Metatranscriptome Analysis of the Vaginal Microbiota Reveals Potential Mechanisms for Protection against Metronidazole in Bacterial Vaginosis

Zhi-Luo Deng,a Cornelia Gottschick,a* Sabin Bhuju,b Clarissa Masur,c Christoph Abels,c Irene Wagner-Döblera

aResearch Group Microbial Communication, Department of Medical Microbiology, Helmholtz Centre for Infection Research, Braunschweig, Germany
bGenome Analytics, Helmholtz Centre for Infection Research, Braunschweig, Germany
cDr. August Wolff GmbH & Co. KG Arzneimittel, Bielefeld, Germany

ABSTRACT Bacterial vaginosis (BV) is a prevalent multifactorial disease of women in their reproductive years characterized by a shift from the *Lactobacillus* species-dominated microbial community toward a taxonomically diverse anaerobic community. For unknown reasons, some women do not respond to therapy. In our recent clinical study, among 37 women diagnosed with BV, 31 were successfully treated with metronidazole, while 6 still had BV after treatment. To discover possible reasons for the lack of response in those patients, we performed a metatranscriptome analysis of their vaginal microbiota, comparing them to the patients who responded. Seven of 8 clustered regularly interspaced short palindromic repeat (CRISPR)-associated (Cas) genes of *Gardnerella vaginalis* were highly upregulated in nonresponding patients. Cas genes, in addition to protecting against phages, might be involved in DNA repair, thus mitigating the bactericidal effect of DNA-damaging agents such as metronidazole. In the second part of our study, we analyzed the vaginal metatranscriptomes of four patients over 3 months and showed high *in vivo* expression of genes for pore-forming toxins in *L. iners* and of genes encoding enzymes for the production of hydrogen peroxide and D-lactate in *L. crispatus*.

IMPORTANCE Bacterial vaginosis is a serious issue for women in their reproductive years. Although it can usually be cured by antibiotics, the recurrence rate is very high, and some women do not respond to antibiotic therapy. The reasons for that are not known. Therefore, we undertook a study to detect the activity of the complete microbiota in the vaginal fluid of women who responded to antibiotic therapy and compared it to the activity of the microbiota in women who did not respond. We found that one of the most important pathogens in bacterial vaginosis, *Gardnerella vaginalis*, has activated genes that can repair the DNA damage caused by the antibiotic in those women that do not respond to therapy. Suppressing these genes might be a possibility to improve the antibiotic therapy of bacterial vaginosis.

KEYWORDS bacterial vaginosis, antibiotic resistance, metatranscriptome, vaginal microbiota

The healthy vaginal microbiome is characterized by low pH and low diversity and can be categorized into community state types (CSTs) that are dominated by different *Lactobacillus* spp. such as *L. crispatus*, *L. iners*, *L. gasseri*, and, less frequently, *L. jensenii* or a more diverse community (1). Bacterial vaginosis (BV) is a frequent multifactorial disease of women in their reproductive years that is characterized by a shift of the *Lactobacillus* species-dominated bacterial community to a community of various, mostly anaerobic bacteria (2). BV is associated with a higher risk of preterm birth and of acquiring sexually transmitted infections such as HIV (3). The most common bacteria

found in BV, identified by 16S rRNA gene sequencing, are *Gardnerella*, *Atopobium*, *Prevotella*, *Bacteroides*, *Peptostreptococcus*, *Mobiluncus*, *Sneathia*, *Leptotrichia*, *Mycoplasma*, and BV-associated bacterium 1 (BVAB1) to BVAB3 of the order Clostridiales. Recently, three CSTs dominated by *Gardnerella vaginalis*, *Lachnospiraceae*, and *Sneathia sanguinegens*, respectively, have been described (4). In our recent clinical study, *S. amnii* was identified as the best biomarker for BV (5).

The most important pathogen in BV is *Gardnerella vaginalis* (6). It is currently the only described species in the genus *Gardnerella*, but genome comparisons suggest that it can be separated into four genetically isolated subspecies (7). While they cannot be resolved by 16S rRNA gene sequencing, the universal target from the chaperonin-60 (cp-60) gene separates the species into the same four subgroups (group A, clade 4; subgroup B, clade 2; subgroup C, clade 1; subgroup D, clade 3) (6, 8, 9). All four subgroups of *G. vaginalis* can be detected in the vaginal microbiota of healthy women throughout the menstrual cycle (10). Subgroups A and C define distinct CSTs in health (11). Isolates from the four subgroups of *G. vaginalis* differ in their virulence as well as in their resistance against metronidazole. The sialidase activity of *G. vaginalis* is an important virulence factor, and it was detected in all isolates from subgroup B and in a few isolates from subgroup C but not in isolates from subgroups A and D (12). The presence of sialidase activity is used for diagnosis of BV in a commercial kit (13). Resistance against metronidazole was found in isolates from subgroups A and D, while those from subgroups B and C were highly susceptible (14).

Metronidazole is a widely applied chemotherapeutic agent used to treat infectious diseases caused by anaerobic bacteria, and it is the first-line antibiotic for treating BV (15, 16). Metronidazole is a prodrug which requires enzymatic reduction within the cell, which occurs under anaerobic conditions only, to transform it into an active form (17). Activated metronidazole acts by covalently binding to DNA, disrupting its helical structure and causing single- and double-strand breaks that lead to DNA degradation and death of the pathogens (17). Resistance can therefore be mediated by lack of activation of the prodrug, or by repair of DNA damage, and has been studied in various pathogens. In *Helicobacter pylori* and *Campylobacter* spp., ferredoxin, ferredoxin/ ferredoxin-NADP reductase (FNR), and nitroreductase contribute to metronidazole resistance (17). In *Bacteroides fragilis*, genes responsible for DNA repair such as *recA* and the *recA*-mediated autopeptidase (Rma) gene and a gene named the nitroimidazole resistance (*nim*) gene encoding a nitroimidazole reductase were shown to confer resistance against metronidazole (18, 19). Failure of BV treatment by metronidazole is relatively rare (5, 20). It is unclear if it is caused by resistance of the BV pathogens to metronidazole and which mechanisms are acting *in vivo*. A recent study has demonstrated that failure of treatment of BV with metronidazole is not associated with higher loads of *G. vaginalis* and *Atopobium vaginae* (21). Isolates from *G. vaginalis* subgroups A and D are intrinsically resistant against metronidazole, but the underlying mechanism is unknown (14).

Until now, the majority of studies regarding the vaginal microbiota have focused on 16S rRNA gene sequencing, answering only questions on the taxonomic composition of bacterial communities and not on their functions (2). A metatranscriptome analysis comparing vaginal swabs from two women with BV to vaginal swabs from two healthy subjects showed that *L. iners* upregulates transcription of the cholesterol-dependent cytolysin (CDC) and of genes belonging to the clustered regularly interspaced short palindromic repeat (CRISPR) system in BV (22). No study has investigated the activity shifts of the vaginal microbiota during antibiotic treatment of BV.

We had previously analyzed the vaginal microbiota in the context of a clinical trial using 16S rRNA gene sequencing (5). Of 37 patients diagnosed with BV and included in this study, 31 were initially cured by a single oral dose of metronidazole. Six patients did not respond; i.e., they were still diagnosed with BV according to the Nugent score after antibiotic therapy. Here we asked if differences in the activity of the microbiota might be responsible for the lack of response in those six patients. We therefore analyzed their metatranscriptomes at the time of diagnosis of BV (visit 1) and after treatment with

metronidazole (visit 2) and compared them to those of 8 patients that responded to treatment according to the Nugent score.

The high rate of recurrence is another crucial problem for BV treatment. The 1-year recurrence rate of BV ranges from 40% to 80% after therapy with metronidazole (23) or clindamycin cream (24). CSTs dominated by *L. iners* might have an increased probability to shift to a dysbiotic state (22, 25, 26). In the second part of our study, we therefore followed the activity of the microbiota of four of the patients that initially responded to metronidazole treatment over a period of 3 months (visits 3 to 5) and analyzed gene expression of *L. crispatus* and *L. iners in vivo*.

We show the importance of *G. vaginalis* for BV, which can be massively underestimated using 16S rRNA gene sequencing. The relative abundances of the four subgroups of *G. vaginalis* could be determined in responders and nonresponders. Transcripts potentially leading to a lack of response to metronidazole treatment were identified. CRISPR-Cas genes are suggested to be a novel mechanism of *G. vaginalis* to mitigate the DNA-damaging effect of metronidazole. *L. iners* highly expressed genes for pore-forming toxins *in vivo*, and the transcripts that were most highly expressed in *L. crispatus in vivo* encoded enzymes for D-lactate and hydrogen peroxide production.

RESULTS

Study population and overview of sequencing results. We studied the vaginal microbiome of 14 patients before, during, and after metronidazole treatment of BV using metatranscriptome sequencing (Fig. 1A). Patients were part of a clinical trial described elsewhere (5). In the first part of the study, we analyzed samples from two time points (diagnosis of BV, visit 1) and after metronidazole treatment (visit 2). Eight patients responded to treatment, and six patients did not respond to treatment with the antibiotic and thus were still BV positive according to the Nugent score at visit 2. In the second part of our study, three additional time points were analyzed for four of the patients that initially responded to metronidazole therapy, covering a total period of 3 months. Those four patients belonged to the lactic acid arm of the clinical study. Two of them experienced recurrence, while the other two were stably non-BV after treatment according to the Nugent score (see details in Data Set S1, sheet 1, in the supplemental material). In total, we analyzed 40 vaginal fluid samples, totaling 22 with BV status and 18 without. Metatranscriptome sequencing resulted in a total of 1,879,945,342 reads. Of these, 1,377,516,082 reads (73%) were left after quality filtering and removal of rRNA (Data Set S1, sheet 2). On average, 34 million reads were analyzed per sample.

