy of metronidazole, and patients who had higher pretreatment abundance of BV–associated species relative to *Lactobacillus spp.* experienced a boosted metronidazole efficacy leading to a favorable outcome [67].

This study focuses on recurrent BV patients as an extension of our study initially reported in 2019 in which we saw three basic responses to treatment with oral metronidazole: refractory, recurrent, and remission [15, 70]. This qPCR-based study found that *Gardnerella Gsp07* and *G. swidsinskii / G. leopoldii* were strongly associated with refractory and recurrent responses, and a more robust recovery of *Lactobacillus* species, most often *L. iners*, was associated with remission patients compared to recurrent patients, even though both groups were in remission at the posttreatment visit [70]. The current study examined a subset of the initial patient samples by next-generation sequencing of the V4 domain of the 16S ribosomal RNA to provide a deeper understanding of changes to the vaginal microbiome before and after oral metronidazole. Our hypothesis is that there are key species differences in the pretreatment vaginal microbiota of remission patients, in addition to lower abundance of the virulent *Gardnerella* species, that are associated with a more enduring remission, and that key species differences at the posttreatment visit are prognostic of remission versus recurrent outcomes.

Materials and methods

Patient cohort

This study employed a portion of the patient cohort involved in prior studies conducted by the laboratories of Dr. Robert Akins and Dr. Jack Sobel at the School of Medicine and the Vaginitis Clinic at Wayne State University [70, 71]. This study used de-identified samples from women who had signed written, informed consent after discussion and witnessing by the research nurse. This and post-study data analysis using these samples were approved by the WSU IRB board, approval number 040314M1F. The following requirements were fulfilled when enrolling recurrent BV patients: three or more episodes of symptomatic BV in the preceding year, presence of acute symptoms of BV at enrollment, minimum age of 18, premenopausal, absence of mixed vaginal infections, withholding of vaginal products during the course of the trial, heterosexuality, use of condoms during sexual intercourse during the trial period, disclosure of any unprotected sexual activity, voluntary sexual refrainment 48 hours prior to any visit, and alcohol abstinence. A patient was further considered a BV–positive if they tested positive for at least three of the four Amsel criteria [72], and if they presented with symptoms such as malodor, vaginal discharge, and vaginal itching. (S1 Table in S1 File). Other exclusions are described in the original studies. The standard of care (SOC) treatment called for oral

metronidazole therapy, twice a day, over a span of 7 days. These patients then returned for clinical assessment approximately 7–14 days post treatment for their second visit, and then once every month for up to 9 months. Along with these monthly visits, patients were required to perform daily vaginal swabs and communicate any engagement in sexual intercourse, use of sexual protection, and time of menses [70].

Patients with persistent BV symptoms and who tested positive for three out of four Amsel criteria at the second visit were classified as refractory BV patients. Patients who showed no BV symptoms at the second visit but later returned with BV symptoms and a positive Amsel criteria score were recognized as recurrent BV patients. Lastly, those who neither showed BV symptoms nor satisfied the Amsel criteria and were able to maintain this state for at least three months, were considered to have reached long-term remission.

Our initial study utilized 43 (18 refractory, 16 recurrent, and 9 remission) of the 74 patients from the Akins-Sobel study to assess the patient clinical outcomes during days 1–14 (7 days of treatment and 7 days posttreatment) in terms of *Gardnerella* and *Lactobacillus* species relative abundance by qPCR [70]. This study will be making use of 41 patients (18 refractory, 12 recurrent, and 11 remission) to assess the vaginal microbiome relative to patient clinical outcomes before and after treatment (Fig 1). Limited resources prevented us from using all available patient samples for these analyses.

DNA sequencing

DNA from vaginal swabs was prepared as described [15]. Next-generation sequencing of the 16S ribosomal RNA gene, V4 region. Sequencing was performed at the Michigan State University Research Technology Support Facility Genomics Core (Lansing MI). The V4 hypervariable region of the 16S rRNA gene was amplified using dual indexed Illumina compatible
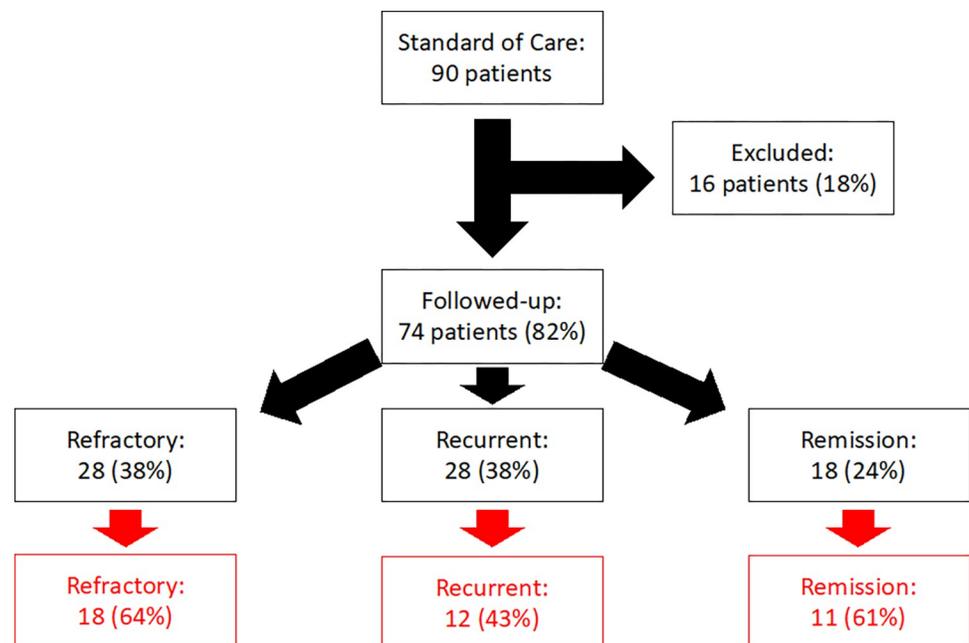
Fig 1. Patient enrollment flow diagram. This study sequenced 41 of 74 patients (represented by the red boxes) from the Akins-Sobel study [71]. Refractory patients returned to the second visit with BV symptoms and tested positive for 3 out of the 4 Amsel criteria. Recurrent patients briefly achieved remission, at the second visit, prior to recurrence. Remission patients reached clinical remission, at visit two, and were able to maintain this state for at least 3 months.

primers 515f/806r as described [73]. PCR products were normalized using Invitrogen Sequal-Prep DNA Normalization plates and the products recovered from the plates were pooled. This pool was size-selected with AMPureXP magnetic SPRI beads and tested for quality and quantity using a combination of Qubit dsDNA HS, Caliper LabChipGX HS DNA and Kapa Illumina Library Quantification qPCR assays. Sequencing of the pooled amplicons was performed on an Illumina MiSeq v2 standard flow cell using a 500 cycle v2 reagent cartridge. Custom Sequencing and index primers were added to appropriate wells of the reagent cartridge as described [73]. Base calling was done by Illumina Real Time Analysis (RTA) v1.18.54 and output of RTA was demultiplexed and converted to FastQ format with Illumina Bcl2fastq v2.18.0. There was an average of 34,154 SD 13,754 paired-end reads per sample with QC scores >30, typically 251–253 bp per trimmed sequence.

The processing of these sequences from FASTQ files was performed with Mothur software V1.35.0 [74, 75] following the default protocol for the MiSeq platform, including pairing, trimming of ends, exclusion of low-quality reads, removal of chimeras, de novo clustering into operational taxonomic units (OTU's) at a 3% variation threshold, and identification of a representative sequences of each OTU at the midpoint of polymorphemic sequences within the OTU. 5011 OTU's were assigned to genus or species levels in most cases, using the *assignTaxonomy* function from the DADA2 package within the R programming language [76] using the reference Silva training set file, "silva_nr99_v138.1_wSpecies_train_set.fa.gz" [77, 78], downloaded from the Github website [79] classified 2673 OTUs (out of 5011) to the species level.

To assign the best possible classification to the remaining OTUs, they were grouped and aligned by phylum to the 16S rRNA database in the Ribosomal Database Project (RDP) [80]. These sequences, along with the public representative isolates from the RDP, were then incorporated into phylogenic trees using the Maximum Likelihood (ML) algorithm, bootstrap phylogenetic method, and the Tamura-Nei model within the MEGA X software [81, 82]. A total of 16 different phylogenic trees were made–one for each phylum assessed (S1 Fig). Once these trees were made, the unassigned OTUs were renamed according to their location on the phylogenic tree relative to nearby cultured representative isolates. A summary table of all OTU sequences and taxonomic labels, related sequences in the databases, and a phylogenetic tree that positions all included reads and best hits are provided as S2 Table in S1 File and S2 Fig. Representative reads that were not assigned to species level were assigned a species or lowest possible taxonomic label if the RDP alignment was >97% to a reference species. In instances where multiple related species shared the same V4 sequence, or where there was heterogeneity within a phylogenetic branch, a representative species was used along with an asterisk symbol (*); species included by these * symbols are provided in S3 Table in S1 File. Phylogenies that were less characterized, e.g. Acidobacteria, were grouped by the RDP Classifier tool [83] with a confidence threshold marker of 60%. OTUs that could be assigned to the species level were classified to the lowest higher taxonomy level and labeled with a number (Taxa_#) to indicate that it belonged to a distinct phylogenetic branch with that level. The numbers were assigned at the end of the taxa to indicate separate groups of the higher order. Once renaming of the unassigned OTUs to the lowest taxonomy level was completed, these RDP Seqmatch results were compared to the classifications done by the *assignTaxonomy* algorithm. For those names, where both naming algorithms disagreed, a decision was made based on the RDP Classifier tool and the NCBI MOLE-BLAST tool. The highest common taxonomy level was retained if a decision using the different tool could not be made.

Together, the newly formed taxonomy table, NGS counts table, and patient metadata table associated with the pretreatment and posttreatment samples were loaded into MicrobiomeAnalyst [84, 85], MetaboAnalyst 5.0 [86, 87], and SPSS Statistics (Version 27–28, Armonk, NY: IBM Corp.). Within Microbiome Analyst, data was uploaded in two ways. In the first, the

OTU counts data were filtered to condense OTUs identified as a single species into a representative OTU, yielding 865 OTUs (S4 Table in S1 File). These data, after filtering out OTU's with low abundance, prevalence, and variance in Microbiomeanalyst, yielded 96 species that were the basis for comparative analysis (S5 Table in S1 File). We also input all 5011 OTU counts with a taxonomic classification scheme that filled all levels with the best identification, e.g. if an OTU was identified only to Genus, the species name given was Genus_sp. These OTUs were subjected within Microbiomeanalyst to a stricter filter: minimum 4 reads per OTU, 20% minimum prevalence, and 10% minimum variance among samples, leaving 51 OTUs (S6 Table in S1 File). We used unfiltered data from the 5011 OTU input file for alpha diversity analysis; all other analysis used the 865 OTU input dataset. Filtered raw counts were then normalized, using Total Sum Scaling (TSS), within the software. TSS normalizes the raw counts by dividing individual counts with the maximum sum of all counts and then multiply it with ten million for convenience. Data input into MetaboAnalyst 5.0 was generated in Microbiome Analyst as normalized abundance data, which we rescaled to the mean number of reads of all samples (21724), and log10 transformed. Both analytical software packages generate multiple analytical models from which it was determined whether overall bacterial compositions of the patient samples, or key species, at either visit, were associated with clinical outcome.

Statistics. Species that held statistical significance (p–value <0.05 and FDR < 0.1) were flagged by the different analytical tools such as classical univariate analysis, sparse high–throughput sequencing analysis, differential abundance analysis, and linear discriminate analysis effect size, and random forest in MicrobiomeAnalyst as potential diagnostic and prognostic markers. Using these flagged species, univariate and multivariate receiver operating characteristic (ROC) analyses were conducted in MetaboAnalyst for each outcome comparison. Species were removed multivariate ROC analysis, one at a time, eliminated if the exclusion did not result in a lower area under the curve (AUC) value, but replaced if it did lower the AUC value. This process was repeated until a maximum AUC value was achieved. This process identified the smallest combination of species that provided the largest significant AUC values. ROC–derived thresholds, sometimes adjusted manually, were determined for each species contributing to the optimized multivariant ROC analysis. Patient samples that were above or below these threshold values were summed and analyzed by the Fisher's exact tests, conducted in Graph Pad Prism, which determined prognostic values (version 6.07 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com).

Results

Differences in alpha diversity associated with clinical outcome in both pre- and posttreatment samples

We detected 329 species among all 39 pretreatment patient samples, averaging 57 SD 18 species per sample. Among all 38 posttreatment patients we detected 298 species, for an average of 49 SD 17 species per sample.

