Unraveling Mechanisms of Bacterial Vaginosis through Metabolic Modeling and Metabolomics Analysis

Signature page

Abstract

Bacterial vaginosis (BV) is a common vaginal condition that has a significant impact on women's health. However, our understanding of its microbial community structure and metabolic interactions is limited. In this dissertation, I focus on investigating the metabolic pathways within the *Gardnerella* pangenome and defining the functional metabolic relationships between *Gardnerella* and other vaginal microbial species associated with BV. By constructing metabolic network models and conducting *in silico* analysis, I uncover both conserved and unique metabolic mechanisms within the *Gardnerella* pangenome. Notably, I find that genetic similarity does not always correspond to metabolic functional similarity.

To further explore the dynamics of *Gardnerella* strains in the vaginal metabolic environment, I used flux balance analysis to identify essential genes and potential drug targets. Through *in silico* simulations of pair-wise bacterial interactions, I observe significant clustering of *Gardnerella* species based on mutualistic benefits, underscoring the complex nature of these interactions. To validate my findings, I integrate clinical data, *in silico* analysis, *in vitro* experiments, and metabolomics to unravel the intricate BV-associated bacterial community structures.

Overall, this research enhances our understanding of BV and provides insights for personalized treatments and novel intervention development. By elucidating the metabolic relationships within the *Gardnerella* pangenome and unraveling the complex BV-associated bacterial community structures, we can improve the management and outcomes of this common vaginal condition.

Acknowledgements

Many thanks to my committee chair, Dr. Charles Farber and all of my committee members – Dr. Alison Criss, Dr. Ani Manichaikul, and Dr. Mete Civelek. You all have been continued pillars of support, and have helped me create a body of work I am truly proud of. I could not have undertaken this journey without my advisor Dr. Jason Papin, who has taught me how to lead with kindness and respect. Jason, you have helped me grow into not only a better scientist, but also a better person. Glynis, I smile just when I hear your name. You are the only reason I have survived wet-lab and maybe even had a teensy bit of fun. I'd like to acknowledge my funding sources: NIH T32 Biotechnology Training Program, NSF Expand Fellowship, Raven Fellowship, GIDI-UP, and Evvy Inc. I had the pleasure of collaborating with Dr. Amanda Lewis, Pita Navarro, Fiorella Wever, and Dr. Krystal Thomas-White. I am thankful for my mentor Dr. Jess Couch who shares not only my love of science and research equity, but also my love of yoga. Thank you to my parents, who have always supported my love for science and my very long academic journey. I would be remiss to not mention my lab mates, past and present, and especially Emma and Joe, for the innumerable laughs and coffee breaks, y'all made grad school fun. To my partner Helena, I feel truly lucky to be loved by someone who will read all of my papers, and listen to all of my hairbrained research ideas. You make me feel cared for and ground me when my thoughts go too fast. I could write a whole other thesis on why you are simply without question the best person and partner. To Debian, I think we have been best friends and sisters in every lifetime. Thank you for never saying no to MB, always letting me cry on your shoulder, knowing when to check-in, and loving me even when I can be pure chaos. You're truly a cozy cup of tea to everyone you meet. There are so many more people I could thank: Mary Nance, Becca, Clark and Lyle, Holly and Tucker, Sheba and Phineas, Shaylyn and all of QIMS, Dr. Ann Matthysse, Dr. Terry Furey, and so many more. Lastly, I want to thank every woman who has given their time, energy, and selves to make the science I do even possible.

Abstract ii

Acknowledgements iii

Chapter 1: New Horizons in the Study of the Vaginal Microbiome: Leveraging Metabolic Models and Systems Biology 1

Vaginal Health and Dysbiosis 1

Genome Scale Metabolic Modeling 2

Bacterial Community Modeling 5

Modeling the Vaginal Microbiome 8

Specific Aims 10

Chapter 2: Exploring the *Gardnerella* Pangenome's Interaction with the Vaginal Environment through Metabolic Network Models 20

Context 20

Synopsis 20

Introduction 21

Methods 22

Model construction and contextualization 22

Model comparisons 23

Gene essentiality 24

Flux comparison 24

Results 25

Strain comparisons 25

Genetic and reaction conservation 27

Gene essentiality 28

Model flux comparisons 30

Discussion 32

Pangenome content comparison 32

Gene essentiality 34

Flux analysis 35

Conclusion 35

Chapter 3: Competition and mutualism in the dysbiotic vaginal microbiome 50

Context 50

Synopsis 50

Introduction 51

Methods 53

***Gardnerella* and co-occurring species comparison** 53
Co-occurrence frequency 53
Average nucleotide identity for speciation determination 54
Dendrogram construction 54
Model construction and contextualization 55
Competition and mutualism simulation 56
Spent Media Analysis 57
Metabolomics Sample Collection 57
Mass Spectrometry Sample Preparation and Analysis 58
Growth Analysis 61
Supplemented Growth Analysis 61
62
Results 62
***In vivo* BV Co-Occurring Species** 62
Pairwise *in silico* bacterial interactions 64
Pairwise *in vitro* bacterial interactions 68
Growth Curves 68
Metabolomics 69
Supplemented Growth 70
Discussion 71
Conclusion 73
Chapter 4: Beyond the Dissertation 87
Capstone Team 87
Introduction 87
Materials and Methods 89
Biofilm Growth 89
Scanning Electron Microscopy Sample Preparation 89
Lectin Staining 90
Biofilm Carbon Source Growth 90
Biofilm Quantification 91
Enzyme Biofilm Disruption 91
Results 91
SEM Visualization 91

Lectin Staining 92
Biofilm Carbon Source Utilization 92
G. vaginalis Biofilm Enzyme Disruption 93
Discussion 94
Limitations 97
Chapter 5: Future Directions 109
Analysis Next Steps 109
Caffeate: Role as a host-microbiome intermediary 109
Identifying conserved BV metabolic units 110
Translational application 111
Defining BV Sub-categories & Improving Diagnosis 111
Developing Therapeutic Bacterial Consortia 112
Personalizing BV Treatment Regimens 114
Conclusion 114
Supplemental: Understanding COVID-19 Disease Severity - Insights from Metabolic Modeling 116
Context 116
Synopsis 117
Introduction 117
Methods 118
Patients 118
Patient plasma preparation 119
Mass spectrometry preparation and analysis 120
Biomarker identification 120
Pathway Analysis 122
Genome-scale metabolic modeling 122
Results 123
Biomarker Identification 123
Pathway Analysis 126
Genome scale metabolic modeling 128
Non-Acute COVID-19 Metabolism 130
Severe COVID-19 Metabolism 132

Chapter 1: New Horizons in the Study of the Vaginal Microbiome: Leveraging Metabolic Models and Systems Biology

Portions of the text for this chapter has been previously published as research articles here:

Dillard LR, Payne DD[*], & Papin JA (June 2021). Mechanistic models of microbial community metabolism. *Molecular Omics*. https://doi.org/10.1039/D0MO00154F

Dillard LR, Lee CY[*], Arnold KB, Papin JA (Oct. 2022) New perspectives into the vaginal microbiome with systems biology. *Trends in Microbiology.* https://doi.org/10.1016/j.tim.2022.09.011

Vaginal Health and Dysbiosis

The vaginal microbiome (VMB) is an essential component of female reproductive health, playing a critical role in fertility, pregnancy, and preventing pelvic inflammatory disease among other various infections (Ceccarani et al., 2019; Feehily et al., 2020; Houdt et al., 2018; Kong et al., 2020; Lewis & Gilbert, 2020; Mcmillan et al., 2015; Moreno et al., 2016). Despite its importance, the mechanisms that connect the composition and function of the VMB to physiological outcomes are only beginning to be understood. Currently, an optimal VMB is defined as having a non-*iners Lactobacillus* dominant bacterial population, maintaining an acidic pH to inhibit pathogenic growth, and a mild, musky scent (Gajer et al., 2012; Ravel et al., 2011). An optimal VMB promotes the health and turnover of vaginal epithelial cells (van de Wijgert, 2017). Conversely, non-optimal VMB states, particularly those associated with bacterial vaginosis (BV), involve the overgrowth of

harmful anaerobes, typically belonging to the *Gardnerella* genus, an alkaline pH, an itchy or burning sensation, and a fishy odor, which is linked to volatile organic compound breakdown (Wijgert et al., 2014). BV affects nearly one-third of reproductive-age women and increases the risk of contracting sexually transmitted infections, preterm birth, pelvic inflammatory disease, surgical complications, and decreases quality of life (Koumans et al., 2007, pp. 2001–200; Mitra et al., 2015). Current treatment options, including metronidazole and clindamycin, are associated with high rates of BV recurrence. Non-antibiotic therapeutics, including vaginal probiotics and boric acid suppositories, are rapidly developing (A. Powell et al., 2019). Most vaginal probiotic research focuses on *Lactobacillus* species (Ahire et al., 2023; Makarova et al., 2006). However, there is a lack of information regarding the optimal dose or duration of treatment, and current trials have only been efficacious within limited sub-populations. In short, there is significant need for not only improved therapeutic interventions, but also increased understanding of the community level structure and diversity that encapsulates the heterogeneity of BV. *In silico* bacterial metabolic modeling can be used to expand our understanding of intra- and inter-bacterial interactions in BV, while circumventing the fastidious nature of many vaginal species that make *in vitro* analysis difficult and time consuming.

Genome Scale Metabolic Modeling

Genome-scale metabolic network reconstructions (GENREs) are repositories of high-quality biochemical networks converted to a computationally-interpretable format (Fang et al., 2020; Gu et al., 2019). Metabolic network reconstruction begins with the annotated

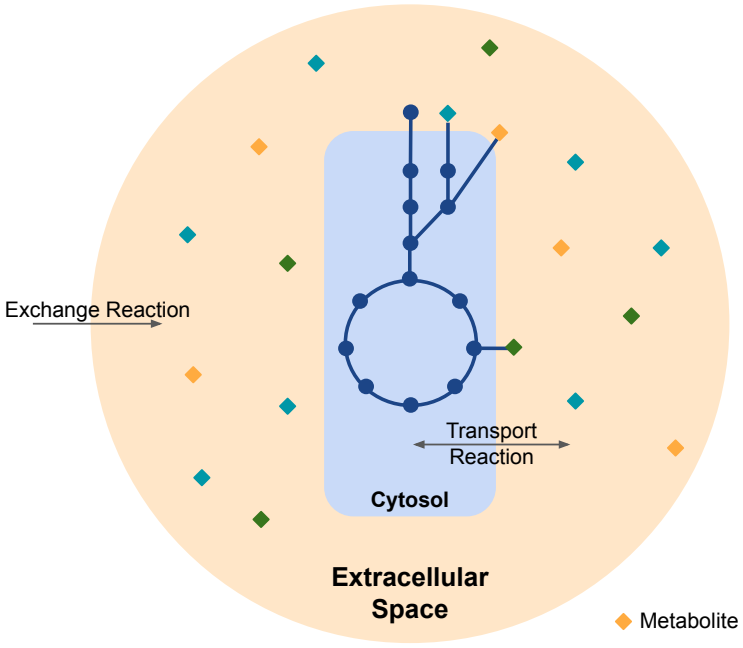


Figure 1: Visualization of GENRE

genome; the metabolic genes an organism contains determine the proteins it can synthesize and the metabolic reactions it can catalyze (**Figure 1**). These associations are stored as gene-protein-reaction (GPR) rules, with the reactants and products of each reaction cataloged in a stoichiometric matrix. Reaction bounds capture the kinetic constraints and reversibility of reactions by dictating the amount and direction of flux that an individual reaction can carry. Metabolites in the reconstruction are assigned to compartments that simulate biologicall y discrete spaces, such as the cytosol and the extracellular space. Exchange reactions introduce and remove metabolites from the extracellular space, allowing designated metabolites to be accessible for transport into the modeled organism. Transport reactions allow metabolites to flow between the extracellular and cytosolic compartments. As the metabolic network takes shape, objective functions (OFs) that represent the metabolic goal of the organism are added to the model. Specifically, biomass synthesis OFs account for all the metabolic components that must be synthesized for growth, making biomass a common OF formula.

Even with known constraints on reaction fluxes, there remains a large set of potential metabolic states of the network. Experimental data can be used to contextualize the model, trimming possible states to only those that are most biologically accurate in a given environment or experimental condition. RIPTiDe and TIMBR are both integration algorithms that utilize transcriptomic data to inform reaction weights, which then impact predicted metabolic outputs (Jenior et al., 2020; Rawls et al., 2019). RIPTiDe additionally allows for transcriptomic data to be used to "prune" reactions and metabolites from a model that do not have strong transcriptional support and are not necessary for an OF to carry minimal flux. However, due to the varying residence time of RNA (~3 minutes) and proteins (0.5–35 hours), these data offer different temporal snapshots of which metabolic functions are occurring in the organisms (Schmittgen et al., 2000). Metabolomic profiles of supernatants from growth culture experiments provide evidence for the metabolites an organism consumes and produces as it grows. This data can be integrated as bounds on the exchange reactions that force the model to mimic the uptake and secretion of metabolites that were observed *in vitro*. Metabolomics can also be analyzed as evidence for what metabolic pathways are being utilized and therefore integrated into the model as reaction weights (Hadadi et al., 2020).

As the field of metabolic modeling continues to rapidly evolve, there is a growing need for model standardization as well as metrics to assess model quality. The metabolic modeling community developed MEMOTE as a tool to conduct a uniform set of tests to assess both biological accuracy and ensure model standardization (Lieven et al., 2020). This tool provides a detailed report with quality scores and identifies gaps in the model that serve as

points for curation. Quality assessments are not the end point of the model development process, but rather a feedback mechanism to inform further curation. The iterative process of model curation is essential to obtain the most relevant and accurate biological insights.

Bacterial Community Modeling

There is currently a limited understanding of microbial metabolism at the community level, as bacterial behavior in natural, multi-species environments differs from that observed in isolation in typical laboratory experiments. The rules that govern polymicrobial community interactions remain poorly understood. To better explore the complex interactions that take place in these systems, researchers often isolate bacteria of interest and co-culture them with other bacterial species to evaluate community effects. However, the high-throughput capability of *in silico* modeling provides a powerful alternative approach for studying bacterial community metabolism.

GENREs and genome-scale metabolic models (GEMs) are valuable tools for studying bacterial community metabolism. There are different approaches to modeling bacterial community metabolism, and the best approach depends on the types of questions a community model is trying to answer. These approaches can be broken down into two main decision points. Once individual models have been reconstructed and then benchmarked using MEMOTE, the first decision point is determining how individual models will simulate bacterial interactions. Bacterial compartmentalization determines how individual bacterial species models interact with one another and with the surrounding environment. In the discrete approach, each bacterial model exists in a separate cytosolic compartment, and individual bacteria interact metabolically through a shared extracellular space where

species-specific transport reactions move communal metabolites into and out of the bacterial cytoplasm. Using this approach, Pacheco et al. simulated the metabolic relationships that emerge when pairs of microbes are grown in co-culture, identifying key cross-fed metabolites and how they impact the growth capability of each microbe (Pacheco et al., 2019).

Alternatively, there is the “bag-of-genes” approach. The bag-of-genes approach combines all the genetic makeup for the integrated models into one “super” bacterium (Faria et al., 2017; Roume et al., 2015). This approach has been used to construct metabolic models of microbial communities belonging to the human microbiome from metagenomic shotgun sequencing data (Abubucker et al., 2012). The authors explored community metabolic capabilities without the challenge of sorting reads into single-organism genomes. While such an analysis enables an interrogation of potential metabolic capabilities of a microbial community, this approach does not take into account the natural competition that occurs between bacteria as it allows for ubiquitous metabolite availability. The collective nature of the bag-of-genes approach does not consider species-specific transporters, which hinders the model's ability to accurately recapitulate physiological community interactions.

The second decision is choosing an objective function (OF). There are three broad categories that encompass approaches for defining potential OFs: individual level, community level, or multi-scale. An individual-focused OF seeks to maximize an objective such as biomass synthesis at the individual cell level, rather than being concerned with overall polymicrobial community biomass production. Community level OFs often

optimize biomass production across all species, at the expense of the biomass production of some individual species. The argument for individual-focused OF is that bacteria have not necessarily evolved for the optimization of a community biomass objective, even if participating in a community structure may be beneficial with respect to specific evolutionary objectives (Celiker & Gore, 2012; Cordero et al., 2012; Özkaya et al., 2017).

Multi-scale OFs often seek to optimize biomass production at both the individual and community level. Multi-scale OFs allow for the investigation of the trade-offs of bacteria sacrificing individual-level growth in exchange for greater community growth. These OFs strike a balance between the evolutionary concept that bacterial species focus on individual growth while also considering the evolutionary benefits of community-level metabolic interdependence. OptCom is a method that breaks down these two optimization equations into inner and outer problems (Zomorrodi & Maranas, 2012). The inner problem is defined as individual-level biomass synthesis maximization, while the outer problem is community biomass synthesis maximization. The inner problem is optimized first in order to identify the maximum biomass production of each bacterial species. The outer problem is then optimized, but individual bacteria are allowed to grow at a proportion of their optimal biomass in order to simulate sacrifice of an individual species' growth for community biomass synthesis maximization. The multi-scale tool MICOM offers the additional flexibility of manually adjusting the trade-off values assigned to individual bacteria, while also being able to integrate metagenomic community data (Diener et al., 2020).

After integrating GEMs of various bacteria into one simulated community, the metabolic interactions of a polymicrobial community can be analyzed using Flux Balance Analysis (FBA) and related modeling methods. FBA-based simulations can provide information on what genes, metabolic pathways, and metabolites are utilized by community members in simulated environments. Integrating omic data into the community model provides a layer of physiological insight in relation to gene transcription, metabolite usage, and protein synthesis, depending on the data used. Through a better understanding of community metabolism, we can improve predictions of bacterial evolution and host-pathogen interactions.

Modeling the Vaginal Microbiome

The complexity of multispecies and host-microbiome interactions makes analysis of microbiome data challenging. Therefore, quantitative systems biology approaches have been used to account for complexity while distilling key drivers of community behavior and function. Promising nascent applications of systems biology in VMB research show potential for discovering new insight into the complex roles that vaginal microbes play in women's health and disease.

The VMB presents specific research challenges, such as its dynamic nature and lack of adequate experimental and clinical models. In particular, cross-sectional data, which are commonly used to study the relationship between microbiome composition, function, and disease, may be problematic for the VMB due to its ability to shift between optimal and non-optimal states over a short period of time (Gajer et al., 2012; Song et al., 2020; Srinivasan et al., 2010). Specifically, during menses when *Lactobacillus* tends to decline

creating an environment more amenable to infection. Additionally, both *in vitro* and *in vivo* studies have limitations in accurately modeling the human VMB (Holm et al., 2020; Miller et al., 2016; Yang et al., 2018; Yildirim et al., 2014). This leads to heavy reliance on clinical measurements of microbial abundance that can be difficult to interpret due to confounding variables including: collection time point relative to menses cycle, race, age, and sexual history. In addition, the inherent characteristics of 16S rRNA gene and metagenomic data — limited resolution, abundance bias, contamination — require implementation of feature selection methods and awareness of correlation bias (Friedman & Alm, 2012; Kurtz et al., 2015). The lack of experimental models also limits the ability to investigate vaginal microbiota under controlled conditions.

The field of systems biology offers unique opportunities to address these VMB specific challenges. Data-driven models can infer microbial signatures associated with health or disease states when there are large amounts of high-throughput data available. The VMB has been extensively studied using data-driven systems biology approaches to classify and define community state types (CSTs) (Ma et al., 2012; Nunn & Forney, 2016). Hierarchical clustering methods have been used to group microbial abundance data into five distinct CSTs, three of which are associated with health, one with community transitions, and one with bacterial vaginosis (Ravel et al., 2011). Data-driven studies on CSTs and sexually-transmitted infection (STI) acquisition have also been essential for implicating non-optimal bacteria in increased STI susceptibility and linking other *Lactobacillus* to an optimal VMB and protection from STIs (Brotman et al., 2014; De Seta et al., 2019; Houdt et al., 2018; Nelson et al., 2015). Alternatively, theory-driven models provide direct mechanistic insight

into the cause-effect relationships that drive biological function. However, fewer theory-driven mechanistic models have been implemented to study the VMB despite their use for understanding microbiomes at other sites, including the gastrointestinal tract. Mechanistic modeling has the potential to define community dynamics within the VMB and link them to variability in disease-related outcomes.

Specific Aims

Aim 1. Delineate key metabolic pathways in the *Gardnerella* pangenome

1.1 *Construct and curate a metabolic network reconstruction of* Gardnerella

1.2 *Identify key metabolic pathways that are conserved vs. differential across the* Gardnerella *pangenome*

Aim 2. Define functional metabolic relationship between *Gardnerella* and other vaginal microbial species associated with dysbiosis

2.1 *Define metabolic dynamics within the* Gardnerella *pangenome*

2.2 *Define the metabolic relationship between* Gardnerella *and bacteria associated with symptomatic BV*

2.3 *Define the metabolic relationship between* Gardnerella *and bacteria associated with asymptomatic BV*

References

Abubucker, S., Segata, N., Goll, J., Schubert, A. M., Izard, J., Cantarel, B. L., Rodriguez-Mueller, B., Zucker, J., Thiagarajan, M., Henrissat, B., White, O., Kelley, S. T., Methé, B., Schloss, P. D., Gevers, D., Mitreva, M., & Huttenhower, C. (2012). Metabolic Reconstruction for Metagenomic Data and Its Application to the Human Microbiome. *PLOS Computational Biology*, *8*(6), e1002358. https://doi.org/10.1371/journal.pcbi.1002358

Ahire, J. J., Sahoo, S., Kashikar, M. S., Heerekar, A., Lakshmi, S. G., & Madempudi, R. S. (2023). In Vitro Assessment of Lactobacillus crispatus UBLCp01, Lactobacillus gasseri UBLG36, and Lactobacillus johnsonii UBLJ01 as a Potential Vaginal Probiotic Candidate. *Probiotics and Antimicrobial Proteins*, *15*(2), 275–286. https://doi.org/10.1007/s12602-021-09838-9

Bacterial Vaginosis and Anaerobic Bacteria Are Associated with Endometritis | Clinical Infectious Diseases | Oxford Academic. (n.d.). Retrieved April 20, 2023, from https://academic.oup.com/cid/article/39/7/990/495902

Brotman, R. M., Shardell, M. D., Gajer, P., Fadrosh, D., Chang, K., Silver, M. I., Viscidi, R. P., Burke, A. E., Ravel, J., & Gravitt, P. E. (2014). Association between the vaginal microbiota, menopause status, and signs of vulvovaginal atrophy. *Menopause (New York, N.Y.)*, *21*(5), 450–458. https://doi.org/10.1097/GME.0b013e3182a4690b

Ceccarani, C., Foschi, C., Parolin, C., D'Antuono, A., Gaspari, V., Consolandi, C., Laghi, L., Camboni, T., Vitali, B., Severgnini, M., & Marangoni, A. (2019). Diversity of

vaginal microbiome and metabolome during genital infections. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-50410-x

Celiker, H., & Gore, J. (2012). Competition between species can stabilize public-goods cooperation within a species. *Molecular Systems Biology*, *8*, 621. https://doi.org/10.1038/msb.2012.54

Cordero, O. X., Ventouras, L.-A., DeLong, E. F., & Polz, M. F. (2012). Public good dynamics drive evolution of iron acquisition strategies in natural bacterioplankton populations. *Proceedings of the National Academy of Sciences*, *109*(49), 20059–20064. https://doi.org/10.1073/pnas.1213344109

De Seta, F., Campisciano, G., Zanotta, N., Ricci, G., & Comar, M. (2019). The Vaginal Community State Types Microbiome-Immune Network as Key Factor for Bacterial Vaginosis and Aerobic Vaginitis. *Frontiers in Microbiology*, *10*, 2451. https://doi.org/10.3389/fmicb.2019.02451

Diener, C., Gibbons, S. M., & Resendis-Antonio, O. (2020). MICOM: Metagenome-Scale Modeling To Infer Metabolic Interactions in the Gut Microbiota. *MSystems*, *5*(1), e00606-19. https://doi.org/10.1128/mSystems.00606-19

Fang, X., Lloyd, C. J., & Palsson, B. O. (2020). Reconstructing organisms in silico: Genome-scale models and their emerging applications. *Nature Reviews Microbiology*, *18*(12), Article 12. https://doi.org/10.1038/s41579-020-00440-4

Faria, J. P., Khazaei, T., Edirisinghe, J. N., Weisenhorn, P., Seaver, S. M. D., Conrad, N., Harris, N., DeJongh, M., & Henry, C. S. (2017). Constructing and Analyzing Metabolic Flux Models of Microbial Communities. In T. J. McGenity, K. N. Timmis, & B. Nogales (Eds.), *Hydrocarbon and Lipid Microbiology Protocols:*

Genetic, Genomic and System Analyses of Communities (pp. 247–273). Springer. https://doi.org/10.1007/8623_2016_215

Feehily, C., Crosby, D., Walsh, C. J., Lawton, E. M., Higgins, S., McAuliffe, F. M., & Cotter, P. D. (2020). Shotgun sequencing of the vaginal microbiome reveals both a species and functional potential signature of preterm birth. *Npj Biofilms and Microbiomes*, *6*(1), 1–9. https://doi.org/10.1038/s41522-020-00162-8

Friedman, J., & Alm, E. J. (2012). Inferring Correlation Networks from Genomic Survey Data. *PLOS Computational Biology*, *8*(9), e1002687. https://doi.org/10.1371/journal.pcbi.1002687

Gajer, P., Brotman, R. M., Bai, G., Sakamoto, J., Schütte, U. M. E., Zhong, X., Koenig, S. S. K., Fu, L., Ma, Z., Zhou, X., Abdo, Z., Forney, L. J., & Ravel, J. (2012). Temporal dynamics of the human vaginal microbiota. *Science Translational Medicine*, *4*(132). https://doi.org/10.1126/SCITRANSLMED.3003605

Gu, C., Kim, G. B., Kim, W. J., Kim, H. U., & Lee, S. Y. (2019). Current status and applications of genome-scale metabolic models. *Genome Biology*, *20*(1), 121. https://doi.org/10.1186/s13059-019-1730-3

