



# Pregnancy's Stronghold on the Vaginal Microbiome

Marina R. S. Walther-Antônio<sup>1</sup>, Patricio Jeraldo<sup>1,2</sup>, Margret E. Berg Miller<sup>2</sup>, Carl J. Yeoman<sup>3</sup>, Karen E. Nelson<sup>4</sup>, Brenda A. Wilson<sup>2,5</sup>, Bryan A. White<sup>2,6</sup>, Nicholas Chia<sup>1,2,7</sup>, Douglas J. Creedon<sup>8\*</sup>

**1** Department of Surgical Research, Mayo Clinic, Rochester, Minnesota, United States of America, **2** Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States of America, **3** Department of Animal and Range Sciences, Montana State University, Bozeman, Montana, United States of America, **4** Genomic Medicine Group, J. Craig Venter Institute, Rockville, Maryland, United States of America, **5** Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States of America, **6** Animal Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States of America, **7** Department of Physiology and Biomedical Engineering, Mayo College, Rochester, Minnesota, United States of America, **8** Obstetrics and Gynecology, Mayo Clinic, Rochester, Minnesota, United States of America

## Abstract

**Objective:** To assess the vaginal microbiome throughout full-term uncomplicated pregnancy.

**Methods:** Vaginal swabs were obtained from twelve pregnant women at 8-week intervals throughout their uncomplicated pregnancies. Patients with symptoms of vaginal infection or with recent antibiotic use were excluded. Swabs were obtained from the posterior fornix and cervix at 8–12, 17–21, 27–31, and 36–38 weeks of gestation. The microbial community was profiled using hypervariable tag sequencing of the V3–V5 region of the 16S rRNA gene, producing approximately 8 million reads on the Illumina MiSeq.

**Results:** Samples were dominated by a single genus, *Lactobacillus*, and exhibited low species diversity. For a majority of the patients (n=8), the vaginal microbiome was dominated by *Lactobacillus crispatus* throughout pregnancy. Two patients showed *Lactobacillus iners* dominance during the course of pregnancy, and two showed a shift between the first and second trimester from *L. crispatus* to *L. iners* dominance. In all of the samples only these two species were identified, and were found at an abundance of higher than 1% in this study. Comparative analyses also showed that the vaginal microbiome during pregnancy is characterized by a marked dominance of *Lactobacillus* species in both Caucasian and African-American subjects. In addition, our Caucasian subject population clustered by trimester and progressed towards a common attractor while African-American women clustered by subject instead and did not progress towards a common attractor.

**Conclusion:** Our analyses indicate normal pregnancy is characterized by a microbiome that has low diversity and high stability. While *Lactobacillus* species strongly dominate the vaginal environment during pregnancy across the two studied ethnicities, observed differences between the longitudinal dynamics of the analyzed populations may contribute to divergent risk for pregnancy complications. This helps establish a baseline for investigating the role of the microbiome in complications of pregnancy such as preterm labor and preterm delivery.

**Citation:** Walther-Antônio MRS, Jeraldo P, Berg Miller ME, Yeoman CJ, Nelson KE, et al. (2014) Pregnancy's Stronghold on the Vaginal Microbiome. PLoS ONE 9(6): e98514. doi:10.1371/journal.pone.0098514

**Editor:** Jack Anthony Gilbert, Argonne National Laboratory, United States of America

**Received:** February 6, 2014; **Accepted:** May 2, 2014; **Published:** June 4, 2014

**Copyright:** © 2014 Walther-Antônio et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was funded by the National Institutes of Health (NIH) grant R56 AI085096-01, the Center for Individualized Medicine, and the Mayo-Illinois Strategic Alliance for Technology-Based Healthcare. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: creedon.douglas@mayo.edu

## Introduction

The Human Microbiome Project (HMP) has helped define variations in the microbiome of healthy individuals and in providing a basis for assessing the role of the microbiome in disease [1]. Prior to HMP, our understanding of vaginal microbiota was based primarily on clinical diagnosis and culture techniques [2]. These do not always correlate to a particular microbiome and may be of limited use in guiding treatment [3]. With high-throughput sequencing, microbial community profiling [4] has allowed detailed resolution of the vaginal microbiome.

Gajer *et al.* [5] measured changes in microbial diversity and stability over time. Ecological diversity measures were consistent

with past reports of low microbial diversity [6–8], however, a dynamic vaginal microbiome with marked changes through time was found. Given the stability and resilience manifested by multiple microbiome sites [9], the expectation for a healthy vaginal microbiome is one of stability and resilience as well. An unstable microbiome could be susceptible to invasion and dysbiosis.

Perturbations in the vaginal microbiome have been implicated in complications of pregnancy [10–15]. A study of the vaginal microbiome in pregnant and non-pregnant women has shown that pregnancy is characterized by a stable *Lactobacillus* dominated community [16,17]. However, previous studies either did not examine time-longitudinal dynamics of the microbiome during

pregnancy [16] or were restricted largely to particular minority groups [17]. Here, we provide evidence in support of *Lactobacillus* dominated microbiomes and higher stability during pregnancy in a time-longitudinal cohort that complements these existing studies.

## Materials and Methods

### Ethics Statement

Subjects were consented under IRB #10-006257, which was approved and reviewed by the Mayo Clinic Institutional Review Board. Written consent was provided by all subjects.

### Subject Enrollment

Here we report the results from 12 subjects enrolled at the Obstetric Division, Mayo Clinic, Rochester, MN under an IRB approval protocol (Vaginal Microbiome – Protocol 10-006257). The inclusion criteria were the following:  $\geq 18$  years of age; no known pregnancy complications at the first obstetrics visit (uncomplicated pregnancy); ability to provide written informed consent, willingness to participate in the mandatory translational research component of the study; and weight greater than 110 pounds (50 kilograms, a standard requirement in obstetrics studies that include blood draws). Patients having the following criteria were excluded from the study: known immunodeficiency; chronic, active viral infections, including HIV-1/2, HTLV-1/2, hepatitis B/C; known autoimmune disease, such as, rheumatoid arthritis or systemic lupus erythematosus; solid organ or transplant recipient; and multiple gestations. Upon enrollment the subjects were requested to fill a questionnaire about sexual and reproductive health and history. The metadata from the questionnaires was stored at REDCap [18]. Here we present the results from the first 12 patients with a healthy pregnancy progression and outcome.

### Sample Collection

Two Dacron swabs were used by the obstetrician or certified nurse midwife to sample the posterior fornix and cervix at time points: 8–12, 17–21, 27–31, and 36–38 weeks of gestation (Figure 1) and placed in a NAT (Nucleic Acid Transport, CentraCare Laboratory Services, St. Cloud, MN) collection tube. The last sample for one of the subjects (#110) was not collected. After collection the samples were kept at  $-80^{\circ}\text{C}$  until processing to preserve the DNA material from degradation.

