Elucidating mechanisms that drive changes in the composition of the human vaginal microbiome

A Dissertation
Presented in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
with a
Major in Bioinformatics and Computational Biology
in the
College of Graduate Studies
University of Idaho
by
Kenetta Leigh Nunn

Major Professor: Larry J. Forney, Ph.D.
Committee Members: James A. Foster, Ph.D.; Tanya Miura, Ph.D.; Christine Parent, Ph.D.;
Chris Remien, Ph.D.
Department Administrator: David C. Tank, Ph.D.

May 2020

AUTHORIZATION TO SUBMIT DISSERTATION

This dissertation of Kenetta Leigh Nunn submitted for the degree of Doctor of Philosophy with a Major in Bioinformatics and Computational Biology and titled "Elucidating mechanisms that drive changes in the composition of the human vaginal microbiome" has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor: ______________________________ Date: __________
Larry J. Forney, Ph.D.

Committee Members: ______________________________ Date: __________
James A. Foster, Ph.D.

______________________________ Date: __________
Tanya Miura, Ph.D.

______________________________ Date: __________
Christine Parent, Ph.D.

______________________________ Date: __________
Chris Remien, Ph.D.

Department Administrator: ______________________________ Date: __________
David C. Tank, Ph.D.

ABSTRACT

The influence of the vaginal microbiome on women's relative risk to disease and reproductive complications has been well documented. In the last decade, extensive cross-sectional and longitudinal studies have shown that many types of vaginal communities are found in healthy women, and that the composition of these communities can change over a matter of days. Currently, we have a poor understanding of the factors that cause compositional differences observed between women and temporal changes within women. Thus, the goal of my work is to fill this gap in knowledge. In the opening chapter, I briefly review what is known about the composition of vaginal communities in reproductive age women. Due to the positive association with *Lactobacillus* in the vagina and vaginal health, I narrowed the focus on species within this genus. Additionally, I discuss two host factors, namely estrogen and glycogen that have been associated with changes in the abundances of lactobacilli over a woman's lifespan.

Next, I characterized the composition of vaginal bacterial communities in a cohort of black adolescent women based on differences in the total abundances of *Lactobacillus*. Then I explore the relationships between estrogen, glycogen, stress, and vaginal community composition. I confirmed previous findings that glycogen levels are associated with high abundances of lactobacilli in vaginal communities. I also report that estradiol levels do not correlate with vaginal glycogen measurements, nor vaginal community composition. This finding is noteworthy as it suggests the association between estrogen, glycogen, and vaginal lactobacilli is more complex than previously thought. This study addresses potential developmental factors that influence the composition of vaginal communities and provides new insights into the kinds of communities that are present among black adolescent women.

In my third chapter, I differentiated putative amylases in the human vagina, which likely break down glycogen into simpler sugars that are fermented by vaginal lactobacilli. For this work, I designed a pilot study to collect vaginal samples from 23 reproductive age women. Using metagenomics and proteomics, I identified putative amylases expressed in vaginal fluids that mapped to the metagenomes of vaginal bacteria. I report novel findings that there are multiple amylases in the human vagina that are produced by both the host and bacteria within the community.

Finally, I explore the transition from pregnancy to postpartum in a cohort of 48 women. I document findings that are consistent with previous studies that showed the composition of vaginal communities changes from pregnancy to postpartum and is marked by reduced abundances of *Lactobacillus* and increased alpha diversity. I report new findings of host-associated compounds that are significantly associated with the changes observed during postpartum.

ACKNOWLEDGEMENTS

Wow! The process of obtaining my PhD has been an amazing journey and there are so many people that have played an integral role in me getting to this point. Prior to graduate school, I had gained a lot of experience working at the bench and performing experiments. I pursued a PhD so that I could develop skills to think critically in order to contribute to what science was being done in addition to how it was being done. I truly got what I came for. For this, I thank my advisor Dr. Larry J. Forney. I am grateful for all of all of his encouragement and support over the past five and a half years. Larry knew exactly when to push, how much to push, but also knew when to let up and give me the space to grow on my own as a scientist. I also enjoyed working together on manuscripts because his feedback over the years made me a much better writer. Thank you for trusting me to represent this work.

I also must thank my committee: Dr. James Foster, Dr. Tanya Miura, Dr. Christine Parent, and Dr. Chris Remien. In addition to Larry, these four individuals greatly contributed to helping improve my work by asking insightful questions and providing helpful feedback through yearly and one-on-one meetings. I so much enjoyed having coffee with Christine at One World discussing concepts in community ecology, philosophical discussions and Ale House brewery takeovers with James, going through notebooks and experimental protocols with Tanya, and talking about statistical models with Chris. Thank you all for your time, patience, and interest in my work.

I am a firm believer in choosing mentors wherever we need them. I have definitely acquired many more here at the University of Idaho. Dr. Ben Ridenhour has been a sounding board to double check my use of math models to analyze data, provide feedback on other analyses that I've done, and has contributed to making presentations and job talks better. Dr. Celeste Brown has been a wonderful mentor to discuss science, women in science, and life with. There were many days spent in her office chair that I will truly miss. Dr. Jodie Nicotra, who took me on as a teaching assistant, was always so inspiring and just a great person to have in my corner. Dr. Lee Fortunato was always available to talk about new experiments, life, and to help me stress a little less. To Dr. Dorah Mtui, Dr. Onesmo Balemba and Jeremiah, I am so grateful to them for opening their home and for making Hunter and I feel like family while in Idaho.

Thank you to all of my lab mates, Dr. Roxana Hickey, Dr. Michael France, Dr. Karol Gliniewicz, and Maria Schneider. Maria has been so much more than a lab manager, she has become a friend/adopted family, and has been a tremendous support for me here in Idaho. Thank you to all of the students that I have worked with: Biá Freire Costa, Genesis Sanchez-Lopez, Zoila Álvarez, Kristen Frafjord, and Micah Woodard. Special thanks to everyone in the Genomics Resources Core (Dr. Sam

Hunter, Dan New, Matt Flagnan, Dr. Kim Andrews), Dr. Benji Oswald of the Computational Resources Core, and everyone at IBEST. Your support has made all of the research in this dissertation possible.

Over the years, I've been fortunate to collaborate with many people. I thank Dr. Dennis Fortenberry, Dr. Emily Chester, Dr. Virginia Vitzthum, and Dr. Ben Ridenhour for contributing to the project presented in Chapter 2. Thank you to Dr. Geremy Clair, Dr. Joshua Atkins, Kristin Engbrecht, and Thomas Fillmore for contributing to the amylase study in Chapter 3. In addition, thank you to Dr. Celeste Brown for discussing methods to detect amylase activity, Jennifer Walker for helping me through the University's IRB process, Russ McClanahan for providing me with space to work on this study, and Cody Muir for helping me to design a flyer to advertise for the study. Thank you to Dr. Steve Witkin, Dr. Dimitrios Nasioudis, Dr. Evelyn Minis, Dr. Karol Gliniewicz, and Maria Schneider for contributing to the pregnancy study in Chapter 4. I would like to especially thank Dennis and Steve for trusting me to take the lead on the black adolescent study and the pregnancy study, respectively. Most of all, thank you to all of the study participants that made Chapters 2 – 4 possible.

In addition to collaborations on projects, I've had many people who have provided me with lab/office space to use, ensured that I had things that I needed as a student (funding, supplies, room reservations, food, feedback on presentations). Thank you to Dr. Jim Nagler, Gina Tingley, Kristi Accola, Dr. Martina Ederer, Dr. Trish Hartzell, Dr. Onesmo Balemba, Dr. Luke Harmon, Dr. JT Van Leuven, Ann Norton, Audrey Harris, Dr. Holly Wickman, and Luann Scott. I thank Dr. Dave Tank and Dr. Eva Top who both served as one of my BCB directors, in addition to the BCB board. For funding and support, I thank Epscor for the President's Doctoral Scholars award, IBEST, BCB, BEACON, NIH, and Glen E. Nielsen and Jean K for the Science Scholarship at the University of Idaho.

Moving to Moscow, Idaho as a woman of color from Durham, North Carolina was quite the challenge. This process was made more enjoyable by the many students, postdocs, and friends that I've met/acquired while being here. Thank you to all of the Biology and BCB students, and postdocs, members of my roller derby team (Palouse River Rollers), and the people of St. Mark's Family. Thank you for helping me to explore the Pacific and Inland Northwest. I must also thank Lisha Abendroth who has been the MOST amazing program manager and friend that anyone could ever ask for. Lisha played a huge role in me wanting to stay in Idaho and not flee for NC. Thank you for being so dope!

Most of all, thank you to my family (blood and adopted), friends, and my husband whose encouragement keep me strong.

DEDICATION

To my family (Hayes, Evans, Nunn, Morgan, Reid, Eaton, Reid-Eaton) for supporting and believing in me, and for encouraging me to push past the limits that society naturally poses on people of color. To my mother, Deborah L. Hayes for teaching me work ethic and how to be a strong, resilient yet kind person, for always lending a safe space, and for reminding me to be myself. To my father, Kennett L. Nunn for helping me to stay focused, constantly telling me to — "remember your goals K. Nunn."

Most of all to my husband, Hunter Reid-Eaton, for being my biggest cheerleader and friend, my rock, for making me laugh, and for trusting me to direct our journey. Dear, I know this has not been easy and the uncertainty at times was unsettling. I truly appreciate your support and understanding more than you will ever know.

TABLE OF CONTENTS

Authorization to Submit Dissertation ii

Abstract iii

Acknowledgements iv

Dedication vi

Table of Contents vii

List of Tables ix

List of Figures x

Statement of Contribution xiv

Chapter 1: Unraveling the dynamics of the human vaginal microbiome 1

- Abstract 1
- Introduction 1
- Vaginal community composition varies over time 4
- Drivers of community composition and change 4
- Glycogen *per se* is not a key nutrient for vaginal lactobacilli 6
- Concluding remarks 7
- Literature Cited 10

Chapter 2: Vaginal glycogen, not estradiol, is associated with vaginal bacterial community composition in black adolescent women 14

- Abstract 14
- Introduction 15
- Methods and Materials 16
- Results 19
- Discussion 27
- Literature Cited 30

Chapter 3: Amylases in the human vagina 34

- Abstract 34

Introduction.. 34
Methods and Materials.. 36
Results.. 39
Discussion.. 48
Literature Cited.. 51
Chapter 4: The postpartum vaginal microbiome.. 56
Abstract.. 56
Introduction.. 56
Methods and Materials.. 58
Results.. 60
Discussion.. 68
Literature Cited.. 72
Chapter 5: Conclusion and future directions.. 76
Literature Cited.. 83
Appendix A - SUPPLEMENTAL MATERIALS FOR CHAPTER 1.. 85
Appendix B - SUPPLEMENTAL MATERIALS FOR CHAPTER 2.. 86
Appendix C - SUPPLEMENTAL MATERIALS FOR CHAPTER 3.. 100
Appendix D - SUPPLEMENTAL MATERIALS FOR CHAPTER 4.. 107

LIST OF TABLES

Table 2-1. Stress, anxiety, and depression in community groups A, B, and C. 25

Table 3-1 Putative amylase proteins annotated in the metagenomic assembled genomes of bacteria in human vaginal samples. 45

Table 4-1. Study participant demographics 61

Table 4-2. Mean relative proportions for the most abundant taxa in each stage. The following 10 taxa below were identified by taking the top ten ranked taxa for each stage, collapsing them together, and keeping only those greater than 1% in all samples. 66

Table 4-3. Transitions between dominant bacteria among 32 women with all three time points. 67

LIST OF FIGURES

Figure 1.1 Composition and structure of vaginal bacterial communities found in 396 reproductive age women. The bacterial populations in each sample were classified based on partial 16S rRNA genes sequences and the communities were clustered based on the relative abundances of these bacterial populations. Major groups of communities were used to define community state types I to V. This heatmap shows the relative abundance (see color key) of bacterial taxa (listed on the left) in each community. [Reprinted from (3) with the permission of the Proceedings of the National Academy of Sciences USA] 3

Figure 1.2 Temporal dynamics of vaginal bacterial communities. Heat maps (top) and interpolated bar plots (bottom) depict changes in the relative abundances of bacterial taxa in the vaginal communities of four women over 16 weeks. The color code for each taxon is shown below the figure. [Reprinted from (8) with the permission of *Science Translational Medicine.]* 5

Figure 1.3 Schemes for the production and use of glycogen-derived resources by vaginal bacterial communities. Panel A depicts our current understanding of how glycogen-derived resources are produced and used by vaginal bacterial communities. Host estrogen stimulates the production and accumulation of glycogen, which is degraded by human α-amylase to produce simpler sugars that are consumed by vaginal lactobacilli and fermented to produce lactic acid. In Panel B we propose that bacterial populations in vaginal communities could also be sources of α-amylase in addition to (or instead of) the host. The resulting simpler sugars could serve as a ‘common good’ that is available to the entire bacterial community, thus setting the stage for interspecies competition for these resources. Both panels illustrate ecological networks that include species of Lactobacillus, various other bacterial populations and the host. 9

Figure 2.1 Heatmap based on the relative proportions of the 20 most abundant taxa in vaginal communities of black adolescent women. The columns of the heatmap include 146 samples collected from 21 young women over a 6-month period. The corresponding dendrogram represents complete-linkage hierarchical clustering of samples based on alt-Gower distances. The colored bar immediately below the dendrogram indicates which clusters were combined to form three groups (A, B, and C). 20

Figure 2.2 Examples of changes in community composition over time. The stacked bar charts represent proportions of bacteria in each community over 7 monthly visits where zero was the baseline visit. Colors for each taxon are shown in the legend below the figure. The community group for each visit is highlighted by a colored diamond below each chart and the corresponding legend is listed in the bottom right. The profiles of these subjects were chosen to illustrate the temporal variability in community composition within subjects. 22

Figure 2.3 Differences in estradiol, glycogen, Nugent score, and vaginal pH measurements among groups. The boxplots represent log-transformed estradiol (panel A), log-transformed glycogen (panel B), Nugent scores (panel C), and log-transformed vaginal pH (panel D) for samples in groups A, B, and C. Statistical significance (* $p < 0.05$, *** $p < 0.001$) is indicated above the bars. 23

Figure 2.4 The relationship between log-transformed estradiol and glycogen concentrations in black adolescent women sampled longitudinally. The amounts of salivary estradiol in each sample were log-transformed and modeled over corresponding glycogen measurements including subject as a random effect. The resulting linear model equation and p-value (shown in the upper right corner) were obtained by computing an analysis of variance (ANOVA) on the linear model. Each dot represents the log estrogen concentration and the corresponding glycogen value for a given sample. Dots are colored according to community group as shown in the legend to the right of the graph................................ 26

Figure 3.1 Scatter plot of glycogen measurements in mg/ml as a function of amylase activity measured in U/mg of total protein. The points in the graph represent glycogen measurements for each sample, plotted with the corresponding amylase activity for the sample. The regression line was generated by fitting a generalized linear model to these data points. The dashed lines outside of the regression line represent the 95% confidence interval for the model fit. The pseudo-R^2 and the P value for the model are shown in the upper right-hand corner of the figure.. 40

Figure 3.2 Amylase activity relative to bacterial species composition in the human vagina. (A) Amylase activity expressed in enzyme units (U) per milligram (mg) of total protein in the sample is arranged in increasing order with the corresponding stacked bar (B) below it. The dashed grey line in panel A represents the mean of amylase activity across samples (0.55 U/mg). The stacked bars represent the relative proportions of bacteria in each community. Colors for each taxon are shown in the legend below the figure. 42

Figure 3.3 Amylase enzymes in vaginal fluids collected from donors F02, F06, and F08 separated by native PAGE. These columns represent image slices taken from Figure C.1 (containing all of the images of native PAGE gels completed for all samples) for donor samples F02, F06, and F08. A total of 20 µg of total protein was loaded for each sample. After the amylase proteins were separated in the gels, gels were incubated in 1% starch to enable starch hydrolysis and stained with iodine solution to observe where hydrolysis took place. The images that the slices were extracted from were inverted so that the dark background shows up white and the clear zones indicating hydrolysis of starch show up dark. Thus, the dark bands reflect amylase isozymes that were resolved in each donor sample. PA (human pancreatic α-amylase) and GA (Rhizopus spp. glucoamylase) are two commercial amylases that were used as positive controls for the assay. 43

Figure 3.4 Amylase proteins expressed in vaginal fluids. Four putative amylases from bacteria in addition to human alpha-amylase were detected in vaginal fluids using LC-MS/MS. The relative abundance of the protein (listed on the far left) is shown in individual heatmaps below each sample. Relative abundances are indicated by the z-scores shown in the legend below the plot. Black squares indicate that the protein was not detected in a sample. The set of gray columns to the right of the heatmaps show where the sequence of the protein originated and is indicated by an orange square with a plus sign in the middle. The set of dark gray columns to the far-right show for which organism within the vaginal community the protein sequence was found in. To the far right of the figure lists the protein annotations from the assembled metagenomes of bacteria. Replicates were performed based on sample volume. 47

Figure 4.1. Relative proportions of bacteria in the vaginal communities of 48 pregnant women separated by stage and cluster. The stacked bars represent the proportions of bacterial taxa within one sample. Bars are separated by the pregnancy stage in which the sample was collected (right headings), and the cluster a sample was grouped into (top headings). Clusters were determined using complete-linkage hierarchical clustering of alt Gower distances calculated from the bacterial abundance data. Taxa colors are indicated in the legend below the figure. "Other" represents the sum of all bacterial taxa that were not present at 1% in at least two women. 63

Figure 4.2. Rank abundance and α-diversity profiles of samples collected during the first trimester, third trimester, and postpartum. Panel A shows the counts of the top ranked taxa among the first trimester, third trimester, and postpartum samples. The skew for each set of data was calculated using the R package e1071 (v. 1.7.2) and is indicated within the plot. Skewness represents the degree to which a sample distribution varies from a normal distribution. Panel B shows the distribution of Shannon and Simpson diversity measures for each stage. The thick band within the box represents the median for the data and the lower and upper boxes represent the 25th and 75th quartiles, respectively. The upper and lower whiskers represent the 95% confidence interval for the median. Individual dots above represent outliers that were not removed. Differences in group means were evaluated using hypothesis testing of linear mixed effects models in R. Statistical significance is indicated above the lines ("***" = $p < 0.001$). 64

Figure 4.3. Differences in *Lactobacillus* species by stage. The dot plot shows the distribution of *Lactobacillus* relative proportions for the first trimester, third trimester, and postpartum. Each dot is one sample from a woman. Differences in group means were evaluated using hypothesis testing of linear mixed effects models in R. Statistical significance is indicated above the lines ("***" = $p < 0.001$). 65

Figure 4.4. Differences in the means of measured compounds by stage. The boxplots in this figure represent the distributions of various compounds measured grouped by stage. The thick band within the box represents the median for the data and the lower and upper boxes represent the 25th and 75th quartiles, respectively. The upper and lower whiskers represent the 95% confidence interval for the median. Individual dots above represent outliers that were not removed. Differences in group means were evaluated using hypothesis testing of linear mixed effects models or nonparametric T-test where appropriate in R. Statistical significance is indicated above the lines as follows: "*", $p < 0.05$; "**", $p < 0.01$; "***" = $p < 0.001$. 69

Figure 5.1. Schematic for the relationship between estrogen, glycogen, and *Lactobacillus* in the human vagina. 76

Figure 5.2. *Lactobacillus crispatus* cultured in (A) MRS media and (B) NYC-III media with out sugar (red), with glucose (blue), or with glycogen (grey). Four replicates were performed for each. 79

Figure 5.3. *Lactobacillus crispatus* cultured in MRS without sugar (red), without sugar plus horse serum (orange), with glycogen (grey), with glycogen plus horse serum (green). Four replicates were performed for each. 80

Figure 5.4. Schematic for VK2 cell and *L. crispatus* co-culture, RNA-seq experiment. 81

STATEMENT OF CONTRIBUTION

With this statement, I confirm that the research in this dissertation is my own and that I am the primary author for Chapters 1 – 4. My contributions in addition to all co-author contributions are listed below.

Author list for Chapter 1: Kenetta L Nunn (KLN) and Larry J Forney (LJF)
Author Contributions for Chapter 1: KLN and LJF wrote the manuscript.

Author list for Chapter 2: Kenetta L Nunn (KLN), Benjamin J Ridenhour (BJR), Emily M Chester (EMC), Virginia J Vitzthum (VJV), J Dennis Fortenberry (JDF), and Larry J Forney (LJF)
Author Contributions for Chapter 2: JDF and LJF designed the research and provided feedback on data analysis and interpretation. EMC and VJV both helped with data collection and analysis. BJH helped with statistical data analysis and interpretation. **KLN** performed data collection, data analysis and interpretation, and wrote the manuscript. All authors contributed to editing the manuscript prior to submission.

Author list for Chapter 3: Kenetta L Nunn (KLN), Geremy C Clair (GCC), Joshua Adkins (JA), Kristin Engbrecht (KE), Thomas Fillmore (TF), and Larry J Forney (LJF)
Author Contributions for Chapter 3: KLN conceptualized and designed the study with feedback from LJF. JA and GCC helped with experimental protocol design, and KE and TF helped with data collection for the proteomics work. GCC performed data analysis for and oversaw the proteomics work. **KLN** performed data collection, data analysis and interpretation, and wrote the manuscript with feedback from LJF. All authors will have the opportunity to contribute to editing the manuscript prior to submission.

Author list for Chapter 4: Kenetta L Nunn (KLN), Steven S. Witkin (SSW), G Maria Schneider (GMS), Allison Boester (AB), Dimitrios Nasioudis (DN), Evelyn Minis (EM), Karol Gliniewicz (KG), and Larry J Forney (LJF)
Author Contributions for Chapter 4: SSW conceptualized and designed the study with feedback from LJF. GMS, AB, DN, EM, and KG helped with experimental protocol design and data collection. **KLN** performed data collection, data analysis and interpretation, and wrote the manuscript with feedback from LJF. All authors will have the opportunity to contribute to editing the manuscript prior to submission.

CHAPTER 1: UNRAVELING THE DYNAMICS OF THE HUMAN VAGINAL MICROBIOME[1]

ABSTRACT

Four *Lactobacillus* species, namely *L. crispatus*, *L. iners*, *L. gasseri*, and *L. jensenii*, commonly dominate the vaginal communities of most reproductive-age women. It is unclear why these particular species, and not others, are so prevalent. Historically, estrogen-induced glycogen production by the vaginal epithelium has been proffered as being key to supporting the proliferation of vaginal lactobacilli. However, the 'fly in the ointment' (that has been largely ignored) is that the species of *Lactobacillus* commonly found in the human vagina cannot directly metabolize glycogen. It would appear that this riddle has been solved as studies have demonstrated that vaginal lactobacilli can metabolize the products of glycogen depolymerization by α-amylase, and fortunately, amylase activity is found in vaginal secretions. These amylases are presumed to be host-derived, but we suggest that other bacterial populations in vaginal communities could also be sources of amylase in addition to (or instead of) the host. Here we briefly review what is known about human vaginal bacterial communities and discuss how glycogen-derived resources and resource competition might shape the composition and structure of these communities.

INTRODUCTION

Previous studies have shown that while species of *Lactobacillus* tend to dominate the vaginal communities of most healthy reproductive-aged women [1-5], these communities are quite personalized in terms of species composition and temporal dynamics [1-3,6-9]. These lactobacilli are thought to have a protective role by maintaining an acidic environment through the production of lactic acid thereby restricting the growth of pathogenic organisms and evidence to support this has been extensively reviewed in the literature [10-15]. In a study of black and white women done by Zhou *et al.,* samples from 144 women were analyzed and eight major types of vaginal communities were found [2]. Eighty percent of these women had communities dominated by various *Lactobacillus* species, with *L. iners* being the most prevalent and found in 66% of the women. The other dominant *Lactobacillus* species included *L. crispatus*, *L. gasseri*, and *L. jensenii*. These same four lactobacilli were dominant in four of the five major groups of bacterial communities found in a later study that involved a larger

[1] This chapter is a review paper that was previously published as: Nunn KL, Forney LJ. Unraveling the Dynamics of the Human Vaginal Microbiome. Yale J Biol Med. 2016; 89(3):331-337.

cohort of 396 reproductive-age women [3]. These groups, which are referred to as community state types (CSTs), were determined by clustering vaginal bacterial communities based on bacterial composition and relative abundance. In that study *L. crispatus* was dominant in CST I, *L. gasseri* was dominant in CST II, while *L. iners* and *L. jensenii* were dominant in CST III and CST V, respectively (Figure 1.1) [3]. However, not all communities were dominated by lactobacilli [1,3] as the fifth CST (CST IV) lacked high proportions of *Lactobacillus* and instead was characterized by an increased relative abundance of strict anaerobes that included *Gardnerella vaginalis*, *Prevotella*, *Atopobium*, *Megasphaera* and others [3].

Both of these studies revealed significant differences in vaginal community composition between women of different ethnic groups. For example, Ravel *et al*. showed that *Lactobacillus* species were dominant in 80.2% and 89.7% of Asian and white women, but this was the case in only 59.6% and 61.9% of black and Hispanic women. Similarly, Zhou *et al*. found *Lactobacillus* species to be dominant in 91% of White women but only 68% of black women. These differences in vaginal community composition between women of different ethnicities were confirmed in subsequent studies done to characterize the vaginal communities of African American women and women of European ancestry [2,16]. Likewise, a more recent study found four kinds of vaginal communities in a population of South African women [17]. Of these women, only 37% had *Lactobacillus* dominated communities with *L. iners* being the most common species found.

Given that there are more than 130 species of *Lactobacillus* known it is unclear why just four of these species dominate most vaginal communities. At the very least this observation suggests that *L. crispatus*, *L. gasseri*, *L. jensenii*, and *L. iners* are especially well adapted to the vaginal environment and have specific traits that allow them to colonize this habitat. When surveyed for niche-specific traits, comparative genomic analysis of 25 *Lactobacillus* species showed that vaginal lactobacilli had significantly smaller genomes and lower G+C content than non-vaginal species [18]. Efforts to identify a set of common traits that might account for their shared ability to successfully colonize the human vagina were unsuccessful. However, there were a number of species-specific traits identified, which hints at the possibility that each species has unique characteristics that enable them to effectively compete in the vaginal ecosystem. A more in-depth functional analysis of these traits will be necessary to define their relative importance in mediating interactions between lactobacilli and the host.

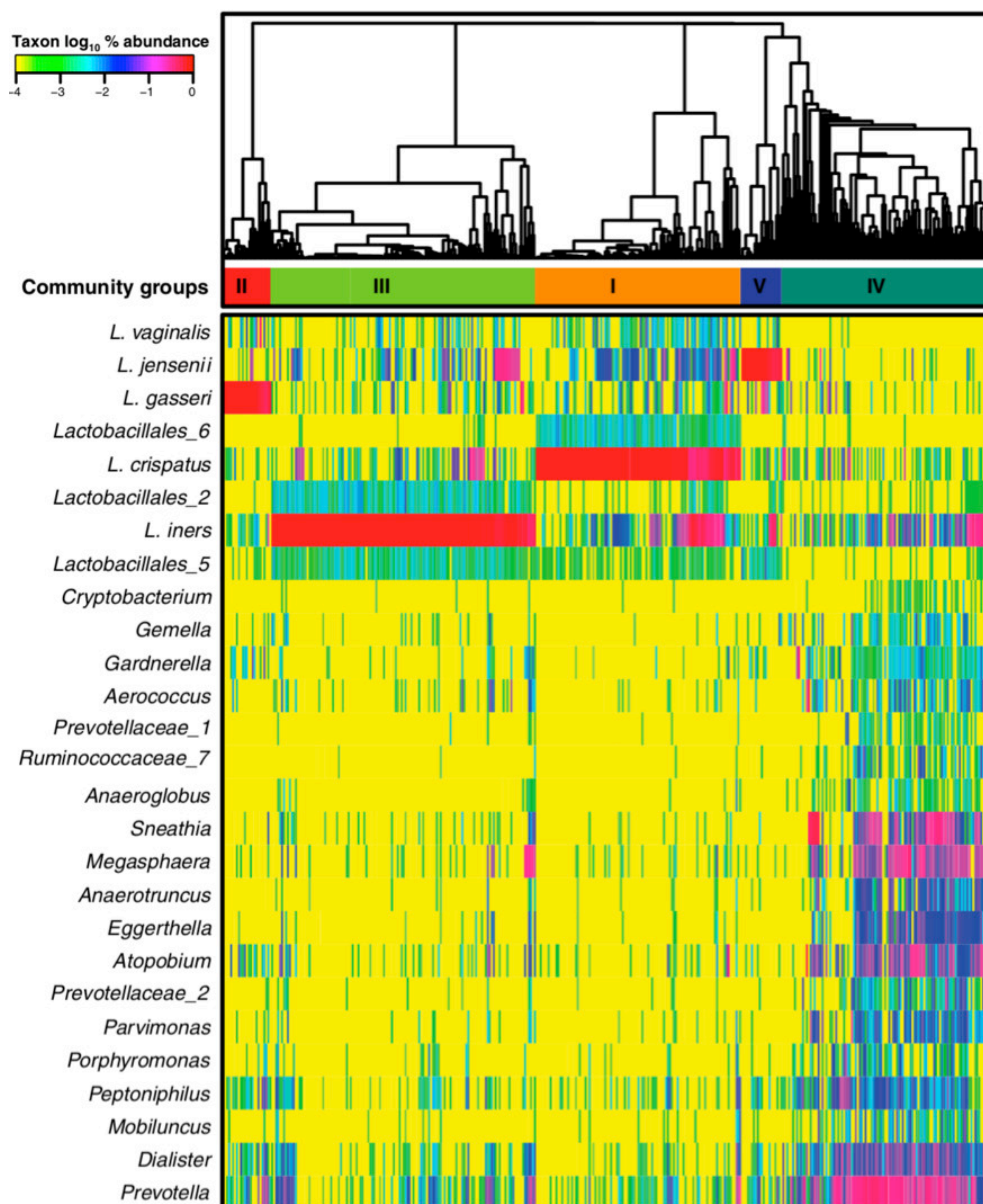


Figure 1.1 Composition and structure of vaginal bacterial communities found in 396 reproductive age women. The bacterial populations in each sample were classified based on partial 16S rRNA genes sequences and the communities were clustered based on the relative abundances of these bacterial populations. Major groups of communities were used to define community state types I to V. This heatmap shows the relative abundance (see color key) of bacterial taxa (listed on the left) in each community. [Reprinted from (3) with the permission of the Proceedings of the National Academy of Sciences USA]

The stark differences in vaginal communities within and between women, especially those of different ethnicities, suggest that the host may have a prominent role in shaping the species composition of these communities. Some key questions include: 1) why do only certain species of *Lactobacillus* colonize the vagina; 2) why does the composition of vaginal bacterial communities differ between women; and 3) what drives communities to change in composition over time? In this review we explore what is known about how the host environment and competition for resources may influence the composition and stability of these communities.

VAGINAL COMMUNITY COMPOSITION VARIES OVER TIME

Studies on the longitudinal dynamics of the vaginal microbiome in healthy reproductive-age women have shown that these communities often change over relatively short periods of time [8,9,19-21]. Gajer *et al*. analyzed samples that were self-collected twice weekly over 16 weeks from 32 women [8]. The community composition in each sample from each time point was classified into CSTs, and the patterns of CSTs over time were clustered, demonstrating five temporal patterns. Of the *Lactobacillus* dominated CSTs, only CST I-III were found; probably because too few women were included in the study. Some communities changed in composition while others remained relatively stable over time (Figure 1.2). Moreover, not all transitions between CSTs were equally likely to occur. This study showed that community dynamics vary greatly between individuals. Unfortunately, the mechanisms that drive these dynamics in community composition are not well understood.

DRIVERS OF COMMUNITY COMPOSITION AND CHANGE

All of the resources necessary to support vaginal bacterial communities are ultimately derived from the host. It follows that differences in host physiology that alter the kinds and abundances of resources present might influence which bacterial species will successfully colonize the vagina. Vaginal secretions and vaginal epithelial cells that are sloughed and subsequently lyse are thought to be the principle sources of nutrients that support bacterial growth. Vaginal secretions contain nutrients that include proteins, carbohydrates, amino acids, and cervicovaginal mucus, which is rich in mucins and glycoproteins [22,23]. The vaginal epithelium is rich in glycogen, which has long been thought to be a key nutrient for vaginal lactobacilli. However, this assertion is primarily based on the positive correlation that exists between the levels of free glycogen and the abundance of vaginal lactobacilli [24].

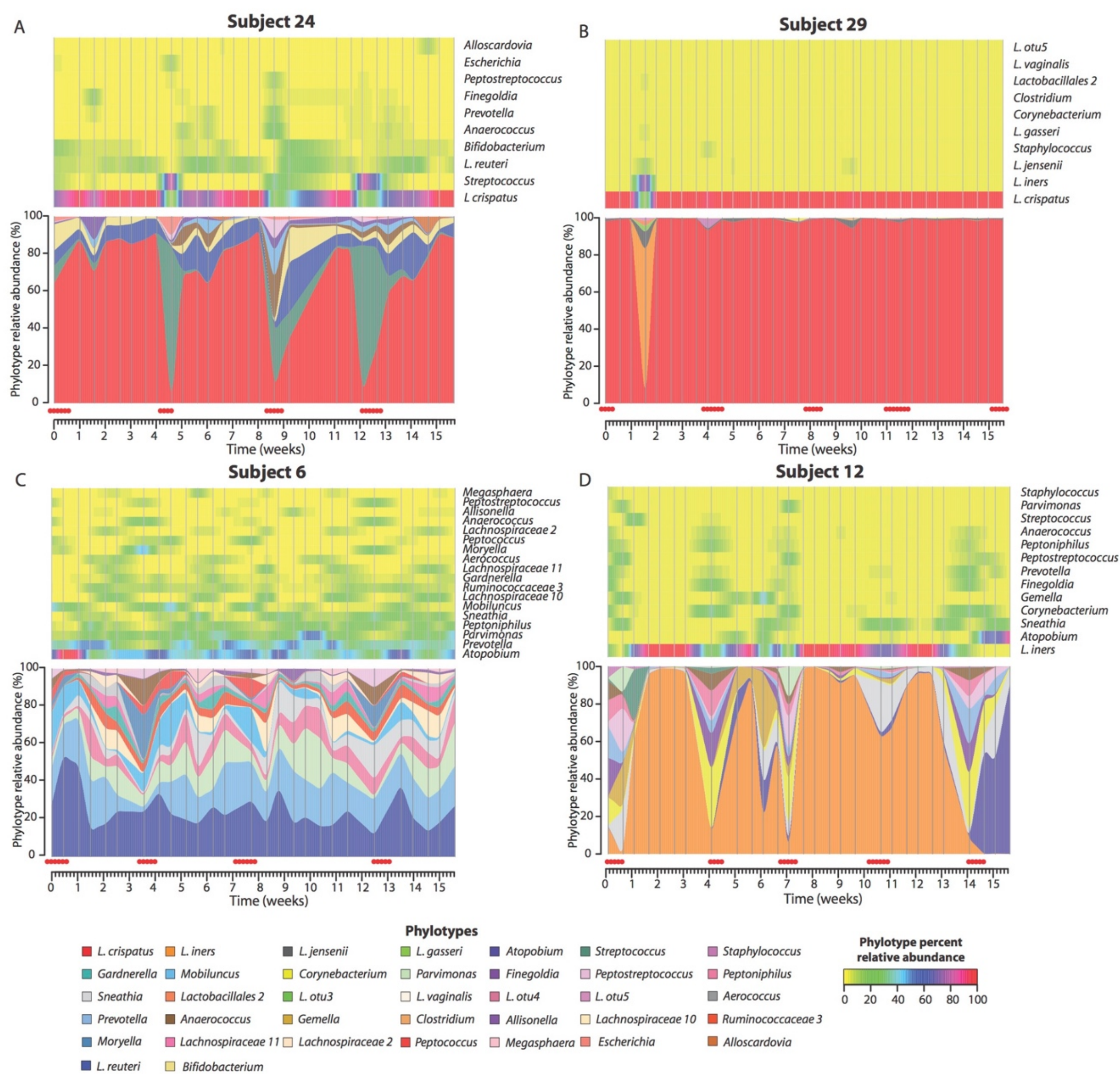


Figure 1.2 Temporal dynamics of vaginal bacterial communities. Heat maps (top) and interpolated bar plots (bottom) depict changes in the relative abundances of bacterial taxa in the vaginal communities of four women over 16 weeks. The color code for each taxon is shown below the figure. [Reprinted from (8) with the permission of *Science Translational Medicine*.]