Hadadi, N., Pandey, V., Chiappino-Pepe, A., Morales, M., Gallart-Ayala, H., Mehl, F., Ivanisevic, J., Sentchilo, V., & Meer, J. R. van der. (2020). Mechanistic insights into bacterial metabolic reprogramming from omics-integrated genome-scale models. *Npj Systems Biology and Applications*, *6*(1), Article 1. https://doi.org/10.1038/s41540-019-0121-4

Holm, J. B., France, M. T., Ma, B., McComb, E., Robinson, C. K., Mehta, A., Tallon, L. J., Brotman, R. M., & Ravel, J. (2020). Comparative Metagenome-Assembled

Genome Analysis of “Candidatus Lachnocurva vaginae”, Formerly Known as Bacterial Vaginosis-Associated Bacterium−1 (BVAB1). *Frontiers in Cellular and Infection Microbiology*, *10*, 117. https://doi.org/10.3389/fcimb.2020.00117

Houdt, R. van, Ma, B., Bruisten, S. M., Speksnijder, A. G. C. L., Ravel, J., & Vries, H. J. C. de. (2018). Lactobacillus iners-dominated vaginal microbiota is associated with increased susceptibility to Chlamydia trachomatis infection in Dutch women: A case–control study. *Sexually Transmitted Infections*, *94*(2), 117–123. https://doi.org/10.1136/sextrans-2017-053133

Jenior, M. L., Moutinho, T. J., Dougherty, B. V., & Papin, J. A. (2020). Transcriptome-guided parsimonious flux analysis improves predictions with metabolic networks in complex environments. *PLoS Computational Biology*, *16*(4). https://doi.org/10.1371/JOURNAL.PCBI.1007099

Kong, Y., Liu, Z., Shang, Q., Gao, Y., Li, X., Zheng, C., Deng, X., & Chen, T. (2020). The Disordered Vaginal Microbiota Is a Potential Indicator for a Higher Failure of in vitro Fertilization. *Frontiers in Medicine*, *7*. https://www.frontiersin.org/articles/10.3389/fmed.2020.00217

Koumans, E. H., Sternberg, M., Bruce, C., McQuillan, G., Kendrick, J., Sutton, M., & Markowitz, L. E. (2007). The prevalence of bacterial vaginosis in the United States, 2001-2004; associations with symptoms, sexual behaviors, and reproductive health. *Sexually Transmitted Diseases*, *34*(11), 864–869. https://doi.org/10.1097/OLQ.0B013E318074E565

Kurtz, Z. D., Müller, C. L., Miraldi, E. R., Littman, D. R., Blaser, M. J., & Bonneau, R. A. (2015). Sparse and Compositionally Robust Inference of Microbial Ecological

Networks. *PLOS Computational Biology*, *11*(5), e1004226. https://doi.org/10.1371/journal.pcbi.1004226

Lewis, A. L., & Gilbert, N. M. (2020). Roles of the vagina and the vaginal microbiota in urinary tract infection: Evidence from clinical correlations and experimental models. *GMS Infectious Diseases*, *8*, Doc02. https://doi.org/10.3205/id000046

Lieven, C., Beber, M. E., Olivier, B. G., Bergmann, F. T., Ataman, M., Babaei, P., Bartell, J. A., Blank, L. M., Chauhan, S., Correia, K., Diener, C., Dräger, A., Ebert, B. E., Edirisinghe, J. N., Faria, J. P., Feist, A. M., Fengos, G., Fleming, R. M. T., García-Jiménez, B., … Zhang, C. (2020). MEMOTE for standardized genome-scale metabolic model testing. *Nature Biotechnology*, *38*(3), 272–276. https://doi.org/10.1038/s41587-020-0446-y

Ma, B., Forney, L. J., & Ravel, J. (2012). Vaginal microbiome: Rethinking health and disease. *Annual Review of Microbiology*, *66*, 371–389. https://doi.org/10.1146/annurev-micro-092611-150157

Makarova, K., Slesarev, A., Wolf, Y., Sorokin, A., Mirkin, B., Koonin, E., Pavlov, A., Pavlova, N., Karamychev, V., Polouchine, N., Shakhova, V., Grigoriev, I., Lou, Y., Rohksar, D., Lucas, S., Huang, K., Goodstein, D. M., Hawkins, T., Plengvidhya, V., … Mills, D. (2006). Comparative genomics of the lactic acid bacteria. *Proceedings of the National Academy of Sciences*, *103*(42), 15611–15616. https://doi.org/10.1073/pnas.0607117103

Mcmillan, A., Rulisa, S., Sumarah, M., Macklaim, J. M., Renaud, J., Bisanz, J. E., Gloor, G. B., & Reid, G. (2015). A multi-platform metabolomics approach identifies highly specific biomarkers of bacterial diversity in the vagina of pregnant and

non-pregnant women OPEN. *Nature Publishing Group*, *5*, 14174. https://doi.org/10.1038/srep14174

Miller, E. A., Beasley, D. E., Dunn, R. R., & Archie, E. A. (2016). Lactobacilli Dominance and Vaginal pH: Why Is the Human Vaginal Microbiome Unique? *Frontiers in Microbiology*, *7*, 1936. https://doi.org/10.3389/fmicb.2016.01936

Mitra, A., MacIntyre, D. A., Lee, Y. S., Smith, A., Marchesi, J. R., Lehne, B., Bhatia, R., Lyons, D., Paraskevaidis, E., Li, J. V., Holmes, E., Nicholson, J. K., Bennett, P. R., & Kyrgiou, M. (2015). Cervical intraepithelial neoplasia disease progression is associated with increased vaginal microbiome diversity. *Scientific Reports*, *5*(1), Article 1. https://doi.org/10.1038/srep16865

Moreno, I., Codoñer, F. M., Vilella, F., Valbuena, D., Martinez-Blanch, J. F., Jimenez-Almazán, J., Alonso, R., Alamá, P., Remohí, J., Pellicer, A., Ramon, D., & Simon, C. (2016). Evidence that the endometrial microbiota has an effect on implantation success or failure. *American Journal of Obstetrics and Gynecology*, *215*(6), 684–703. https://doi.org/10.1016/j.ajog.2016.09.075

Nelson, T. M., Borgogna, J.-L. C., Brotman, R. M., Ravel, J., Walk, S. T., & Yeoman, C. J. (2015). Vaginal biogenic amines: Biomarkers of bacterial vaginosis or precursors to vaginal dysbiosis? *Frontiers in Physiology*, *6*, 253. https://doi.org/10.3389/fphys.2015.00253

Nunn, K. L., & Forney, L. J. (2016). Unraveling the Dynamics of the Human Vaginal Microbiome. *The Yale Journal of Biology and Medicine*, *89*(3), 331–337.

Özkaya, Ö., Xavier, K. B., Dionisio, F., & Balbontín, R. (2017). Maintenance of Microbial Cooperation Mediated by Public Goods in Single- and Multiple-Trait

Scenarios. *Journal of Bacteriology*, *199*(22), e00297-17. https://doi.org/10.1128/JB.00297-17

Pacheco, A. R., Moel, M., & Segrè, D. (2019). Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-018-07946-9

Powell, A., Ghanem, K. G., Rogers, L., Zinalabedini, A., Brotman, R. M., Zenilman, J., & Tuddenham, S. (2019). Clinicians' use of Intravaginal Boric Acid Maintenance Therapy for Recurrent Vulvovaginal Candidiasis and Bacterial Vaginosis. *Sexually Transmitted Diseases*, *46*(12), 810–812. https://doi.org/10.1097/OLQ.0000000000001063

Ravel, J., Gajer, P., Abdo, Z., Schneider, G. M., Koenig, S. S. K., McCulle, S. L., Karlebach, S., Gorle, R., Russell, J., Tacket, C. O., Brotman, R. M., Davis, C. C., Ault, K., Peralta, L., & Forney, L. J. (2011). Vaginal microbiome of reproductive-age women. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(SUPPL. 1), 4680–4687. https://doi.org/10.1073/pnas.1002611107

Rawls, K. D., Blais, E. M., Dougherty, B. V., Vinnakota, K. C., Pannala, V. R., Wallqvist, A., Kolling, G. L., & Papin, J. A. (2019). Genome-Scale Characterization of Toxicity-Induced Metabolic Alterations in Primary Hepatocytes. *Toxicological Sciences*, *172*(2), 279. https://doi.org/10.1093/TOXSCI/KFZ197

Roume, H., Heintz-Buschart, A., Muller, E. E. L., May, P., Satagopam, V. P., Laczny, C. C., Narayanasamy, S., Lebrun, L. A., Hoopmann, M. R., Schupp, J. M., Gillece, J.

D., Hicks, N. D., Engelthaler, D. M., Sauter, T., Keim, P. S., Moritz, R. L., & Wilmes, P. (2015). Comparative integrated omics: Identification of key functionalities in microbial community-wide metabolic networks. *Npj Biofilms and Microbiomes*, *1*(1), Article 1. https://doi.org/10.1038/npjbiofilms.2015.7

Schmittgen, T. D., Zakrajsek, B. A., Mills, A. G., Gorn, V., Singer, M. J., & Reed, M. W. (2000). Quantitative Reverse Transcription–Polymerase Chain Reaction to Study mRNA Decay: Comparison of Endpoint and Real-Time Methods. *Analytical Biochemistry*, *285*(2), 194–204. https://doi.org/10.1006/abio.2000.4753

Song, S. D., Acharya, K. D., Zhu, J. E., Deveney, C. M., Walther-Antonio, M. R. S., Tetel, M. J., & Chia, N. (2020). Daily Vaginal Microbiota Fluctuations Associated with Natural Hormonal Cycle, Contraceptives, Diet, and Exercise. *MSphere*, *5*(4), e00593-20. https://doi.org/10.1128/mSphere.00593-20

Srinivasan, S., Liu, C., Mitchell, C. M., Fiedler, T. L., Thomas, K. K., Agnew, K. J., Marrazzo, J. M., & Fredricks, D. N. (2010). Temporal variability of human vaginal bacteria and relationship with bacterial vaginosis. *PloS One*, *5*(4), e10197. https://doi.org/10.1371/journal.pone.0010197

van de Wijgert, J. H. H. M. (2017). The vaginal microbiome and sexually transmitted infections are interlinked: Consequences for treatment and prevention. *PLoS Medicine*, *14*(12), e1002478. https://doi.org/10.1371/journal.pmed.1002478

Wijgert, J. H. H. M. van de, Borgdorff, H., Verhelst, R., Crucitti, T., Francis, S., Verstraelen, H., & Jespers, V. (2014). The Vaginal Microbiota: What Have We Learned after a Decade of Molecular Characterization? *PLOS ONE*, *9*(8), e105998. https://doi.org/10.1371/journal.pone.0105998

Yang, E., Fan, L., Yan, J., Jiang, Y., Doucette, C., Fillmore, S., & Walker, B. (2018). Influence of culture media, pH and temperature on growth and bacteriocin production of bacteriocinogenic lactic acid bacteria. *AMB Express*, *8*(1), 10. https://doi.org/10.1186/s13568-018-0536-0

Yildirim, S., Yeoman, C. J., Janga, S. C., Thomas, S. M., Ho, M., Leigh, S. R., Primate Microbiome Consortium, White, B. A., Wilson, B. A., & Stumpf, R. M. (2014). Primate vaginal microbiomes exhibit species specificity without universal Lactobacillus dominance. *The ISME Journal*, *8*(12), 2431–2444. https://doi.org/10.1038/ismej.2014.90

Zomorrodi, A. R., & Maranas, C. D. (2012). OptCom: A Multi-Level Optimization Framework for the Metabolic Modeling and Analysis of Microbial Communities. *PLOS Computational Biology*, *8*(2), e1002363.

Chapter 2: Exploring the *Gardnerella* Pangenome's Interaction with the Vaginal Environment through Metabolic Network Models

The text for this chapter has been previously published as a research article here:

Dillard LR, Glass EM, Lewis AL, Thomas-White K, Papin JA (Dec. 2022) Metabolic network models of the *Gardnerella* pangenome identify key interactions in the vaginal environment. *mSystems.* https://doi.org/10.1128/msystems.00689-22

Context

After establishing my interest in studying the vaginal microbiome and inter-bacterial interactions, I began with the question "What do we know about dysbiotic vaginal bacterial communities?". My focus was on *Gardnerella*, because it is a primary biomass contributor in bacterial vaginosis. Due to the gender health gap and the rapidly evolving availability of whole genome sequences, little was known regarding metabolic diversity within the *Gardnerella* genus, but the BV-BRC data base provided sequences to begin exploring. This precipitated my first dissertation-focused project which was to define the functional metabolic capacity and diversity of the *Gardnerella* pangenome.

Synopsis

We present the first *Gardnerella* genome-scale metabolic network reconstructions repository. We used this toolbox of GENREs to analyze metabolic conservation and diversity, as well as predicting conserved essential genes within cervicovaginal fluid

context. Additionally, we investigated how strains of *Gardnerella* interact differently with their extracellular environment via flux balance analysis and dimensionality reduction.

Introduction

Bacterial vaginosis (BV) is a commonly occurring vaginal condition in reproductive-age women with vaginal complaints, characterized by low levels of *Lactobacillus*, high levels of diverse anaerobes, a vaginal pH of >4.5, thin vaginal discharge, and a fishy odor (Schwiertz et al., 2006; J. Wang, 2000). BV disproportionately impacts women of color in North America, with an estimated annual healthcare-associated cost globally of $4.8 billion and an additional $9.6 billion when accounting for BV-associated HIV infection and BV-associated preterm birth (Allsworth et al., 2008; Culhane et al., 2002; Koumans et al., 2007; Muzny & Kardas, 2020; Peebles et al., 2019). Despite BV's pervasiveness and its impact on women's health, treatment options are limited and often ineffective (Bradshaw & Sobel, 2016). The bacterial etiology of BV and the mechanisms of pathogenic outcomes remain largely ill-defined. *Gardnerella* has consistently been reported as being one of the dominant genera in the vagina during BV, thus understanding the functional metabolism of *Gardnerella* is critical (Schwebke et al., 2014). However, the *Gardnerella* pangenome remains largely uncharacterized, as noted by the rapidly evolving species classifications within this genus (Qin & Xiao, 2022). Metabolic predictions via *in silico* analysis offer a unique opportunity to study taxonomic relatedness based on inferred function. By characterizing the *in silico* models, which reflect the protein coding genetic content and metabolism of the *Gardnerella* pangenome, we can identify potential antibiotic targets, both strain-specific and conserved, and make predictions regarding differential pathogenesis. This study aims to define the conserved metabolic functions and strain-level

variation within the *Gardnerella* pangenome, making testable predictions about microbial physiology and providing structure to the heterogeneous nature of BV.

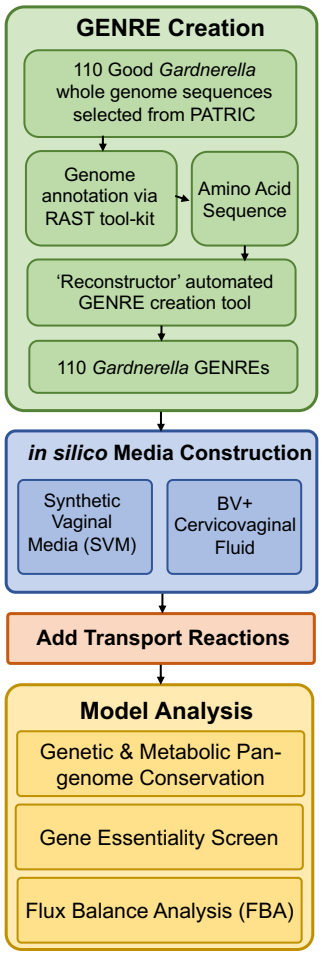


Figure 1: Analysis workflow for pangenome model reconstruction, contextualization, and analysis

Methods

Model construction and contextualization

To conduct our *in silico* analysis of the *Gardnerella* pangenome, we first identified 110 whole-genome sequences from the BV-BRC 3.6.12 database that met the quality criteria of being sufficiently complete (80%) and low contamination (10%), with amino acid sequences that were at least 87% consistent with known protein sequences (Wattam et al., 2017). We then annotated the corresponding amino acid sequences using RAST 2.0 (Aziz et al., 2008; Brettin et al., 2015; Overbeek et al., 2014). We used the Reconstructor algorithm to create genome-scale metabolic models for each of the 110 strains (**Figure 1**) (Jenior et al., Under Review). To assess the quality of the models, we used MEMOTE score, which is a widely used field-standard quality assessment tool (Lieven et al., 2020). The results are available on the associated GitHub repository (https://github.com/emmamglass/Gardnerella_Pangenome). We defined two *in silico* media conditions: synthetic vaginal media (SVM) and bacterial vaginosis-positive cervicovaginal fluid media (BVCFM). The SVM condition was based on previously defined *in vitro* media used for the growth of vaginal microflora, which has

been shown to support *Gardnerella* growth (Geshnizgani & Onderdonk, 1992). The BVCFM condition was based on previous metabolomics analyses of cervicovaginal fluid samples from both healthy and BV-positive patients, and was enriched for metabolites that were found at significantly higher levels in BV-positive samples (Vitali et al., 2015). We added transport reactions to the reconstructions as necessary to enable *in silico* utilization of media metabolites.

Model comparisons

We utilized BLASTp output annotations to determine gene presence for each strain, and constructed binary matrices indicating the presence or absence of each protein coding gene and model reaction for each strain (Madden, 2002). The dendrograms of gene and reaction presence were constructed via the dendextend R package using the Sørensen-Dice method for dissimilarity matrices construction and hierarchical clustering using the Ward method (Dice, 1945; Galili, 2015; Ward, 1963). Entanglement values were calculated, and k-means clusters were represented by branch coloring. We classified genes and reactions as core (>75%), peripheral (25-75%), or unique (<25%) based on their presence across models.

We also used the Resistance Gene Identifier 5.2.1 platform to predict antibiotic resistance based on the amino acid sequences of each strain, considering genes with >80% regional match based on protein sequence to have a conserved mutant allele (Alcock et al., 2020). These methods allowed us to characterize the *Gardnerella* pangenome, and provide insight into functional relatedness of strains and potential antibiotic resistance.

Gene essentiality

To contextualize the metabolic models, the corresponding exchange reactions were opened with flux bounds of -1,000 to 1,000 for components of SVM and BVCFM (**Table S1**). This resulted in the creation of two contextualized models for each of the 110 strains. To determine gene essentiality, we used the gene essentiality function in the COBRApy toolbox (Ebrahim et al., 2013). This function simulates single-gene deletions for every gene in the model, and if a deletion results in a >80% reduction in flux through the objective function (biomass synthesis), the gene is categorized as essential. The KEGG ortholog values for each essential gene were identified and further analyzed. After running the gene essentiality screen, we generated a heatmap using the pheatmap package in R, which utilizes Euclidean distance and complete clustering to determine hierarchical structure (Kolde, 2019).

Flux comparison

To assess the variability of metabolic flux across the 110 *Gardnerella* strains in cervicovaginal fluid medium, we collected 100 flux samples for each strain using the COBRApy-compatible GAPSPLIT function (Keaty et al., 2020). To visualize the clustering of strains based on their simulated flux distributions, we used dimensionality reduction via t-distributed stochastic neighbor embedding (tSNE) on the collected flux samples. The tSNE analysis was performed using the sci-kit learn sklearn.manifold package in Python with default parameters (Maaten et al., 2008; Pedregosa et al., 2012). We utilized strain metadata from BV-BRC to map the sample isolation sources for tSNE visualization. Furthermore, we extracted transport reactions from the flux sampling data, and created a heatmap by using the median values of the top 25 most variable transport

reaction fluxes across models. The pheatmap library in R was used to construct the heatmap, which utilizes Euclidean distance and complete clustering to determine hierarchical structure.

Results

Strain comparisons

Our analysis revealed that the number of protein-encoding genes varied greatly across the pangenome, ranging from 434 to 1,012, with a median value of 471. Similarly, the number of genes in the metabolic models ranged from 431 to 688, with a median value of 468. The number of model metabolites and reactions also showed significant variation, ranging from 782 to 1,077 and 752 to 1,012, respectively, with median values of 873 and 818.

Interestingly, our analysis identified six outlier strains across all four categories (**Figure 2A**). To further investigate the relationship between the genetic and metabolic content of the *Gardnerella* strains, we performed hierarchical clustering based on protein coding genes and model reaction content. The resulting dendrograms showed a dissimilarity of 61%, suggesting that there is a significant degree of genetic and metabolic diversity among the *Gardnerella* strains (**Figure 2B**).

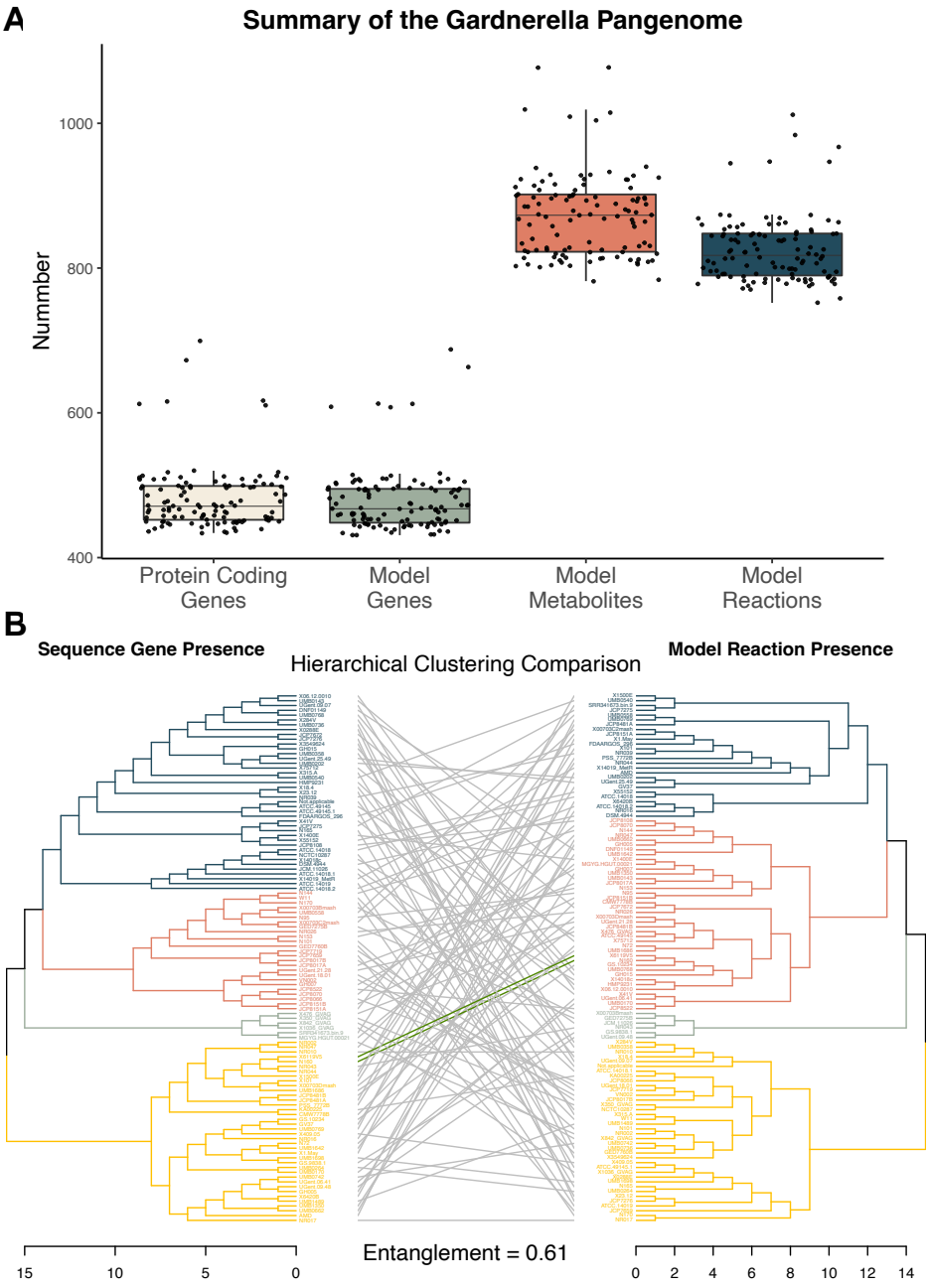


Figure 2: A) summary of pangenome sequence content and corresponding metabolic model content **B)** Comparison of hierarchical clustering based on protein coding gene presence (left) and model reaction presence (right). Branch color indicates the associated K-means group. Green entanglement lines indicate the strain was placed in the same branch for both dendrograms.

Genetic and reaction conservation

We also analyzed the *Gardnerella* pangenome to identify core, peripheral, and unique genes. Our analysis revealed that 318 genes were considered core, 90 genes were considered peripheral, and 359 genes were considered unique (**Figure 3A**). Furthermore, we found a high degree of conservation of genes implicated in antibiotic resistance across the pangenome, with rifamycin resistance genes being present in 98% of the strains. Other resistance genes, such as those implicated in mupirocin, streptogramin, lincosamide, and pleuromutilin resistance, were also commonly found, whereas tetracycline, macrolide, and aminoglycoside resistance genes were present in a minority of strains (**Figure 3B**).

In addition to analyzing genetic content, we also analyzed the reaction content of our 110 *Gardnerella* strain metabolic models. Our analysis revealed that 695 reactions were considered core, 221 were peripheral, and 919 were unique (**Figure 3C**). We also found that certain metabolic categories were enriched for either unique or core reactions. For instance, amino acid metabolism was enriched for unique reactions, whereas glycan biosynthesis and metabolism was enriched for core reactions.