Construction of the reference genome and gene databases for taxonomic and activity profiling. Human reads comprised ~11% (BV) versus ~56% (non-BV) of the total putative mRNA reads based on the standard Kraken (27) database (Data Set S1, sheet 2). This suggests that the bacterial load is much lower in non-BV than in BV since the human contamination is much higher in non-BV. Using the standard Kraken database, only 41% of total putative microbial (nonhuman) mRNA reads could be assigned taxonomically (Data Set S1, sheet 2). To improve the fraction of taxonomically assignable reads, we then constructed a refined database (ref_Genome) which combined the urogenital subset of the Human Microbiome Project (HMP) (28) database (147 genomes) and all species which are not included in the urogenital subset of the HMP database but are detected by the standard Kraken database with an abundance of >1% (7 genomes). We also added *S. amnii* and *S. sanguinegens*, which had previously been shown to be highly abundant based on 16S rRNA gene sequencing (5) but were not contained in either the HMP database or the standard Kraken reference database. There are four *G. vaginalis* strains in the HMP database, of which one belongs to subgroup A and three belong to subgroup C. Given the importance and high intraspecies diversity of *G. vaginalis*, we added 5 additional *G. vaginalis* genomes based on the genome tree reported in the NCBI database and the completeness of the genome assembly; those five strains cover all four subgroups. We added the genomes of *Gardnerella* sp. strain 26-12 and *Gardnerella* sp. strain 30-4, which were isolated from
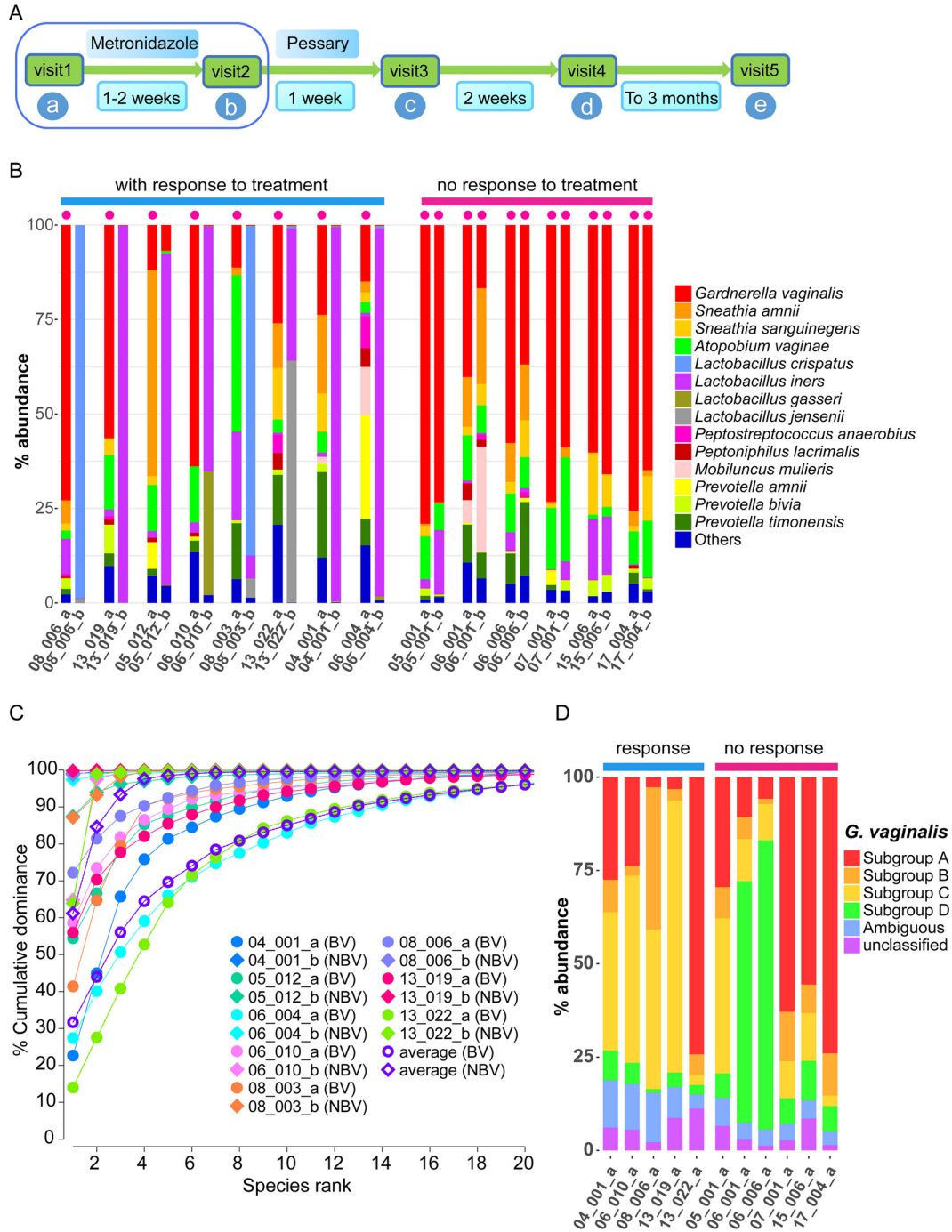
FIG 1 Study design and taxonomic composition of vaginal fluid metatranscriptomes in BV during and after treatment with metronidazole. (A) Time course of the clinical study. (B) Taxonomic composition of the metatranscriptome at visit 1 (diagnosis) and visit 2 (after metronidazole therapy). (C) Cumulative dominance of the vaginal microbiota in BV and non-BV. (D) The subspecies composition of the *G. vaginalis* subcommunity. Species with an average relative abundance lower than 0.5% were grouped into "Others." A red dot at the top of a sample column in panel B indicates BV. The digits indicate the patient ID, while the letters a and b denote visit 1 and visit 2. After the first sampling at visit 1, the patients were treated with metronidazole. Total putative bacterial mRNA reads were mapped to the ref_Genome database using Kraken (see Materials and Methods for details). BV status was determined by Nugent score. The "BV" and "NBV" in parentheses in panel C indicate BV and non-BV, respectively.

the bladder recently (29). They were classified into *G. vaginalis* subgroup A based on sequence homology (29). In total, the database contained 163 bacterial genomes from 105 species (Data Set S1, sheet 3). Using the database, the rate of taxonomically classified putative microbial mRNA reads could be improved to 86% on average (Data Set S1, sheet 2).

For functional assignment, we constructed a reference gene database (ref_Gene) (Data Set S1, sheet 4). It was based on the same genomes as the ref_Genome database, except that the seven additional *Gardnerella* species genomes were not included because of the low quality of the annotation of coding sequences. The ref_Gene database contained 301,323 genes. To investigate the activity shifts of the communities, we mapped the cleaned metatranscriptomic reads to the ref_Gene database using Burrows-Wheeler Aligner (BWA). In total, 78% of total putative microbial mRNA reads could be mapped to the ref_Gene database. Per sample, an average of 8.9 million microbial mRNA reads could be mapped with a mapping quality (MAPQ) value of >10 (Data Set S1, sheet 2).

Shifts in the taxonomic composition of the active community following metronidazole treatment. The taxonomic composition of transcripts was determined using Kraken and the ref_Genome database. Figure 1B shows that in all communities with BV status, the most abundant species were *G. vaginalis*, *A. vaginae*, *S. amnii*, and *Prevotella timonensis*. In the posttreatment communities, the metatranscriptomes from responders (non-BV; Nugent score of <6) were dominated by *L. crispatus*, *L. iners*, and *L. jensenii*, representing typical CSTs of the healthy female microbiota.

On average, fewer than 14 species contributed >90% of the mapped reads in BV and 3 species accounted for >90% of the mapped reads in non-BV communities (Fig. 1C). The individual dominance plots showed the same pattern, where 10 species contributed >90% of the metatranscriptomes for most BV patients. In non-BV subjects, this number was 2 for most patients and the dominance curves were extremely steep. For comparison, in the periodontal metatranscriptome, more than 100 species were required to cover 90% of mapped reads (30). These data show that the active microbiota in BV is much less diverse than is suggested by 16S rRNA gene sequencing.

G. vaginalis was the most dominant active species in BV. To estimate the relative abundances of the four subgroups of *G. vaginalis*, we extracted all reads assigned to *G. vaginalis* from the metatranscriptomes and assigned them to the four strains representing subgroups A, B, C, and D (409-05, 00703Bmah, HMP9231, and 00703Bmash, respectively) using Kraken. For this analysis, we used samples from visit 1 where *G. vaginalis* reads comprised at least 20% of all reads, which included all 6 patients without response to treatment and 5 of the 8 patients that responded to treatment. Figure 1D shows that, on average, >95% of the reads could be mapped to the four subgroups and that, on average, only 7% were assigned ambiguously. In those patients that did not respond to treatment, subgroups A and D comprised 68.5% $\pm$ 17.2% of all reads, while they accounted for 30.5% $\pm$ 29.3% of all reads in patients that responded to treatment (Wilcoxon test $P = 0.0520$). We observed that *Gardnerella* spp. previously isolated from the bladder (*Gardnerella* sp. strain 26-12 and *Gardnerella* sp. strain 30-4) (29) contributed on average 6% of all taxonomically assigned reads in BV (see Fig. S1 in the supplemental material).