The availability of resources in the vagina seems to be driven by estrogen levels. Estrogen is known to increase the volume of vaginal secretions. Moreover, elevation of estrogen levels induces thickening of the vaginal epithelium and prompts the accumulation of glycogen. Together these are thought to create an environment that stimulates the proliferation of *Lactobacillus* [25]. Furthermore, changes in the relative abundance of vaginal lactobacilli is also associated with estrogen levels and glycogen content [26,27] over a woman's life span. During reproductive years, estrogen levels are known to vary throughout the menstrual cycle. Estrogen levels are low during menses (< 50 pg/ml), peak before ovulation (200-250 pg/ml), decline shortly thereafter and peak again around day 21 (150 pg/ml) [8]. Changes in glycogen content have been shown to parallel these changes [28]. Variation in estrogen levels and glycogen content over the menstrual cycle might partly explain why there are differences in community composition within women over time.

In an effort to understand the effect of the menstrual cycle on changes in community composition, Gajer *et al*. modeled the rate of change in a community's state as a function of the menstrual cycle. Based on their findings, community composition appears to be more stable when estrogen levels are highest [8]. This observation is also seen during pregnancy, a condition where estrogen levels are high. Estrogen levels increase dramatically during pregnancy due to additional estrogen produced by the placenta [29]. Studies evaluating the vaginal microbiome during pregnancy have shown that the vaginal communities of pregnant women are more stable and have higher relative abundances of lactobacilli than those of non-pregnant women [30,31]. Moreover, vaginal communities in pregnant women tend to be dominated by *L. crispatus* or *L. iners*, while CST IV (in which *Lactobacillus* is not dominant) is rarely observed [31,32]. Estrogen levels decrease post-partum and concomitant changes in communities are observed, with lactobacilli becoming less dominant [31]. These studies suggest a strong correlation between estrogen and community composition.

GLYCOGEN *PER SE* IS NOT A KEY NUTRIENT FOR VAGINAL LACTOBACILLI

While it is casually suggested that glycogen directly enriches for *Lactobacillus*, it should be noted that vaginal lactobacilli cannot directly metabolize glycogen. In 1964, Stewart-Tull isolated 36 strains of *Lactobacillus* from the vagina and cervix of pregnant women, and showed that none of the strains could metabolize glycogen [33]. Similarly, years later, Whylie and Henderson isolated 42 strains of *Lactobacillus* from the vaginas of pregnant women [34], of which 11 were strains of *L. acidophilus*. One of these could metabolize glycogen from oysters and two could metabolize glycogen isolated from the human vagina. Additional evidence that few or no vaginal lactobacilli are capable of directly using

glycogen as a resource was obtained in a study by Martín *et al*. in which none of the *L. crispatus*, *L. gasseri*, and *L. jensenii* strains tested were able to metabolize glycogen [35]. Recently, Spear *et al*. confirmed these results with *L. gasseri* and *L. jensenii* [36].

While vaginal *Lactobacillus* species cannot metabolize glycogen, they have been shown to grow on smaller oligomers of glucose produced through the depolymerization of glycogen by α-amylase [32]. α-Amylase is an endoglycosidase that cleaves α-(1,4) glycosidic bonds [37] to produce maltose, maltotriose, as well as α-(1,4) and α-(1,6)-dextrins [38]. In humans, α-amylase is reportedly only found in saliva and the pancreas, but definitive evidence for its presence in vaginal secretions is lacking. Spear *et al*. showed that *L. gasseri* and *L. jensenii* could grow in media containing glycogen if saliva is first added [36]. The growth profiles in media containing glycogen and saliva were similar to those seen when these species were grown in media containing glucose alone, and likewise, both species were able to grow on maltose [36]. These investigators also showed that vaginal secretions exhibited amylase-like activity and could depolymerize glycogen. It is not clear whether the amylase (or multiple amylases) in vaginal secretions originated from the human host, from one or more bacterial populations in vaginal communities, or both.

With the premise that vaginal lactobacilli play a key role in the maintenance of health, and that α-amylase promotes the dominance of *Lactobacillus* species in the vagina, Nasioudis *et al*. conducted a study to evaluate α-amylase levels in women with bacterial vaginosis or vulvo-vaginal candidiasis [39]. Their results showed α-amylase levels were lowest in women with bacterial vaginosis and highest in healthy women without either vaginal disorder. This study did not analyze the composition of the vaginal communities; therefore, it is difficult to directly associate α-amylase levels with the dominance of *Lactobacillus* in the vaginal communities of these women. We can, however, suggest that high levels of α-amylase might be associated with healthy conditions in the vagina.

CONCLUDING REMARKS

The evidence available can be used to paint a simple picture (Figure 1.3A). Glycogen levels in the vagina are driven by estrogen in the host and this resource is depolymerized by human α-amylase to produce simpler sugars. These carbohydrates are readily fermented by species of vaginal lactobacilli to produce lactic acid, which lowers the pH of the vagina creating an environment that putatively restricts the growth of non-indigenous organisms. The dominance of *L. crispatus*, *L. gasseri*, *L. iners*, and *L. jensenii* in the vaginal communities of most women suggests that these species may be

particularly adept at competing for glycogen-derived resources. The key elements of this scenario – estrogen, glycogen and amylase – are all governed by the host and may vary over time due to physiological changes within an individual. Moreover, these elements might differ between individuals due to genetic or behavioral (e.g. dietary) differences between individuals and account for temporal variation in vaginal community composition.

Here we suggest a few wrinkles that might be considered to refine our understanding. One concerns the source of amylases that mediate the depolymerization of glycogen. In humans, α-amylase has only been shown to be present in saliva and the pancreas, and we await confirmation that it is produced in the vagina. Aside from human α-amylase, it is plausible that various bacterial populations in vaginal communities also produce α-amylase (Figure 1.3B) since various anaerobic taxa are known to produce amylases and ferment the breakdown products of glycogen [40-44]. In theory, multiple sources of amylases would increase the redundancy of this function in vaginal communities and increase the probability that this critical function is maintained in the face of perturbations, thus rendering the community more stable. Secondly, there is no reason *a priori* to think that only lactobacilli use the glycogen-derived simpler sugars (Figure 1.3B). The hydrolysis of glycogen by these amylases almost certainly occurs in the extracellular environment, which produces resources (e.g. maltose) that are 'common goods' available to all members of the bacterial community. These glycogen-derived resources might be broadly shared and support the growth of other community members. While any community member capable of utilizing these resources can compete for them, perhaps vaginal lactobacilli more effectively compete for and sequester these resources. Within this conceptual framework, both glycogen levels and the functional redundancy of amylase production might be key to the maintenance of relatively stable bacterial communities that are dominated by *Lactobacillus* species.

Combining the simple picture and the wrinkles that we describe above there are at least four possibilities that may account for the dominance of lactobacilli and overall composition of bacterial communities in the human vagina. First, the numerical dominance of lactobacilli suggests that they simply out-compete other taxa for glycogen-derived resources. A second alternative is that other taxa use glycogen and/or glycogen-derived resources more efficiently than lactobacilli, but they in turn produce key nutrients (e.g. vitamins, amino acids) that spur the proliferation vaginal lactobacilli. This would imply that the prevalence of vaginal lactobacilli could be largely based on interactions with and metabolic dependence on other community members rather than direct interactions with the host. Third, while emphasizing the positive correlation between the levels of glycogen and the abundance of vaginal lactobacilli there

is a tendency to gloss over the fact that lactobacilli cannot use glycogen directly. It could well be that high levels of glycogen develop precisely because vaginal lactobacilli cannot use it as a resource. Meanwhile, a high relative abundance of lactobacilli only allows for low proportions of other taxa that might metabolize glycogen. If rate of glycogen metabolism does not exceed the rate at which it is produced this might explain why we continue to see high levels of glycogen in the presence of lactobacilli. Finally, the correlation between estrogen, glycogen and the selection of specific species of *Lactobacillus* could be a 'red herring' if there is no causal relationship between these factors. This tangle of possibilities will need to be sorted out in future research. Meanwhile researchers should avoid over simplifying what is most likely a complex ecological network that includes both the host and various bacterial species in these communities.

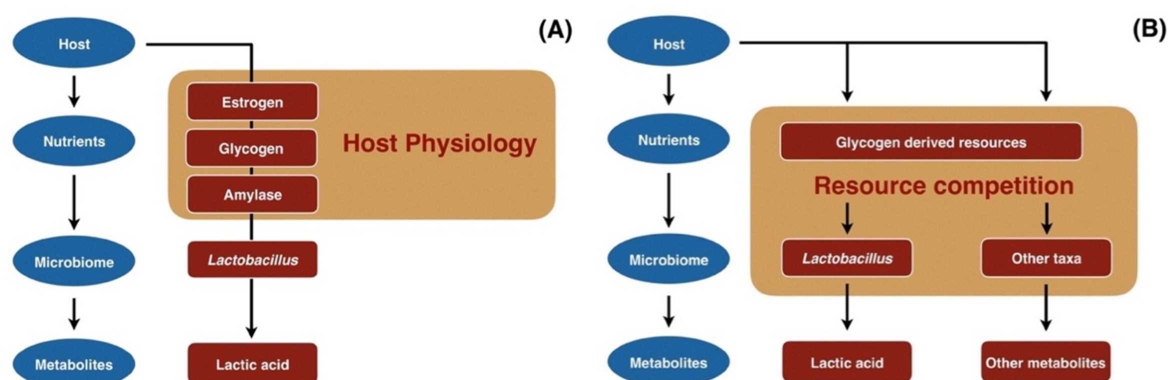


Figure 1.3 Schemes for the production and use of glycogen-derived resources by vaginal bacterial communities. Panel A depicts our current understanding of how glycogen-derived resources are produced and used by vaginal bacterial communities. Host estrogen stimulates the production and accumulation of glycogen, which is degraded by human α-amylase to produce simpler sugars that are consumed by vaginal lactobacilli and fermented to produce lactic acid. In Panel B we propose that bacterial populations in vaginal communities could also be sources of α-amylase in addition to (or instead of) the host. The resulting simpler sugars could serve as a 'common good' that is available to the entire bacterial community, thus setting the stage for interspecies competition for these resources. Both panels illustrate ecological networks that include species of *Lactobacillus*, various other bacterial populations and the host.

LITERATURE CITED

[1] Zhou X, Bent SJ, Schneider MG, Davis CC, Islam MR, Forney LJ. Characterization of vaginal microbial communities in adult healthy women using cultivation-independent methods. Microbiology. 2004;150(Pt 8):2565–2573.

[2] Zhou X, Brown CJ, Abdo Z, Davis CC, Hansmann MA, Joyce P, Foster JA, Forney LJ. Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women. ISME J. 2007 Jun;1(2):121–133.

[3] Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, *et al*. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci USA. 201;108 Suppl 1:4680–4687.

[4] Antonio MA, Hawes SE, Hillier SL. The identification of vaginal *Lactobacillus* species and the demographic and microbiologic characteristics of women colonized by these species. J Infec Dis. 1999;180(6):1950–1956.

[5] Burton JP, Cadieux PA, Reid G. Improved understanding of the bacterial vaginal microbiota of women before and after probiotic instillation. Appl Environ Microbiol. 2003;69(1):97–101.

[6] Hyman RW, Fukushima M, Diamond L, Kumm J, Giudice LC, Davis RW. Microbes on the human vaginal epithelium. Proc Natl Acad Sci USA. 2005;102(22):7952–7957.

[7] Zhou X, Hansmann MA, Davis CC, Suzuki H, Brown CJ, Schütte U, *et al*. The vaginal bacterial communities of Japanese women resemble those of women in other racial groups. FEMS Immunol Med Microbiol. 2010;58(2):169–181.

[8] Gajer P, Brotman RM, Bai G, Sakamoto J, Schutte UME, Zhong X, *et al*. Temporal dynamics of the human vaginal microbiota. Sci Transl Med. 2012;4(132):132ra52–132ra52.

[9] Brotman RM, Ravel J, Cone RA, Zenilman JM. Rapid fluctuation of the vaginal microbiota measured by Gram stain analysis. Sex Transm Infect. 2010;86(4):297–302.

[10] Ma B, Forney LJ, Ravel J. Vaginal microbiome: rethinking health and disease. Annu Rev Microbiol. 2012; 66(1):371–389.

[11] Hickey RJ, Zhou X, Pierson JD, Ravel J, Forney LJ. Understanding vaginal microbiome complexity from an ecological perspective. Transl Res. 2012;160(4):267–282.

[12] Larsen B, Monif GR. Understanding the bacterial flora of the female genital tract. Clin Infect Dis. 2001;32(4):e69–77.

[13] Huang B, Fettweis JM, Brooks JP, Jefferson KK, Buck GA. The changing landscape of the vaginal microbiome. Clin Lab Med. 2014;34(4):747–761.

[14] van de Wijgert JHHM, Borgdorff H, Verhelst R, Crucitti T, Francis S, Verstraelen H, *et al*. The vaginal microbiota: what have we learned after a decade of molecular characterization? PLoS ONE. 2014;9(8):e105998–10.

[15] Redondo-Lopez V, Cook RL, Sobel JD. Emerging role of lactobacilli in the control and maintenance of the vaginal bacterial microflora. Rev Infect Dis. 1990;12(5):856–872.

[16] Fettweis JM, Brooks JP, Serrano MG, Sheth NU, Girerd PH, Edwards DJ, *et al*. Differences in vaginal microbiome in African American women versus women of European ancestry. Microbiology. 2014;160(Pt_10):2272–2282.

[17] Anahtar MN, Byrne EH, Doherty KE, Bowman BA, Yamamoto HS, Soumillon M, *et al*. Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. Immunity. 2015;42(5):965–976.

[18] Mendes-Soares H, Suzuki H, Hickey RJ, Forney LJ. Comparative functional genomics of *Lactobacillus* spp. reveals possible mechanisms for specialization of vaginal lactobacilli to their environment. J Bacteriol. 2014;196(7):1458–1470.

[19] Srinivasan S, Liu C, Mitchell CM, Fiedler TL, Thomas KK, Agnew KJ, *et al*. Temporal variability of human vaginal bacteria and relationship with bacterial vaginosis. PLoS ONE. 2010;5(4):e10197–8.

[20] Keane FE, Ison CA, Taylor-Robinson D. A longitudinal study of the vaginal flora over a menstrual cycle. Int J STD AIDS. 2008;8(8):1–6.

[21] Schwebke JR, Richey CM, Weiss HL. Correlation of behaviors with microbiological changes in vaginal flora. J Infect Dis. 1999;180(5):1–5.

[22] Owen DH, Katz DF. A vaginal fluid simulant. Contraception. 1999;59(2):91–95.

[23] Geshnizgani AM, Onderdonk AB. Defined medium simulating genital tract secretions for growth of vaginal microflora. J Clin Microbiol. 1992;30(5):1323–1326.

[24] Mirmonsef P, Hotton AL, Gilbert D, Burgad D, Landay A, Weber KM, *et al*. Free glycogen in vaginal fluids is associated with *Lactobacillus* colonization and low vaginal pH. PLoS ONE; 2014;9(7):e102467.

[25] Farage M, Maibach H. Lifetime changes in the vulva and vagina. Arch Gynecol Obstet. 2006;273(4):195–202.

[26] Cruickshank R, Sharman A. The biology of the vagina in the human subject. BJOG. 1934;41(2):208–226.

[27] Paavonen J. Physiology and ecology of the vagina. Scand J Infect Dis Suppl. 1983;40:31–35.

[28] Sjöberg I, Cajander S, Rylander E. Morphometric characteristics of the vaginal epithelium during the menstrual cycle. Gynecol Obstet Invest. 1988;26(2):136–144.

[29] Jones RERE. Human reproduction and sexual behavior. Englewood Cliffs, N.J.: Prentice-Hall; 1984.

[30] MacIntyre DA, Chandiramani M, Lee YS, Kindinger L, Smith A, Angelopoulos N, *et al*. The vaginal microbiome during pregnancy and the postpartum period in a European population. Sci Rep. 2015;5:8988–9.

[31] Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Nikita L, *et al*. The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. Microbiome. 2014;2(1):1–19.

[32] Walther-António MRS, Jeraldo P, Berg Miller ME, Yeoman CJ, Nelson KE, Wilson BA, *et al*. Pregnancy's stronghold on the vaginal microbiome. PLoS ONE. 2014;9(6):e98514.

[33] Stewart-Tull DE. Evidence that vaginal lactobacilli do not ferment glycogen. Am J Obstet Gynecol. 1964; 88:676–9.

[34] Wylie JG, Henderson A. Identity and glycogen-fermenting ability of lactobacilli isolated from the vagina of pregnant women. J Med Microbiol. 1969;2(3):363–366.

[35] Martín R, Soberón N, Vaneechoutte M, Flórez AB, Vázquez F, Suárez JE. Characterization of indigenous vaginal lactobacilli from healthy women as probiotic candidates. Int Microbiol. 2008;11(4):261–266.

[36] Spear GT, French AL, Gilbert D, Zariffard MR, Mirmonsef P, Sullivan TH, *et al*. Human α-amylase present in lower-genital-tract mucosal fluid processes glycogen to support vaginal colonization by *Lactobacillus*. J Infect Dis. 2014;210(7):1019–28.

[37] Brayer GD, Luo Y, Withers SG. The structure of human pancreatic alpha-amylase at 1.8 A resolution and comparisons with related enzymes. Protein Sci. 1995;4(9):1730–1742.

[38] Roberts PJ, Whelan WJ. The mechanism of carbohydrase action. 5. Action of human salivary alpha-amylase on amylopectin and glycogen. Biochem J. 1960;76:246–53.

[39] Nasioudis D, Beghini J, Bongiovanni AM, Giraldo PC, Linhares IM, Witkin SS. α-Amylase in vaginal fluid: association with conditions favorable to dominance of *Lactobacillus*. Reprod Sci; 2015;22(11):1393–1398.

[40] Gupta R, Gigras P, Mohapatra H, Goswami VK, Chauhan B. Microbial α-amylases: a biotechnological perspective. Process Biochem. 2003;38(11):1599–1616.

[41] Pandey A, Nigam P, Soccol CR, Soccol VT, Singh D, Mohan R. Advances in microbial amylases. Biotechnol Appl Biochem. 2000;31(2):135–19.

[42] Reddy, NS, Nimmagadda A, Samasiva Rao KRS. An overview of the microbial α-amylase family. Afr J Biotechnol. 2003; 2(12):645-648.

[43] MacGregor EA, Janecek S, Svensson B. Relationship of sequence and structure to specificity in the alpha-amylase family of enzymes. Biochim Biophys Acta. 2001; 1546(1):1–20.

[44] Elbein AD. Microbial Glycobiology. 1st ed. San Diego: Academic Press; c2010. Chapter 11, Cytoplasmic carbohydrate molecules: trehalose and glycogen. p. 185–201.

CHAPTER 2: VAGINAL GLYCOGEN, NOT ESTRADIOL, IS ASSOCIATED WITH VAGINAL BACTERIAL COMMUNITY COMPOSITION IN BLACK ADOLESCENT WOMEN[2]

ABSTRACT

Purpose

The purpose of this study was to characterize the composition of vaginal bacterial communities in a cohort of black adolescent women and to determine how the species composition of these communities correlate with levels of estradiol, glycogen, and stress.

Methods

Twenty-one black adolescent women were sampled longitudinally. The composition of their vaginal communities was determined by analyzing the sequences of the V1-V3 region of 16S rRNA genes and they were grouped based on patterns in species abundances. The relationships between estradiol, glycogen, psychosocial stress, and the composition of these communities were assessed.

Results

Vaginal communities could be distinguished and classified into three groups that differed in the abundances of *Lactobacillus*. Eighty-one percent of study participants had communities dominated by species of *Lactobacillus*. Glycogen levels were higher in communities dominated by one or multiple species of *Lactobacillus* as compared to those having low proportions of *Lactobacillus*. Estradiol and psychosocial stress measurements did not differ among the three groups, while estradiol and glycogen exhibited a weak positive relationship that was not statistically significant.

Conclusions

The findings of this pilot study suggest that glycogen levels are associated with vaginal community composition in young black women; however, estradiol and psychosocial stress are not. Additionally, the results suggest there is no simple relationship between levels of estradiol and the production of vaginal glycogen.

[2] This chapter was previously published as: Nunn KL, Ridenhour BJ, Chester EM, Vitzthum VJ, Fortenberry JD, Forney LJ. Vaginal glycogen, not estradiol, is associated with vaginal bacterial community composition in black adolescent women. *J Adolesc Health*. 2019; DOI: 10.1016/j.jadohealth.2019.01.010.

Implications and Contribution

The positive association between *Lactobacillus*-dominance and glycogen levels suggest that factors affecting vaginal glycogen levels may influence clinical conditions such as bacterial vaginosis and alter susceptibility to other urogenital infections. Future studies might elucidate these factors and explore whether glycogen levels could serve as a surrogate that predicts community composition.

INTRODUCTION

Sexually transmitted infections (STIs) are highly prevalent among adolescent women [1]. In 2016 alone, young women ages 10 – 19 y comprised nearly one fifth of chlamydia and gonorrhea cases [2], and these STIs increase HIV risk in adolescents [3]. Rates are especially high among young black women [2]. While various physiological, behavioral, and psychosocial factors are thought to contribute [4-6], it is unclear why young black women have such high risk. Poor sexual and reproductive health outcomes in women have been linked to the bacterial composition of the vaginal microbiome of older women, but we know very little about the vaginal microbiomes of black adolescent women [7,8].

In pre-menopausal women older than 18 y, high proportions of *Lactobacillus* are associated with a lower prevalence of HIV and other STIs [9]. *Lactobacillus* species produce lactic acid, which is thought to inhibit colonization by pathogenic organisms [10-13]. However, approximately 25% of reproductive age women have vaginal communities that are depleted of *Lactobacillus* [11]. Moreover, the prevalence of communities dominated by *Lactobacillus* is greater in Asian and white women (80.2% and 89.7%, respectively) as compared to black and Hispanic women (61.9% and 59.6%, respectively) [14]. In addition to these observed differences between women, point estimates of vaginal bacterial community composition can change rapidly [15-17], sometimes resulting in low proportions of lactobacilli that may expose women to windows of risk for acquiring infections. Currently, the factors accounting for the compositional differences observed between women and temporal changes within women are poorly understood. Thus, there is a need to understand drivers of changes in vaginal community composition, especially in young black women.

The abundances of *Lactobacillus* species have been positively associated with circulating estrogen and vaginal glycogen content (reviewed in [11]). It has also been proposed that other factors such as psychosocial stress may influence vaginal community composition and account for the racial/ethnic differences in vaginal communities seen among adult women [18]. Thus, the main objective of this

study was to determine whether vaginal bacterial community composition in black adolescent women is associated with stress, estradiol, and glycogen. To explore this, we sampled the vaginas of twenty-one 14-year-old women at baseline and then monthly for six months. We determined the species composition of their vaginal bacterial communities, assessed levels of estrogen and glycogen, measured vaginal pH and Nugent scores, and assessed psychosocial stress. With these data, we addressed three questions: First, what kinds of vaginal bacterial communities do young black women have? Second, do the levels of estradiol, glycogen, and stress differ among women with different kinds of communities? And third, what are the relationships between estradiol, glycogen, stress, and key species of the vaginal communities?

METHODS AND MATERIALS

Study Design

Twenty-five self-identified black women (ages 14.01 – 14.99 years) were recruited from neighborhood clinics in Indianapolis, where most participants received primary care, to participate in a longitudinal study to assess the relationships between stress, estrogen, and vaginal community composition. This study was approved by the Institutional Review Board at Indiana University. We focused on 14-year-olds because this age is developmentally meaningful interval in which hygiene practices (such as pad or tampon use, douching, and pubic hair removal) become common and the nature of sexual interpersonal relationships changes rapidly and studying this group fills gaps in existing research that cannot otherwise be addressed. Written, informed consent of both participants and a parent were obtained prior to enrollment. Exclusion criteria at enrollment included structural abnormalities of the vagina, chronic medical conditions that could alter the vaginal microbiome, pregnancy, immune deficiency conditions, and antibiotic use within the previous 90 days. Two participants were excluded due to antibiotic use, and two were excluded due to health complications or non-compliance. In total, 21 participants provided seven monthly, self-collected vaginal swab and saliva samples, and self-assessments of psychosocial stress, menses, and sexual behaviors. All study participants were post-menarcheal.

Estrogen and glycogen measurements

Saliva samples were used to measure estradiol levels because the method of sample collection is non-invasive, well accepted assays exist [19,20], and levels of salivary estradiol are well correlated with serum estradiol concentrations [20,21]. For three days up to and including the scheduled day of vaginal swab sampling, once-per-day salivary samples were self-collected by passive drool into polypropylene

cryotubes and frozen shortly after collection. Salivary estradiol concentrations were determined with a commercially available enzyme-linked immunosorbent assay kit (17β-Estradiol high sensitivity ELISA kit ADI-901-174**,** Enzo Life Sciences). The standard curve was prepared per manufacturer's instructions but extended to eight standards with 7.8 pg/ml as the lowest value. The intra-assay variation was 4.2% and the inter-assay variation was 7.8%. During extraction, samples were concentrated 4X in assay buffer, and each was measured in duplicate using the protocol available at http://hdl.handle.net/2022/21883.

Glycogen in vaginal swab samples was quantified using the EnzyChrom Glycogen Assay Kit (BioAssay Systems) according to the manufacturer's instructions. To measure vaginal pH, subjects inserted a gloved finger into the vagina for ten seconds then rolled the finger over a commercially available pH stick (pH-EcoCare^{TM} Comfort; Merete Medical GmbH). The pH measurements were confirmed by a trained research associate and recorded in 0.5 increments. Self-obtained vaginal swabs (eSwab^{TM} Copan Diagnostics Inc.) were used to assess Nugent score, and provide vaginal samples for STI testing and microbial community analyses. Nugent score is a Gram stain scoring system used to diagnose bacterial vaginosis (BV) [22]. To assess Nugent scores, vaginal swabs were rolled onto glass microscope slides, air-dried in the field, and transported to the Infectious Diseases Laboratory at Indiana University. Slides were then stained according to standard procedures, and scored 0 – 3 (normal), 4 – 6 (intermediate), and 7-10 (abnormal/BV) [22]. STI testing for chlamydia and gonorrhea was conducted using Abbott Realtime CT/NG assay on the Abbott m2000 platform (Abbott Molecular, DesPlaines, IL). *Trichomonas vaginalis* testing was performed using a validated real time PCR assay on the Abbott m2000 platform. The remainder of swab samples were stored at -80^{0}C until DNA sequencing was done at the University of Idaho.

Measurement of psychosocial factors

Psychosocial factors were assessed using audio computer assisted self-interview [23], obtained at baseline, month three, and month six. The standardized scales (all previously validated in adolescent populations) addressed depression (the Patient Health Questionnaire [PHQ] [24]), stress (Perceived Stress Scale [PSS] [25]), and anxiety (the Brief Symptom Inventory [BSI] [26]).

Determination of menstrual cycle phase

Throughout the study self-reported menses data were used to determine the start and end date of menstrual periods for subjects not on hormonal birth control. We assumed an average luteal phase of 14 days, one day for ovulation, and the remaining days to be follicular. Based on these criteria, we divided menstrual period into phases, and matched sample collection dates accordingly.

Microbial community analysis

Total genomic DNA was extracted from vaginal swab samples using chemical and mechanical lysis, and purified using QIAamp DNA mini kits (Qiagen) as described previously [27]. The V1-V3 region of 16S rRNA genes were amplified using a two-step PCR protocol, first amplifying the gene region using universal primers 27F and 534R, and then adding sample barcodes and sequence adapters. Amplicons were sequenced using an Illumina MiSeq platform in the Genomics Resources Core facility at the University of Idaho. High quality reads were obtained from all samples except one, resulting in 146 samples total. Forward and reverse reads were paired using FLASH [28], processed through DADA2 [29] to identify unique sequences, and these were classified to genus and species levels using SPINGO [30]. A total of 422 taxa were identified. Of those, 44 taxa were present at a minimum of 1% in at least two individuals or at least 5% in one individual. Using this filter, the remaining taxa (378) were considered uncommon and therefore grouped into an “other” category that was included in subsequent analyses. The relative abundances of the taxa found in these communities are reported in Table B.1.

To group communities on the basis of similarities and differences in composition we performed complete-linkage hierarchical clustering on alt-Gower distances computed from taxon relative abundance data. Silhouette information was used to define nine clusters. Clusters were assigned to groups A, B, and C based on whether they were dominated by one species of *Lactobacillus* (group A), dominated by multiple species of *Lactobacillus* (group B), or had low proportions of lactobacilli (group C).

Statistics

Linear and linear mixed effects models were used for multiple analyses, including 1) modeling the means of the response variables estradiol, glycogen, vaginal pH, and psychosocial stress (BSI, PHQ, PSS) between groups; 2) modeling the means of estradiol between menstrual cycle phase; and 3)

characterizing the linear relationship between estradiol and glycogen. Response variables were transformed where appropriate to avoid violating model assumptions. Statistical significance of these models was determined using analysis of variance (ANOVA). Statistical comparisons were performed by testing general linear hypotheses and multiple comparisons of the means using Tukey's test. A Kruskal-Wallis rank sum test was used to evaluate group significant differences in Nugent score, with post-hoc analysis of multiple comparisons using a Dunn's test with Bonferroni adjustment. Finally, Pearson correlation coefficients were used to explore relationships between metadata and key taxa. More details on these analyses can be found in the Supplemental Material.

RESULTS

Study participants and metadata collection

We determined the relationships of stress, estrogen, and vaginal community composition in 21 black women who averaged 14.6 ± 0.3 years of age at the time of enrollment (Table B.2). The mean vaginal pH and Nugent score across all individuals were 5.2 ± 0.9 and 1.3 ± 2.5, respectively. This vaginal pH is higher than that reported for older reproductive age women. Vaginal glycogen levels across all samples showed substantial variability (mean glycogen 324.7 ± 399.1 µg /mL). Mean salivary estradiol across samples was 7.3 ± 4.5 pg/mL. Finally, across all measurements, the mean PHQ score was 5.9 ± 4.8 (range 0-25), the mean BSI score was 0.5 ± 0.7 (range 0-3.7), and the mean PSS score was 18.5 ± 6.5 (range 7-38). These scores indicate that on average the women of this cohort experienced mild to moderate depression, low anxiety, and moderate to high stress.

Vaginal bacterial community composition

To characterize the composition of vaginal communities we sequenced the V1 to V3 regions of 16S rRNA genes. Most participants (81%; 17 of 21) had communities that were dominated by *Lactobacillus*. Figure 2.1 shows a heat map of relative abundance data for the 20 most abundant taxa. The dendrogram in Figure 2.1 was used to identify nine clusters of communities that differed in composition. We observed that we could further group communities (i.e., combine clusters) based on the abundances of *Lactobacillus* species. Seven of the nine clusters had more than 50% total *Lactobacillus*, leaving two clusters in which the relative abundances of lactobacilli were less than 50%. The use of a 50% threshold to define "dominated" and "not dominated" was adopted from Klatt et al. [31] who used the same criterion in a study of how the abundance of lactobacilli was positively correlated with the efficacy of tenofovir. Of the seven clusters that were dominated by *Lactobacillus*, four were dominated by a single

species of *Lactobacillus* and the remaining three had mixtures of *Lactobacillus* species. These became groups A and B, respectively. All clusters that had communities with low proportions of *Lactobacillus* were aggregated into group C. In addition to low proportions of lactobacilli, these communities had higher proportions of *G. vaginalis* and mixtures of other bacteria such as *Atopobium vaginae*, *Corynebacterium spp.*, *Prevotella spp.*, *Peptoniphilus spp.*, *Streptococcus spp.*, and *Anaerococcus prevotii* (Figure 2.1). The mean relative abundances of key taxa in groups A-C are shown in Table B.3.

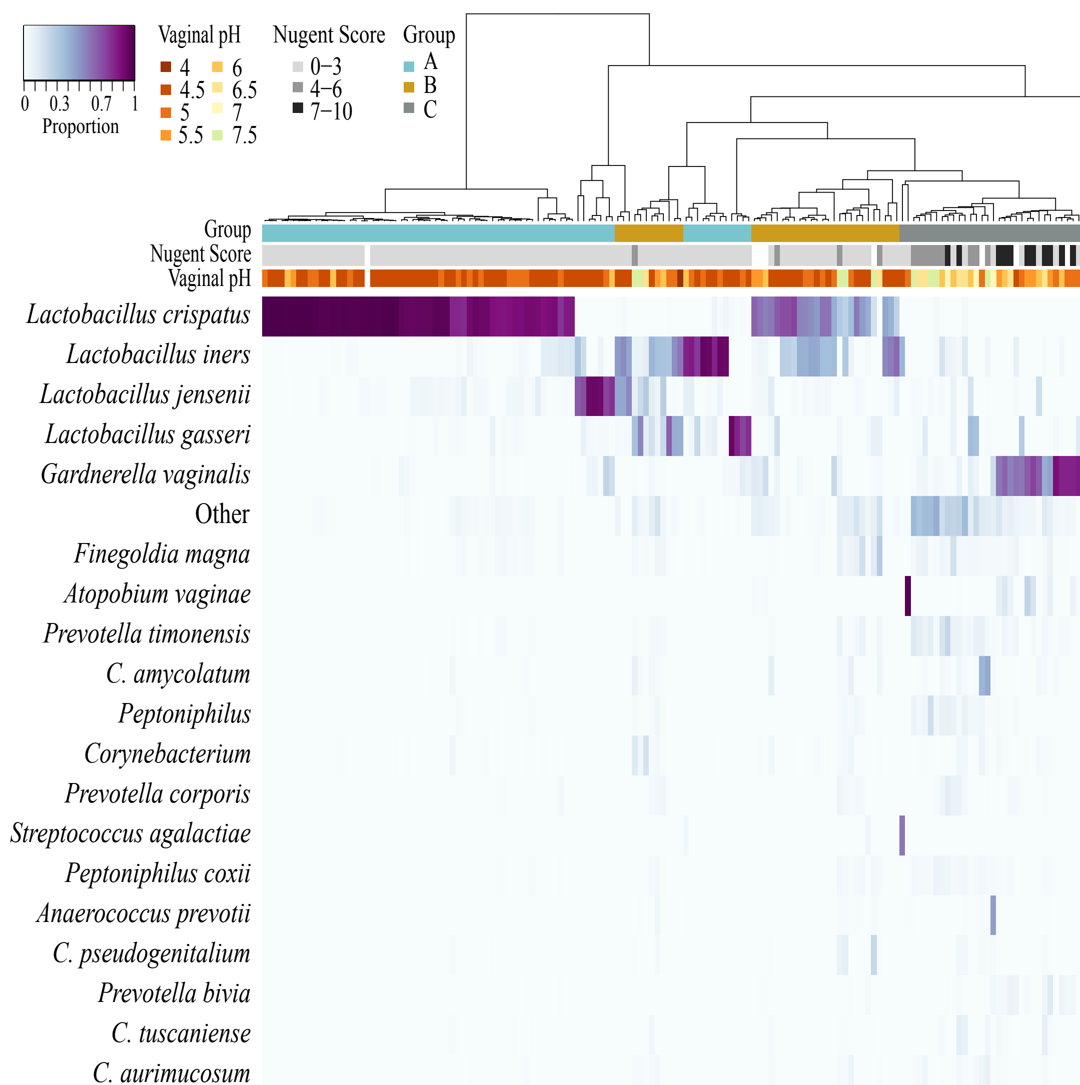


Figure 2.1 Heatmap based on the relative proportions of the 20 most abundant taxa in vaginal communities of black adolescent women. The columns of the heatmap include 146 samples collected from 21 young women over a 6-month period. The corresponding dendrogram represents complete-linkage hierarchical clustering of samples based on alt-Gower distances. The colored bar immediately below the dendrogram indicates which clusters were combined to form three groups (A, B, and C).