### Sample Processing

Once thawed the samples were centrifuged for 10 minutes at 10,000 *g* to collect the bacterial cells, and the supernatant was discarded. Genomic DNA extraction was performed by using the MoBio Ultraclean Soil Kit (MoBio Laboratories, Inc., Carlsbad, CA); however, instead of vortexing, MP FastPrep (MP Biomedicals, Solon, OH) for 40 seconds at 6.0 m/s was substituted to obtain a more effective and rapid lysis of the cells. Incubation period was done for a minimum of 30 minutes and a maximum of overnight. After extraction the DNA content was measured using High Sensitivity Qubit (Life Technologies Corporation, Carlsbad, CA) with the results ranging from 25–60 ng/ $\mu\text{l}$ . The V3–V5 region of the 16S rRNA was then amplified through a polymerase chain reaction (PCR) as follows: 25  $\mu\text{l}$  of Kapa HiFi (Kapa Biosystems, Woburn, MA), 1.5  $\mu\text{l}$  (10  $\mu\text{M}$ ) forward primer, 1.5  $\mu\text{l}$  (10  $\mu\text{M}$ ) reverse primer, 50 ng of DNA with the remaining volume being added by molecular grade water (up to a final volume of 50  $\mu\text{l}$  per reaction). The forward primer was the universal primer 357F (5'GTCCTACGGGAGGCAGCAG3') with the added construct on the 5' end of the 5' Illumina Adapter (5'AATGATACGGCGACCACCGAGATCTACAC3') + Forward Primer

Pad (5'TATGGTAATT3') to a total sequence: 5'AATGATACGGCGACCACCGAGATCTACACTATGGTAATTGT-CCTACGGGAGGCAGCAG3' and the universal bacterial reverse primer was 926R (5'CCGTCAATTCMTTTRAGT3') with an added construct on the 5' end of the reverse complement of 3' Illumina adapter (5'CAAGCAGAAGACGGCATAACGATGATGCCGCATTCGAT3') + Barcode (12 base pairs) to a total sequence: 5'CAAGCAGAAGACGGCATAACGATGATGCCGCATTCGATXXXXXXXXXXXXCCGTCAATTCMTTTRAGT3'. The barcode introduced in the reverse primer construct was unique to each sample, functioning as a genetic ID for sequencing. The PCR cycle was the following:  $95^{\circ}\text{C}$  for 3 min,  $98^{\circ}\text{C}$  for 20 sec,  $70^{\circ}\text{C}$  for 15 sec,  $72^{\circ}\text{C}$  for 15 sec, cycle repeated for 34 times and  $72^{\circ}\text{C}$  for 5 min. The products of the amplification were verified by TapeStation D1K Tape (2200 TapeStation Instrument, Agilent Technologies, Santa Clara, CA) to be free of contamination and the expected amplification size, approximately 700 base pairs. Upon verification the PCR products were purified using Agencourt AMPure (Beckman Coulter, Brea, CA). After purification the concentrations were measured using Qubit High Sensitivity and were found to be  $> 1$  ng/ $\mu\text{l}$ . The first samples 8–12 weeks of gestation were processed at the University of Illinois Urbana-Champaign with an identical protocol. Subsequently, all of the samples were then collected and sent to the University of Illinois Urbana-Champaign for 16S rRNA sequencing using a high-throughput next-generation Illumina MiSeq (250 PE, San Diego, CA) sequencing platform.

### Sequencing Outcome

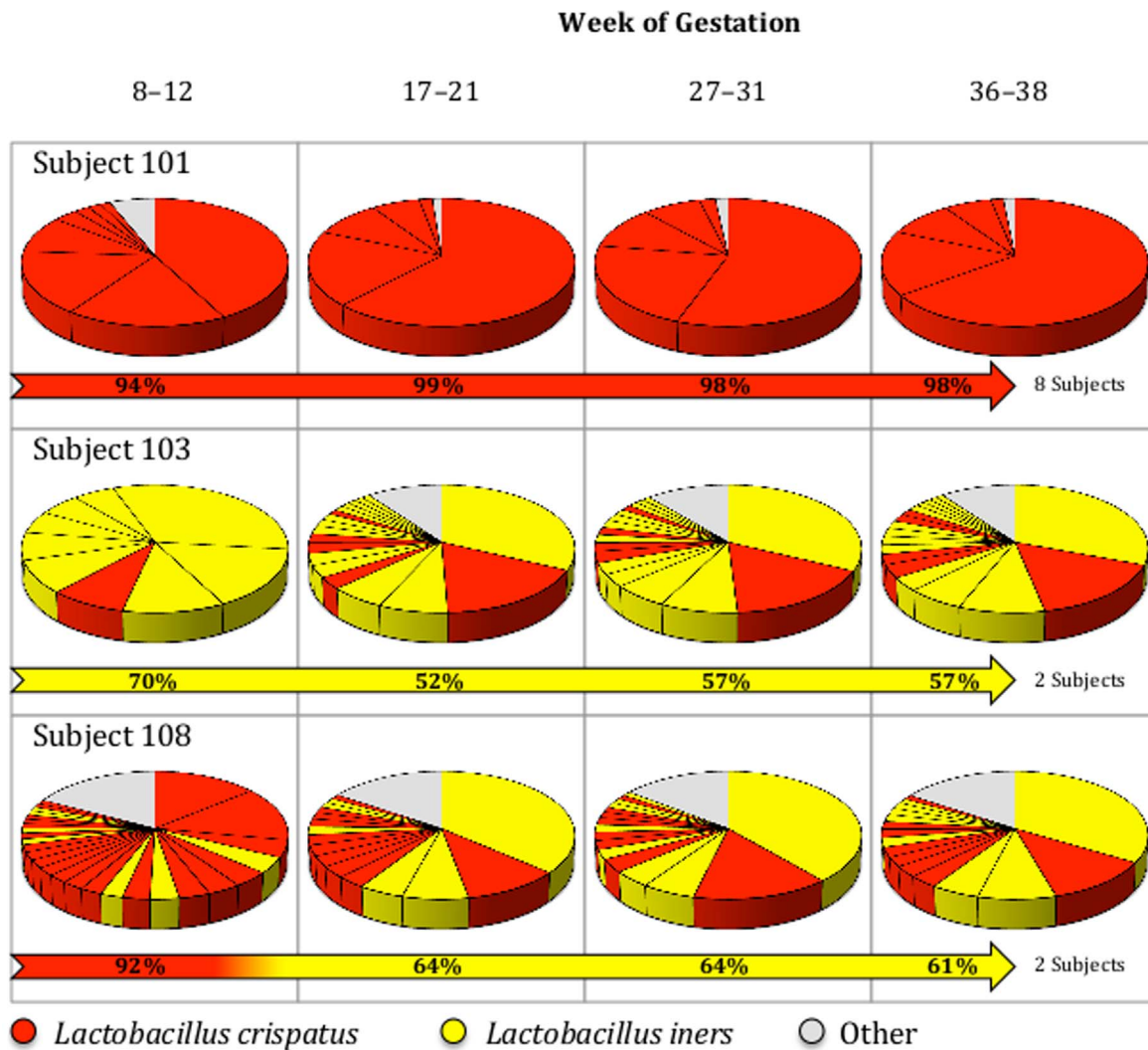
A total of approximately 3,721,000 sequence reads (6,000 to 300,000 reads per sample) were obtained. After quality assessment (sequence reads with less than 187 bp in both R1 and R2 were discarded), 2,424,872 sequence reads (71 to 235,133 sequences per sample) remained and were found to be suitable for further analysis. Despite similar DNA concentrations the number of sequence reads was significantly different (lower) between the first collection time point and the last collection time point. For this reason these samples were repeatedly sequenced both at the University of Illinois Urbana-Champaign and the Mayo Clinic Medical Genome Center with similar results replicated. The reasons for the observed difference are unknown at this point, although they do not appear to have impacted the results.