Comparison of the taxonomic compositions of vaginal fluid samples between metatranscriptome and 16S rRNA gene sequencing. We compared the taxonomic compositions determined using 16S rRNA sequencing previously (5) to the taxonomic composition of the metatranscriptome determined here by Kraken with the ref_Genome database. In non-BV samples, we did not observe large differences for the four most abundant species (Data Set S1, sheet 6), while in BV samples, large differences between the two data sets were found. Figure 2 shows the top 12 most abundant taxa identified using 16S rRNA gene sequencing or metatranscriptomics. Most of the abundant species identified in the mRNA sequencing data set were also identified using 16S rRNA gene amplicon sequencing, although usually at different abundances. For example, *A. vaginae* comprised 13% of all reads based on 16S rRNA gene sequencing but only 11% in the metatranscriptome data set. The most pronounced difference was observed for *G. vaginalis*, which comprised, on average, 47% of the relative abundance in the metatranscriptome and, on average, only 5% in the 16S rRNA sequencing data. Several additional differences were found. Higher-level taxa such as Veillonellaceae (family) or *Parvimonas* (genus) are not listed among the top 12 taxa of
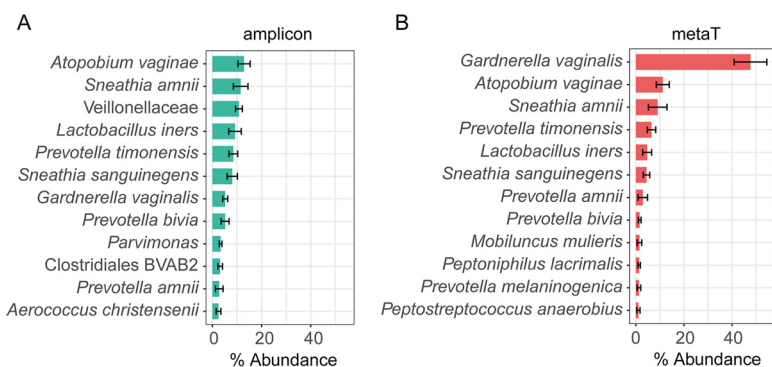
FIG 2 Average taxonomic composition of vaginal fluid samples in BV determined by 16S rRNA amplicon sequencing (A) and metatranscriptome (metaT) sequencing (B). (A) Amplicon sequencing was performed as described for our previous study (5) using primers V1-V2. (B) The taxonomy was assigned based on all cleaned reads after removal of human reads using Kraken and the ref_Genome database. The top 12 most abundant taxa for each approach are shown in panels A and B. Relative average abundance was calculated based on all mapped reads. Means and standard errors are shown.

the metatranscriptome, because mapping occurred to the species level, and the abundance of individual species of these higher-order taxa was too low for them to be found among the top 12 taxa (Data Set S1, sheet 6). BVAB2 is readily detected by PCR and is an important indicator of BV, but it has not yet been cultivated and so there is no genome available to map the reads against.

Global community profiling in non-BV and BV. In order to profile the function of the communities, all cleaned putative mRNA reads (Data Set S1, sheet 2) were mapped using BWA onto the ref_Gene database annotated with KEGG ortholog (KO) genes. We used principal-component analysis (PCA) to visualize the difference between the microbiota in BV and the microbiota in non-BV on the levels of taxonomy (16S rRNA gene) (Fig. 3A), taxonomic composition of expressed genes (metatranscriptome) (Fig. 3B), and functional annotation of transcripts to KEGG orthologues (KO genes) (Fig. 3C). Figure 3A shows that the non-BV communities form a tight cluster on the level of the 16S rRNA sequencing, while the BV communities vary, in accordance with the studies using amplicon sequencing of BV. *L. iners*, *Prevotella* spp., *G. vaginalis*, *A. vaginae*, and *S. amnii* drive the separation between non-BV and BV. On the level of the taxonomic composition of the metatranscriptomes (Fig. 3B), this pattern was reversed; samples from non-BV communities were much more heterogeneous than those from BV communities. The non-BV communities clustered into two groups dominated by *L. iners* and *L. crispatus*, respectively, whereas *G. vaginalis*, *A. vaginae*, and *S. amnii* were abundant in the BV communities. This reversal is even stronger on the level of KO genes (Fig. 3C). Samples from BV form a tight cluster, while those from non-BV vary widely. The pattern is opposite that found for the phylogenetic marker gene. The KO genes that contributed most to these differences in the non-BV communities were phosphofructokinase isozyme gene *pfkA* (31) and ribosomal protein coding genes *rpsI*, *rpmF*, and *rplU* (32, 33). In the BV communities, the *msmE*, *cycB*, and *pflD* genes that encode proteins involved in carbohydrate uptake and metabolism (34, 35) were stably more highly expressed.

***In vivo* expression of putative metronidazole resistance-associated genes in *G. vaginalis*.** To clarify the possible contribution of genes related to metronidazole resistance in Gram-positive pathogens to the differences in the responses to treatment of the vaginal microbiota, we examined their expression (Data Set S1, sheet 9) in *G. vaginalis*. For this analysis, BV communities from 11 patients at visit 1 were analyzed, and the level of *G. vaginalis* transcripts was >20%. Six of these patients did not respond to treatment, and five responded. Although *A. vaginae* and *S. amnii* are also key players in BV, we could not analyze them here since there were too few samples dominated by them. As shown in Fig. 4A, there was no clear expression pattern for most of these
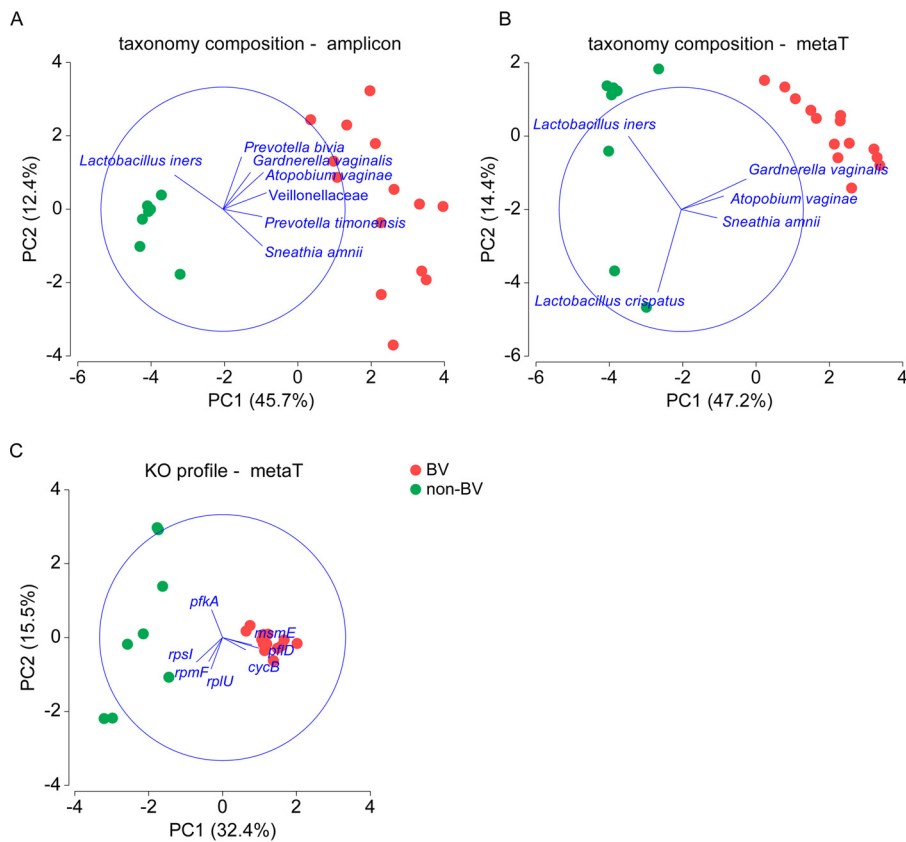
FIG 3 PCA based on taxonomic profiles and activity profiles in BV during and after metronidazole therapy (non-BV). (A) The PCA plot is based on the taxonomic profile determined using 16S rRNA gene sequencing. (B) PCA based on taxonomic composition determined by metatranscriptome. (C) PCA plot based on KO gene expression profile. The communities at visit 1 (BV) and visit 2 (after metronidazole therapy, non-BV) from 14 patients are indicated. In the PCA biplots of taxonomy composition (panels A and B), the taxa with multivariate (multiple) correlation values higher than 0.3 are illustrated, while for PCA of KO profiles (panel C), the KO genes with correlation values of >0.2 are shown.

genes (detailed data are provided in Data Set S1, sheet 9). The only significantly changed gene expression was that of the gene encoding ferredoxin, which was less active in *G. vaginalis* in nonresponding patients (fold change = 1.67; Wilcoxon test P = 0.00866 based on relative abundance [read count of given genes of *G. vaginalis*/read count of *G. vaginalis* percentage]).