Species richness, the estimated median number of species by the Chao index (Fig 2A) ranged from 31–68 species. The Chao index distributions were not significantly different between outcome groups, but the downward trend among posttreatment samples, refractory>recurrent>remission, was significant. There was a significant posttreatment decrease in the Chao index only among remission patients. The Shannon index comparisons among pretreatment patients, while not significantly different by ANOVA, showed a significant downward trend, refractory > recurrent > remission (Fig 2B). This decrease was more pronounced and achieved significance at the posttreatment visit. Shannon indices decreased significantly from pre- to posttreatment for both recurrent and remission patient samples, but not for refractory
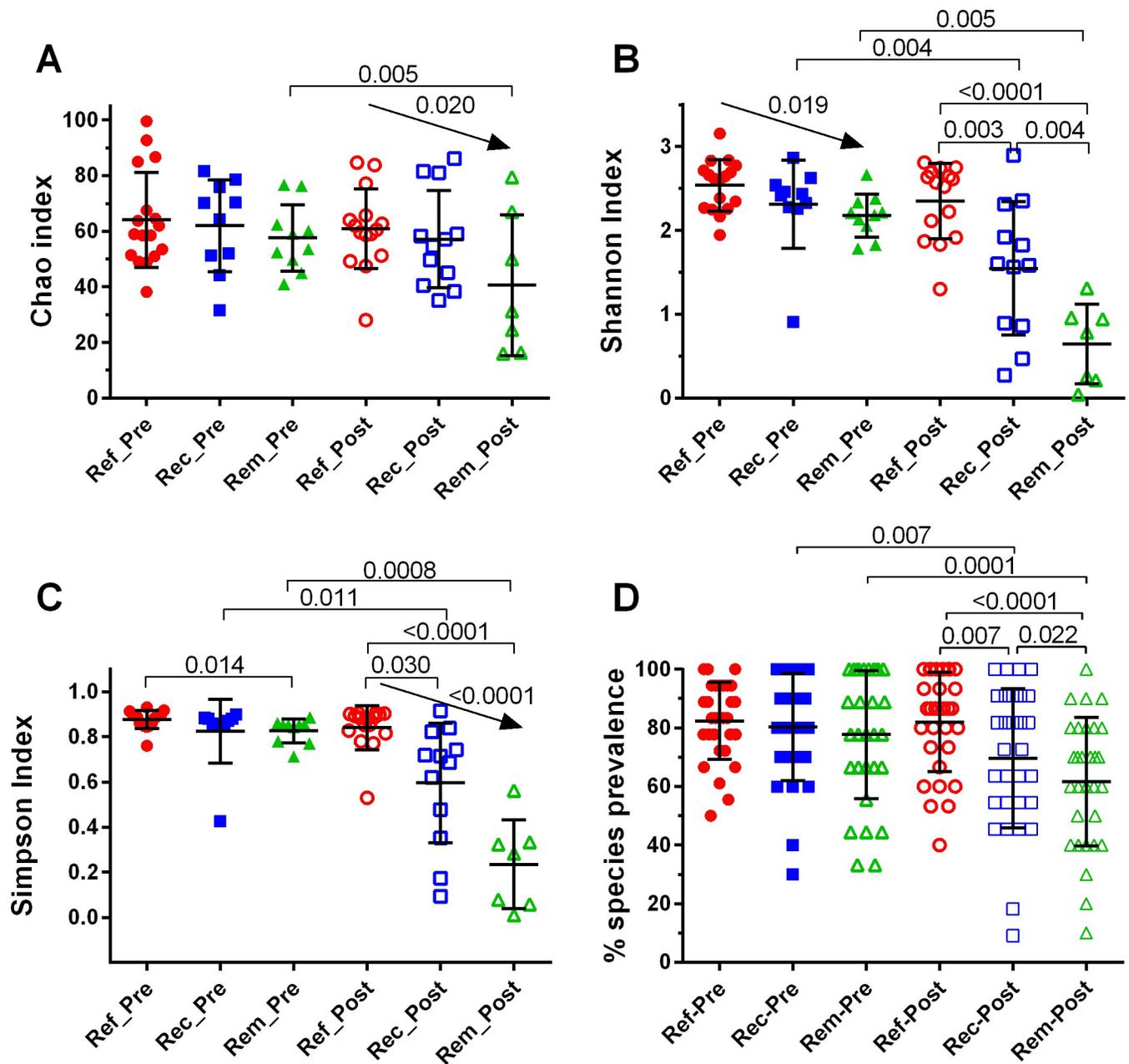
Fig 2. Alpha diversity indices among clinical outcome groups pre- and post-treatment. Index values for Chao, Shannon, and Simpson were determined using Microbiomeanalyst.ca with the unfiltered 5011 OTU dataset and Total Sum Scaling (TSS). % species prevalence was calculated based on the percent of species-positive samples versus total samples within each group and treatment, including only the 30 most prevalent species. Scatterplots showing mean and standard deviation, and statistical analyses were performed in GraphPad Prism; only significant p values or trends are shown. Between outcome comparisons of normally distributed Chao indices were analyzed by ANOVA with Holm-Sidak's multiple comparisons test; non-normal Shannon and Simpson indices were analyzed by Kruskal-Wallis tests, with Dunn's multiple comparisons tests. Pretreatment versus posttreatment comparisons were performed for Chao and % species prevalence by paired t-tests and for Shannon and Simpson by Wilcoxon matched-pairs signed rank tests (GraphPad Prism).

https://doi.org/10.1371/journal.pone.0272012.g002

patient samples. Comparisons with the Simpson index generated conclusions like those from the Shannon index (Fig 2C). Another metric of richness, the % prevalence of the top 30 most prevalent species, was also like comparisons using the Shannon index, except that pretreatment outcome groups were not differentiated (Fig 2D)

Differences in diversity metrics were sufficiently pronounced to have moderate prognostic value. Pretreatment patients above a ROC-based Shannon index threshold of 2.23 with an AUC of 0.815 (Metaboanalyst.ca), that is, those with elevated diversity and evenness, were predicted to have a refractory or recurrent outcome with 84% accuracy (95% CI 68% to 94%, Fisher's exact test, MedCalc software). Recurrent or remission posttreatment patient samples above a ROC-based Shannon index threshold of 1.49 were predicted to have a recurrent rather than a remission outcome with an accuracy of 76% (95% CI 53% to 92%). That is, those with elevated diversity and evenness, were predicted to have a refractory or recurrent outcome with 84% accuracy, 95% CI 68% to 94%, by Fisher's exact test [88]. Recurrent or remission posttreatment patient samples above a ROC-based Shannon index threshold of 1.43 (AUC = 0.845, Metaboanalyst.ca) were predicted to have a recurrent rather than a remission outcome with an accuracy of 79%, 95% CI 54% to 94 [88] (S7 Table in S2 File).

Overall differences in composition (beta diversity) were only evident between posttreatment outcome groups

Principal coordinate analysis (PCoA) of pretreatment samples did not separate patients into distinct groups consistent with clinical outcomes (Fig 3A). By PCoA, only 6 of 39 samples from refractory and recurrent patient samples were placed outside of a shared single cluster (95% confidence interval zone) that included the remaining 33 patient samples from all outcome groups. In contrast, most posttreatment refractory and remission samples were readily separable into two distinct groups by PCoA (Fig 3B). Samples from recurrent patients were split, approximately half clustered with refractory and half with remission samples.

In an alternative approach to investigate diversity among outcome groups, we manually classified patients by their most abundant genus (MAG; Fig 4). This process identified 5 MAG groups. *Gardnerella*, *Prevotella*, *Shuttleworthia* (including *Lachnospiraceae_1(BVAB1)*, and *Sneathia* groups were almost exclusively from symptomatic patients, but the 4 groups were not
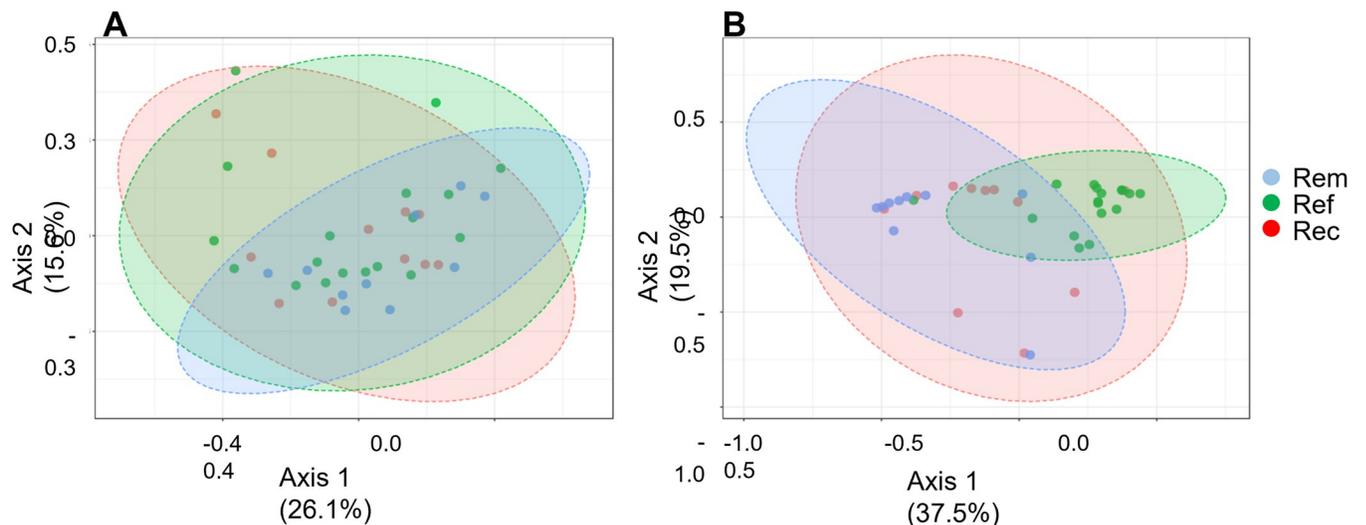
Fig 3. Grouping by overall compositions of patients by PCoA. PCoA analyses were performed separately for pretreatment (A.) and posttreatment (B.) using the MicrobiomeAnalyst.ca platform, inputting raw counts, normalizing with Total Sum Scaling (TSS) without further transformation, and labeling samples after plotting by clinical outcome group. Sample distance was measured using the non-phylogenic Bray-Curtis distance method. The shaded zones represent the range of two-component values with the 95% confidence intervals of observed data for each outcome group. Statistical significance was measured using Permutational ANOVA (PERMANOVA). (A) F-value: 0.96757;R-squared: 0.051012; p-value < 0.504 (top left). (B) F-value: 5.6556;R-squared: 0.24424; p-value < 0.001 (top right).

https://doi.org/10.1371/journal.pone.0272012.g003
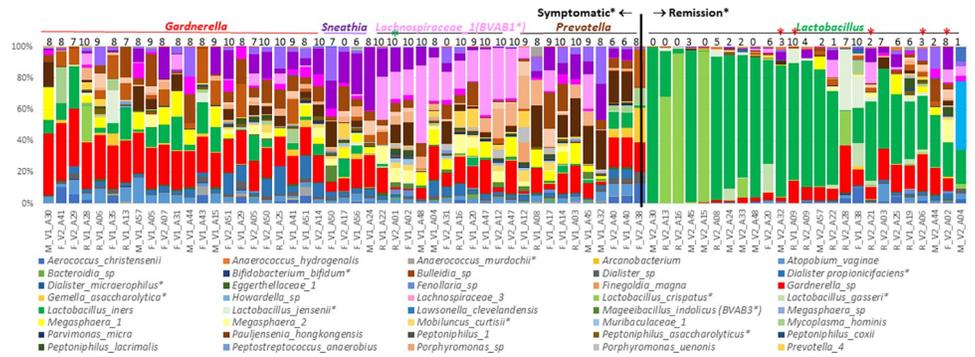
Fig 4. Manual clustering of patient samples by their most abundant genus (MAG). Relative abundances of genera were plotted as a stacked barchart in Excel and sorted by the 5 most MAG's, color coded as shades of their title fonts. Within each group, samples were sorted by relative abundance of the MAG. Nugent scores are displayed above each sample. *: exceptional sample, Amsel-negative but non-*Lactobacillus* MAG sample. *: exceptional samples, symptomatic *Lactobacillus* MAG samples. Label codes: F = refractory patient, V1 = pretreatment sample, V2 = posttreatment sample, A## = patient identifier.

https://doi.org/10.1371/journal.pone.0272012.g004

associated with clinical outcome. In contrast, the fifth group, *Lactobacillus*, was composed mostly of samples from asymptomatic posttreatment patients. Within this group, recurrent patients were typically those in the lower range of % *Lactobacillus* compared to remission patient samples. However, there was no association of incidence of clinical outcome with pretreatment MAGs (chi sq p = 0.876). This association at posttreatment was significant (chi sq p = 0.002) and this was driven by the high incidence of Amsel-negative recurrent and remission patients in the *Lactobacillus* MAG: 9 of 10 posttreatment recurrent patients, and 10 of 10 posttreatment remission patients, were in the *Lactobacillus* MAG group. We observed no significant differences of Nugent scores between the MAG categories per outcome groups at pre- or post-treatment (Kruskal-Wallis, Dunn's multiple comparisons tests). Mean Nugent scores of post-treatment Lactobacillus MAG groups varied (refractory 7 SD 1.4; recurrent 3.5 SD 2.7; remission 1.5 SD 1.2) but these differences were not significant due to intragroup variability and low numbers of refractory patients.