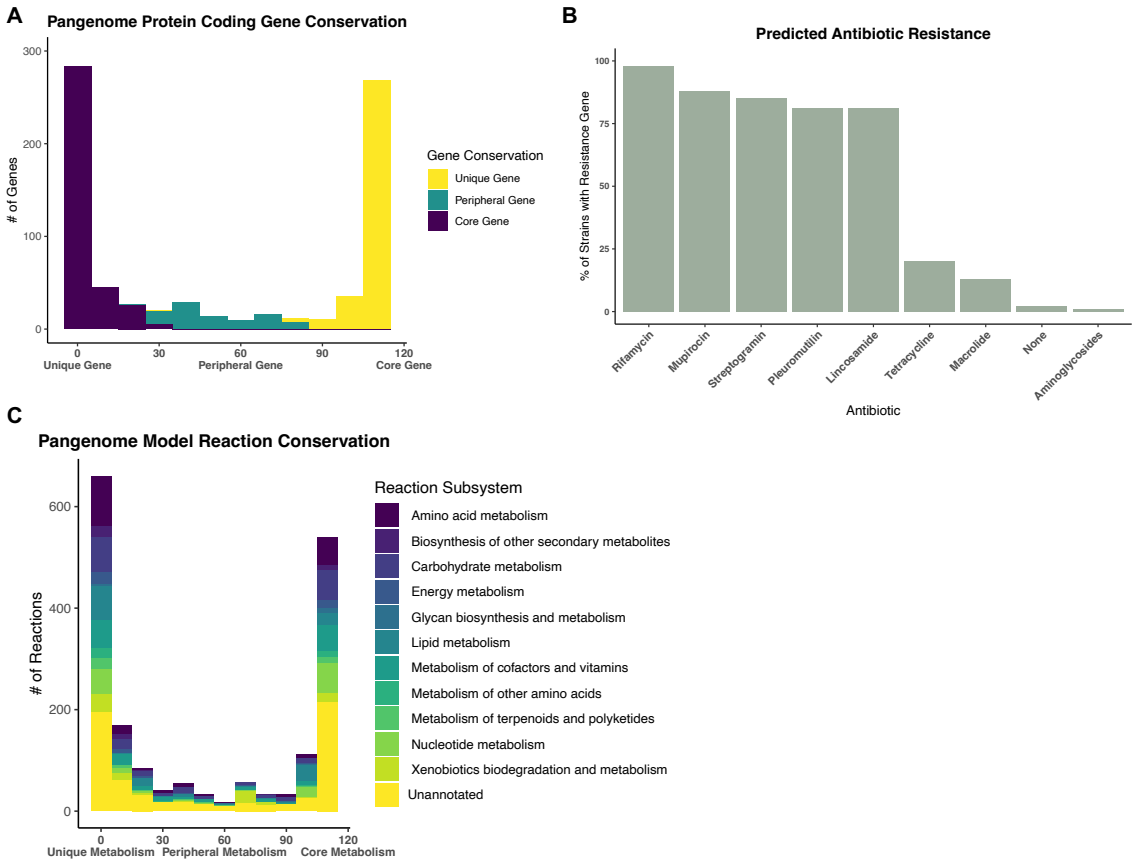


Figure 3: Analysis of pangenome **A)** genetic conservation based on protein coding gene presence **B**) predicted prevalence of antibiotic resistance genes based on Resistance Gene Identifier protein sequence annotation and **C)** pangenome metabolic conservation based on model reaction presence

Gene essentiality

After evaluating model gene essentiality in both BVCFM and SVM, 57 genes were found to be essential across the *Gardnerella* pangenome (**Figure 4**). Among these, four genes showed near-universal essentiality, namely, *gpsA* (K00057), *fas* (K11533), *suhB* (K01092), and *psd* (K01613). Based on KEGG annotations, *gpsA* is involved in glycerophospholipid metabolism, *fas* in fatty acid biosynthesis, *suhB* in inositol phosphate metabolism, and *psd* in glycerophospholipid metabolism. Our analysis using the DrugBank repository revealed two potential compounds, pyrazinamide and pretomanid, capable of targeting the essential fatty acid synthesis gene *fas* (Wishart et al., 2018). Notably, both drugs are currently

approved for treating tuberculosis but have not been studied for their efficacy against *Gardnerella* infections or BV.

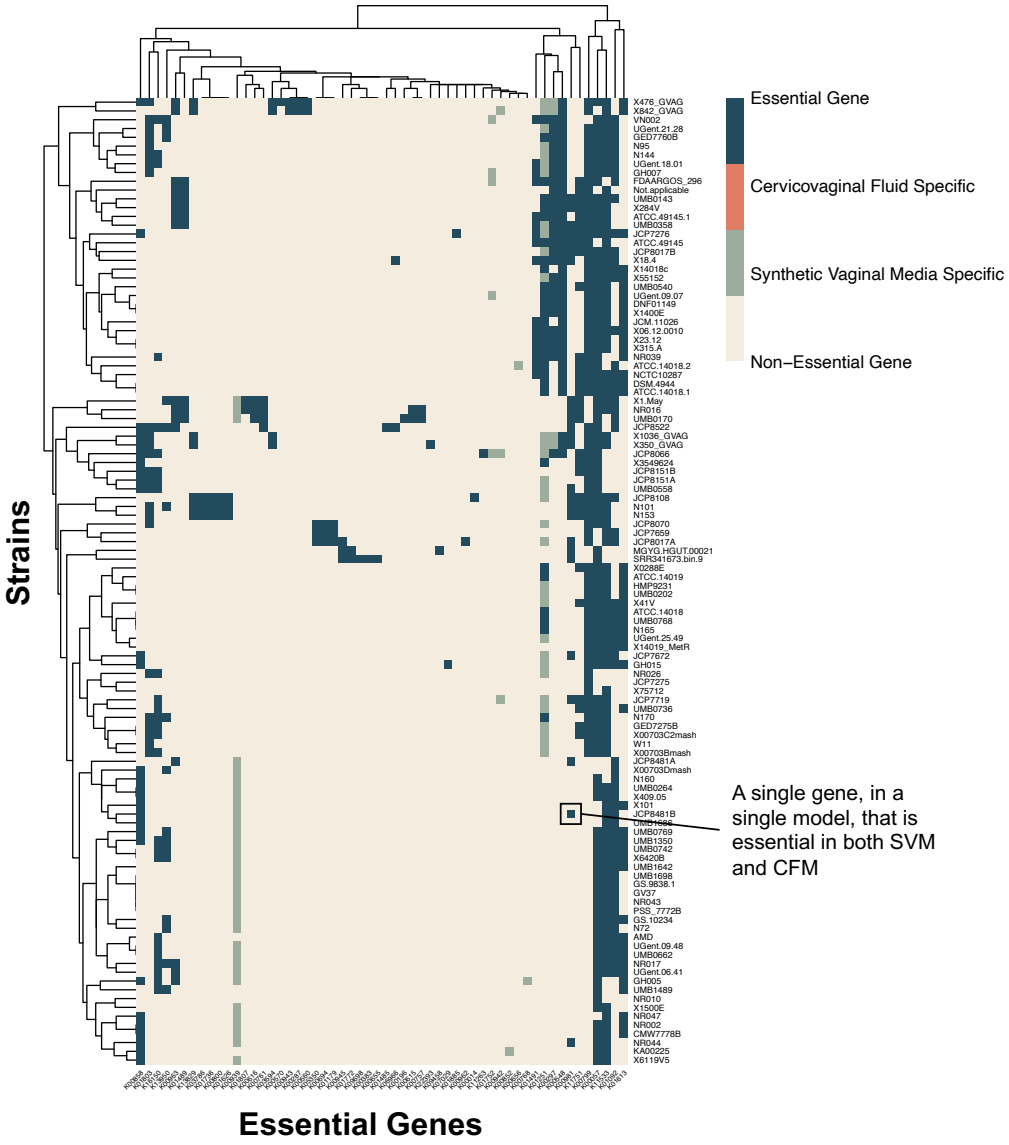
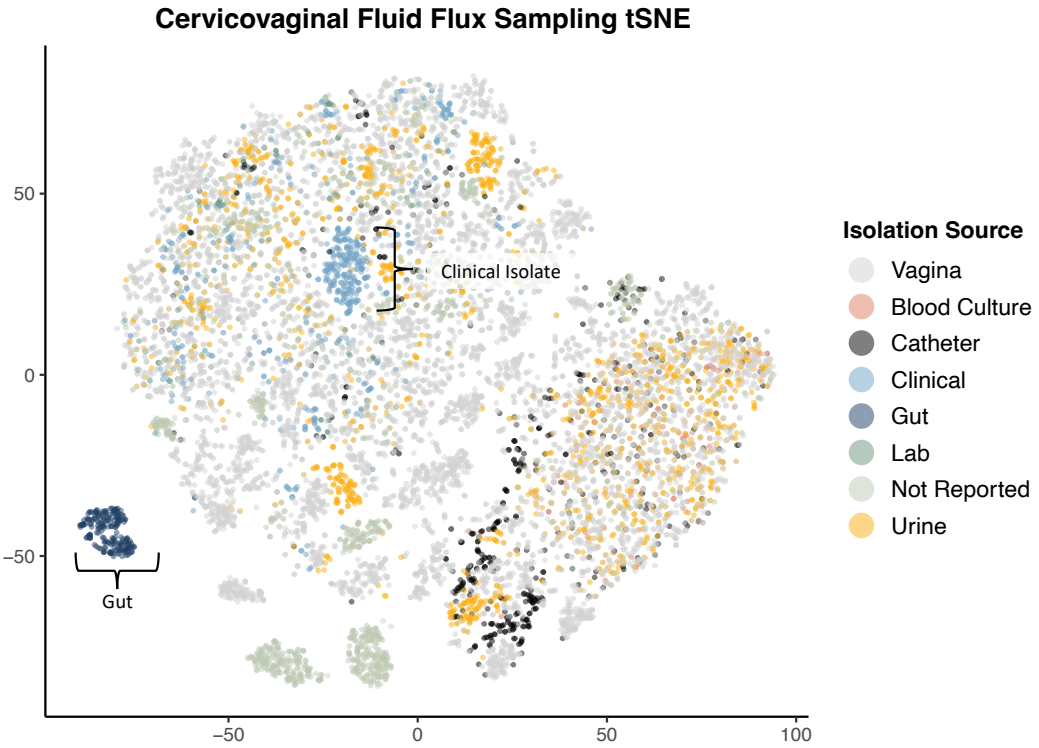


Figure 4: Gene essentiality screen across 110 models, contextualized to synthetic vaginal media (SVM) or BV positive cervicovaginal fluid (CFM)



Blood Culture	Catheter	Clinical	Gut	Lab	Urine	Vagina	Not Reported
1	5	6	2	1	15	65	15

Figure 5: Model flux comparison based on 100 flux samples, across all reactions via tSNE dimensionality reduction and associated number of strains per isolation source.

Model flux comparisons

When comparing the 110 metabolic models using flux sampling and t-SNE for dimensionality reduction, we observed limited clustering based on sample isolation source (**Figure 5**). However, clustering was observed for samples isolated from the gut and a subset of clinical isolates. Further analysis of transport flux values using t-SNE revealed pronounced clustering of gut isolates, blood culture isolates, and the laboratory 14019_MetR strain (**Figure 6A**). Heatmap comparison of transport reactions with the most varied flux values revealed that a subset of models exporting l-threonine, chloride, l-valine,

and aspartate glutamate were also importing galactose and sodium (**Figure 6B**). Additionally, a small subset of models was found to be significantly importing mannose-6-phosphate (MAN6P).

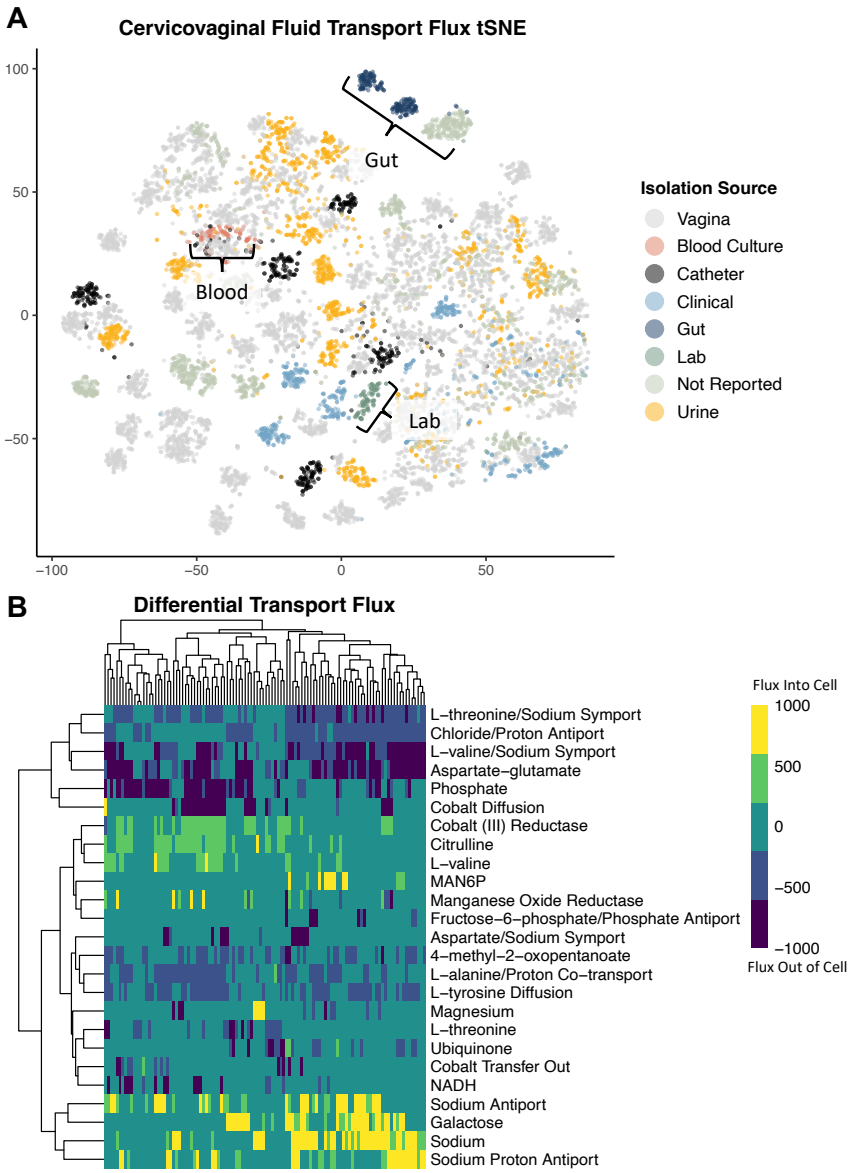


Figure 6: **A)** Model flux comparison based on 100 flux samples, across transport reactions via tSNE dimensionality reduction **B)** Heatmap of median flux values for each model, based on the top 25 most varied transport flux reaction values. Determined based on 100 flux samples collected via Gapsplit.

Discussion

The characterization of the *Gardnerella* pangenome is crucial in understanding the functional differences of this pathogen. Despite the prevalence of a significant number of *Gardnerella* strains, the majority remain uncharacterized, which limits the development of effective treatments for BV. We present the largest set of *Gardnerella* metabolic network reconstructions, which provides a functional characterization of the known *Gardnerella* pangenome. Our findings highlight conserved and unique metabolic mechanisms that could serve as valuable resources for the development of therapeutic strategies. This study sheds light on the metabolic functionality of *Gardnerella* and provides new insights into its potential drug targets, ultimately contributing to the reduction of BV recurrence rates.

Pangenome content comparison

We aimed to explore the relationship between genetic relatedness and functional relatedness in the *Gardnerella* pangenome. Our analysis revealed a significant dissimilarity (61%) between the dendrograms representing genetic content and metabolic functionality, suggesting that genetic content similarity does not necessarily correlate with metabolic functional similarity in *Gardnerella* strains. This result aligns with the concept that even small genetic differences can have a significant impact on metabolic functionality, as a single gene can be involved in multiple reactions. Therefore, to obtain a more accurate representation of phylogenetic relatedness within the *Gardnerella* pangenome, genetic expression profiles, which may more closely mirror metabolic functionality, should be explored (Eisen et al., 1998; Tavazoie et al., 1999; Zaslaver et al., 2004).

In terms of antibiotic resistance gene conservation, we identified the most highly conserved antibiotic resistance genes associated with drug classes such as rifamycin, mupirocin, streptogramin, lincosamide, and pleuromutilin. Interestingly, nitroimidazole class antibiotics, which include metronidazole, did not show significant conservation of associated antibiotic resistance genes within the *Gardnerella* pangenome. Clindamycin, an antibiotic of the lincosamide class, could be less effective in treating BV due to potential resistance. These findings could be important in informing treatment approaches for BV, which is known to have high recurrence rates ("Lincosamides," 2016).

We also observed a high proportion of unique reactions associated with amino acid metabolism, which aligns with previous literature indicating that differential amino acid metabolism can distinguish *Gardnerella* subgroups (Khan et al., 2021). The large number of unique reactions associated with xenobiotic biodegradation suggests that the *Gardnerella* genus is capable of differential interactions with pharmaceutical treatments as well as nonendogenous probiotics (C. Li et al., 2019; Z. Wang & He, n.d.). This finding highlights the need for understanding which *Gardnerella* strains are present in BV in order to develop patient-specific treatments. The large number of unique reactions associated with terpenoid and polyketide metabolism could offer insight into why some women experience persistent and odorous BV while others remain asymptomatic. Symptom variation could be due to the antimicrobial properties of polyketide metabolism, which allow the odor producing bacteria to outcompete their microbial competitors (Ridley et al., 2008). Some *Gardnerella* strains may be better equipped to outcompete other vaginal microbes. Finally, we found that glycan-related metabolism is uniquely enriched in the

pangenome core metabolism, indicating potential coevolution of *Gardnerella* with sialic acid-catabolizing microbes such as *Fusobacterium* (Agarwal et al., 2020). This finding could be useful in understanding the intra-pangenome coevolution of *Gardnerella* based on differences in sialidase activity across strains (W. G. Lewis et al., 2013). Overall, our study highlights the importance of considering functional relatedness in addition to genetic relatedness when investigating the phylogenetic relatedness of *Gardnerella* strains.

Gene essentiality

Conserved essential genes are potential targets for drug development, because disruption can inhibit virulence and adaptation. For example, the *gpsA* gene, which is involved in phospholipid synthesis, has been identified as a virulence factor in Lyme disease and nasopharyngeal colonization by *Streptococcus pneumoniae* (Drecktrah et al., 2022; Green et al., 2021). Similarly, *suhB*, an inositol monophosphatase, regulates virulence factors in *Pseudomonas aeruginosa* and is essential for *Burkholderia cenocepacia* biofilm formation, motility, and antibiotic resistance (K. Li et al., 2013; Rosales-Reyes et al., 2012). Both *gpsA* and *suhB* may be universally essential for driving *Gardnerella* virulence and adaptation to the vaginal mucosal environment. The *psd* gene, which plays a role in bacterial membrane biogenesis, has been successfully inhibited in *Plasmodium falciparum* using 4-quinolinamine compounds and may serve as a starting point for novel BV treatment development (Choi et al., 2016; Voelker, 1997). Fatty acid biosynthesis, which is essential for bacterial membrane construction, may be targeted by the *fas* gene, as fatty acid synthase type II (FASII) is bacterium specific.

Flux analysis

Investigating transport-specific flux values in *Gardnerella* strains allows for insights into how strains interact with their vaginal metabolic environment. The use of dimensionality reduction and visualization techniques, such as tSNE, revealed that laboratory strains differ in their metabolic functionality from strains collected from body sites, highlighting the need to characterize metabolic functionality of non-laboratory vaginal microbiome strains. Furthermore, the dispersed nature of vaginal isolates indicates wide variation in functional metabolism. Functional metabolic differences are further supported by differential galactose import, suggesting that there is strain-level variation in energy sources and metabolic pathways involved in BV (Benito et al., 1986). Of note, a small set of strains import high levels of mannose-6-phosphate, which is an essential ligand for the mannose-6-phosphate enzyme, which is key for lysosomal function (Das Purkayastha et al., 2019; Gary-Bobo et al., 2007; Oh et al., 2007). Some strains of *Gardnerella* may sequester mannose-6-phosphate as a mechanism of evading host lysosomal clearance, which could result in disordered vaginal epithelial cell function due to the lack of waste removal and increased inflammation and oxidative stress (Ferreira & Gahl, 2022; Vuolo et al., 2021).

Conclusion

In conclusion, our study provides important insights into the *Gardnerella* pangenome and its metabolic function. We discovered that genetic relatedness does not necessarily translate to functional relatedness among *Gardnerella* strains. Our findings suggest that BV research should focus on understanding the functional metabolic differences among strains to design effective interventions at a strain-specific level. The identified conserved gene essentiality predictions of *gpsA*, *fas*, *suhB*, and *psd* could inform the development of

novel drugs targeting this diverse genus. Additionally, *Gardnerella* strains interact differently with their vaginal metabolic environment, indicating the potential for metabolic niche development within the pangenome. These discoveries have implications for developing a deeper understanding of patient-level variation in BV and its impact on health outcomes and infection, which can ultimately lead to personalized therapeutic approaches. Overall, our work highlights the importance of integrating computational and experimental approaches to better understand the diversity and functionality of the *Gardnerella* pangenome.

Supplementary

S1A: BVCFM *in silico* formula

Metabolite	*in silico* Media Inclusion	Model Seed ID	FluxBounds
myristoleate	TRUE	cpd05237	[-1000,1000]
kynurenate	TRUE	cpd01182	[-1000,1000]
pentadecanoate	TRUE	cpd15622	[-1000,1000]
2-O-methylguanosine	TRUE	cpd34555	[-1000,1000]
3-dephosphocoenzymeA	TRUE	cpd00655	[-1000,1000]
N2,N2-dimethylguanine	TRUE	cpd28562	[-1000,1000]
flavinadeninedinucleotide(FAD)	TRUE	cpd00015	[-1000,1000]
alanine	TRUE	cpd00035	[-1000,1000]
valine	TRUE	cpd00156	[-1000,1000]
N-acetylalanine	TRUE	cpd33748	[-1000,1000]
3-methyl-2-oxovalerate	TRUE	cpd00508	[-1000,1000]
3-methyl-2-oxobutyrate	TRUE	cpd00123	[-1000,1000]
tricarballylate	TRUE	cpd16654	[-1000,1000]
4-methyl-2-oxopentanoate	TRUE	cpd00200	[-1000,1000]
N-acetylneuraminate	TRUE	cpd00232	[-1000,1000]
citrulline	TRUE	cpd00274	[-1000,1000]
2-aminobutyrate	TRUE	cpd01573	[-1000,1000]
phenylacetate	TRUE	cpd00430	[-1000,1000]
sarcosine(N-Methylglycine)	TRUE	cpd00183	[-1000,1000]

nicotinate	TRUE	cpd00218	[-1000,1000]
palmitoylethanolamide	TRUE	cpd16300	[-1000,1000]
threitol	TRUE	cpd17172	[-1000,1000]
galactose	TRUE	cpd00108	[-1000,1000]
N6-acetyllysine	TRUE	cpd01770	[-1000,1000]
1-phenylethanamine	TRUE	cpd34388	[-1000,1000]
alpha-hydroxyisocaproate	TRUE	cpd33351	[-1000,1000]
2-Hydroxybutyrate	TRUE	cpd03561	[-1000,1000]
succinate	TRUE	cpd00036	[-1000,1000]
2-hydroxyglutarate	TRUE	cpd02041	[-1000,1000]
3-phenylpropionate(hydrocinnamate)	TRUE	cpd03343	[-1000,1000]
thymine	TRUE	cpd00151	[-1000,1000]
agmatine	TRUE	cpd00152	[-1000,1000]
N-acetylputrescine	TRUE	cpd01758	[-1000,1000]
4-hydroxybutyrate	TRUE	cpd00728	[-1000,1000]
tyramine	TRUE	cpd00374	[-1000,1000]
putrescine	TRUE	cpd00118	[-1000,1000]
deoxycarnitine	TRUE	cpd00870	[-1000,1000]
tryptamine	TRUE	cpd00318	[-1000,1000]
cadaverine	TRUE	cpd01155	[-1000,1000]
5-aminovalerate	TRUE	cpd00339	[-1000,1000]
4-Hydroxyphenylacetate	TRUE	cpd00489	[-1000,1000]
pipecolate	TRUE	cpd00323	[-1000,1000]

pelargonate	FALSE		
indolepropionate	FALSE		
myristate	FALSE		
13-methylmyristicacid	FALSE		
N-acetylvaline	FALSE		
N-acetylphenylalanine	FALSE		
N-acetylleucine	FALSE		
N-acetylglutamate	FALSE		
13-HODE	FALSE		
oleicethanolamide	FALSE		
N-acetylaspartate(NAA)	FALSE		
12-HETE	FALSE		
alpha-hydroxyisovalerate	FALSE		
2-hydroxy-3-methylvalerate	FALSE		
3-(4-hydroxyphenyl)propionate	FALSE		

S1B: SVM *in silico* formula

Metabolite	***in silico* Media Inclusion**	**Model Seed ID**	**FluxBounds**
sodium	TRUE	cpd00971	[-1000,1000]
potassium	TRUE	cpd00205	[-1000,1000]
dextrose	TRUE	cpd00027	[-1000,1000]
cysteine	TRUE	cpd00084	[-1000,1000]
glycogen	TRUE	cpd00155	[-1000,1000]

mucin	TRUE	cpd00984	[-1000,1000]
urea	TRUE	cpd00073	[-1000,1000]
phytomenadione	TRUE	cpd01401	[-1000,1000]
heme	TRUE	cpd00028	[-1000,1000]
magnesium	TRUE	cpd00254	[-1000,1000]
sulfate	TRUE	cpd00048	[-1000,1000]
bicarbonate	TRUE	cpd00242	[-1000,1000]
biotin	TRUE	cpd00104	[-1000,1000]
myo-inositol	TRUE	cpd00121	[-1000,1000]
niacinamide	TRUE	cpd00133	[-1000,1000]
pyridoxine	TRUE	cpd00263	[-1000,1000]
thiamine	TRUE	cpd00305	[-1000,1000]
D-Calciumpantothenate	TRUE	cpd19112	[-1000,1000]
folate	TRUE	cpd00393	[-1000,1000]
choline	TRUE	cpd00098	[-1000,1000]
riboflavin	TRUE	cpd00220	[-1000,1000]
l-ascorbicacid	TRUE	cpd00059	[-1000,1000]
retinol	TRUE	cpd00365	[-1000,1000]
Cyanocobalamin	TRUE	cpd01826	[-1000,1000]
water	TRUE	cpd00001	[-1000,1000]
albumin	FALSE		
p-Aminobenzoic	FALSE		
acid	FALSE		

calciferol	FALSE		
tween20	FALSE		
K2HPO4buffer	FALSE		
KHSPO4buffer	FALSE		

References

Agarwal, K., Robinson, L. S., Aggarwal, S., Foster, L. R., Hernandez-Leyva, A., Lin, H., Tortelli, B. A., O'Brien, V. P., Miller, L., Kau, A. L., Reno, H., Gilbert, N. M., Lewis, W. G., & Lewis, A. L. (2020). Glycan cross-feeding supports mutualism between Fusobacterium and the vaginal microbiota. *PLoS Biology*, *18*(8). https://doi.org/10.1371/JOURNAL.PBIO.3000788

Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A. L. V., Cheng, A. A., Liu, S., Min, S. Y., Miroshnichenko, A., Tran, H. K., Werfalli, R. E., Nasir, J. A., Oloni, M., Speicher, D. J., Florescu, A., Singh, B., … McArthur, A. G. (2020). CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Research*, *48*(D1), D517–D525. https://doi.org/10.1093/NAR/GKZ935

Allsworth, J. E., Lewis, V. A., & Peipert, J. F. (2008). Viral sexually transmitted infections and bacterial vaginosis: 2001-2004 National Health and Nutrition Examination Survey data. *Sexually Transmitted Diseases*, *35*(9), 791–796. https://doi.org/10.1097/OLQ.0B013E3181788301

Aziz, R. K., Bartels, D., Best, A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman,

A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., … Zagnitko, O. (2008). The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics*, *9*. https://doi.org/10.1186/1471-2164-9-75

Benito, R., Vazquez, J. A., Berron, S., Fenoll, A., & Saez-Neito, J. A. (1986). A modified scheme for biotyping Gardnerella vaginalis. *Journal of Medical Microbiology*, *21*(4), 357–359. https://doi.org/10.1099/00222615-21-4-357/CITE/REFWORKS

Bradshaw, C. S., & Sobel, J. D. (2016). Current Treatment of Bacterial Vaginosis-Limitations and Need for Innovation. *Journal of Infectious Diseases*, *214*(Suppl 1), S14–S20. https://doi.org/10.1093/infdis/jiw159

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., Olson, R., Overbeek, R., Parrello, B., Pusch, G. D., Shukla, M., Thomason, J. A., Stevens, R., Vonstein, V., Wattam, A. R., & Xia, F. (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports*, *5*. https://doi.org/10.1038/SREP08365

Choi, J. Y., Kumar, V., Pachikara, N., Garg, A., Lawres, L., Toh, J. Y., Voelker, D. R., & Ben Mamoun, C. (2016). Characterization of Plasmodium phosphatidylserine decarboxylase expressed in yeast and application for inhibitor screening. *Molecular Microbiology*, *99*(6), 999. https://doi.org/10.1111/MMI.13280

Culhane, J. F., Rauh, V., McCollum, K. F., Elo, I. T., & Hogan, V. (2002). Exposure to chronic stress and ethnic differences in rates of bacterial vaginosis among pregnant women. *American Journal of Obstetrics and Gynecology*, *187*(5), 1272–1276. https://doi.org/10.1067/MOB.2002.127311

Das Purkayastha, S., Bhattacharya, M. K., Prasad, H. K., Upadhyaya, H., Lala, S. D., Pal, K., Das, M., Sharma, G. D., & Bhattacharjee, M. J. (2019). Contrasting diversity of vaginal lactobacilli among the females of Northeast India. *BMC Microbiology*, *19*(1). https://doi.org/10.1186/S12866-019-1568-6

Dice, L. R. (1945). Measures of the Amount of Ecologic Association Between Species. *Ecology*, *26*(3), 297–302. https://doi.org/10.2307/1932409

Drecktrah, D., Hall, L. S., Crouse, B., Schwarz, B., Richards, C., Bohrnsen, E., Wulf, M., Long, B., Bailey, J., Gherardini, F., Bosio, C. M., Lybecker, M. C., & Samuels, D. S. (2022). The glycerol-3-phosphate dehydrogenases GpsA and GlpD constitute the oxidoreductive metabolic linchpin for Lyme disease spirochete host infectivity and persistence in the tick. *PLoS Pathogens*, *18*(3). https://doi.org/10.1371/JOURNAL.PPAT.1010385

Ebrahim, A., Lerman, J. A., Palsson, B. O., & Hyduke, D. R. (2013). COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Systems Biology*, *7*. https://doi.org/10.1186/1752-0509-7-74

Eisen, M. B., Spellman, P. T., Brown, P. O., & Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(25), 14863–14868. https://doi.org/10.1073/PNAS.95.25.14863

Ferreira, C. R., & Gahl, W. A. (2022). Lysosomal Storage Disease. *Metabolic Diseases: Foundations of Clinical Management, Genetics, and Pathology*, 367–440. https://doi.org/10.3233/978-1-61499-718-4-367

Galili, T. (2015). dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics (Oxford, England)*, *31*(22), 3718–3720. https://doi.org/10.1093/BIOINFORMATICS/BTV428

Gary-Bobo, M., Nirdé, P., Jeanjean, A., Morère, A., & Garcia, M. (2007). Mannose 6-phosphate receptor targeting and its applications in human diseases. *Current Medicinal Chemistry*, *14*(28), 2945. https://doi.org/10.2174/092986707782794005

Geshnizgani, A. M., & Onderdonk, A. B. (1992). Defined medium simulating genital tract secretions for growth of vaginal microflora. *Journal of Clinical Microbiology*, *30*(5), 1323–1326. https://doi.org/10.1128/jcm.30.5.1323-1326.1992

Green, A. E., Howarth, D., Chaguza, C., Echlin, H., Langendonk, R. F., Munro, C., Barton, T. E., Hinton, J. C. D., Bentley, S. D., Rosch, J. W., & Neill, D. R. (2021). Pneumococcal Colonization and Virulence Factors Identified Via Experimental Evolution in Infection Models. *Molecular Biology and Evolution*, *38*(6), 2209. https://doi.org/10.1093/MOLBEV/MSAB018

Jenior, M. L., Glass, E. M., & Papin, J. A. (2022). *Reconstructor: A COBRApy compatible tool for automated genome-scale metabolic network reconstruction with parsimonious flux-based gap-filling*. http://github.com/emmamglass/reconstructor.

Keaty, T. C., Keaty, T. C., Jensen, P. A., Jensen, P. A., & Jensen, P. A. (2020). Gapsplit: Efficient random sampling for non-convex constraint-based models.

Bioinformatics (Oxford, England), *36*(8), 2623–2625. https://doi.org/10.1093/BIOINFORMATICS/BTZ971

Khan, S., Vancuren, S. J., & Hill, J. E. (2021). A Generalist Lifestyle Allows Rare Gardnerella spp. To Persist at Low Levels in the Vaginal Microbiome. *Microbial Ecology*, *82*(4), 1048–1060. https://doi.org/10.1007/S00248-020-01643-1

Kolde, R. (2019). Package "pheatmap": Pretty heatmaps. *Version 1.0.12*, 1–8.

Koumans, E. H., Sternberg, M., Bruce, C., McQuillan, G., Kendrick, J., Sutton, M., & Markowitz, L. E. (2007). The prevalence of bacterial vaginosis in the United States, 2001-2004; associations with symptoms, sexual behaviors, and reproductive health. *Sexually Transmitted Diseases*, *34*(11), 864–869. https://doi.org/10.1097/OLQ.0B013E318074E565

Lewis, W. G., Robinson, L. S., Gilbert, N. M., Perry, J. C., & Lewis, A. L. (2013). Degradation, Foraging, and Depletion of Mucus Sialoglycans by the Vagina-adapted Actinobacterium Gardnerella vaginalis. *The Journal of Biological Chemistry*, *288*(17), 12067. https://doi.org/10.1074/JBC.M113.453654

Li, C., Wang, T., Li, Y., Zhang, T., Wang, Q., He, J., Wang, L., & Li, L. (2019). Probiotics for the treatment of women with bacterial vaginosis: A systematic review and meta-analysis of randomized clinical trials. *European Journal of Pharmacology*, *864*(April), 172660. https://doi.org/10.1016/j.ejphar.2019.172660

Li, K., Xu, C., Jin, Y., Sun, Z., Liu, C., Shi, J., Chen, G., Chen, R., Jin, S., & Wu, W. (2013). SuhB is a regulator of multiple virulence genes and essential for pathogenesis of Pseudomonas aeruginosa. *MBio*, *4*(6). https://doi.org/10.1128/MBIO.00419-13

Lieven, C., Beber, M. E., Olivier, B. G., Bergmann, F. T., Ataman, M., Babaei, P., Bartell, J. A., Blank, L. M., Chauhan, S., Correia, K., Diener, C., Dräger, A., Ebert, B. E., Edirisinghe, J. N., Faria, J. P., Feist, A. M., Fengos, G., Fleming, R. M. T., García-Jiménez, B., … Zhang, C. (2020). MEMOTE for standardized genome-scale metabolic model testing. *Nature Biotechnology*, *38*(3), 272–276. https://doi.org/10.1038/s41587-020-0446-y

Lincosamides. (2016). *Meyler's Side Effects of Drugs*, 581–588. https://doi.org/10.1016/B978-0-444-53717-1.00980-X

Maaten, L. V. der, research, G. H.-J. of machine learning, & 2008, undefined. (2008). Visualizing data using t-SNE. *Jmlr.Org*, *9*, 2579–2605.

Madden, T. (2002, October 9). *The BLAST Sequence Analysis Tool*. The NCBI Handbook. https://www.ncbi.nlm.nih.gov/books/NBK21097/

Muzny, C. A., & Kardas, P. (2020). A Narrative Review of Current Challenges in the Diagnosis and Management of Bacterial Vaginosis. *Sexually Transmitted Diseases*, *47*(7), 441–446. https://doi.org/10.1097/OLQ.0000000000001178

Oh, Y. K., Palsson, B. O., Park, S. M., Schilling, C. H., & Mahadevan, R. (2007). Genome-scale reconstruction of metabolic network in Bacillus subtilis based on high-throughput phenotyping and gene essentiality data. *The Journal of Biological Chemistry*, *282*(39), 28791–28799. https://doi.org/10.1074/JBC.M703759200

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A. R., Xia, F., & Stevens, R. (2014). The SEED and the Rapid Annotation of microbial genomes

using Subsystems Technology (RAST). *Nucleic Acids Research*, *42*(Database issue). https://doi.org/10.1093/NAR/GKT1226

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2012). *Scikit-learn: Machine Learning in Python*. http://arxiv.org/abs/1201.0490

Peebles, K., Velloza, J., Balkus, J. E., McClelland, R. S., & Barnabas, R. V. (2019). High Global Burden and Costs of Bacterial Vaginosis: A Systematic Review and Meta-Analysis. *Sexually Transmitted Diseases*, *46*(5), 304–311. https://doi.org/10.1097/OLQ.0000000000000972

Qin, H., & Xiao, B. (2022). Research Progress on the Correlation Between Gardnerella Typing and Bacterial Vaginosis. *Frontiers in Cellular and Infection Microbiology*, *12*. https://doi.org/10.3389/FCIMB.2022.858155

Ridley, C. P., Ho, Y. L., & Khosla, C. (2008). Evolution of polyketide synthases in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(12), 4595–4600. https://doi.org/10.1073/PNAS.0710107105/SUPPL_FILE/10107FIG9.PDF

Rosales-Reyes, R., Saldías, M. S., Aubert, D. F., El-Halfawy, O. M., & Valvano, M. A. (2012). The suhB gene of Burkholderia cenocepacia is required for protein secretion, biofilm formation, motility and polymyxin B resistance. *Microbiology (Reading, England)*, *158*(Pt 9), 2315–2324. https://doi.org/10.1099/MIC.0.060988-0

Schwebke, J. R., Muzny, C. A., & Josey, W. E. (2014). Role of Gardnerella vaginalis in the pathogenesis of bacterial vaginosis: A conceptual model. *Journal of Infectious Diseases*, *210*(3), 338–343. https://doi.org/10.1093/infdis/jiu089

Schwiertz, A., Taras, D., Rusch, K., & Rusch, V. (2006). Throwing the dice for the diagnosis of vaginal complaints? *Annals of Clinical Microbiology and Antimicrobials*, *5*. https://doi.org/10.1186/1476-0711-5-4

Tavazoie, S., Hughes, J. D., Campbell, M. J., Cho, R. J., & Church, G. M. (1999). Systematic determination of genetic network architecture. *Nature Genetics 1999 22:3*, *22*(3), 281–285. https://doi.org/10.1038/10343

Vitali, B., Cruciani, & F., Picone, G., Parolin, & C., Donders, & G., & Laghi, & L. (2015). *Vaginal microbiome and metabolome highlight specific signatures of bacterial vaginosis*. https://doi.org/10.1007/s10096-015-2490-y

Voelker, D. R. (1997). Phosphatidylserine decarboxylase. *Biochimica et Biophysica Acta*, *1348*(1–2), 236–244. https://doi.org/10.1016/S0005-2760(97)00101-X

Vuolo, D., Do Nascimento, C. C., & D'Almeida, V. (2021). Reproduction in Animal Models of Lysosomal Storage Diseases: A Scoping Review. *Frontiers in Molecular Biosciences*, *8*, 1113. https://doi.org/10.3389/FMOLB.2021.773384/BIBTEX

Wang, J. (2000). Bacterial vaginosis. *Primary Care Update for OB/GYNS*, *7*(5), 181–185. https://doi.org/10.1016/S1068-607X(00)00043-3

Wang, Z., & He, Y. (n.d.). *Probiotics for the Treatment of Bacterial Vaginosis: A Meta-Analysis*. 1–13.

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, *58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Wattam, A. R., Davis, J. J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., Conrad, N., Dietrich, E. M., Disz, T., Gabbard, J. L., Gerdes, S., Henry, C. S., Kenyon, R. W., Machi, D., Mao, C., Nordberg, E. K., Olsen, G. J., Murphy-Olson, D. E., Olson, R., … Stevens, R. L. (2017). Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Research*, *45*(D1), D535–D542. https://doi.org/10.1093/NAR/GKW1017

Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., MacIejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, Di., … Wilson, M. (2018). DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Research*, *46*(Database issue), D1074. https://doi.org/10.1093/NAR/GKX1037

Zaslaver, A., Mayo, A. E., Rosenberg, R., Bashkin, P., Sberro, H., Tsalyuk, M., Surette, M. G., & Alon, U. (2004). Just-in-time transcription program in metabolic pathways. *Nature Genetics 2004 36:5*, *36*(5), 486–491. https://doi.org/10.1038/ng1348

Chapter 3: Competition and mutualism in the dysbiotic vaginal microbiome

The text for this chapter is currently under-review:

Dillard LR, Kolling GL, Thomas-White K, Wever F, Glass EM, Papin JA (*under review)* Competition and Mutualism in the Dysbiotic Vaginal Microbiome. *Nature Microbiology*

Context

My pangenome project created a base of understanding surrounding the metabolic diversity within the *Gardnerella* genus. I wanted to build upon this foundation and delve into the inter-bacterial complexity of bacterial vaginosis (BV). Through my partnership with Evvy, a vaginal microbiome startup company, I was able to identify common non-*Gardnerella* bacterial species that are associated with BV. Using these data, along with former Papin lab post-doc Dr. Matt Jenior's algorithm to simulate inter-bacterial competition and mutualism, the second chapter of my dissertation was born: defining inter-bacterial metabolic interactions within BV.

Synopsis

Bacterial vaginosis (BV) is the most prevalent vaginal condition among reproductive-age women experiencing vaginal complaints. Despite its significant impact on women's health, limited knowledge exists regarding the microbial community structure and metabolic interactions associated with BV. In this study, we analyzed metagenomic data obtained from human vaginal swabs to generate *in silico* predictions of BV-associated bacterial metabolic interactions via genome-scale metabolic network reconstructions (GENREs).

Our *in silico* simulations revealed distinct functional metabolic relatedness compared to genetic relatedness within the *Gardnerella* genus. We grew the most common co-occurring bacteria on the spent media of *Gardnerella* species and performed metabolomics to identify potential mechanisms of metabolic interaction. Notably, we identified some BV associated bacteria significantly produce caffeate, a compound implicated in estrogen receptor binding. These findings underscore the complex and diverse nature of BV-associated bacterial community structures.

Introduction

Bacterial vaginosis (BV) is the most common state of vaginal dysbiosis among reproductive age women with vaginal complaints (Schwiertz et al., 2006). BV is a polymicrobial condition characterized by low levels of *Lactobacillus*, high levels of diverse anaerobes, a vaginal pH greater than 4.5, thin vaginal discharge, and a fishy odor (Wang, 2000). Various factors, such as sexual activity, menstruation, antibiotics, and douching can alter a healthy, acidic, *Lactobacillus*-dominant, vaginal microbiome, leading to the development of BV (Gajer et al., 2012; Lopes dos Santos Santiago et al., 2012; Mayer et al., 2015; Tachedjian et al., 2017; Tortelli et al., 2020). BV is disproportionally prevalent among women of color, affecting 33-64% of Black women, and 31-32% of Hispanic women, compared to 23-35% of White women (Allsworth et al., 2008; Culhane et al., 2002; Koumans et al., 2007, p. 20; Muzny & Kardas, 2020). BV increases risk of contracting sexually transmitted diseases, including HIV, and is also associated with risk of preterm birth (Feehily et al., 2020; Fettweis et al., 2019; Shimaoka et al., 2019). BV accounts for an estimated cost of $14.4 billion USD annually when considering both BV

treatment costs and BV-associated healthcare costs in the United States alone (Peebles et al., 2019).

Despite the significant impact of BV on women's health, there is a dearth of information regarding the microbial community structure and metabolic interactions associated with this condition (Bradshaw & Sobel, 2016). By investigating the competitive and mutualistic relationships among BV-associated bacteria, we can identify potential targets for therapeutic interventions. The inclusive list of bacteria involved in BV has yet to be defined; however, certain bacterial species are commonly associated with this condition, including: *Gardnerella* species, *Prevotella bivia, Prevotella amnii, Prevotella buccalis, Hoylesella timonensis, Lactobacillus iners, Fannyhessea vaginae,* and *Aerrococcus christenssii* (**Figure 1**) (Forsum et al., 2005; Marrazzo et al., 2009, 2009; Srinivasan et al., 2012). Although the *Gardnerella* genus is the primary contributor to BV, the large number of distinct *Gardnerella* species associated with BV make experimental analysis challenging (Castro et al., 2019; Dillard et al., 2022; Horrocks et al., n.d.; Janulaitiene et al., 2017). Computational modeling allows us to simulate thousands of pairwise interactions between bacterial species to analyze possible competitive and mutualistic behaviors.

In this study, we conducted an analysis of metagenomic data obtained from human vaginal swabs in order to ascertain the microbial composition of both symptomatic and asymptomatic cases of BV. Through the application of these data, we generated *in silico* predictions of BV-associated bacterial metabolic interactions using genome-scale

metabolic network reconstructions (GENREs). To validate these predictions, we carried out *in vitro* growth experiments. Furthermore, we collected bacterial supernatants to identify metabolites that putatively underlie BV-associated bacterial interactions. We subsequently conducted follow-up experiments to confirm the role of metabolites that were implicated both *in silico* and *in vitro*, in contributing to competitive interactions. By elucidating the underlying mutualistic and competitive interactions between bacterial pairs, we aim to enhance our understanding of the intricate microbial community associated with BV, ultimately leading to the development of more effective treatments for this prevalent condition.

Methods

Gardnerella and co-occurring species comparison

Co-occurrence frequency

Samples were split into three different groups: symptomatic BV (N=212), asymptomatic BV (N=504) and a healthy cohort (N=154) (**Table 1**). The symptomatic and asymptomatic BV groups both consisted of samples of Community State Type (CST) 4 and were dominated by *Gardnerella* (>= 50% relative abundance) (De Seta et al., 2019). However, the symptomatic BV group experienced excessive discharge and itchiness (either internal or external), while these symptoms were absent in the asymptomatic BV group. The healthy cohort consisted of samples of CST 1, 2 or 5, dominated by lactobacilli, with a relative abundance of >= 50% *L. crispatus, L. gasseri, L. paragasseri, L. pensenii, L. mulieris* and did not experience excessive discharge nor itchiness. Relative frequency of

species that co-occur with *Gardnerella vaginalis* within each subgroup was calculated using the following formula:

$$\frac{\text{\# of samples with >= 2\% relative abundance of co-occurring species}}{\text{\# of samples with >= 2 \% relative abundance of } \textit{G. vaginalis}}$$

Average nucleotide identity for speciation determination

The nucleotide similarity between the *Gardnerella* genomes was determined using an average nucleotide identity (ANI) comparison (Goris et al., 2007). 222 *Gardnerella* genomes were input to pyani, a Python implementation tool for ANI analysis (Pritchard et al., 2015). After running an initial comparison with pyani using the ANIb method, 13 genomes were removed from the analysis as they were too divergent from the majority of genomes. The remaining 209 *Gardnerella* genomes were input to pyani and the ANIb method which aligns 1020 nt fragments of the input sequences using BLASTN+ was used (Madden, 2002; Madden & Camacho, 2021). The generated percentage identity heatmap was color coded based on the taxonomic classifications given by the NCBI records of the genomes (Schoch et al., 2020).

Dendrogram construction

The pyani analysis generated a distance matrix for percentage identity, which was used to generate a dendrogram. The distance matrix was clustered via hierarchical clustering using the "complete" method via the TAPE R package (Zhuang et al., 2022). The clustered output was then converted to a tree object in Newick tree format. The dendrogram tree file was visualized using iTOL v6.7.1 and color coded based on the taxonomic classifications of the genomes' NCBI records (Letunic & Bork, 2021).

Model construction and contextualization

We analyzed potential metabolic interactions among *Gardnerella* strains by identifying 208 *Gardnerella* whole genome sequences from the BV-BRC 3.28.21 database that meet the "good" quality criteria defined as a genome with at least 80% completeness, less than 10% contamination, and at least 87% consistency with known protein sequences (Olson et al., 2023). Thirteen genomes were removed due to ANI clustering indicating either incomplete genome sequence or chimera status. We used the remaining 195 *Gardnerella* strains' annotated amino acid sequences to generate GENREs using the Reconstructor algorithm (Jenior et al., 2023). Additionally, we created GENREs for seven co-occurring species of interest: *Lactobacillus iners, Prevotella bivia, Prevotella amnii, Prevotella buccalis, Hoylesella timonensis, Fannyhessea vaginae,* and *Aerococcus christensii*. The quality of the GENREs was assessed using MEMOTE, which is the standard metric in computational metabolic modeling (Lieven et al., 2020). The GENREs and corresponding MEMOTE scores are available at the project's GitHub repository. (https://github.com/emmamglass/Gardnerella-Interactions).

After constructing the GENREs, we created a list of minimal metabolites required for 20% of maximal biomass production (Min20) for each individual GENRE using the COBRApy toolbox minimal_medium built-in function. We combined those minimal metabolite requirement lists to form a consensus list of minimal metabolites across all GENREs (Min20). To simulate vaginal dysbiosis, we also composed an *in silico* BVpositive cervicovaginal fluid media (BV-CFM) based on previous metabolomics analysis of cervicovaginal fluid collected from both healthy and BV-positive patients (Vitali et al.,

2015). This BV-CFM *in silico* media was enriched for metabolites found to be significantly elevated in BV-positive cervicovaginal fluid. We added transport reactions, which were absent from initial model construction but required for *in silico* media metabolite usage, in addition to respective exchange reactions. The final *in silico* media used for model simulation consisted of Min20 + BV-CFM.

Competition and mutualism simulation

To analyze the metabolic interactions between *Gardnerella* strains and other co-occurring species, we used a metabolic interaction prediction algorithm developed previously (Jenior et al., Under Review) and applied it to the 202 GENREs that we constructed. This method allowed us to simulate approximately 20,000 pairwise interactions and calculate the impact on biomass production and extracellular metabolites. We used a heatmap generated in R's pheatmap package to visualize the increase or decrease in biomass flux due to mutualism or competition, respectively (Kolde, 2019). Columns and rows were hierarchically clustered using the built-in Ward's minimum variance method, and associated distance matrix using the built-in Euclidean distance method (Lele & Richtsmeier, 1995; Ward, 1963). We also calculated net-flux values and visualized them using the same methodology.

To better understand the clustering of strains based on mutualistic benefit, competitive cost, and net outcome, we used t-SNE for dimensionality reduction and visualization. We generated three plots via the R Rtsne package using mutualism biomass flux increase values, competition biomass flux decrease values, and net-flux values, respectively (iteration = 50,000; perplexity = 14) (Van der Maaten & Hinton, 2008). ANI-calculated

clading values were used to color *Gardnerella* data points and highlight strain level inter-clade clustering.

To understand the relationship between mutualistic benefit and competition cost, we calculated the average decrease in biomass flux due to competition and the average increase in biomass flux due to mutualism for each bacterial strain. We then graphed all 202 bacteria of interest as depicted in **Figure 2**. We identified the top five most competitive and most mutualistic bacteria using a one-tailed t-test and corresponding Bonferroni multiple test correction using the p.adjust R function (Bonferroni, 1936; Student, 1908). Finally, we calculated the percent of total interactions a metabolite was either competed for and/or shared between GENREs.

Spent Media Analysis

Metabolomics Sample Collection

The following describes the experimental setup and sample preparation for metabolomics analysis of *Gardnerella vaginalis* (strain ATCC 14018) and *Gardnerella piotti* (strain JCP8151B) primary spent media. The bacteria were grown overnight under anaerobic conditions in NYCIII + 2% FBS media, and inoculated into T25 polystyrene tissue culture flasks. After 18 hours of incubation, samples were collected from each flask, spun down (15 minutes; 7,500 RPM), and filter-sterilized. The filtered supernatant was stored at -80°C until metabolomics processing. OD_{600} was recorded prior to the spin down step. Pooled spent media samples for both *G. vaginalis* and *G. piotti* were also aliquoted into microcentrifuge tubes and stored at -80°C. The remaining spent media for *G. vaginalis* and *G. piotti* was stored at -20°C.