### Sequence Analysis

Sequence reads were aligned with our own custom multiple alignment tool known as the Illinois-Mayo Taxon Operations for RNA Dataset Organization (IM-TORNADO) that merges paired end reads into a single multiple alignment and obtains taxa calls [19]. IM-TORNADO then clusters sequences into operational taxonomic units (OTUs) using AbundantOTU+ [20]. Further processing for visualization was performed using QIIME [21]. The sequence reads are publicly available in MG-RAST (<http://metagenomics.anl.gov>), sequence IDs: 4563804.3–4563899.3.

## Results

Our cohort consisted of 12 Caucasian subjects, 5 multiparous patients plus 7 primigravids who ranged in age from 24–36 years with a mean age of  $29 \pm 4$ . The participants all self-reported having experienced regular periods prior to pregnancy and to be heterosexual. Within the cohort, there were no complications of pregnancy, including pre-term birth or gestational diabetes.



**Figure 1. Pie charts representing the bacterial distribution found in each of the samples represented.** Subject 101 is representative of 7 other patients with similar profiles (Subjects 104, 105, 106, 107, 109, 113 and 114). Subject 103 and Subject 110 have similar profiles. Subject 108 and Subject 111 have similar profiles. Divisions within the pie charts represent different OTUs (Operational Taxonomic Units – Strain level). “Other” represents OTUs with  $\leq 1\%$  abundance that belong to various species, excluding *L. crispatus* and *L. iners*. doi:10.1371/journal.pone.0098514.g001

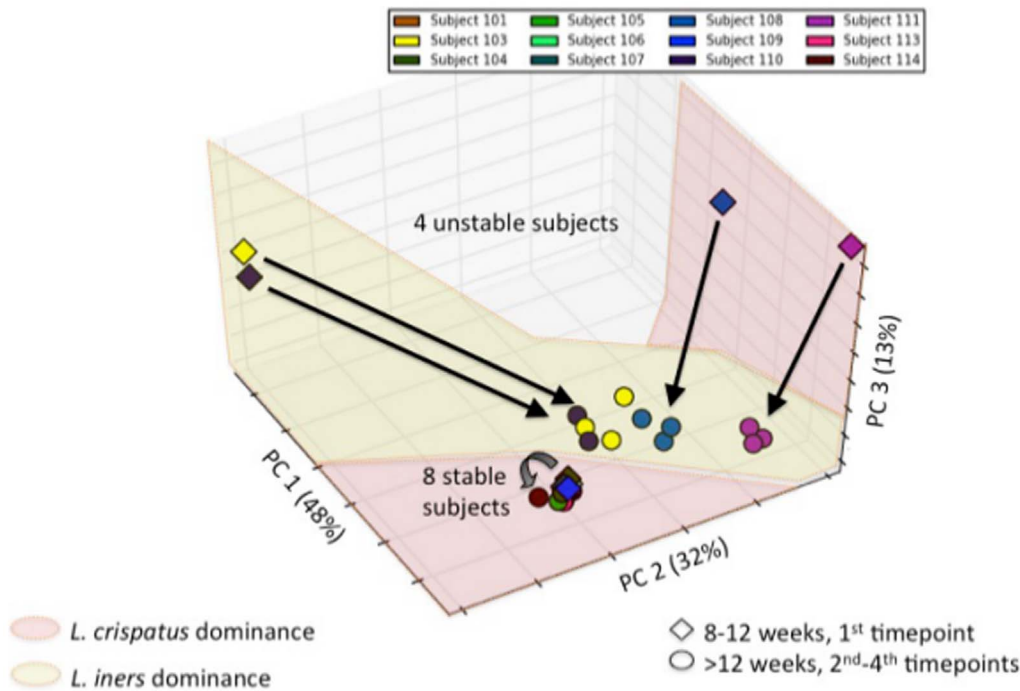
The analysis of the sequence reads revealed that two species dominated the microbial content ( $>1\%$  representation) of samples from the entire cohort (Table S2). The identified species were *L. crispatus* and *L. iners*. Among the 12 patients there were 3 profiles that could be distinguished (Figure 1): Eight of the subjects showed a high prevalence ( $>90\%$ ) of *L. crispatus* throughout pregnancy; two of the subjects showed a prevalence of *L. iners* (92–61%); and the remaining 2 subjects showed a transition in dominance after the first trimester of gestation from *L. crispatus* (70%) to *L. iners* (52–57%). Principal component analysis (PCA) confirmed the existence of the three profiles (Figure 2). Participants dominated by *L. iners* throughout pregnancy were significantly older ( $35 \pm 0$  years old vs.  $28 \pm 3$  years old and  $28 \pm 0$  for the *L. crispatus* and transitory profiles, respectively; Table 1). All other metadata parameters were found not be significant (Metadata S1).

Shannon's diversity indexes [22] (Figure 3) showed that the profile characterized by a dominance of *L. crispatus* throughout pregnancy is less diverse than the other profiles (except during the first trimester).

### Meta-Analysis

In order to allow for a better interpretation of the results we performed a meta-analysis comparing our results with those recently published by Romero *et al.* [17]. In this published work 22 pregnant women (19 African-American, 2 Caucasian and 1 Hispanic) were sampled throughout their pregnancy. Since all of our subjects were Caucasian, the comparison between both studies allows for an investigation of putative ethnic differences. It is well known that African-American healthy vaginal microbiotas exhibit a higher degree of inter-subject diversity than observed among Caucasian subjects [23]. Since the risk factor for the development of both gynecological [23] and obstetrical [24] morbidities is higher in African-Americans [23] it becomes of interest to characterize the healthy state of the vaginal microbiome in both groupings in order to investigate if the higher inter-subject variability is maintained in African-Americans or is homogenized by pregnancy.