Collectively, these data indicate that overall compositions of symptomatic recurrent BV patients at pretreatment were not predictive of clinical outcome.

Partial least squares discriminant analysis (PLSDA) identifies candidate biomarkers of treatment outcome

We removed species whose variation was not associated with clinical outcome to focus on a short list of species that were associated by performing reiterative PLSDA analysis. For pretreatment samples, supervised classification modeling approaches, such as partial least squares discriminant analysis (PLSDA) or sparse PLSDA, did not differentiate samples by outcome group, as they had misclassification rates greater than 59%, indicating poor cross-validation (S8a Table in S3 File). While reiterative linear discriminate modelling by sPLSDA (repeating the analysis using the most differentiating species from the initial analysis) yielded some improvements, classification error rates were still > 41%) (S8b Table in S3 File).

We sought to improve discrimination of pretreatment samples by integrating data on the proportion of total *Gardnerella*, identified by V4 sequences, with previous qPCR data that partially resolved *Gardnerella* to species levels. In our previous qPCR study [15], clade-specific primers targeting the *Gardnerella* cpn60 gene were able to quantify individual or subsets of genomospecies among some of the sample samples subjected to NGS sequencing in the
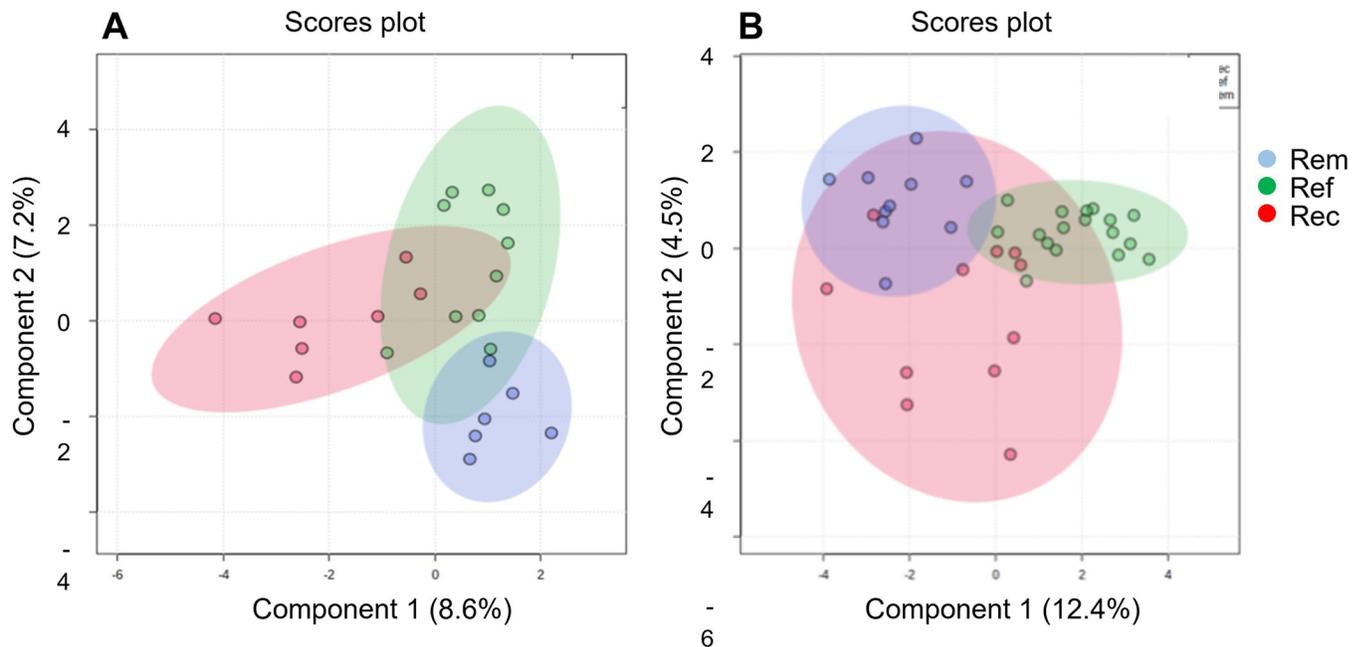
Fig 5. Sparse-PLSDA analysis of pre- and posttreatment samples by clinical outcome group. Normalized, unfiltered data was log10 transformed in MetaboAnalyst.ca platform. (A) pretreatment reiterative sPLSDA analysis using the integration of the major *Gardnerella* species. The top 23 species S8a Table in S3 File from the initial sPLSDA analysis were input into a second sPLSDA analysis to generate the result shown here. (B) posttreatment reiterative sPLSDA analysis using the unintegrated NGS counts dataset. The shaded zones represent the range of two-component values with the 95% confidence intervals of observed data for each outcome group.

https://doi.org/10.1371/journal.pone.0272012.g005

current study. To incorporate the qPCR data into our sequencing dataset, normalized relative abundance per patient was calculated as the relative abundance by qPCR of each *Gsp* species relative to the total bacterial abundance, based on qPCR with broad-spectrum primers [15], times the normalized *Gardnerella* counts present in our sequence reads (S9 Table in S1 File). Performing two cycles of sPLSDA modeling on this integrated dataset, differentiated pretreatment patient samples by outcome group and yielded a sPLSDA classification error rate of 26% (Fig 5A). Key species for this resolution were *Gardnerella*_Gsp07, *Megasphaera lornae*, and *Aerococcus christensenii*. (S8c, S8d Table in S3 File). These conclusions of pretreatment samples are provisional, however, due to low sample numbers; posttreatment sample numbers were too low in the integrated *Gsp* dataset to allow an analysis using these major *Gardnerella* species at the posttreatment visit, could not be performed.

In contrast to pretreatment samples, sparse PLSDA separated refractory from most recurrent posttreatment patient samples without resorting to the *Gsp* subset of data (Fig 5B). This analysis had a classification error rate of only 15.8% and pointed to a small number of species that best differentiated the outcome groups. A subset, but not all, of these species, *A. vaginae*, *M. hominis*, *Gardnerella*, *Dialister sp*, *Gemella asaccharolytica**, and *L. jensenii** were components of multivariant biomarkers that were prognostic of clinical outcomes. Differentiation of refractory patient samples at posttreatment was expected, since recurrent and remission patients were both in clinical remission and typically show prevalence or dominance of one or two species of *Lactobacillus*. However, this early detection of elevated abundance of species of patients who were transiently in remission versus those who achieved long-term remission was not a given; examples include *Lachnospiraceae_1(BVAB1)*, *Corynebacterium spp*, and *Ureaplasma urealyticum* (Fig 5B, S8e, S8f Table in S3 File).

Overview of core microbiota by clinical outcome and treatment status

We characterized the core microbiomes of these groups, defined as species present in at least 20% of samples a > 1% of sample abundance for any outcome group, using heat maps that categorize prevalence levels at varying ranges of percent abundance (Fig 6). For example, 75–100% of pretreatment refractory patient samples have *Gardnerella* at ≥ 0.1% abundance but none of these samples have *Gardnerella* at ≥ 32% prevalence. First, overall compositions of the
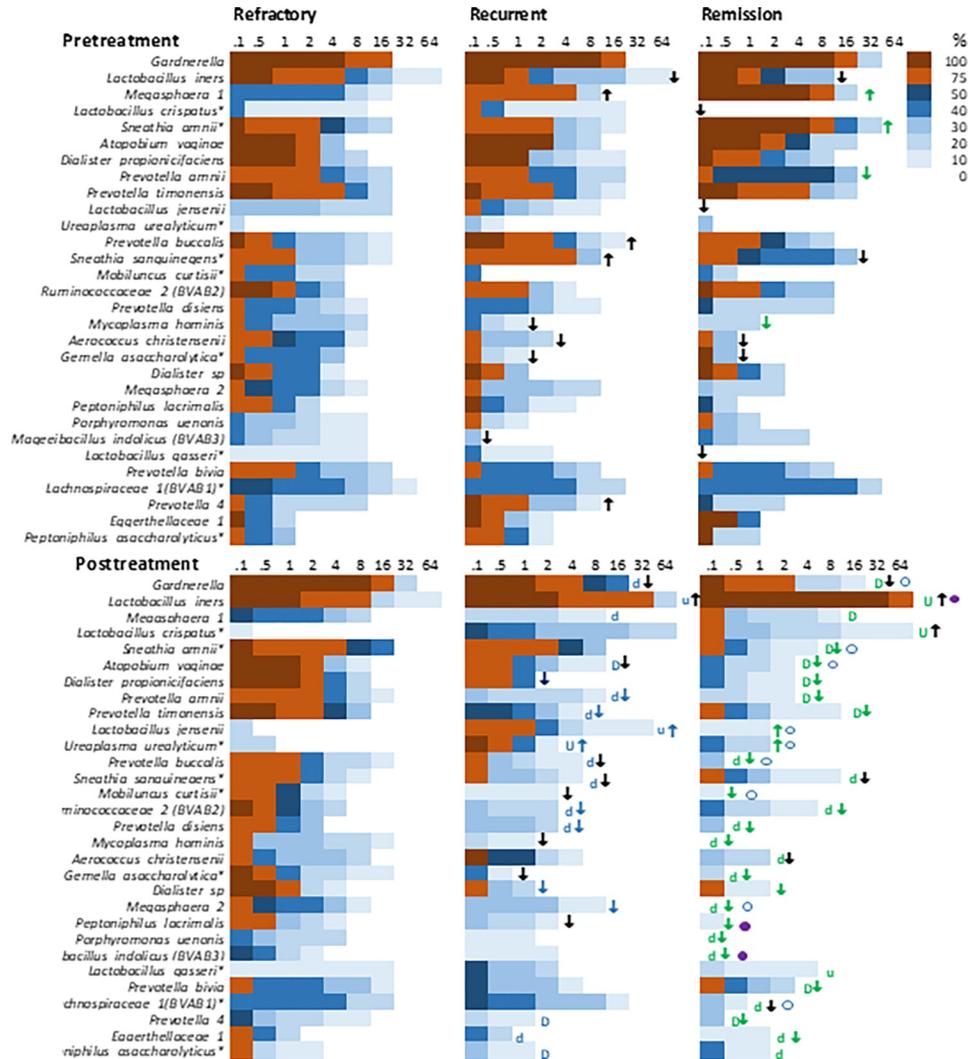
Fig 6. Heat maps of abundance and prevalence of core microbiota of refractory, recurrent and remission patient samples pre- and posttreatment. The color scale classifies upper limits of % prevalence categories. Percent abundance is binned as indicated on the x-axis. Outcome group comparisons (horizontal): ↑ or ↓ indicates a large visual up or down change in either abundance or prevalence of the species relative to that seen in the refractory patient, color coded if the difference was significant (p value <0.05 by abundance using the Kruskal Wallis test, Bonferroni corrected, or by prevalence using ChiSq or Fisher's exact tests, SPSS). The open circle indicates a large visual change in abundance or prevalence of a remission sample species relative to recurrent samples; a closed circle indicates that the difference was significant (p <0.05, by abundance using Kruskal Wallis test, Bonferroni corrected, or by prevalence using ChiSq or Fisher's exact tests, SPSS). Pre- versus posttreatment comparisons of change in relative abundance (top versus bottom panel): U and D denote large up or down changes, capitalization indicates the difference was significant (p < 0.05 by abundance assessed by Wilconox rank match pairs tests or by prevalence assessed by paired samples proportions tests with Wald's test of significance (SPSS).

https://doi.org/10.1371/journal.pone.0272012.g006

3 outcome groups at pretreatment were similar, since only 4 species differed significantly among remission versus refractory patients; other species varied but did not achieve statistical significance. Secondly, overall changes in composition among refractory patients did not change significantly pre- versus posttreatment. In contrast, many species among remission patients declined in prevalence and/or abundance after treatment, and recurrent patient samples were intermediate, where many species decreased posttreatment but were less likely than remission patients to achieve statistical significance. Thirdly, many species were less prevalent and/or less abundant among posttreatment recurrent patients and especially those in remission, compared to refractory patients, and many were also lower among remission than recurrent patients.