Vaginal community composition is known to change over time in older, reproductive age women. To determine whether similar trends occur in young black women, we created bar plots representing vaginal bacterial community composition over time for each subject (Fig. 2.2, Fig. B.1-B.2). Figure 2.2 shows plots for four subjects that illustrate the variability in community composition found within and between women of the cohort. Communities with high stability were seen in women of different groups. For example, Subject 35 (group A) maintained high proportions of *L. crispatus* over time, whereas Subject 11 (group C) maintained 50% or more *G. vaginalis* over time. In contrast, Subject 31 had high proportions of *G. vaginalis* (group C) during the first three months, then transitioned once to a community dominated by *L. gasseri* (group A). Subject 24 started with a community dominated by *L. jensenii* (group A) and transitioned to different groups five times over the course of the study.

Differences between groups

We determined if levels of estradiol, glycogen, vaginal pH, and Nugent score differed among groups A, B, and C. Vaginal pH and Nugent score were included in the analysis because they are well accepted correlates of vaginal community composition. We fit linear mixed effects models accounting for variation due to subject and performed non-parametric analyses where appropriate. The distributions of log-transformed estradiol and glycogen levels for each group are depicted by the boxplots in Figure 2.3a and 2.3b, respectively. There were no significant differences in estradiol levels among groups (χ^2 = 4.4, p = 0.1). Estradiol levels are known to vary over the menstrual cycle; therefore, we compared estradiol measurements between menstrual cycle phase to test our ability to detect differences in estradiol levels (Figure B.3). Mean estradiol concentrations were lowest in follicular samples and highest in peri-ovulatory samples; these differences were not significant (χ^2= 5.1, p = 0.08). Although we were unable to discriminate between estradiol levels, we did find that vaginal glycogen differed significantly between groups A and C (z = -4.1, $p < 0.001$).

Figure 2.3c and 2.3d show the distribution of Nugent scores and log-transformed vaginal pH for each group. As expected, group C had significantly higher Nugent scores and vaginal pH than did samples in groups A (Nugent: z = -7.6, pH: z = 6.5; $p < 0.001$ for both) and B (Nugent: z = -5.0, pH: z = 3.8, $p <$ 0.001 for both). Moreover, groups A and B, both marked by high proportions of *Lactobacillus*, had similar Nugent scores yet differed significantly in vaginal pH (Nugent: z = -1.9, p = 0.2; pH: z = 2.9; p = 0.01).

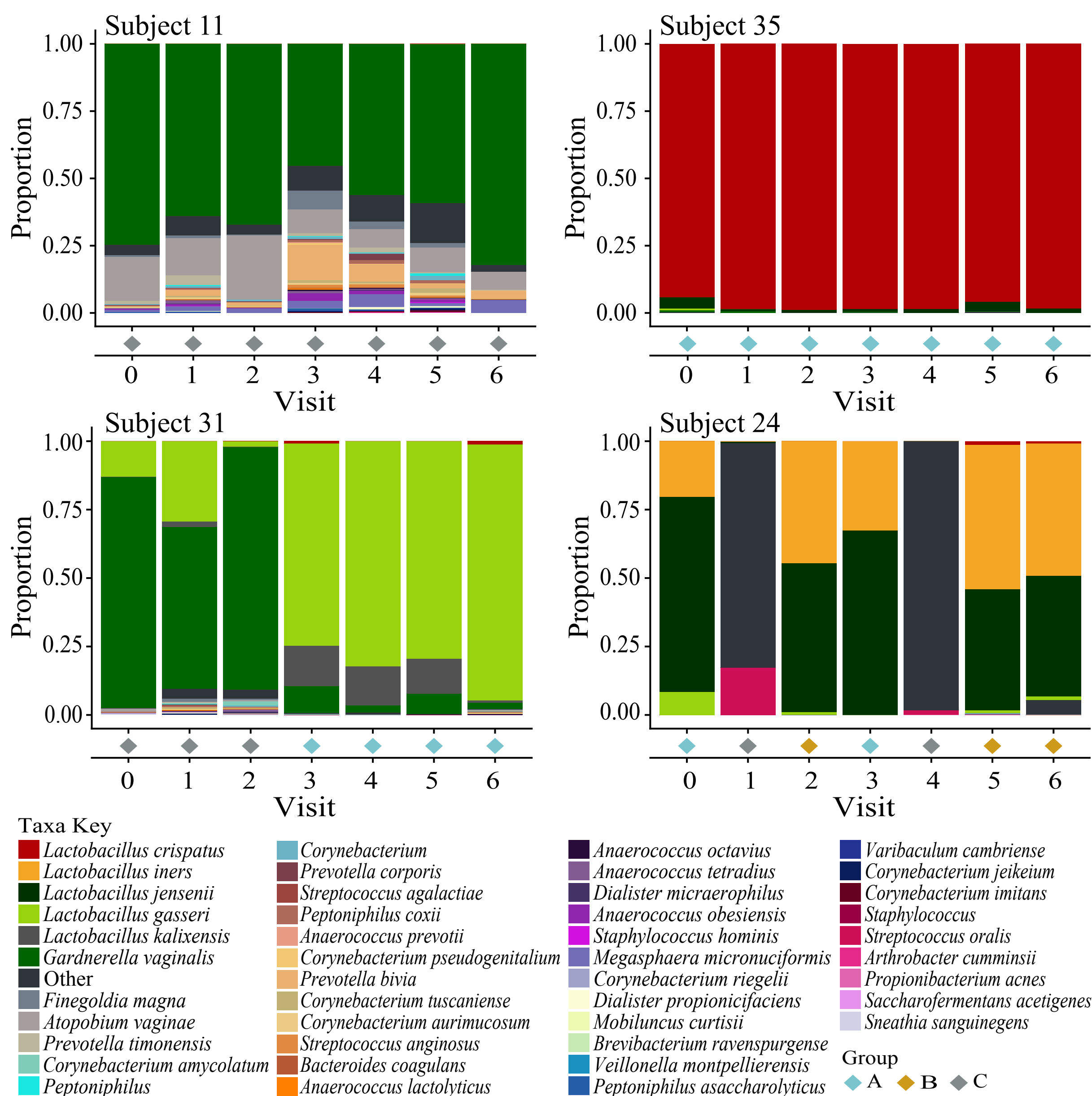


Figure 2.2 Examples of changes in community composition over time. The stacked bar charts represent proportions of bacteria in each community over 7 monthly visits where zero was the baseline visit. Colors for each taxon are shown in the legend below the figure. The community group for each visit is highlighted by a colored diamond below each chart and the corresponding legend is listed in the bottom right. The profiles of these subjects were chosen to illustrate the temporal variability in community composition within subjects.

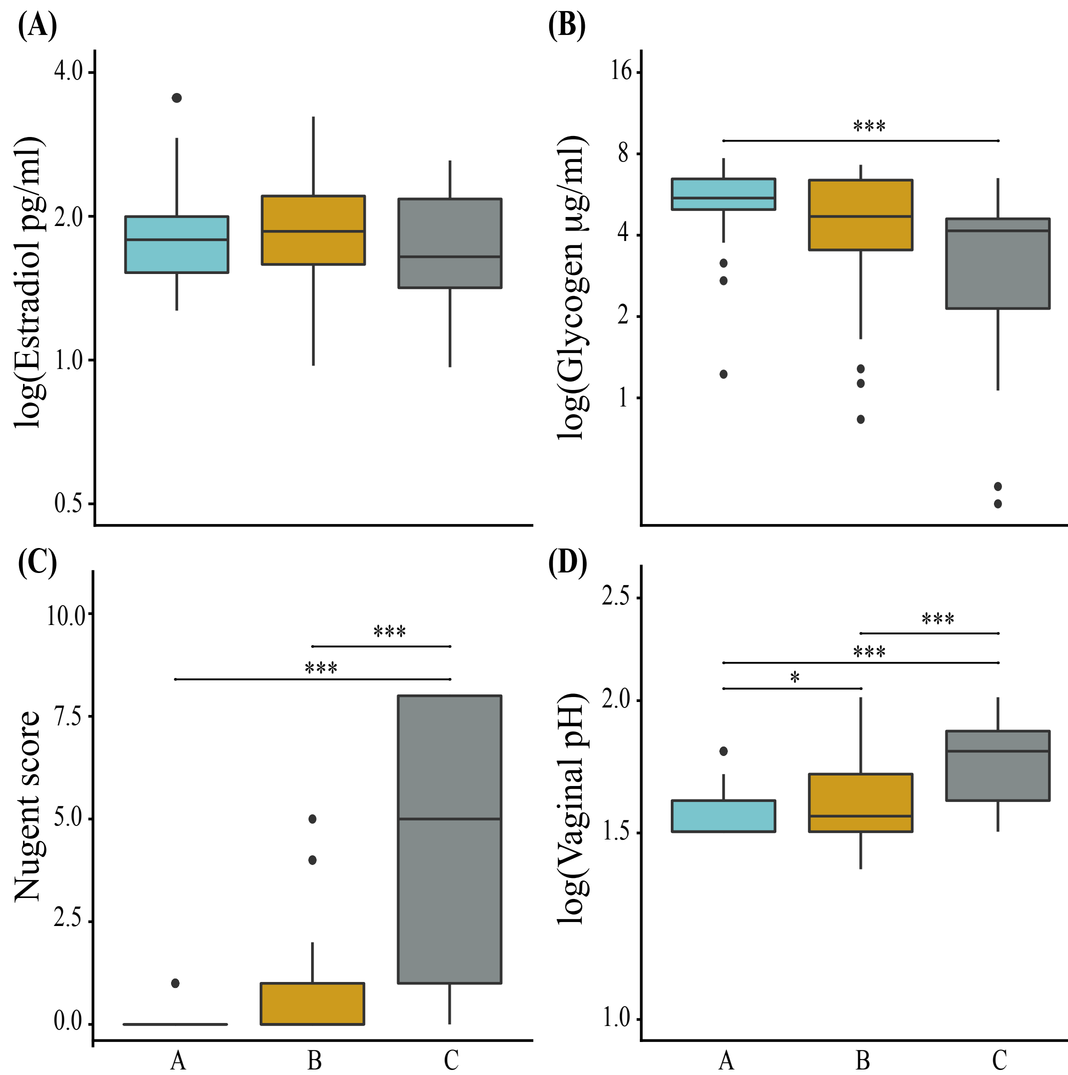


Figure 2.3 Differences in estradiol, glycogen, Nugent score, and vaginal pH measurements among groups. The boxplots represent log-transformed estradiol (panel A), log-transformed glycogen (panel B), Nugent scores (panel C), and log-transformed vaginal pH (panel D) for samples in groups A, B, and C. Statistical significance (* $p < 0.05$, *** $p < 0.001$) is indicated above the bars.

To gain insight to whether measurements of psychosocial factors differed between groups, we used linear mixed effects models, incorporating variation due to subject when necessary, to test for differences in the mean values of perceived stress, anxiety, and depression. Summary statistics for the models are shown in Table 2.1. No significant differences in perceived stress (F = 0.04, $p = 1$), anxiety ($\chi^2 = 0.7$, $p = 0.7$), or depression ($\chi^2 = 3.2$, $p = 0.2$) were identified.

Correlations between estradiol, glycogen, stress, and key taxa

We investigated the relationships between estradiol, glycogen, stress, and key taxa in the vaginal community by plotting Pearson correlation coefficients (Fig. B.4). As expected, vaginal pH and Nugent score were positively correlated (r = 0.63), and both pH and Nugent score were negatively correlated with glycogen (r = -0.39 and -0.34, respectively) (Fig. B.4 panel A). Vaginal glycogen was positively correlated with *L. crispatus* (r = 0.17) and *L. jensenii* (r = 0.26) (Fig. B.4 panel B). Estradiol was positively correlated with *L. iners* (r = 0.37) but no other *Lactobacillus* species (Fig. B.4 panel B). Furthermore, we did not observe a statistically significant correlation between levels of estradiol and glycogen.

To further evaluate the relationship between estradiol and glycogen, we fit a linear mixed effects model incorporating variation due to subject. Figure 2.4 shows a scatterplot of log estradiol vs. glycogen content in all samples. Our model suggests a positive relationship between estradiol and glycogen (m = 1.4 e-4); however, in agreement with the correlation plots, this relationship was not statistically significant ($p = 0.2$) (Fig. 2.4).

Table 2-1. Stress, anxiety, and depression in community groups A, B, and C.

Variable	Measure	Group						Statistic[d]	*p - value* [e]
		A		B		C			
		N	Mean (SD)	N	Mean (SD)	N	Mean (SD)		
PHQ score[a]	Depression	29	2.4 (0.2)	19	2.0 (0.3)	15	1.9 (0.3)	$\chi 2 = 3.2$	0.2
BSI score[b]	Anxiety	29	0.5 (0.1)	19	0.5 (0.1)	15	0.6 (0.2)	$\chi 2 = 0.7$	0.7
PSS score[c]	Stress	29	4.3 (0.1)	19	4.2 (0.2)	15	4.2 (0.2)	F = 0.04	1

[a] PHQ (Personal Health Questionnaire - 9) is a self-report questionnaire consisting of 10 questions (9 asking about specific symptoms, and the final asking how impactful those symptoms are to assess the severity of depression).

[b] BSI (Brief Symptom Inventory) – the anxiety subscale used here – is a self-report questionnaire consisting of 6 questions designed to clinically assess the level of anxiety in individuals.

[c] PSS (Perceived Stress Scale) is a self-report questionnaire used to evaluate the degree to which particular situations in one's life are deemed stressful.

[d] The test statistic listed results from the linear mixed effects models conducted for PHQ and BSI, and the linear model conducted for PSS to test differences in their means between groups A, B, and C. The means and standard error (SE) that are reported above result from the modeled means (betas) and SE. Variation due to subject was incorporated where appropriate. Type II Wald chisquare tests were calculated for PHQ and BSI, and the F-test was calculated for PSS.

[e] The p-value listed in this table is associated with the test statistic to the left.

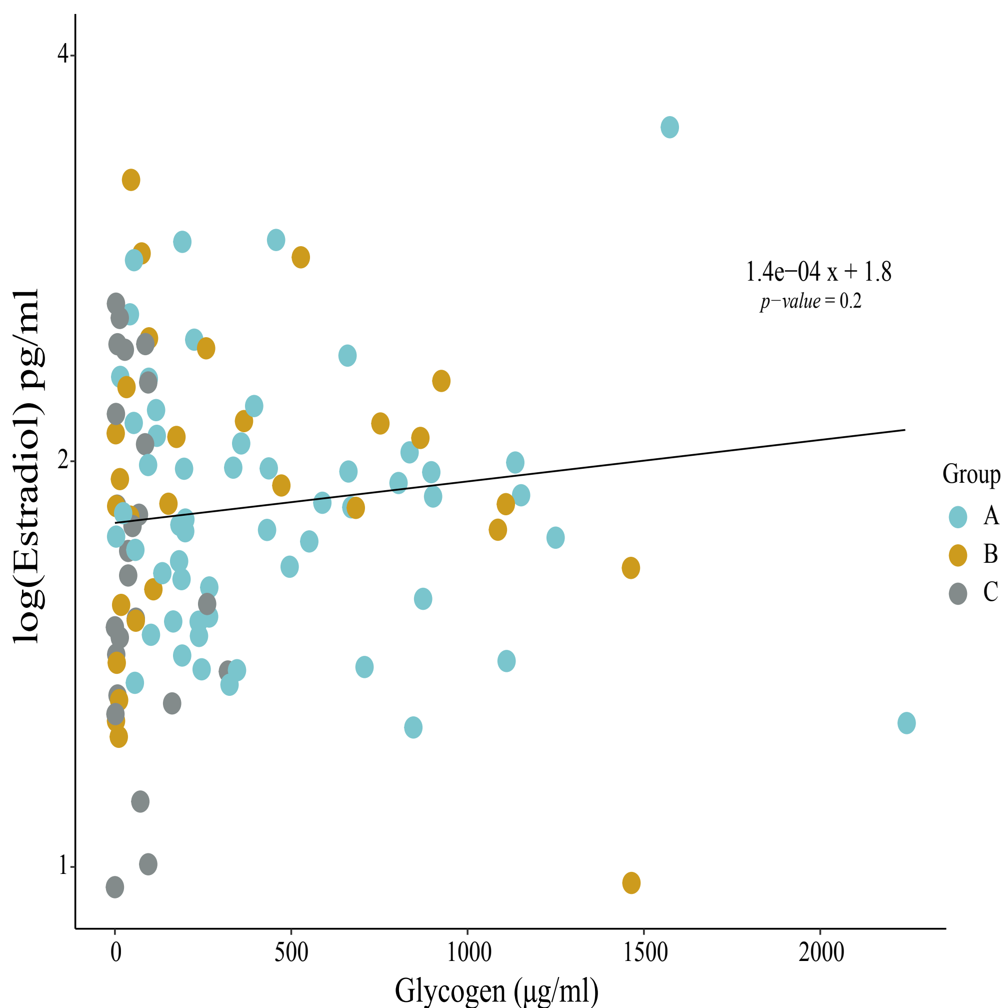


Figure 2.4 The relationship between log-transformed estradiol and glycogen concentrations in black adolescent women sampled longitudinally. The amounts of salivary estradiol in each sample were log-transformed and modeled over corresponding glycogen measurements including subject as a random effect. The resulting linear model equation and p-value (shown in the upper right corner) were obtained by computing an analysis of variance (ANOVA) on the linear model. Each dot represents the log estrogen concentration and the corresponding glycogen value for a given sample. Dots are colored according to community group as shown in the legend to the right of the graph.

DISCUSSION

We evaluated the relationships between mental health (perceived stress, depression, and anxiety), salivary estradiol, vaginal glycogen, and vaginal community composition in a small cohort of 14-year-old black women. We showed that most of the cohort (81%) had communities that were dominated by species of *Lactobacillus*. The kind and abundances of *Lactobacillus* species served as a basis for the classification of communities into three groups: A, B, and C. Consistent with reports in older reproductive age women [11], some women had stable vaginal communities whereas others transitioned from one community group to another. Importantly, differences in groups were marked by differences in vaginal glycogen levels but not salivary estradiol or any of the mental health measures.

Few studies have used culture-independent methods to characterize vaginal community composition in adolescent women. A cross-sectional study found young women, ages 13–18, to have either vaginal communities that were dominated by *L. iners* or *L. crispatus*, or contained a mixture of *L. crispatus*, *L. jensenii*, and *L. gasseri*, or were heterogeneous in composition with low proportions of *Lactobacillus* [7]. Another study showed *Lactobacillus* species to be prominent members in the adolescent vaginal microbiome early in puberty, even prior to menarche [8]. The high prevalence of *Lactobacillus* in our cohort of young black women (81%) is in agreement with previous studies, but is higher than that reported for older black women (61.9%) [14]. As noted above, reproductive age black women are less likely than white and Asian women to have *Lactobacillus* dominant communities [11,14], and more likely to have communities that have been associated with BV [32]. The fact that more than three quarters of this cohort of young black women had mostly *Lactobacillus* in their communities suggests there could be an age-associated transitional period in these vaginal communities. To understand this, we need more extensive longitudinal studies that evaluate the normal development of the vaginal microbiome in black women over time. This will broaden our understanding of what is healthy and what leads to the health disparities observed later in the lives of black women. In future studies the variability seen within ethnic groups might be lessened if host genetics were used to classify individuals instead of self-reported ethnicity.

Mental health such as chronic psychosocial stress has been associated with recurrent vulvo-vaginal candidiasis [33] and increased odds of vaginal conditions such as BV [34]. The correlation between chronic stress and BV has been demonstrated to be even more prominent in pregnant women [18]. Culhane et al. found that pregnant women who experienced high stress were 2.2 times more likely to have BV than those who experienced low stress [18]. These associations between stress and BV likely

result from stress-induced reduction in proteins involved in immune homeostasis that is associated with a decrease in the abundance of vaginal lactobacilli [35]. Although we sampled vaginal microbial communities that varied in terms of the abundance of *Lactobacillus* and other species, we did not observe significant differences in psychosocial stress, depression, or anxiety among the women studied. The sample size of our pilot study could have contributed to our inability to detect differences in stress and this is the most likely explanation for the discrepancies between the findings of our study and others [18,34]. However, it is also plausible that differences in the cohorts that were sampled in addition to variation in the perception of stress among cohorts of women could have played a role. Future research to evaluate the influence of psychosocial stress on vaginal community composition might include measuring a common biomarker of stress such as cortisol in addition to self-reported measures.

The positive association between estrogen, glycogen, and *Lactobacillus* over a woman's lifespan [11] has led to an assumption that there is a simple linear relationship between estrogen and the levels of glycogen in the vagina. Our results suggest otherwise and are in general agreement with those of Mirmonsef et al. [36] who sampled older reproductive age women over time and found no relationship between estrogen levels and vaginal glycogen. In an effort to understand this, one might assume that the rate of glycogen production (and release from cells) is counterbalanced by the rate of glycogen metabolism by members of the vaginal community. The resulting pseudo-steady state could result in some relatively constant level of glycogen in vaginal secretions in instances where the rate of glycogen production exceeds the rate of consumption. In contrast, the steady-state concentration of glycogen would be near zero if the rate of glycogen consumption is equal to or greater than the rate of glycogen production. This reasoning could also be extended to explain how high levels of glycogen can persist in low estrogen environments [37]. The relationship between the levels of estrogen and glycogen in the vagina might be further complicated by the fact that both the rates of glycogen production by the host and the rates of glycogen consumption by vaginal bacteria probably vary among women and over time within a woman. A second scenario emerges from the results of a study done by Pessina et al. [38] to evaluate the effects of steroidal hormones on the structure of the vaginal tissues of rats which showed that sub-physiological levels of estradiol thickened the vaginal epithelium more than physiological doses. Given this, it could well be that a positive linear relationship between estradiol and glycogen may only exist only at low concentrations of estradiol. As the estradiol level increases, its effect on glycogen production could be diminished or perhaps even saturated. The result would be a nonlinear relationship between estradiol levels and glycogen production in which glycogen is mainly influenced by the number of glycogen producing cells in the thickened vaginal epithelium.

In this study we sampled a relatively small number of individuals (21) infrequently over time and this limited our ability to resolve temporal changes in the vaginal microbiome and potential correlates of change (e.g., estradiol and stress), or to assess the effects of potentially confounding factors such as sexual activity or methods of birth control. Additionally, the sampling regimen precluded knowing the ovulatory status of subjects in a menstrual cycle. Since adolescents have a high frequency of anovulatory cycles [39] this might impact the vaginal environment in ways not commonly observed in adult women and confounded efforts to demonstrate an association between estradiol and microbiome composition. While these limitations exist, this is the first study done to evaluate estradiol, glycogen and stress as potential drivers of vaginal community composition in black adolescent women. Our results support previous findings that vaginal glycogen content is positively correlated with *Lactobacillus* dominance in vaginal secretions from older reproductive age women [40] and confirm that there is no simple relationship between levels of estradiol and the production of vaginal glycogen.

LITERATURE CITED

[1] Forhan SE, Gottlieb SL, Sternberg MR, et al. Prevalence of sexually transmitted infections among female adolescents aged 14 to 19 in the United States. Pediatrics 2009;124(6):1505–12. DOI: 10.1542/peds.2009-0674.

[2] Centers for Disease Control and Prevention. Sexually Transmitted Disease Surveillance 2016. Atlanta: U.S. Department of Health and Human Services; 2017.

[3] Newbern EC, Anschuetz GL, Eberhart MG, et al. Adolescent Sexually Transmitted Infections and Risk for Subsequent HIV. Am J Public Health 2013;103(10):1874–81. DOI: 10.2105/AJPH.2013.301463.

[4] Dellar RC, Dlamini S, Karim QA. Adolescent girls and young women: key populations for HIV epidemic control. J Int AIDS Soc 2015;18(2 Suppl 1):19408. DOI: 10.7448/IAS.18.2.19408.

[5] Hulland EN, Brown JL, Swartzendruber AL, et al. The association between stress, coping, and sexual risk behaviors over 24 months among African-American female adolescents. Psychol Health Med 2014;20(4):443–56. DOI: 10.1080/13548506.2014.951369.

[6] Meyer IH. Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. Psychol Sex Orient Gender Divers 2013;1(S):3–26.

[7] Yamamoto T, Zhou X, Williams CJ, et al. Bacterial populations in the vaginas of healthy adolescent women. J Pediatr Adolesc Gynecol 2009;22(1):11–8. DOI: 10.1016/j.jpag.2008.01.073.

[8] Hickey RJ, Zhou X, Settles ML, et al. Vaginal microbiota of adolescent girls prior to the onset of menarche resemble those of reproductive-age women. mBio 2015;6(2):e00097–15–14. DOI: 10.1128/mBio.00097-15.

[9] Borgdorff H, Tsivtsivadze E, Verhelst R, et al. *Lactobacillus*-dominated cervicovaginal microbiota associated with reduced HIV/STI prevalence and genital HIV viral load in African women. ISME J 2014;8(9):1781–93. DOI: 10.1038/ismej.2014.26.

[10] Boskey ER, Cone RA, Whaley KJ, et al. Origins of vaginal acidity: high D/L lactate ratio is consistent with bacteria being the primary source. Hum Reprod 2001;16(9):1809–13.

[11] Hickey RJ, Zhou X, Pierson JD, et al. Understanding vaginal microbiome complexity from an ecological perspective. Transl Res 2012;160(4):267–82. DOI: 10.1016/j.trsl.2012.02.008.

[12] Ma B, Forney LJ, Ravel J. Vaginal Microbiome: Rethinking Health and Disease. Annu Rev Microbiol 2012;66:371–89. DOI: 10.1146/annurev-micro-092611-150157.

[13] van de Wijgert JHHM, Borgdorff H, Verhelst R, et al. The Vaginal Microbiota: What Have We Learned after a Decade of Molecular Characterization? PLoS ONE 2014;9(8):e105998–10. DOI: 10.1371/journal.pone.0105998.

[14] Ravel J, Gajer P, Abdo Z, et al. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci USA 2011;108 Suppl 1:4680–7. DOI: 10.1073/pnas.1002611107.

[15] Gajer P, Brotman RM, Bai G, et al. Temporal Dynamics of the Human Vaginal Microbiota. Sci Transl Med 2012;4(132):132ra52–2. DOI: 10.1126/scitranslmed.3003605.

[16] Brotman RM, Ravel J, Cone RA, et al. Rapid fluctuation of the vaginal microbiota measured by Gram stain analysis. Sex Transm Infect 2010;86(4):297–302. DOI: 10.1136/sti.2009.040592.

[17] Keane FE, Ison CA, Taylor-Robinson D. A longitudinal study of the vaginal flora over a menstrual cycle. Int J STD AIDS 1997,8(8):489-94.

[18] Culhane J, Rauh J, McCollum KF, et al. Maternal Stress is Associated With Bacterial Vaginosis in Human Pregnancy. Matern Child Health J 2001;5(2):127–134.

[19] Gray SH, Ebe LK, Feldman HA, et al. Salivary Progesterone Levels Before Menarche: A Prospective Study of Adolescent Girls. J Clin Endocrinol Metab 2010;95:3507–11. DOI: 10.1210/jc.2010-0125.

[20] Worthman CM, Stallings JF, Hofman LF. Sensitive salivary estradiol assay for monitoring ovarian function. Clin Chem 1990;36:1769–73.

[21] Dielen C, Fiers T, Somers S, et al. Correlation between saliva and serum concentrations of estradiol in women undergoing ovarian hyperstimulation with gonadotropins for IVF/ICSI. Facts Views Vis Obgyn 2017;9(2):85–91. DOI: 10.1002/14651858.CD005289.pub3.

[22] Nugent RP, Krohn MA, Hillier SL. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. J Clin Microbiol 1991;29(2):297–301.

[23] Gutiérrez JP, Torres-Pereda P. Acceptability and reliability of an adolescent risk behavior questionnaire administered with audio and computer support. Rev Panam Salud Publica 2009;25(5):418–22.

[24] Kroenke K and Spitzer RL. The PHQ-9: A New Depression Diagnostic and Severity Measure. Psych Annals 2002; 32(9), 509–515.

[25] Cohen S, Kamarck T, Mermelstein R. A Global Measure of Perceived Stress. J Health Soc Behav 1983;24(4):385–396.

[26] Derogatis L and Melisaratos N. The Brief Symptom Inventory: an introductory report. Psychol Med 1983;13(3):595–605.

[27] Yuan S, Cohen DB, Ravel J, et al. Evaluation of methods for the extraction and purification of DNA from the human microbiome. PLoS ONE 2012;7(3):e33865. DOI: 10.1371/journal.pone.0033865.

[28] Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 2011;27(21):2957–63. DOI: 10.1093/bioinformatics/btr507.

[29] Callahan BJ, McMurdie PJ, Rosen MJ, et al. DADA2: High resolution sample inference from amplicon data. Nat Meth 2016;13(7):581–583. DOI: 10.1038/nmeth.3869.

[30] Allard G, Ryan FJ, Jeffery IB, et al. SPINGO: a rapid species-classifier for microbial amplicon sequences. BMC Bioinformatics 2015;16:324. DOI: 10.1186/s12859-015-0747-1.

[31] Klatt NR, Cheu R, Birse K, et al. Vaginal bacteria modify HIV tenofovir microbicide efficacy in African women. Science 2017;356(6341):938–45. DOI: 10.1126/science.aai9383.

[32] Fettweis JM, Brooks JP, Serrano MG, et al. Differences in vaginal microbiome in African American women versus women of European ancestry. Microbiol 2014;160(Pt_10):2272–82. DOI: 10.1099/mic.0.081034-0.

[33] Akimoto-Gunther L, Bonfim-Mendonça P de S, Takahachi G, et al. Highlights Regarding Host Predisposing Factors to Recurrent Vulvovaginal Candidiasis: Chronic Stress and Reduced Antioxidant Capacity. PLoS ONE 2016;11(7):e0158870. DOI: 10.1371/journal.pone.0158870.

[34] Nansel TR, Riggs MA, Yu K-F, Andrews WW, et al. The association of psychosocial stress and bacterial vaginosis in a longitudinal cohort. Am J Obstet Gynecol 2006;194(2):381–6. DOI: 10.1016/j.ajog.2005.07.047.

[35] Jašarević E, Howerton CL, Howard CD, Bale TL. Alterations in the Vaginal Microbiome by Maternal Stress Are Associated With Metabolic Reprogramming of the Offspring Gut and Brain. Endocrinology 2015;156(9):3265–76. DOI: 10.1210/en.2015-1177.

[36] Mirmonsef P, Hotton AL, Gilbert D, et al. Glycogen Levels in Undiluted Genital Fluid and Their Relationship to Vaginal pH, Estrogen, and Progesterone. PLoS ONE 2016;11(4):e0153553. DOI: 10.1371/journal.pone.0153553.

[37] Ayre WB. The glycogen-estrogen relationship in the vaginal tract. J Clin Endocrinol 1951;11:103–110. DOI: 10.1210/jcem-11-1-103.

[38] Pessina MA, Hoyt RF, Goldstein I, Traish AM. Differential effects of estradiol, progesterone, and testosterone on vaginal structural integrity. Endocrinology 2006;147(1):61–9. DOI: 10.1210/en.2005-0870.

[39] Rosenfield RL. Clinical review: Adolescent anovulation: maturational mechanisms and implications. J Clin Endocrinol Metab 2013;98(9):3572–83. DOI: 10.1210/jc.2013-1770.

[40] Mirmonsef P, Hotton AL, Gilbert D, et al. Free glycogen in vaginal fluids is associated with *Lactobacillus* colonization and low vaginal pH. PLoS ONE 2014;9(7):e102467. DOI: 10.1371/journal.pone.0102467.

CHAPTER 3: AMYLASES IN THE HUMAN VAGINA[3]

ABSTRACT

Dominance of *Lactobacillus species* in vaginal communities is a hallmark of healthy conditions in the female genital tract. Sugars produced by α-amylase degradation of glycogen in the vagina are thought to support the proliferation of lactobacilli. While α-amylase activity has been demonstrated in vaginal fluids, it is unclear whether α-amylase is produced by the host, bacteria in the vaginal community, or both. The goal of this study was to use proteomics to identify amylases in the vagina and when possible, determined the likely source(s) of these enzymes. We collected and characterized cervicovaginal mucus (CVM) from 23 reproductive age women and measured vaginal pH and the levels of amylase activity, glycogen, and lactic acid. The species composition of vaginal communities was determined by analyzing partial 16S rRNA gene sequences. Based on differences in pH, amylase activity, lactic acid levels, and community composition, four samples were selected for metagenomic and proteomic analyses. We identified eight putative amylase proteins among the assembled bacterial metagenomes, four of which were determined to be expressed in vaginal fluids. Specifically, the following proteins were detected in vaginal fluids: α-amylase from *Bifidobacterium lacrimalis*; pullulanase from *Lactobacillus iners*, *Bifidobacterium vaginale*, and *B. lacrimalis*; oligo-1,6-glucosidase from *L. iners*; and intracellular maltogenic amylase from *Lactobacillus crispatus*. Moreover, there were multiple amylases detected in each of the four samples that were analyzed. We also detected human α-amylase in vaginal fluids but were unable to discern whether the source was salivary or pancreatic. These findings provide evidence that there are multiple amylases present in the vagina of any given woman, which are produced by several bacteria in the vaginal community and the host.

INTRODUCTION

Bacterial communities in the human vagina play an integral role in maintaining reproductive health. Within these communities, *Lactobacillus* species provide a key ecosystem service by producing lactic acid [1-4], which is thought to restrict pathogenic organisms from colonizing the vagina [3]. The abundances of lactobacilli and the overall composition of vaginal microbiota differs markedly between women [5-10]. Moreover, some women have vaginal communities that consistently lack appreciable

[3] This chapter is to be submitted for publication as: Nunn KL, Claire GC, Adkins J, Engbrecht, K, Fillmore T, and Forney, LJ. Amylases in the Human vagina.

numbers of lactobacilli whereas others experience periods of time where their communities transition to states where the abundances of lactobacilli are low [11-14]. All instances in which there are low abundances of lactobacilli constitute windows of elevated risk for disease and adverse reproductive outcomes. Thus far, the key drivers of vaginal community composition are poorly understood.

The abundances of *Lactobacillus spp*. in the vagina are positively associated with the levels of estrogen and vaginal glycogen content [15,16]. Increased estrogen levels cause the vaginal epithelium to thicken, thereby prompting the accumulation of glycogen within epithelial cells. These events correlate with increases in the absolute abundances of bacteria and in the proportions of *Lactobacillus spp.* in the vagina. In ways that are not fully understood estrogen and glycogen are thought to create an environment that stimulates the proliferation of vaginal lactobacilli, possibly because it serves as a source of carbon [17].

Glycogen is a regular repeating glucose polymer, linearly connected by α-1,4-glycosidic bonds with branching at roughly every eight to ten residues by way of α-1,6-glycosidic bonds (reviewed in [18]). Alpha-amylase is an extracellular glycoside hydrolase that cleaves α-1,4-glycosidic bonds, producing maltose, maltotriose, and α-limit dextrins [19]. This enzyme, which investigators have presumed to be host-derived, is likely required to first depolymerize glycogen to form simpler sugars that can be transported into cells where it is further catabolized by *Lactobacillus* to produce lactic acid [20]. Humans express α-amylase in the salivary glands (AMY1) and in the pancreas (AMY2) [21,22]. Meanwhile, bacterial species have been shown to produce α-amylase in various habitats as well as other kinds of amylases, including but not limited to β-amylase, glucoamylase, α-glucosidase and pullulanase [23-25]. Currently, it is not known whether bacterial species that reside in the human vagina produce amylases. Moreover, it has not been shown whether either of the human α-amylases are present in vaginal fluids.

Enzymes often overlap in catalytic function, and according to three dimensional structure and function, α-amylases and other glycoside hydrolases are grouped into many different protein families [26]. The largest glycoside hydrolase family, glycoside hydrolase 13 (GH13), is colloquially known as the α-amylase family. GH13 includes various types of hydrolases and due to its complexity has been further divided into at least 35 subfamilies of proteins [26,27]. Given this, we hypothesize there are multiple kinds and sources of amylase in the human vagina, and most if not all of these are produced by bacteria.

To explore this, we obtained self-collected cervicovaginal mucus samples (CVM) from reproductive age women and used proteomics to identify amylases in CVM. We detected both host and bacterial enzymes in vaginal fluids with the functional capability to hydrolyze both α-1,4- and α-1,6-glycosidic bond in glycogen. Further, we determined that multiple amylases were present in the vagina of any given woman.