We subjected data from Romero *et al.* [17] to our analysis, which focuses on the microbial community dynamics between trimesters,



**Figure 2. Principal Component Analysis (PCA) of temporal dynamics of the vaginal microbiome during pregnancy.** Arrows indicate shifts from the first timepoint, taken at 8–12 weeks, and the subsequent timepoints. Notice that all temporal profiles that involve *L. iners* dominance involve a shift in the microbiome profile while purely *L. crispatus* dominated temporal profiles do not. Visualization carried out using Matplotlib version 1.2.1 (The MathWorks, Inc.).  
doi:10.1371/journal.pone.0098514.g002

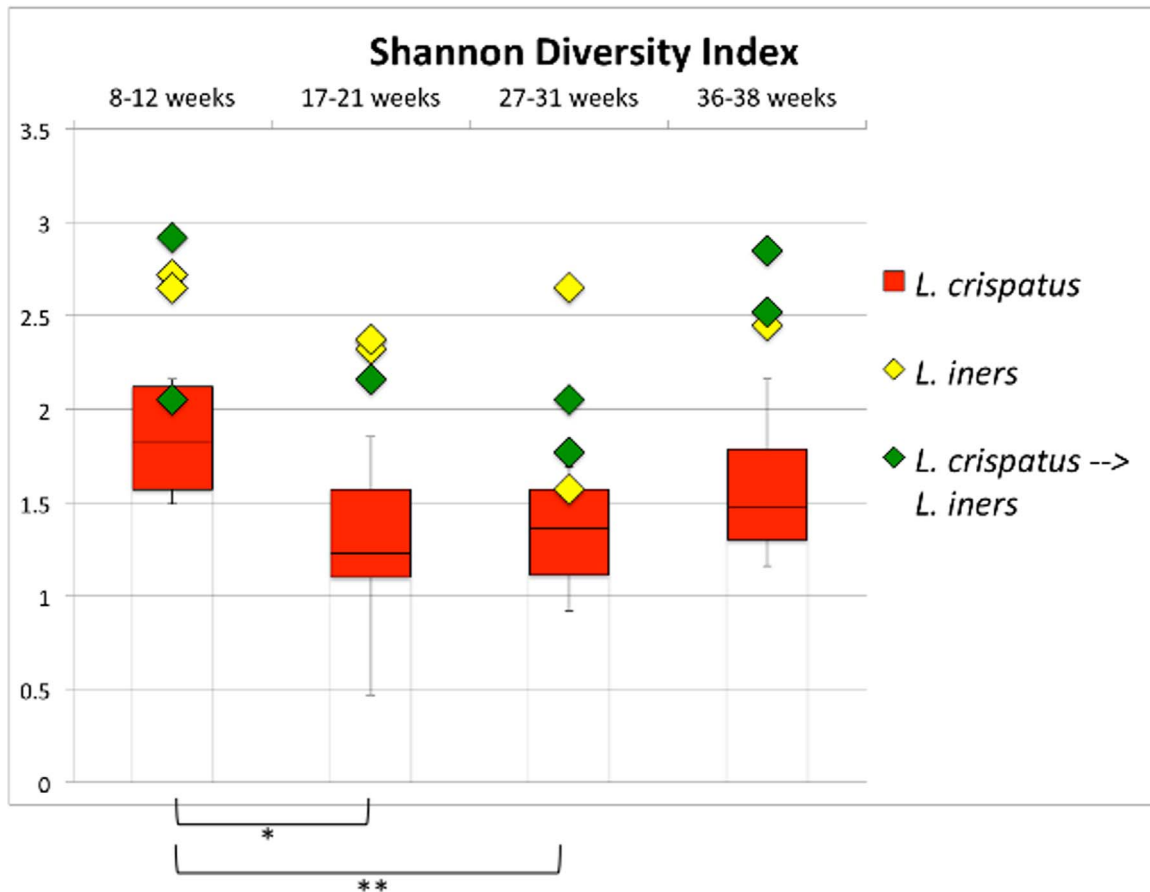
**Table 1. Significant parameters found in the metadata.**

Subject	Age (years)
<b>Crispatus</b>	
101	32
104	24
105	30
106	26
107	32
109	27
113	27
114	27
<b>Average</b>	<b>28</b>
<b>Standard Deviation</b>	3
<b>Iners</b>	
103	36
110	34
<b>Average</b>	<b>35*</b>
<b>Standard Deviation</b>	1
<b>Crispatus-Iners</b>	
108	28
111	28
<b>Average</b>	<b>28</b>
<b>Standard Deviation</b>	0

Significant differences from the *L. crispatus* profile, as determined by t-test, are highlighted by \*.  
doi:10.1371/journal.pone.0098514.t001

in order to compare results with this concurrent study. Due to the different 16S rRNA region amplified in the different studies (V1–V3 in Romero’s study and V3–V5 in this study) and the different sequencing platforms (454 in Romero’s study and Illumina in this study) the comparison is limited (See Methods S1 for procedures and limitations; Figure S1; Figure S2, and Table S1 for the parameters analyzed). Despite limitations due to the different 16S rRNA primers and sequencing platform utilized in the Romero *et al.* study, some meaningful comparative analyses were possible through the use of reference OTUs [21]. The results show that in both ethnicities pregnancy significantly decreases  $\alpha$ -diversity (Figure 4). Comparison in overall diversity between ethnicities did not reveal any significant differences regardless of pregnancy state (Figure S3). Overall, we found divergence between subjects to be higher between non-pregnant subjects in Caucasians than between pregnant ones ( $p < 0.01$ , Figure 5). The opposite was observed between non-pregnant and pregnant African-Americans, where inter-subject variability was lower than between pregnant individuals ( $p < 0.01$ , Figure 5). Comparison of the relative abundance of the genus *Lactobacillus* in the non-pregnant state is significantly higher in Caucasians than African-Americans but this proportion becomes non-significant when pregnant subjects are compared (Figure 6).

While many meaningful results arose from our meta-analysis, certain examinations were unfortunately obscured by the effect of different sequencing platforms. (Table S1; Figure S2). For the pregnant subjects, the effects of ethnicity on  $\beta$ -diversity appear to be overwhelmed by the effects of primer choice and platform. A major limiting factor was present in side-by-side time-longitudinal analyses and ethnicity as we identified only 3 African-American subjects with longitudinal data across all 3 trimesters (all other subjects did not have samples collected during the first trimester).