In order to collect metabolomics on spent media conditions, six of the co-occurring species - *L. iners, P. amnii, P. buccalis, H. timonensis, F. vaginae* and *A. christensii* - were grown to turbidity overnight in enriched media (NYCIII+2% FBS, BHI, or MRS) under anaerobic conditions. Twelve-well plates with 2.5 mL of either *G. piotti* or *G. vaginalis* spent media were inoculated with 100 uL of co-occurring species inoculum. After inoculation, samples were grown to turbidity, as confirmed via OD_{600}. Samples were then collected from each flask, spun down (5 minutes; 5,000 RPM), filter-sterilized and stored at -80°C.

Mass Spectrometry Sample Preparation and Analysis

Untargeted metabolomics was performed by the University of Virginia Biomolecular Analysis Facility Core. Sample preparation involved adding 2000 µL of 80% methanol (-20 °C) to 500 µL of each culture medium. The mixture was vortexed for one minute and incubated at -20 °C for two hours to induce protein precipitation. Subsequently, the tubes were centrifuged at 14,000 g for 30 minutes at 4 °C. The resulting supernatants were transferred to new tubes, and a quality control (QC) sample was prepared by combining 35 µL from each sample. All samples were then dried in a speed vacuum for approximately four hours and stored at -80 °C until further processing.

For mass spectrometry analysis, the dried samples were reconstituted in 100 µL of 0.1% formic acid along with 100X diluted Metabolomics QReSS heavy-labeled standards (Percy et al., 2022; Percy, Andrew et al., 2021). A dilution of 1:800 was applied to both samples and QC sample. Subsequently, 10 µL of each diluted sample was injected for analysis.

Mass spectrometry data acquisition was performed using a Thermo Orbitrap IDX MS connected to a Vanquish UPLC system. A Waters BEH C18 column (100 x 2.1 mm, 1.9 μm) was employed for soluble metabolite separation. The column temperature was set at 30 °C, and a flow rate of 250 μL/min was maintained. The mobile phase consisted of 0.1% formic acid in water (mobile phase A) and 0.1% formic acid in methanol (mobile phase B). The mass scan range was set from 67 to 1000 with a resolution of 120,000 and a scan range of 0.6 sec.

Data-dependent MS2 scans were obtained using a real-time precursor exclusion strategy with AcquireX mode. Prior to the analysis of the samples, initial runs were performed with solvent blank samples in full scan mode (67-1000 mass range) to generate an exclusion list containing all the peaks present in the blank samples. A QC sample was then run in full scan mode, and all the peaks were saved as an inclusion list. Four data-dependent acquisition (DDA) injections were performed, with the exclusion list automatically updated based on the selected precursor.

After data acquisition, the samples were subjected to analysis using the open-source software MS-DIAL (Tsugawa et al., 2015, 2019). To identify and remove background ions, five blank samples were included in the analysis. The samples acquired in full scan MS1 mode were used for quantification, while the samples acquired in DDA mode were utilized for spectral identification of metabolites. The MS1 tolerance was set to 0.01 Da, and the MS2 tolerance was set to 0.025 Da. For peak picking, a mass slice width of 0.1 was

employed. Peak alignment was performed with a maximum retention time tolerance of 0.2 min and an MS1 tolerance of 0.015.

Peak identification was carried out by searching the MS2 spectra against the MS-DIAL public database (January 2023). A mass tolerance of 0.01 Da for MS1 and 0.05 Da for MSMS was applied, with an identification score cutoff of 60%. Additionally, the peaks were searched against an in-house IORA library in both positive and negative mode, using a mass tolerance of 0.01 Da and an identification cutoff of 80%. The data was manually inspected, and identifications without corresponding MS2 spectra were filtered out, except for IORA. Data analysis was performed using Metaboanalyst 5.0 (Pang et al., 2021). The samples were normalized by median and subjected to log transformation. Subsequently, a fold change analysis was performed to detect dysregulated metabolites between secondary spent media compared to primary spent media. Statistical analysis was conducted using a t-test to determine the significance of group differences. Over representation analysis (ORA) was conducted via Metaboanalyst, which calculates enrichment ratio derived from comparing between actual versus predicted number of metabolite hits per super-class metabolite category (Pang et al., 2021). The statistical significance of the enrichment ratios is calculated using hypothesis testing using a binomial distribution to assign a p-value, which is adjusted using Holm's multiple comparison test (Holm, 1979). Data was visualized using volcano plots generated in R's ggplot2 package, with a p-value cutoff of <0.01 and a $\log_2$(fold change) cut off of $> |2|$.

Growth Analysis

Overnight cultures of *L. iners, P. amnii, P. buccalis, H. timonensis, F. vaginae, A. christensii, G. vaginalis,* and *G. piotti* were cultured under anaerobic conditions for 24-48h in NYCIII media + 2% FBS, OD600 measured, pelleted by centrifugation (5,000 rpm, 5 min) and resuspended in fresh NYCIII. Cultures were then diluted in spent media and aliquoted into a 96-well plate (starting OD_{600} ~0.05) containing spent or blank media for a kinetic growth assay under anaerobic conditions without shaking. *G. piotti* and *A. christensii's* flocculant phenotype necessitated the use of shaking.

Supplemented Growth Analysis

Cultures of *G. piotti* and *H. timonensis* were grown under anaerobic conditions for 24-48h in NYCIII media + 2% FBS, OD_{600} measured, pelleted by centrifugation (5,000 rpm, 5 min) and diluted to an OD_{600} of 1 in fresh NYCIII. Cultures were then aliquoted into a 96-well plate containing media conditions specified below (starting OD_{600} 0.1) for a kinetic growth assay under anaerobic conditions without shaking. *G. piotti* was cultured in spent *G. piotti* media and spent *G. piotti* media supplemented with 0.1% filter sterilized L-histidine. *H. timonensis* was cultured in spent *G. vaginalis* media and spent *G. vaginalis* media supplemented with 0.01% filter sterilized propionic acid.

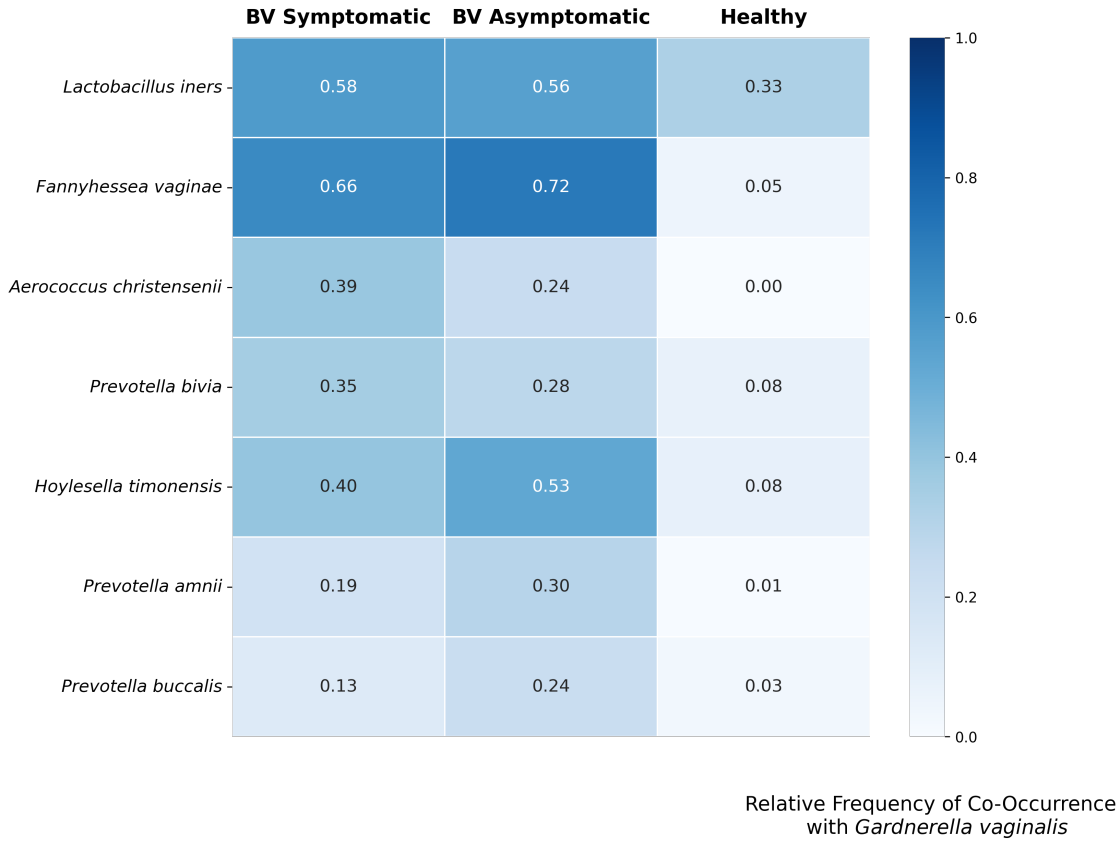


Figure 1 Heatmap illustrating relative frequency of *Gardnerella vaginalis* co-occurrence with non-*Gardnerella* species in BV symptomatic, BV asymptomatic, and Healthy samples

Results

In vivo BV Co-Occurring Species

L. iners, G. vaginalis, F. vaginae, and *H. timonensis* are frequently found to co-occur in both symptomatic and asymptomatic samples of BV (**Figure 1**). *L. iners* was observed to co-occur with *G. vaginalis* in 33% of healthy samples, whereas the co-occurrence of *F. vaginae* and *H. timonensis* was only 5% and 8%, respectively. *A. christensenii* and *P. bivia* were identified as more common co-occurring species in symptomatic samples. On the other hand, *P. amnii, P. buccalis*, and *H. timonensis* were more commonly observed as co-

occurring species with *G. vaginalis* in asymptomatic samples. Demographic features collected from sample metadata did not appear to correlate with differences in symptomatic, asymptomatic and healthy samples (**Table 1**).

Demographic and Clinical Characteristics			
	Symptomatic BV	**Asymptomatic BV**	**Healthy**
Number of Tests	212	504	154
Number of Users (n =) **note: this is used as the denominator for most of the following stats*	208	468	147
Age (years) mean (standard deviation)	34.6 (7.8)	37.8 (9.4)	35.6 (9.6)
Body Mass Index mean (standard deviation)	25.1 (6.4)	24.9 (5.9)	24 (5.4)
Race/Ethnicity number (%) * multiselect			
White	139 (66.8%)	319 (68%)	114 (77.6%)
Black or African American	33 (15.9%)	71 (15%)	13 (8.8%)
Hispanic or Latino	39 (18.8%)	57 (12%)	15 (10.2%)
Asian	11 (5.3%)	32 (6.8%)	6 (4.1%)
American Indian or Alaska Native	6 (2.9%)	11 (2.4%)	7 (4.8%)
Southeast Asian	4 (1.9%)	4 (0.8%)	1 (0.7%)
Middle Eastern	2 (0.9%)	7 (1.5%)	3 (2%)
South Asian	0 (0%)	4 (0.8%)	1 (0.7%)
Other	1 (0.5%)	6 (1.3%)	3 (2%)
Prefer not to say	3 (1.4%)	9 (1.9%)	2 (1.4%)
Which of these describes your sexual activity? number (%) *multiselect			
One partner	135 (65%)	327 (70%)	99 (67%)
Penetrative vaginal sex	128 (62%)	269 (57%)	80 (54%)
Male partner(s) *assigned at birth	102 (49%)	209 (45%)	68 (46%)
Receiving oral sex	96 (46%)	209 (45%)	61 (41%)
Sex toys	76 (37%)	179 (38%)	53 (36%)
Multiple partners	18 (9%)	47 (10%)	15 (10%)
Receiving anal sex	15 (7%)	42 (9%)	11 (7.5%)
Female partner(s) *assigned at birth	12 (6%)	20 (4.3%)	3 (2%)
Prefer not to say	0 (0%)	6 (1.3%)	0 (0%)
When were you last sexually active? number (%)			
Within the past 5 days	35 (17%)	130 (28%)	44 (30%)
Within the past 2 weeks	23 (11%)	74 (16%)	14 (9.5%)
Within the past 30 days	88 (42%)	165 (35%)	51 (35%)
Previously sexually active (over a month ago)	58 (28%)	83 (18%)	38 (26%)
Never sexually active	2 (1%)	7 (1.5%)	0 (0%)
Prefer not to say	2 (1%)	9 (1.9%)	0 (0%)

Table 1 Demographics and clinical characteristics of vaginal metagenomic samples

ANI of *Gardnerella* genomes identified 13 genetically distinct species defined as 95% genetic similarity (**S1**). Four species are already named (*G. vaginalis, G. piotii, G. swidsinkii*, and *G. leopoldii*) with the remainder unnamed and therefore given a number (species 1-9). Within these species there were

also distinct genetic clades for both *G. vaginalis* and *G. piotii* designated as clades A and B in each (**S1**).

Pairwise *in silico* bacterial interactions

We analyzed mutualism, competition, and net interactions at the single bacterial species, single interaction level. Our investigation revealed a consistent lack of mutualism benefit across two *Prevotella* species and *H. timonensis*, which cluster closely at the primary bacteria level (**Figure 2A**). In contrast, *A. christensii* and *L. iners* showed significant mutualistic benefits in pairwise simulations. In terms of competition, a small subset of *Gardnerella* strains at the bottom-most of the heatmap were repeatedly outcompeted, as evidenced by high biomass flux decrease across almost all interactions (**Figure 2B**). At the net flux level, *L. iners* and *A. christensiin* at the primary bacteria level flux values indicated consistent biomass benefit from mutualistic interactions and low biomass cost due to competition (**Figure 2C**). Finally, *L. iner*s, all *Prevotella* species, *H. timonensis, F.*

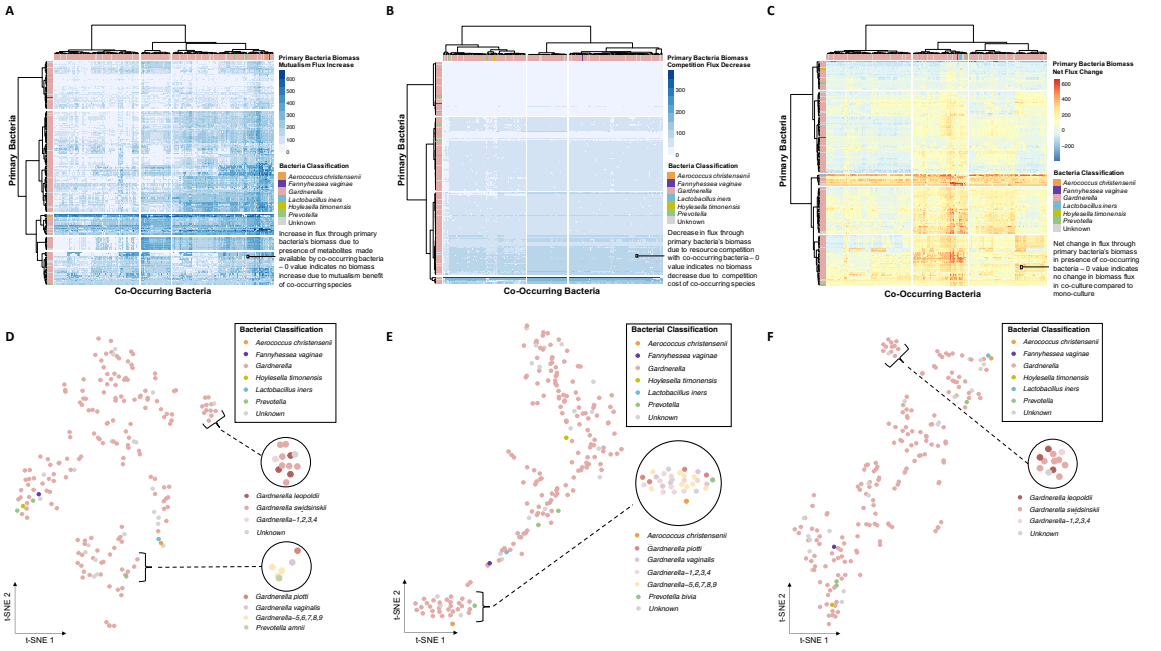


Figure 2 Heatmap illustrating primary bacteria biomass flux change in the presence of co-occurring bacteria when simulating (A) mutualism, (B) competition, (C) net interactions; t-SNE plots of primary bacteria biomass flux change reduced across all co-occurring bacteria when simulating (D) mutualism, (E) competition, (F) net interactions

vaginae, A. christensii, and some *Gardnerella* strains played significant mutualistic roles in approximately half of the primary bacterial strain interactions, while playing a more neutral to competitive role in the other half of primary bacterial strain interactions.

Using t-SNE, we assessed single bacterial species across interactions – mutualistic, competitive, and net – to examine high-level patterns of bacterial similarity. We found inter-*Gardnerella* clade clustering during mutualism, including *G. leopoldii, G. swidsinksii, G. 1,2,3,4,* and unknown strains (**Figure 2D**). Additionally, we observed clustering across *G. vaginalis, G. piotti, G. 5,6,7,8,9,* and *P. amnii*. Specifically, *P. amnii* clustered independently from *P. bivia* and *P. buccalis*. Conversely, there was less prevalent local structure observed at the competition level including clustering of *A. christensenii, G. piotti, G. vaginalis, G. 1,2,3,4, G. 5,6,7,8,9, P. bivia* and an unknown *Gardnerella* (**Figure 2E**). Looking at net flux values, we see analogous inter-*Gardnerella* clustering of species clustered in mutualism (**Figure 2F**).

Based on our analysis of average mutualism benefit and average competition cost as quantified via change in flux through biomass, we found that *L. iners* and *A. christensenii* both significantly benefited from pairwise mutualism compared to low competition cost (**Figure 3A**). Conversely, *H. timonensis* showed high competition cost relative to its low mutualism benefit. Our analyses showed that most of the investigated strains fell within

the low to medium mutualism benefit range, with a similarly low to medium competition cost, except for one *Gardnerella* strain with significant competition cost.

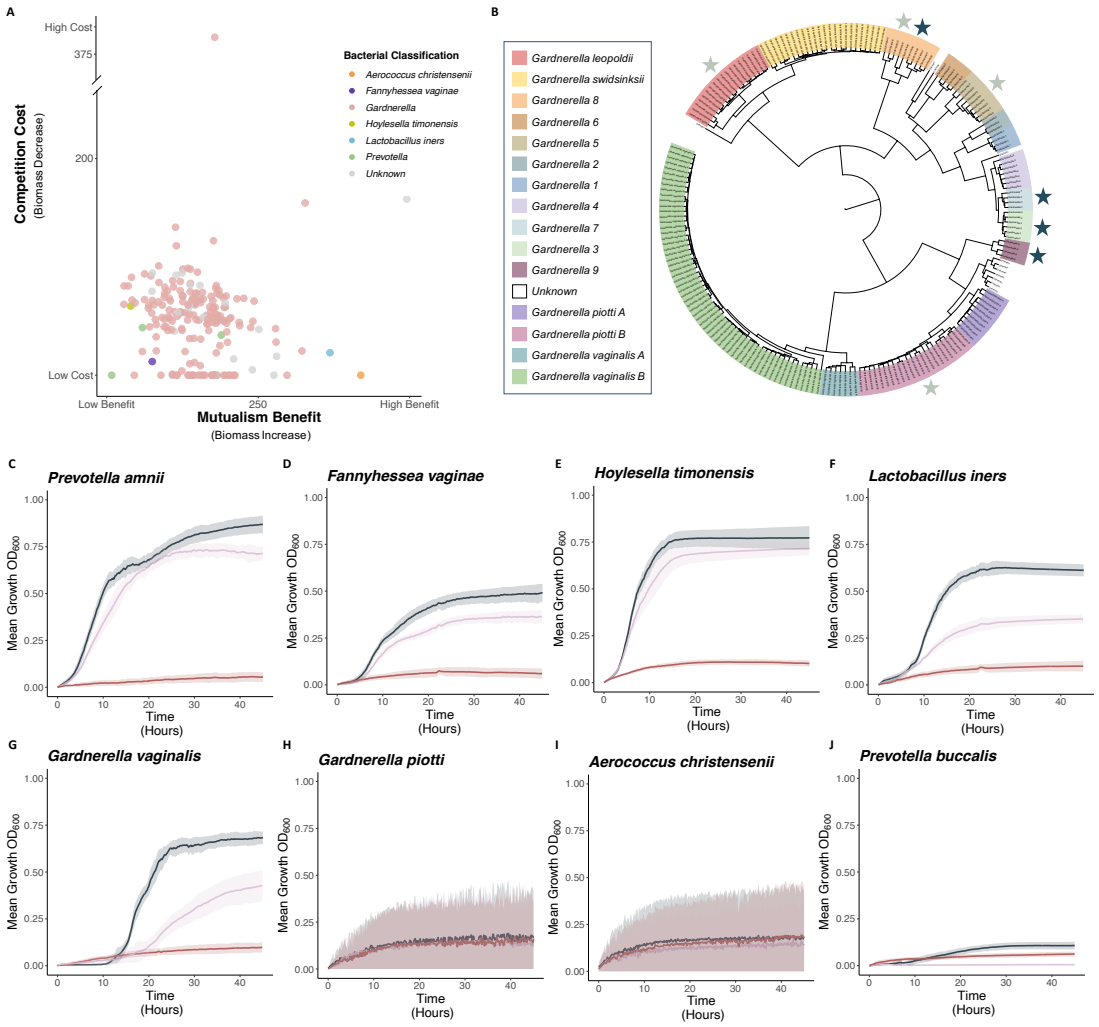
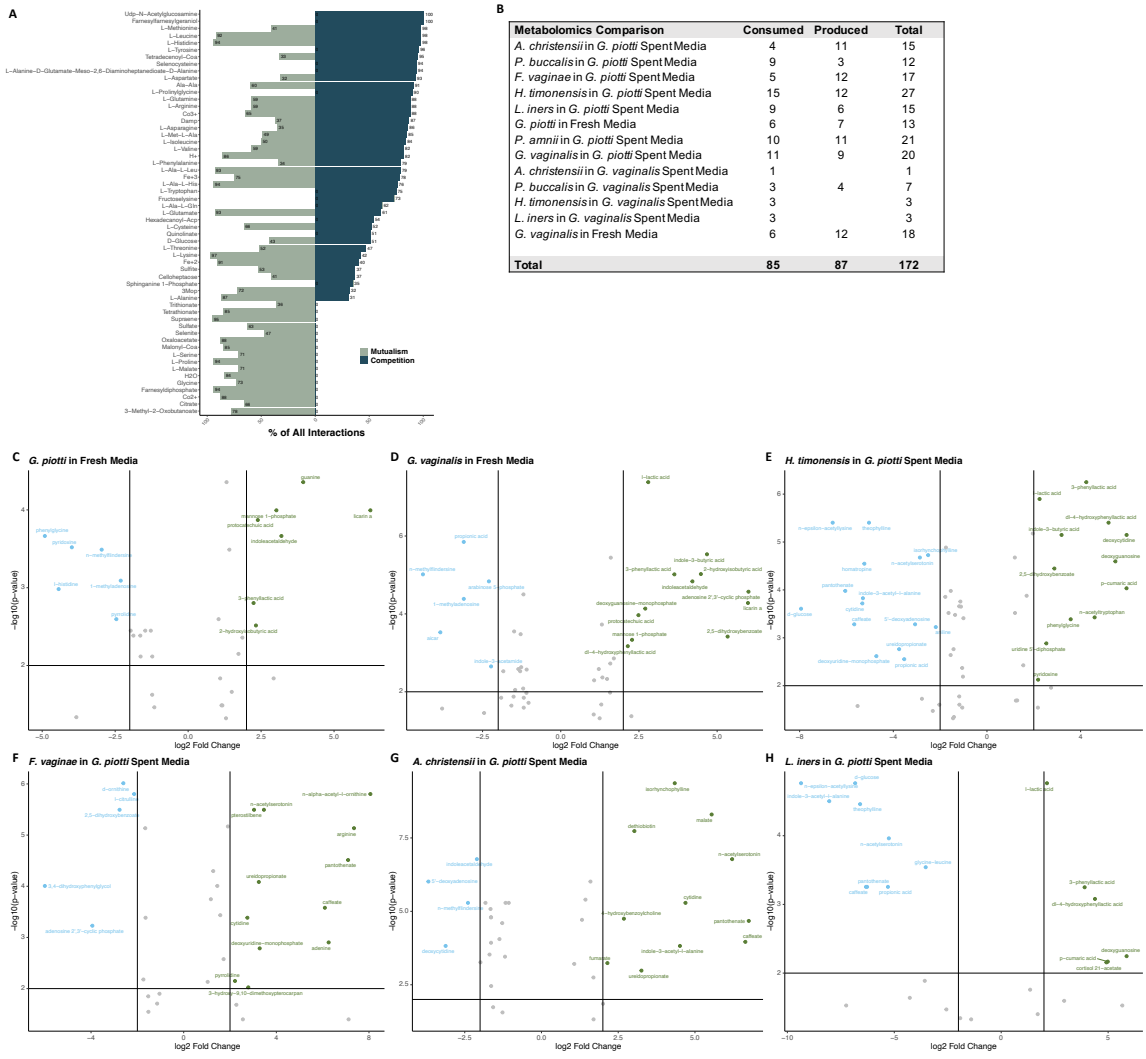


Figure 3 (A) average biomass increase due to mutualism benefit vs. average biomass decrease due to competition cost for each bacteria (B) dendrogram of *Gardnerella* species, with light stars denoting top four most mutualistic bacteria and dark stars indicating top four most competitive bacteria across all interactions simulations (C-J) growth curves in NYC III enriched media (black), spent *G. piotti* media (light pink), and spent *G. vaginalis* media (dark pink/red), with mean and two standard deviation spread

Using a one-tailed t-test, followed by Bonferroni multiple test corrections, we highlighted secondary bacteria that most significantly benefited primary bacteria with which they co-occurred. The top five most mutualistic bacteria belonged to the following *Gardnerella* clades: *G. 8* (p-value: 2.18 x 10^{-99}, t-statistic: 42.6), unknown strain (p-value: 4.59 x 10^{-93}, t-statistic: 39.2), *G. piotti B* (p-value: 7.12 x 10^{-85}, t-statistic: 35.2), *G. leopoldii* (p-value:

3.50 x 10^{-84}, t-statistic: 34.8), and *G. 5* (p-value: 4.31 x 10^{-80}, t-statistic: 32.9). Conversely, the top bacterial competitors were defined as secondary bacteria that most significantly outcompeted primary bacteria, resulting in a decrease in biomass during pairwise simulations. The top five most competitive bacteria fell under the following *Gardnerella* clades: *G. 8* (p-value: 4.82 x 10^{-51}, t-statistic: 21.3), unknown species (p-value: 3.77 x 10^{-44}, 18.9), *G. 9* (p-value: 1.57 x 10^{-43}, t-statistic: 18.7). *G. 3* (p-value: 1.74 x 10^{-43}, t-statistic: 18.7), and *G. 7* (p-value: 1.74 x 10^{-43}, t-statistic: 18.7). We overlaid the most competitive and most mutualistic bacteria onto the *Gardnerella* ANI dendrogram and observed a greater genetic diversity among top mutualistic bacteria compared to the top competitive bacterial strains (**Figure 3B**).