Looking at a more detailed level at the most abundant species, *Gardnerella* was not differentially distributed among outcome groups at pretreatment, and the observed reductions among posttreatment recurrent and remission patients relative to refractory patients did not achieve statistical significance. *L. iners* trended higher among pretreatment refractory patients but not significantly so; increases at posttreatment were significant among remission patients, but intermediate among recurrent patients. Among less abundance species at pretreatment, the biggest difference in species composition between outcome groups was the higher prevalence of abundant *Megasphaera lornae* among pretreatment recurrent and especially remission compared to refractory patient samples. Also of note was the higher prevalence and abundance of *Sneathia amnii*, and lower prevalence of *Prevotella amnii**, among pretreatment remission patients compared to refractory and recurrent patients. In pretreatment samples, several species were less abundant or prevalent among recurrent and remission patient samples compared to refractory, but none of these achieved statistical significance.

Among posttreatment samples, there was a significant downward trend in the relative abundance of *Dialister sp* and *M. curtisii** among the posttreatment refractory>recurrent>remission samples. Specifically, the relative abundance of *Dialister sp* was significantly higher in refractory patients compared to the other two outcome groups and the relative abundances of *M. curtisii** and *G. asaccharolytica** were significantly higher in refractory patients than in remission patients. While *Megasphaera lornae* was reduced in prevalence among posttreatment remission patients, the difference between remission and recurrent patients was statistically significant. Though the relative abundance of *A. vaginae* and *A. christensenii* were lower among remission patients compared to recurrent patients, the difference was not statistically significant at the posttreatment visit. However, there was a significant downward trend in the abundance of A.vaginae at the posttreatment visit, across refractory>recurrent>remission samples. Like *A. vaginae*, other signature species of symptomatic BV, such as *Prevotella spp* and *Sneathia spp* were differentially distributed refractory > recurrent > remission patient samples. Several species dropped from detectable at pretreatment to below the limits of detection among all posttreatment remission samples, including *M. hominis*, *Dialister* 1, and *P. uenonis*, however, most species were not eliminated. Conversely, *L. crispatus*, and *L. jensenii* were below detection among almost all pretreatment remission patient samples but were detected at posttreatment, most dramatically *L. crispatus*: in ~10% of the posttreatment samples *L. crispatus* had risen to 64% of total bacteria. At the posttreatment visit, while *L. jensenii* was reduced in remission patients, there was no significant difference compared to the recurrent patients.

Differential response of common species to metronidazole between outcome groups

No species in refractory patient samples changed dramatically in abundance or prevalence after metronidazole treatment, whereas those same species typically decreased among

recurrent and especially remission patient samples (Fig 6). Consistently, alpha diversity metrics showed that there were no significant changes in pre- versus posttreatment refractory patient samples, whereas overall prevalence of species among posttreatment recurrent and remission samples decreased significantly, by 15–20% (Fig 2). This point is augmented in before-after plots of individual species per refractory patient (S4 Fig). Among refractory patient samples, no species showed a robust or prevalent up or down trend following treatment. However, some of these same species shifted dramatically after treatment among recurrent and remission patient samples. Most notable among remission patient samples were decreases in abundance of *Gardnerella*, *Megasphaera lornae*, *Sneathia amnii*, with less dramatic but significant decreases in *A.vaginae*, *Prevotella timonensis*, and *Dialister sp*, and robust and pervasive increases in *L. iners*. These changes were also noted among recurrent samples, but to lesser extents and incomplete prevalence. While no upward or downward trend was found in the abundance of *U. urealyticum* and *P. asaccharolytica** among refractory and remission patients, these two species increased and decreased significantly, respectively, after metronidazole treatment among recurrent patients. Species that persistently decreased or did not increase after treatment in association with remission but not recurrence: *U. urealyticum*, *Gardnerella*, *S. amnii*, and *Megasphaera lornae*, along with a more robust increase in *L. iners*. Species that persistently decreased in association with a non-refractory response: *A. vaginae*, *Prevotella_4*, *Gardnerella*, *S. amnii*, *Megasphaera lornae*, *P. timonensis**, and *P. asaccharolytica**.

Key species were associated with outcome among pretreatment sample groups

Despite the lack of distinction of overall compositions of outcome groups at pretreatment (Fig 6), small numbers of key species were significantly different or trended differently and were therefore candidates for prognostic markers or etiologic agents, alone or in combinations. These species were preliminarily identified by multiple statistical approaches since no single approach is definitive (see Methods) and by visualizations on scatterplots (S3 Fig) and on heat tree maps if species comparisons generated a Wilcoxon Rank Sum test p value ≤ 0.05 (S10 Table in S4 File). We combined species identified by all approaches with those from a filtering approach in which 96 species derived by filtering 868 vaginal species in Microbiomeanalyst with the criteria of 2 reads minimum, 10% minimum prevalence, and 5% minimum variance across samples. These species were analyzed by multivariate ROC analysis; species were removed from inclusion in the multivariate ROC analysis one at a time and eliminated if the exclusion did not result in a lower AUC value. Patient samples above or below ROC thresholds were counted as input for Fisher's exact tests to generate the prognostic statistics in Table 1. Data indicated that pretreatment abundance of 3 species (*M. indolicus*, *M. curtisii**, and *M. hominis*) above thresholds predicted a refractory response with 80.6% accuracy. Pretreatment elevated levels of *Finegoldia magna* and *Gardnerella Gsp07*, combined with lower levels of *Megasphaera lornae*, predicted a refractory or recurrent response with 87% accuracy. At posttreatment, elevated levels of *Gardnerella*, *Dialister sp*, *G. asaccharolytica**, and *M. curtisii**, combined with lower levels of *L. iners*, were diagnostic of the refractory response to treatment with 97.3% accuracy. Among Amsel-negative patients at posttreatment, elevated levels of 5 species, including *Gardnerella* and *A. vaginae*, with lower levels of *Megasphaera lornae*, predicted a subsequent recurrence with 95% accuracy.

Previously, we reported that higher pretreatment abundance of *Gardnerella Gsp07* was associated with refractory or recurrent outcomes based on qPCR data [15]. This we confirmed with our current sequencing data (S9 Table in S1 File). Specifically, *Gardnerella Gsp07* was

Table 1. Species combinations with significant prognostic value.

	Pretreatment				Posttreatment		
Group 1	Ref[1]	RefRec[1]	Ref Hi[1]		Ref[2]		Rec[2]
Group 2	RecRem	Rem	Ref Lo		RecRem		Rem
Species 1	*Mageeibacillus indolicus*(BVAB3) (0.003)	*Finegoldia magna* (0)	*Atopobium vaginae* (2.3)	*Mageeibacillus indolicus*(BVAB3) (3.49)	*Mageeibacillus indolicus*(BVAB3) (3.49)	*Dialister sp* (1.4)	*Aerococcus christensenii* (0.092)
Species 2	*Mobiluncus curtisii** 0.228)	*Gardnerella Gsp07* (0.007)	**Megasphaera 1 (0.009)**	*Anaerococcus mediterraneensis* (2.51)	*Prevotella timonensis* (5.63)	*Gardnerella* (9.2)	*Atopobium vaginae* (0.460)
Species 3	*Mycoplasma hominis* (0.003)	**Megasphaera1 (8.3)**	*Prevotella timonensis* (4.6)			*Gemella asaccharolytica** (0.276)	*Gardnerella* (6.9)
Species 4			**Sneathia sanguinegens (0.921)**			**Lactobacillus iners (27.6)**	*Lactobacillus jensenii** (0.092)
Species 5						*Mobiluncus curtisii** (0.138)	**Megasphaera 1 (0.018)**
Species 6							*Megasphaera 2* (0)
ROC AUC	0.759	0.964	0.689	0.93	0.887	0.947	0.86
95% CI	0.553–0.94	0.816–1	0.327–1	0.492–1	0.333–1	0.833–1	0.561–1
Sensitivity	0.824	0.882	1	1	0.833	0.938	1
95% CI	0.566–0.962	0.636–0.985	0.715–1	0.541–1	0.359–0.996	0.698–0.998	0.715–1
Specificity	0.789	0.833	0.833	1	1	1	0.9
95% CI	0.544–0.939	.359–.996	0.359–0.996	0.398–1	0.398–1	0.839–1	0.555–0.997
PPV	0.778	0.938	0.917	1	1	1	0.917
95% CI	0.524–0.936	0.698–0.998	0.615–0.998	0.541–1	0.478–1	0.782–1	0.615–0.998
NPV	0.833	0.714	1	1	0.8	0.955	1
95% CI	0.586–0.964	0.290–0.963	0.478–1	0.398–1	0.284–0.995	0.772–0.999	0.664–1
Odds Ratio	17.5	37.5	84.3	117	33	444	146
95% CI	0.31–92.5	2.77–508	2.93–2425	1.94–7067	1.06–1024	16.9–11660	5.29–4008
Accuracy	0.806	0.870	0.941	1.00	0.9	0.973	0.952
95% CI	0.640–0.918	0.664–0.972	0.713–0.999	0.69–1	0.55–1.00	0.858–0.999	0.762–0.999
p-value	0.0006	0.0034	0.001	0.0048	0.0476	<0.0001	<0.0001

Thresholds in () after each species is the % abundance value from ROC analysis used to count patient samples; normal font of species indicates that Group 1 abundance is ≥ the threshold and bold font indicates that group 1 abundance is < the threshold. ROC analysis performed in Metaboanalyst.ca. Accuracies were calculated at MedCalc. Fisher tests and statistics were performed with GraphPad Prism. [1] Combination of two out of total number of key species identified to give best predictive values. [2] Combination of three out of total number of key species identified to give best predictive values. Three alternative sets of biomarker species were determined for the Ref Hi versus Ref Lo groups. Abbreviations: Ref: refractory; Rec: recurrent; Rem: remission; RefRec: patients from both refractory and recurrent outcome groups; RecRem: patients from both recurrent and remission outcome groups; Ref Hi: subset of refractory patients with high abundance (>1%) of Gardnerella Gsp07; Ref Lo: subset of refractory patients with low abundance (≤1%) of Gardnerella Gsp07; CI: confidence interval. Three alternative sets of biomarker species were determined for the Ref Hi versus Ref Lo groups.

https://doi.org/10.1371/journal.pone.0272012.t001

significantly more abundant among refractory samples than in remission samples and there was a downward trend in the abundance of *Gardnerella Gsp07* from refractory>recurrent>remission samples (S3 Fig). However, there were exceptions, i.e. some refractory patients presented with low *Gardnerella Gsp07* abundance similar to recurrent or remission patients [15]. In this study, we saw that other species were differentially distributed among the refractory patients with abundant *Gardnerella Gsp07* (GspHi, >1% relative abundance) versus

those low, <1% Gsp07 (GspLo). sPLSDA modelling revealed a clear distinction between the two groups, with classification error rates as low as 20%, implying significant amounts of variation between the two groups (S8g Table in S3 File). Notably, refractory GspLo patients had higher abundance levels of *Mageeibacillus indolicus (BVAB3)*, *P. timonensis*, and *Anaerococcus mediterraneensis*, and reduced abundance of *Gardnerella swidsinski_leopoldii* and, by definition, lower *Gardnerella Gsp07*. These patients may have an alternative form of BV in which species other than Gsp07 marks or determines their status and their refractory response.

Increased relative abundance of *L. iners* or total *Lactobacillus* species at the posttreatment visit was associated with satisfactory early clinical response but did not predict later recurrence

The relative abundance of *L. iners* and combined *Lactobacillus* species increased significantly after treatment among recurrent and remission patient samples by 3.5 to 21-fold, respectively, but not among refractory samples (Fig 7). Mean abundance of *L. crispatus* trended upward after treatment among remission and recurrent patient samples but did not achieve statistical significance; *L. crispatus* decreased significantly among refractory patient samples after treatment. Although mean abundance among remission posttreatment patient samples was higher than recurrent posttreatment samples, the differences were not significant (p = 0.343) and therefore this metric alone was not prognostic for recurrence. Importantly, species of *Lactobacillus* individually or collectively were not significantly different in abundance among outcome groups at the pretreatment visit, indicating that these levels were not a major determinant of whether a patient will have a successful outcome to treatment. Additionally, pretreatment *Gardnerella/Lactobacillus* ratios did not generate significant or useful predictive values, in contrast to other studies [67]. Other individual species of *Lactobacillus* were not significantly different in abundance among outcome groups or visits.