**Figure 3. Shannon Diversity Index for the vaginal microbiome samples.** The profile characterized by a dominance of *L. crispatus* throughout pregnancy (shared by 8 of the subjects) is significantly less diverse than the two other profiles (except during the first trimester). \*  $p < 0.05$ ; \*\*  $p < 0.005$  on paired t-test.

doi:10.1371/journal.pone.0098514.g003

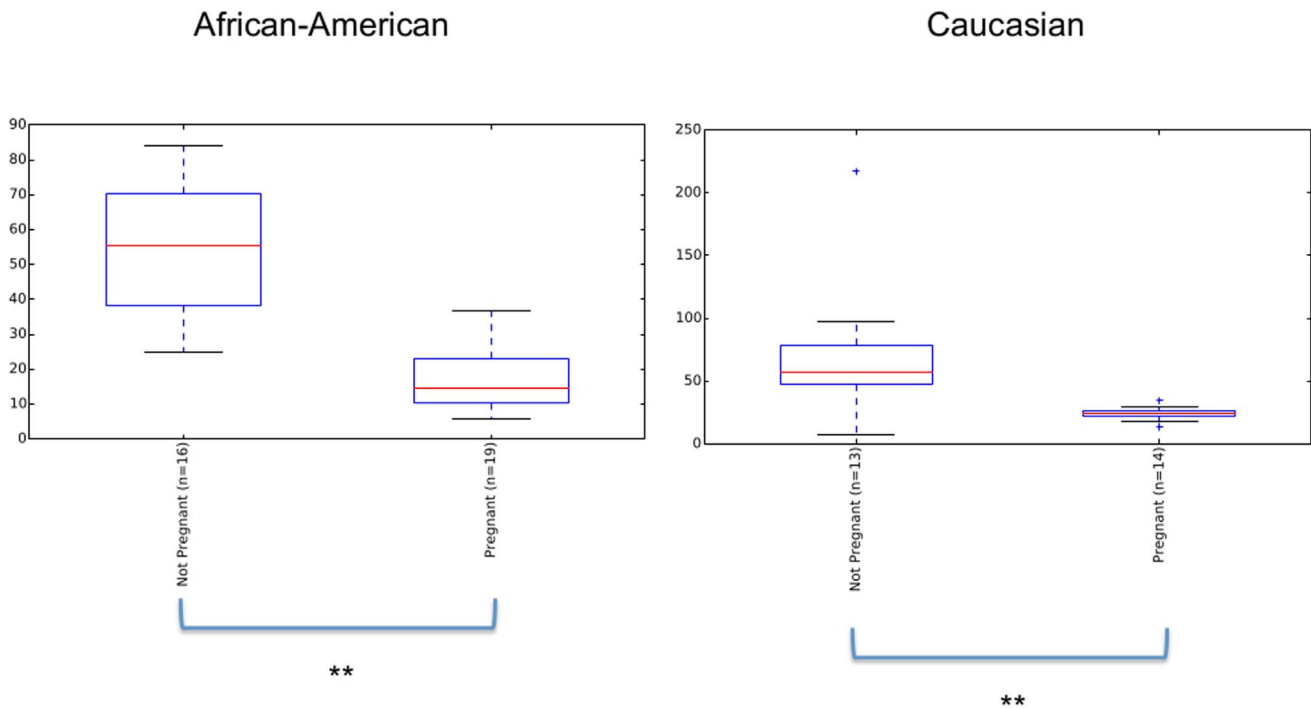
Though small in size, we nonetheless compared the results from these 3 African-American subjects and our 12 Caucasian subjects and showed that the diversity observed between the African-American subjects varied widely across these 3 cases in comparison to the 12 Caucasian subjects (Figure 7). The PCA plots shown in Figure 8 further showed that the African-American subjects clustered by subject and did not show a common attractor as the pregnancies progressed. This is in contrast with what was observed in the Caucasian subjects, which did not cluster by subject, but instead showed a strong common attractor towards the later stages of the pregnancy.

## Discussion

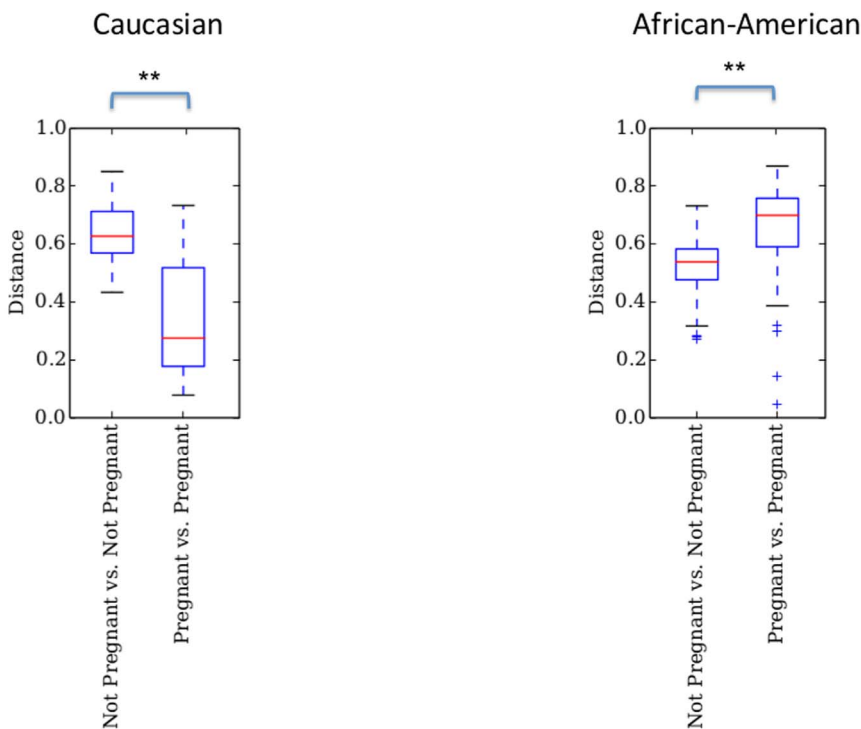
The vaginal microbiome seems unique among the sites sampled for the HMP in terms of its low diversity [1,8,16]. A recent cross-sectional study by Aagaard *et al.* [16] found enrichment for *Lactobacilli* in the vaginal microbiome of pregnant women. Our results are consistent with this finding and identify only two species that comprised  $>1\%$  of the microbial population in our samples—*L. crispatus* and *L. iners*. Furthermore, our longitudinal sampling reveals that this low diversity is also matched by a high stability across the course of pregnancy. This contrasts with non-pregnant women who have been shown to exhibit fluctuations with menstrual cycles [5].

Numerous factors may contribute to the stability of the vaginal microbiome during pregnancy, including: a lack of cyclic fluctuations in hormones, absence of menstrual flow, absence of alterations in the cervical or vaginal secretions associated with the reproductive cycle, or changes in sexual activity during pregnancy [25,26–30]. Some or all of these factors could be implicated in the stability of the vaginal microbiome during pregnancy [31].