CRISPR-associated protein coding genes of *G. vaginalis* were strongly upregulated in vaginal fluids of patients not responding to treatment. We then performed a global analysis of differential expression (DE) of KO genes of *G. vaginalis* in the same communities (visit 1, 11 BV samples with >20% transcripts from *G. vaginalis*, including 6 patients that did not respond to treatment and 5 that responded). We observed that there were 9 KO genes highly upregulated with a false-discovery rate (FDR) of ≤0.05 (log$_2$ fold change of up to 9.46) in communities without response. Strikingly, among the most strongly upregulated KO genes, seven were *cas* genes (36) (Fig. 4B). In total, there were 8 different *G. vaginalis* CRISPR-associated (Cas) genes found in the genomes, namely, *cas1* to *cas3* and *casA* to *casE*, of which 7 (*cas1* to *cas3* and *casA* to *casD*) were upregulated.

There were 9 KO genes that were downregulated, but the fold change values were not as high as for the upregulated genes. The *fucP* (fucose permease) gene was identified as the most strongly downregulated gene, with a log$_2$ fold change of −4.24.

Time course of activity profiles and recurrence. In the second part of our study, we analyzed the metatranscriptome of vaginal fluid samples from four of the patients that initially responded to therapy with metronidazole for the complete duration of the
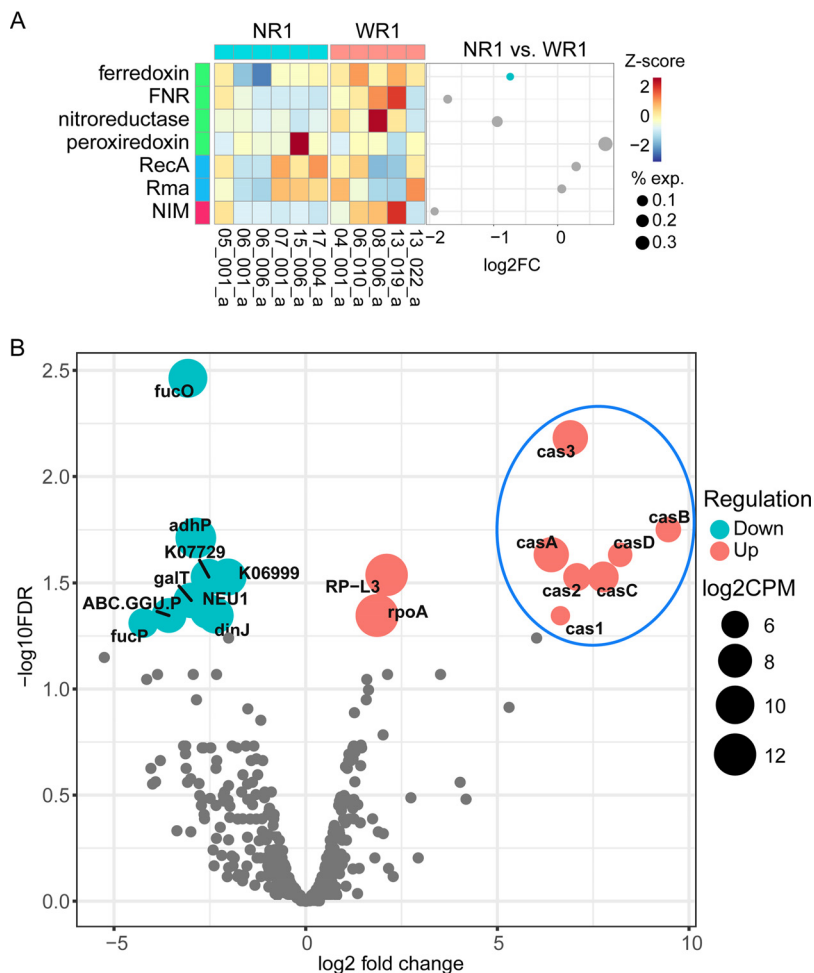
FIG 4 Changes in gene expression of *G. vaginalis* in patients responding to antibiotic treatment compared to nonresponders. (A) Expression of putative metronidazole resistance-associated genes of *G. vaginalis* in vaginal fluid microbiota. (B) Differential expression of KO genes. Seven *cas* genes of *G. vaginalis* were highly upregulated in communities from patients who did not respond to the treatment. (A) The expression values were calculated based on the relative abundances of reads mapped onto *G. vaginalis* using BWA. "NR1" (No response 1) indicates the BV samples from six patients that did not respond to metronidazole treatment; "WR1" (With response 1) represents the BV samples from four patients who afterward responded to metronidazole. The dot plot illustrates the log2 fold change (log2FC) values of the corresponding activities in comparisons between *G. vaginalis* from nonresponders and *G. vaginalis* from responders. The values in the heat map were scaled using the Z-score. In the figure key, "exp." indicates the relative expression level. (B) "NR1" samples were compared with "WR1" samples. KO genes with an FDR of ≤0.05 are colored in red or turquoise (significantly differentially regulated), while those with an FDR of >0.05 are in colored in gray.

clinical trial. Two of these patients experienced recurrence of BV, and two remained stably non-BV. Five time points were analyzed, the first two of which are shown in Fig. 1A as described above. They represented acute BV at the time of diagnosis (visit 1) and after metronidazole therapy (visit 2). Here, we also show data from visits 3 to 5, which were all visits by non-BV subjects, with the exception of recurrence at visit 5 in patient 04_001 and at visit 3 in patient 06_004. Figure 5A shows the taxonomic composition of the communities. *L. crispatus* dominated the microbiota in one of the two patients that stably maintained a non-BV status, and *L. iners* dominated the microbiota in the other. The results of principal-component analysis of the activity profiles are shown in Fig. 5B. In acute BV, samples from all four patients clustered together (red circle). After the treatment, the data corresponding to samples from patient 08_006, who was stably non-BV, moved into a very dense and distinct cluster (illustrated by the arrow 1 and enclosed by a green circle). Samples from patient

mSphere®
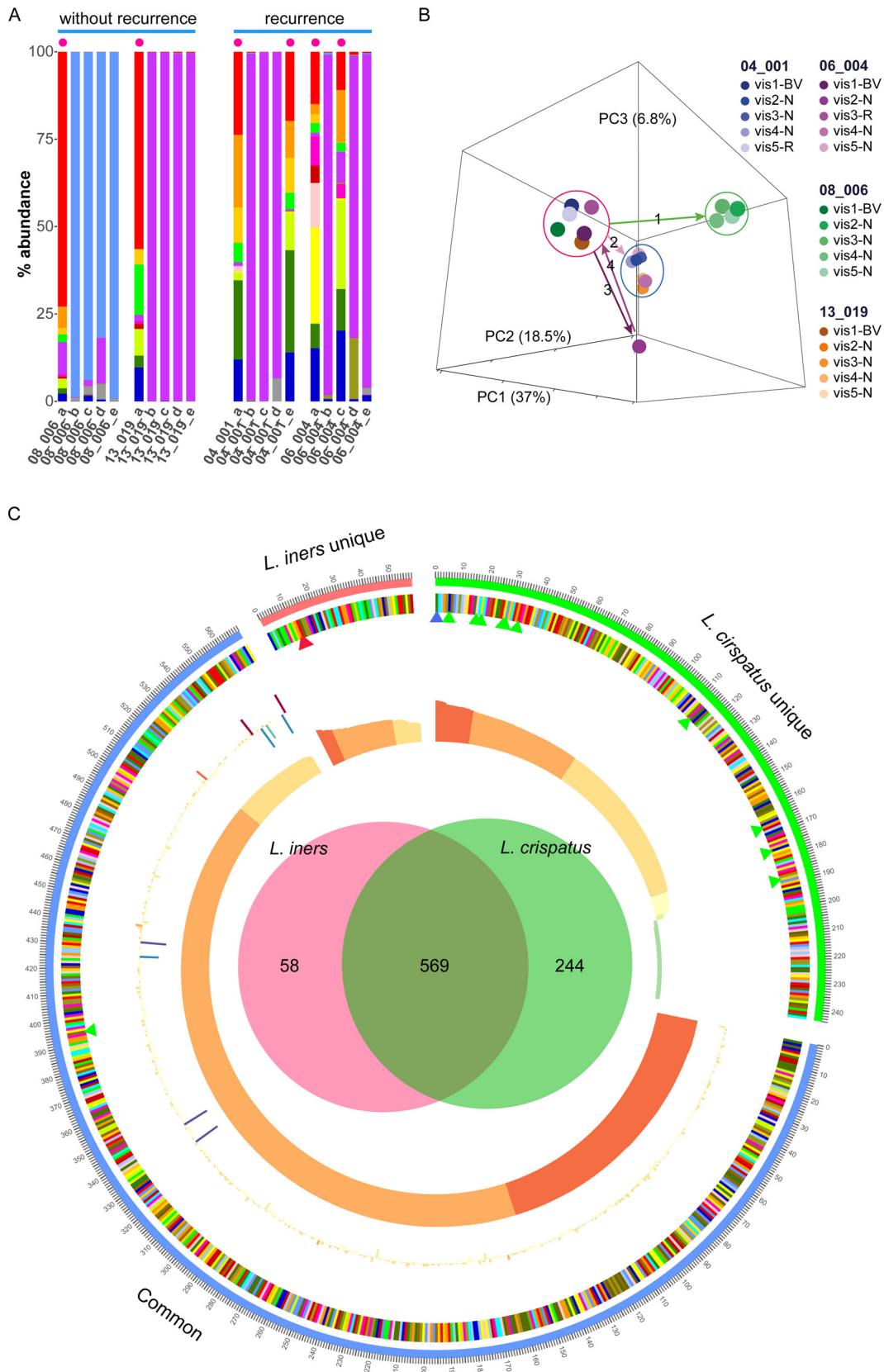

FIG 5 Shifts in the vaginal microbiome over 3 months. (A) Taxonomic composition of the metatranscriptome in two patients that were stably non-BV (without recurrence) and two patients that experienced recurrence. Acute BV and recurrence according to the Nugent score are indicated as red dots. (B) PCA of activity profiles based on KO genes from the same patients.