Discussion

This study found many associations of potential therapeutic significance and with implications for pathobiology involved in poor responses to oral metronidazole, summarized in Table 2. Decreased evenness, that is, presence of fewer and more dominant species, was associated with better clinical outcomes of both pretreatment and posttreatment samples, indicating that increased dysbiosis is less likely to respond well to treatment. This is consistent with results of an independent study [89]. PCoA data indicated that overall composition of pretreatment samples was not associated with clinical outcome, likely because variance of most species is not related to outcome. In contrast, this data indicated that overall composition of posttreatment samples differentiated refractory from remission as expected, showing also that some recurrent samples were remission-like, and others were refractory-like in composition. SPLSDA analysis, together with other approaches (Classical Univariate Analysis, Metagenomeseq, EdgeR, DEseq2, and LefSe) generated a short list from which key species were identified by multivariate ROC analysis to predict clinical outcome, with modest accuracy at the pretreatment visit, and with high accuracy in predicting recurrence at the posttreatment visit. These species constitute potential prognostic biomarkers. We also observed a lack of response of all core species to oral metronidazole among refractory patients, within the limits of our sample times. This and the biomarker species have implications for the pathogenesis of recurrent BV.

At two extremes of this cohort, we have patients who were refractory to oral metronidazole versus those who achieved long term remission, despite their history of recurrent BV. A signature, prognostic feature of most refractory patients before treatment was a combination of
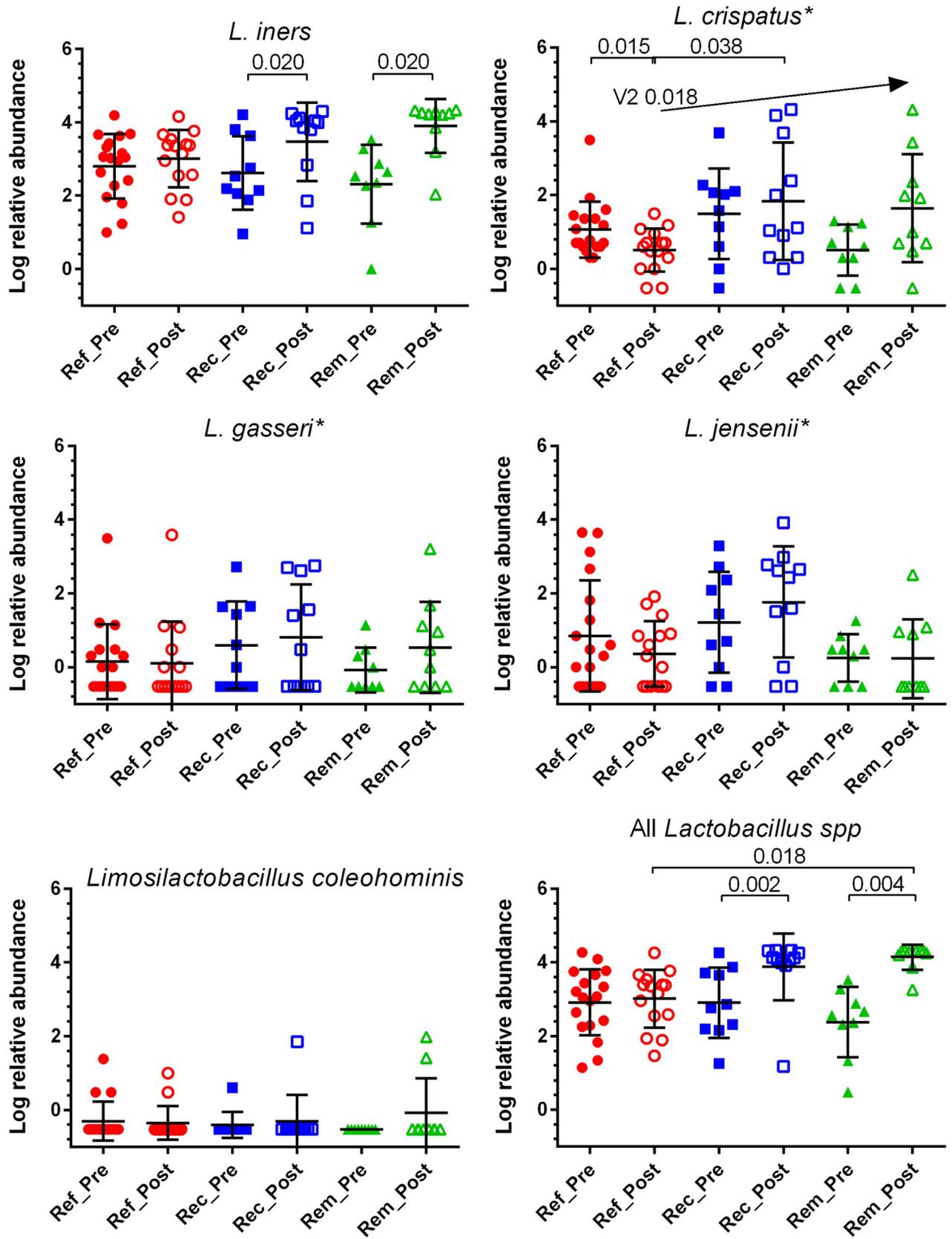
Fig 7. Relative abundance of *Lactobacillus* species pre- and posttreatment among outcome groups. Normalized relative abundances of all *Lactobacillus* species were summed per outcome group. Pre- versus posttreatment samples for each of the 3 outcome groups were compared by the related samples Wilconox signed rank test. Cross-group comparisons among pre- and posttreatment samples were analyzed by Kruskal-Wallis tests of independent samples with post-hoc pairwise tests; p values were adjusted for multiple comparisons with the Bonferroni correction. * indicates significant difference. Analyses were performed with SPSS V28, graphed with GraphPad Prism V6.07.

higher prevalence and abundance of *Mageeibacillus indolicus (BVAB3)*, *Mycoplasma hominis*, or *Mobiluncus curtisii** compared to recurrent or remission patients. A signature, prognostic feature of pretreatment remission patients was higher abundance of *Megasphaera lornae* and lower abundance of *Gardnerella Gsp07* and *Finegoldia magna*. Most of these species considered alone were not accurate predictors of clinical outcome at pretreatment; the sole exception was that elevated *Megasphaera lornae* predicted remission with 78% accuracy.

This study adds perspective with the finding that patients achieving remission were characterized at pretreatment by higher abundance and 100% prevalence of *Megasphaera Type 1*, now reclassified as *Megasphaera lornae* sp. nov. [90]. This is partly supported by a previous report that *Megasphaera* was associated with a successful clinical outcome to vaginal metronidazole, although this study used intravaginal metronidazole, was short-term, and did not focus

Table 2. Summary of major findings.

Outcome	Pretreatment	Posttreatment
Refractory	High species diversity and evenness	Persistently high species diversity and evenness
	Prognostic (81%): ↑of 2 of 3 of *Mageeibacillus indolicus (BVAB3)*, *Mycoplasma hominis*, and *Mobiluncus curtisii**	**Diagnostic (97%)**: Higher abundance in 3 of 5 of *Gardnerella*, *Dialister sp*, *G. asaccharolytica*, and *Mobiluncus curtisii** with ↓ of *Lactobacillus iners*
	↑ of *Gardnerella Gsp07* OR ↑of *Mageeibacillus indolicus (BVAB3) with Prevotella timonensis* or *Anaerococcus mediterraneensis*	No decreases in abundance of core species, including *Lactobacillus spp.*
	Overall microbial composition not associated with clinical response	Distinct microbial composition from remission patients, overlapping composition with some recurrent patients
	Lactobacillus spp. composition not associated with clinical response	
Recurrent	High species diversity and intermediate evenness	Moderately decreased species diversity and evenness
	Prognostic: absence of refractory and remission prognostic biomarkers	**Prognostic (95%)**: ↑ in 3 of 5 species, including *Gardnerella*, *Atopobium vaginae*, *Aerococcus christensenii*, and *Megasphaera_2*, and ↓ of *Megasphaera lornae*
	Overall microbial composition not associated with clinical response	Intermediate decrease in abundance of many species which were unchanged in refractory patients; associated with an increased abundance of *Ureaplasma urealyticum* and a decreased abundance of *Peptoniphilus asaccharolytica**
	Lactobacillus spp. composition not associated with clinical response	Intermediate increase in abundance of *Lactobacillus iners*
Remission		Species that persistently decreased in association with a non-refractory response: *Atopobium vaginae*, *Prevotella_4*, *Gardnerella*, *Sneathia amnii*, *Megasphaera lornae*, *Prevotella timonensis**, and *Peptoniphilus asaccharolytica**
	High species diversity and less evenness	Largest decrease in species diversity and evenness
	Prognostic (87%): ↑ of *Megasphaera lornae* and lower abundance of *Gardnerella Gsp07* or *Finegoldia magna*	**Prognostic (95%)**: ↓ in 3 of 5 species, including *Gardnerella*, *Atopobium vaginae*, *Aerococcus christensenii*, and *Megasphaera_2*, and ↑ of *Megasphaera lornae*
	Prognostic (78%): ↑ of *Megasphaera lornae* alone predicts a remission response	Largest decreases in species abundance, uniquely *Ureaplasma urealyticum*, *Gardnerella*, *Sneathia amnii*, and *Megasphaera lornae*
	Overall microbial composition not associated with clinical response	Distinct microbial composition from refractory patients, overlapping composition with some recurrent patients
	Lactobacillus spp. composition not associated with clinical response	Highest increase in abundance of *Lactobacillus iners*

↑: higher abundance relative to other outcome groups; ↓: lower abundance relative to other outcome groups. Increase or decrease: posttreatment changes in abundance relative to pretreatment. Times New Roman font: true of both posttreatment recurrent and remission samples. Prognostic or diagnostic features and % accuracies in bold.

on recurrent BV patients [54]. As a working model, we suggest that there are at least two sub-types of recurrent BV; recurrent BV among remission patients is driven by *Megasphaera lornae* and perhaps *Sneathia amnii*, rather than by more virulent *Gardnerella* species. One mechanism by which *Megasphaera lornae* may drive symptomatic BV is its ability to deplete vaginal lactic acid, shown in vitro [90]. This a property shared with species of *Megasphaera* in the gut [91, 92] or in vaginal models [93] and would reverse the lactic acid-mediated inhibition of most BV-associated species normally imposed by *Lactobacillus* [94–97]. *Megasphaera lornae* might also interfere with biofilm production, likely initiated by *Gardnerella*, which would make vaginal species more susceptible to metronidazole [7–12]. However, evidence for involvement of *Megasphaera* in vaginal biofilm is limited and sporadic [98, 99]. Patients whose BV is marked by *Megasphaera lornae* may be more likely to achieve long term remission because it is more susceptible to metronidazole than *Gardnerella* species [68, 100–102], evidenced here by its decline in pairwise comparisons of posttreatment to pretreatment abundance (Fig 7). Conversely, refractory patients are more likely to fail to respond to oral metronidazole because their BV is caused by more virulent and resistant strains such as *Gardnerella Gsp07*, instead of by *Megasphaera lornae*; consistently, *Gardnerella* species did not decline in our pairwise comparisons. Virulent *Gardnerella* species, or a covariant, may also contribute to resistance in novel ways. Recurrent patients are intermediates in this model; their lower abundance of *Gardnerella* and *Megasphaera lornae* allowed only transient remission.

It was a surprising feature of the comparisons of pre- versus posttreatment abundance, that no species showed significant declines among refractory patients whereas many of these same species decreased significantly among recurrent or remission patients after treatment. Since samples were not analyzed over the ~14 days between pretreatment and posttreatment, species in refractory samples may have declined then rebounded; these would then be uniformly tolerant rather than resistant. It is possible, but seems unlikely, that all species in these refractory patients were resistant or tolerant to metronidazole and were of different lineages than the same species in remission patients. It is more plausible that metronidazole in refractory patients was being inactivated or sequestered, as previously suggested [67]. However, in our cohort there was no correlation between abundance of pretreatment *Lactobacillus* abundance and clinical outcome, so that aspect of the published model does not apply here. We suggest that instead the inactivation/sequestration of metronidazole may be due to species or combinations of species that were notably more abundant and prevalent in the refractory group. *Gardnerella Gsp07* was significantly more prevalent and abundant in pretreatment refractory patients [15], so it or its covariants may be responsible for this process. Multivariate ROC analysis suggests that a combination of *Mageeibacillus indolicus (BVAB3)*, *Mycoplasma hominis*, or *Mobilucus curtisii** are possible sources of this activity. Of these, *Gardnerella* was shown to be capable of inactivation of metronidazole in vitro [67]. However, *Gardnerella Gsp07* alone cannot account for the lack of responsiveness of core species to metronidazole, because it was observed in both the Gsp07Hi and Gsp07Lo refractory patient subgroups.