The Shannon Diversity Index revealed that the *L. crispatus*-dominant profile was associated with a lower diversity, suggesting that *L. crispatus* is more exclusionary when dominant than *L. iners*. Our findings are consistent with those reported in a cross-sectional study by Ravel *et al.* [8] that profiled the vaginal microbiome in non-pregnant women. The relative amount of dominance exhibited by *L. crispatus* and *L. iners* may be a factor in the susceptibility to changes in the microbiota. In addition, *L. iners* has been identified as a transitional vaginal microbe, potentially indicating susceptibility to dysbiosis or a recovery from it [32]. In this study we did not verify any difference in the pregnancy progression or outcome of women whose vaginal microflora was dominated by *L. iners*, although the fact that these profiles were more diverse than the ones dominated by *L. crispatus* may indicate a higher susceptibility to change. Overall, it is well known that *Lactobacillus* species actively protect themselves and the vaginal environment from invaders by the production of lactic acid which acidifies the vaginal pH as well as the production of  $H_2O_2$  which prevents ascending infection [33]. At least 25% of pre-term births

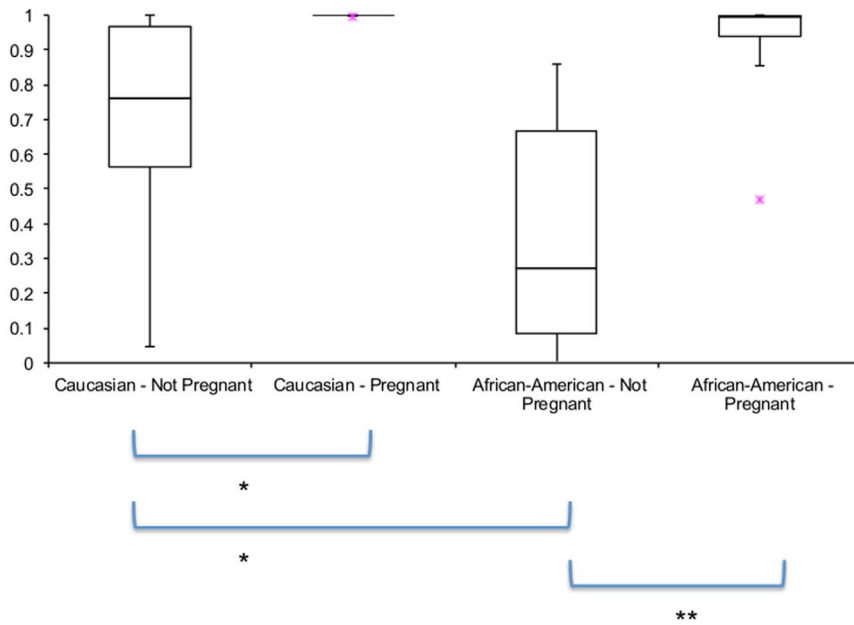


**Figure 4. Pregnancy effect in African-American and Caucasian subjects as measured by Chao1 diversity Index.** Diversity is significantly reduced during pregnancy in both ethnicities (\*\* $p < 0.01$ , Monte Carlo analyses, 999 permutations). doi:10.1371/journal.pone.0098514.g004



**Figure 5. Unweighted Unifrac beta-diversity between pregnant and non-pregnant subjects within each ethnicity.** Caucasians show significant convergence as a result of pregnancy (results confounded by platform effects – See Table S1) while African-Americans show significant divergence in pregnancy when compared to non-pregnant subjects. (\*\* $p < 0.001$ , Monte Carlo analyses, 999 permutations). doi:10.1371/journal.pone.0098514.g005

# Lactobacillus

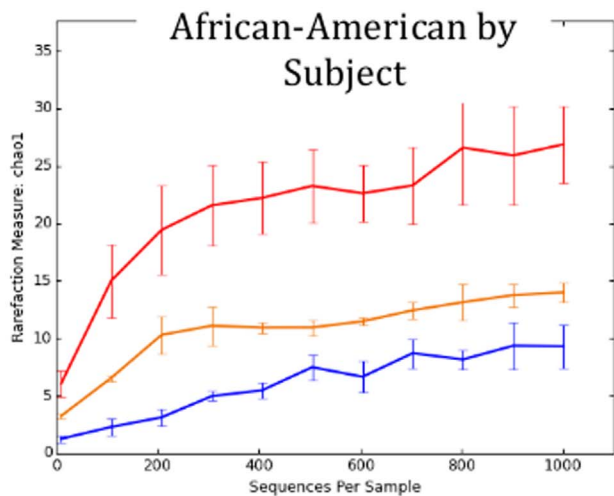


**Figure 6. *Lactobacillus* species frequencies across ethnicities and pregnancy states.** Significant differences were found between all categories except between pregnant Caucasians and pregnant African-Americans (\* $p < 0.05$ ; \*\* $p < 0.01$ ; Monte Carlo analysis, 999 permutations). doi:10.1371/journal.pone.0098514.g006

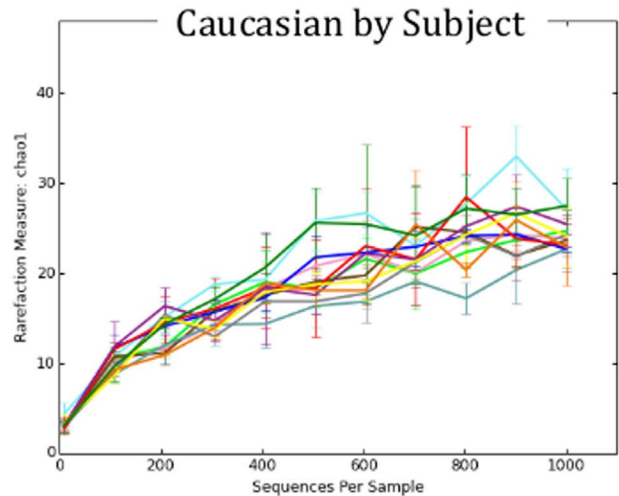
are associated with infection, and a healthy vaginal microbiome may be a stronghold against potential microbial invaders [34].

One of the limitations of this study was the small sample size and the homogeneous population. All women in this study were Caucasian. However, the large majority of participants in the

complementary study from Romero *et al.* [17] are from African-American women. Although still limited by numbers, our meta-analysis and comparison allowed us to draw inferences in these two populations. Perhaps, the most important limiting factor in the meta-analysis are the differences in hypervariable 16S rRNA