(Continued on next page)

13_019, who also remained stably non-BV, moved to a different cluster after treatment, shown by arrow 2 and encircled in blue.

The activity shifts in patient 06_004, who experienced recurrence at visit 3, were especially noteworthy. After treatment, the community moved toward an activity profile distinct from all others (arrow 3). The recurrence of BV caused the community to shift back to the BV cluster (red circle, arrow 4). At visit 4, the community moved to the non-BV cluster (blue circle) and the patient became non-BV according to the Nugent score. We speculate that there was an unknown intervention after visit 3 that changed the microbiome but which was not recorded. Interestingly, the other case of recurrence (patient 04_001) had a different progression. From visit 2 to visit 4, patient 04_001 was non-BV and these samples clustered together in the "non-BV" cluster indicated by the blue circle. At visit 5, however, patient 04_001 had recurrent BV and the community shifted back again to the BV cluster.

Transcriptomics of *L. iners* and *L. crispatus in vivo*. The stable colonization of the vaginal fluid of two patients with either *L. iners* or *L. crispatus* that responded to antibiotic therapy allowed us to profile their gene expression *in vivo* to gain more understanding of their different roles in the vaginal microbiota. We extracted the reads mapped on *L. crispatus* in patient 08_006 (time points b to e in Fig. 1A) and *L. iners* in patient 13_019 (time points b to e) and performed a differential expression (DE) analysis using edgeR to compare their activity profiles based on their KO genes, comparing the expression of KO genes of *L. crispatus* in *L. crispatus*-dominated samples ($n = 4$) with the expression of KO genes of *L. iners* in *L. iners*-dominated samples ($n = 4$). The Venn diagram in Fig. 5C shows that the two species share 569 KO genes, while 58 are unique to *L. iners* and 244 are unique for *L. crispatus*, indicating that *L. crispatus* possesses a far greater number of diverse functions than *L. iners*. The DE analysis identified 654 significantly differentially expressed KO genes, of which 393 were upregulated in *L. crispatus* (Data Set S1, sheet 8). Among the top 100 most differentially expressed KO genes in terms of FDR value, 64 were upregulated in *L. crispatus*. Remarkably, genes encoding enzymes involved in the production of H_2O_2 (pyruvate oxidase, NADH oxidase, glycolate oxidase) (37, 38) were highly expressed in *L. crispatus* (Data Set S1, sheet 8, KO genes colored in blue). The D-lactate dehydrogenase gene (K03778) was the most highly expressed gene in *L. crispatus* (log 2 counts per min [$\log_2$CPM] = 13.3) (Data Set S1, sheet 8, colored in light green), and this gene is absent in the genome of *L. iners*. On the other hand, we found that inerolysin (INY) was highly expressed ($\log_2$CPM = 9.8) in *L. iners* but was absent in the genome of *L. crispatus* (Data Set S1, sheet 8, colored in red). Interestingly, we also found the gene orthologous to the inerolysin gene known as the vaginolysin gene (K11031) (39) to be highly expressed in *G. vaginalis* (details in Data Set S1, sheet 7). Hemolysin C, another pore-forming toxin, was highly expressed ($\log_2$CPM = 9.8) in *L. iners* but absent in *L. crispatus*.

DISCUSSION

The aim of this study was to identify activity patterns in the vaginal fluid microbiota in BV and after metronidazole therapy. In particular, we compared the transcriptional profiles of *G. vaginalis* in the vaginal microbiotas of patients who did and did not respond to metronidazole treatment. This is the first study to have investigated the activity alterations of the vaginal microbiota in patients with BV during treatment with

FIG 5 Legend (Continued)

Data from two women with recurrence (pink and blue color range) and two women without recurrence (green and orange color range) are shown. In the figure key, "BV" indicates the time point with BV, "R" indicates the recurrence time point, and "H" represents health. The green and blue circles highlight healthy clusters, while the red circle highlights samples from BV. The arrows denote the temporal shifts of the communities during the treatment. (C) Gene expression *in vivo* of *L. crispatus* and *L. iners*. The Venn diagram indicates the unique KO genes of *L. crispatus* and *L. iners* as well as their shared KO genes. The innermost ring denotes the expression of KO genes quantified by log2CPM; the outer ring illustrates the fold change of the expression of KO genes between *L. crispatus*-dominated communities and *L. iners*-dominated communities quantified by log2FC. The KO genes are shown in descending order based on log2CPM. The small red triangles mark the inerolysin and hemolysin C genes, while blue and green triangles mark the genes encoding proteins involved in the production of D-lactic acid and hydrogen peroxide, respectively.

the antibiotic metronidazole using the metatranscriptomics approach. We found several changes in gene expression in nonresponding patients that might contribute to resistance against metronidazole by either not activating the prodrug or repairing DNA damage.

G. vaginalis was the most dominant active species in BV. *G. vaginalis* can be divided into four phylogenetic subgroups which may in the future be described as subspecies and which differ in virulence and in susceptibility to metronidazole (6, 12). We found transcripts from all four subgroups in all patients, as previously shown based on sequencing of the universal target cp-60 gene (6, 9, 11). Interestingly, in those patients that did not respond to treatment, *Gardnerella* subgroups A and D, whose members are resistant to metronidazole (14), were slightly more abundant.

Sequencing of phylogenetic marker genes such as the 16S rRNA gene or the cp-60 universal target is a fast and sensitive method to profile the microbiota composition, but it does not provide functional information and is prone to PCR bias. Moreover, DNA from dead cells might also be detected. Therefore, we compared the taxonomic composition of the transcripts with that of the 16S rRNA genes determined previously in those samples (5). We observed that *G. vaginalis* comprised on average 47% of all transcripts in BV, while only 5% of 16S rRNA genes were assigned to this species. This suggests that *G. vaginalis* is transcriptionally more active than other vaginal bacteria; moreover, the commonly used 27F primer was previously shown to underrepresent *G. vaginalis* (40). Other differences between the two methods are caused by the low taxonomic resolution of the 16S rRNA gene, especially of short amplicons, where a large fraction of 16S rRNA reads is assigned to higher-level taxa, e.g., genus or family. In contrast, the metatranscriptome reads are mapped to genomes and so have species-level resolution. Finally, transcripts can only be mapped if a genome is available. If the species in question has not yet been cultivated, as, for example, in the case of the BVAB strains, then reads cannot be assigned. In the periodontal pocket microbiota, about 50% of all reads cannot be mapped to any bacterial genome (30). In contrast, the vaginal microbiota is much less diverse and most of its representatives have been cultivated; using the improved ref_Genome database, we were able to map 86% of all reads, indicating that uncultivated taxa did not contribute very significantly to the active community in BV and after metronidazole therapy.

Low diversity in health and high diversity in BV are hallmarks of BV, and it is so striking that it has even been suggested to use diversity indices based on PCR-amplified 16S rRNA genes in addition to the clinical diagnosis based on Amsel criteria and the Nugent score (4, 41–43). Our comparison between communities in BV and after metronidazole therapy on the levels of (i) the 16S rRNA gene, (ii) the taxonomic composition of total transcripts, and (iii) functional profiling based on KO genes shows a reversal of this observation: BV communities, although highly diverse on the taxonomic level, cluster tightly together on the functional level of KO genes. In contrast, non-BV communities are similar on the taxonomic level but are highly diverse among individuals on the functional level.

In our metatranscriptome analysis we found evidence for mechanisms that hinder the activation of the metronidazole prodrug or that mitigate the damage that metronidazole inflicts on DNA and thus could be important reasons for the lack of response in some women.

We show that the ferredoxin gene of *G. vaginalis* was less active in those patients that did not respond to metronidazole. As an electron carrier, ferredoxin is downregulated in *H. pylori* bacteria grown in the presence of metronidazole (17). It is required for activation of the prodrug in *H. pylori* (44) and might have a similar role in *G. vaginalis*. Lack of response might result from lack of activation of the prodrug. Unexpectedly, the nitroimidazole resistance (*nim*) gene, which has been shown to mediate resistance to metronidazole in *B. fragilis* by transforming metronidazole to a nontoxic amino derivative (16), was not highly expressed in nonresponders. This could have been due to technical problems, since the Nim protein sequence contains only partial coding DNA sequences (CDS) (https://www.ebi.ac.uk/ena/data/view/AGN03877). Moreover, *nim-*

negative strains of *B. fragilis* can tolerate high levels of metronidazole, indicating the importance of other mechanisms of resistance (16).

Remarkably, *cas* genes of *G. vaginalis* were highly upregulated in samples from patients that did not respond to metronidazole treatment. The CRISPR-Cas genes are present in about half of all *Bacteria* and most *Archaea* (45); they represent a mechanism of adaptive immunity which protects the prokaryotic cell against foreign DNA and has been developed into a universal tool for genome editing (46). The *cas* genes of *G. vaginalis* belong to the *Escherichia coli* subtype and were found in about half of the clinical isolates (36). Their upregulation might reflect increased phage attacks in BV. Phages have been hypothesized to be crucial for the etiology of BV by causing the collapse of *Lactobacillus* populations (47); accordingly, *L. iners* upregulates its CRISPR-Cas system in BV (22). More than 400 annotated prophage sequences were found in 39 *Gardnerella* strains (29). They might be induced to enter the lytic cycle by the change in pH accompanying the shift to BV. However, the viral transcripts contributed 0.1% of the total metatranscriptome in both nonresponders and responders before treatment.