Our analysis of the most commonly shared and/or competed for metabolites highlighted the physiological relevance of four metabolites in the BV-positive *in silico* bacterial interaction simulation (**Figure 4A**). Specifically, L-histidine (Competitive interactions: 98%; Mutualistic interactions: 94%), which can be decarboxylated to form histamine, a key host inflammatory response immune regulator (Branco et al., 2018). Selenocysteine (Competitive interactions: 94%), is an essential component of selenoproteins, which play a role in host immune function (Avery & Hoffmann, 2018; Rayman, 2000). Sphinganine 1-phosphate (S1P) (Competitive interactions: 35%) has been associated with pro-inflammatory properties and implicated in tissue damage, but rarely implicated in bacterial metabolism (Ledgerwood et al., 2008; Liu et al., 2021). Iron (III) (Competitive interactions: 78%; Mutualistic interactions: 75%) is an essential *Gardnerella* growth micronutrient (Jarosik et al., 1998).



Metabolomics Comparison	Consumed	Produced	Total
A. christensii in *G. piotti* Spent Media	4	11	15
P. buccalis in *G. piotti* Spent Media	9	3	12
F. vaginae in *G. piotti* Spent Media	5	12	17
H. timonensis in *G. piotti* Spent Media	15	12	27
L. iners in *G. piotti* Spent Media	9	6	15
G. piotti in Fresh Media	6	7	13
P. amnii in *G. piotti* Spent Media	10	11	21
G. vaginalis in *G. piotti* Spent Media	11	9	20
A. christensii in *G. vaginalis* Spent Media	1		1
P. buccalis in *G. vaginalis* Spent Media	3	4	7
H. timonensis in *G. vaginalis* Spent Media	3		3
L. iners in *G. vaginalis* Spent Media	3		3
G. vaginalis in Fresh Media	6	12	18
Total	**85**	**87**	**172**

Figure 4 (A) Most highly competed for (competition) and/or shared (mutualism) metabolites across all interaction simulations (C-H) Volcano plots of differential metabolites (blue: consumed; green: produced; grey: not significant) from co-occurring species grown in *G. vaginalis* spent media

Pairwise *in vitro* bacterial interactions

Growth Curves

Because of the fastidious nature of *P. bivia,* it was not used for *in vitro* analysis. *P. amnii, F. vaginae,* and *H. timonensis* all showed similar growth capacity in NYCIII enriched media compared to *G. piotti* spent media (**Figure 4C-E**). *L. iners* and *G. vaginalis* had 44%

and 37% reduced growth in *G. piotti* spent media compared to enriched media (**Figure 4F-G)**). All species had minimal growth in *G. vaginalis* spent media. *P. buccalis, A. christensii,* and *G. piotti* had minimal growth across all three conditions (**Figure H-J**).

Metabolomics

Both *P. amnii* and *G. piotti* showed no growth in *G. vaginalis* spent media and were not submitted for metabolomics analysis. After filtering for p-value and $\log_2$(fold change) (<0.01, > |2|) 85 metabolites were identified as significantly consumed and 87 metabolites were identified as significantly produced (**Figure 4B**). Using ORA we identified the top three metabolite super-classes that were most significantly enriched in both consumed in produced metabolites based on enrichment ratio. Consumed metabolites were enriched for nucleic acids (Enrichment-ratio: 272.7; Holm p-value: 1.95×10^{-12}), carbohydrates, (Enrichment-ratio: 182.9; Holm p-value: 1.57×10^{-5}), and organic acids, (Enrichment-ratio: 54.5; Holm p-value: 3.05×10^{-8}). Produced metabolites were enriched for nucleic acids (Enrichment-ratio: 263.2; Holm p-value: 2.0×10^{-14}), carbohydrates (Enrichment-ratio: 100.5; Holm p-value: 0.004), and organoheterocyclic compounds (Enrichment-ratio: 58.4; Holm p-value: 1.74×10^{-5}).

G. piotti significantly consumed L-histidine when grown in fresh media (**Figure 4C**). Both *G. vaginalis* and *G. piotti* produced protocatechuic acid (PCA) and 2-hydroxyisobutryic acid (2-HIBA) (**Figure 4C-D**). PCA has been shown to have beneficial anti-inflammatory effects in the gut (Murota et al., 2018). Conversely, previous research has identified 2-HIBA as a disease biomarker for diabetes, adiposity, as well as autoimmune diseases (Elliott et al., 2015; Li et al., 2009; Tsoukalas et al., 2020).Both *G. vaginalis* and *H.*

timonensis significantly consumed the volatile fatty acid, propionic acid, which has been found to be elevated in BV positive cervicovaginal fluid (**Figure 4D-E**) (Al-Mushrif et al., 2000; Delgado-Diaz et al., 2020; Gajer et al., 2012; Mirmonsef et al., 2011; Spiegel et al., 1983). *F. vaginae* grown in *G. vaginalis* spent media did not have any significantly different metabolites. Caffeate is produced by both *A. christensii* and *F. vaginae* when grown in *G. piotti.* Caffeate is able to bind to estrogen receptor alpha (ERα), which is located throughout the vaginal tissue (**Figure 4F-G**) (Rezaei-Seresht et al., 2019). *L. iners* significantly produced lactic acid only when grown in *G. piotti* spent media (**Figure 4H**).

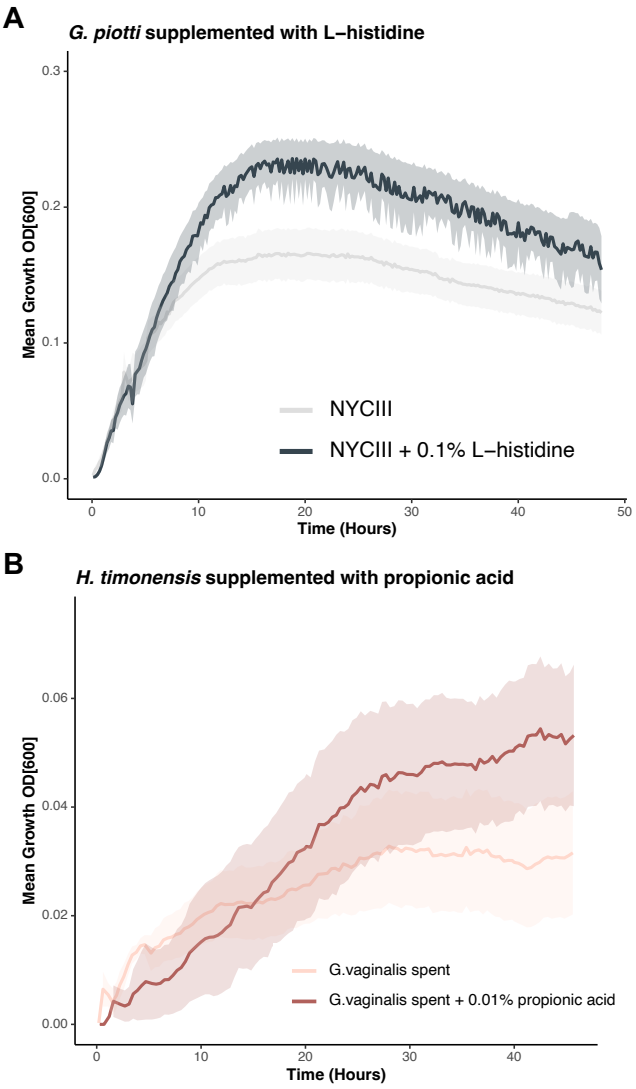


Figure 5: Growth curves of A) *G. piotti* and B) *H. timonensis* in supplemented and non-supplemented media

Supplemented Growth

In silico analysis revealed that L-histidine was consumed through both mutualism and competition, a finding further supported by metabolomics data demonstrating significant consumption of L-histidine by *G. piotti in vitro*. To investigate the growth benefits of L-histidine, we conducted a direct comparison of *G. piotti* growth in

NYCIII media and NYCIII media supplemented with 0.1% L-histidine. The presence of 0.1% L-histidine resulted in a significant enhancement of *G. piotti* growth, as confirmed by a one-tailed t-test of OD600 readings at the 48-hour time-point (t-statistic: 2.2, p-value < 0.05) (**Figure 5A**).

Considering the enrichment of propionic acid in BV positive cervicovaginal fluid, along with the significant consumption of propionic acid by both *H. timonensis* in *G. piotti* spent media and by *G. vaginalis* in NYCIII media, we hypothesized that *H. timonensis* and *G. vaginalis* compete for propionic acid. To test this hypothesis, we cultured *H. timonensis* in *G. vaginalis* spent media and in *G. vaginalis* spent media supplemented with 0.01% propionic acid. At the 35-hour mark, we observed a borderline significant enhancement of *H. timonensis* growth (t-statistic: 1.4, p-value: 0.09), although overall growth remained minimal (**Figure 5B**).

Discussion

Despite the significant impact of BV on both acute and chronic health, there is a lack of information regarding the microbial community structures and metabolic interactions underlying this condition. Our study employs clinical data, *in silico,* and *in vitro* analyses to investigate the variations in *Gardnerella* co-occurring species between symptomatic and asymptomatic BV, as well as explore the differential interactions of *Gardnerella* strains with each other and non-*Gardnerella* co-occurring species. Additionally, we compare *in silico* predictions with *in vitro* metabolomics of spent media to further understand the metabolites defining inter-bacterial interactions.

Using *in vivo* metagenomic analysis of vaginal swabs, we observed variations in BV community structure between symptomatic and asymptomatic samples. Our findings indicate a higher prevalence of *A. christensii* and *P. bivia* in symptomatic samples, while *P. amnii, P. buccalis*, and *H. timonensis* were more frequently found in asymptomatic samples. These differences in community structures underscore the need for further sub-categorization of BV. By establishing more specific definitions, development of more targeted treatments that move away from the current broad-spectrum antibiotic approach are possible (Kedaigle & Fraenkel, 2018; Sun et al., 2014; Zhang et al., 2022).

The complexity of BV is further supported by *in silico* simulations of pair-wise bacterial interactions. Our analysis reveals significant clustering between *Gardnerella* species based on assessed mutualistic benefit, indicating that genetic similarity does not necessarily correlate with functional metabolic similarities. Additionally, we observe a bifurcated metabolic benefit versus metabolic cost of *L. iners*, all *Prevotella* species, *H. timonensis, F. vaginae, A. christensii*, and certain *Gardnerella* strains when co-occurring with *Gardnerella* strains. This result suggests that a strain's metabolic relationship may be beneficial in one inter-bacterial context but detrimental in another. Overall, *A. christensii* and *L. iners* appear to consistently benefit the most from inter-bacterial interactions *in silico*. *In vitro* experiments reveal that *A. christensii* does not experience growth benefits when cultured in spent media of either *G. vaginalis* or *G. piotti*. These findings differ from our *in silico* predictions and demonstrate that pair-wise simulations cannot fully capture the complexity of the polymicrobial communities present in BV.

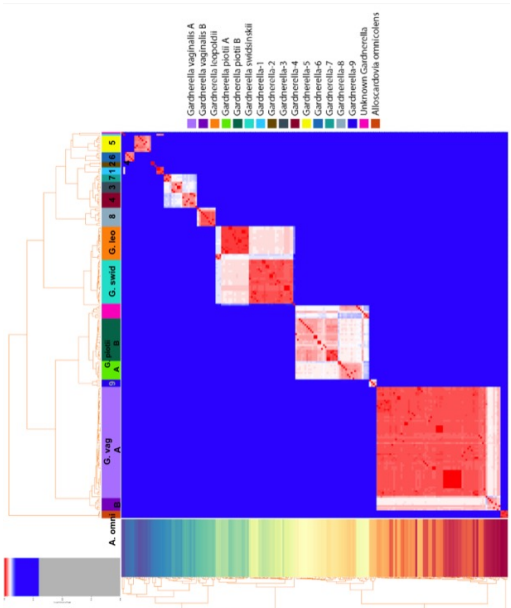
Using metabolomics analysis, we investigated what specific metabolites were being consumed and produced by co-occurring species in BV. *G. piotti* significantly consumed L-histidine *in vitro,* which recapitulated our *in silico* prediction that L-histidine is highly competed over. Our metabolomics analysis was also able to reveal metabolite production and consumption that was not demonstrated in our *in silico* analysis. Specifically, *A. christensii* and *F. vaginae* significantly produced caffeate when grown in *G. piotti* spent media. Previous studies have shown that caffeate can bind to ER$\boldsymbol{\alpha}$, which regulates estrogen sensitive gene expression. Reduced circulating estrogen levels results in vaginal epithelial atrophy, as seen in post-menopausal women (Jelinsky et al., 2008). These findings point to the importance of ER$\boldsymbol{\alpha}$ in supporting vaginal epithelial health. Our data suggest that metabolic byproducts of BV may alter vaginal epithelial regulation. Previous research has shown BV-associated metabolites can alter vaginal immune response (Eade et al., 2012; Fichorova et al., 2013).

Conclusion

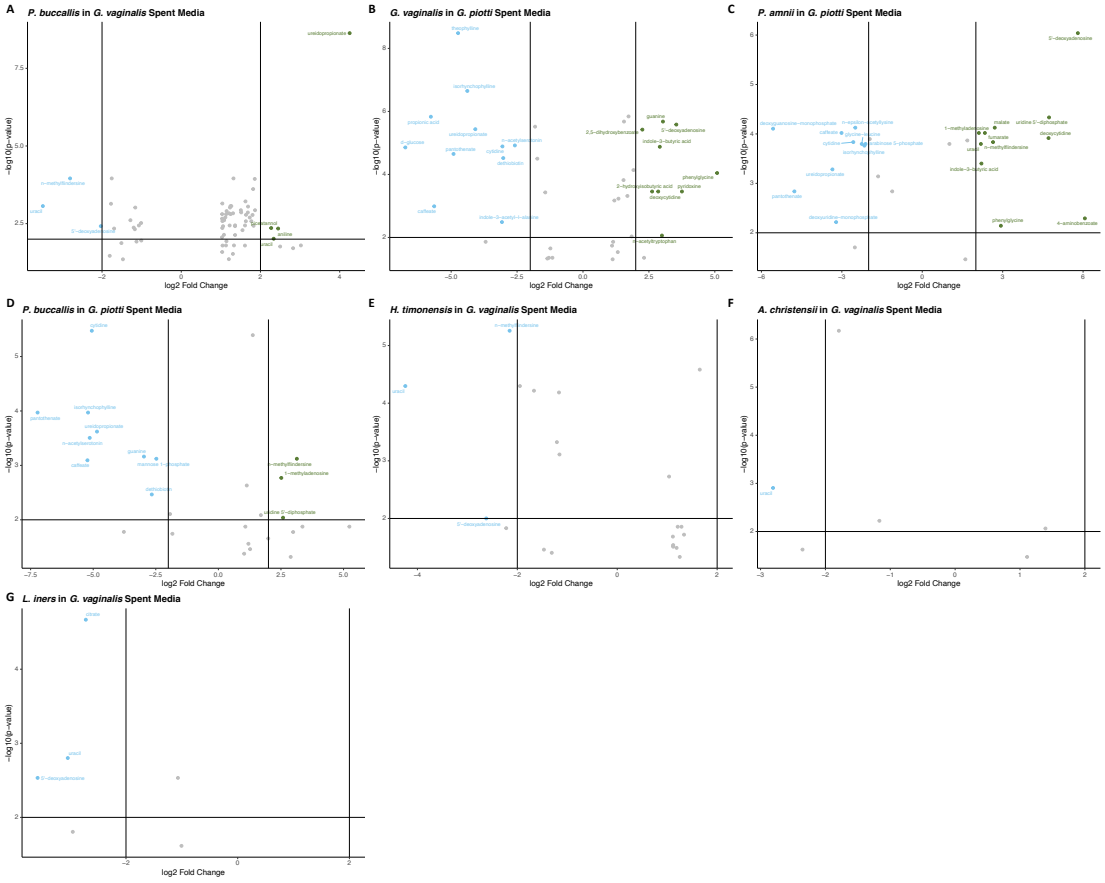
In summary, our study sheds light on the varied community composition and metabolic interactions associated with BV, a prevalent vaginal condition that significantly impacts women's health. By analyzing vaginal metagenomic data and constructing pair-wise GENRE simulations of BV-associated bacteria, we highlighted the context dependent nature of these competitive and mutualistic relationships. Through *in vitro* metabolomics analysis, we further investigated the role of specific metabolites in these inter-bacterial interactions. Our findings further emphasize the complex and diverse nature of BV-associated bacterial community structures. Additionally, we need to develop intricate

simulations that go beyond pairwise interactions to more accurately recapitulate *in vivo* community dynamics. Further research in this field will lead to improved treatments for BV, as well as a deeper understanding of women's health.

Supplementary



Supplementary 1 *Gardnerella* average nucleotide identity (ANI) hierarchically clustered distance matrix



Supplementary 2 (A-G) Volcano plots of differential metabolites (blue: consumed; green: produced; grey: not significant) from co-occurring species grown in *G. vaginalis* spent media

References

Allsworth, J. E., Lewis, V. A., & Peipert, J. F. (2008). Viral sexually transmitted infections and bacterial vaginosis: 2001-2004 National Health and Nutrition Examination Survey data. *Sexually Transmitted Diseases*, *35*(9), 791–796. https://doi.org/10.1097/OLQ.0B013E3181788301

Al-Mushrif, S., Eley, A., & Jones, B. M. (2000). Inhibition of chemotaxis by organic acids from anaerobes may prevent a purulent response in bacterial vaginosis. *Journal of Medical Microbiology*, *49*(11), 1023–1030. https://doi.org/10.1099/0022-1317-49-11-1023

Avery, J. C., & Hoffmann, P. R. (2018). Selenium, Selenoproteins, and Immunity. *Nutrients*, *10*(9), 1203. https://doi.org/10.3390/nu10091203

Bonferroni, C. E. (1936). Teoria statistica delle classi e calcolo delle probabilità. *R Istituto Superiore Di Scienze Economiche e Commerciali Di Firenze*, *8*, 3–62.

Bradshaw, C. S., & Sobel, J. D. (2016). Current Treatment of Bacterial Vaginosis-Limitations and Need for Innovation. *Journal of Infectious Diseases*, *214*(Suppl 1), S14–S20. https://doi.org/10.1093/infdis/jiw159

Branco, A. C. C. C., Yoshikawa, F. S. Y., Pietrobon, A. J., & Sato, M. N. (2018). Role of Histamine in Modulating the Immune Response and Inflammation. *Mediators of Inflammation*, *2018*, 9524075. https://doi.org/10.1155/2018/9524075

Castro, J., Jefferson, K. K., & Cerca, N. (2019). Innate immune components affect growth and virulence traits of bacterial-vaginosis-associated and non-bacterial-vaginosis-associated Gardnerella vaginalis strains similarly. *Pathogens and Disease*, *76*(9), 1–7. https://doi.org/10.1093/femspd/fty089

Culhane, J. F., Rauh, V., McCollum, K. F., Elo, I. T., & Hogan, V. (2002). Exposure to chronic stress and ethnic differences in rates of bacterial vaginosis among pregnant women. *American Journal of Obstetrics and Gynecology*, *187*(5), 1272–1276. https://doi.org/10.1067/MOB.2002.127311

De Seta, F., Campisciano, G., Zanotta, N., Ricci, G., & Comar, M. (2019). The Vaginal Community State Types Microbiome-Immune Network as Key Factor for Bacterial Vaginosis and Aerobic Vaginitis. *Frontiers in Microbiology*, *10*, 2451. https://doi.org/10.3389/fmicb.2019.02451

Delgado-Diaz, D. J., Tyssen, D., Hayward, J. A., Gugasyan, R., Hearps, A. C., & Tachedjian, G. (2020). Distinct Immune Responses Elicited From Cervicovaginal Epithelial Cells by Lactic Acid and Short Chain Fatty Acids Associated With Optimal and Non-optimal Vaginal Microbiota. *Frontiers in Cellular and Infection Microbiology*, *9*. https://www.frontiersin.org/articles/10.3389/fcimb.2019.00446

Dillard, L. R., Glass, E. M., Lewis, A. L., Thomas-White, K., & Papin, J. A. (2022). Metabolic Network Models of the Gardnerella Pangenome Identify Key Interactions with the Vaginal Environment. *MSystems*, *8*(1), e00689-22. https://doi.org/10.1128/msystems.00689-22

Eade, C. R., Diaz, C., Wood, M. P., Anastos, K., Patterson, B. K., Gupta, P., Cole, A. L., & Cole, A. M. (2012). Identification and Characterization of Bacterial Vaginosis-Associated Pathogens Using a Comprehensive Cervical-Vaginal Epithelial Coculture Assay. *PLoS ONE*, *7*(11). https://doi.org/10.1371/journal.pone.0050106

Feehily, C., Crosby, D., Walsh, C. J., Lawton, E. M., Higgins, S., McAuliffe, F. M., & Cotter, P. D. (2020). Shotgun sequencing of the vaginal microbiome reveals both

a species and functional potential signature of preterm birth. *Npj Biofilms and Microbiomes*, *6*(1), 1–9. https://doi.org/10.1038/s41522-020-00162-8

Fettweis, J. M., Serrano, M. G., Brooks, J. P., Edwards, D. J., Girerd, P. H., Parikh, H. I., Huang, B., Arodz, T. J., Edupuganti, L., Glascock, A. L., Xu, J., Jimenez, N. R., Vivadelli, S. C., Fong, S. S., Sheth, N. U., Jean, S., Lee, V., Bokhari, Y. A., Lara, A. M., … Buck, G. A. (2019). The vaginal microbiome and preterm birth. *Nature Medicine*, *25*(6), Article 6. https://doi.org/10.1038/s41591-019-0450-2

Fichorova, R. N., Buck, O. R., Yamamoto, H. S., Fashemi, T., Dawood, H. Y., Fashemi, B., Hayes, G. R., Beach, D. H., Takagi, Y., Delaney, M. L., Nibert, M. L., Singh, B. N., & Onderdonk, A. B. (2013). The villain team-up or how Trichomonas vaginalis and bacterial vaginosis alter innate immunity in concert. *Sexually Transmitted Infections*, *89*(6), 460–466. https://doi.org/10.1136/sextrans-2013-051052

Forsum, U., Holst, E., Larsson, P. G., Vasquez, A., Jakobsson, T., & Mattsby-Baltzer, I. (2005). Bacterial vaginosis—A microbiological and immunological enigma. *APMIS*, *113*(2), 81–90. https://doi.org/10.1111/j.1600-0463.2005.apm1130201.x

Gajer, P., Brotman, R. M., Bai, G., Sakamoto, J., Schütte, U. M. E., Zhong, X., Koenig, S. S. K., Fu, L., Ma, Z., Zhou, X., Abdo, Z., Forney, L. J., & Ravel, J. (2012). Temporal dynamics of the human vaginal microbiota. *Science Translational Medicine*, *4*(132). https://doi.org/10.1126/SCITRANSLMED.3003605

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., & Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to

whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology*, *57*(1), 81–91. https://doi.org/10.1099/ijs.0.64483-0

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Horrocks, V., Hind, C. K., Wand, M. E., Fady, P.-E., Chan, J., Hopkins, J. C., Houston, G. L., Tribe, R. M., Sutton, J. M., & Mason, A. J. (n.d.). Nuclear Magnetic Resonance Metabolomics of Symbioses between Bacterial Vaginosis-Associated Bacteria. *MSphere*, *7*(3), e00166-22. https://doi.org/10.1128/msphere.00166-22

Janulaitiene, M., Paliulyte, V., Grinceviciene, S., Zakareviciene, J., Vladisauskiene, A., Marcinkute, A., & Pleckaityte, M. (2017). Prevalence and distribution of Gardnerella vaginalis subgroups in women with and without bacterial vaginosis. *BMC Infectious Diseases*, *17*(1), 1–9. https://doi.org/10.1186/S12879-017-2501-Y/TABLES/1

Jarosik, G. P., Land, C. B., Duhon, P., Chandler, R., & Mercer, T. (1998). Acquisition of Iron by Gardnerella vaginalis. *Infection and Immunity*, *66*(10), 5041–5047.

Jelinsky, S. A., Choe, S. E., Crabtree, J. S., Cotreau, M. M., Wilson, E., Saraf, K., Dorner, A. J., Brown, E. L., Peano, B. J., Zhang, X., Winneker, R. C., & Harris, H. A. (2008). Molecular analysis of the vaginal response to estrogens in the ovariectomized rat and postmenopausal woman. *BMC Medical Genomics*, *1*(1), 27. https://doi.org/10.1186/1755-8794-1-27

Jenior, M. L., Glass, E. M., & Papin, J. A. (2023). Reconstructor: A COBRApy compatible tool for automated genome-scale metabolic network reconstruction

with parsimonious flux-based gap-filling. *Bioinformatics*, btad367. https://doi.org/10.1093/bioinformatics/btad367

Jenior, M. L., Leslie, J., Kolling, G. L., Archbald-Pannone, L., Powers, D., Petri, W. A., & Papin, J. A. (Under Review). Systems-ecology designed bacterial consortium protects from severe Clostridioides difficile infection. *Nature Biotechnology*.