METHODS AND MATERIALS

Study Design

Reproductive age women were recruited to participate in a pilot study designed to learn more about the amylases present in the human vagina. This study was approved by the Institutional Review Board at the University of Idaho (IRB #18-118). Written informed consent was obtained from each study participant prior to sample collection. Women were enrolled if they were not pregnant, not experiencing vaginal bleeding, did not have an intrauterine device, not taking oral or topical antibiotics to treat vaginal infections, and not experiencing vaginal symptoms that might indicate the presence of a sexually transmitted infection or disease. Moreover, participants were asked to refrain from having vaginal intercourse and using vaginal lubricants within 48 hours of sample collection, as those events could influence the vaginal microbiome [28-30]. Donors were given a unique identifier starting with the letter F followed by a number, and these designations are used throughout this work.

Self-collected vaginal samples were obtained from study participants as previously described [31]. Briefly, participants were provided a 50 ml conical tube and a commercially available device used to collect menstrual fluid called the Softcup®. Participants were asked to insert the Softcup® into their vagina, leave it there for one minute, remove it, and then place it into the 50 ml conical tube. Cervicovaginal mucus (CVM) was collected at the bottom of the tube by centrifugation at 2,000 rpm for five minutes and then transferred to a 1.5 ml tube. The volume, weight, and pH of the samples were recorded in addition to observations on sample consistency and appearance. To measure pH, ten microliters of mucus were spread over a commercially available pH stick (EMD Millipore) and based on the color change, pH was recorded in 0.3 increments. Samples were then aliquoted into 1.5 ml tubes and stored at -80 ºC.

Amylase activity measurements and resolution by native polyacrylamide gel electrophoresis (PAGE)

CVM samples were centrifuged at 14,000 rpm for five minutes to separate the mucus into a solid and aqueous phase. The aqueous phase was aspirated and transferred to a clean 0.5 ml tube and stored at -20ºC until further use. Amylase activity in CVM was measured using the commercial EnzChek® Ultra Amylase Assay Kit (Molecular Probes) following manufacturer's instructions. *Bacillus* sp. α-amylase (Sigma, A-6380) was used to generate a standard curve. Duplicate measurements of amylase activity were recorded in U/ml. These measurements were normalized by total protein in samples, measured using the Bradford assay, and recorded in U/mg of total protein.

We resolved amylase proteins using the aqueous phase of samples prepared above by native PAGE. Briefly, 20 µg of total protein was diluted in native sample buffer (BioRad) and loaded on 12% Criterion™ TGX™ precast midi protein gels (BioRad). Human pancreatic α-amylase and *Rhizopus spp*. glucoamylase were loaded with samples to serve as positive controls. PAGE was performed in 1X Tris-Glycine buffer and gels received 36 V for one hour to allow samples to settle into the stacking portion of the gel and then ~150 V for 4.5 hours. Gels were washed in deionized water for two minutes and incubated for one hour at 37ºC in 1% starch solution in 0.02M Tris-Cl with 1mM $CaCl_2$, pH 7.4. Afterwards, gels were washed in deionized water for two minutes, incubated in 0.02M Tris-Cl pH 7.4 for ten minutes at 37ºC, and rinsed in deionized water again. Gels were developed by overlaying Miracloth (Calbiochem) previously soaked in 10mM iodine/14mM potassium iodine on the gel for 10 minutes and destained with deionized water. Then, gels were fixed in 1% acetic acid and imaged on the BioRad Gel Doc™ XR+. This method allows us to determine whether amylase is present in a sample and visualize the pattern of isoenzymes produced because iodine stains portions of the gel with intact starch dark blue, leaving segments of the gel in which starch has been hydrolyzed clear.

Glycogen and lactic acid measurements

CVM samples were diluted 5-fold (w/w) in 1X PBS, homogenized by vortexing for one minute, and centrifuged at 14,000 rpm for five minutes. The supernatant was transferred to a clean 1.5 ml tube and stored at -20ºC until further use. Glycogen was quantified using the EnzyChrom Glycogen Assay Kit (BioAssay Systems) according to the manufacturer's instructions. Duplicate measurements for glycogen were recorded in mg/ml. To measure the concentrations of lactic acid in the samples, we used

commercially available D- and L- lactic acid kits (Bioassays) following the manufacturer's protocol. Duplicate measurements of D-, L-, and total lactic acid were recorded in mM.

Bacterial community analysis

The species composition of vaginal microbial communities was determined by classifying partial 16S rRNA gene sequences. We used methods that have been previously described [32] with a modification in the sample volume. In brief, total genomic DNA was extracted from 25 µl of CVM using chemical and mechanical lysis and purified using QIAamp DNA mini kits (Qiagen). Genomic DNA concentrations were determined using the Quant-iT™ PicoGreen™ dsDNA assay kit (Invitrogen). For amplicon sequencing, the V1-V3 16S rRNA gene regions were amplified using a two-step PCR protocol, first amplifying the gene region using universal primers 27F and 534R, and then adding sample barcodes and sequence adapters. Amplicons were sequenced using the Illumina MiSeq at the University of Idaho. Forward and reverse reads were paired using FLASH [33], processed through DADA2 v 1.12.1 [34] to identify distinct sequences, and the distinct sequences were classified to genus and species level using SPINGO [35].

For whole genome shotgun sequencing, DNA libraries were prepared using the Nextera DNA library kit (Illumina) and pooled for sequencing at the IBEST Genomics Resources Core at the University of Idaho. Whole genome shotgun sequencing was performed using the HiSeq 4000 at the University of Oregon. Sequences were quality trimmed then filtered, assembled, and annotated using the metaWRAP pipeline [36]. The resulting metagenomic assembled genomes (MAGs) were classified to species level using gtdbtk [37].

Proteomic analysis of vaginal fluids

To analyze the proteins in vaginal fluids, CVM was processed using the MPLeX method which has been previously described [38,39]. Briefly, this method uses a solvent based extraction that incorporates a mixture of water, chloroform, and methanol to extract proteins, lipids, and polar metabolites in three different fractions. Once extracted, proteins were reduced with dithiothreitol, alkylated using iodoacetamide, and digested with trypsin. Digested proteins were then concentrated and analyzed by reverse-phase liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). Raw MS data were searched in MaxQuant [40] against MAGs generated in this study in addition to the human

Uniprot non-redundant database to identify proteins in the four samples. Proteins were quantified using the intensity-based absolute quantification (iBAQ) method [41].

Statistics

All analyses were performed in R (version 3.6.0) [42]. We used a generalized linear model with a log-link function to assess the relationship between amylase activity and glycogen levels. The pseudo-R^2 for the model was calculated by subtracting the ratio of the model's residual deviance and the null deviance from the value of one. To analyze proteomic data, protein intensities were log2 transformed and normalized by computing z-scores. Hierarchical clustering was performed using Ward's method [43] on Euclidian distances calculated from normalized protein intensities. Statistical significance was set to the level of $P = 0.05$.

RESULTS

Study cohort

The overall goal of this study was to identify amylases in the vagina and if possible, determine their source(s). To accomplish this, we enrolled 23 women in a pilot study to obtain self-collected CVM. A summary of the metadata regarding the cohort is in Table C.1. Study participants ranged from 19 to 45 years of age, with an average age of 27. Women reported that they were in good health, were not pregnant, had their ovaries or uterus, and had not had vaginal intercourse in the 48 hours prior to sample collection. In this cohort, the species composition of vaginal communities were similar to those found in other studies of healthy reproductive age women (Table C.2) [10]. We found measurements of pH, total protein, amylase, glycogen, and lactic acid to vary substantially across donor samples (Table C.3).

All of the samples collected in this study were used to develop a better understanding of vaginal amylases, specifically to characterize the relationship between amylase activity and glycogen levels and to determine whether amylase activity is associated with a particular bacterial community composition in the vagina. However, only a subset of samples was selected for additional analyses to identify amylase proteins using proteomics and to assemble bacterial metagenomes. This subset was chosen based on differences in pH, amylase activity, glycogen levels, and lactic acid measurements. Essentially, we chose donor samples based on whether or not they were below, at, or above the mean value for each measurement. We also factored in differences in vaginal community composition. In

total, we selected four CVM from donors F02, F06, F08, and F12 for proteomics and metagenomics. The corresponding metadata for these samples are in Table C.4.

Glycogen levels are negatively correlated with amylase activity

Given that amylase is thought to depolymerize glycogen into simpler sugars that are metabolized by bacteria in the vaginal community [20], we expected there to be a negative association between amylase activity and vaginal glycogen levels. To evaluate this relationship, we measured amylase activity and glycogen levels in CVM and fit a generalized linear model using a log link function to these data. The resulting scatter plot of glycogen as the response variable and amylase activity as the predictor variable is show in Figure 3.1. There was considerable variability in glycogen measurements (range 3.7 – 32.8 mg/ml) when the corresponding amylase activity was less than 0.5 U/mg total protein. However, the model indicated there was a negative relationship between amylase activity and glycogen levels (R^2 = 0.21, p = 0.02), which implies that glycogen levels tend to decrease with increasing amylase activity in the vagina.

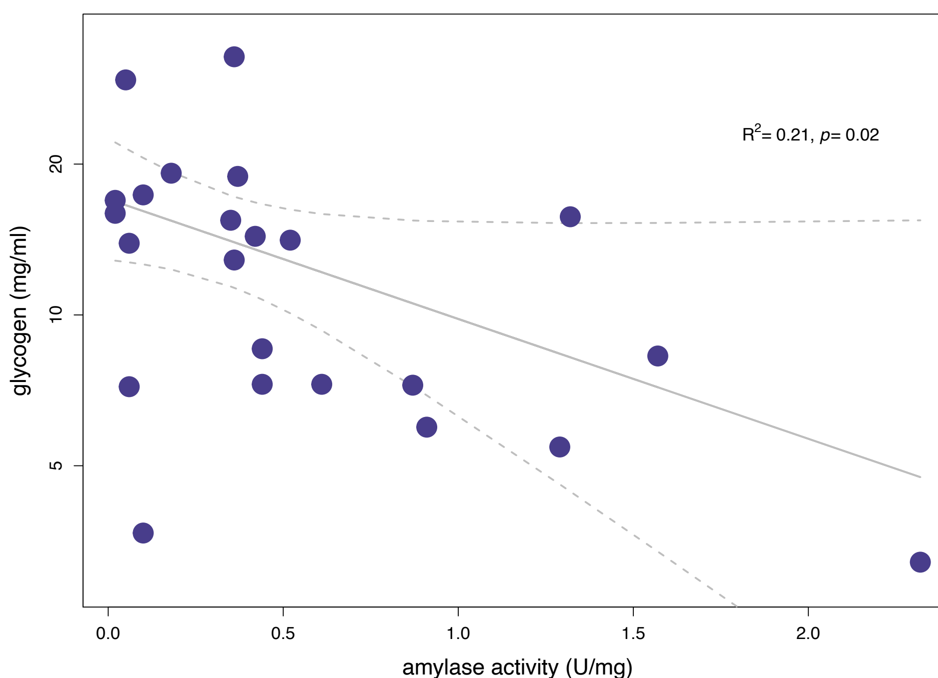


Figure 3.1 Scatter plot of glycogen measurements in mg/ml as a function of amylase activity measured in U/mg of total protein. The points in the graph represent glycogen measurements for each sample, plotted with the corresponding amylase activity for the sample. The regression line was generated by fitting a generalized linear model to these data points. The dashed lines outside of the regression line represent the 95% confidence interval for the model fit. The pseudo-R^2 and the P value for the model are shown in the upper right-hand corner of the figure.

Amylase activity is not associated with bacterial species composition in the vagina

To get a sense as to whether amylase activity was associated with the composition of vaginal communities, we prepared a summary plot of bacterial species composition for each sample with the corresponding amylase activity measurement (Fig 3.2). Fig 3.2A shows a dot plot of amylase activity measured in U/mg of total protein by donor sample. Below each dot in Fig 3.2A is a stacked bar (Fig 3.2B) that represents the relative proportion of bacteria in that sample. Of the 23 women studied, 21 of them had communities dominated by species of *Lactobacillus*. One woman had a community dominated by *Gardnerella vaginalis*, and the remaining woman had a community dominated by *Bifidobacterium pseudocatenulatum*. The two communities that were dominated by *L. jensenii* had average amylase activity. On the other hand, the lowest and highest values of amylase activity were observed in communities that were dominated by either *L. crispatus* or *L. iners*. Despite the differences in community composition, there were no obvious distinctions in amylase activity. Thus, from these data, we cannot conclude whether the presence of any particular bacterium is related to the amylase detected in the vagina.

The electrophoretic mobility of amylase isozymes varies across donors

Native polyacrylamide gel electrophoresis (PAGE) has been widely used to resolve native amylase proteins and can distinguish the two forms of human α-amylase (reviewed in [22]). Here, we used native PAGE to observe whether we could detect different amylases in CVM. We found the electrophoretic mobility of amylase isozymes to differ between these samples (Figure 3.3, Figure C.1). Figure 3.3 shows images of amylase proteins in CVM from donors F02, F06, and F08 that were separated on 12% Criterion™ pre-cast gels. We observed two isozymes in CVM from donor F02, six were visible in F06, and seven were detected in F08. We did not observe amylase isozymes in CVM from F12 (Figure C.1). Given that we were able to measure low levels of amylase activity in CVM from F12, we proceeded with including it for further analyses. The differences in electrophoretic mobility and number of isozymes observed suggests that quantity and type of amylase enzyme likely differs between women. Furthermore, the zones of starch hydrolysis reflected by the dark banding pattern in the image varied in intensity, which implies there is variation in enzyme activity among the different amylases.

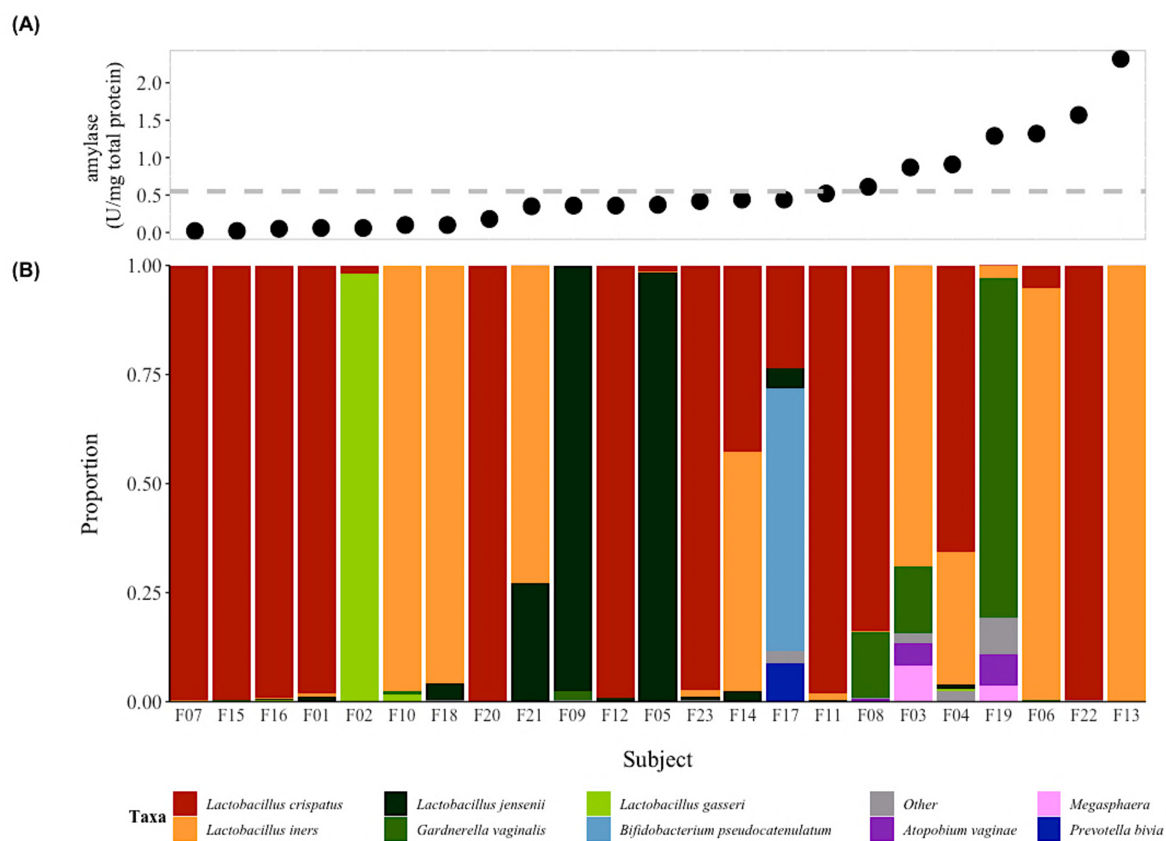


Figure 3.2 Amylase activity relative to bacterial species composition in the human vagina. (A) Amylase activity expressed in enzyme units (U) per milligram (mg) of total protein in the sample is arranged in increasing order with the corresponding stacked bar (B) below it. The dashed grey line in panel A represents the mean of amylase activity across samples (0.55 U/mg). The stacked bars represent the relative proportions of bacteria in each community. Colors for each taxon are shown in the legend below the figure.

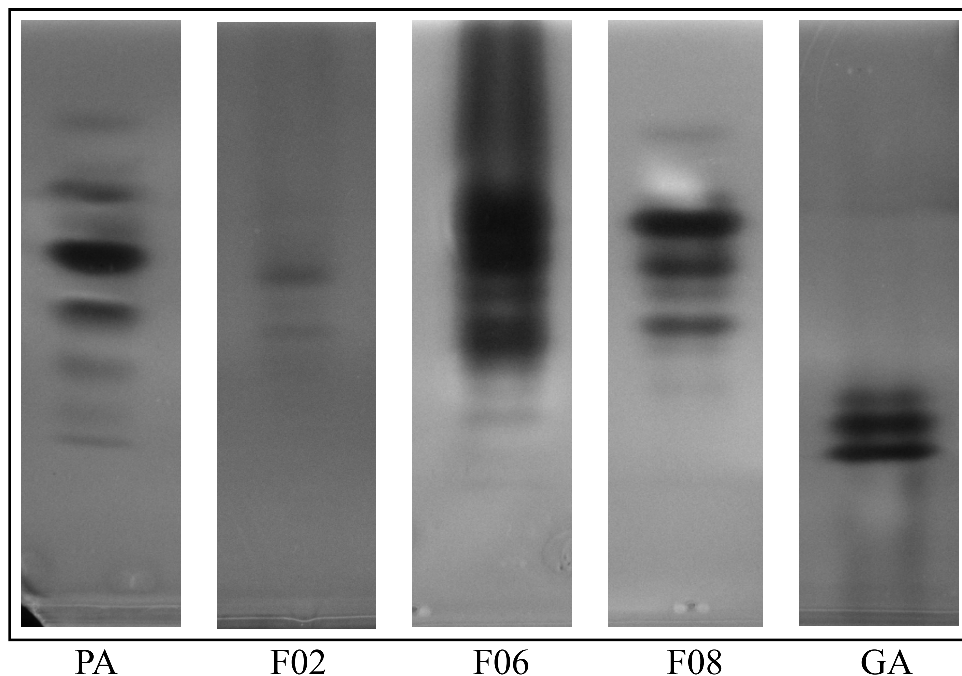


Figure 3.3 Amylase enzymes in vaginal fluids collected from donors F02, F06, and F08 separated by native PAGE. These columns represent image slices taken from Figure C.1 (containing all of the images of native PAGE gels completed for all samples) for donor samples F02, F06, and F08. A total of 20 μg of total protein was loaded for each sample. After the amylase proteins were separated in the gels, gels were incubated in 1% starch to enable starch hydrolysis and stained with iodine solution to observe where hydrolysis took place. The images that the slices were extracted from were inverted so that the dark background shows up white and the clear zones indicating hydrolysis of starch show up dark. Thus, the dark bands reflect amylase isozymes that were resolved in each donor sample. PA (human pancreatic α-amylase) and GA (Rhizopus spp. glucoamylase) are two commercial amylases that were used as positive controls for the assay.

Several putative amylase proteins are present in vaginal bacterial metagenomes

To identify putative amylase genes in vaginal bacteria we performed shotgun metagenomics on genomic DNA obtained from donors F02, F06, F08, and F12. We assembled, annotated, and classified metagenome assembled genomes (MAGs) to the species level. We recovered two MAGs from the metagenome of donor F02 and found they were derived from *Lactobacillus gasseri* and *Bifidobacterium vaginale*. One MAG was recovered from F06 and identified as *Lactobacillus iners*. Five MAGs were recovered from F08 and identified as *Fannyhessea vaginae*, *Ureplasma parvum*, *Lactobacillus crispatus*, *Bifidobacterium vaginale*, and *Bifidobacterium lacrimalis*. Finally, three MAGs were recovered from F12 and identified as *Lactobacillus crispatus*, *Lactobacillus jensenii*, and *Bifidobacterium vaginale*. MAGs were classified using the GTDB taxonomy database [37], and some of the species designations were cross listed using a name that differed from that in the NCBI taxonomy database. In the NCBI taxonomy database *Bifidobacterium vaginale* is listed as *Gardnerella vaginalis*, *Bifidobacterium lacrimalis* is listed as *Peptoniphilus lacrimalis*, and *Fannyhessea vaginae* is listed as *Atopobium vaginae*. Due to the high prevalence of horizontal gene transfer in human associated communities [44], we must emphasize that these taxonomic designations for the MAGs should be considered putative hosts.

The GH13 protein family includes both transferases and hydrolases, but for this work, we were only interested in detecting proteins with hydrolase activity. Thus, we filtered the annotations to identify hydrolase proteins within the GH13 protein family (EC number 3.2.1.-). We found a total of eight different GH13 hydrolase proteins that were annotated as α-amylase (EC 3.2.1.1), glucan 1,6-α-glucosidase (EC 3.2.1.70), oligo-1,6-glucosidase (3.2.1.10), intracellular maltogenic amylase (EC 3.2.1.-), trehalose synthase/amylase TreS (EC 3.2.1.1), glycogen operon protein glgX (EC 3.2.1.-), pullulanase (EC 3.2.1.41), and neopullulanase (EC 3.2.1.135). These proteins, which all have the ability to cleave either α-1,4- and/or α-1,6-glycosidic linkages (Table 3.1), were detected in multiple MAGS. We did not detect any of these proteins in the MAGs designated as *F. vaginae* or *U. parvum*.

Table 3-1 Putative amylase proteins annotated in the metagenomic assembled genomes of bacteria in human vaginal samples.

Sample	Taxa[a,b]	Protein	EC #	Reaction	Gene(s)
F02	*Bifidobacterium vaginale*	Oligo-1,6-glucosidase	3.2.1.10	α-1,6-glycosidic linkages	malL
		Pullulanase	3.2.1.41	α-1,6-glycosidic linkages	pulA
		Trehalose synthase/amylase TreS	3.2.1.1	α-1,4-glycosidic linkages	treS
		α-amylase	3.2.1.1	α-1,4-glycosidic linkages	aml
		Neopullulanase	3.2.135	α-1,4-, and α-1,6-glycosidic linkages	nplT
	Lactobacillus gasseri	Glucan 1,6-α-glucosidase	3.2.1.70	α-1,6-glycosidic linkages	dexB
		Intracellular maltogenic amylase	3.2.1.-	α-1,4-glycosidic linkages	bbmA
		Oligo-1,6-glucosidase	3.2.1.10	α-1,6-glycosidic linkages	malL
F06	*Lactobacillus iners*	Glucan 1,6-α-glucosidase	3.2.1.70	α-1,6-glycosidic linkages	dexB
		Intracellular maltogenic amylase	3.2.1.-	α-1,4-glycosidic linkages	bbmA
		Oligo-1,6-glucosidase	3.2.1.10	α-1,6-glycosidic linkages	malL
		Pullulanase	3.2.1.41	α-1,6-glycosidic linkages	amyX_1; amyX_2; amyX_3
F08	*Bifidobacterium lacrimalis*	Glycogen operon protein GlgX	3.2.1.-	α-1,6-glycosidic linkages	glgX
		Oligo-1,6-glucosidase	3.2.1.10	α-1,6-glycosidic linkages	malL_2
		Neopullulanase 2	3.2.135	α-1,4-, and α-1,6-glycosidic linkages	tvaII
		Oligo-1,6-glucosidase 1	3.2.1.10	α-1,6-glycosidic linkages	malL_1
		Pullulanase	3.2.1.41	α-1,6-glycosidic linkages	amyX; pulA
	Bifidobacterium vaginale	Neopullulanase	3.2.135	α-1,4-, and α-1,6-glycosidic linkages	nplT
		Oligo-1,6-glucosidase	3.2.1.10	α-1,6-glycosidic linkages	malL
		Pullulanase	3.2.1.41	α-1,6-glycosidic linkages	pulA
		Trehalose synthase/amylase TreS	3.2.1.1	α-1,4-glycosidic linkages	treS
	Lactobacillus crispatus	Glucan 1,6-α-glucosidase	3.2.1.70	α-1,6-glycosidic linkages	dexB
		Intracellular maltogenic amylase	3.2.1.-	α-1,4-glycosidic linkages	bbmA
		Oligo-1,6-glucosidase	3.2.1.10	α-1,6-glycosidic linkages	malL_1; malL_2
		Pullulanase	3.2.1.41	α-1,6-glycosidic linkages	amyX
F12	*Bifidobacterium vaginale*	Glycogen operon protein GlgX	3.2.1.-	α-1,6-glycosidic linkages	glgX
		Neopullulanase	3.2.135	α-1,4-, and α-1,6-glycosidic linkages	nplT
		Oligo-1,6-glucosidase	3.2.1.10	α-1,6-glycosidic linkages	malL
		Pullulanase	3.2.1.41	α-1,6-glycosidic linkages	pulA
		Trehalose synthase/amylase TreS	3.2.1.1	α-1,4-glycosidic linkages	treS
	Lactobacillus crispatus	Glucan 1,6-α-glucosidase	3.2.1.70	α-1,6-glycosidic linkages	dexB
		Intracellular maltogenic amylase	3.2.1.-	α-1,4-glycosidic linkages	bbmA
		Oligo-1,6-glucosidase	3.2.1.10	α-1,6-glycosidic linkages	malL
		Pullulanase	3.2.1.41	α-1,6-glycosidic linkages	amyX
	Lactobacillus jensenii	Glucan 1,6-α-glucosidase	3.2.1.70	α-1,6-glycosidic linkages	dexB
		Intracellular maltogenic amylase	3.2.1.-	α-1,4-glycosidic linkages	bbmA
		Oligo-1,6-glucosidase	3.2.1.10	α-1,6-glycosidic linkages	malL

[a]*Bifidobacterium vaginale* is referenced from the GTDBTK taxonomy database and reported here but listed as G*ardnerella vaginalis* in the NCBI taxonomy database.

[a]*Bifidobacterium lacrimalis* is referenced from the GTDBTK taxonomy database and reported here but listed as *Peptoniphilus lacrimalis in* the NCBI taxonomy database.

Vaginal amylases are produced by both the host and bacteria in the vaginal community

The amylase assays and metagenomic analysis strongly suggest there are multiple amylases produced by bacteria in the vagina. Next, we used proteomics to determine which of the putative amylases were expressed *in vivo*. To do this, proteins were extracted from CVM samples, analyzed by LC-MS/MS, and identified by searching identified protein sequences against the protein annotations for MAGs produced in this study in addition to the human proteome. Based on hierarchical clustering analysis of all of the human and bacterial proteins that were detected, we observed that the proteomes of F02 and F06 were similar but distinct from F08 and F12, and F08 and F12 differed from one another (Figure C.2). Similarly, the bacterial proteins that were highly abundant in CVM were associated with individual MAGs identified in the donors (Figure C.3).

We identified four different proteins in the proteomes of our samples that matched translated proteins sequences identified in the corresponding bacterial metagenomes. The bacterial proteins identified included α-amylase from *B. vaginale* (F02), intracellular maltogenic amylase from *L. crispatus* (F08 and F12), oligo-1,6-glucosidase from *L. iners* (F06), and pullulanase from *L. iners* (F06), *B. vaginale* (F02 and F08), and *B. lacrimalis* (F08). Human α-amylases were identified at varying levels in all samples; however, we were unable to distinguish between the salivary and pancreatic form. Normalized protein intensities were higher for the bacterial amylases than human amylases detected in CVM from donors F02, F06, and F12. The opposite was true for donor F08.

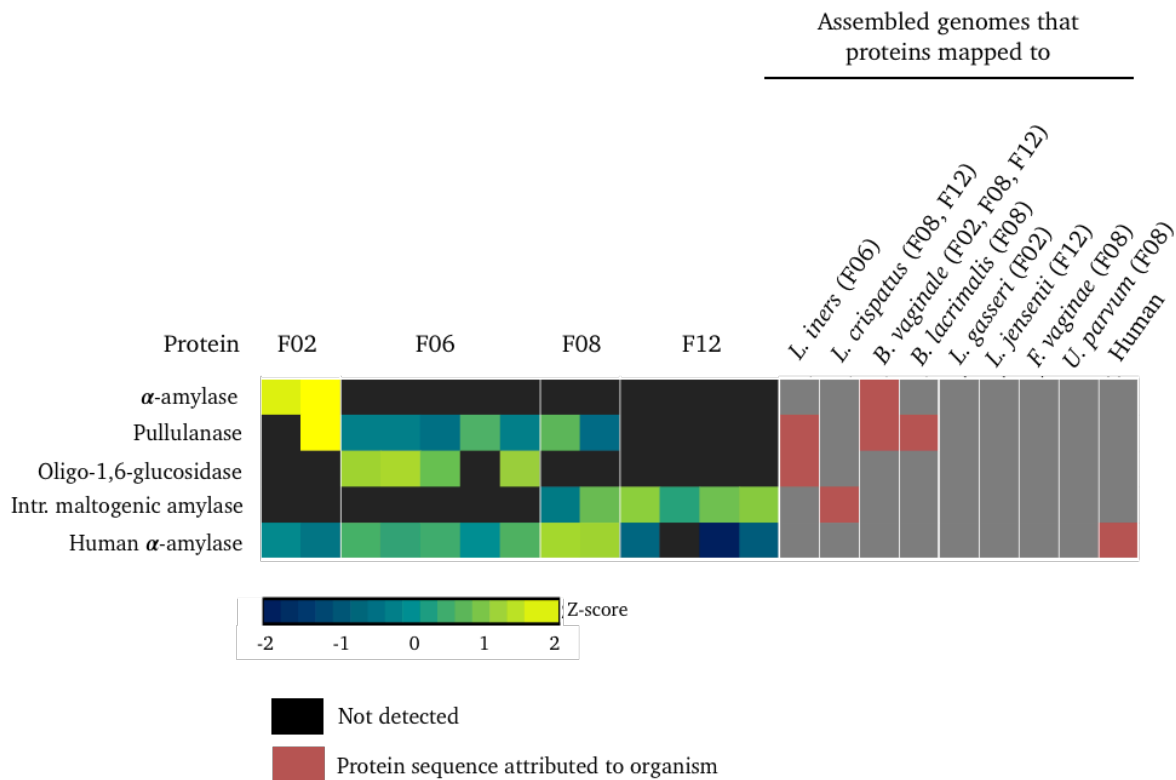


Figure 3.4 Amylase proteins expressed in vaginal fluids. Four putative amylases from bacteria in addition to human alpha-amylase were detected in vaginal fluids using LC-MS/MS. The relative abundance of the protein (listed on the far left) is shown in individual heatmaps below each sample. Relative abundances are indicated by the z-scores shown in the legend below the plot. Black squares indicate that the protein was not detected in a sample. The set of dark gray columns to the far-right show for which organism within the vaginal community the protein sequence was found in and is indicated by the dark coral squares. Replicates were performed based on sample volume.

DISCUSSION

Alpha-amylase is a key enzyme in the vaginal ecosystem that depolymerizes glycogen into simpler oligosaccharides, which are presumed to be fermented by *Lactobacillus* to produce lactic acid [20,45-47]. Previous research has measured amylase activity in vaginal fluids [20,48,49] and looked at the expression of amylases in other genital tract tissues [50,51]. To our knowledge, this is the first study that has used methods to identify the types and sources of amylase in the vagina. We found five different amylases within the α-amylase family of proteins in vaginal fluids from a subset of women in our cohort. The four bacterial amylases matched protein annotations from the assembled metagenomes of *L. crispatus*, *L. iners*, *B. lacrimalis*, and *B. vaginale*. The fifth amylase that we detected was human α-amylase.

In each donor, we detected a combination of bacterial and human amylases. This finding could explain why differences in amylase activity was not reflected in the types of vaginal communities that were observed in our cohort. A previously published study reported that α-amylase activity was lower in women with bacterial vaginosis (BV), a condition in which *Lactobacillus* abundances are low, when compared to women without BV [49]. Of the women in our cohort, only two of them had communities that were not dominated by vaginal lactobacilli. Although this is a small number in comparison, there was still no difference in amylase activity. Nasioudis *et al*. specifically measured human pancreatic α-amylase in vaginal fluids and it could be that differences in human α-amylase activity correlate with different types of vaginal communities. However, the assay that we used is not specific to any one amylase and we measured any amylase that could be detected in CVM, reflecting both bacterial and human amylases. This, along with differences in substrate specificity and catalytic rate constants, might have also contributed to the variability observed in glycogen measurements when the corresponding amylase activity measured below 0.5 U/mg protein.

We identified amylase proteins in vaginal fluids that could hydrolyze either α-1,4- or α-1,6-glycosidic bonds. Alpha-amylase and intracellular maltogenic amylase hydrolyze α-1,4-glycosidic bonds suggesting that it can break linear glucose polymers. Meanwhile, oligo-1,6-glucosidase and pullulanase hydrolyze α-1,6-glycosidic bonds indicating the ability to break branched glucose linkages. Due to the combination of bacterial and human amylases, each sample except for F12 contained enzymes that could perform both reactions. This suggests that within a vaginal community, multiple enzymes might be employed to metabolize glycogen by first cleaving branched glucose chains followed by the hydrolysis of linear glucose monomers. In F12 we only detected amylases that break linear glucose

chains, which included intracellular maltogenic amylase from *L. crispatus* and human α-amylase. Interestingly, F12 had two to four times the amount of glycogen measured in samples F02, F06, and F08. Perhaps this is attributed to the inability to completely break down glycogen due to the lack of an enzyme that can hydrolyze branched glucose chains in glycogen. Because of our limited sample size, we cannot address this further. To provide a mechanistic understanding of the conditions that are required to depolymerize glycogen *in vivo*, future studies might compare the levels of glycogen in the vagina with the presence or absence of amylase proteins that hydrolyze α-1,4- and α-1,6-glycosidic bonds.

Degradation of glycogen by amylase occurs in the extracellular environment and the resulting resources (i.e. maltose, maltotriose) are 'common goods' that are available to all species in the vaginal community. Here we identified multiple bacterial amylases within a sample. Coinciding with differences in vaginal community composition, the kinds of bacterial amylases differed between women. Given the temporal changes that are observed in vaginal community composition [45,52,53], we expect that the kinds of bacterial amylases present will also change during these times. Differences in amylase proteins and changes in these proteins over time likely influence the pool of resources that are broadly shared and support the growth of many members of a vaginal community. Because *Lactobacillus spp*. tend to dominate the communities of most reproductive age women in rank and abundance, it is possible that lactobacilli are more effective at competing for and sequestering these resources in the vagina.

The question still remains whether *Lactobacillus spp*. metabolize glycogen or simpler sugars in the vagina. Until recently, studies have suggested that *Lactobacillus* cannot metabolize glycogen because when tested, lactobacilli did not grow on media containing glycogen in the laboratory [20,54-56]. van de Veer *et al.* cultured *L. crispatus* isolates in media supplemented with glycogen [25]. Those that could not grow were reported to have a mutation in a pullulanase type I gene [25]. We did not detect pullulanase from *L. crispatus* in the proteins that were isolated from vaginal fluids. However, we did detect pullulanase from *L. iners*. Two other *Lactobacillus* amylases were also identified (i.e. oligo-1,6-glucosidase and intracellular maltogenic amylase). These results would imply that lactobacilli do have the functional capability to metabolize glycogen *in vivo*.