The upregulation of CRISPR-Cas system genes in *G. vaginalis* from those patients that did not respond to treatment by metronidazole suggests that the CRISPR-Cas system might have a role in mitigating the DNA-damaging effect of metronidazole. In addition to providing adaptive immunity, CRISPR-Cas systems can have various additional functions (48), and it was previously shown that they can protect the cell against DNA-damaging agents (49). The Cas1 enzyme of *E. coli* (YgbT) physically and genetically interacts with the DNA repair system (RecBC, RuvB) and is recruited to DNA double-strand breaks; moreover, YgbT is necessary for resistance of *E. coli* to DNA damage caused by the genotoxic antibiotic mitomycin C or UV light (49). Our findings suggest that the CRISPR-Cas system may protect the vaginal microbiota against the DNA-damaging effect of metronidazole. If experimentally confirmed, this finding might open a new path for fighting bacterial resistance against DNA-damaging agents. For example, it would be worth testing if the susceptibility to metronidazole can be modified in *Gardnerella* isolates and possibly other vaginal pathogens according to the expression level of *cas* genes. It is not known how upregulation of *cas* genes is regulated in the vaginal microbiota. It might be a response to phage attack; thus, by suppressing *cas* genes, the susceptibility to phages might be increased simultaneously with the susceptibility to metronidazole. Using CRISPR engineered phages for therapy of dysbiotic communities has been considered to be one of many options of new therapeutic strategies based on a deeper understanding of the human microbiome (50).

L. iners and *L. crispatus* dominate their respective CSTs in the healthy vaginal microbiota. There were many factors observed by laboratory or genomic studies (37, 51, 52) which suggest that more protection against dysbiosis is provided by *L. crispatus* than by *L. iners*. Here we analyzed which genes are actually highly expressed *in vivo*; we observed that genes encoding proteins for the production of H_2O_2 and D-lactic acid were highly expressed in *L. crispatus*. H_2O_2 inhibits BV-associated bacteria, but it has been questioned if its level in the vaginal milieu is high enough, and it was suggested that lactic acid is more protective (53). In *L. iners*, the genes for the pore-forming toxins inerolysin and hemolysin C were highly active, supporting the hypothesis that *L. iners* may play an ambiguous role in the vaginal econiche and is associated with vaginal dysbiosis (26, 54).

Conclusions. This first study of the *in vivo* transcriptional activity of the vaginal fluid microbiota during metronidazole treatment of BV focused on possible reasons for the lack of response to antibiotic therapy in some patients. Genes related to activation of the prodrug and repairing the DNA damage caused by metronidazole were shown to be differentially expressed in responders and nonresponders. A completely new role for Cas proteins is hypothesized which warrants closer inspection and may help to develop more-efficient novel therapies to improve the treatment of BV.

MATERIALS AND METHODS

Study design. The vaginal fluid samples obtained from the women analyzed here were a subset of the samples obtained during a randomized controlled clinical trial described previously (5). The trial protocol was approved by the local ethics committee (Ärztekammer Nordrhein—Medical Association North Rhine), and written consent was obtained from all participants. The clinical trial was conducted in accordance with the Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects. Principles and guidelines for good clinical practice were followed. The study was registered on ClinicalTrials.gov under identifier NCT02687789. Briefly, women were included in the clinical trial if their samples were BV positive according to Amsel criteria and the Nugent score and were biofilm positive on vaginal epithelial cells and positive for extracellular polysaccharides (EPS) in urine. For treatment of acute BV, they received 2 g of metronidazole orally and were afterward treated with an intravaginal pessary twice a week for 3 weeks. Samples were taken during acute BV (visit 1), after administration of metronidazole 7 to 28 days after visit 1 (visit 2), after pessary application 1 week after visit 2 (visit 3), after continued pessary application 2 weeks after visit 3 (visit 4), and during follow-up 3 months after visit 4 (visit 5).

The aim of the clinical trial had been to compare the levels of effectiveness of two different types of pessary. The results and the taxonomic composition of the vaginal microbial communities have been previously reported (5). For the metatranscriptome analysis reported here, we chose a subset of 14 patients from the clinical trial. These 14 patients consisted of two groups named "with response to treatment" ($n = 8$) and "no response to treatment" ($n = 6$) (Fig. 1). Among the eight patients who responded to treatment, six had no recurrence and two experienced recurrence during the month 3 follow-up. For the analysis of lack of response to metronidazole, samples from all 14 patients were analyzed at two time points: the acute BV time point (visit 1) and 7 to 28 days after antibiotic treatment (visit 2) (28 samples in total). For the analysis of recurrence, samples from all 5 visits were analyzed for 4 patients (2 without recurrence and 2 with) (20 samples in total). These four patients all received the commercially available lactic acid pessary after metronidazole therapy at visit 3 and visit 4. A Nugent score of >6 was used to determine BV status since it is considered the gold standard for BV diagnosis (55) (see Data Set S1, sheet 1, in the supplemental material). The BV status at visit 5 was determined by Amsel criteria as there was no Nugent score available at that time point. The sample identifier (ID) was obtained by concatenating the patient ID and letters "a" to "e," indicating visits 1 to 5.

Sample collection and transport. Vaginal fluid was obtained by infusing 2 ml of saline solution into the vagina followed by rotation against the vaginal wall with a speculum and then collecting the vaginal fluid with a syringe. Approximately 700 μl of the fluid was immediately transferred to a tube containing 2 ml RNAprotect (Qiagen, Germany). The tubes were immediately frozen at −20°C, transported at −20°C within a week, and stored at −70°C.

RNA extraction and mRNA enrichment. RNA was extracted from 1 ml vaginal fluid suspension using a Mo Bio PowerMicrobiome RNA Isolation kit (Qiagen, Germany) with pretreatment (vaginal fluid was centrifuged at 13,000 rpm for 1 min). The pellet was resuspended in MoBio lysis buffer, and the suspension was added to the supplied bead tubes filled with 500 μl ice-cold phenol:chloroform:isoamyl alcohol solution (Carl Roth, Germany). The bead-suspension mix was shaken at 5 m/s for 1 min in 3 intervals which were 2 min apart using a Mo Bio PowerLyzer (Qiagen, Germany). After centrifugation for 1 min at 13,000 rpm and 4°C, the upper phase containing the RNA was further processed according to the manufacturer's instructions, including DNase I treatment. RNA was eluted in 100-μl nuclease-free water and vacuum concentrated to 50 μl. A Ribo-Zero Gold rRNA Removal kit (Epidemiology) (Illumina, USA) was then used for mRNA enrichment with ethanol precipitation according to the manufacturer's instructions. Integrity of RNA was evaluated using a model 2100 Bioanalyzer (Agilent, Germany).

Library preparation, sequencing, and preprocessing of sequencing data. Paired-end mRNA Illumina sequencing libraries were constructed using a ScriptSeq kit (Illumina). Strand-specific paired-end sequencing was performed on a HiSeq 2500 Sequencer to yield 2 × 110-bp paired-end reads. Primers and sequencing adaptors were removed from raw sequencing data, followed by clipping the bases with a quality score of <20 from the reads using Fastq-Mcf (56). After clipping, the remaining reads that were shorter than 50 nucleotides were removed. Thereafter, the rRNA reads were eliminated using SortMeRNA v2.0 (57) with the default parameters.

Taxonomy assignment using Kraken. Kraken (27), an accurate and ultrafast taxonomy assignment tool for metagenomes, was used to determine the taxonomic composition of the metatranscriptome data. Kraken uses the K-mer strategy and the lowest common ancestor (LCA) algorithm to affiliate a given read with a taxon. The standard Kraken database was used with addition of the human genome to identify human reads. The standard database consists of prokaryote genomes ($n = 2,786$) and virus genomes ($n = 4,418$). The human genome (Ver. GRCh38) was additionally downloaded from NCBI.

The ref_Genome database contained 163 bacterial genomes from 105 species of bacteria, including 147 genomes from the urogenital subset of the HMP reference genome sequence data (HMRGD). The complete list of reference genomes in the ref_Genome database can be found in Data Set S1, sheet 3. All results pertaining to taxonomic composition in this study were achieved based on this database.

Short-read alignment by BWA. Kraken was used for taxonomy classification, while BWA was applied to determine the expression of genes. A reference gene database named ref_Gene was constructed which contained the genes from the urogenital tract subset of the HMRGD and the genes from 9 additional genomes (Data Set S1, sheet 4). The genes of *Gardnerella* sp. strain 26-12 and *Gardnerella* sp. strain 30-4 could not be included because only less than half of their CDS are available (around 1,000 genes). The short-read alignments were performed using BWA with the BWA-MEM (58) algorithm. A mapping seed length of 31, which is much longer than the default seed length of 19, was applied to

achieve reliable alignments. Reads that mapped with a mapping quality score (MAPQ) lower than 10 were excluded. MAPQ contains the Phred-scaled posterior probability that the mapping position is wrong (59).