Others have reported that the relative abundance of pretreatment *Lactobacillus*, or the ratio of *Gardnerella* to *Lactobacillus* 16S DNA [67, 103] or RNA [104], was associated with clinical outcome. Neither of these associations were significant in our cohort, nor in our larger cohort analyzed by qPCR [15]; abundance of total *Lactobacillus spp* or *L. iners* at pretreatment or among Amsel-negative posttreatment patients was not prognostic of clinical outcome. In another study, BV recurrence at one month was associated with *Enterococcus*, *L. crispatus*, *Ureaplasma*, *Aerococcus*, and *L. jensenii* [55], only minimally overlapping our associated key species, We attribute these differences to the apples to oranges comparison of the studies; cohorts in the published studies did not have prior histories of BV recurrence whereas ours

were recurrent BV patients, and we are assigning outcomes based on much longer monitoring (3–9 months) of recurrent and remission status. Similarly, our comparisons of posttreatment patient samples obtained from refractory versus combined Amsel-negative patients in recurrent and remission outcome groups, is very different from comparing healthy women to BV patients who acquire BV only sporadically.

We also saw at posttreatment that combinations of key species differed between recurrent and remission patients, both in clinical remission at that time. Combinations of higher levels of 3 of 5 species, including *Gardnerella* and *A. vaginae*, and lower levels of *Megasphaera lornae*, predicted later recurrence with 95% accuracy. Species in recurrent samples at this stage were also significantly more evenly distributed. Upon validation, these biomarkers could be used to continue or modify the standard treatment regimen.

This study has limitations, chiefly in the small numbers and single location of patients in each group, which means that conclusions, while significant, are contingent on confirmation in a larger and multi-site study. The patients are predominantly African-American with histories of frequent recurrence, so conclusions may not apply to other groups. Our recurrent patients were slightly older than remission patients, so an argument that they are different not by vaginal bacterial composition but by age-related factors will need to be addressed when data from more patients in each group are available. The OTU-based approach to species clustering ignores sequence heterogeneity within single or related OTU's, which could mask associations particularly among less abundant variants. The 16S rRNA, V4 domain sequence does not resolve all vaginal species, although it is likely that species that share V4 sequences are related. ROC-based analysis of the association of key species combinations with clinical outcome is susceptible to overfitting, something which will require a larger and more diverse study to resolve. Any contributory role of vaginal epithelial cell biofilm in women with BV was not explored. This study was focused on oral metronidazole therapy, so conclusions cannot be extrapolated to patients treated with vaginal metronidazole or oral clindamycin. The study was limited to heterosexual women due to power considerations, given our limited enrollment, and therefore conclusions cannot be extrapolated to homosexual or bisexual women. We did not have measures to ensure compliance to the patient protocol, so it is possible that refractory patients simply did not adhere to the oral metronidazole regimen; this is made less likely given that these patients performed and submitted daily swab samples in this interval, along with information such as menses and coitus. Finally, the arguments pertaining to the apparent lack of response to metronidazole among refractory patients are limited by lack of sequencing data on daily samples from the end of treatment at day 7 to our "posttreatment" visit at day 14. This limitation means that we cannot distinguish between resistance, which would have resulted in no decreases in relative abundance all throughout the 7 day course of oral metronidazole, from tolerance, which could result in reduced relative abundance during treatment, but followed by recovery beginning after treatment through the second clinic visit at day 14.

Evidence is presented that vaginal microbial composition or microbiota measured during florid BV and immediately following recommended metronidazole therapy has prognostic significance and is potentially useful in selecting therapeutic antimicrobial agents, both with regard to dose and duration of therapy. Of note is evidence of antimicrobial resistance or tolerance suggestive of metronidazole inactivation or sequestration among refractory patients.

Supporting information

S1 Fig. Phylogenetic trees of 16 individual phyla.
(PDF)

S2 Fig. Phylogenetic trees all representative OTUs.
(PDF)

S3 Fig. Relative pre- and post-treatment abundances.
(PDF)

S4 Fig. Percent abundance of species pre- & posttreatment.
(PDF)

S1 File. S1 Meta Data, S2 Phylogenetic trees all representative OTUs, S3 Lists of species included in species* groups, S4 Counts of sequences condensed to 865 species/OTUs for compositional analyses, S5 Counts of sequences filtered to 96 species/OTUs for comparative analyses, S6 Counts of sequences filtered to 51 species/OTUs for comparative analyses, S9 Counts of sequences of a subset of pretreatment samples that distribute *Gardnerella* counts into groups of *Gardnerella* species.
(XLSX)

S2 File.
(XLSX)

S3 File.
(XLSX)

S4 File.
(XLSX)

Author Contributions

Conceptualization: Ashomathi Mollin, Jack D. Sobel, Robert A. Akins.

Data curation: Robert A. Akins.

Formal analysis: Ashomathi Mollin, Mounika Katta, Jack D. Sobel, Robert A. Akins.

Funding acquisition: Jack D. Sobel, Robert A. Akins.

Investigation: Jack D. Sobel, Robert A. Akins.

Methodology: Ashomathi Mollin, Mounika Katta, Jack D. Sobel, Robert A. Akins.

Project administration: Jack D. Sobel, Robert A. Akins.

Resources: Jack D. Sobel, Robert A. Akins.

Supervision: Robert A. Akins.

Validation: Ashomathi Mollin, Jack D. Sobel, Robert A. Akins.

Visualization: Ashomathi Mollin, Jack D. Sobel, Robert A. Akins.

Writing – original draft: Ashomathi Mollin, Mounika Katta, Jack D. Sobel, Robert A. Akins.

Writing – review & editing: Ashomathi Mollin, Mounika Katta, Jack D. Sobel, Robert A. Akins.

References

1. Jung HS, Ehlers MM, Lombaard H, Redelinghuys MJ, Kock MM. Etiology of bacterial vaginosis and polymicrobial biofilm formation. Crit Rev Microbiol. 2017; 43(6):651–67. https://doi.org/10.1080/1040841X.2017.1291579 PMID: 28358585

2. Chen X, Lu Y, Chen T, Li R. The Female Vaginal Microbiome in Health and Bacterial Vaginosis. Front Cell Infect Microbiol. 2021; 11:631972. https://doi.org/10.3389/fcimb.2021.631972 PMID: 33898328

3. Muzny CA, Taylor CM, Swords WE, Tamhane A, Chattopadhyay D, Cerca N, et al. An Updated Conceptual Model on the Pathogenesis of Bacterial Vaginosis. J Infect Dis. 2019; 220(9):1399–405. https://doi.org/10.1093/infdis/jiz342 PMID: 31369673

4. Muzny CA, Schwebke JR. Pathogenesis of Bacterial Vaginosis: Discussion of Current Hypotheses. J Infect Dis. 2016; 214 Suppl 1:S1–5. https://doi.org/10.1093/infdis/jiw121 PMID: 27449868

5. Schwebke JR, Muzny CA, Josey WE. Role of *Gardnerella vaginalis* in the pathogenesis of bacterial vaginosis: a conceptual model. J Infect Dis. 2014; 210(3):338–43. https://doi.org/10.1093/infdis/jiu089 PMID: 24511102

6. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, et al. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci U S A. 2011; 108 Suppl 1:4680–7. https://doi.org/10.1073/pnas.1002611107 PMID: 20534435

7. Zhou X, Brown CJ, Abdo Z, Davis CC, Hansmann MA, Joyce P, et al. Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women. The ISME journal. 2007; 1(2):121–33. https://doi.org/10.1038/ismej.2007.12 PMID: 18043622

8. 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. Microbiol (Reading, England). 2014; 160(Pt 10):2272–82. https://doi.org/10.1099/mic.0.081034-0 PMID: 25073854

9. Gajer P, Brotman RM, Bai G, Sakamoto J, Schutte UM, Zhong X, et al. Temporal dynamics of the human vaginal microbiota. Sci Transl Med. 2012; 4(132):132ra52. https://doi.org/10.1126/scitranslmed.3003605 PMID: 22553250

10. Vodstrcil LA, Muzny CA, Plummer EL, Sobel JD, Bradshaw CS. Bacterial vaginosis: drivers of recurrence and challenges and opportunities in partner treatment. BMC Med. 2021; 19(1):194. https://doi.org/10.1186/s12916-021-02077-3 PMID: 34470644

11. Sobel JD. Recurrent bacterial vaginosis, relapse or reinfection: the role of sexual transmission. BJOG. 2021; 128(4):768. https://doi.org/10.1111/1471-0528.16471 PMID: 33241645

12. Mayer BT, Srinivasan S, Fiedler TL, Marrazzo JM, Fredricks DN, Schiffer JT. Rapid and Profound Shifts in the Vaginal Microbiota Following Antibiotic Treatment for Bacterial Vaginosis. J Infect Dis. 2015; 212(5):793–802. https://doi.org/10.1093/infdis/jiv079 PMID: 25676470

13. Mtshali A, San JE, Osman F, Garrett N, Balle C, Giandhari J, et al. Temporal Changes in Vaginal Microbiota and Genital Tract Cytokines Among South African Women Treated for Bacterial Vaginosis. Front Immunol. 2021; 12:730986. https://doi.org/10.3389/fimmu.2021.730986 PMID: 34594336

14. Verwijs MC, Agaba SK, Darby AC, van de Wijgert J. Impact of oral metronidazole treatment on the vaginal microbiota and correlates of treatment failure. Am J Obstet Gynecol. 2020; 222(2):157.e1–.e13. https://doi.org/10.1016/j.ajog.2019.08.008 PMID: 31404542

15. Turner E, Sobel JD, Akins RA. Prognosis of recurrent bacterial vaginosis based on longitudinal changes in abundance of *Lactobacillus* and specific species of *Gardnerella*. PloS one. 2021; 16(8): e0256445. https://doi.org/10.1371/journal.pone.0256445 PMID: 34424942

16. Ling Z, Liu X, Chen W, Luo Y, Yuan L, Xia Y, et al. The restoration of the vaginal microbiota after treatment for bacterial vaginosis with metronidazole or probiotics. Microbial ecol. 2013; 65(3):773–80. https://doi.org/10.1007/s00248-012-0154-3 PMID: 23250116

17. Janulaitiene M, Paliulyte V, Grinceviciene S, Zakareviciene J, Vladisauskiene A, Marcinkute A, et al. Prevalence and distribution of *Gardnerella vaginalis* subgroups in women with and without bacterial vaginosis. BMC Infect Dis. 2017; 17(1):394. https://doi.org/10.1186/s12879-017-2501-y PMID: 28583109

18. Piot P, Van Dyck E, Peeters M, Hale J, Totten PA, Holmes KK. Biotypes of *Gardnerella vaginalis*. J Clin Microbiol. 1984; 20(4):677–9. https://doi.org/10.1128/jcm.20.4.677-679.1984 PMID: 6333436

19. Ingianni A, Petruzzelli S, Morandotti G, Pompei R. Genotypic differentiation of *Gardnerella vaginalis* by amplified ribosomal DNA restriction analysis (ARDRA). FEMS Immunol Med Microbiol. 1997; 18 (1):61–6. https://doi.org/10.1111/j.1574-695X.1997.tb01028.x PMID: 9215588

20. Hill JE, Goh SH, Money DM, Doyle M, Li A, Crosby WL, et al. Characterization of vaginal microflora of healthy, nonpregnant women by chaperonin-60 sequence-based methods. Am J Obstet Gynecol. 2005; 193(3 Pt 1):682–92. https://doi.org/10.1016/j.ajog.2005.02.094 PMID: 16150261

21. Schellenberg JJ, Links MG, Hill JE, Dumonceaux TJ, Kimani J, Jaoko W, et al. Molecular definition of vaginal microbiota in East African commercial sex workers. Appl Environ Microbiol. 2011; 77 (12):4066–74. https://doi.org/10.1128/AEM.02943-10 PMID: 21531840

22. Paramel Jayaprakash T, Schellenberg JJ, Hill JE. Resolution and characterization of distinct cpn60-based subgroups of *Gardnerella vaginalis* in the vaginal microbiota. PloS one. 2012; 7(8):e43009. https://doi.org/10.1371/journal.pone.0043009 PMID: 22900080