Previous studies have detected amylase genes in vaginal microbes [23,57,58], but until now, it has not been confirmed whether these proteins are actually expressed in the vagina. While we detected multiple

amylases in the vaginas of four women, we did not determine whether or not they actually break down glycogen. Future studies might aim to biochemically characterize these enzymes to understand their substrate specificities, the factors that regulate their expression, and how the metabolic products (oligosaccharides, maltose and glucose) are partitioned among community members.

LITERATURE CITED

[1] Boskey ER, Telsch KM, Whaley KJ, et al. Acid production by vaginal flora in vitro is consistent with the rate and extent of vaginal acidification. Infect Immun 1999;67:5170–5.

[2] Boskey ER, Cone RA, Whaley KJ, et al. Origins of vaginal acidity: high D/L lactate ratio is consistent with bacteria being the primary source. Hum Reprod 2001;16:1809–13.

[3] O'Hanlon DE, Moench TR, Cone RA. In vaginal fluid, bacteria associated with bacterial vaginosis can be suppressed with lactic acid but not hydrogen peroxide. BMC Infect Dis 2011;11:251.

[4] O'Hanlon DE, Moench TR, Cone RA. Vaginal pH and Microbicidal Lactic Acid When Lactobacilli Dominate the Microbiota. PLoS ONE 2013;8:e80074–8.

[5] Antonio MA, Hawes SE, Hillier SL. The identification of vaginal *Lactobacillus* species and the demographic and microbiologic characteristics of women colonized by these species. J Infect Dis 1999;180:1950–6.

[6] Burton JP, Cadieux PA, Reid G. Improved Understanding of the Bacterial Vaginal Microbiota of Women before and after Probiotic Instillation. Appl Environ Microbiol 2003;69:97–101.

[7] Zhou X, Bent SJ, Schneider MG, et al. Characterization of vaginal microbial communities in adult healthy women using cultivation-independent methods. Microbiology 2004;150:2565–73.

[8] Zhou X, Brown CJ, Abdo Z, et al. Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women. ISME J 2007;1:121–33.

[9] Zhou X, Hansmann MA, Davis CC, et al. The vaginal bacterial communities of Japanese women resemble those of women in other racial groups. FEMS Immunol Med Microbiol 2010;58:169–81.

[10] Ravel J, Gajer P, Abdo Z, et al. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci USa 2011;108 Suppl 1:4680–7.

[11] Brotman RM, Ravel J, Cone RA, et al. Rapid fluctuation of the vaginal microbiota measured by Gram stain analysis. Sex Trans Infect 2010;86:297–302.

[12] Gajer P, Brotman RM, Bai G, et al. Temporal Dynamics of the Human Vaginal Microbiota. Sci Transl Med 2012;4:132ra52–2.

[13] Srinivasan S, Liu C, Mitchell CM, et al. Temporal Variability of Human Vaginal Bacteria and Relationship with Bacterial Vaginosis. PLoS ONE 2010;5:e10197–8.

[14] Keane FEA, Ison CA, Taylor-Robinson D. A longitudinal study of the vaginal flora over a menstrual cycle. Int J STD AIDS 1997; 8:1–6.

[15] Farage M, Maibach H. Lifetime changes in the vulva and vagina. Arch Gynecol Obstet 2005;273:195–202.

[16] Mirmonsef P, Hotton AL, Gilbert D, et al. Free glycogen in vaginal fluids is associated with *Lactobacillus* colonization and low vaginal pH. PLoS ONE 2014;9:e102467.

[17] Cruickshank R, Sharman A. The Biology of the Vagina in the Human Subject. BJOG 1934:208–26.

[18] Smythe C, Cohen P. The discovery of glycogenin and the priming mechanism for glycogen biogenesis. Eur J Biochem 1991;200:625–31.

[19] Roberts PJ, Whelan WJ. The mechanism of carbohydrase action. 5. Action of human salivary alpha-amylase on amylopectin and glycogen. Biochem J 1960;76:246–53.

[20] Spear GT, French AL, Gilbert D, et al. Human α-amylase Present in Lower-Genital-Tract Mucosal Fluid Processes Glycogen to Support Vaginal Colonization by *Lactobacillus*. J Infect Dis 2014;210:1019–28.

[21] Merritt AD, Rivas ML, Bixler D, et al. Salivary and Pancreatic Amylase: Electrophoretic Characterizations and Genetic Studies. Am J Hum Genet 1973;25:510–22.

[22] Merritt AD, Karn RC. The human amylases. Adv Hum Genet 1977; 135–234.

[23] Huffman RD, Nawrocki LD, Wilson WA, et al. Digestion of glycogen by a glucosidase released by Trichomonas vaginalis. Exp Parasitol 2015;159:151–9.

[24] Pandey A, Nigam P, Soccol CR, et al. Advances in microbial amylases. Biotechnol Appl Biochem 2000;31:135–52.

[25] van der Veer C, Hertzberger RY, Bruisten SM, et al. Comparative genomics of human *Lactobacillus crispatus* isolates reveals genes for glycosylation and glycogen degradation: implications for in vivo dominance of the vaginal microbiota. Microbiome 2019;7:49.

[26] MacGregor EA, Janeček Š, Svensson B. Relationship of sequence and structure to specificity in the α-amylase family of enzymes. Biochim Biophys Acta 2001;1549:1–20.

[27] Stam MR, Danchin EGJ, Rancurel C, et al. Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of α-amylase-related proteins. Protein Eng Des Sel 2006;19:555–62.

[28] Brotman RM, Ravel J, Cone RA, et al. Rapid fluctuation of the vaginal microbiota measured by Gram stain analysis. Sex Trans Infect 2010;86:297–302.

[29] Fashemi B, Delaney ML, Onderdonk AB, et al. Effects of feminine hygiene products on the vaginal mucosal biome. Microb Ecol Health Dis 2013;24:807–6.

[30] Vodstrcil LA, Twin J, Garland SM, et al. The influence of sexual activity on the vaginal microbiota and *Gardnerella vaginalis* clade diversity in young women. PLoS ONE 2017;12:e0171856–15.

[31] Nunn KL, Wang Y-Y, Harit D, et al. Enhanced Trapping of HIV-1 by Human Cervicovaginal Mucus Is Associated with *Lactobacillus crispatus*-Dominant Microbiota. mBio 2015;6:e01084–15–15.

[32] Yuan S, Cohen DB, Ravel J, et al. Evaluation of methods for the extraction and purification of DNA from the human microbiome. PLoS ONE 2012;7:e33865.

[33] Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 2011;27:2957–63.

[34] Callahan BJ, McMurdie PJ, Rosen MJ, et al. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Meth n.d.;13:581EP–.

[35] Allard G, Ryan FJ, Jeffery IB, et al. SPINGO: a rapid species-classifier for microbial amplicon sequences. BMC Bioinformatics 2015;16:1–8.

[36] Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome 2018; 6:158.

[37] Parks DH, Chuvochina M, Waite DW, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 2018;36:996–1004.

[38] Folch J, Lees M, Stanley GHS. A simple method for the isolation and purification of total lipides form animal tissues. J Biol Chem 2002;226:497–509.

[39] Nakayasu ES, Nicora CD, Sims AC, et al. MPLEx: a Robust and Universal Protocol for Single-Sample Integrative Proteomic, Metabolomic, and Lipidomic Analyses. mSystems 2016;1:558–14.

[40] Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnol 2008;26:1367–72.

[41] Schwanhäusser B, Busse D, Li N, et al. Global quantification of mammalian gene expression control. Nature 2011;473:337–42.

[42] Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. 2019:1–2673.

[43] Ward JH. Hierarchical Grouping to Optimize an Objective Function. J Am Stat Assoc 1963;58:236–44.

[44] Jeong H, Arif B, s GC-AX, et al. Horizontal gene transfer in human- associated microorganisms inferred by phylogenetic reconstruction and reconciliation. Sci Rep 2019:1–18.

[45] Nunn KL, Forney LJ. Unraveling the Dynamics of the Human Vaginal Microbiome. Yale J Biol Med 2016;89:331–7.

[46] Witkin SS, Linhares IM. Why do lactobacilli dominate the human vaginal microbiota? BJOG 2016:1–6.

[47] Amabebe E, Anumba DOC. The Vaginal Microenvironment: The Physiologic Role of Lactobacilli. Front Med 2018;5:164–11.

[48] Skerlavay M, Epstein JA, Sobrero AJ. Cervical Mucus Amylase Levels in Normal Menstrual Cycles. Fert Steril 1968;19:726–30.

[49] Nasioudis D, Beghini J, Bongiovanni AM, et al. α-Amylase in Vaginal Fluid: Association With Conditions Favorable to Dominance of *Lactobacillus*. Reprod Sci 2015;22:1393–8.

[50] Skude G, Mårdh PA, Weström L. Amylases of the genital tract. I. Isoamylases of genital tract tissue homogenates and peritoneal fluid. Am J Obstet Gynecol 1976;126:652–6.

[51] Lee YS, Raju GC. The expression and localization of amylase in normal and malignant glands of the endometrium and endocervix. J Pathol 1988;155:201–5.

[52] Hickey RJ, Zhou X, Pierson JD, et al. Understanding vaginal microbiome complexity from an ecological perspective. Transl Res 2012;160:267–82.

[53] van de Wijgert JHHM, Borgdorff H, Verhelst R, et al. The Vaginal Microbiota: What Have We Learned after a Decade of Molecular Characterization? PLoS ONE 2014;9:e105998–10.

[54] STEWART-TULL DE. Evidence that vaginal lactobacilli do not ferment glycogen. Am J Obstet Gynecol 1964;88:676–9.

[55] Identity and glycogen-fermenting ability of lactobacilli isolated from the vagina of pregnant women. J Med Microbiol 1969;2:363–6.

[56] Martin R, Soberon N, Vaneechoutte M, et al. Characterization of indigenous vaginal lactobacilli from healthy women as probiotic candidates. Int Microbiol 2008;11:261–6.

[57] Macklaim JM, Fernandes AD, Di Bella JM, et al. Comparative meta-RNA-seq of the vaginal microbiota and differential expression by *Lactobacillus iners* in health and dysbiosis. Microbiome 2013;1:12.

[58] Štáfková J, Rada P, Meloni D, et al. Dynamic secretome of *Trichomonas vaginalis*: Case study of β-amylases. Mol Cell Proteomics 2018;17:304–20.

CHAPTER 4: THE POSTPARTUM VAGINAL MICROBIOME[4]

ABSTRACT

The composition of the vaginal microbiome changes dramatically following pregnancy. Vaginal bacterial communities during the postpartum period can be discriminated from those during pregnancy by having reduced abundances of *Lactobacillus* species with an increase in bacteria associated with bacterial vaginosis and increased α-diversity. This diverse community persists for some time post-delivery and has the potential to increase health risks in women. Currently, we know very little about this transitional period and the factors driving the observed changes. In this study vaginal fluid samples from 48 women were obtained during the first and third trimesters and postpartum. We characterized the species composition of vaginal communities at each time point and measured the levels of NGAL, p62, Hsp70, hyaluronan, EMMPRIN, MMP8, α-amylase, and lactic acid. We used linear mixed effects models to determine associations with vaginal compounds and the postpartum period. In accordance with previous studies we found that the vaginal microbiome postpartum is more diverse and has lower abundances of *Lactobacillus*, but increased abundances of *Streptococcus anginosus* and *Prevotella bivia*. As compared to samples from pregnant women, postpartum samples showed increased levels of hyaluronan and Hsp70 and decreased levels of both lactic acid isomers. We posit that lower concentrations of lactic acid combined with increased levels of Hsp70 enables diverse communities to persist during the postpartum period.

INTRODUCTION

Considerable changes in the composition of the vaginal microbiome occur immediately following pregnancy. During a healthy pregnancy, the vaginal microbiome is characterized by low bacterial species diversity [1,2], and is typically dominated by one or multiple species of *Lactobacillus* [3]. However, in the transition from pregnancy to the postpartum period, the composition of the vaginal microbiome shifts to lower abundances of lactobacilli and increased microbial diversity [4,5]. *Lactobacillus* species are thought to help maintain good vaginal health by producing lactic acid [6,7]. Lactic acid is associated with many protective mechanisms in the vagina including the suppression of pathogens [8-10], promoting DNA repair in vaginal epithelial cells [11], and enhancing innate and adaptive immunity [12]. Thus, this transition in community composition during the postpartum period could have profound effects on women's health. Currently, we know very little about the composition

[4] This chapter is to be submitted for publication as: Nunn KL, Witkin SS, Schneider GM, Boester A, Nasioudis D, Minis E, Gliniewicz K, Forney LJ. The postpartum vaginal microbiome.

of the vaginal microbiome during the postpartum stage. Likewise, the mechanisms driving these observed changes are poorly understood.

Only a few cross-sectional [13] and longitudinal studies [4,5,14] have compared the vaginal microbiome during and shortly after pregnancy. During this transition, there is reportedly an overall decrease in *Lactobacillus* [4,13,14] with a concomitant increase in diverse taxa including but not limited to Clostridia (*Peptoniphilus* and *Anaerococcus*), Bacteroidia, and bacteria associated with bacterial vaginosis (BV; including *Prevotella*, *Veillonella*, *Porphyromonas*, and *Megasphaera*) [4,5,13,14]. BV is a clinical syndrome typified by an overgrowth of diverse anaerobic bacteria and a reduction of vaginal lactobacilli. If it is symptomatic BV can be accompanied by inflammation, malodorous vaginal discharge, itching and burning (reviewed in [15]). Moreover, BV is often associated with increased susceptibility to sexually transmitted infections and other adverse reproductive outcomes [16-18]. The increase in diverse taxa during the postpartum period is paralleled by increases in α-diversity [4,5,13,14]. Doyle *et al*. obtained vaginal samples from a cohort of Malawi women anywhere from five to 583 days post-delivery and demonstrated that *Lactobacillus spp*. were present in less than a third of the women [19,20]. Further, postpartum vaginal communities resembled those of community state type (CST) III and IV as defined by Ravel *et al*. [19,20]. CST-III consists of communities in which *Lactobacillus iners* is dominant. On the other hand, CST-IV vaginal communities are heterogenous in composition with low proportions of lactobacilli and an assortment of anaerobic bacteria. Doyle and colleagues found the abundances of *L. iners* increased with time throughout the postpartum stage [19,20]. In comparison to the other commonly found vaginal lactobacilli, communities dominated by *L. iners* tend to be more diverse and less stable [21-23] as they often transition to communities resembling CST-IV [24]. These characteristics have led to a debate as to whether or not the prevalence of *L. iners* in the vaginal microbiome is beneficial to women's health [21]. It has been reported that the observed changes in the bacterial community composition can occur as early as the onset of labor [14,25] and may persist for up to one year [4].

While changes in vaginal community composition are occurring after delivery, simultaneous alterations in host immune response that include inflammation and suppression of cell-mediated immunity are observed [26,27]. To complicate things further, it could take up to four months for normal immune function to be restored [26]. The combination of diminished abundances of lactobacilli, increased abundances of BV-associated bacteria and diversity, inflammation, and inadequate immune function amplifies health risks in women at times following delivery. To mitigate these risks, we need to better understand the transition in bacterial communities in the vagina from pregnancy to postpartum, and

what causes changes in vaginal community composition to occur. Therefore, in this study, we obtained vaginal samples from 48 women during early and late pregnancy, and the postpartum stage. We determined the species composition of vaginal communities at each time point and evaluated associations with the postpartum stage and levels of lactic acid, neutrophil gelatinase-associated lipocalin (NGAL), p62, inducible 70-kDa heat shock protein (Hsp70), hyaluronan, extracellular matrix metalloproteinase inducer (EMMPRIN), metalloproteinase-8 (MMP), and α-amylase. We showed that postpartum samples had elevated levels of Hsp70 and lower levels of both D- and L-lactic acid. Low concentrations of lactic acid lessen the degree of protection conferred to the vagina. This combined with Hsp70, which inhibits autophagy, likely enables diverse communities to persist during the postpartum period.

METHODS AND MATERIALS

Study participants and sample collection

The prospective study included pregnant women who were seen for outpatient obstetrics service at Weill Cornell Medicine. Exclusion criteria were signs or symptoms of a gynecological disorder or infection at the time of examination, multifetal gestation, antibiotic treatment in the previous four weeks, vaginal bleeding or the inability to give informed consent. The study was approved by the Institutional Review Board at Weill Cornell Medicine and all subjects gave informed, written consent. In total, 48 women were enrolled and provided three vaginal samples each. Samples of vaginal secretions of women during the first (≤ 12 weeks) and third trimesters (28 – 38 weeks), and 28 – 45 days postpartum from the posterior vagina were obtained with a cotton swab. The swabs were vigorously shaken in a sterile tube containing 1 ml of sterile 1X phosphate-buffered saline (PBS), centrifuged, and aliquots of the supernatant were stored at -80°C.

Compound measurements

Vaginal levels of D- and L-lactic acid isomers were quantitated by colorimetric assays using the EnzyChrom D-lactic acid and L-lactic acid kits (BioAssay Systems, Haywood, CA). The levels of MMP-8 (R&D Systems, Minneapolis, MN), EMMPRIN (R&D Systems), NGAL (R&D Systems), hyaluronan (R&D Systems), α-amylase (human pancreatic α-amylase, Abcam, Cambridge, United Kingdom), Hsp70 (R&D Systems), and p62 (Enzo Life Sciences, Farmingdale, NY) were determined using commercially available ELISA kits. Hsp70 and p62 were measured from the supernatant of lysed epithelial cells while all of the other compounds were measured directly from vaginal secretions. To obtain the epithelial cell lysate, vaginal secretions were centrifuged, and the epithelial cell pellet was

immediately lysed in a detergent containing a protease inhibitor cocktail. The lysate was then centrifuged, and the supernatant was stored in aliquots at -80 ºC. Total protein levels in samples were measured using a colorimetric assay (Thermo-Fisher Scientific, Waltham, MA). Concentrations were calculated from a standard curve that was generated in parallel to test samples and converted to millimolar, units per milliliter, or picograms per milliliter. When applicable, measurements were normalized by the total amount of protein in the sample and reported in picograms, nanograms, or units per microgram of total protein. All assays were performed by staff blinded to all clinical information.

Determining vaginal community composition

The species composition of vaginal microbial communities was determined by classifying partial 16S rRNA gene sequences as previously described [28]. In brief, total genomic DNA was extracted from 250 µl of vaginal swabs stored in 1X PBS using chemical and mechanical lysis and purified using QIAamp DNA mini kits (Qiagen). Genomic DNA concentrations were determined using the Quant-iT™ PicoGreen™ dsDNA assay kit (Invitrogen). For amplicon sequencing, the V1-V3 16S rRNA gene regions were amplified using a two-step PCR protocol, first amplifying the gene region using universal primers 27F and 534R, and then adding sample barcodes and sequence adapters. Amplicons were sequenced using the Illumina MiSeq at the University of Idaho. Forward and reverse reads were paired using FLASH [29], processed through DADA2 v 1.12.1 [30] to identify distinct sequences, and the distinct sequences were classified to genus and species level using SPINGO [31].

A total of 358 taxa were identified and these data were used to analyze rank abundance of taxa and measure α-diversity. For subsequent analyses including visualizing community composition, we created a smaller dataset of taxon abundance for 57 taxa. Fifty-six of the 358 taxa were present at a minimum of 1% in at least two individuals. Using this filter, the remaining taxa (302) were considered rare with respect to sampling and therefore grouped into the "other" category. The corresponding relative abundance data are reported in Table D.1.

Bioinformatic and statistical analyses

All analyses were performed using R v 3.6.0 [32]. In addition to the base packages, the following packages were used for organizing data, performing analyses, and to produce figures: cluster [33], e1071 [34], FSA [35], ggplot2 [36], ggpubr [37], lme4 [38], lmerTest [39], multcomp [40], plyr [41], tidyr [42], vegan [43]. To determine what kinds of communities were present in this cohort, we performed complete-linkage hierarchical clustering on alt-Gower distances computed from taxon

relative abundance data. We then used silhouette information to identify ten clusters, numbered one through ten. Alpha diversity was measured using the Shannon and Simpson diversity indices.

Linear mixed effects models were used to model the means of response variables between pregnancy stage to evaluate significant differences. From this point forward, stage refers to first and third trimester, and postpartum. In these models, subject was included as a random effect because the same subjects were represented in each stage. The various response variables included bacterial species abundances, measures of α-diversity, and compounds measured in vaginal fluids. Statistical comparisons were performed by testing general linear hypotheses and multiple comparisons of the means using Tukey's test. If the linear mixed effects model resulted in a singular fit, meaning there was no variation explained by including subject as a random effect, we used a Kruskal-Wallis rank sum test to evaluate significant differences in means of the response variables between pregnancy stage. A significant Kruskal-Wallis test was followed with post-hoc analysis of multiple comparisons using a Dunn's test with Bonferroni adjustment. To test associations with changes in Shannon and Simpson diversity, we created a base linear model for both measures including the compounds that differed significantly by stage and stage as fixed effects. Subject was included as a random effect to account for differences in subject across stage. Stepwise linear regression was then used to determine which variables should be included in the model. A summary for each final model was generated to determine associations. Significance was set to the level of $P = 0.05$.

RESULTS

Forty-eight women were enrolled to participate in a study to elucidate factors associated with changes in the composition of the vaginal microbiome from pregnancy to the postpartum stage. All of the women enrolled in this study had full-term pregnancies. The demographics of the study population are shown in Table 4.1.

Comparison of the vaginal microbiome from pregnancy to postpartum

To characterize changes in the vaginal microbiome that occur following pregnancy, we collected vaginal swab samples during the first trimester, third trimester, and following delivery. The composition of bacterial communities in these samples were determined by sequencing partial 16S rRNA gene amplicons. Samples that did not have at least 3,000 sequence reads were discarded from further analysis. This resulted in 47, 45, and 34 samples for the first trimester, third trimester, and postpartum stages, respectively. Thirty-two women had all three timepoints, 14 women had two

timepoints, and two women had only one time point. The relative proportions of 57 taxa identified in each stage of pregnancy are shown in Fig D.1. The findings of this analysis suggest that vaginal communities transition from being dominated by species of *Lactobacillus* during the first and third trimester to having low proportions of lactobacilli and more diverse communities during postpartum.

Table 4-1. Study participant demographics

Characteristic	No. of women or %	Mean	SD[a]
Age at delivery (y)	48	33.44	3.61
BMI (kg/m^2)	44	22.30	2.58
Weeks of gestation	48	39.54	0.96
Race	48		
white	56.25%		
black	2.08%		
Asian	12.50%		
Hispanic	4.17%		
other	25.00%		
Gravida[b]	48	2.15	1.09
First Birth	15		
Prior Births	33		
Term	23		
Pre-term	2		
Spontaneous miscarriage	11		

[a] SD = standard deviation

[b] Gravida = the total number of confirmed pregnancies

Next, we wanted to determine the kinds of communities that were present in the women of this cohort at each stage. Thus, we grouped communities based on similarities and differences using complete linkage hierarchical clustering. In total, we identified ten clusters of community types. Figure 4.1 shows a stacked bar chart where samples were separated by cluster and pregnancy stage. We found the kinds of vaginal communities to be similar between the first and third trimesters. Roughly 79% of the first trimester samples and 78% of the third trimester samples had vaginal communities in which *Lactobacillus spp.* were dominant. In contrast, postpartum vaginal communities were more heterogenous in composition. Approximately 77% of the postpartum vaginal communities were instead typified by low proportions of *Lactobacillus spp*. and a diverse array of bacteria including *G. vaginalis*, *Prevotella spp.*, *Streptococcus spp*., and others.

To visualize the distribution of taxa and compare diversity between stage, we first prepared rank abundance profiles for the top 20 taxa in each stage and measured skewness of these data (Figure 4.2a). Skewness measures the degree to which a sample distribution varies from normal. Compositional microbiome data typically have a long-tailed distribution and are positively skewed. This results from communities having few species that are highly abundant and many that are present at low abundances. Third trimester samples had the highest skew (15.84) followed by first trimester (13.44) and postpartum samples (9.27). This suggests postpartum samples had a distribution of taxa that were more even than the first and third trimester samples. Next, we measured α-diversity using Shannon and Simpson diversity indices and used linear mixed effects models to determine whether α-diversity differed between stage (Figure 4.2b). Shannon diversity increased from first to third trimester and again to postpartum, simply because in each stage there were more species present. There were 44 species found in first trimester samples, 112 in third trimester samples, and 225 in postpartum samples. Measures of Simpson diversity show a similar trend. We found no significant change in α-diversity from early to late pregnancy. There was a significant increase in Shannon ($p < 0.001$, first – postpartum; third – postpartum) and Simpson diversity indices ($p < 0.001$, first – postpartum; third – postpartum) in postpartum communities.

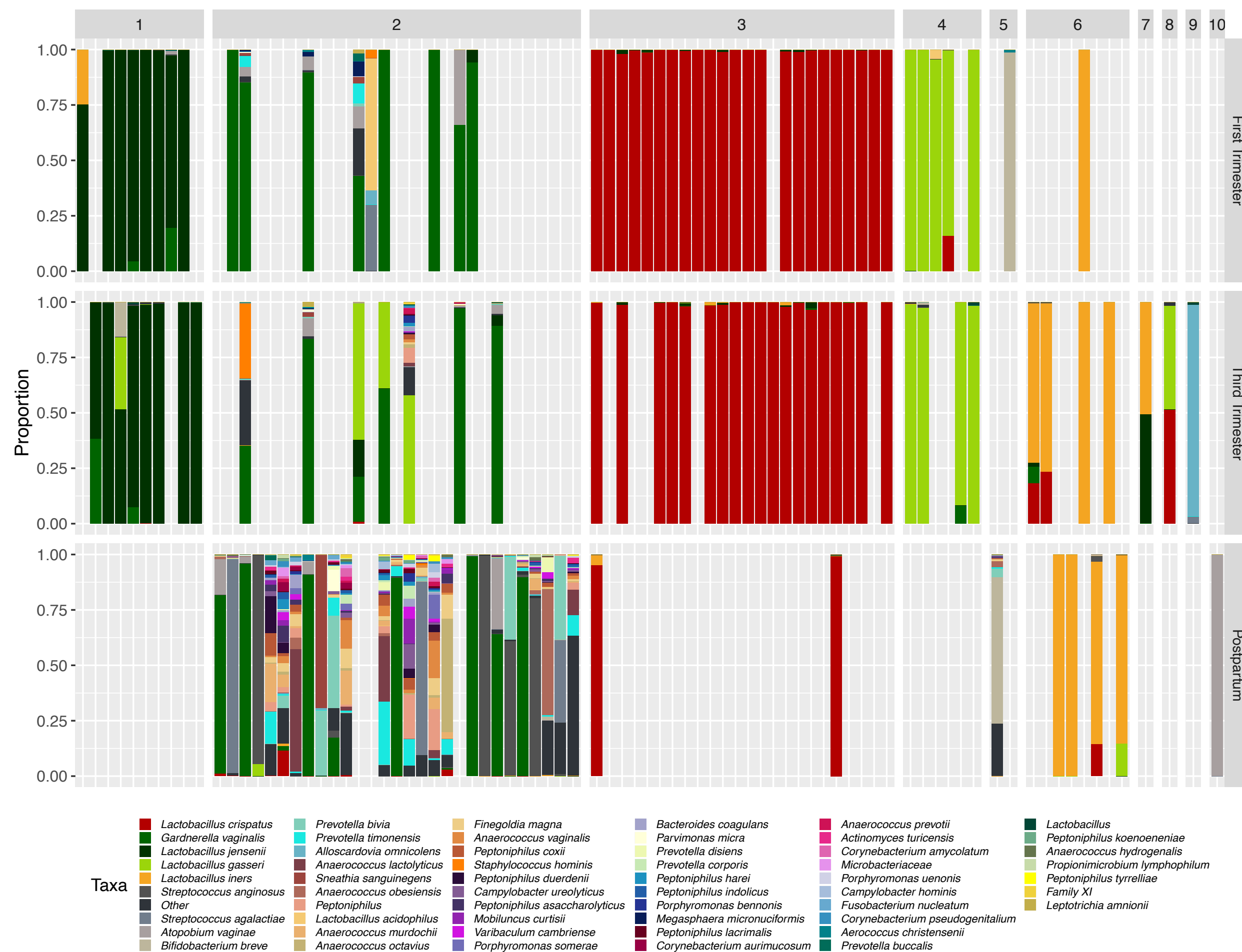


Figure 4.1. Relative proportions of bacteria in the vaginal communities of 48 pregnant women separated by stage and cluster. The stacked bars represent the proportions of bacterial taxa within one sample. Bars are separated by the pregnancy stage in which the sample was collected (right headings), and the cluster a sample was grouped into (top headings). Clusters were determined using complete-linkage hierarchical clustering of alt Gower distances calculated from the bacterial abundance data. Taxa colors are indicated in the legend below the figure. "Other" represents the sum of all bacterial taxa that were not present at 1% in at least two women.

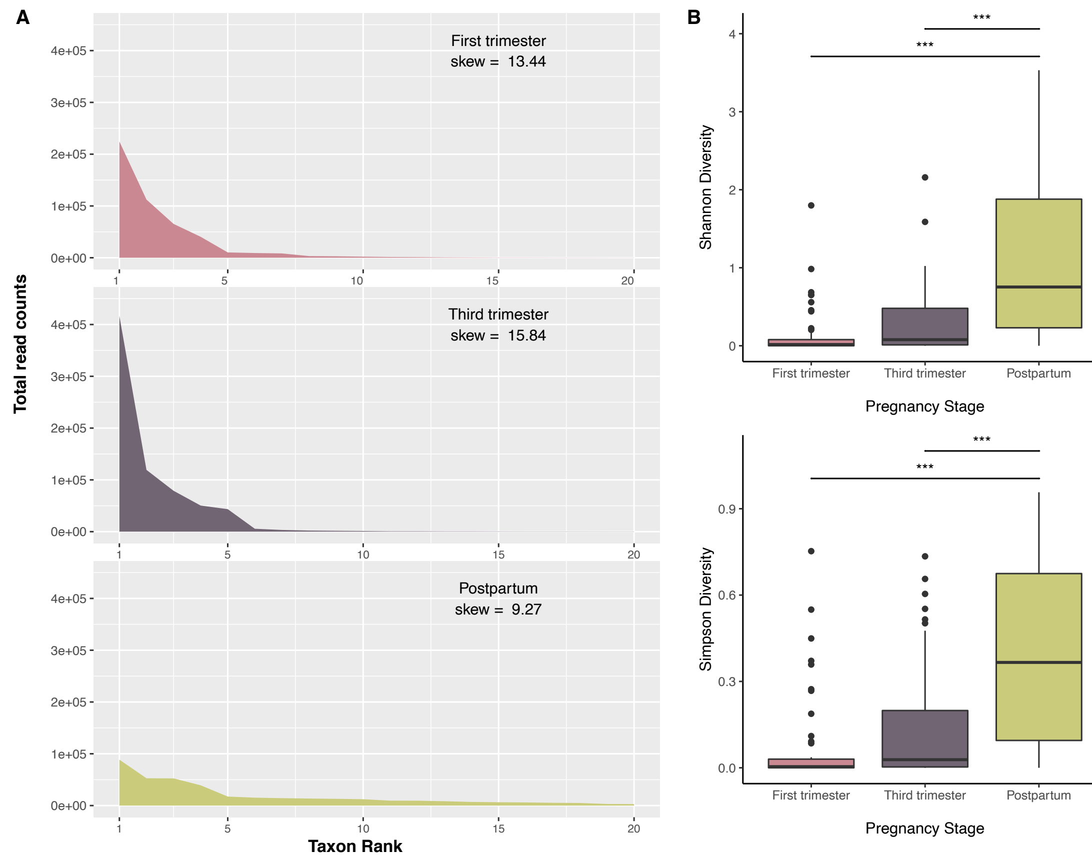


Figure 4.2. Rank abundance and α-diversity profiles of samples collected during the first trimester, third trimester, and postpartum. Panel A shows the counts of the top ranked taxa among the first trimester, third trimester, and postpartum samples. The skew for each set of data was calculated using the R package e1071 (v. 1.7.2) and is indicated within the plot. Skewness represents the degree to which a sample distribution varies from a normal distribution. Panel B shows the distribution of Shannon and Simpson diversity measures for each stage. The thick band within the box represents the median for the data and the lower and upper boxes represent the 25th and 75th quartiles, respectively. The upper and lower whiskers represent the 95% confidence interval for the median. Individual dots above represent outliers that were not removed. Differences in group means were evaluated using hypothesis testing of linear mixed effects models in R. Statistical significance is indicated above the lines (“***” = $p < 0.001$).

Changes in key taxa from pregnancy to postpartum

To quantify the change in *Lactobacillus* during pregnancy and afterwards, we used linear mixed effects models to model the means of *Lactobacillus* relative proportions between stages and controlled for variation within subject. The distribution of relative abundances of *Lactobacillus* by stage are shown in Figure 4.3. There was no change in the mean relative abundance of *Lactobacillus* over gestation. However, there was a sharp decrease in the abundances of lactobacilli during the postpartum stage (p < 0.001, for both comparisons: first - postpartum and third - postpartum). Next, we evaluated whether there were significant changes in individual lactobacilli and other top-ranked taxa (Table 4.2, Table D.2). We found a significant decrease in the relative abundances of *L. crispatus* (p < 0.001, first – postpartum, third – postpartum), *L. jensenii* (p < 0.01, first – postpartum; p < 0.001, third – postpartum), and *L. gasseri* (p < 0.01, third – postpartum), but not *L. iners*. Postpartum communities had significantly higher abundances of *Streptococcus anginosus* (p < 0.001, first – postpartum, third – postpartum) and *Prevotella bivia* (p < 0.001, first – postpartum, third – postpartum). *Streptococcus agalactiae* was present at higher abundances in a few samples postpartum but overall these differences were not significant. Likewise, there was relatively no change in *G. vaginalis*, *Atopobium vaginae*, or *Bifidobacterium breve* across stages.

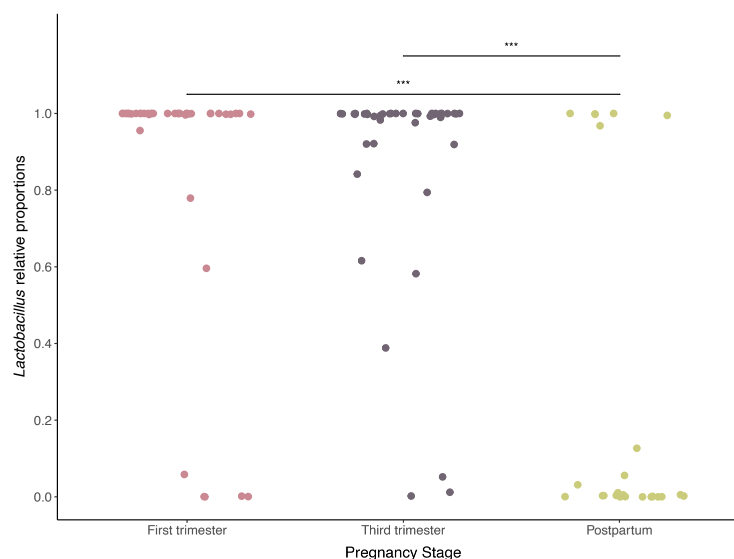


Figure 4.3. Differences in *Lactobacillus* species by stage. The dot plot shows the distribution of *Lactobacillus* relative proportions for the first trimester, third trimester, and postpartum. Each dot is one sample from a woman. Differences in group means were evaluated using hypothesis testing of linear mixed effects models in R. Statistical significance is indicated above the lines ("***" = $p < 0.001$).

Table 4-2. Mean relative proportions for the most abundant taxa in each stage. The following 10 taxa below were identified by taking the top ten ranked taxa for each stage, collapsing them together, and keeping only those greater than 1% in all samples.