KEGG ortholog (KO) gene annotation of ref_Gene database. The ref_Gene database was annotated using KEGG prokaryote protein sequences. The KEGG prokaryote protein sequence database represents a nonredundant protein data set of *Bacteria* and Archaea at the species level and contains about 7 million nonredundant peptide sequences grouped into 14,390 distinct KO genes. A KO gene contains several genes from different species with similar functions. DIAMOND (60), a much faster alternative to BLASTX, was applied to map the ref_Gene sequences against the KEGG prokaryote protein sequence database with its "more sensitive mode." To obtain reliable annotation, only alignments with sequence identity of ≥ 50 and E values of $\leq 1e-5$ and query coverage of $\geq 70\%$ were taken into account. By annotating the genes in the ref_Gene database to KO genes, we were able to determine the expression profile of KO genes and to investigate the activity shifts in BV based on differential expression analysis of KO genes. The cumulative dominance analysis and PCA were carried out using primer 7 (61).

Differential expression (DE) analysis. All differential expression (DE) analyses were performed using the R package edgeR (62). The Benjamini-Hochberg (BH) method was used to correct the P value of the results of DE analysis with the false-discovery rate (FDR) for multiple comparisons. Genes with a FDR of lower than 0.05 were considered significantly differentially regulated. The sample groups defined for each comparison are listed in Data Set S1, sheet 1.

Detection of putative metronidazole resistance-related genes. To detect the expression of previously reported putative metronidazole resistance genes such as *recA* and the *recA*-mediated autopeptidase (Rma), peroxiredoxin, nitroimidazole resistance protein (NIM), ferredoxin/ferredoxin-NADP reductase (FNR), nitroreductase, and ferredoxin genes, we examined the expression level of these genes for *G. vaginalis* in the vaginal community from patients without response to metronidazole treatment ($n = 6$) as well as in the vaginal community from patients with response ($n = 4$). As most of these genes do not have corresponding KO genes, we annotated the ref_Gene database based on the sequences of these genes using BLASTN. The sequences were retrieved from ENA by key words of each of "ferredoxin," "NADPH flavin oxidoreductase," "nitroreductase," "peroxiredoxin," "pyruvate ferredoxin oxidoreductase," "recA," and "nitroimidazole resistance" plus "*G. vaginalis*." In total, 155 unique sequences of *G. vaginalis* were obtained for the annotation of ref_Gene database. The identification of duplicate sequences was done by SeqKit (63).

Data availability. The sequencing data have been deposited in the European Nucleotide Archive with accession number PRJEB21446.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mSphereDirect.00262-18.

FIG S1, TIF file, 1.3 MB.

DATA SET S1, XLSX file, 17.4 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, Brotman RM, Davis CC, Ault K, Peralta L, Forney LJ. 2011. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci U S A 108(Suppl 1):4680–4687. https://doi.org/10.1073/pnas.1002611107.

2. van de Wijgert JH, Borgdorff H, Verhelst R, Crucitti T, Francis S, Verstraelen H, Jespers V. 2014. The vaginal microbiota: what have we learned after a decade of molecular characterization? PLoS One 9:e105998. https://doi.org/10.1371/journal.pone.0105998.

3. Onderdonk AB, Delaney ML, Fichorova RN. 2016. The human microbiome during bacterial vaginosis. Clin Microbiol Rev 29:223–238. https://doi.org/10.1128/CMR.00075-15.

4. Dols JA, Molenaar D, van der Helm JJ, Caspers MP, de Kat Angelino-Bart A, Schuren FH, Speksnijder AG, Westerhoff HV, Richardus JH, Boon ME,

Reid G, de Vries HJ, Kort R. 2016. Molecular assessment of bacterial vaginosis by Lactobacillus abundance and species diversity. BMC Infect Dis 16:180. https://doi.org/10.1186/s12879-016-1513-3.

5. Gottschick C, Deng ZL, Vital M, Masur C, Abels C, Pieper DH, Rohde M, Mendling W, Wagner-Döbler I. 2017. Treatment of biofilms in bacterial vaginosis by an amphoteric tenside pessary—clinical study and microbiome analysis. Microbiome 5:119. https://doi.org/10.1186/s40168-017-0326-y.

6. Schellenberg JJ, Patterson MH, Hill JE. 2017. Gardnerella vaginalis diversity and ecology in relation to vaginal symptoms. Res Microbiol 168:837–844. https://doi.org/10.1016/j.resmic.2017.02.011.

7. Ahmed A, Earl J, Retchless A, Hillier SL, Rabe LK, Cherpes TL, Powell E, Janto B, Eutsey R, Hiller NL, Boissy R, Dahlgren ME, Hall BG, Costerton JW, Post JC, Hu FZ, Ehrlich GD. 2012. Comparative genomic analyses of 17

clinical isolates of Gardnerella vaginalis provide evidence of multiple genetically isolated clades consistent with subspeciation into genovars. J Bacteriol 194:3922–3937. https://doi.org/10.1128/JB.00056-12.

8. Hill JE, Goh SH, Money DM, Doyle M, Li A, Crosby WL, Links M, Leung A, Chan D, Hemmingsen SM. 2005. Characterization of vaginal microflora of healthy, nonpregnant women by chaperonin-60 sequence-based methods. Am J Obstet Gynecol 193:682–692. https://doi.org/10.1016/j.ajog.2005.02.094.

9. Paramel Jayaprakash T, Schellenberg JJ, Hill JE. 2012. Resolution and characterization of distinct cpn60-based subgroups of Gardnerella vaginalis in the vaginal microbiota. PLoS One 7:e43009. https://doi.org/10.1371/journal.pone.0043009.

10. Chaban B, Links MG, Jayaprakash TP, Wagner EC, Bourque DK, Lohn Z, Albert AY, van Schalkwyk J, Reid G, Hemmingsen SM, Hill JE, Money DM. 2014. Characterization of the vaginal microbiota of healthy Canadian women through the menstrual cycle. Microbiome 2:23. https://doi.org/10.1186/2049-2618-2-23.

11. Albert AY, Chaban B, Wagner EC, Schellenberg JJ, Links MG, van Schalkwyk J, Reid G, Hemmingsen SM, Hill JE, Money D; VOGUE Research Group. 2015. A study of the vaginal microbiome in healthy Canadian women utilizing cpn60-based molecular profiling reveals distinct Gardnerella subgroup community state types. PLoS One 10:e0135620. https://doi.org/10.1371/journal.pone.0135620.

12. Schellenberg JJ, Paramel Jayaprakash T, Withana Gamage N, Patterson MH, Vaneechoutte M, Hill JE. 2016. Gardnerella vaginalis subgroups defined by cpn60 sequencing and sialidase activity in isolates from Canada, Belgium and Kenya. PLoS One 11:e0146510. https://doi.org/10.1371/journal.pone.0146510.

13. Bradshaw CS, Morton AN, Garland SM, Horvath LB, Kuzevska I, Fairley CK. 2005. Evaluation of a point-of-care test, BVBlue, and clinical and laboratory criteria for diagnosis of bacterial vaginosis. J Clin Microbiol 43:1304–1308. https://doi.org/10.1128/JCM.43.3.1304-1308.2005.

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

15. Bradshaw CS, Pirotta M, De Guingand D, Hocking JS, Morton AN, Garland SM, Fehler G, Morrow A, Walker S, Vodstrcil LA, Fairley CK. 2012. Efficacy of oral metronidazole with vaginal clindamycin or vaginal probiotic for bacterial vaginosis: randomised placebo-controlled double-blind trial. PLoS One 7:e34540. https://doi.org/10.1371/journal.pone.0034540.

16. Löfmark S, Edlund C, Nord CE. 2010. Metronidazole is still the drug of choice for treatment of anaerobic infections. Clin Infect Dis 50(Suppl 1):S16–S23. https://doi.org/10.1086/647939.

17. Mendz GL, Mégraud F. 2002. Is the molecular basis of metronidazole resistance in microaerophilic organisms understood? Trends Microbiol 10:370–375. https://doi.org/10.1016/S0966-842X(02)02405-8.

18. Löfmark S, Fang H, Hedberg M, Edlund C. 2005. Inducible metronidazole resistance and nim genes in clinical Bacteroides fragilis group isolates. Antimicrob Agents Chemother 49:1253–1256. https://doi.org/10.1128/AAC.49.3.1253-1256.2005.

19. Steffens LS, Nicholson S, Paul LV, Nord CE, Patrick S, Abratt VR. 2010. Bacteroides fragilis RecA protein overexpression causes resistance to metronidazole. Res Microbiol 161:346–354. https://doi.org/10.1016/j.resmic.2010.04.003.

20. Hilbert DW, Smith WL, Paulish-Miller TE, Chadwick SG, Toner G, Mordechai E, Adelson ME, Sobel JD, Gygax SE. 2016. Utilization of molecular methods to identify prognostic markers for recurrent bacterial vaginosis. Diagn Microbiol Infect Dis 86:231–242. https://doi.org/10.1016/j.diagmicrobio.2016.07.003.

21. Ferreira CST, Donders GG, Parada CMGL, Tristão ADR, Fernandes T, da Silva MG, Marconi C. 2017. Treatment failure of bacterial vaginosis is not associated with higher loads of Atopobium vaginae and Gardnerella vaginalis. J Med Microbiol 66:1217–1224. https://doi.org/10.1099/jmm.0.000561.

22. Macklaim JM, Fernandes AD, Di Bella JM, Hammond JA, Reid G, Gloor GB. 2013. Comparative meta-RNA-seq of the vaginal microbiota and differential expression by Lactobacillus iners in health and dysbiosis. Microbiome 1:12. https://doi.org/10.1186/2049-2618-1-12.