23. Ahmed A, Earl J, Retchless A, Hillier SL, Rabe LK, Cherpes TL, et al. Comparative genomic analyses of 17 clinical isolates of *Gardnerella vaginalis* provide evidence of multiple genetically isolated clades consistent with subspeciation into genovars. J Bacteriol. 2012; 194(15):3922–37. https://doi.org/10.1128/JB.00056-12 PMID: 22609915

24. Balashov SV, Mordechai E, Adelson ME, Gygax SE. Identification, quantification and subtyping of *Gardnerella vaginalis* in noncultured clinical vaginal samples by quantitative PCR. J Med Microbiol. 2014; 63(Pt 2):162–75. https://doi.org/10.1099/jmm.0.066407-0 PMID: 24200640

25. Vaneechoutte M, Guschin A, Van Simaey L, Gansemans Y, Van Nieuwerburgh F, Cools P. Emended description of *Gardnerella vaginalis* and description of *Gardnerella leopoldii sp. nov*., *Gardnerella piotii sp. nov*. and *Gardnerella swidsinskii sp. nov*., with delineation of 13 genomic species within the genus *Gardnerella*. Int J Syst Evol Microbiol. 2019; 69(3):679–87. https://doi.org/10.1099/ijsem.0.003200 PMID: 30648938

26. Potter RF, Burnham CD, Dantas G. In Silico Analysis of *Gardnerella* Genomospecies Detected in the Setting of Bacterial Vaginosis. Clin Chem. 2019; 65(11):1375–87. https://doi.org/10.1373/clinchem.2019.305474 PMID: 31462445

27. Swidsinski A, Mendling W, Loening-Baucke V, Swidsinski S, Dorffel Y, Scholze J, et al. An adherent *Gardnerella vaginalis* biofilm persists on the vaginal epithelium after standard therapy with oral metronidazole. Am J Obstet Gynecol. 2008; 198(1):97 e1–6. https://doi.org/10.1016/j.ajog.2007.06.039 PMID: 18005928

28. Janulaitiene M, Gegzna V, Baranauskiene L, Bulavaitė A, Simanavicius M, Pleckaityte M. Phenotypic characterization of *Gardnerella vaginalis* subgroups suggests differences in their virulence potential. PloS one. 2018; 13(7):e0200625. https://doi.org/10.1371/journal.pone.0200625 PMID: 30001418

29. Cauci S, Culhane JF, Di Santolo M, McCollum K. Among pregnant women with bacterial vaginosis, the hydrolytic enzymes sialidase and prolidase are positively associated with interleukin-1beta. Am J Obstet Gynecol. 2008; 198(1):132 e1–7. S0002-9378(07)00674-6 [pii]. https://doi.org/10.1016/j.ajog.2007.05.035 PMID: 17714681

30. Cauci S, McGregor J, Thorsen P, Grove J, Guaschino S. Combination of vaginal pH with vaginal sialidase and prolidase activities for prediction of low birth weight and preterm birth. Am J Obstet Gynecol. 2005; 192(2):489–96. https://doi.org/10.1016/j.ajog.2004.07.023 PMID: 15695992

31. Pleckaityte M, Janulaitiene M, Lasickiene R, Zvirbliene A. Genetic and biochemical diversity of *Gardnerella vaginalis* strains isolated from women with bacterial vaginosis. FEMS Immunol Med Microbiol. 2012; 65(1):69–77. https://doi.org/10.1111/j.1574-695X.2012.00940.x PMID: 22309200

32. Gelber SE, Aguilar JL, Lewis KL, Ratner AJ. Functional and phylogenetic characterization of Vaginolysin, the human-specific cytolysin from *Gardnerella vaginalis*. J Bacteriol. 2008; 190(11):3896–903. JB.01965-07 [pii] https://doi.org/10.1128/JB.01965-07 PMID: 18390664

33. Garcia EM, Serrano MG, Edupuganti L, Edwards DJ, Buck GA, Jefferson KK. Sequence Comparison of Vaginolysin from Different *Gardnerella* Species. Pathogens (Basel, Switzerland). 2021; 10(2). https://doi.org/10.3390/pathogens10020086 PMID: 33498226

34. Santiago GL, Deschaght P, El Aila N, Kiama TN, Verstraelen H, Jefferson KK, et al. *Gardnerella vaginalis* comprises three distinct genotypes of which only two produce sialidase. Am J Obstet Gynecol. 2011; 204(5):450.e1–7. https://doi.org/10.1016/j.ajog.2010.12.061 PMID: 21444061

35. Machado D, Castro J, Palmeira-de-Oliveira A, Martinez-de-Oliveira J, Cerca N. Bacterial Vaginosis Biofilms: Challenges to Current Therapies and Emerging Solutions. Front Microbiol. 2015; 6:1528. https://doi.org/10.3389/fmicb.2015.01528 PMID: 26834706

36. Castro J, Alves P, Sousa C, Cereija T, França Â, Jefferson KK, et al. Using an in-vitro biofilm model to assess the virulence potential of bacterial vaginosis or non-bacterial vaginosis *Gardnerella vaginalis* isolates. Sci Reports. 2015; 5:11640. https://doi.org/10.1038/srep11640 PMID: 26113465

37. Schellenberg JJ, Paramel Jayaprakash T, Withana Gamage N, Patterson MH, Vaneechoutte M, Hill JE. *Gardnerella vaginalis* Subgroups Defined by cpn60 Sequencing and Sialidase Activity in Isolates from Canada, Belgium and Kenya. PloS one. 2016; 11(1):e0146510. https://doi.org/10.1371/journal.pone.0146510 PMID: 26751374

38. Shipitsyna E, Krysanova A, Khayrullina G, Shalepo K, Savicheva A, Guschin A, et al. Quantitation of all Four *Gardnerella vaginalis* Clades Detects Abnormal Vaginal Microbiota Characteristic of Bacterial Vaginosis More Accurately than Putative G. vaginalis Sialidase A Gene Count. Mol Diagn Ther. 2019; 23(1):139–47. https://doi.org/10.1007/s40291-019-00382-5 PMID: 30721449

39. Albert AY, Chaban B, Wagner EC, Schellenberg JJ, Links MG, van Schalkwyk J, et al. A Study of the Vaginal Microbiome in Healthy Canadian Women Utilizing cpn60-Based Molecular Profiling Reveals

Distinct *Gardnerella* Subgroup Community State Types. PloS one. 2015; 10(8):e0135620. https://doi.org/10.1371/journal.pone.0135620 PMID: 26266808

40. Wells JS, Chandler R, Dunn A, Brewster G. The Vaginal Microbiome in U.S. Black Women: A Systematic Review. J Women's Health (2002). 2020; 29(3):362–75. https://doi.org/10.1089/jwh.2019.7717 PMID: 32109181

41. Srinivasan S, Hoffman NG, Morgan MT, Matsen FA, 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(6):e37818. https://doi.org/10.1371/journal.pone.0037818 PMID: 22719852

42. Lambert JA, John S, Sobel JD, Akins RA. Longitudinal analysis of vaginal microbiome dynamics in women with recurrent bacterial vaginosis: recognition of the conversion process. PloS one. 2013; 8 (12):e82599. https://doi.org/10.1371/journal.pone.0082599 PMID: 24376552

43. Muzny CA, Blanchard E, Taylor CM, Aaron KJ, Talluri R, Griswold ME, et al. Identification of Key Bacteria Involved in the Induction of Incident Bacterial Vaginosis: A Prospective Study. J Infect Dis. 2018; 218(6):966–78. https://doi.org/10.1093/infdis/jiy243 PMID: 29718358

44. Brotman RM, Ravel J, Cone RA, Zenilman JM. Rapid fluctuation of the vaginal microbiota measured by Gram stain analysis. Sex Transm Infect. 2010; 86(4):297–302. https://doi.org/10.1136/sti.2009.040592 PMID: 20660593

45. Fredricks DN. Molecular methods to describe the spectrum and dynamics of the vaginal microbiota. Anaerobe. 2011; 17(4):191–5. https://doi.org/10.1016/j.anaerobe.2011.01.001 PMID: 21376827

46. Tettamanti Boshier FA, Srinivasan S, Lopez A, Hoffman NG, Proll S, Fredricks DN, et al. Complementing 16S rRNA Gene Amplicon Sequencing with Total Bacterial Load To Infer Absolute Species Concentrations in the Vaginal Microbiome. mSystems. 2020; 5(2). https://doi.org/10.1128/mSystems.00777-19 PMID: 32265316

47. Lee S, Oh KY, Hong H, Jin CH, Shim E, Kim SH, et al. Community State Types of Vaginal Microbiota and Four Types of Abnormal Vaginal Microbiota in Pregnant Korean Women. Front Public Health. 2020; 8:507024. https://doi.org/10.3389/fpubh.2020.507024 PMID: 33194936

48. Ling Z, Kong J, Liu F, Zhu H, Chen X, Wang Y, et al. Molecular analysis of the diversity of vaginal microbiota associated with bacterial vaginosis. BMC Genomics. 2010; 11:488. https://doi.org/10.1186/1471-2164-11-488 PMID: 20819230

49. Ravel J, Brotman RM, Gajer P, Ma B, Nandy M, Fadrosh DW, et al. Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. Microbiome. 2013; 1(1):29–. https://doi.org/10.1186/2049-2618-1-29 PMID: 24451163

50. Fredricks DN, Fiedler TL, Marrazzo JM. Molecular identification of bacteria associated with bacterial vaginosis. N Engl J Med. 2005; 353(18):1899–911. https://doi.org/10.1056/NEJMoa043802 PMID: 16267321

51. Hilbert DW, Smith WL, Chadwick SG, Toner G, Mordechai E, Adelson ME, et al. Development and Validation of a Highly Accurate Quantitative Real-Time PCR Assay for Diagnosis of Bacterial Vaginosis. J Clin Microbiol. 2016; 54(4):1017–24. https://doi.org/10.1128/JCM.03104-15 PMID: 26818677

52. Hilbert DW, Smith WL, Paulish-Miller TE, Chadwick SG, Toner G, Mordechai E, et al. Utilization of molecular methods to identify prognostic markers for recurrent bacterial vaginosis. Diagn Microbiol Infect Dis. 2016; 86(2):231–42. https://doi.org/10.1016/j.diagmicrobio.2016.07.003 PMID: 27431434

53. Marrazzo JM, Thomas KK, Fiedler TL, Ringwood K, Fredricks DN. Relationship of specific vaginal bacteria and bacterial vaginosis treatment failure in women who have sex with women. Ann Intern Med. 2008; 149(1):20–8. 149/1/20 [pii]. https://doi.org/10.7326/0003-4819-149-1-200807010-00006 PMID: 18591634

54. Wang B, Xiao BB, Shang CG, Wang K, Na RS, Nu XX, et al. Molecular analysis of the relationship between specific vaginal bacteria and bacterial vaginosis metronidazole therapy failure. Eur J Clin Microbiol Infect Dis. 2014; 33(10):1749–56. https://doi.org/10.1007/s10096-014-2128-5 PMID: 24816815

55. Xiao B, Wu C, Song W, Niu X, Qin N, Liu Z, et al. Association Analysis on Recurrence of Bacterial Vaginosis Revealed Microbes and Clinical Variables Important for Treatment Outcome. Front Cell Infect Microbiol. 2019; 9:189. https://doi.org/10.3389/fcimb.2019.00189 PMID: 31245300

56. Hay P. Recurrent bacterial vaginosis. Curr Op Infect Dis. 2009; 22(1):82–6. https://doi.org/10.1097/QCO.0b013e32832180c6 PMID: 19532085

57. Bradshaw CS, Morton AN, Hocking J, Garland SM, Morris MB, Moss LM, et al. High recurrence rates of bacterial vaginosis over the course of 12 months after oral metronidazole therapy and factors associated with recurrence. J Infect Dis. 2006; 193(11):1478–86. https://doi.org/10.1086/503780 PMID: 16652274

58. Bradshaw CS, Sobel JD. Current Treatment of Bacterial Vaginosis-Limitations and Need for Innovation. J Infect Dis. 2016; 214 Suppl 1:S14–20. https://doi.org/10.1093/infdis/jiw159 PMID: 27449869

59. Eschenbach DA. Bacterial vaginosis: resistance, recurrence, and/or reinfection? Clin Infect Dis. 2007; 44(2):220–1. https://doi.org/10.1086/509584 PMID: 17173220