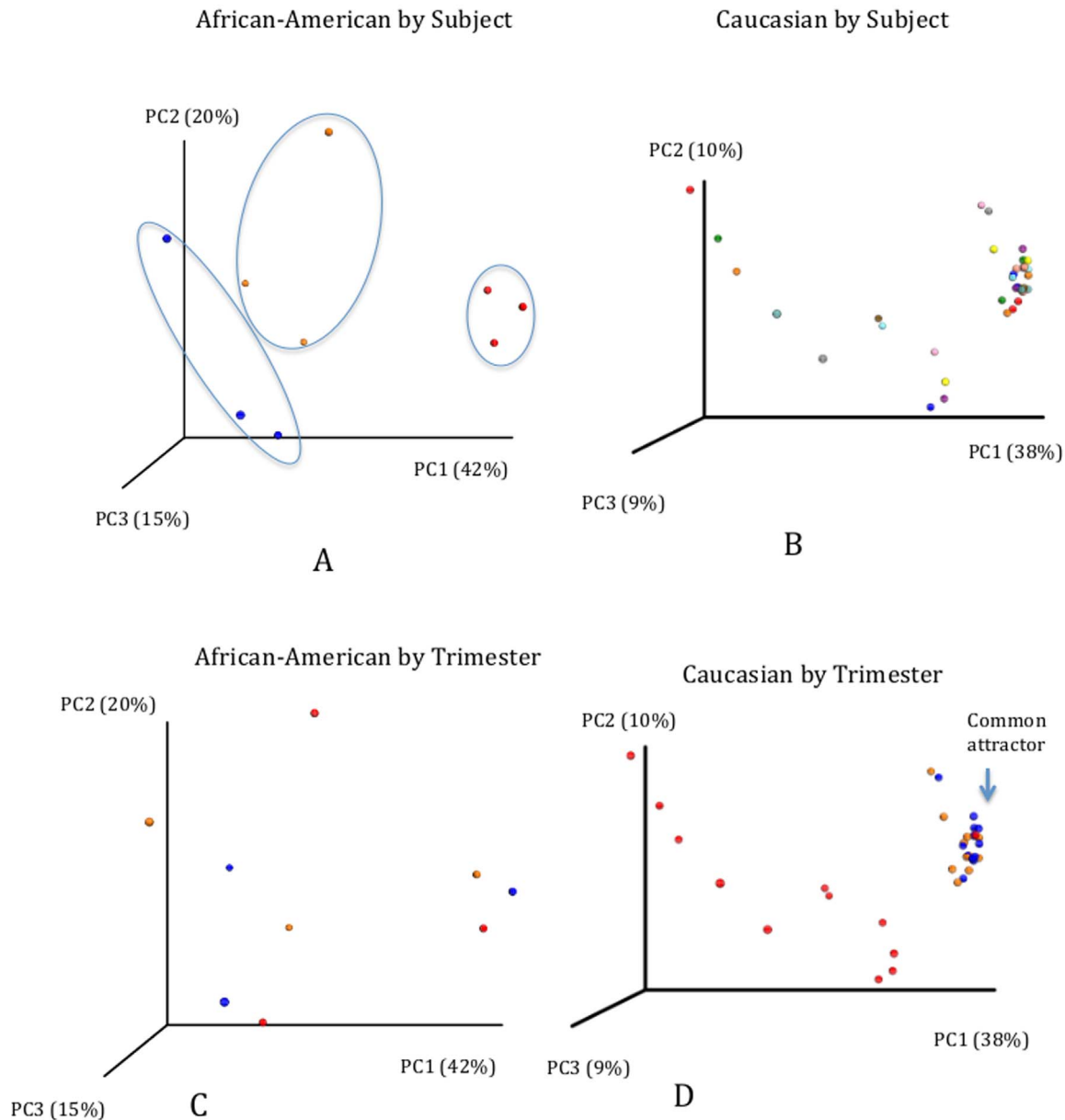


A



B

**Figure 7. Estimation of species richness according to Chao 1 index.** A – African-American subjects (Red – N009, Orange – N018, Blue – N017). Diversity between subjects is significant and therefore does not allow for the grouping of the subjects for analytical purposes. B – Caucasian subjects. No significant differences in diversity were found among subjects. doi:10.1371/journal.pone.0098514.g007



**Figure 8. Unweighted PCA analysis by Subject and Trimester.** A – African-American by Subject. Clustering by Subject can be observed. B – Caucasian by Subject. No clear clustering is observed. C – African-American by Trimester. No common attractor can be identified. D – Caucasian by Trimester. A common attractor can be observed. Plots C and D – 1<sup>st</sup> Trimester – Red; 2<sup>nd</sup> Trimester – Blue; 3<sup>rd</sup> Trimester – Orange. doi:10.1371/journal.pone.0098514.g008

regions amplified (V1-V2 in Romero’s dataset and V3-V5 in our study) and sequencing platforms used (454 in Romero’s and Illumina in ours). Attempts to single out the biases and artifacts introduced by these factors were made and the result of such comparisons determined that these artifacts made many of the direct comparisons between both datasets inconclusive (details presented in the Methods S1, Figure S2 and Table S1). This further highlights the importance of standards in the burgeoning field of microbiome research [35] in the women’s reproductive tract.

Despite some of these technical challenges important conclusions can be inferred from the results. It is well known that there are differences between the microbiotas of African-American and Caucasian women, with the latter being more often reported to

have a higher representation of *Lactobacillus* species [23]. The vaginal health benefits of a *Lactobacillus* dominated flora are well-described, and the fact that African-American women are more susceptible to gynecological morbidities such as bacterial vaginosis and vaginitis may be linked to a flora depleted in *Lactobacillus* [23] or a functional substitute that will maintain the acidity and peroxide contents that shield the vaginal canal from a multitude of pathogens. The dominance of *Lactobacillus* in Caucasian women seen on our cohort is telling when coupled with the fact that pregnancy complications leading to preterm labor and delivery are higher in African-American women, at a rate of 16–18% of all deliveries, when compared to 5–9% in Caucasian [24].

A significant difference was found in the dynamics of the vaginal microbiome where Caucasians experienced a diminished diver-



gence of the microbiota as pregnancy progressed, while African-American subjects exhibited an augmented divergence between subjects. Ravel et al. [8] have shown that African-American subjects are more likely to have a vaginal microbiome not dominated by *Lactobacillus*. However, Zhou et al. [23] have shown that Caucasians vaginal microbiota is commonly dominated by similar numbers of more than one species of *Lactobacillus* while this was rarely observed in African-Americans. Hence, both these observations confound diversity estimations between the two ethnic groups. We have not found a statistically significant difference in the vaginal microbiome diversity between African-Americans and Caucasians either non-pregnant or pregnant (Figure S3). What we did find was that the non-pregnant state is significantly more diverse than the pregnant state regardless of ethnicity (Figure 4). In accordance with Romero et al. [23] we verified that the proportion of *Lactobacillus* dominated communities is different between the two ethnicities when not pregnant (higher in Caucasians) but becomes statistically indistinguishable during pregnancy (Figure 6). This finding is an indicator that pregnancy has a marked effect in the vaginal microbiome, which experiences a complete dominance by *Lactobacillus* species regardless of ethnicity.

Interestingly, examination of the microbial community dynamics using principal coordinate analysis reveals that Caucasian women cluster by trimester towards a common attractor (Figure 8), suggesting that these subjects share a common microbiome dynamic. On the other hand, African-American women cluster by subject and do not show a common attractor (Figure 8). This is also supported by the wider analysis shown on Figure 5. It is tempting to speculate that these differences in microbial dynamics may underlie the increased risk of pregnancy complications in particular individuals in the African-American population. However, due to the multitude of other factors that may also vary across these two populations, it is difficult to isolate the cause. Nonetheless, it is reasonable to consider the role of the microbiome in preterm birth and other pregnancy complications. This study highlights the importance of larger studies comparing

subjects with different ethnical backgrounds that also analyze specific pregnancy complications such as preterm labor.