	Mean relative proportions				P values for comparison of the means[b]		
Taxa	**First Trimester (N [a] = 47)**	**Third Trimester (N = 45)**	**Postpartum (N = 34)**	**All Samples (N = 126)**	**P value [c]**	**First - Postpartum[d]**	**Third - Postpartum[d]**
Lactobacillus crispatus	49.1	44.0	6.7	35.8	**0.00**	***	***
Gardnerella vaginalis	15.0	10.0	18.6	14.1	0.29		
Lactobacillus jensenii	16.2	17.4	0.0	12.2	**0.00**	**	***
Lactobacillus gasseri	10.2	13.9	0.6	8.9	**0.01**	·	**
Lactobacillus iners	2.7	8.9	11.0	7.2	0.21		
Streptococcus anginosus	0.0	0.0	10.1	2.7	**0.00**	***	***
Streptococcus agalactiae	0.6	0.1	6.2	1.9	0.11		
Atopobium vaginae	1.2	0.3	4.6	1.8	0.35		
Bifidobacterium breve	2.1	0.4	2.0	1.5	0.69		
Prevotella bivia	0.0	0.0	4.6	1.3	**0.00**	***	***

[a] N = the number of samples/observations in each category

[b] The asterisks correspond to the following levels of significance for the P values: '.', P < 0.1; *, P < 0.05; **, P < 0.01, ***, P < 0.001

[c] The *P* values in this column result from running linear mixed effects models or Kruskal-Wallis tests to test for significant differences between the means of each stage.

[d] The *P* values in this column result from evaluating multiple comparisons between group means using Tukey's test for linear models or the Dunn's test with a Bonferroni adjustment.

Next, we sought to compare the number of women that transitioned to and from a community that was dominated by *Lactobacillus* spp. Therefore, we identified the dominant bacterial species in the vaginal communities of study participants at each stage sampled (Table D.3). Here, dominance is defined as having a relative proportion greater than 50% within the community. A community was labeled "mixed" if the community contained a more even representation of taxa and was not dominated by any one species. Among the three stages, *Lactobacillus* was dominant in 35, 38, and six communities of the first trimester, third trimester, and postpartum samples, respectively. To evaluate transitions between stage, we only considered the 32 women with complete data, meaning that they had a sample for each stage. Transitions from third trimester to postpartum occurred more frequently than from first to third trimester (Table 4.3). Roughly a third (N = 10, 31%) of the women transitioned from first to third trimester, whereas the majority transitioned from third trimester to postpartum (N = 28, 88%). From early to late pregnancy, transitions to or between communities dominated by *Lactobacillus* were more common. Having a community dominated by *Lactobacillus* did not influence whether or not a woman would transition from third trimester to the postpartum stage as the majority transitioned from a community that was dominated by *Lactobacillus* to one that was not.

Table 4-3. Transitions between dominant bacteria among 32 women with all three time points.

Transition[a]	First ➡ Third	Third ➡ Postpartum
LB ➡ LB	4	4
LB ➡ non LB	1	22
non LB ➡ LB	4	0
non LB ➡ non LB	1	2
Total # of transitions	10	28
% transitions[b]	31	88

[a] Transitions are defined as LB for *Lactobacillus* dominant, or non LB for non *Lactobacillus* dominant communities.

[b] % transitions = (total # of transitions/# of women)*100

Compounds associated with the postpartum stage

We assessed differences in compounds that were measured in vaginal fluids and epithelial cells using linear mixed effects models (Figure 4.4). In comparison to both early and late pregnancy, postpartum samples were characterized by increased levels of hyaluronan ($p < 0.001$) and Hsp70 ($p < 0.01$), and reduced levels of D- ($p < 0.001$) and L-lactic acid ($p < 0.001$, first – postpartum; $p < 0.05$ third – postpartum). Next, we regressed α-diversity measures against these four compounds with subject as a random effect. Using stepwise linear regression, we selected only those that were significantly associated with an increase in α-diversity observed in the transition from pregnancy to the postpartum stage (Table D.4). We found that the increase in α-diversity was negatively associated with levels of L-lactic acid ($p < 0.05$, Shannon diversity; $p < 0.001$, Simpson diversity), but positively associated with hyaluronan ($p < 0.001$, for both), and the postpartum stage ($p < 0.01$, Shannon diversity).

DISCUSSION

We demonstrated that in the transition from pregnancy to the postpartum period there was a sharp decline in the abundances of *Lactobacillus*, namely *L. crispatus*, *L. gasseri*, and *L. jensenii*. This coincided with an increase in the relative abundances of bacteria associated with asymptomatic BV, *S. anginosus* and *P. bivia*, and communities that were more diverse overall. These findings are consistent with previous published work that found the composition of vaginal communities to change considerably from pregnancy to the postpartum stage [4,5,14]. In addition to compositional changes, we detected a significant increase in the levels of hyaluronan and Hsp70, and a significant decrease in the levels of D- and L-lactic acid during the postpartum period. Of these compounds, L-lactic acid and hyaluronan significantly correlated with a decrease or increase in α-diversity, respectively.

The majority of the women in our study that had communities dominated by vaginal lactobacilli during late pregnancy transitioned to communities with low proportions of lactobacilli during the postpartum stage. Similar results were reported in a study that surveyed European women over gestation and six weeks after birth. MacIntyre *et al*. reported that 60% of the postpartum communities shifted to CST-IV from a community previously dominated by *Lactobacillus* during pregnancy. That study included a limited number of samples (N=15) and we've confirmed these findings in a larger cohort (N = 32). Nonetheless, this implies that despite the stability of *Lactobacillus*-dominant communities during pregnancy [44], transitions to more diverse communities after birth are prone to occur.

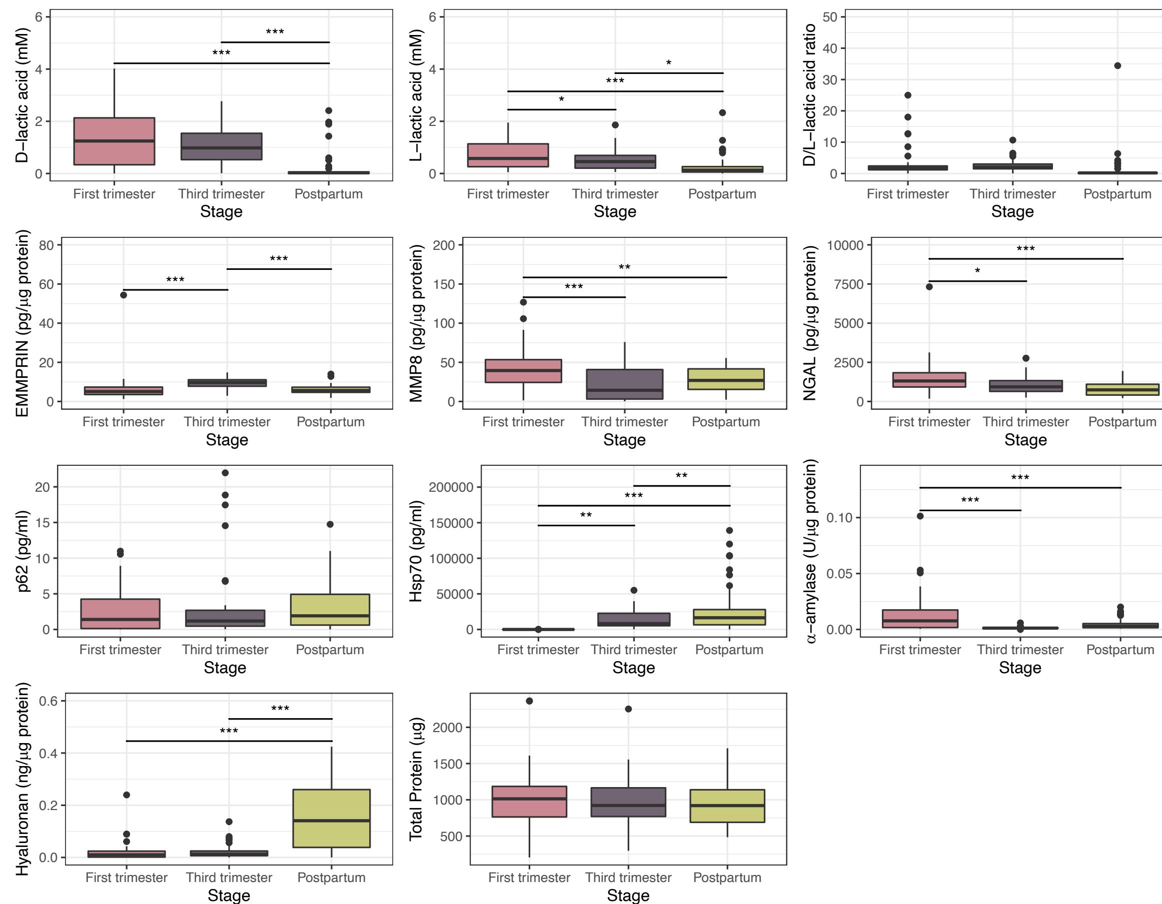


Figure 4.4. Differences in the means of measured compounds by stage. The boxplots in this figure represent the distributions of various compounds measured grouped by stage. The thick band within the box represents the median for the data and the lower and upper boxes represent the 25th and 75th quartiles, respectively. The upper and lower whiskers represent the 95% confidence interval for the median. Individual dots above represent outliers that were not removed. Differences in group means were evaluated using hypothesis testing of linear mixed effects models or nonparametric T-test where appropriate in R. Statistical significance is indicated above the lines as follows: "*", $p < 0.05$; "**", $p < 0.01$; "***" = $p < 0.001$.

The abundances of *Lactobacillus* in the vagina have been positively associated with the levels of estrogen in women (reviewed in [45,46]). Thus, the decrease in the abundances of *Lactobacillus* may well result from reduced levels of estrogen post-delivery (reviewed in [47]). During pregnancy, the placenta produces two forms of estrogen, namely estradiol and estriol, and this causes a sharp increase in estrogen levels with estradiol and estriol increasing two to three orders of magnitude, respectively (reviewed in [47]). This is paralleled by higher abundances of *Lactobacillus* during pregnancy (reviewed in [48]). Once the placenta is removed after birth, estrogen levels are thought to fall precipitously (reviewed in [47]). However, it is unclear exactly how much the levels of estrogen change. One study reported estradiol levels were roughly 7,000 pg/ml at 36 weeks' gestation [49]. Meanwhile, another study determined estradiol levels to be 20 ng/ml during the third trimester [50]. During the postpartum period, estradiol levels were reported as 20 pg/ml at one week and increased slightly above 40 pg/ml after six months [26]. All three studies measured serum estradiol levels. If the decline in estrogen levels is as severe as indicated and if there is a lag time for estrogen levels to rebound after birth, this could lead to lower abundances of lactobacilli in vaginal communities for quite some time post-delivery [4,5]. A limitation of our study is that we did not directly measure estrogen levels to evaluate how this might contribute to changes in vaginal community composition observed in our cohort.

In addition to the decline in estrogen levels, normal physiological changes that occur during labor could influence the dramatic shift in vaginal community composition. Immediately prior to and during parturition, innate immune cells are activated to initiate an inflammatory process that enables uterine contractions, dilation of the cervix, and rupturing of fetal membranes to occur [51]. Avershina *et al*. determined shifts in vaginal community composition can occur as soon as labor begins [25]. To get a sense as to whether there were any host factors associated with changes in community composition, we measured vaginal fluid levels of lactic acid, EMMPRIN, MMP8, NGAL, hyaluronan, α-amylase, in addition to Hsp70 and p62 in vaginal epithelial cell lysates and compared them across stage. Hyaluronan and Hsp70 were significantly higher in postpartum samples, while both D- and L-lactic acid were significantly lower. Hyaluronan is a high molecular weight glycosaminoglycan found in the extracellular matrix in animal tissues [52]. Immediately before labor, hyaluronan levels increase to facilitate softening and dilation of the cervix, and decrease after delivery [53]. Given that we collected postpartum samples within a month of delivery, the association between hyaluronan and the postpartum period observed in our cohort is most likely the result of residual hyaluronan detected in vaginal samples and is not of any significance. Hsp70 is induced in response to stress and downregulates autophagy [54] and has been linked to group B *Streptococcus* persistence in the vaginal communities

of pregnant women [55]. Although we did not measure vaginal pH, we expect that low levels of lactic acid do not acidify the vagina to levels that prevent harmful bacteria from colonizing. Merging high levels of Hsp70 with low levels of lactic acid could very well contribute to the persistence of diverse assemblages of bacteria in the vaginal microbiome post-delivery.

The shift to a microbially diverse community after pregnancy occurs in most women irrespective of ethnic background [5]. It is unclear why these changes are so widespread. However, the postpartum community type appears to be a byproduct of changes in host physiology during and after parturition, which raises several important questions. First, what are the consequences for harboring a diverse vaginal community during the postpartum period? Second, are women who have just given birth at higher risk for acquiring infections? Third, if a woman conceives within a year of giving birth, are her chances for reproductive complications increased? To answer these questions, we need to develop a better understanding about the transitional period after pregnancy and the role of the vaginal microbiome in postpartum health.

LITERATURE CITED

[1] Aagaard K, Riehle K, Ma J, *et al*. A Metagenomic Approach to Characterization of the Vaginal Microbiome Signature in Pregnancy. PLoS ONE 2012;7:e36466.

[2] Freitas AC, Chaban B, Bocking A, *et al*. The vaginal microbiome of pregnant women is less rich and diverse, with lower prevalence of *Mollicutes*, compared to non-pregnant women. Sci Rep 2017:1–16.

[3] Hernández-Rodríguez C, Romero-González R, Albani-Campanario M, *et al*. Vaginal microbiota of healthy pregnant Mexican women is constituted by four *Lactobacillus* species and several vaginosis-associated bacteria. Infect Dis Obstet Gynecol 2011;2011:851485.

[4] DiGiulio DB, Callahan BJ, McMurdie PJ, *et al*. Temporal and spatial variation of the human microbiota during pregnancy. Proc Natl Acad Sci USA 2015;112:11060–5.

[5] MacIntyre DA, Chandiramani M, Lee YS, *et al*. The vaginal microbiome during pregnancy and the postpartum period in a European population. Sci Rep 2015;5:e36466–9.

[6] Boskey ER, Telsch KM, Whaley KJ, *et al*. Acid production by vaginal flora in vitro is consistent with the rate and extent of vaginal acidification. Infect Immun 1999;67:5170–5.

[7] Boskey ER, Cone RA, Whaley KJ, *et al*. Origins of vaginal acidity: high D/L lactate ratio is consistent with bacteria being the primary source. Hum Reprod 2001;16:1809–13.

[8] O'Hanlon DE, Moench TR, Cone RA. In vaginal fluid, bacteria associated with bacterial vaginosis can be suppressed with lactic acid but not hydrogen peroxide. BMC Infect Dis 2011;11:251.

[9] O'Hanlon DE, Moench TR, Cone RA. Vaginal pH and Microbicidal Lactic Acid When Lactobacilli Dominate the Microbiota. PLoS ONE 2013;8:e80074–8.

[10] Aldunate M, Tyssen D, Johnson A, *et al*. Vaginal concentrations of lactic acid potently inactivate HIV. J Antimicrob Chemother 2013;68:2015–25.

[11] Wagner W, Ciszewski WM, Kania KD. L- and D-lactate enhance DNA repair and modulate the resistance of cervical carcinoma cells to anticancer drugs via histone deacetylase inhibition and hydroxycarboxylic acid receptor 1 activation. Cell Commun Sig 2015:1–16.

[12] Witkin SS. The vaginal microbiome, vaginal anti-microbial defense mechanisms and the clinical challenge of reducing infection-related preterm birth. BJOG 2015;122:213–9.

[13] Huang Y-E, Wang Y, He Y, *et al*. Homogeneity of the Vaginal Microbiome at the Cervix, Posterior Fornix, and Vaginal Canal in Pregnant Chinese Women. Microb Ecol 2014;69:407–14.

[14] Bisanz JE, Enos MK, PrayGod G, *et al*. Microbiota at Multiple Body Sites during Pregnancy in a Rural Tanzanian Population and Effects of Moringa-Supplemented Probiotic Yogurt. Appl Environ Microbiol 2015;81:4965–75.

[15] Onderdonk AB, Delaney ML, Fichorova RN. The Human Microbiome during Bacterial Vaginosis. Clin Microbiol Rev 2016;29:223–38.

[16] Atashili J, Poole C, Ndumbe PM, *et al*. Bacterial vaginosis and HIV acquisition: a meta-analysis of published studies. AIDS 2008;22:1493–501.

[17] van de Wijgert JHHM, Borgdorff H, Verhelst R, *et al*. The Vaginal Microbiota: What Have We Learned after a Decade of Molecular Characterization? PLoS ONE 2014;9:e105998–10.

[18] Nasioudis D, Linhares IM, Ledger WJ, *et al*. Bacterial vaginosis: a critical analysis of current knowledge. BJOG 2017; 124(1):61-69.

[19] Ravel J, Gajer P, Abdo Z, *et al*. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci USA 2011;108 Suppl 1:4680–7.

[20] Doyle R, Gondwe A, Fan Y-M, *et al*. A *Lactobacillus*-Deficient Vaginal Microbiota Dominates Postpartum Women in Rural Malawi. Appl Environ Microbiol 2018;84:4680–12.

[21] Petrova MI, Reid G, Vaneechoutte M, *et al*. *Lactobacillus iners*: Friend or Foe? Trends Microbiol 2017;25:182–91.

[22] Kindinger LM, Bennett PR, Lee YS, *et al*. The interaction between vaginal microbiota, cervical length, and vaginal progesterone treatment for preterm birth risk. Microbiome 2017; 5(1):6.

[23] Goltsman DSA, Sun CL, Proctor DM, *et al*. Metagenomic analysis with strain-level resolution reveals fine-scale variation in the human pregnancy microbiome. Genome Res 2018;28:1467–80.

[24] Gajer P, Brotman RM, Bai G, *et al*. Temporal Dynamics of the Human Vaginal Microbiota. Sci Transl Med 2012;4:132ra52–2.

[25] Avershina E, Slangsvold S, Simpson MR, *et al*. Diversity of vaginal microbiota increases by the time of labor onset. Sci Rep 2017; 7(1):17558.

[26] Groer ME, Jevitt C, Ji M. Immune Changes and Dysphoric Moods Across the Postpartum. Am J Reprod Immunol 2014;73:193–8.

[27] Aghaeepour N, Ganio E, Mcilwain D, *et al*. An immune clock of human pregnancy. Sci Immunol 2017:eean2946.

[28] Yuan S, Cohen DB, Ravel J, *et al*. Evaluation of methods for the extraction and purification of DNA from the human microbiome. PLoS ONE 2012;7:e33865.

[29] Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 2011;27:2957–63.

[30] Callahan BJ, McMurdie PJ, Rosen MJ, *et al*. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Meth n.d.;13:581EP–.

[31] Allard G, Ryan FJ, Jeffery IB, *et al*. SPINGO: a rapid species-classifier for microbial amplicon sequences. BMC Bioinformatics 2015;16:1–8.

[32] Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. 2019:1–2673.

[33] Maechler M, Rousseeuw P, Struyf A, *et al*. cluster: Cluster Analysis Basics and Extensions 2019.

[34] Meyer D, Dimitriadou E, Hornik K, *et al*. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien 2019.

[35] Ogle DH, Wheeler P, Dinno A. Fisheries Stock Analysis 2019.

[36] Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag; 2016.

[37] Kassambara A. ggpubr: “ggplot2” Based Publication Ready Plots 2019:1–126.

[38] Bates D, Mächler M, Bolker B, *et al*. Fitting Linear Mixed-Effects Models Using lme4. J Stat Soft 2015;67:1–48.

[39] Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTestPackage: Tests in Linear Mixed Effects Models. J Stat Soft 2017;82:1–26.

[40] Hothorn T, Bretz F, Westfall P. Simultaneous Inference in General Parametric Models. Biom J 2008;50:346–63.

[41] Wickham H. The Split-Apply-Combine Strategy for Data Analysis. J Stat Soft 2011;40:1–29.

[42] Wickham H, Henry L. tidyr: Tidy Messy Data 2019.

[43] Oksanen J, Blanchet FG, Friendly M, *et al*. vegan: Community Ecology Package. R package version 2.5-6 2019:1–295.

[44] Romero R, Hassan SS, Gajer P, *et al*. The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. Microbiome 2014;2(1):10.

[45] Hickey RJ, Zhou X, Pierson JD, *et al*. Understanding vaginal microbiome complexity from an ecological perspective. Transl Res 2012;160:267–82.

[46] Nunn KL, Forney LJ. Unraveling the Dynamics of the Human Vaginal Microbiome. Yale J Biol Med 2016;89:331–7.

[47] D VHM, D LLAM, D RSM. Review Article Hormonal Changes in the Postpartum and Implications for Postpartum Depression. Psychosomatics 2011;39:93–101.

[48] Amabebe E, Anumba DOC. The Vaginal Microenvironment: The Physiologic Role of Lactobacilli. Front Med 2018;5:164–11.

[49] Glynn LM, Davis EP, Sandman CA, *et al*. Gestational hormone profiles predict human maternal behavior at 1-year postpartum. Hormones and Behavior 2016;85:19–25.

[50] Schock H, Zeleniuch-Jacquotte A, Lundin E, *et al*. Hormone concentrations throughout uncomplicated pregnancies: a longitudinal study. BMC Pregnancy and Childbirth 2016;16:146.

[51] Gomez-Lopez N, StLouis D, Lehr MA, *et al*. Immune cells in term and preterm labor. Cell Mol Immunol 2014;11(6):571–81.

[52] Fouladi-Nashta AA, Raheem KA, Marei WF, *et al*. Regulation and roles of the hyaluronan system in mammalian reproduction. Reproduction 2016;153:R43–R58.

[53] Marandy EE, Kanayama N, Kobayashi H, *et al*. The role of hyaluronic acid as a mediator and regulator of cervical ripening. Hum Reprod 1997;12(5):1080–8.

[54] Kanninen TT, Sisti G, Witkin SS. Induction of the 70 kDa heat shock protein stress response inhibits autophagy: possible consequences for pregnancy outcome. J Matern Fetal Neonatal Med 2014;29:159–62.

[55] Scholl J, Nasioudis D, Boester A, *et al*. Group B streptococcus alters properties of vaginal epithelial cells in pregnant women. Am J Obstet Gynecol 2016;214(3):383.e1–5.

CHAPTER 5: CONCLUSION AND FUTURE DIRECTIONS

The goal of my research has been to gain insight to why changes in the composition of the vaginal microbiome occur, with an emphasis on *Lactobacillus* species. The scientific premise behind my dissertation stems from previous research that showed a correlation between changes in the relative abundances of vaginal lactobacilli and the levels of estrogen and vaginal glycogen in women [1-4]. As reviewed in the opening chapter, this relationship between estrogen, glycogen, and lactobacilli has been consistently observed throughout the reproductive stages of a woman's lifetime (Figure 5.1). Briefly, increases in estrogen levels cause the vaginal epithelium to thicken, which prompts the accumulation of glycogen within the vaginal epithelium. This glycogen is released from sloughed epithelial cells whereupon it is thought to be depolymerized by host α-amylase into smaller subunits of oligosaccharides, maltose and glucose that can be fermented by *Lactobacillus* to produce lactic acid. This acidifies the vaginal environment which provides protection against invading pathogens.

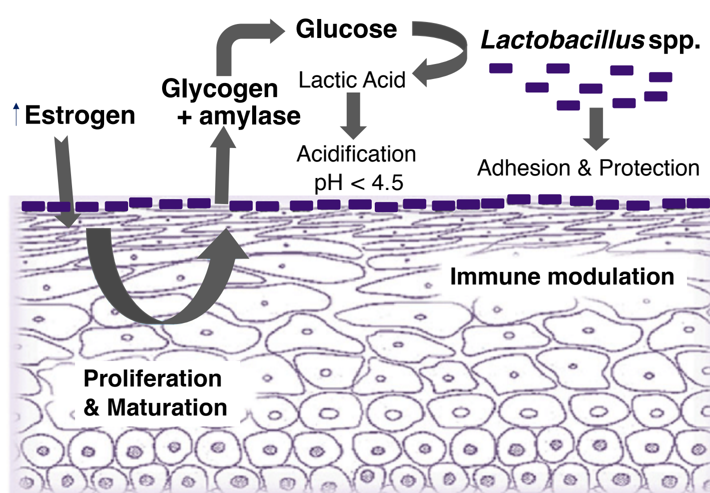


Figure 5.1. Schematic for the relationship between estrogen, glycogen, and *Lactobacillus* in the human vagina. Adapted/modified from [18].

While these correlations exist, there is scant direct evidence that documents these relationships and the studies in Chapters 2 and 3 were the first to dissect these associations. My findings indicate that while estrogen levels are positively associated with increased abundances of *Lactobacillus* in the vagina, this relationship is much more complex than drawing a straight line from estrogen to glycogen to lactobacilli. As mentioned earlier, once glycogen is released into the vaginal lumen, it is presumed to be broken down by α-amylase produced by the host. Prior to the study in Chapter 3, there was no

definitive evidence to confirm the source of α-amylase in the vagina. I present new evidence that both host and bacteria amylases are found in vaginal fluids of reproductive age women. Some of the bacterial amylases that we detected were produced by vaginal lactobacilli. In addition, the profiles of bacterial amylases and the abundances of human amylases differ between women. These differences might contribute to variation in the kinds and amounts of glycogen-derived resources in the vagina. Moreover, competition for these resources could determine which bacterial species will inhabit the vagina.

It is known that changes in community composition occur during pregnancy and these transitions are thought to be driven by estrogenic compounds. Moreover, the levels of estrogen decrease precipitously following delivery [5]. In Chapter 4, I observed a remarkable shift in vaginal community composition that occurred in the majority of women in the study cohort concurrent with the transition from pregnancy to the postpartum period. Communities that were previously dominated by species of *Lactobacillus* shifted to ones that had low proportions of *Lactobacillus*. In addition, we report a significant increase in Hsp70, and significant decreases in both lactic acid isomers during this transition from pregnancy to postpartum. We hypothesize that increased levels of Hsp70, with low levels of lactic acid, may enable BV-associated bacteria to persist in the vagina for an extended period during the postpartum stage. The study in Chapter 4 is the first to assess the relationships between host associated vaginal compounds and changes in vaginal community composition immediately following pregnancy.

These studies contribute novel findings that help fill gaps in knowledge regarding the roles of estrogen and glycogen in shaping the composition of vaginal communities. At the same time, several unanswered questions remain, and I focus on two of these below. First, *Lactobacillus* species produce amylases that are expressed in vaginal fluids, which suggests they have the metabolic potential to catabolize glycogen. Until recently, attempts to culture vaginal lactobacilli on glycogen have been unsuccessful [6-9]. Why one group was successful [10] while the others were not is unclear. What is the cause of this discrepancy? Second, the positive correlation between estrogen, glycogen, and vaginal lactobacilli is well documented [11,12]. In this dissertation, I reported there is no simple relationship between the levels of estrogen, glycogen and lactobacilli. In addition to prompting maturation of the vaginal epithelium, estrogen has been shown to alter many aspects of host gene expression [13-15]. Could estrogen-induced changes in gene expression influence the abundances of lactobacilli in the vagina?

Glycogen as a resource for Lactobacillus in the vagina

Due to the positive association with estrogen, glycogen and the abundances of *Lactobacillus* in the human vagina, glycogen had long been thought to be a key resource for lactobacilli. Currently, there is a debate as to whether vaginal lactobacilli could metabolize glycogen. The idea that lactobacilli cannot metabolize glycogen is supported by published work dating back to the 1960's, perhaps even earlier, where researchers were unable to culture lactobacilli in media containing glycogen [6-9]. Contrary to this, a recent study that included 33 *Lactobacillus crispatus* vaginal isolates and showed that 24 of them could grow in media supplemented with glycogen [10]. In accordance with these findings I found *Lactobacillus* amylases expressed in vaginal fluids. I wanted to investigate the source of these conflicting results.

Comparison of Lactobacillus growth in different sources of media containing glycogen

I expect the conflicting results described above result from differences in the culture media used. van der Veer and colleagues [10], who observed growth of *L. crispatus* in media containing glycogen used NYC-III media. This media is supplemented with horse serum, which has been reported to contain amylase enzymes [16]. Researchers who were unable to culture lactobacilli on glycogen used de Man, Rogosa, and Sharpe (MRS) media [17], which does not contain horse serum. I designed a growth experiment to determine whether the addition of horse serum to MRS medium would facilitate the growth of *L. crispatus* on glycogen.

I first aimed to reproduce and compare the results of *L. crispatus* cultured in MRS and NYC-III media. To do this, I cultured the type strain for *L. crispatus* (ATCC 33820) overnight in either MRS media containing 2% glucose or NYC-III media containing 0.5% glucose. Overnight cultures were diluted back 10 - 20% to re-establish log phase growth. Cells were harvested after at least three doublings, washed once in 1X PBS, and resuspended in MRS or NYC-III supplemented with glucose as a positive control or glycogen (both 2% for MRS or 0.5% for NYC-III as per the media recipe), or no sugar. Media without sugar added served as a control for growth on basal media components. Optical density (OD) measurements at 600nm were recorded at times 0, 24, and 48 hours. The results showed that *L. crispatus* 33820 did not grow in MRS media supplemented with glycogen (Figure 5.2A), but it did grow in NYC-III media supplemented with glycogen (Figure 5.2B). Here, I was able to reproduce the results that suggest *Lactobacillus* could not grow on glycogen and those that suggest the opposite.

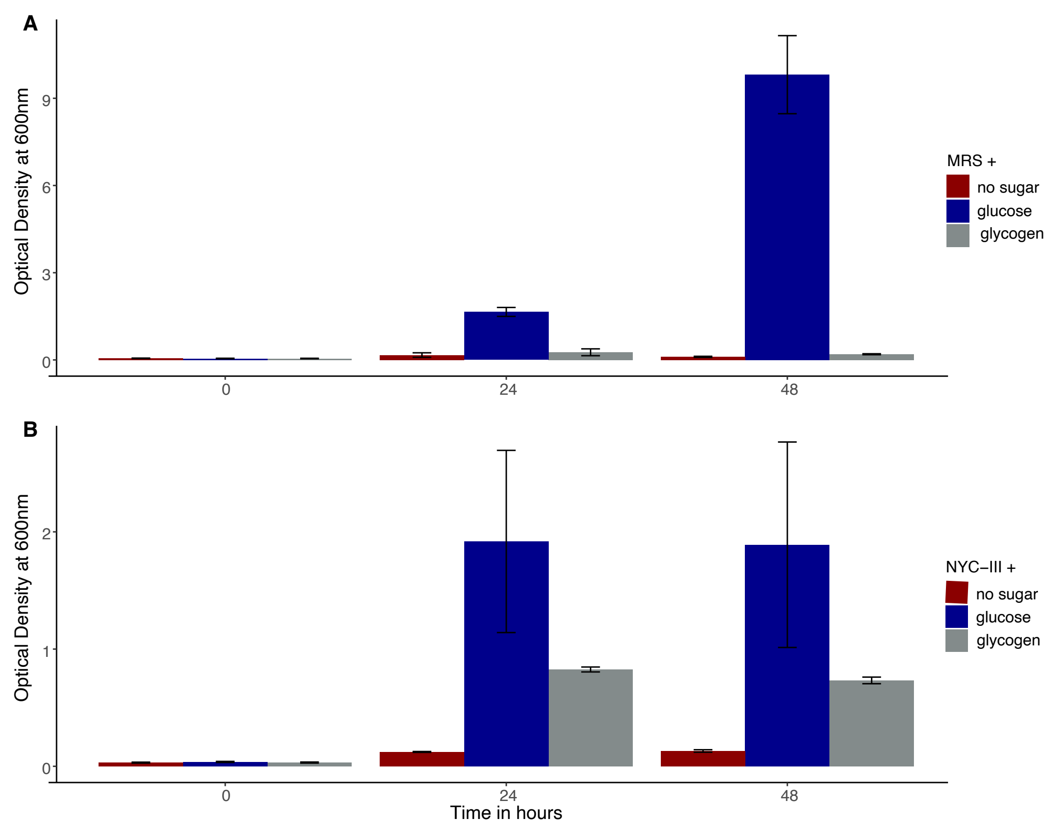


Figure 5.2. *Lactobacillus crispatus* cultured in (A) MRS media and (B) NYC-III media with out sugar (red), with glucose (blue), or with glycogen (grey). Four replicates were performed for each.

I hypothesize the horse serum component of NYC-III contains an enzyme that breaks down glycogen into simpler sugars that *L. crispatus* are able to grow on. Similarly, if I add horse serum to MRS media supplemented with glycogen, I should be able to achieve similar growth when compared to *L. crispatus* cultured in NYC-III. Thus, I added horse serum to MRS supplemented with glycogen and used this media to culture *L. crispatus*. *L. crispatus* cultures were initiated as described above in MRS media supplemented with glucose. I compared *L. crispatus* growth in the following media: MRS without sugar, MRS without sugar plus 10% horse serum, MRS with glycogen, and MRS with glycogen and 10% horse serum. I observed growth in MRS media supplemented with glycogen and 10% horse serum. The final OD measurements were similar to those when *L. crispatus* was cultured in NYC-III media with glycogen (Figure 5.3). These results suggest that horse serum has amylase activity that can break down glycogen, which could have enabled *L. crispatus* to grow in this NYC-III media but not MRS. To test this further, one could inhibit amylase activity in horse serum and evaluate how this would affect the growth of *L. crispatus* in NYC-III media supplemented with glycogen. Additionally, more work could be done to determine the products that are produced in media containing horse serum and glycogen and to characterize the growth of *L. crispatus* on the resulting substrates.

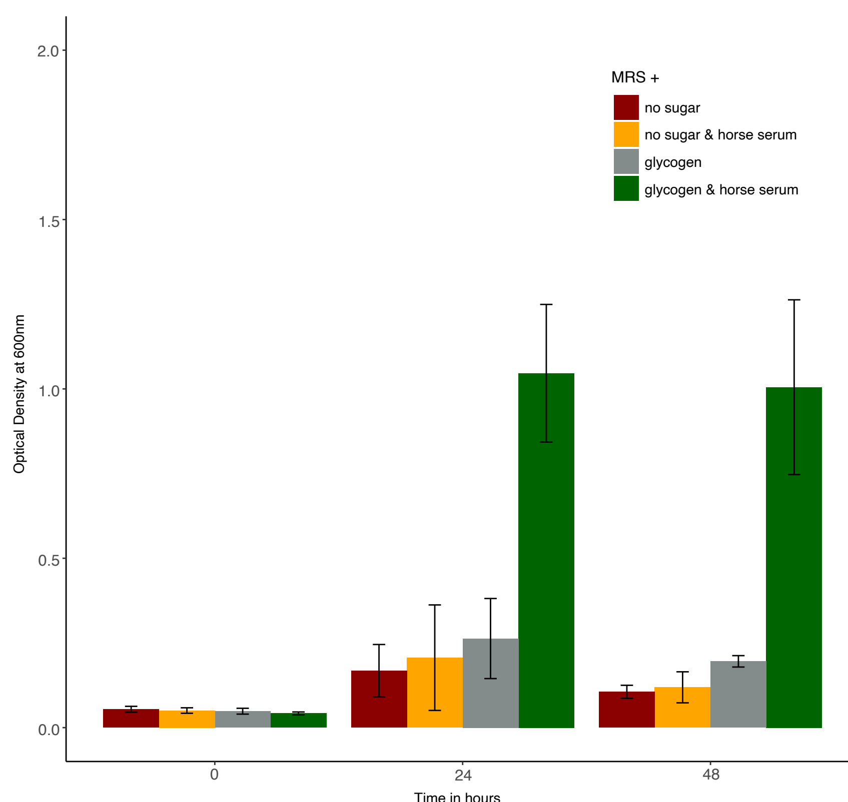


Figure 5.3. *Lactobacillus crispatus* cultured in MRS without sugar (red), without sugar plus horse serum (orange), with glycogen (grey), with glycogen plus horse serum (green). Four replicates were performed for each.

This pilot experiment suggests the findings of previous study by van der Veer *et al*. [10] could have been compromised by inadequate controls and attention to experimental variables, which could also include the source of glycogen. Whylie and Henderson isolated 42 strains of *Lactobacillus* from the vaginas of pregnant women and showed that only one strain could metabolize oyster glycogen and another two could metabolize vaginal glycogen [7]. Thus, perhaps the source of glycogen also matters. To confirm whether or not lactobacilli can metabolize glycogen, future studies might also test different sources of glycogen.