Kedaigle, A., & Fraenkel, E. (2018). Turning omics data into therapeutic insights. *Current Opinion in Pharmacology*, *42*, 95–101. https://doi.org/10.1016/j.coph.2018.08.006

Kolde, R. (2019). Package "pheatmap": Pretty heatmaps. *Version 1.0.12*, 1–8.

Koumans, E. H., Sternberg, M., Bruce, C., McQuillan, G., Kendrick, J., Sutton, M., & Markowitz, L. E. (2007). The prevalence of bacterial vaginosis in the United States, 2001-2004; associations with symptoms, sexual behaviors, and reproductive health. *Sexually Transmitted Diseases*, *34*(11), 864–869. https://doi.org/10.1097/OLQ.0B013E318074E565

Ledgerwood, L. G., Lal, G., Zhang, N., Garin, A., Esses, S. J., Ginhoux, F., Merad, M., Peche, H., Lira, S. A., Ding, Y., Yang, Y., He, X., Schuchman, E. H., Allende, M. L., Ochando, J. C., & Bromberg, J. S. (2008). The sphingosine 1-phosphate receptor 1 causes tissue retention by inhibiting the entry of peripheral tissue T lymphocytes into afferent lymphatics. *Nature Immunology*, *9*(1), Article 1. https://doi.org/10.1038/ni1534

Lele, S., & Richtsmeier, J. T. (1995). Euclidean distance matrix analysis: Confidence intervals for form and growth differences. *American Journal of Physical Anthropology*, *98*(1), 73–86. https://doi.org/10.1002/ajpa.1330980107

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, *49*(W1), W293–W296. https://doi.org/10.1093/nar/gkab301

Lieven, C., Beber, M. E., Olivier, B. G., Bergmann, F. T., Ataman, M., Babaei, P., Bartell, J. A., Blank, L. M., Chauhan, S., Correia, K., Diener, C., Dräger, A., Ebert, B. E., Edirisinghe, J. N., Faria, J. P., Feist, A. M., Fengos, G., Fleming, R. M. T., García-Jiménez, B., Zhang, C. (2020). MEMOTE for standardized genome-scale metabolic model testing. *Nature Biotechnology*, *38*(3), 272–276. https://doi.org/10.1038/s41587-020-0446-y

Liu, Z., Li, L., Fang, Z., Lee, Y., Zhao, J., Zhang, H., Chen, W., Li, H., & Lu, W. (2021). Integration of Transcriptome and Metabolome Reveals the Genes and Metabolites Involved in Bifidobacterium bifidum Biofilm Formation. *International Journal of Molecular Sciences*, *22*(14), 7596. https://doi.org/10.3390/ijms22147596

Lopes dos Santos Santiago, G., Tency, I., Verstraelen, H., Verhelst, R., Trog, M., Temmerman, M., Vancoillie, L., Decat, E., Cools, P., & Vaneechoutte, M. (2012). Longitudinal qPCR study of the dynamics of L. crispatus, L. iners, A. vaginae, (sialidase positive) G. vaginalis, and P. bivia in the vagina. *PloS One*, *7*(9). https://doi.org/10.1371/JOURNAL.PONE.0045281

Madden, T. (2002, October 9). *The BLAST Sequence Analysis Tool*. The NCBI Handbook. https://www.ncbi.nlm.nih.gov/books/NBK21097/

Madden, T., & Camacho, C. (2021). BLAST+ features. In *BLAST® Command Line Applications User Manual [Internet]*. National Center for Biotechnology Information (US). https://www.ncbi.nlm.nih.gov/books/NBK569839/

Marrazzo, J. M., Antonio, M., Agnew, K., & Hillier, S. L. (2009). Distribution of Genital Lactobacillus Strains Shared by Female Sex Partners. *The Journal of Infectious Diseases*, *199*(5), 680–683. https://doi.org/10.1086/596632

Mayer, B. T., Srinivasan, S., Fiedler, T. L., Marrazzo, J. M., Fredricks, D. N., & Schiffer, J. T. (2015). Rapid and Profound Shifts in the Vaginal Microbiota Following Antibiotic Treatment for Bacterial Vaginosis. *The Journal of Infectious Diseases*, *212*(5), 793–802. https://doi.org/10.1093/INFDIS/JIV079

Mirmonsef, P., Gilbert, D., Zariffard, M., Hamaker, B., Kaur, A., Landay, A., & Spear, G. (2011). The effects of commensal bacteria on innate immune responses in the female genital tract. *American Journal of Reproductive Immunology (New York, N.Y. : 1989)*, *65*(3), 190–195. https://doi.org/10.1111/j.1600-0897.2010.00943.x

Murota, K., Nakamura, Y., & Uehara, M. (2018). Flavonoid metabolism: The interaction of metabolites and gut microbiota. *Bioscience, Biotechnology, and Biochemistry*, *82*(4), 600–610. https://doi.org/10.1080/09168451.2018.1444467

Muzny, C. A., & Kardas, P. (2020). A Narrative Review of Current Challenges in the Diagnosis and Management of Bacterial Vaginosis. *Sexually Transmitted Diseases*, *47*(7), 441–446. https://doi.org/10.1097/OLQ.0000000000001178

Olson, R. D., Assaf, R., Brettin, T., Conrad, N., Cucinell, C., Davis, J. J., Dempsey, D. M., Dickerman, A., Dietrich, E. M., Kenyon, R. W., Kuscuoglu, M., Lefkowitz, E. J., Lu, J., Machi, D., Macken, C., Mao, C., Niewiadomska, A., Nguyen, M., Olsen, G. J., … Stevens, R. L. (2023). Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): A resource combining PATRIC,

IRD and ViPR. *Nucleic Acids Research*, *51*(D1), D678–D689. https://doi.org/10.1093/nar/gkac1003

Pang, Z., Chong, J., Zhou, G., de Lima Morais, D. A., Chang, L., Barrette, M., Gauthier, C., Jacques, P.-É., Li, S., & Xia, J. (2021). MetaboAnalyst 5.0: Narrowing the gap between raw spectra and functional insights. *Nucleic Acids Research*, *49*(W1), W388–W396. https://doi.org/10.1093/nar/gkab382

Peebles, K., Velloza, J., Balkus, J. E., McClelland, R. S., & Barnabas, R. V. (2019). High Global Burden and Costs of Bacterial Vaginosis: A Systematic Review and Meta-Analysis. *Sexually Transmitted Diseases*, *46*(5), 304–311. https://doi.org/10.1097/OLQ.0000000000000972

Percy, A., Souza, Amanda, Ntai, Ioanna, & Amer, Bashar. (2022). From QC to Quantitation: Utility of QReSS™ Metabolites in FBS Measurements. *Cambridge Isotope Laboratories, Inc.*, *Application Note 51*. https://cil.showpad.com/share/PfylguI9IIFzYtChsvFrL

Percy, Andrew, Proos, Robert, & Demianova, Zuzana. (2021). Standardizing Quantitative Metabolomics Analyses Through the QReSS™ Kit. *Cambridge Isotope Laboratories, Inc.*, *Application Note 49*.

Pritchard, L., Glover, R. H., Humphris, S., Elphinstone, J. G., & Toth, I. K. (2015). Genomics and taxonomy in diagnostics for food security: Soft-rotting enterobacterial plant pathogens. *Analytical Methods*, *8*(1), 12–24. https://doi.org/10.1039/C5AY02550H

Rayman, M. P. (2000). The importance of selenium to human health. *The Lancet*, *356*(9225), 233–241. https://doi.org/10.1016/S0140-6736(00)02490-9

Rezaei-Seresht, H., Cheshomi, H., Falanji, F., Movahedi-Motlagh, F., Hashemian, M., & Mireskandari, E. (2019). Cytotoxic activity of caffeic acid and gallic acid against MCF-7 human breast cancer cells: An in silico and in vitro study. *Avicenna Journal of Phytomedicine*, *9*(6), 574–586. https://doi.org/10.22038/AJP.2019.13475

Schoch, C. L., Ciufo, S., Domrachev, M., Hotton, C. L., Kannan, S., Khovanskaya, R., Leipe, D., Mcveigh, R., O'Neill, K., Robbertse, B., Sharma, S., Soussov, V., Sullivan, J. P., Sun, L., Turner, S., & Karsch-Mizrachi, I. (2020). NCBI Taxonomy: A comprehensive update on curation, resources and tools. *Database: The Journal of Biological Databases and Curation*, *2020*, baaa062. https://doi.org/10.1093/database/baaa062

Schwiertz, A., Taras, D., Rusch, K., & Rusch, V. (2006). Throwing the dice for the diagnosis of vaginal complaints? *Annals of Clinical Microbiology and Antimicrobials*, *5*. https://doi.org/10.1186/1476-0711-5-4

Shimaoka, M., Yo, Y., Doh, K., Kotani, Y., Suzuki, A., Tsuji, I., Mandai, M., & Matsumura, N. (2019). Association between preterm delivery and bacterial vaginosis with or without treatment. *Scientific Reports*, *9*(1), 1–8. https://doi.org/10.1038/s41598-018-36964-2

Spiegel, C. A., Amsel, R., & Holmes, K. K. (1983). Diagnosis of bacterial vaginosis by direct gram stain of vaginal fluid. *Journal of Clinical Microbiology*, *18*(1), 170–177. https://doi.org/10.1128/jcm.18.1.170-177.1983

Srinivasan, S., Hoffman, N. G., Morgan, M. T., Matsen, F. A., Fiedler, T. L., Hall, R. W., Ross, F. J., McCoy, C. O., Bumgarner, R., Marrazzo, J. M., & Fredricks, D. N.

(2012). Bacterial communities in women with bacterial vaginosis: High resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. *PloS One*, *7*(6), e37818. https://doi.org/10.1371/journal.pone.0037818

Student. (1908). The Probable Error of a Mean. *Biometrika*, *6*(1), 1–25. https://doi.org/10.2307/2331554

Sun, J., Bi, J., & Kranzler, H. R. (2014). Multi-view singular value decomposition for disease subtyping and genetic associations. *BMC Genetics*, *15*, 73. https://doi.org/10.1186/1471-2156-15-73

Tachedjian, G., Aldunate, M., Bradshaw, C. S., & Cone, R. A. (2017). The role of lactic acid production by probiotic Lactobacillus species in vaginal health. *Research in Microbiologoy*, *168*(9–10), 782–792. https://doi.org/10.1016/j.resmic.2017.04.001

Tortelli, B. A., Lewis, W. G., Allsworth, J. E., Member-Meneh, N., Foster, L. R., Reno, H. E., Peipert, J. F., Fay, J. C., & Lewis, A. L. (2020). Associations between the vaginal microbiome and Candida colonization in women of reproductive age. *American Journal of Obstetrics and Gynecology*, *222*(5), 471.e1-471.e9. https://doi.org/10.1016/J.AJOG.2019.10.008

Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., Kanazawa, M., VanderGheynst, J., Fiehn, O., & Arita, M. (2015). MS-DIAL: Data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nature Methods*, *12*(6), Article 6. https://doi.org/10.1038/nmeth.3393

Tsugawa, H., Nakabayashi, R., Mori, T., Yamada, Y., Takahashi, M., Rai, A., Sugiyama, R., Yamamoto, H., Nakaya, T., Yamazaki, M., Kooke, R., Bac-Molenaar, J. A., Oztolan-Erol, N., Keurentjes, J. J. B., Arita, M., & Saito, K. (2019). A

cheminformatics approach to characterize metabolomes in stable-isotope-labeled organisms. *Nature Methods*, *16*(4), Article 4. https://doi.org/10.1038/s41592-019-0358-2

Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, *9*(11).

Vitali, B., Cruciani, & F., Picone, G., Parolin, & C., Donders, & G., & Laghi, & L. (2015). *Vaginal microbiome and metabolome highlight specific signatures of bacterial vaginosis*. https://doi.org/10.1007/s10096-015-2490-y

Wang, J. (2000). Bacterial vaginosis. *Primary Care Update for OB/GYNS*, *7*(5), 181–185. https://doi.org/10.1016/S1068-607X(00)00043-3

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, *58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Zhang, S., Guo, Z., Tian, E., Liu, D., Wang, J., & Kong, W. (2022). Meniere disease subtyping: The direction of diagnosis and treatment in the future. *Expert Review of Neurotherapeutics*, *22*(2), 115–127. https://doi.org/10.1080/14737175.2022.2030221

Zhuang, Y., Wolford, B. N., Nam, K., Bi, W., Zhou, W., Willer, C. J., Mukherjee, B., & Lee, S. (2022). Incorporating family disease history and controlling case–control imbalance for population-based genetic association studies. *Bioinformatics*, *38*(18), 4337–4343. https://doi.org/10.1093/bioinformatics/btac459

Chapter 4: Beyond the Dissertation

Capstone Team

Context: During the 3rd year of my PhD I began mentoring undergraduate Christina D. George. We developed a research project that focused on understanding how changing available carbon sources impacted *Gardnerella* biofilm production. Christina's work developed into a capstone project titled "Carbon Source and Biofilm Formation: Implications for Bacterial Vaginosis Treatment Strategies". I was able to mentor Christina's team, which included Kaitlyn M. Gray and Peyton C. Johnston. The following portion of my dissertation summarizes the team's research and findings.

Introduction

Bacterial vaginosis (BV) is a polymicrobial vaginal condition characterized by an increase in pathogenic anaerobic bacteria, particularly the dominant genus *Gardnerella*, and a decrease in beneficial *Lactobacillus* species (Nasioudis et al., 2017; Onderdonk et al., 2016; Srinivasan et al., 2010). BV is the most prevalent vaginal condition among reproductive age women, accounting for over 60% of vulvovaginal infections (Marrazzo et al., 2009). The current treatment options for BV include antibiotics such as metronidazole, clindamycin, tinidazole, and secnidazole, as well as boric acid suppositories and other non-antibiotic treatments (Bradshaw et al., 2012; Heczko et al., 2015; C. Li et al., 2019). However, traditional antibiotic regimens have shown high rates of BV recurrence, with approximately 58% of women experiencing a recurrent infection within one year of initial treatment (Bradshaw et al., 2006, 2013). Recurring BV infections pose

an increased risk of antibiotic resistance, as BV-associated can accumulate mutations in response to successive antibiotic treatments.

Some research suggests that the ability of *Gardnerella* to form biofilms may contribute to the high rates of BV recurrence due to reduced susceptibility to antibiotics (T. Li et al., 2020; Swidsinski et al., 2008). *Gardnerella* biofilms are composed of a complex extracellular polymeric substance (EPS), consisting of sticky carbohydrates, proteins, DNA, and nucleic acids, which create a physical barrier around the enclosed bacteria (Machado & Cerca, 2015; Morrill et al., 2020; Ravel et al., 2011; Schwebke et al., 2014). This EPS hinders the penetration of antibiotics, thereby reducing their effectiveness in eliminating the bacteria associated with BV. Despite the potential role of biofilms in recurrent BV, there is limited knowledge regarding the metabolic mechanisms and composition of *Gardnerella* biofilms.

To address this knowledge gap, we conducted a study to examine the composition of biofilms formed by *Gardnerella vaginalis* and *Gardnerella piotti* using Scanning Electron Microscopy, Lectin staining, and enzyme disruption techniques. Additionally, we investigated how different carbon sources may affect biofilm production. Through our analysis, we gained insights into the biofilm composition of these two *Gardnerella* species and identified enzymes that specifically disrupt BV-associated biofilms. By characterizing *Gardnerella* biofilms, we aim to enhance BV treatment strategies and reduce the occurrence of recurrent infections.

Materials and Methods

Biofilm Growth

Gardnerella vaginalis (strain ATCC 14018) and *Gardnerella piotti* (strain JCP8151B) were cultured under anaerobic conditions in NYCIII enriched media at 37℃. After 18 hours of incubation, the samples were collected and centrifuged at 7,500 RPM for 5 minutes. The supernatant was discarded, and the pellet was resuspended in 1x PBS. This wash step was repeated three times. After the final wash, the bacteria were resuspended in 10 mL of NYCIII media and incubated anaerobically for 1 hour. A 50 µL inoculum was then added to 200 µL of reduced NYCIII in a 96-well tissue-culture treated plate. The plates were incubated for 48 hours under anaerobic conditions at 37℃.

Scanning Electron Microscopy Sample Preparation

The incubation and wash steps described above were used to prepare the inoculum of *G. vaginalis* and *G. piotti*. A 50 µL inoculum and 100 µL of NYCIII media were added to glass coverslips placed in six-well plates. Care was taken to maintain the surface tension of both the inoculum and media. The plates were incubated for 48 hours and then washed three times with 1x PBS. Each sample was fixed with 2% glutaraldehyde for 30 minutes. After fixation, the samples underwent repeated wash steps. The samples were dehydrated using graded ethanol, sequentially immersed in 30%, 50%, 70%, 80%, 90%, and 100% ethanol for five minutes each. The final dehydration step involved a five-minute immersion in Hexamethyldisilane (HMDS). The glass coverslips were then coated with gold, mounted on specimen stubs, and visualized using scanning electron microscopy (SEM).

Lectin Staining

FITC-conjugated WGA and TRITC-conjugated UEA lectins were diluted to a concentration of 100 µg/mL in sodium bicarbonate buffer. After a 48-hour incubation, the media was removed from mature *G. vaginalis* and *G. piotti* biofilms, and the wells were washed three times with deionized water. A blocking buffer of 200 µL of non-fat dry milk (NFDM) was added and incubated for 30 minutes at room temperature to prevent non-specific binding. The samples were then rinsed with a sodium bicarbonate buffer, followed by the addition of 200 µL of the respective lectin solution. The samples were incubated under tinfoil for 15 minutes. After removing the lectin solution, the samples were washed three times with the sodium bicarbonate buffer. The samples were resuspended in 200 µL of buffer, and absorbance was measured using a fluorescent plate reader. FITC samples were excited at 492 nm wavelengths, and TRITC samples were excited at 554 nm wavelengths. The absorbance values were collected at emission wavelengths of 517 nm and 570 nm, respectively.

Biofilm Carbon Source Growth

Following the biofilm growth procedures, a 96-well plate was prepared with 200 µL of synthetic vaginal media supplemented with 1% of a carbon source, including mannose, pyruvate, sucrose, acetate, glutamate, lactate, succinate, galactose, maltose, or mannitol. These carbon sources are all associated with BV, *Gardnerella*, and/or other similar biofilm forming organisms (Srinivasan et al., 2015). Next, 50 µL of the inoculum was added to each well, and the plate was incubated at 37°C under anaerobic conditions for 48 hours.

Biofilm Quantification

Biofilm quantification was performed using crystal violet. The samples were washed with a 0.9% sodium chloride solution and dried for 15 minutes at room temperature. Then, 200 µL of 100% methanol was added to each well and incubated at room temperature for 20 minutes. Following fixation, 200 µL of 1% crystal violet was added to each well, and the samples were incubated for 20 minutes at room temperature. The samples were washed with deionized water until non-specific staining was no longer observed. Subsequently, the samples were dried for 15 minutes. Finally, 200 µL of 33% acetic acid was added to each sample, and the plate was shaken at 300 rpm for ten minutes. The absorbance of the samples was measured at 595 nm using a Tecan plate reader.

Enzyme Biofilm Disruption

Based on prior research on *Gardnerella* adjacent organisms, five enzymes were chosen that were hypothesized to disrupt *G. vaginalis* biofilm including: cellulase, DNase I, α-amylase, lipase, and proteinase-k. Enzymes were diluted to a concentration of 2 U/mL in PBS. Mature biofilms were washed twice with PBS and then incubated with the respective enzyme solution for two hours. The samples were washed twice with PBS, and the biofilm quantification procedure described above was followed.

Results

SEM Visualization

Scanning electron microscopy (SEM) was used to visualize the surface topography and biofilm composition of the extracellular polymeric substance (EPS). No apparent visual differences were observed between the biofilms of *G. vaginalis* and *G. piotti* (**Figure 1**).

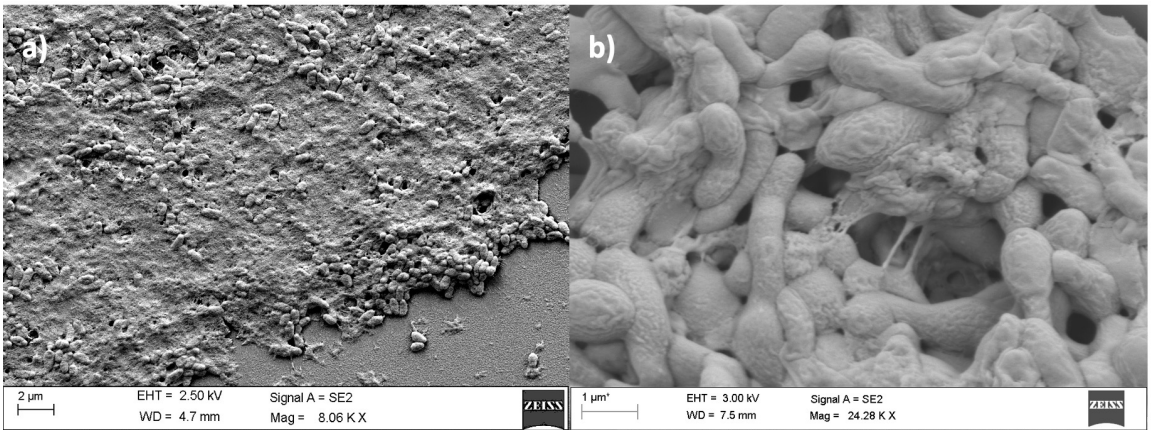


Figure 1: SEM images of A) *G. vaginalis* and B) *G. piotti* biofilm

Lectin Staining

Wheat germ agglutinin (WGA) lectin has a high affinity for the monosaccharides N-acetylneuraminic acid (sialic acid) and N-acetylglucosamine, which are components of the peptidoglycan layer of bacterial cell walls. Ulex europaeus agglutinin (UEA) lectin selectively binds to the monosaccharide L-fucose. Our results showed no significant increase in absorbance of either FITC-conjugated lectin in mature biofilms of *G. piotti* or *G. vaginalis* compared to non-*Gardnerella* biofilm controls, indicating no major differences in the presence of these carbohydrates.

Biofilm Carbon Source Utilization

To compare biofilm formation in synthetic vaginal media (SVM) with and without carbon source supplementation, a two-tailed t-test was performed. Both *G. vaginalis* and *G. piotti* showed a significant decrease in biofilm formation when comparing SVM to SVM supplemented with 1%: acetate (p-value: 5.65×10^{-10}, 1.44×10^{-9}), glutamate (p-value: 7.34×10^{-9}, 0.01), lactate (p-value 5.81×10^{-9}, 4.44×10^{-10}), and succinate ($p = 3.39 \times 10^{-10}$, 3.46×10^{-9}) (**Figure 2**). *G. piotti* biofilm formation significantly increased in the presence of SVM supplemented with 1% mannitol (p-value: 5.53×10^{-5}), but this effect was not

significant for *G. vaginalis*.

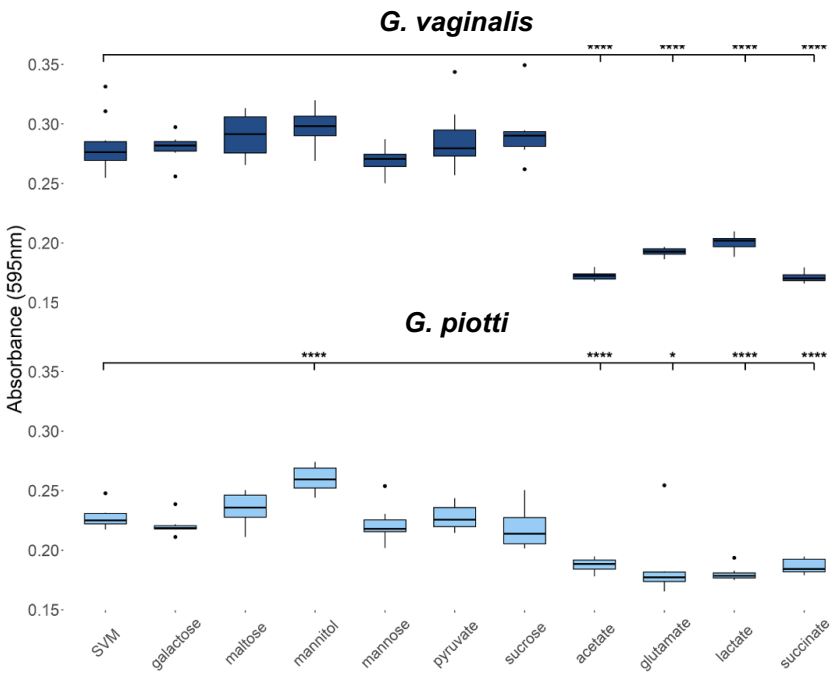


Figure 2: Biofilm growth via crystal violet staining across carbon sources relative to SVM + 1% FBS

G. vaginalis Biofilm Enzyme Disruption

Using a two-tailed t-test, we compared untreated *G. vaginalis* mature biofilm negative controls with enzyme-treated *G. vaginalis* mature biofilms using crystal violet biofilm quantification. All five enzymes significantly disrupted the biofilm: cellulase (p-value: 6.97×10^{-5}), DNase I (p-value: 6.16×10^{-6}), α-amylase (p-value: 1.76×10^{-6}), lipase (p-value: 1.15×10^{-5}), and proteinase-k (p-value: 2.58×10^{-6}) (**Figure 3**). Proteinase-k was the most effective biofilm disruptor, resulting in a five-fold decrease in biofilm amount compared to the untreated control.