## Supporting Information

**Figure S1 PCA showing the clustering by sequencing platform.**

(TIF)

**Figure S2 Venn diagram of samples analyzed.**

(TIF)

**Figure S3 Chao diversity index in non pregnant and pregnant subjects (results confounded by the platform effect in pregnant subjects - see Table S1).** No significance was found between the two groups of subjects ( $p > 0.05$ , Monte Carlo analysis, 999 permutations).

(TIF)

**Table S1**

(PDF)

**Table S2 Supplemental OTU Table.**

(XLSX)

**Metadata S1**

(CSV)

**Methods S1**

(DOCX)

## Acknowledgments

The authors wish to acknowledge clinical, diagnostic and clerical staff at Mayo Clinic with special thanks to Maureen Lemens for procurement of samples.

## Author Contributions

Conceived and designed the experiments: B.White DC. Performed the experiments: MWA MM. Analyzed the data: MWA PJ NC. Contributed reagents/materials/analysis tools: PJ NC B.White CY KN B.Wilson. Wrote the paper: MWA NC DC.

## References

- Human Microbiome Project Consortium (2012) Structure, function and diversity of the healthy human microbiome. *Nature* 486: 207–14.
- White BA, Creedon DJ, Nelson KE, Wilson BA (2011) The vaginal microbiome in health and disease. *Trends Endocrinol Metab* 22: 389–93.
- Yeoman CJ, Thomas SM, Miller ME, Ulanov AV, Torralba M, et al. (2013) A multi-omic systems-based approach reveals metabolic markers of bacterial vaginosis and insight into the disease. *PLoS One* 8: e56111.
- Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, et al. (1985) Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc Natl Acad Sci U S A* 82: 6955–9.
- Gajer P, Brotman RM, Bai G, Sakamoto J, Schutte UM, et al. (2012) Temporal dynamics of the human vaginal microbiota. *Sci Transl Med* 4: 132ra52.
- Zhou X, Bent SJ, Schneider MG, Davis CC, Islam MR, et al. (2004) Characterization of vaginal microbial communities in adult healthy women using cultivation-independent methods. *Microbiology* 150: 2565–73.
- 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: 4898–909.
- Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, et al. (2011) Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* 108: 4680–7.
- Relman DA (2012) The human microbiome: ecosystem resilience and health. *Nutr Rev* 70: S2–9.
- McDonald HM, O'Loughlin JA, Jolley P, Vigneswaran R, McDonald PJ (1991) Vaginal infection and preterm labour. *Br J Obstet Gynaecol* 98: 427–35.
- Platz-Christensen JJ, Pernevi P, Hagmar B, Andersson E, Brandberg A, et al. (1993) longitudinal follow-up of bacterial vaginosis during pregnancy. *Acta Obstet Gynecol Scand* 72: 99–102.
- Hay PE, Lamont RF, Taylor-Robinson D, Morgan DJ, Ison C, et al. (1994) Abnormal bacterial colonisation of the genital tract and subsequent preterm delivery and late miscarriage. *BMJ* 308: 295–8.
- Hauth JC, Goldenberg RL, Andrews WW, DuBard MB, Copper RL (1995) Reduced incidence of preterm delivery with metronidazole and erythromycin in women with bacterial vaginosis. *N Engl J Med* 333: 1732–6.
- Oakeshott P, Hay P, Hay S, Steinke F, Rink E, et al. (2002) Association between bacterial vaginosis or chlamydial infection and miscarriage before 16 weeks' gestation: prospective community based cohort study. *BMJ* 325: 1334.
- Leitch H, Bodner-Adler B, Brunbauer M, Kaider A, Egarter C, et al. (2003) Bacterial vaginosis as a risk factor for preterm delivery: a meta-analysis. *Am J Obstet Gynecol* 189: 139–47.
- Aagaard K, Riehle K, Ma J, Segata N, Mistretta TA, et al. (2012) A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. *PLoS One* 7: e36466.
- Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, et al. (2014) The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. *Microbiome* 2: 4.
- Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, et al. (2009) Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 42: 377–81.
- Sipos M, Jeraldo P, Chia N, Qu A, Dhillon AS, et al. (2010) Robust computational analysis of rRNA hypervariable tag datasets. *PLoS One* 5: e15220.
- Ye Y (2011) Identification and Quantification of Abundant Species from Pyrosequences of 16S rRNA by Consensus Alignment. *Proceedings (IEEE Int Conf Bioinformatics Biomed)* 2010: 153–7.
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7: 335–6.
- Magurran AE (2004) *Measuring Biological Diversity*. Oxford, UK: Blackwell Science Ltd.

23. Zhou X, Brown CJ, Abdo Z, Davis CC, Hansmann MA, et al. (2007). Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women. *The ISME journal* 1: 121–133.
24. Goldenberg RL, Culhane JF, Iams JD, Romero R (2008). Epidemiology and causes of preterm birth. *The lancet* 371: 75–84.
25. Fredricks DN, Fiedler TL, Mrazek JM (2005) Molecular identification of bacteria associated with bacterial vaginosis. *N Engl J Med* 353: 1899–911.
26. Hay PE, Ugwumadu A, Chowns J (1997) Sex, thrush and bacterial vaginosis. *Int J STD AIDS* 8: 603–8.
27. Huppertz B, Kingdom J, Caniggia I, Desoye G, Black S, et al. (2003) Hypoxia favours necrotic versus apoptotic shedding of placental syncytiotrophoblast into the maternal circulation. *Placenta* 24: 181–90.
28. Red-Horse K, Zhou Y, Genbacev O, Prakobphol A, Foulk R, et al. (2004) Trophoblast differentiation during embryo implantation and formation of the maternal-fetal interface. *J Clin Invest* 114: 744–54.
29. Redman CW, Sargent IL (2005) Latest advances in understanding preeclampsia. *Science* 308: 1592–4.
30. Saito S, Shiozaki A, Nakashima A, Sakai M, Sasaki Y (2007) The role of the immune system in preeclampsia. *Mol Aspects Med* 28: 192–209.
31. McGregor JA, French JI (2000) Bacterial vaginosis in pregnancy. *Obstet Gynecol Surv* 55: S1–19.
32. Jakobsson T, Forsum U (2007) *Lactobacillus iners*: a marker of changes in the vaginal flora? *J Clin Microbiol* 45: 3145–3145.
33. Wilks M, Wiggins R, Whiley A, Hennessy E, Warwick S, et al. (2004) Identification and H<sub>2</sub>O<sub>2</sub> production of vaginal lactobacilli from pregnant women at high risk of preterm birth and relation with outcome. *J Clin Microbiol* 42: 713–717.
34. Donders GG, Van Calsteren K, Bellen G, Reybrouck R, Van den Bosch T, et al. (2009) Predictive value for preterm birth of abnormal vaginal flora, bacterial vaginosis and aerobic vaginitis during the first trimester of pregnancy. *BJOG* 116: 1315–24.
35. Jumpstart Consortium Human Microbiome Project Data Generation Working Group (2012). Evaluation of 16S rRNA-based community profiling for human microbiome research. *PLoS One* 7: e39315.