Linking estrogen to increased abundance of Lactobacillus *in the vagina*

There is compelling evidence that estrogen, glycogen and lactobacilli are, in some way, intertwined. However, the underlying mechanisms are unknown. One possibility that no one has studied is the

likelihood of indirect effects – for example, it could be that estrogen alters host gene expression which in turn stimulates the proliferation of *Lactobacillus* in the vagina. One study that assessed the response of vaginal tissues to estradiol reported genes involved in cholesterol biosynthesis, cell growth/development and differentiation, and steroid metabolism were upregulated, while genes involved in host immune response were downregulated [14]. A different study found biological processes involving fatty acid, hormone, pyridine nucleotide, and monosaccharide metabolism in addition to epithelial cell differentiation to be upregulated in the vaginal introitus of pre-menopausal when compared to postmenopausal women [15]. Moreover, the profile of differentially expressed genes in pre-menopausal women were similar to post-menopausal women receiving hormone therapy [15]. I have done preliminary experiments examining patterns of gene expression when vaginal epithelial cells and *L. crispatus* were co-cultured in the presence or absence of estrogen (Figure 5.4).

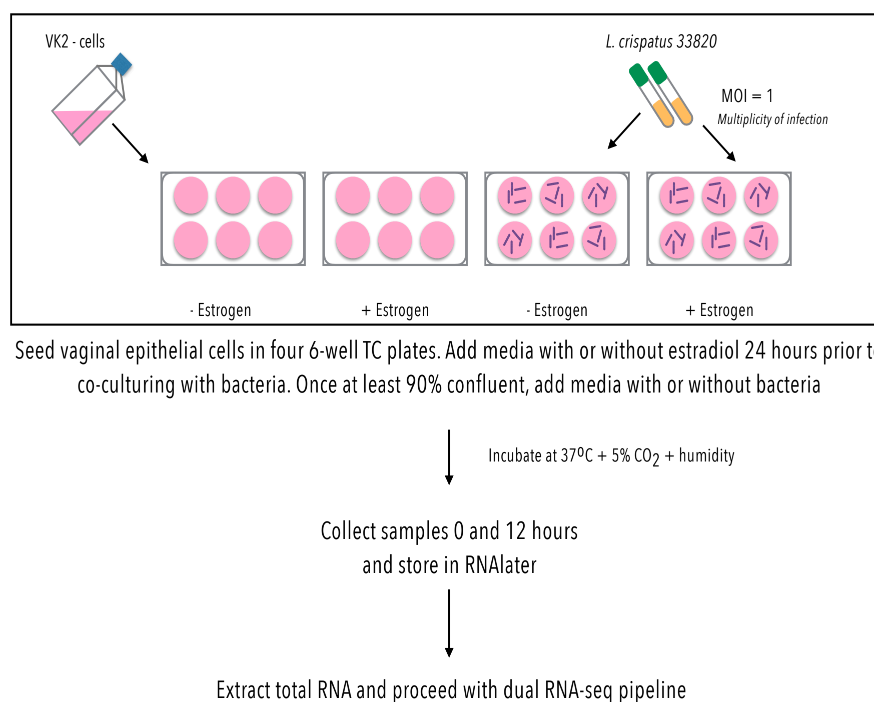


Figure 5.4. Schematic for VK2 cell and *L. crispatus* co-culture, RNA-seq experiment.

To perform these experiments, I used the VK2/E6E7 vaginal epithelial cell line from the American Type Culture Collection. VK2 cells were cultured in keratinocyte serum-free media (KSF) with 0.1 ng/ml human recombinant EGF, 0.05 mg/ml bovine pituitary extract, calcium chloride 0.4 mM, and 1% penicillin-streptomycin. The type strain of *L. crispatus* (ATCC 33820) was cultured in NYC-III

media. Co-culture experiments were performed using a 50:50 mixture of KSF and NYC-III media without antibiotic. For the co-culture experiments, VK2 cells were seeded in wells of a 6-well tissue culture plate (x4) and co-cultured with or without *L. crispatus* and in the presence or absence of 17β-estradiol for 12 hours. Cultures were harvested at time 0 and 12 hours by scraping the wells of the 6-well plate and placed the cell suspensions on ice. Cultures were processed, preserved in RNAlater (Invitrogen), and RNA was extracted using the Rneasy kit (Qiagen). Library preparation and RNA-seq will be performed at the Genomics Resources Core at the University of Idaho. These data will enable us to evaluate the effect of estrogen on host gene expression and host-*Lactobacillus* interactions and hopefully provide insight to how estrogen enables lactobacilli to dominate microbial communities in the human vagina.

Concluding remarks

In summary, the more we try to unravel the role of estrogen and glycogen as potential drivers of the abundances of vaginal lactobacilli, the messier it gets. Many factors likely contribute to shaping the composition of vaginal communities. Some of these include but are not limited to differences in host genetics and immune response, the endocrine system and stress, diet, hygiene, and interactions between members of the vaginal community. As a result, it is almost impossible to account for all of the interactions that occur in order to explain why differences in community composition exist. In this dissertation, I've relied mostly on clinical studies to evaluate various associations with estrogen, glycogen, and vaginal community composition. However, empirical studies are still needed to tease apart basic building blocks of these larger associations.

LITERATURE CITED

[1] Cruickshank R, Sharman A. The Biology of the Vagina in the Human Subject. BJOG 1934:208–26.

[2] Paavonen J. Physiology and ecology of the vagina. Scand J Infect Dis Suppl 1983;40:31–5.

[3] Farage M, Maibach H. Lifetime changes in the vulva and vagina. Arch Gynecol Obstet 2005;273:195–202.

[4] Mirmonsef P, Hotton AL, Gilbert D, et al. Free glycogen in vaginal fluids is associated with *Lactobacillus* colonization and low vaginal pH. PLoS ONE 2014;9:e102467.

[5] Hendrick V, Altshuler LL, Suri R. Hormonal Changes in the Postpartum and Implications for Postpartum Depression. Psychosomatics 2011;39:93–101.

[6] STEWART-TULL DE. Evidence that vaginal lactobacilli do not ferment glycogen. Am J Obstet Gynecol 1964;88:676–9.

[7] Whylie JG, Henderson A. Identity and glycogen-fermenting ability of lactobacilli isolated from the vagina of pregnant women. J Med Microbiol 1969;2:363–6.

[8] Martin R, Soberon N, Vaneechoutte M, et al. Characterization of indigenous vaginal lactobacilli from healthy women as probiotic candidates. Int Microbiol 2008;11:261–6.

[9] Spear GT, French AL, Gilbert D, et al. Human α-amylase Present in Lower-Genital-Tract Mucosal Fluid Processes Glycogen to Support Vaginal Colonization by *Lactobacillus*. J Infect Dis 2014;210:1019–28.

[10] van der Veer C, Hertzberger RY, Bruisten SM, et al. Comparative genomics of human *Lactobacillus crispatus* isolates reveals genes for glycosylation and glycogen degradation: implications for in vivo dominance of the vaginal microbiota. Microbiome 2019;7:49.

[11] Hickey RJ, Zhou X, Pierson JD, et al. Understanding vaginal microbiome complexity from an ecological perspective. Transl Res 2012;160:267–82.

[12] Nunn KL, Forney LJ. Unraveling the Dynamics of the Human Vaginal Microbiome. Yale J Biol Med 2016;89:331–7.

[13] Cotreau MM, Chennathukuzhi VM, Harris HA, et al. A study of 17β-estradiol-regulated genes in the vagina of postmenopausal women with vaginal atrophy. Maturitas 2007;58:366–76.

[14] Jelinsky SA, Choe SE, Crabtree JS, et al. Molecular analysis of the vaginal response to estrogens in the ovariectomized rat and postmenopausal woman. BMC Med Genomics 2008;1:1095–12.

[15] Binder RL, Freedman MA, Sharma KB, et al. Histological and Gene Expression Analysis of the Effects of Menopause Status and Hormone Therapy on the Vaginal Introitus and Labia Majora. J Clin Med Res 2019;11:745–59.

[16] Uto Y, Ikeno T. Amylase activity and the isoenzymes in animals. Jpn J Orl Biol 2010;31:336–9.

[17] De Man JC, Rogosa M, Sharpe ME. A Medium For The Cultivation Of lactobacilli. J Appl Bacteriol 2008;23:130–5.

[18] Mueck AO, Ruan X, Prasauskas V, et al. Treatment of vaginal atrophy with estriol and lactobacilli combination: a clinical review. Climacteric 2018; 21(2): 140–147.

APPENDIX A - SUPPLEMENTAL MATERIALS FOR CHAPTER 1

Permission to reprint YJMB article

Subject: RE: Permission to reprint in dissertation
Date: Friday, December 6, 2019 at 9:12:04 AM Pacific Standard Time
From: Yale Journal of Biology and Medicine
To: Nunn, Kenetta (knunn@uidaho.edu)

Hi Kenetta,

We usually go through Copyright clearance center, but since this is a dissertation, we will waive all fees and forms. All the YJBM asks is that you credit both the author and the journal (YJBM) in reprinting the article.

Best wishes,

Kate

Kate Woodford
Editorial Coordinator
Yale Journal of Biology and Medicine
Edward S. Harkness Building D
367 Cedar Street, Room 321
New Haven, CT 06510
medicine.yale.edu/yjbm

From: Nunn, Kenetta (knunn@uidaho.edu) <knunn@uidaho.edu>
Sent: Thursday, December 5, 2019 1:59 PM
To: Yale Journal of Biology and Medicine <yjbm@yale.edu>
Subject: Permission to reprint in dissertation

Hello,

I would like to reprint the following article in my dissertation for completing my PhD: Nunn KL and Forney LJ. Unraveling the dynamics of the human vaginal microbiome. YJBM, 2016. 89(3):331-337. How would I go about getting permission for that?

Kind regards,
Kenetta

APPENDIX B - SUPPLEMENTAL MATERIALS FOR CHAPTER 2

Permission to reprint JAH article

RightsLink®

Home

Help

Email Support

Sign in

Create Account

Vaginal Glycogen, Not Estradiol, Is Associated With Vaginal Bacterial Community Composition in Black Adolescent Women

Author:
Kenetta L. Nunn,Benjamin J. Ridenhour,Emily M. Chester,Virginia J. Vitzthum,J. Dennis Fortenberry,Larry J. Forney

Publication: Journal of Adolescent Health

Publisher: Elsevier

Date: July 2019

© 2019 Society for Adolescent Health and Medicine. All rights reserved.

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights

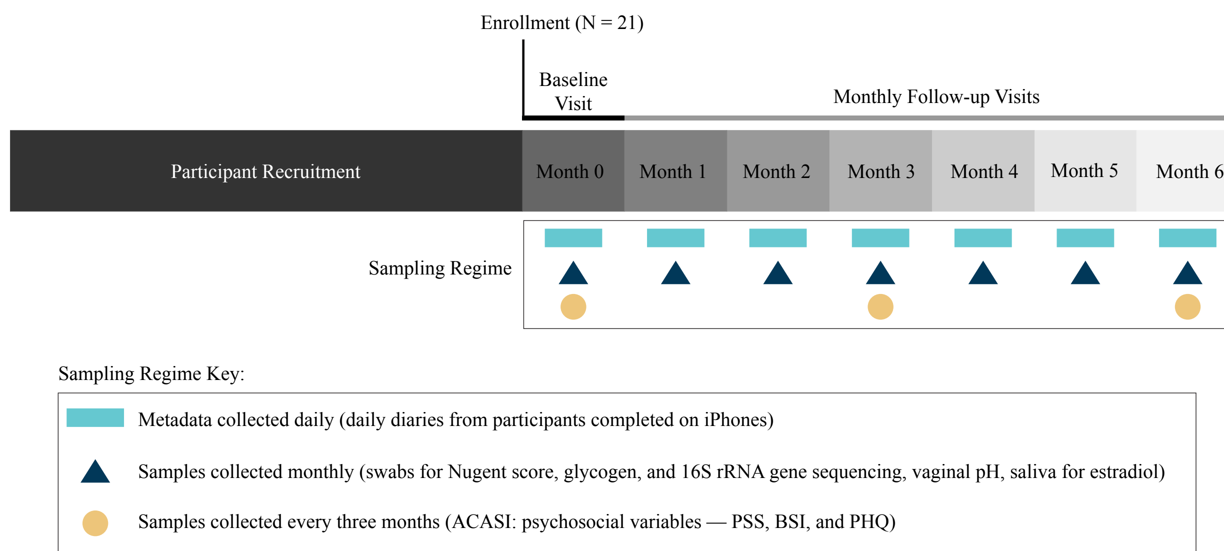
BACK CLOSE WINDOW

© 2019 Copyright - All Rights Reserved | Copyright Clearance Center, Inc. | Privacy statement | Terms and Conditions
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Supplementary Materials and Methods

Sampling regime and data collection

Over the course of the study, 21 study participants completed a total of seven study visits. The first study visit was used as the baseline. After this visit, participants were followed monthly for six months. From the 21 participants, we collected data daily, monthly, and every three months. Daily data included daily diaries consisting of self-reported emotional state, individual activities and group activities, hygiene practices, vaginal and other symptoms, and sexual behaviors. Monthly data included self-reports on menses, vaginal swabs collected for assessing Nugent score, glycogen and sequencing 16S rRNA genes, vaginal pH, and saliva collected by passive drool for measuring estradiol. Saliva samples were also obtained three days up to each monthly visit to represent potential lagging influences on vaginal microbiota. Data collected every three months (starting at the baseline visit) included audio computer assisted interviews (ACASI) for measuring depression using the Patient Health Questionnaire (PHQ), stress using the Perceived Stress Scale (PSS), and anxiety using the Brief Symptom Inventory (BSI). The study design and sampling regime is represented in the diagram below.



Statistical Methods

Microbial community analysis

To determine the kinds of bacterial communities present in women of our cohort, we computed an alt-Gower dissimilarity index (equation 1; [1]) using the R package vegan v2.5-2 [2] from relative abundance data, and then performed complete-linkage hierarchical clustering on this index using the R package stats v.3.4.0 [3].

$$d_{ij} = (1/NZ)\sum\left(abs\left(x_{ij} - y_{ij}\right)\right) \quad (1)$$

Where NZ is the number of columns (non-zero and excluding double-zeros) and xij xik is the observation in column i and rows j and k.

We determined the optimal number of clusters from the hierarchical cluster analysis using the silhouette method [4] in the R package cluster v.2.0.6 [5] and cut the dendrogram shown in Figure 1 into nine clusters based on this result. We observed shared characteristics among these nine clusters with respect to patterns in *Lactobacillus* abundances. For this reason, we combined the nine clusters into groups A, B, and C. Four of the clusters were dominated by one species of *Lactobacillus* (group A). Three of the clusters had high proportions of a mixture of *Lactobacillus* (group B) species. The remaining two clusters exhibited low proportions of *Lactobacillus* (group C).

Differences in measurements between groups A, B, and C

Estradiol, glycogen, and vaginal pH. We modeled the means of the response variables (estradiol, glycogen, or vaginal pH) between groups as the predictor variable using the R package lme4 v.1.1-13 [6]. Estradiol, glycogen, and vaginal pH were all log-transformed to avoid violating model assumptions. For each response variable, we ran two models:

model 1*: log(response variable) ~ group + (1|subject visit) + (1|subject)*

model 2*: log(response variable) ~ group + (1|subject)*

We compared models 1 and 2 for each response variable using analysis of variance (ANOVA) from the R package stats v.3.4.0 [3] and determined there were no significant differences between the models. In each case, however, model 2 had the lower Akaike information criterion (AIC) value. Therefore, we selected model 2 for each response variable. We then performed an ANOVA with type II Wald chi-squared tests on the selected models using the R package car v.2.1-5 [7] to determine if there were any significant differences in estradiol, glycogen, or vaginal pH between groups A, B, and C. If the ANOVA indicated no significant difference, we report the resulting chi-squared test statistic and corresponding *p*-value for the model in the results. If the ANOVA indicated there were significant differences as in the case with glycogen and vaginal pH, we compared group means from the linear mixed-effects models using the R package multcomp v.1.4-8 [8]. Specifically, we performed Tukey's test on the means of groups A, B, and C. We report the *z*-test statistic and corresponding adjusted *p*-value for each comparison in the results. These results are represented in the boxplots in Figure 4.

Stress measurements. Similar to the models selected for estradiol, glycogen, and vaginal pH between groups, we modeled the means of the response variables PHQ score (depression) and BSI score (anxiety) between groups as the predictor variable. Based on the same model selection criteria (ANOVA and AIC), subject was included as a random effect variable in these models. There was no support to include subject as a random effect variable in the model for PSS score (perceived stress); thus, we fit a linear model to model the means of the response variable PSS score between groups as the predictor variable. PSS, PHQ, and BSI scores were transformed by taking the square root of each measurement prior to analysis. We performed an ANOVA with type II Wald chi-squared tests on the selected models to determine if there were any significant differences in PHQ score and BSI score between groups A, B, and C. We report the resulting chi-squared test statistic and corresponding *p*-value for these models in Table 1. To determine if there were any differences in PSS score between groups, we performed an ANOVA with type II *F*-tests on the linear model for PSS score. For this model, we report the resulting *F*-test statistic and corresponding *p*-value in Table 1.

Nugent score. Nugent scores could not be transformed to yield normally distributed values with equal variances; therefore, we performed a Kruskal-Wallis rank sum test modeling the means of Nugent score between groups as the predictor variable. The resulting chi-squared test indicated there were significant differences in Nugent score between groups. Thus, we followed this analysis with a Dunn's test using the R package FSA v.0.8.16 [9], incorporating a Bonferroni adjustment, to evaluate significant

differences in Nugent score between groups. We report the *z*-values and corresponding adjusted *p*-values with these results. These results are represented in the boxplots in Figure 4.

The relationship between estradiol and glycogen

We evaluated the relationship between estradiol and glycogen by modeling log-transformed estradiol measurements over glycogen measurements using linear mixed effects models. Subject was included as a random effect variable to account for any variability due to subject. Statistical significance of this model was evaluated using an ANOVA. These data are representing in the scatter plot in Figure 4 along with the resulting linear model and *p*-value for the model.

Differences in estradiol over the menstrual cycle

We modeled the means of estradiol between menstrual phase as the predictor variable. As in previous analyses, estradiol was log-transformed to avoid violating model assumptions. We ran three models:

phase model 1*: log(estradiol) ~ menstrual phase + (1|subject)*

phase model 2*: log(estradiol) ~ menstrual phase + (1|subject) + (1|subject visit)*

phase model 3*: log(estradiol) ~ menstrual phase + (1+ menstrual phase | subject)*

Using the same model selection criteria as before (ANOVA and AIC) we selected phase model 1. To determine if there were significant differences in estradiol over the menstrual cycle phases, we performed an ANOVA with type II Wald chi-squared tests on the phase model 1. The ANOVA did not detect any significant differences. Thus, we simply report the resulting chi-squared test statistic and corresponding *p*-value in the results. These data are represented in Figure S3.

Correlation plots

We calculated Pearson correlation coefficients for data shown in Figure S4 using the R package Hmisc v.4.0-3 [10]. These correlations were plotted using the R package corrplot v.0.84 [11]. Correlations were deemed significant if the *p*-value was less than 0.05.

LITERATURE CITED:

[1] Anderson MJ, Ellingsen KE, McArdle BH. Multivariate dispersion as a measure of beta diversity. Ecol Lett 2006; 9(6):683–93.

[2] Oksanen J, Blanchet FG, Friendly M, et al. vegan: Community Ecology Package. R package version 2.5-2; 2018. https://CRAN.R-project.org/package=vegan.

[3] Team RC. R: A language and environment for statistical computing. R

Foundation for Statistical Computing, Vienna, Austria; 2017.

URL https://www.R-project.org/.

[4] Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 1987;20:53–65.

[5] Maechler M, Rousseeuw P, Struyf A, et al. cluster: Cluster Analysis Basics and Extensions. R package version 2.0.6; 2017.

[6] Bates D, Mächler M, Bolker B, et al. Fitting Linear Mixed-Effects Models Using lme4. J Stat Soft 2015; 67(1): 1–48.

[7] Fox J, Weisberg S. An {R} Companion to Applied Regression, Second Edition. Thousand Oaks CA: Sage; 2011.

URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

[8] Hothorn T, Bretz F, Westfall P. Simultaneous Inference in General Parametric Models. Biom J 2008; 50(3): 346–63.

[9] Ogle DH. FSA: Fisheries Stock Analysis. R package version 0.8.16. 2017.

[10] Harrell FE Jr. Hmisc: Harrell Miscellaneous. R package version 4.0-3; 2017. https://CRAN.R-project.org/package=Hmisc

[11] Wei T, Simko V. R package “corrplot”: Visualization of a Correlation Matrix (Version 0.84); 2017. Available from https://github.com/taiyun/corrplot

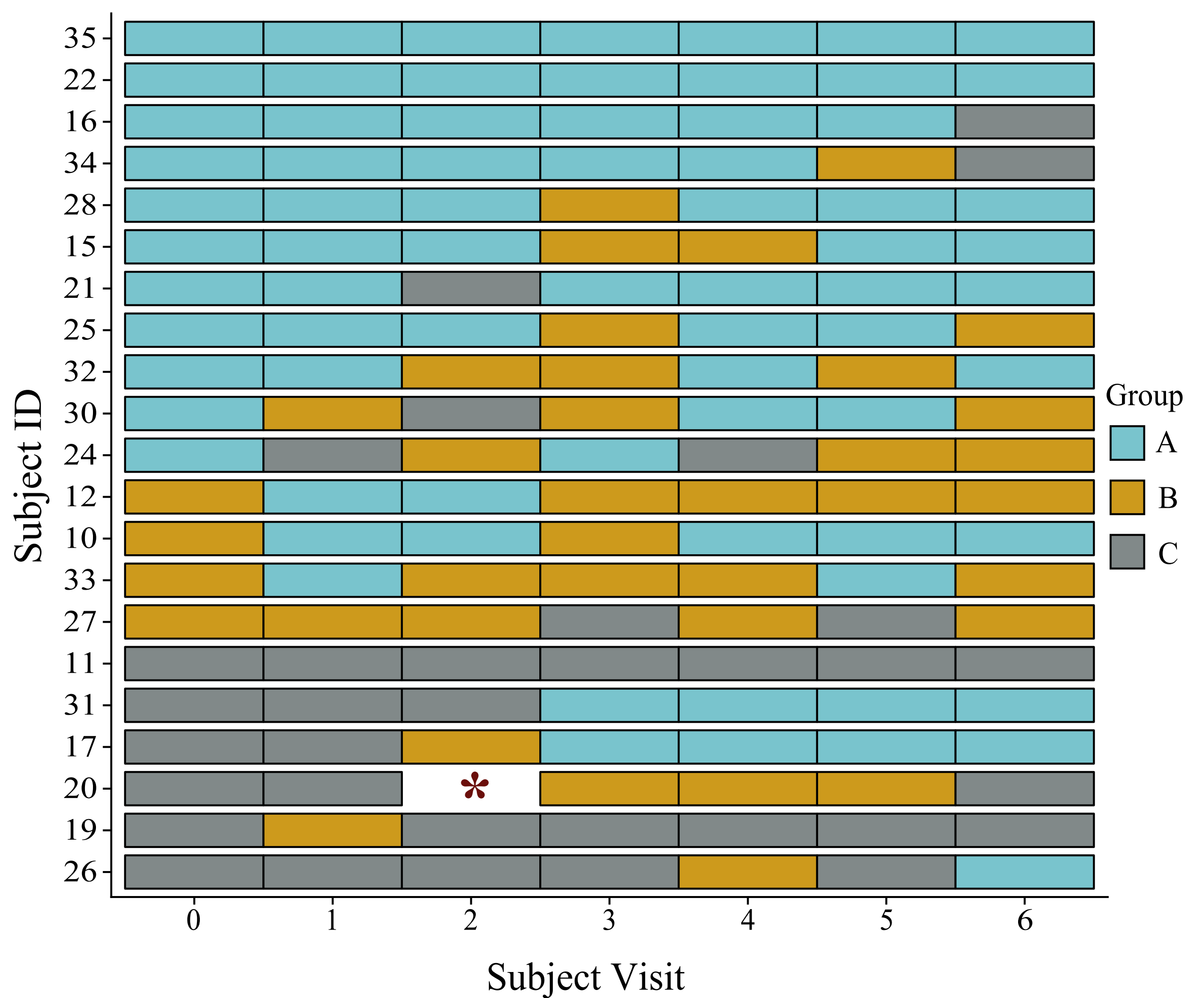


Figure B.1. Temporal changes in vaginal community composition. The occurrence of communities in groups A, B, and C are shown for each subject over time. The legend for the community groups is shown to the right of the plot. The red asterisk for Subject 20 at visit 2 indicates missing data. Subjects are vertically ordered according to the group they belonged to at their baseline visit and the number of transitions that were observed.

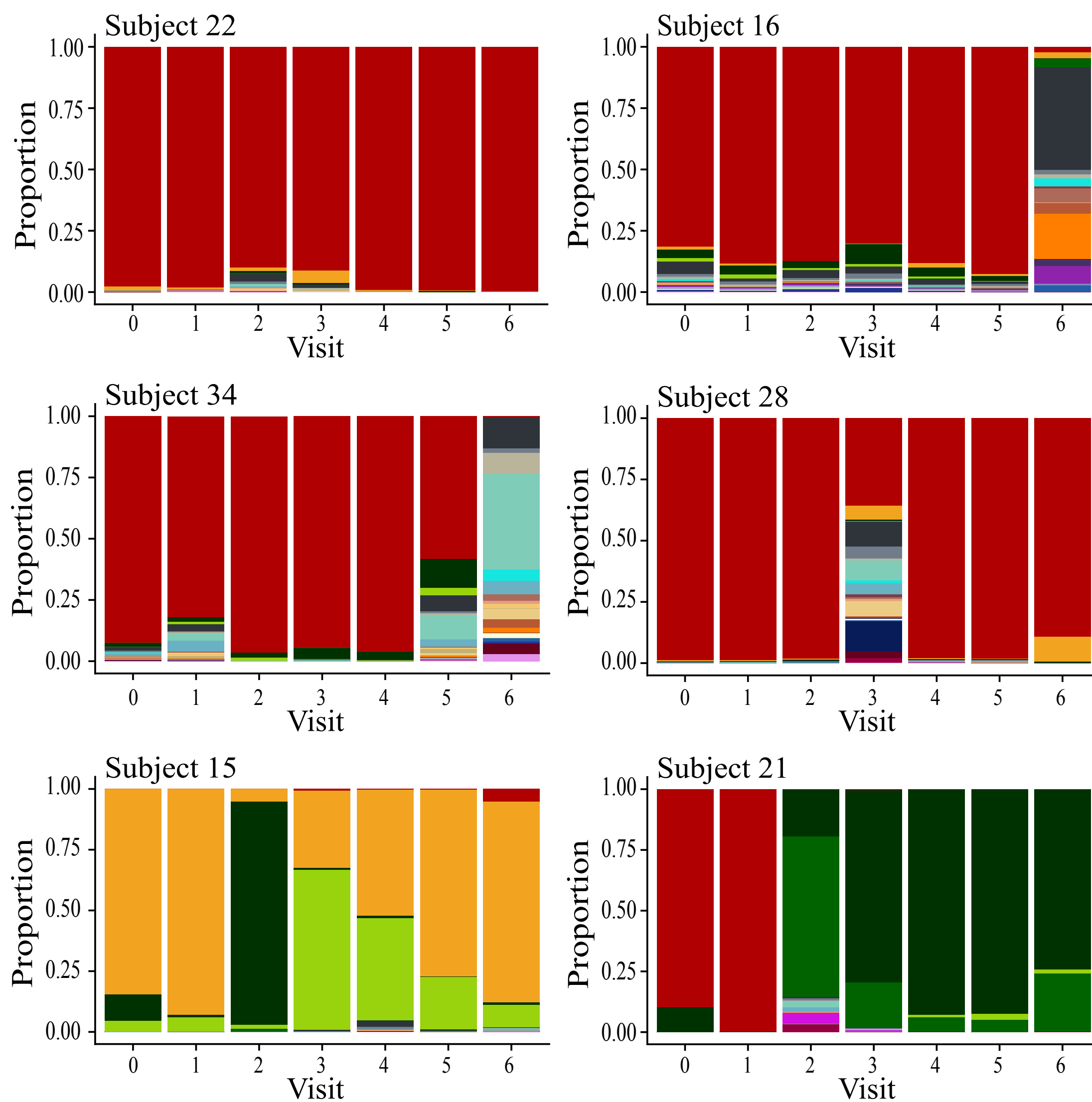
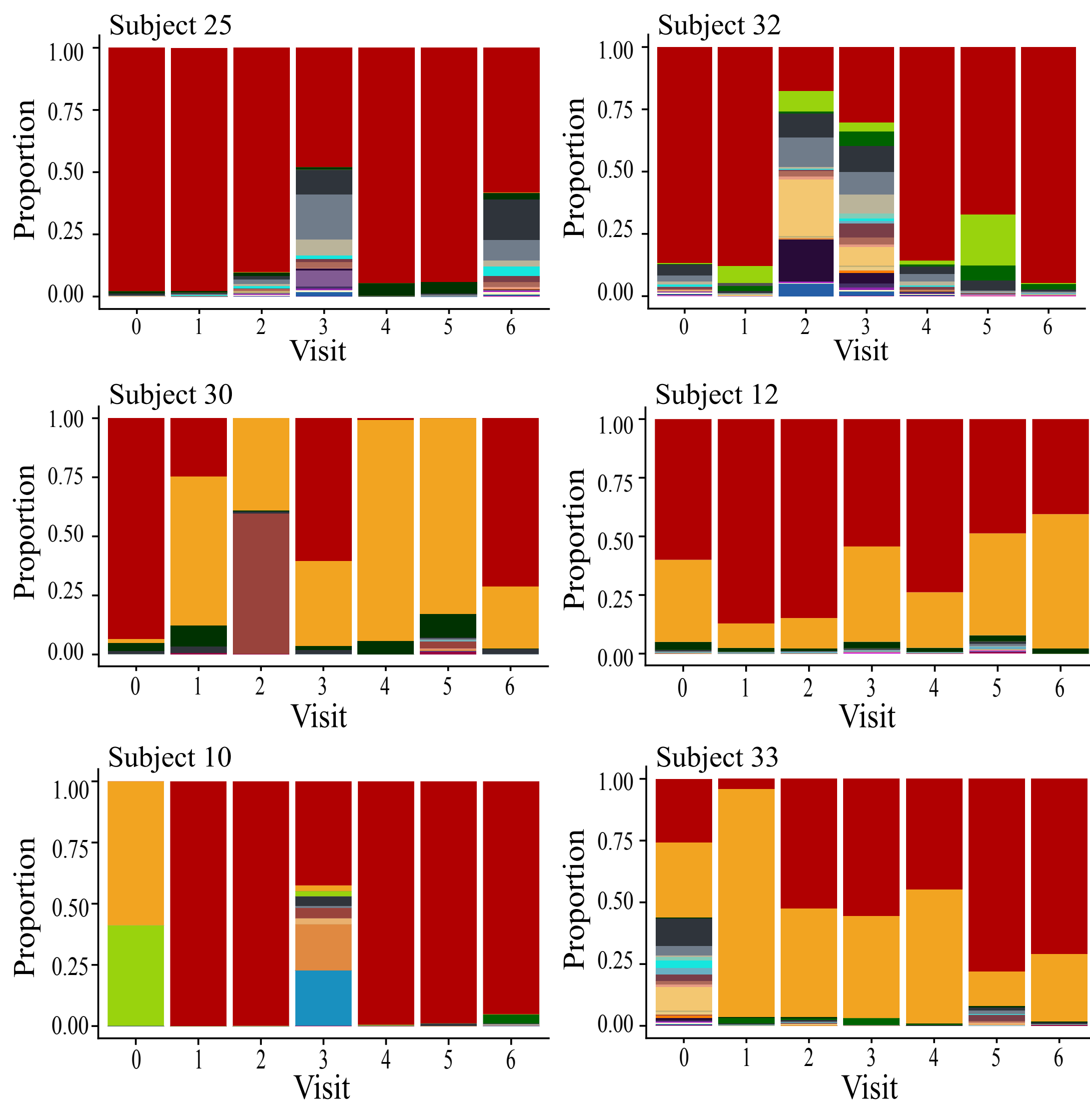
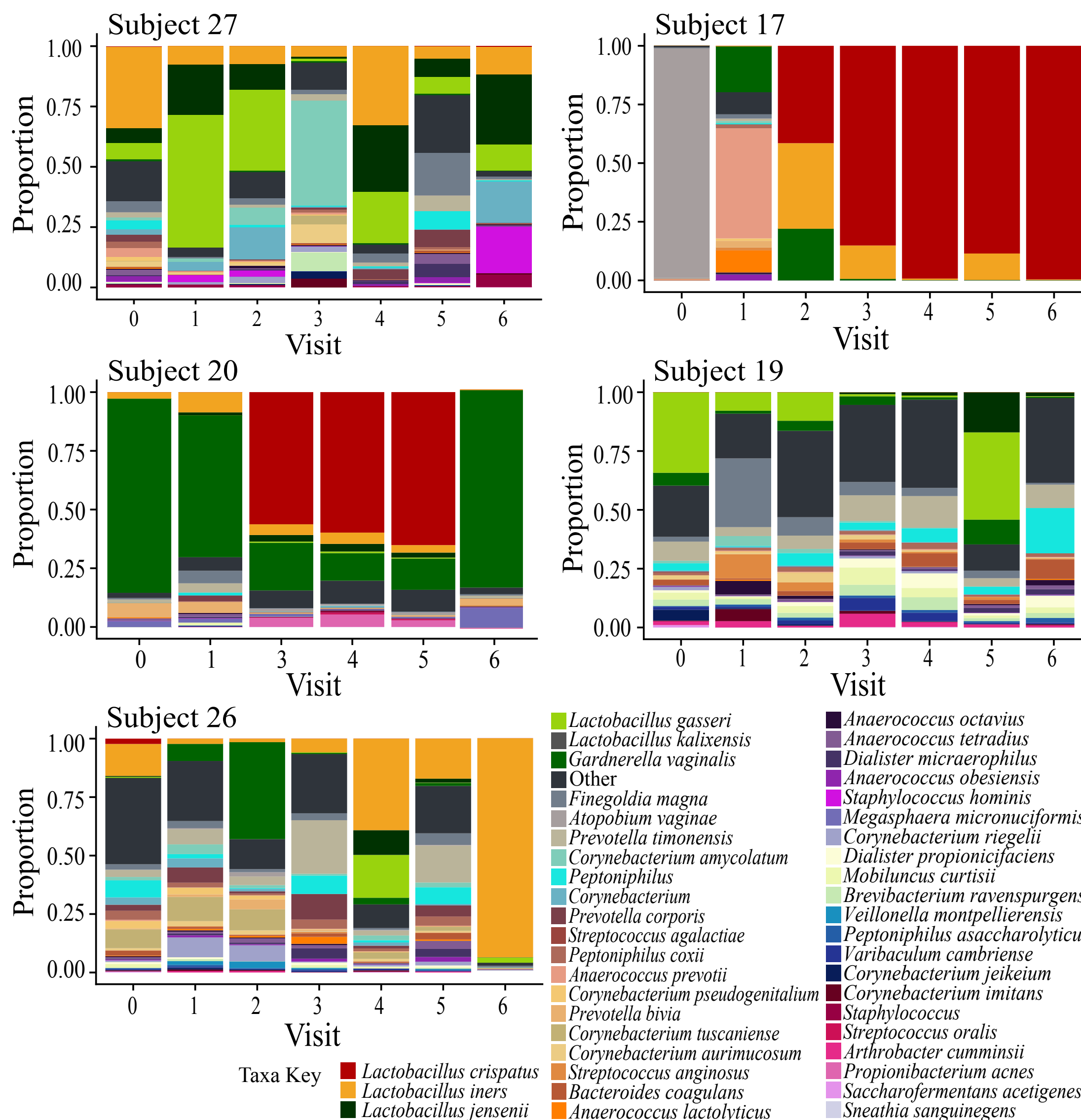


Figure B.2. Vaginal community composition of study participants sampled longitudinally. The stacked bar charts represent proportions of bacteria in each community over 7 visits at monthly intervals, where zero represents the baseline visit. Colors for each taxon are shown in the legend at the lower right corner. The graphs are ordered according to group they belonged to at their baseline visit and the number of transitions that were observed. There was no data for Subject 20 visit 2, thus this subject only has 6 visits. The stacked bar charts for Subjects 11, 24, 31, and 35 can be found in Figure 2.2. This figure continues on the next two pages.