23. Bradshaw CS, Morton AN, Hocking J, Garland SM, Morris MB, Moss LM, Horvath LB, Kuzevska I, Fairley CK. 2006. High recurrence rates of bacterial vaginosis over the course of 12 months after oral metronidazole therapy and factors associated with recurrence. J Infect Dis 193:1478–1486. https://doi.org/10.1086/503780.

24. Sobel JD, Schmitt C, Meriwether C. 1993. Long-term follow-up of patients with bacterial vaginosis treated with oral metronidazole and topical clindamycin. J Infect Dis 167:783–784. https://doi.org/10.1093/infdis/167.3.783.

25. Turovskiy Y, Sutyak Noll K, Chikindas ML. 2011. The aetiology of bacterial vaginosis. J Appl Microbiol 110:1105–1128. https://doi.org/10.1111/j.1365-2672.2011.04977.x.

26. Petrova MI, Reid G, Vaneechoutte M, Lebeer S. 1 December 2016. Lactobacillus iners: friend or foe? Trends Microbiol https://doi.org/10.1016/j.tim.2016.11.007.

27. Wood DE, Salzberg SL. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol 15:R46. https://doi.org/10.1186/gb-2014-15-3-r46.

28. The Human Microbiome Project Consortium. 2012. A framework for human microbiome research. Nature 486:215–221. https://doi.org/10.1038/nature11209.

29. Malki K, Shapiro JW, Price TK, Hilt EE, Thomas-White K, Sircar T, Rosenfeld AB, Kuffel G, Zilliox MJ, Wolfe AJ, Putonti C. 2016. Genomes of Gardnerella strains reveal an abundance of prophages within the bladder microbiome. PLoS One 11:e0166757. https://doi.org/10.1371/journal.pone.0166757.

30. Deng ZL, Szafrański SP, Jarek M, Bhuju S, Wagner-Döbler I. 2017. Dysbiosis in chronic periodontitis: key microbial players and interactions with the human host. Sci Rep 7:3703. https://doi.org/10.1038/s41598-017-03804-8.

31. Hellinga HW, Evans PR. 1985. Nucleotide sequence and high-level expression of the major Escherichia coli phosphofructokinase. Eur J Biochem 149:363–373. https://doi.org/10.1111/j.1432-1033.1985.tb08934.x.

32. Arnold RJ, Reilly JP. 1999. Observation of Escherichia coli ribosomal proteins and their posttranslational modifications by mass spectrometry. Anal Biochem 269:105–112. https://doi.org/10.1006/abio.1998.3077.

33. Bubunenko M, Baker T, Court DL. 2007. Essentiality of ribosomal and transcription antitermination proteins analyzed by systematic gene replacement in Escherichia coli. J Bacteriol 189:2844–2853. https://doi.org/10.1128/JB.01713-06.

34. Webb AJ, Homer KA, Hosie AH. 2008. Two closely related ABC transporters in Streptococcus mutans are involved in disaccharide and/or oligosaccharide uptake. J Bacteriol 190:168–178. https://doi.org/10.1128/JB.01509-07.

35. Shipkowski S, Brenchley JE. 2006. Bioinformatic, genetic, and biochemical evidence that some glycoside hydrolase family 42 beta-galactosidases are arabinogalactan type I oligomer hydrolases. Appl Environ Microbiol 72:7730–7738. https://doi.org/10.1128/AEM.01306-06.

36. Pleckaityte M, Zilnyte M, Zvirbliene A. 2012. Insights into the CRISPR/Cas system of Gardnerella vaginalis. BMC Microbiol 12:301. https://doi.org/10.1186/1471-2180-12-301.

37. Ojala T, Kankainen M, Castro J, Cerca N, Edelman S, Westerlund-Wikström B, Paulin L, Holm L, Auvinen P. 2014. Comparative genomics of Lactobacillus crispatus suggests novel mechanisms for the competitive exclusion of Gardnerella vaginalis. BMC Genomics 15:1070. https://doi.org/10.1186/1471-2164-15-1070.

38. Hertzberger R, Arents J, Dekker HL, Pridmore RD, Gysler C, Kleerebezem M, de Mattos MJ. 2014. H(2)O(2) production in species of the Lactobacillus acidophilus group: a central role for a novel NADH-dependent flavin reductase. Appl Environ Microbiol 80:2229–2239. https://doi.org/10.1128/AEM.04272-13.

39. Gelber SE, Aguilar JL, Lewis KL, Ratner AJ. 2008. Functional and phylogenetic characterization of Vaginolysin, the human-specific cytolysin from Gardnerella vaginalis. J Bacteriol 190:3896–3903. https://doi.org/10.1128/JB.01965-07.

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

41. Srinivasan S, Hoffman NG, Morgan MT, Matsen FA, Fiedler TL, Hall RW, Ross FJ, McCoy CO, Bumgarner R, Marrazzo JM, Fredricks DN. 2012. Bacterial communities in women with bacterial vaginosis: high resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. PLoS One 7:e37818. https://doi.org/10.1371/journal.pone.0037818.

42. Ling Z, Kong J, Liu F, Zhu H, Chen X, Wang Y, Li L, Nelson KE, Xia Y, Xiang C. 2010. Molecular analysis of the diversity of vaginal microbiota associated with bacterial vaginosis. BMC Genomics 11:488. https://doi.org/10.1186/1471-2164-11-488.

43. Oakley BB, Fiedler TL, Marrazzo JM, Fredricks DN. 2008. Diversity of

human vaginal bacterial communities and associations with clinically defined bacterial vaginosis. Appl Environ Microbiol 74:4898–4909. https://doi.org/10.1128/AEM.02884-07.

44. Gerrits MM, van der Wouden EJ, Bax DA, van Zwet AA, van Vliet AH, de Jong A, Kusters JG, Thijs JC, Kuipers EJ. 2004. Role of the rdxA and frxA genes in oxygen-dependent metronidazole resistance of Helicobacter pylori. J Med Microbiol 53:1123–1128. https://doi.org/10.1099/jmm.0.45701-0.

45. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, Barrangou R, Brouns SJ, Charpentier E, Haft DH, Horvath P, Moineau S, Mojica FJ, Terns RM, Terns MP, White MF, Yakunin AF, Garrett RA, van der Oost J, Backofen R, Koonin EV. 2015. An updated evolutionary classification of CRISPR-Cas systems. Nat Rev Microbiol 13:722–736. https://doi.org/10.1038/nrmicro3569.

46. Doudna JA, Charpentier E. 2014. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 346. https://doi.org/10.1126/science.1258096.

47. Blackwell AL. 1999. Vaginal bacterial phaginosis? Sex Transm Infect 75:352–353. https://doi.org/10.1136/sti.75.5.352.

48. Westra ER, Buckling A, Fineran PC. 2014. CRISPR-Cas systems: beyond adaptive immunity. Nat Rev Microbiol 12:317–326. https://doi.org/10.1038/nrmicro3241.

49. Babu M, Beloglazova N, Flick R, Graham C, Skarina T, Nocek B, Gagarinova A, Pogoutse O, Brown G, Binkowski A, Phanse S, Joachimiak A, Koonin EV, Savchenko A, Emili A, Greenblatt J, Edwards AM, Yakunin AF. 2011. A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. Mol Microbiol 79:484–502. https://doi.org/10.1111/j.1365-2958.2010.07465.x.

50. Belizário JE, Napolitano M. 2015. Human microbiomes and their roles in dysbiosis, common diseases, and novel therapeutic approaches. Front Microbiol 6:1050. https://doi.org/10.3389/fmicb.2015.01050.

51. France MT, Mendes-Soares H, Forney LJ. 2016. Genomic comparisons of Lactobacillus crispatus and Lactobacillus iners reveal potential ecological drivers of community composition in the vagina. Appl Environ Microbiol 82:7063–7073. https://doi.org/10.1128/AEM.02385-16.

52. Rampersaud R, Planet PJ, Randis TM, Kulkarni R, Aguilar JL, Lehrer RI,

Ratner AJ. 2011. Inerolysin, a cholesterol-dependent cytolysin produced by Lactobacillus iners. J Bacteriol 193:1034–1041. https://doi.org/10.1128/JB.00694-10.

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

54. Vaneechoutte M. 2017. Lactobacillus iners, the unusual suspect. Res Microbiol 168:826–836. https://doi.org/10.1016/j.resmic.2017.09.003.

55. Mohammadzadeh F, Dolatian M, Jorjani M, Alavi Majd H. 2014. Diagnostic value of Amsel's clinical criteria for diagnosis of bacterial vaginosis. Glob J Health Sci 7:8–14. https://doi.org/10.5539/gjhs.v7n3p8.

56. Aronesty E. 2011. ea-utils: command-line tools for processing biological sequencing data. Expression Analysis, Durham, NC.

57. Kopylova E, Noé L, Touzet H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28:3211–3217. https://doi.org/10.1093/bioinformatics/bts611.

58. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv arXiv:1303.3997 [q-bio.GN]. https://arxiv.org/abs/1303.3997.

59. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

60. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60. https://doi.org/10.1038/nmeth.3176.

61. Clarke KR. 1993. Non-parametric multivariate analyses of changes in community structure. Austral Ecol 18:117–143. https://doi.org/10.1111/j.1442-9993.1993.tb00438.x.

62. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140. https://doi.org/10.1093/bioinformatics/btp616.

63. Shen W, Le S, Li Y, Hu F. 2016. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PLoS One 11:e0163962. https://doi.org/10.1371/journal.pone.0163962.