60. Bradshaw CS, Vodstrcil LA, Hocking JS, Law M, Pirotta M, Garland SM, et al. Recurrence of bacterial vaginosis is significantly associated with posttreatment sexual activities and hormonal contraceptive use. Clin Infect Dis. 2013; 56(6):777–86. https://doi.org/10.1093/cid/cis1030 PMID: 23243173

61. Ratten LK, Plummer EL, Murray GL, Danielewski J, Fairley CK, Garland SM, et al. Sex is associated with the persistence of non-optimal vaginal microbiota following treatment for bacterial vaginosis: a prospective cohort study. BJOG. 2021; 128(4):756–67. https://doi.org/10.1111/1471-0528.16430 PMID: 33480468

62. Plummer EL, Vodstrcil LA, Doyle M, Danielewski JA, Murray GL, Fehler G, et al. A Prospective, Open-Label Pilot Study of Concurrent Male Partner Treatment for Bacterial Vaginosis. mBio. 2021; 12(5): e0232321. https://doi.org/10.1128/mBio.02323-21 PMID: 34663095

63. Schuyler JA, Mordechai E, Adelson ME, Sobel JD, Gygax SE, Hilbert DW. Identification of intrinsically metronidazole-resistant clades of *Gardnerella vaginalis*. Diagn Microbiol Infect Dis. 2016; 84(1):1–3. https://doi.org/10.1016/j.diagmicrobio.2015.10.006 PMID: 26514076

64. Ferris MJ, Masztal A, Aldridge KE, Fortenberry JD, Fidel PL Jr., Martin DH. Association of *Atopobium vaginae*, a recently described metronidazole resistant anaerobe, with bacterial vaginosis. BMC Infect Dis. 2004; 4:5. https://doi.org/10.1186/1471-2334-4-5 PMID: 15018635

65. Deng ZL, Gottschick C, Bhuju S, Masur C, Abels C, Wagner-Dobler I. Metatranscriptome Analysis of the Vaginal Microbiota Reveals Potential Mechanisms for Protection against Metronidazole in Bacterial Vaginosis. mSphere. 2018; 3(3). https://doi.org/10.1128/mSphereDirect.00262-18 PMID: 29875146

66. Gal M, Brazier JS. Metronidazole resistance in Bacteroides spp. carrying nim genes and the selection of slow-growing metronidazole-resistant mutants. J Antimicrob Chemother. 2004; 54(1):109–16. https://doi.org/10.1093/jac/dkh296 PMID: 15190033

67. Lee CY, Cheu RK, Lemke MM, Gustin AT, France MT, Hampel B, et al. Quantitative modeling predicts mechanistic links between pre-treatment microbiome composition and metronidazole efficacy in bacterial vaginosis. Nature commun. 2020; 11(1):6147. https://doi.org/10.1038/s41467-020-19880-w PMID: 33262350

68. Li T, Zhang Z, Wang F, He Y, Zong X, Bai H, et al. Antimicrobial Susceptibility Testing of Metronidazole and Clindamycin against *Gardnerella vaginalis* in Planktonic and Biofilm Formation. Canadian J Infect Dis Med Microbiol. 2020; 2020:1361825. https://doi.org/10.1155/2020/1361825 PMID: 32612729

69. Alves P, Castro J, Sousa C, Cereija TB, Cerca N. *Gardnerella vaginalis* outcompetes 29 other bacterial species isolated from patients with bacterial vaginosis, using in an in vitro biofilm formation model. J Infect Dis. 2014; 210(4):593–6. https://doi.org/10.1093/infdis/jiu131 PMID: 24596283

70. Sobel JD, Kaur N, Woznicki NA, Boikov D, Aguin T, Gill G, et al. Prognostic indicators of recurrence of bacterial vaginosis. J Clin Microbiol. 2019. https://doi.org/10.1128/JCM.00227-19 PMID: 30842235

71. Sobel JD, Kaur N, Woznicki NA, Boikov D, Aguin T, Gill G, et al. Conventional oral and secondary high dose vaginal metronidazole therapy for recurrent bacterial vaginosis: clinical outcomes, impacts of sex and menses. Infect Drug Resist. 2019; 12:2297–307. https://doi.org/10.2147/IDR.S213853 PMID: 31413606

72. Amsel R, Totten PA, Spiegel CA, Chen KC, Eschenbach D, Holmes KK. Nonspecific vaginitis. Diagnostic criteria and microbial and epidemiologic associations. Am J Med. 1983; 74(1):14–22. https://doi.org/10.1016/0002-9343(83)91112-9 PMID: 6600371

73. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol. 2013; 79(17):5112–20. https://doi.org/10.1128/AEM.01043-13 PMID: 23793624

74. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol. 2009; 75(23):7537–41. https://doi.org/10.1128/AEM.01541-09 PMID: 19801464

75. Schloss PD. Reintroducing mothur: 10 Years Later. Appl Environ Microbiol. 2020; 86(2). https://doi.org/10.1128/AEM.02343-19 PMID: 31704678

76. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nature methods. 2016; 13(7):581–3. https://doi.org/10.1038/nmeth.3869 PMID: 27214047

77. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2012; 41(D1): D590–D6. https://doi.org/10.1093/nar/gks1219 PMID: 23193283

78. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. Nucleic Acids Res. 2013; 42(D1):D643–D8. https://doi.org/10.1093/nar/gkt1209 PMID: 24293649

79. McLaren MR, Callahan BJ. Silva 138.1 prokaryotic SSU taxonomic training data formatted for DADA2 [Data set]. 2021. Available from: https://doi.org/10.5281/zenodo.4587955.

80. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 2014; 42(Database issue):D633–42. https://doi.org/10.1093/nar/gkt1244 PMID: 24288368

81. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol. 2018; 35(6):1547–9. https://doi.org/10.1093/molbev/msy096 PMID: 29722887

82. Stecher G, Tamura K, Kumar S. Molecular Evolutionary Genetics Analysis (MEGA) for macOS. Mol Biol Evol. 2020; 37(4):1237–9. https://doi.org/10.1093/molbev/msz312 PMID: 31904846

83. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol. 2007; 73(16):5261–7. https://doi.org/10.1128/AEM.00062-07 PMID: 17586664

84. Dhariwal A, Chong J, Habib S, King IL, Agellon LB, Xia J. MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. Nucleic Acids Res. 2017; 45 (W1):W180–w8. https://doi.org/10.1093/nar/gkx295 PMID: 28449106

85. Chong J, Liu P, Zhou G, Xia J. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. Nat Protoc. 2020; 15(3):799–821. https://doi.org/10.1038/s41596-019-0264-1 PMID: 31942082

86. Xia J, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. Metabolomics. 2013; 9(2):280–99. https://doi.org/10.1007/s11306-012-0482-9 PMID: 23543913

87. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Res. 2021; 49(W1):W388–w96. https://doi.org/10.1093/nar/gkab382 PMID: 34019663

88. Ltd. MS. Diagnostic test evaluation calculator. (Version 20.106; accessed April 5, 2022) Available from: https://www.medcalc.org/calc/diagnostic_test.php.

89. Gustin AT, Thurman AR, Chandra N, Schifanella L, Alcaide M, Fichorova R, et al. Recurrent bacterial vaginosis following metronidazole treatment is associated with microbiota richness at diagnosis. Am J Obstet Gynecol. 2022; 226(2):225.e1–.e15. https://doi.org/10.1016/j.ajog.2021.09.018 PMID: 34560047

90. Srinivasan S, Beamer MA, Fiedler TL, Austin MN, Sizova MV, Strenk SM, et al. *Megasphaera lornae sp. nov.*, *Megasphaera hutchinsoni sp. nov.*, and *Megasphaera vaginalis* sp. nov.: novel bacteria isolated from the female genital tract. Int J Syst Evol Microbiol. 2019; 71(3). https://doi.org/10.1099/ijsem.0.004702 PMID: 33616513

91. Shetty SA, Marathe NP, Lanjekar V, Ranade D, Shouche YS. Comparative Genome Analysis of *Megasphaera sp*. Reveals Niche Specialization and Its Potential Role in the Human Gut. PloS one. 2013; 8(11):e79353. https://doi.org/10.1371/journal.pone.0079353 PMID: 24260205

92. Hashizume K, Tsukahara T, Yamada K, Koyama H, Ushida K. *Megasphaera elsdenii* JCM1772T normalizes hyperlactate production in the large intestine of fructooligosaccharide-fed rats by stimulating butyrate production. J Nutr. 2003; 133(10):3187–90. https://doi.org/10.1093/jn/133.10.3187 PMID: 14519808

93. Salliss ME, Maarsingh JD, Garza C, Łaniewski P, Herbst-Kralovetz MM. Veillonellaceae family members uniquely alter the cervical metabolic microenvironment in a human three-dimensional epithelial model. NPJ Biofilms Microbiomes. 2021; 7(1):57. https://doi.org/10.1038/s41522-021-00229-0 PMID: 34230496

94. Garg KB, Ganguli I, Kriplani A, Lohiya NK, Thulkar J, Talwar GP. Metabolic properties of lactobacilli in women experiencing recurring episodes of bacterial vaginosis with vaginal pH >or = 5. Eur J Clin Microbiol Infect Dis. 2010; 29(1):123–5. https://doi.org/10.1007/s10096-009-0818-1 PMID: 19802749

95. Tachedjian G, Aldunate M, Bradshaw CS, Cone RA. The role of lactic acid production by probiotic *Lactobacillus* species in vaginal health. Res Microbiol. 2017; 168(9–10):782–92. https://doi.org/10.1016/j.resmic.2017.04.001 PMID: 28435139

96. Teixeira GS, Carvalho FP, Arantes RME, Nunes AC, Moreira JLS, Mendonça M, et al. Characteristics of *Lactobacillus* and *Gardnerella vaginalis* from women with or without bacterial vaginosis and their relationships in gnotobiotic mice. J Med Microbiol. 2012; 61(Pt 8):1074–81. https://doi.org/10.1099/jmm.0.041962-0 PMID: 22539000

97. O'Hanlon DE, Moench TR, Cone RA. In vaginal fluid, bacteria associated with bacterial vaginosis can be suppressed with lactic acid but not hydrogen peroxide. BMC Infect Dis. 2011; 11:200. https://doi.org/10.1186/1471-2334-11-200 PMID: 21771337

98. Swidsinski A, Mendling W, Loening-Baucke V, Ladhoff A, Swidsinski S, Hale LP, et al. Adherent biofilms in bacterial vaginosis. Obstet Gynecol. 2005; 106(5 Pt 1):1013–23. https://doi.org/10.1097/01.AOG.0000183594.45524.d2 PMID: 16260520

99. Gottschick C, Deng Z-L, Vital M, Masur C, Abels C, Pieper DH, et al. Treatment of biofilms in bacterial vaginosis by an amphoteric tenside pessary-clinical study and microbiota analysis. Microbiome. 2017; 5(1):119. https://doi.org/10.1186/s40168-017-0326-y PMID: 28903767

100. Ruiz-Perez D, Coudray MS, Colbert B, Krupp K, Kumari H, Stebliankin V, et al. Effect of metronidazole on vaginal microbiota associated with asymptomatic bacterial vaginosis. Access Microbiol. 2021; 3 (5):000226–. https://doi.org/10.1099/acmi.0.000226 PMID: 34151180

101. Mitchell CM, Hitti JE, Agnew KJ, Fredricks DN. Comparison of oral and vaginal metronidazole for treatment of bacterial vaginosis in pregnancy: impact on fastidious bacteria. BMC Infect Dis. 2009; 9:89–. https://doi.org/10.1186/1471-2334-9-89 PMID: 19515236

102. Petrina MAB, Cosentino LA, Rabe LK, Hillier SL. Susceptibility of bacterial vaginosis (BV)-associated bacteria to secnidazole compared to metronidazole, tinidazole and clindamycin. Anaerobe. 2017; 47:115–9. https://doi.org/10.1016/j.anaerobe.2017.05.005 PMID: 28522362

103. Numanovic N, Ribis S, Cukic J, Nenadic D, Zivanovic A, Sazdanovic P, et al. Quantification of *Gardnerella vaginalis*, *Atopobium vaginae* and *Lactobacillus spp.* in bacterial vaginosis. J Infect Dev Ctries. 2021; 15(9):1293–8. https://doi.org/10.3855/jidc.13091 PMID: 34669599

104. Deng T, Shang A, Zheng Y, Zhang L, Sun H, Wang W. Log (*Lactobacillus crispatus*/ *Gardnerella vaginalis*): a new indicator of diagnosing bacterial vaginosis. Bioengineered. 2022; 13(2):2981–91. https://doi.org/10.1080/21655979.2022.2027059 PMID: 35038957