Subject 25
Subject 32
Subject 30
Subject 12
Subject 10
Subject 33
Proportion
Visit
1.00
0.75
0.50
0.25
0.00
0
1
2
3
4
5
6


Subject 27
Subject 17
Subject 20
Subject 19
Subject 26
Proportion
Visit
1.00
0.75
0.50
0.25
0.00
0
1
2
3
4
5
6
Taxa Key
Lactobacillus crispatus
Lactobacillus iners
Lactobacillus jensenii
Lactobacillus gasseri
Lactobacillus kalixensis
Gardnerella vaginalis
Other
Finegoldia magna
Atopobium vaginae
Prevotella timonensis
Corynebacterium amycolatum
Peptoniphilus
Corynebacterium
Prevotella corporis
Streptococcus agalactiae
Peptoniphilus coxii
Anaerococcus prevotii
Corynebacterium pseudogenitalium
Prevotella bivia
Corynebacterium tuscaniense
Corynebacterium aurimucosum
Streptococcus anginosus
Bacteroides coagulans
Anaerococcus lactolyticus
Anaerococcus octavius
Anaerococcus tetradius
Dialister micraerophilus
Anaerococcus obesiensis
Staphylococcus hominis
Megasphaera micronuciformis
Corynebacterium riegelii
Dialister propionicifaciens
Mobiluncus curtisii
Brevibacterium ravenspurgense
Veillonella montpellierensis
Peptoniphilus asaccharolyticus
Varibaculum cambriense
Corynebacterium jeikeium
Corynebacterium imitans
Staphylococcus
Streptococcus oralis
Arthrobacter cumminsii
Propionibacterium acnes
Saccharofermentans acetigenes
Sneathia sanguinegens

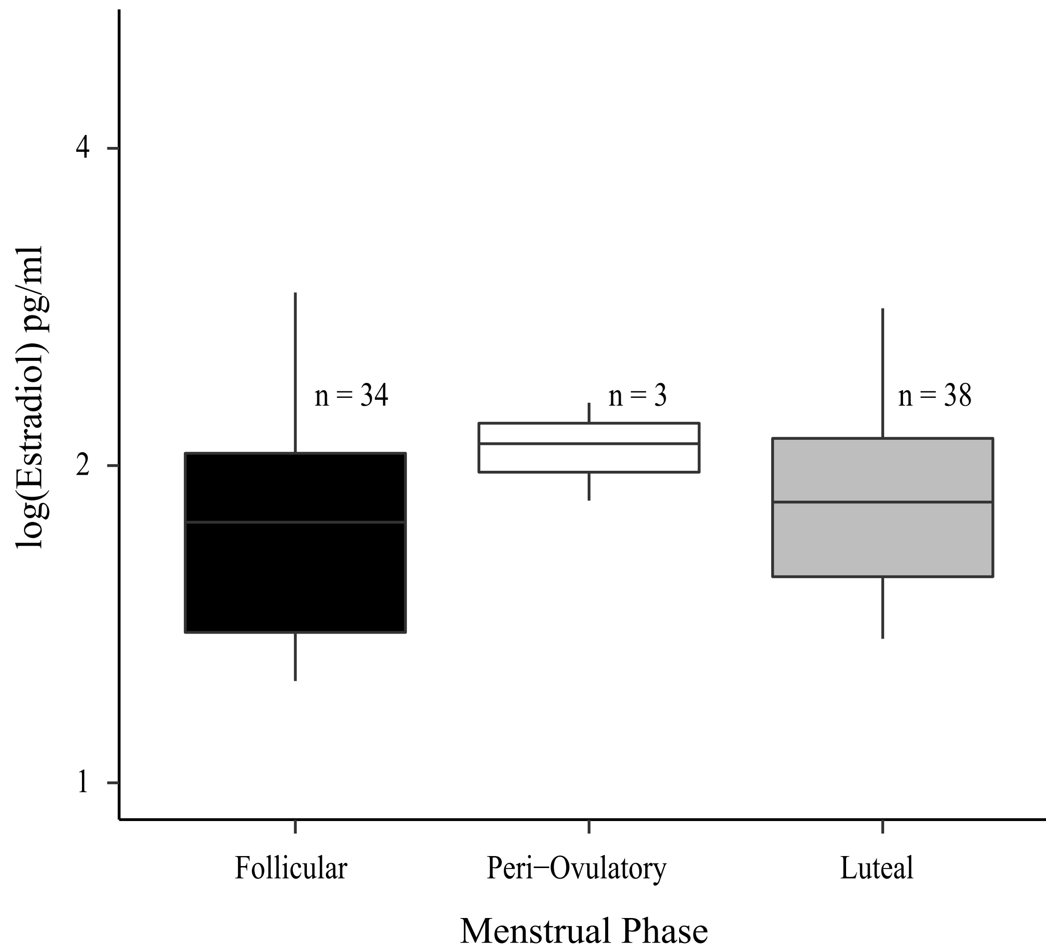


Figure B.3. Changes in estradiol over the menstrual cycle. Samples were categorized by menstrual cycle phase based on self-reported menses data. Log-transformed estradiol concentrations from women in follicular (black), peri-ovulatory (white), and luteal (grey) phases are shown in box plots. The number of samples for each phase is indicated above the corresponding box.

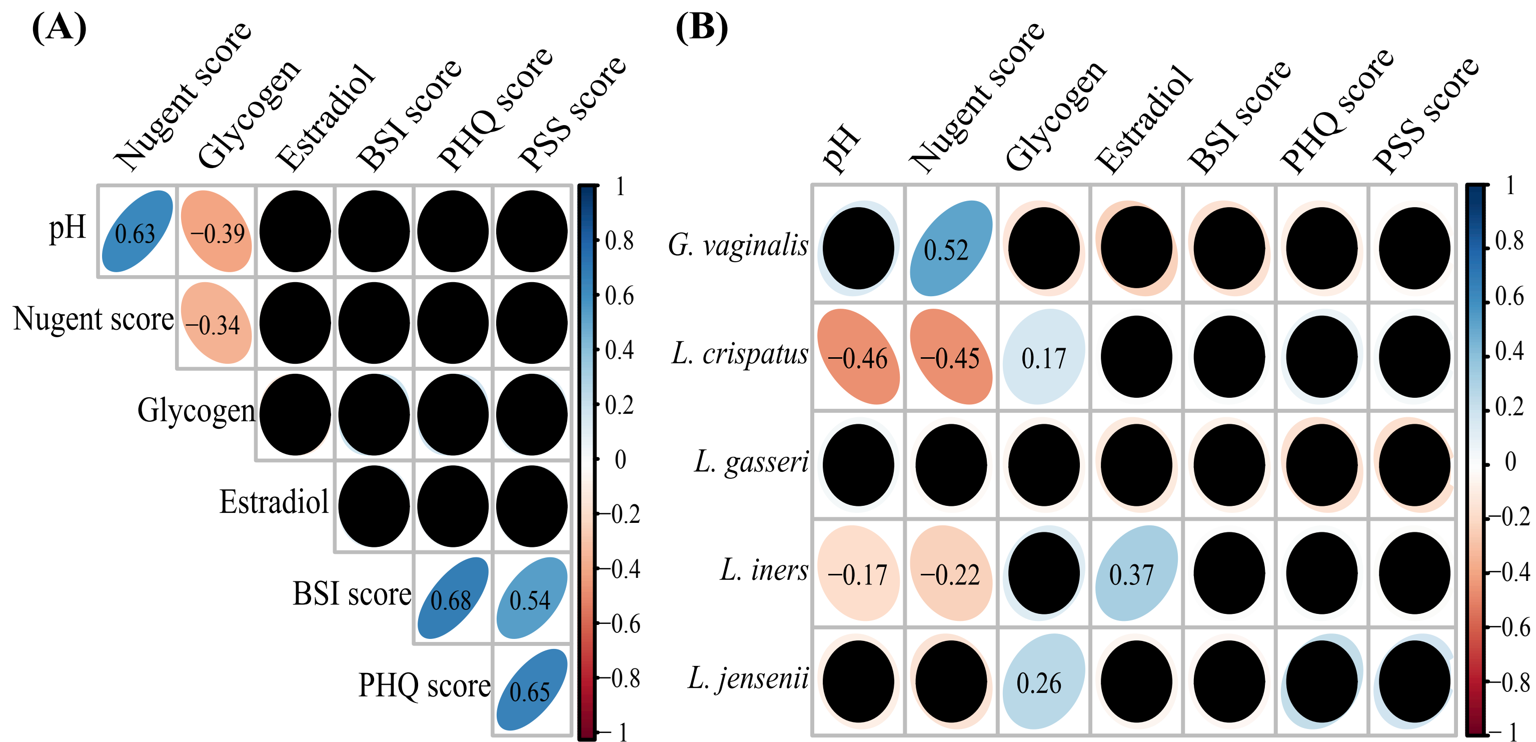


Figure B.4. Correlations between psychosocial factors, metadata, and key taxa. Panel A shows correlations between psychosocial stress factors and sample metadata. Panel B shows correlations between psychosocial stress factors and sample metadata with key taxa in the vaginal microbiome. Pearson correlation coefficients (labeled within each ellipse) were used to build these correlation plots. The sign and magnitude of the correlation is reflected in the color bar to the right of each plot. Shades of blue represent positive correlations, while shades of red represent negative correlations. Solid, black circles represent correlations that were not significant ($p > 0.05$).

Table B.1. Characteristics of study cohort

	Mean (SD)
Age	14.6 (0.3)
Age at menarche	11.8 (1.3)
Vaginal pH	5.2 (0.9)
Nugent Score	1.3 (2.5)
Glycogen (µg/ml)	324.7 (399.1)
Estradiol (pg/ml)	7.3 (4.5)
PSS Score[a]	18.5 (6.5)
BSI Score[b]	0.5 (0.7)
PHQ Score[c]	5.9 (4.8)
	Number of Subjects (%)
Hygiene:	
Douche	0 (0.0)
Birth Control:[d]	
Pill	2 (9.5)
Patch	1 (4.8)
Ring	0 (0.0)
Depo Provera	0 (0.0)
Intrauterine Device	1 (4.8)
Implant (arm rod)	2 (9.5)
Plan B - emergency contraceptive	0 (0.0)
None	16 (76.2)
Sexual Activity:	
Vaginal Sex	6 (28.6)
Oral Sex	
Given	5 (23.8)
Received	7 (33.3)
Anal Sex	0 (0.0)
Postive STI results:[e]	
Chlamydia	1 (4.8)
Gonorrhea	1 (4.8)
Trichomonas	1 (4.8)

[a]PHQ (Personal Health Questionnaire - 9) is a self-report questionnaire consisting of 10 questions (9 asking about specific symptoms, and the final asking how impactful those symptoms are to assess the severity of depression).

[b]BSI (Brief Symptom Inventory) – the anxiety subscale used here – is a self-report questionnaire consisting of 6 questions designed to clinically assess the level of anxiety in individuals.

[c]PSS (Perceived Stress Scale) is a self-report questionnaire used to evaluate the degree to which particular situations in one's life are deemed stressful.

[d]In total, five subjects reported using birth control over the course of the study. One subject reported using both the patch and the implant.

[e]One subject tested positive for all three STIs listed at her baseline visit. There were no other postive STI results from that subject or any other subject throughout the duration of the study.

Table B.2. Relative abundance data can be accessed at the following link:

https://github.com/kenettanunn/KLN_Dissertation/blob/master/Supplemental_File_Tables/NUNN_DISSERTATION_Chapter2_TableS1_TaxaAbundanceTable.xlsx

Table B.3. Mean relative proportions of key taxa in groups A, B, and C. Relative proportions were calculated for each sample individually and then averaged across samples in each group to determine the average relative proportion for key taxa in each community group. Standard deviation is shown in parentheses. Here, "Other" represents the category in which all taxa that were rare with respect to sampling were combined.

	Group		
Key Taxa	**A**	**B**	**C**
L. crispatus	0.70 (0.41)	0.33 (0.27)	0.00 (0.01)
L. iners	0.11 (0.27)	0.26 (0.21)	0.03 (0.08)
L. jensenii	0.09 (0.24)	0.08 (0.14)	0.02 (0.04)
L. gasseri	0.05 (0.19)	0.09 (0.17)	0.04 (0.10)
G. vaginalis	0.01 (0.04)	0.02 (0.05)	0.32 (0.34)
Other	0.01 (0.01)	0.05 (0.05)	0.19 (0.22)

APPENDIX C - SUPPLEMENTAL MATERIALS FOR CHAPTER 3

Approval from the University of Idaho IRB to complete study

University *of* Idaho
Office of Research Assurances
Institutional Review Board
875 Perimeter Drive, MS 3010
Moscow ID 83844-3010
Phone: 208-885-6162
Fax: 208-885-5752
irb@uidaho.edu

To: Larry J. Forney

Cc: Kenetta Nunn

From: Sharon Stoll
Chair, University of Idaho Institutional Review Board

Date: August 14, 2018

Title: Amylases in human vaginal fluids
Project: 18-118

Review Type: Expedited
Approved: 08/14/2018
Renewal: 08/13/2019

On behalf of the Institutional Review Board at the University of Idaho, I am pleased to inform you that the protocol for the research project Amylases in human vaginal fluids is approved as offering no significant risk to human subjects. This approval is valid until 08/13/2019.

This study may be conducted according to the protocol described in the application. Research that has been approved by the IRB may be subject to further appropriate review and approval or disapproval by officials of the Institution. Every effort should be made to ensure that the project is conducted in a manner consistent with the three fundamental principles identified in the Belmont Report: respect for persons; beneficence; and justice. As Principal Investigator, you are responsible for ensuring compliance with all applicable FERPA regulations, University of Idaho policies, state and federal regulations.

Federal regulations require researchers to follow specific procedures in a timely manner. For the protection of all concerned, the IRB calls your attention to the following obligations that you have as Principal Investigator of this study.

1. For any changes to the study (except to protect the safety of participants), an Amendment Application must be submitted to the IRB. The Amendment Application must be reviewed and approved before any changes can take place.

2. Any unanticipated/adverse events or problems occurring as a result of participation in this study must be reported immediately to the IRB.

3. Principal investigators are responsible for ensuring that informed consent is properly documented in accordance with 45 CFR 46.116.

4. A Continuing Renewal Application must be submitted and approved by the IRB prior to the expiration date else automatic termination of this study will occur. If the study expires, all research activities associated with the study must cease and a new application must be approved before any work can continue.

5. Please complete the Continuing Renewal/Closure form in VERAS when the project is completed.

6. Forms can be found at https://veras.uidaho.edu.

Table C.1 can be found at the following link:

https://github.com/kenettanunn/KLN_Dissertation/blob/master/Supplemental_File_Tables/NUNN_DISSERTATION_Chapter3_TableS1.xlsx

Table C.2 can be found at the following link:

https://github.com/kenettanunn/KLN_Dissertation/blob/master/Supplemental_File_Tables/NUNN_DISSERTATION_Chapter3_TableS2_TaxonAbundance.xlsx

Table C.3. Attributes of CVM samples collected from reproductive age women

	N	Mean	CI
Volume (μl)	23	469.6	383.4 - 555.8
pH	23	4.3	4.1 - 4.4
Total protein (mg/ml)	23	3.4	2.7 - 4
amylase activity (U/mg)	23	0.6	0.3 - 0.8
glycogen (mg/ml)	23	7.2	4.1 - 10.3
D-lactic acid (mM)	23	47.5	31.7 - 63.3
L-lactic acid (mM)	23	41.4	27.9 - 54.9
Total Lactic acid (mM)	23	88.9	66.6 - 111.1

Table C.4. Characteristics of samples selected for shotgun metagenomics and proteomic analysis

Subject	pH	Total protein (mg/ml)	amylase activity (U/mg)[a]	glycogen (mg/ml)	D-lactic acid (mM)	L-lactic acid (mM)	Total Lactic acid (mM)	Dominant bacteria[b]
F02	4.5	3.41	0.06	13.90	38.31	26.03	64.35	*Lactobacillus gasseri*
F06	4.4	1.40	1.32	15.70	0.63	35.11	35.75	*Lactobacillus iners*
F08	5	1.73	0.61	7.27	16.16	10.73	26.89	*Lactobacilus crispatus*
F12	3.9	1.45	0.52	32.75	119.82	45.64	165.46	*Lactobacillus crispatus*

[a]amylase activity is expressed in units of amylase (U) per mg of total protein in the sample.

[d]dominant bacteria corresponds to bacteria that have a relative abundance ≥ 50% of the vaginal community.

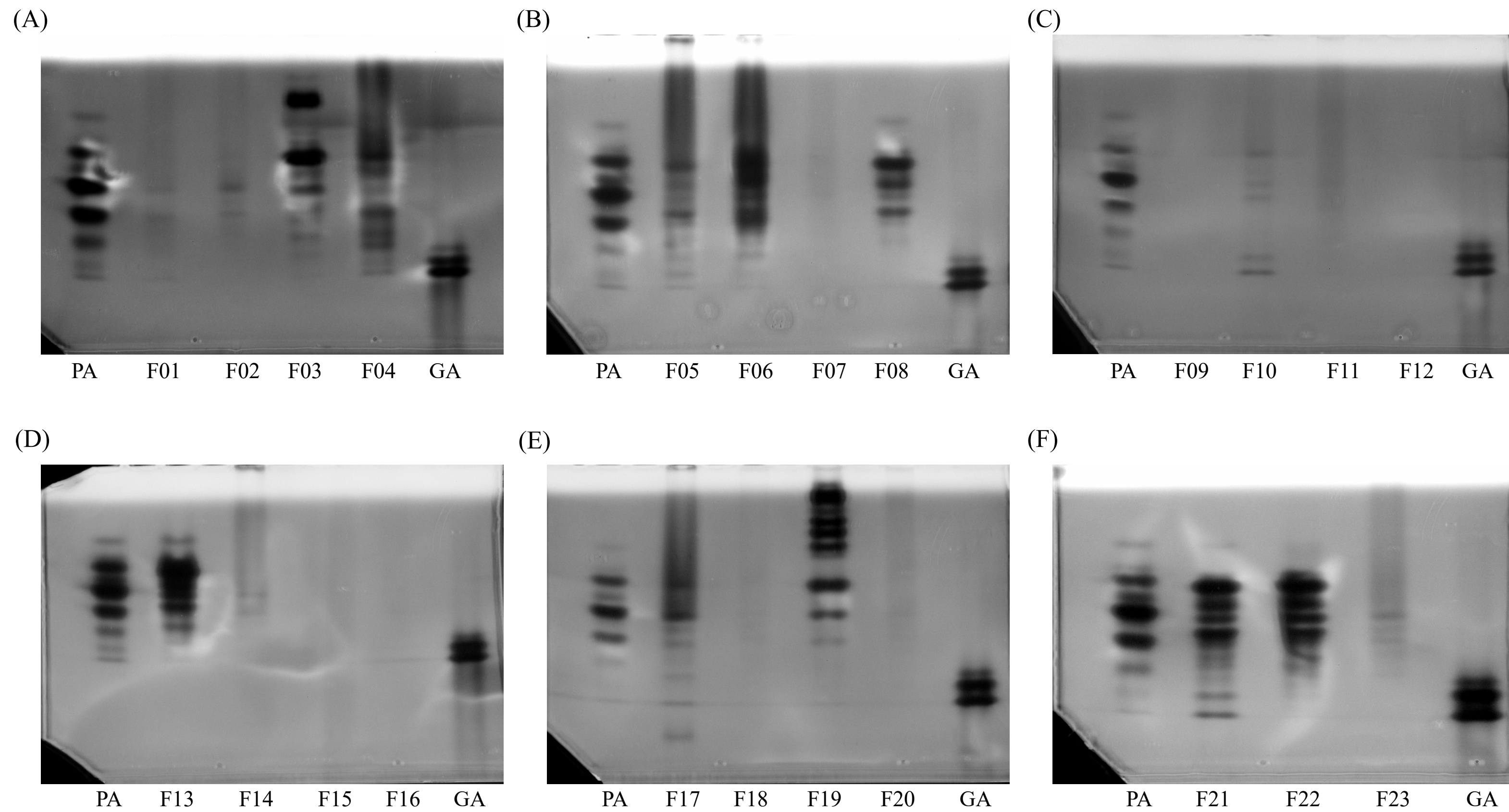


Figure C.1. Amylase enzymes in vaginal fluids collected from reproductive age women separated by native PAGE. In this figure are all of the images obtained from native PAGE gels that were incubated in 1% starch and stained with iodine solution after separating amylase proteins in CVM from all of the women in our study. These images were inverted so that the dark background shows up white and the clear zones indicating hydrolysis of starch show up dark. Thus, the dark bands reflect amylase isozymes that were resolved in each donor sample. The donor ID for each sample is below the gel image. A total of 20 μg of total protein was loaded for each sample. PA (human pancreatic α-amylase) and GA (Rhizopus spp. glucoamylase) are two commercial amylases that were used as positive controls for the assay.

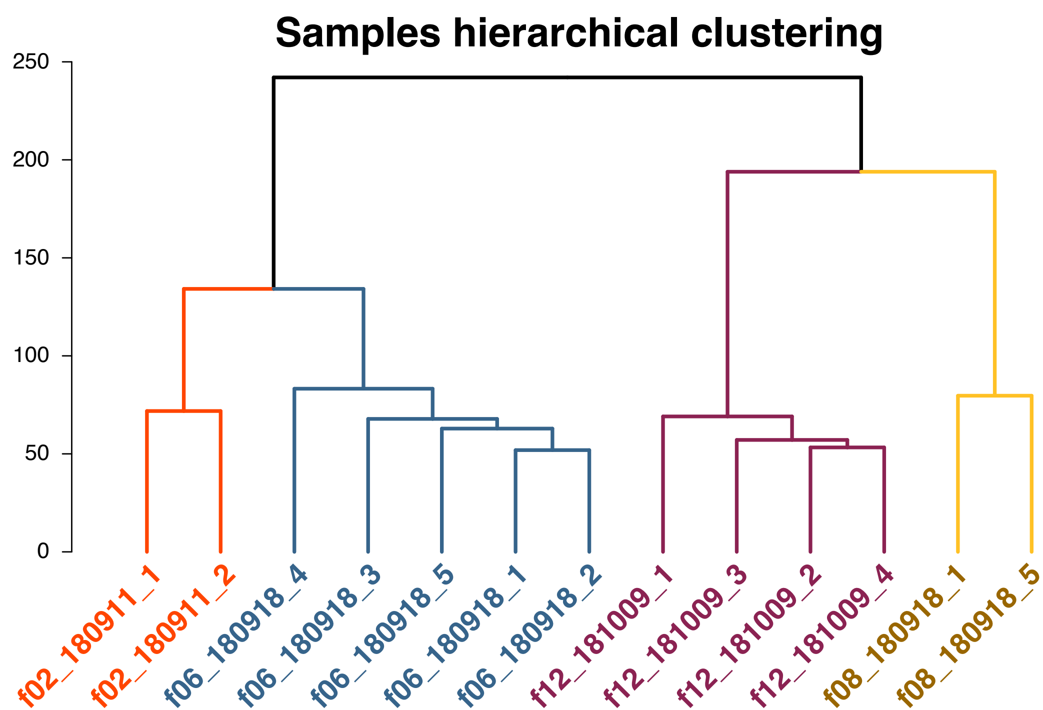


Figure C.2. Dendrogram depicting the clustering of the proteomes of the samples originating from different donors. All of the proteins detected (so a global representation) were included in the clustering analysis. Samples clustered using Ward's method on Euclidean distances calculated from protein abundances. F02 is highlighted in orange, F06 is highlighted in blue, F08 is highlighted in yellow, and F12 is highlighted in purple. Multiple branches per sample indicate the number of replicates that were processed for each. Replicates varied by donor due to sample availability.

Figure C.3. Heatmap of bacterial proteins detected in the vaginal proteome of donors. The top portion of this figure is of a heatmap that reflects the relative abundances of bacterial proteins detected in each sample. Relative abundances are indicated by the z-scores listed below the plot. Blue values indicate low abundances based on the mean of all of the proteins in a sample, whereas yellow indicates high abundances. Grey areas represent proteins that were not detected in a sample. Below the heatmap are orange and green boxes. The orange boxes indicate in which donor genome assembly the proteins were detected, and the green boxes indicate in which MAG (bacteria) the proteins were detected.

APPENDIX D - SUPPLEMENTAL MATERIALS FOR CHAPTER 4

Table D.1, which contains taxon relative abundance data can be accessed at the following link:

https://github.com/kenettanunn/KLN_Dissertation/blob/master/Supplemental_File_Tables/NUNN_DISSERTATION_Chapter4_TableS1.xlsx

Table D.2. Rank abundance of the top 10 taxa in each stage

	First Trimester		Third Trimester		Postpartum	
Rank	**Taxa**	**Count**	**Taxa**	**Count**	**Taxa**	**Count**
1	*Lactobacillus crispatus*	224555	*Lactobacillus crispatus*	416632	*Gardnerella vaginalis*	88691
2	*Gardnerella vaginalis*	112324	*Lactobacillus jensenii*	119300	*Lactobacillus iners*	52707
3	*Lactobacillus jensenii*	65349	*Lactobacillus gasseri*	78973	*Streptococcus anginosus*	52499
4	*Lactobacillus gasseri*	40472	*Lactobacillus iners*	50426	*Lactobacillus crispatus*	39129
5	*Bifidobacterium breve*	10284	*Gardnerella vaginalis*	43474	*Prevotella bivia*	17412
6	*Lactobacillus iners*	9108	*Alloscardovia omnicolens*	5606	*Streptococcus agalactiae*	15220
7	*Atopobium vaginae*	8356	*Staphylococcus hominis*	3290	*Atopobium vaginae*	14326
8	*Lactobacillus acidophilus*	3227	*Bifidobacterium breve*	2159	*Prevotella timonensis*	13572
9	*Prevotella timonensis*	2841	*Atopobium vaginae*	1706	*Sneathia sanguinegens*	13037
10	*Lachnospiraceae*	2096	*Staphylococcus haemolyticus*	1312	*Anaerococcus lactolyticus*	12248

Table D.3. Dominant bacterial species in the vaginal communities of study participants at each pregnancy stage sampled. Dominance is defined as having a relative proportion greater than 50% within the community.

Subject	First Trimester[a,b]	Third Trimester	Postpartum
49	*Lactobacillus jensenii*	*Lactobacillus iners*	*Gardnerella vaginalis*
55	*Gardnerella vaginalis*	*Lactobacillus jensenii*	*Streptococcus agalactiae*
56	*Lactobacillus crispatus*	*Lactobacillus crispatus*	*Lactobacillus crispatus*
57	*Gardnerella vaginalis*	Mixed	*Gardnerella vaginalis*
61	*Lactobacillus gasseri*	*Lactobacillus gasseri*	*Streptococcus anginosus*
63	*Lactobacillus crispatus*	NA	Mixed
65	*Lactobacillus crispatus*	*Lactobacillus crispatus*	Mixed
66	*Lactobacillus jensenii*	*Lactobacillus jensenii*	NA
70	*Lactobacillus gasseri*	*Lactobacillus gasseri*	*Anaerococcus lactolyticus*
71	*Gardnerella vaginalis*	*Gardnerella vaginalis*	*Gardnerella vaginalis*
72	*Lactobacillus gasseri*	NA	NA
74	*Lactobacillus gasseri*	*Lactobacillus crispatus*	*Bifidobacterium breve*
75	*Lactobacillus jensenii*	*Lactobacillus jensenii*	*Sneathia sanguinegens*
76	*Lactobacillus crispatus*	NA	NA
79	*Lactobacillus crispatus*	*Lactobacillus iners*	Mixed
80	*Lactobacillus crispatus*	*Lactobacillus crispatus*	Mixed
81	Mixed	*Lactobacillus gasseri*	NA
82	*Lactobacillus crispatus*	*Lactobacillus crispatus*	NA
83	*Lactobacillus crispatus*	*Lactobacillus crispatus*	*Atopobium vaginae*
84	*Lactobacillus acidophilus*	*Alloscardovia omnicolens*	NA
85	*Gardnerella vaginalis*	*Gardnerella vaginalis*	Mixed
86	*Lactobacillus jensenii*	*Lactobacillus jensenii*	*Gardnerella vaginalis*
87	*Bifidobacterium breve*	*Lactobacillus gasseri*	Mixed
89	*Lactobacillus crispatus*	*Lactobacillus iners*	*Streptococcus agalactiae*
91	*Lactobacillus crispatus*	*Lactobacillus crispatus*	*Lactobacillus iners*
92	*Gardnerella vaginalis*	*Lactobacillus gasseri*	Mixed
93	*Lactobacillus jensenii*	*Lactobacillus jensenii*	NA
94	*Lactobacillus crispatus*	*Lactobacillus crispatus*	*Anaerococcus octavius*
95	*Gardnerella vaginalis*	*Gardnerella vaginalis*	NA
96	*Lactobacillus crispatus*	*Lactobacillus crispatus*	NA
97	*Lactobacillus crispatus*	*Lactobacillus crispatus*	NA
98	*Lactobacillus crispatus*	*Lactobacillus crispatus*	NA
100	*Gardnerella vaginalis*	*Lactobacillus crispatus*	*Gardnerella vaginalis*
104	*Lactobacillus crispatus*	*Lactobacillus crispatus*	*Lactobacillus iners*
109	*Lactobacillus iners*	*Lactobacillus iners*	NA
110	*Lactobacillus jensenii*	*Lactobacillus jensenii*	*Streptococcus anginosus*
111	*Lactobacillus jensenii*	*Gardnerella vaginalis*	*Gardnerella vaginalis*
117	*Lactobacillus crispatus*	*Lactobacillus crispatus*	*Streptococcus anginosus*
118	*Lactobacillus crispatus*	*Lactobacillus crispatus*	*Gardnerella vaginalis*
121	*Lactobacillus crispatus*	*Lactobacillus crispatus*	*Lactobacillus iners*
123	*Lactobacillus crispatus*	*Lactobacillus crispatus*	*Lactobacillus crispatus*
124	NA	*Lactobacillus iners*	*Streptococcus anginosus*
128	*Lactobacillus crispatus*	*Lactobacillus crispatus*	*Anaerococcus obesiensis*
130	*Lactobacillus crispatus*	*Lactobacillus crispatus*	*Lactobacillus iners*
132	*Lactobacillus gasseri*	*Lactobacillus gasseri*	Mixed
144	*Lactobacillus jensenii*	*Lactobacillus jensenii*	*Streptococcus*
145	*Lactobacillus crispatus*	*Lactobacillus jensenii*	NA
157	*Lactobacillus crispatus*	*Lactobacillus crispatus*	NA

[a] NA = missing data; [b] Mixed = the community contained multiple species and was not dominated by any one

Table D.4. Results of linear mixed effects models of α-diversity including random slopes for Subject

Model and Parameters[a]	Model Summary[b]			
Model 1: Shannon diveristy ~ L-lactic acid + Hyaluronan + Stage + (1 | Subject) + Ɛ				
Random effects:			Number of:	
Groups	Variance	SD	Observations	Groups
Subject (Intercept)	0.010	0.102	125	48
Residual	0.215	0.464		
Fixed effects:	Coefficents	SE	df	*P* value
(Intercept)	0.326	0.092	119.475	5.27E-04 ***
L-lactic acid	-0.142	0.058	116.168	1.49E-02 *
Hyaluronan	0.283	0.070	114.766	9.34E-05 ***
Stage Third Trimester	0.073	0.124	82.956	5.56E-01
Stage Postpartum	0.427	0.166	109.714	1.14E-02 *
Model 2: Simpson diversity ~ L-lactic acid + Hyaluronan + (1 | Subject) + Ɛ				
Random effects:			Number of:	
Groups	Variance	SD	Observations	Groups
Subject (Intercept)	0.003	0.054	125	48
Residual	0.046	0.213		
Fixed effects:	Coefficients	SE	df	*P* value
(Intercept)	0.193	0.021	42.815	6.74E-12 ***
L-lactic acid	-0.060	0.021	120.516	5.03E-03 **
Hyaluronan	0.135	0.021	121.686	2.14E-09 ***

[a] Stage = pregnancy stage; Ɛ = random error; fixed effects were chosen using the step() function in the R package lmerTest (v 3.1.0)

[b] SD = standard deviation, SE = standard error; df = degrees of freedom; significance is indicated as follows: "*", $P < 0.05$; "**", $P > 0.01$; "***", $P < 0.001$

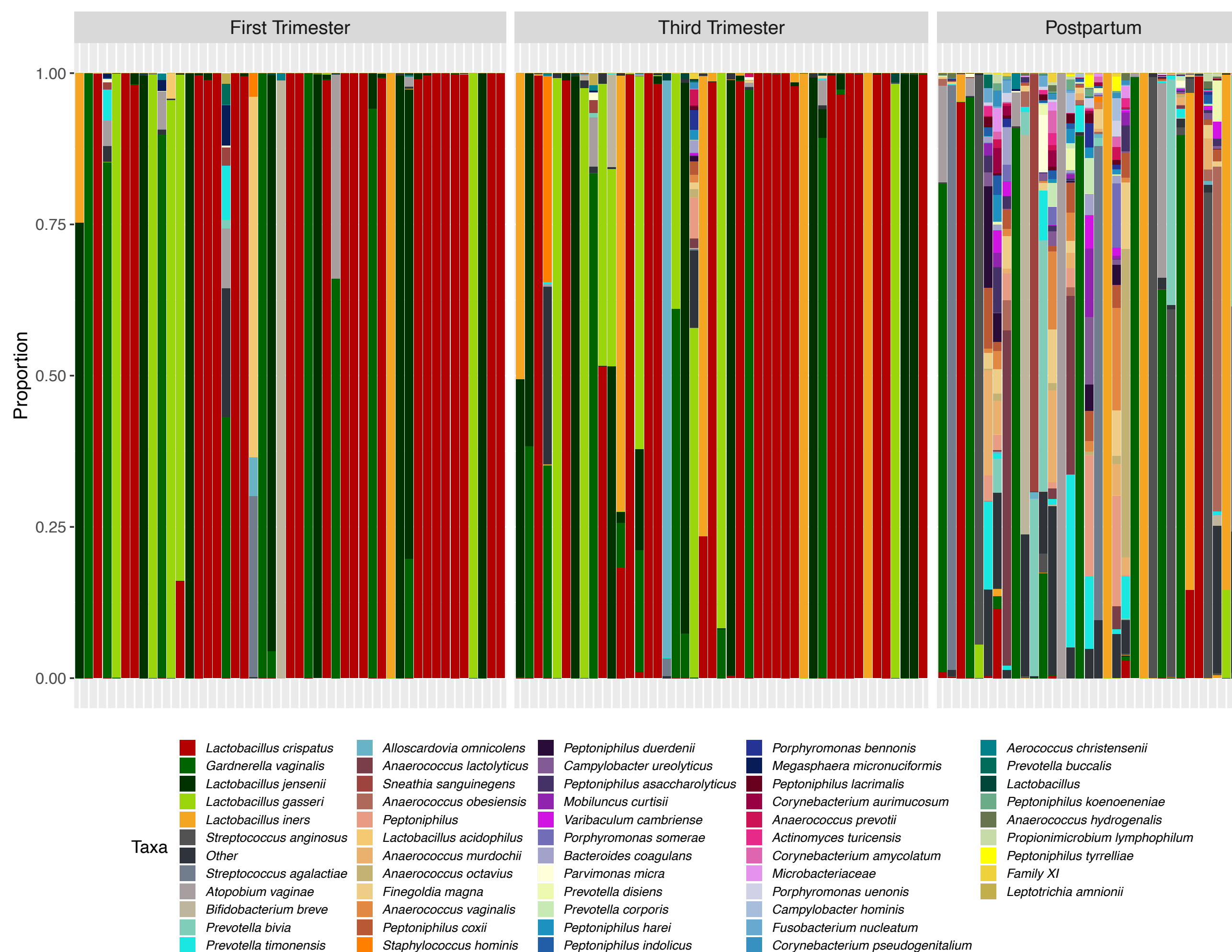


Figure D.1. Relative proportions of bacteria in the vaginal communities of 48 pregnant women separated by stage. The stacked bars represent the proportions of bacterial taxa within one sample. Bars are separated by the pregnancy stage in which the sample was collected (top headings). Taxa colors are indicated in the legend below the figure. “Other” represents the sum of all bacterial taxa that were not present at 1% in at least two women.