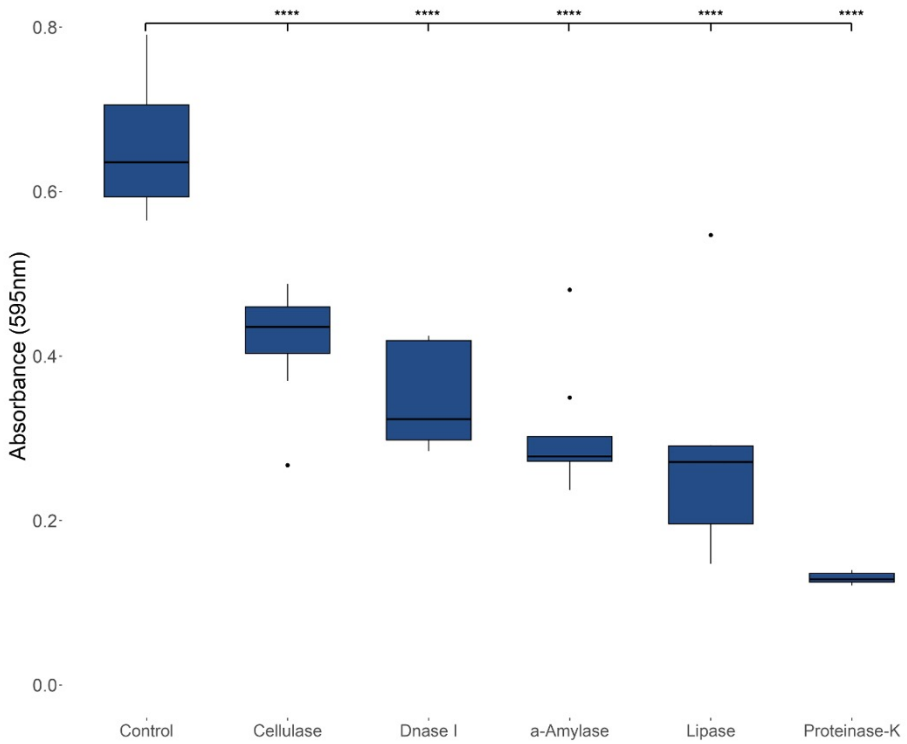


Figure 3: *G. vaginalis* biofilm disruption quantified via crystal violet following enzyme incubation relative to non-treated biofilm

Discussion

Despite the high prevalence, significant health impacts, and frequent recurrence of bacterial vaginosis (BV), there is still a lack of understanding regarding the role of *Gardnerella* biofilm during dysbiosis. *Gardnerella* has the ability to form biofilms, which contribute to reduced susceptibility to antibiotics (by creating a physical barrier that inhibits antibiotic penetrance) and may be a factor in BV recurrence (Khan & Hill, 2021). To advance treatment options, it is crucial to define the composition of *Gardnerella* biofilms and investigate the influence of nutrient availability on their formation. Specifically, our work focused on two primary *Gardnerella* species, *piotti* and *vaginalis*.

Qualitative analysis of biofilm structure using scanning electron microscopy (SEM) did not reveal substantial differences between *G. vaginalis* and *G. piotti*. This finding is expected considering the high genetic similarity (>98.5%) between the two species. The observed images aligned with documented properties of *Gardnerella*, including their ability to adhere and aggregate, which aids in attachment to vaginal epithelial cells and helps evade clearance by vaginal secretions (Limoli et al., 2015; Swidsinski et al., 2008).

Lectin staining results indicated that sialic acid and fucose, which were previously thought to be present in *Gardnerella* biofilms, were not components of *G. vaginalis* or *G. piotti* biofilms (W. G. Lewis et al., 2013; Srinivasan et al., 2015). However, these findings should be interpreted with caution due to the absence of a biofilm fixation step, potentially leading to the removal of the biofilm during the multiple washes. To obtain more accurate assessments of sialic acid and fucose presence in *Gardnerella* biofilms, future optimization of the lectin staining procedure should include biofilm fixation. Additionally, in an *in vivo* environment, free sialic acid would likely be more abundant due to vaginal epithelial cell turnover, increasing its incorporation into the extracellular polymeric substance (EPS) of the biofilm.

To directly evaluate the impact of available vaginal nutrients on biofilm formation, we investigated the influence of relevant carbon sources. Lactate and glutamate are associated with a *Lactobacillus*-dominant healthy vaginal environment and found at reduced levels in BV-positive cervicovaginal fluids (Ceccarani et al., 2019; Chetwin et al., 2019). Based on these data we expected both lactate and glutamate to decrease biofilm formation. Our *in*

vitro experiments confirmed this expectation for both *G. vaginalis* and *G. piotti* (**Figure 3**). Conversely, acetate and succinate, which are enriched in BV-positive cervicovaginal fluid, were predicted to enhance biofilm formation. However, our findings contradicted these predictions, revealing decreased biofilm formation in both *Gardnerella* species. These discrepancies underscore the importance of investigating BV at both the bacterial and community levels, considering the polymicrobial nature of the infection and the presence of other non-*Gardnerella* co-occurring species that influence community dynamics. Furthermore, the increased abundance of a metabolite in BV-positive cervicovaginal fluid does not directly imply the ability of *Gardnerella* to utilize that metabolite, as it could be a vaginal byproduct unrelated to the bacteria.

In our enzymatic biofilm disruption experiment, we explored potential BV biofilm-specific non-antibiotic interventions. All five enzymes significantly disrupted *G. vaginalis* biofilms, with proteinase-K and lipase demonstrating the highest efficacy. These findings differ from studies on non-*Gardnerella* biofilms, where carbohydrates are typically the predominant biofilm component. Proteinase-K targets amino acids, while lipase targets lipids, suggesting that the biofilm composition of *G. vaginalis* may differ from that of biofilms in other body sites, such as *Pseudomonas aeruginosa* (Mathur et al., 2018; Ostapska et al., 2018; L. C. Powell et al., 2018). These findings underscore the need to characterize BV-specific biofilms in order to specifically target their unique composition.

Moving forward, the use of genome-scale metabolic network models could provide a robust approach to investigate *Gardnerella* biofilm formation across hundreds of strains.

These analyses can help predict the diversity of biofilm composition and differential biofilm formation capacity. Additionally, *in silico* modeling would enable the creation of a more physiologically accurate vaginal metabolic environment, facilitating the exploration of potential biofilm-specific essential genes. The data presented in this study contribute to the characterization of *Gardnerella* biofilms and can be utilized to further develop biofilm disruptors, aiming to improve BV treatment methods and reduce recurrence rates.

Limitations

Genome scale metabolic network modeling as an analysis approach presents specific limitations as to how we can probe a bacterial community. Specifically, temporal and spatial aspects are critical considerations when modeling bacterial communities, as these factors play important roles in shaping the interactions between individual species and the overall metabolic network of the community. While flux balance analysis (FBA) solutions can offer insight into a steady-state metabolic snapshot of a given metabolic network, dynamic changes in bacterial communities over time need to be accounted for. Dynamic FBA (dFBA) and Dynamic Multi-Species Metabolic Modeling (DMMM) have been developed to incorporate temporal changes by iteratively calculating flux values and metabolite levels and updating the physiological context of the model for the next simulated time point (Mahadevan et al., 2002; Zhuang et al., 2011). DMMM considers the interplay of growth and death rates to account for fluctuating population sizes and changes in community metabolite composition. However, DMMM does not account for the metabolic changes that occur as a bacterium dies and its resources become available to the community.

Incorporating a spatial component in metabolic modeling allows for the integration of metabolic gradients that determine what metabolites individual bacteria can access. COMETS integrates metabolite diffusion with FBA to evaluate how polymicrobial communities metabolically equilibrate. (Harcombe et al., 2014) Conversely, Agent-based models (ABMs) allow for the integration of spatial considerations to investigate how an organism physically interacts with the surrounding environment. MatNet creates an intersection between GEMs and ABMs, allowing for spatial considerations to be accounted for in metabolic network models (Biggs & Papin, 2013). This approach was used to construct a multiscale model of *Pseudomonas aeruginosa* that recapitulated decreased oxygen accessibility in relation to surface location in the context of biofilm formation. BacArena, an algorithm with a larger focus on individual-centric ABM, was able to more accurately predict *Clostridium beijerinckii* doubling time when compared to COMETS (Bauer et al., 2017; Gonze et al., 2018).

In regards to understanding community dynamics within the vaginal microbiome (VMB) specifically, we can leverage previous mechanistic models that have been applied to other polymicrobial communities in human health. ODE models have been utilized to probe dynamic systems and macro-scale ecological models based on (generalized) Lotka–Volterra equations (Faust et al., 2015; Gibson et al., 2016). These models have been applied to polymicrobial communities to analyze the impact of initial microbial abundance and strength of interspecies interactions on steady-state community species abundance. For example, de Vos *et al*. investigated polymicrobial urinary tract infections to understand the impact of antibiotic treatment on community dynamics and observed that closely related

phylogeny have more negative or neutral interactions compared to more disparate species (de Vos et al., 2017). Additionally, the group observed that the community stability decreased with increased taxa diversity, which was supported by the paper of Coyte *et al*. that studied the gut microbiome. Such methods would be helpful in understanding community stability in the context of both healthy and dysbiotic VMB, which can then help to identify community structures that are optimal for intervention (Coyte et al., 2015).

Lastly, for the past century VMB research has lacked a physiologically accurate *in vivo* model capable of recapitulating the human vaginal microbiome while also portraying host-microbiome interactions. The advent of organoid models and the organ-on-a-chip, has reached VMB as of November 2022. The vagina-on-a-chip model born out of Dr. Donald Ingber's tissue engineering lab represents a major breakthrough for VMB research (Mahajan et al., 2022).

References

Armstrong, E., Hemmerling, A., Miller, S., Burke, K. E., Newmann, S. J., Morris, S. R., Reno, H., Huibner, S., Kulikova, M., Nagelkerke, N., Coburn, B., Cohen, C. R., & Kaul, R. (2022). Sustained effect of LACTIN-V (Lactobacillus crispatus CTV-05) on genital immunology following standard bacterial vaginosis treatment: Results from a randomised, placebo-controlled trial. *The Lancet Microbe*, *3*(6), e435–e442. https://doi.org/10.1016/S2666-5247(22)00043-X

Bauer, E., Zimmermann, J., Baldini, F., Thiele, I., & Kaleta, C. (2017). BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities. *PLoS Computational Biology*, *13*(5), e1005544. https://doi.org/10.1371/journal.pcbi.1005544

Biggs, M. B., & Papin, J. A. (2013). Novel multiscale modeling tool applied to Pseudomonas aeruginosa biofilm formation. *PloS One*, *8*(10), e78011. https://doi.org/10.1371/journal.pone.0078011

Bradshaw, C. S., Morton, A. N., Hocking, J., Garland, S. M., Morris, M. B., Moss, L. M., Horvath, L. B., Kuzevska, I., & Fairley, C. K. (2006). High recurrence rates of bacterial vaginosis over the course of 12 months after oral metronidazole therapy and factors associated with recurrence. *Journal of Infectious Diseases*, *193*(11), 1478–1486. https://doi.org/10.1086/503780

Bradshaw, C. S., Pirotta, M., de Guingand, D., Hocking, J. S., Morton, A. N., Garland, S. M., Fehler, G., Morrow, A., Walker, S., Vodstrcil, L. A., & Fairley, C. K. (2012). Efficacy of Oral Metronidazole with Vaginal Clindamycin or Vaginal Probiotic for Bacterial Vaginosis: Randomised Placebo-Controlled Double-Blind Trial. *PLoS ONE*, *7*(4). https://doi.org/10.1371/JOURNAL.PONE.0034540

Bradshaw, C. S., Vodstrcil, L. A., Hocking, J. S., Law, M., Pirotta, M., Garland, S. M., Guingand, D. D., Morton, A. N., & Fairley, C. K. (2013). *Recurrence of Bacterial Vaginosis Is Significantly Associated With Posttreatment Sexual Activities and Hormonal Contraceptive Use*. *56*, 777–786. https://doi.org/10.1093/cid/cis1030

Ceccarani, C., Foschi, C., Parolin, C., D'Antuono, A., Gaspari, V., Consolandi, C., Laghi, L., Camboni, T., Vitali, B., Severgnini, M., & Marangoni, A. (2019). Diversity of vaginal microbiome and metabolome during genital infections. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-50410-x

Chetwin, E., Manhanzva, M. T., Abrahams, A. G., Froissart, R., Gamieldien, H., Jaspan, H., Jaumdally, S. Z., Barnabas, S. L., Dabee, S., Happel, A. U., Bowers, D., Davids, L., Passmore, J. A. S., & Masson, L. (2019). Antimicrobial and inflammatory properties of South African clinical Lactobacillus isolates and vaginal probiotics. *Scientific Reports*, *9*(1), 1–15. https://doi.org/10.1038/s41598-018-38253-4

Cohen, C. R., Wierzbicki, M. R., French, A. L., Morris, S., Newmann, S., Reno, H., Green, L., Miller, S., Powell, J., Parks, T., & Hemmerling, A. (2020). Randomized Trial of Lactin-V to Prevent Recurrence of Bacterial Vaginosis. *New England Journal of Medicine*, *382*(20), 1906–1915. https://doi.org/10.1056/NEJMoa1915254

Colonna, C., & Steelman, M. (2019). Amsel Criteria. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/pubmed/31194459

Coyte, K. Z., Schluter, J., & Foster, K. R. (2015). The ecology of the microbiome: Networks, competition, and stability. *Science (New York, N.Y.)*, *350*(6261), 663–666. https://doi.org/10.1126/science.aad2602

de Vos, M. G. J., Zagorski, M., McNally, A., & Bollenbach, T. (2017). Interaction networks, ecological stability, and collective antibiotic tolerance in polymicrobial infections. *Proceedings of the National Academy of Sciences*, *114*(40), 10666–10671. https://doi.org/10.1073/pnas.1713372114

Eschenbach, D. A., Hillier, S., Critchlow, C., Stevens, C., DeRouen, T., & Holmes, K. K. (1988). Diagnosis and clinical manifestations of bacterial vaginosis. *American Journal of Obstetrics and Gynecology*, *158*(4), 819–828. https://doi.org/10.1016/0002-9378(88)90078-6

Faust, K., Lahti, L., Gonze, D., de Vos, W. M., & Raes, J. (2015). Metagenomics meets time series analysis: Unraveling microbial community dynamics. *Current Opinion in Microbiology*, *25*, 56–66. https://doi.org/10.1016/j.mib.2015.04.004

Gibson, T. E., Bashan, A., Cao, H.-T., Weiss, S. T., & Liu, Y.-Y. (2016). On the Origins and Control of Community Types in the Human Microbiome. *PLOS Computational Biology*, *12*(2), e1004688. https://doi.org/10.1371/journal.pcbi.1004688

Gonze, D., Coyte, K. Z., Lahti, L., & Faust, K. (2018). Microbial communities as dynamical systems. *Current Opinion in Microbiology*, *44*, 41–49. https://doi.org/10.1016/j.mib.2018.07.004

Harcombe, W. R., Riehl, W. J., Dukovski, I., Granger, B. R., Betts, A., Lang, A. H., Bonilla, G., Kar, A., Leiby, N., Mehta, P., Marx, C. J., & Segrè, D. (2014).

Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell Reports*, *7*(4), 1104–1115. https://doi.org/10.1016/j.celrep.2014.03.070

Heczko, P. B., Tomusiak, A., Adamski, P., Jakimiuk, A. J., Stefanski, G., Mikolajczyk-Cichonska, A., Suda-Szczurek, M., & Strus, M. (2015). Supplementation of standard antibiotic therapy with oral probiotics for bacterial vaginosis and aerobic vaginitis: A randomised, double-blind, placebo-controlled trial. *BMC Women's Health*, *15*(1). https://doi.org/10.1186/S12905-015-0246-6

Khan, S., & Hill, J. E. (2021). Established Gardnerella biofilms can survive metronidazole treatment by reducing metabolic activity. *BioRxiv*, 2021.09.06.459156. https://doi.org/10.1101/2021.09.06.459156

Khan, S., Vancuren, S. J., & Hill, J. E. (2021). A Generalist Lifestyle Allows Rare Gardnerella spp. To Persist at Low Levels in the Vaginal Microbiome. *Microbial Ecology*, *82*(4), 1048–1060. https://doi.org/10.1007/S00248-020-01643-1

Lewis, W. G., Robinson, L. S., Gilbert, N. M., Perry, J. C., & Lewis, A. L. (2013). Degradation, Foraging, and Depletion of Mucus Sialoglycans by the Vagina-adapted Actinobacterium Gardnerella vaginalis. *The Journal of Biological Chemistry*, *288*(17), 12067. https://doi.org/10.1074/JBC.M113.453654

Li, C., Wang, T., Li, Y., Zhang, T., Wang, Q., He, J., Wang, L., & Li, L. (2019). Probiotics for the treatment of women with bacterial vaginosis: A systematic review and meta-analysis of randomized clinical trials. *European Journal of Pharmacology*, *864*(April), 172660. https://doi.org/10.1016/j.ejphar.2019.172660

Li, T., Zhang, Z., Wang, F., He, Y., Zong, X., Bai, H., & Liu, Z. (2020). Antimicrobial Susceptibility Testing of Metronidazole and Clindamycin against Gardnerella vaginalis in Planktonic and Biofilm Formation. *Canadian Journal of Infectious Diseases and Medical Microbiology*, *2020*, 1–7. https://doi.org/10.1155/2020/1361825

Limoli, D. H., Jones, C. J., & Wozniak, D. J. (2015). Bacterial Extracellular Polysaccharides in Biofilm Formation and Function. *Microbiology Spectrum*, *3*(3). https://doi.org/10.1128/MICROBIOLSPEC.MB-0011-2014

Lokken, E. M., Jisuvei, C., Oyaro, B., Shafi, J., Nyaigero, M., Kinuthia, J., Mandaliya, K., Jaoko, W., & McClelland, R. S. (2022). Nugent Score, Amsel's Criteria, and a Point-of-Care Rapid Test for Diagnosis of Bacterial Vaginosis: Performance in a Cohort of Kenyan Women. *Sexually Transmitted Diseases*, *49*(1), e22–e25. https://doi.org/10.1097/OLQ.0000000000001469

Machado, A., & Cerca, N. (2015). Influence of biofilm formation by gardnerella vaginalis and other anaerobes on bacterial vaginosis. *Journal of Infectious Diseases*, *212*(12), 1856–1861. https://doi.org/10.1093/infdis/jiv338

Mahadevan, R., Edwards, J. S., & Doyle, F. J. (2002). Dynamic flux balance analysis of diauxic growth in Escherichia coli. *Biophysical Journal*, *83*(3), 1331–1340.

Mahajan, G., Doherty, E., To, T., Sutherland, A., Grant, J., Junaid, A., Gulati, A., LoGrande, N., Izadifar, Z., Timilsina, S. S., Horváth, V., Plebani, R., France, M., Hood-Pishchany, I., Rakoff-Nahoum, S., Kwon, D. S., Goyal, G., Prantil-Baun, R., Ravel, J., & Ingber, D. E. (2022). Vaginal microbiome-host interactions

modeled in a human vagina-on-a-chip. *Microbiome*, *10*(1), 201. https://doi.org/10.1186/s40168-022-01400-1

Marrazzo, J. M., Antonio, M., Agnew, K., & Hillier, S. L. (2009). Distribution of Genital Lactobacillus Strains Shared by Female Sex Partners. *The Journal of Infectious Diseases*, *199*(5), 680–683. https://doi.org/10.1086/596632

Mathur, H., Field, D., Rea, M. C., Cotter, P. D., Hill, C., & Ross, R. P. (2018). Fighting biofilms with lantibiotics and other groups of bacteriocins. *Npj Biofilms and Microbiomes*, *4*(1), 1–13. https://doi.org/10.1038/s41522-018-0053-6

Menard, J.-P., Fenollar, F., Henry, M., Bretelle, F., & Raoult, D. (2008). Molecular quantification of Gardnerella vaginalis and Atopobium vaginae loads to predict bacterial vaginosis. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *47*(1), 33–43. https://doi.org/10.1086/588661

Morrill, S., Gilbert, N. M., & Lewis, A. L. (2020). Gardnerella vaginalis as a Cause of Bacterial Vaginosis: Appraisal of the Evidence From in vivo Models. *Frontiers in Cellular and Infection Microbiology*, *10*, 168. https://doi.org/10.3389/FCIMB.2020.00168

Nasioudis, D., Linhares, I. M., Ledger, W. J., & Witkin, S. S. (2017). Bacterial vaginosis: A critical analysis of current knowledge. *BJOG: An International Journal of Obstetrics and Gynaecology*, *124*(1), 61–69. https://doi.org/10.1111/1471-0528.14209

Nugent, R. P., Krohn, M. A., & Hillier, S. L. (1991). Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation.

Journal of Clinical Microbiology, *29*(2), 297–301. https://doi.org/10.1128/jcm.29.2.297-301.1991

Onderdonk, A. B., Delaney, M. L., & Fichorova, R. N. (2016). The human microbiome during bacterial vaginosis. *Clinical Microbiology Reviews*, *29*(2), 223–238. https://doi.org/10.1128/CMR.00075-15

Ostapska, H., Howell, P. L., & Sheppard, D. C. (2018). Deacetylated microbial biofilm exopolysaccharides: It pays to be positive. *PLOS Pathogens*, *14*(12), e1007411. https://doi.org/10.1371/JOURNAL.PPAT.1007411

Powell, L. C., Pritchard, M. F., Ferguson, E. L., Powell, K. A., Patel, S. U., Rye, P. D., Sakellakou, S. M., Buurma, N. J., Brilliant, C. D., Copping, J. M., Menzies, G. E., Lewis, P. D., Hill, K. E., & Thomas, D. W. (2018). Targeted disruption of the extracellular polymeric network of Pseudomonas aeruginosa biofilms by alginate oligosaccharides. *Npj Biofilms and Microbiomes*, *4*(1), 13. https://doi.org/10.1038/s41522-018-0056-3

Ravel, J., Gajer, P., Abdo, Z., Schneider, G. M., Koenig, S. S. K., McCulle, S. L., Karlebach, S., Gorle, R., Russell, J., Tacket, C. O., Brotman, R. M., Davis, C. C., Ault, K., Peralta, L., & Forney, L. J. (2011). Vaginal microbiome of reproductive-age women. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(SUPPL. 1), 4680–4687. https://doi.org/10.1073/pnas.1002611107

Schwebke, J. R., Muzny, C. A., & Josey, W. E. (2014). Role of Gardnerella vaginalis in the pathogenesis of bacterial vaginosis: A conceptual model. *Journal of Infectious Diseases*, *210*(3), 338–343. https://doi.org/10.1093/infdis/jiu089

Shipitsyna, E., Roos, A., Datcu, R., Hallén, A., Fredlund, H., Jensen, J. S., Engstrand, L., & Unemo, M. (2013). Composition of the vaginal microbiota in women of reproductive age—Sensitive and specific molecular diagnosis of bacterial vaginosis is possible? *PloS One*, *8*(4), e60670. https://doi.org/10.1371/journal.pone.0060670

Srinivasan, S., Liu, C., Mitchell, C. M., Fiedler, T. L., Thomas, K. K., Agnew, K. J., Marrazzo, J. M., & Fredricks, D. N. (2010). Temporal variability of human vaginal bacteria and relationship with bacterial vaginosis. *PloS One*, *5*(4), e10197. https://doi.org/10.1371/journal.pone.0010197

Srinivasan, S., Morgan, M. T., Fiedler, T. L., Djukovic, D., Hoffman, N. G., Raftery, D., Marrazzo, J. M., & Fredricks, D. N. (2015). Metabolic signatures of bacterial vaginosis. *MBio*, *6*(2), 1–16. https://doi.org/10.1128/mBio.00204-15

Swidsinski, A., Mendling, W., Loening-Baucke, V., Swidsinski, S., Dörffel, Y., Scholze, J., Lochs, H., & Verstraelen, H. (2008). An adherent Gardnerella vaginalis biofilm persists on the vaginal epithelium after standard therapy with oral metronidazole. *American Journal of Obstetrics and Gynecology*, *198*(1), 97.e1-97.e6. https://doi.org/10.1016/j.ajog.2007.06.039

Yeoman, C. J., Yildirim, S., Thomas, S. M., Durkin, A. S., Torralba, M., Sutton, G., Buhay, C. J., Ding, Y., Dugan-Rocha, S. P., Muzny, D. M., Qin, X., Gibbs, R. A., Leigh, S. R., Stumpf, R., White, B. A., Highlander, S. K., Nelson, K. E., & Wilson, B. A. (2010). Comparative genomics of Gardnerella vaginalis strains reveals substantial differences in metabolic and virulence potential. *PLoS ONE*, *5*(8). https://doi.org/10.1371/journal.pone.0012411

Zhuang, K., Izallalen, M., Mouser, P., Richter, H., Risso, C., Mahadevan, R., & Lovley, D. R. (2011). Genome-scale dynamic modeling of the competition between Rhodoferax and Geobacter in anoxic subsurface environments. *The ISME Journal*, *5*(2), 305–316. https://doi.org/10.1038/ismej.2010.117

Chapter 5: Future Directions

Analysis Next Steps

Caffeate: Role as a host-microbiome intermediary

Caffeate has been identified as a binder of estrogen receptor alpha and has thus far only been studied in relation to breast cancer. Our metabolomics analysis has revealed the ability of *F. vaginae* and *A. christensii* to produce caffeate. These findings suggest that caffeate could potentially act as a mediator between the host and the microbiome. Moving forward, we have two main objectives: 1) To determine whether the production of caffeate by *F. vaginae* and *A. christensii* is dependent on BV environment or if it is produced ubiquitously; 2) To analyze whether exposure to caffeate affects the transcription of estrogen-sensitive genes in vaginal epithelial cells.

To address our first objective, we submitted F. vaginae and A. christensii spent media, cultured in either enriched NYCIII or spent G. piotti media, for analysis via high-performance liquid chromatography (HPLC). This analysis will enable us to directly quantify caffeate levels. Moreover, it will help us determine whether these microbes consistently produce caffeate or if production is induced only when they are exposed to byproducts from BV-associated microbes through *G. piotti* spent media.

In addition, we will expose vaginal epithelial cells (specifically, the VK2 cell line) to caffeate, using estradiol as a positive control. By employing quantitative polymerase chain reaction (qPCR), we will measure the transcription levels of three estrogen-sensitive genes:

CCND1, ESR1, and TFF1. These findings will provide insights into the effects of caffeate on the vaginal epithelium of the host.

Identifying conserved BV metabolic units

Identifying functional metabolic units that are conserved in varying BV communities is a crucial step in unraveling the complex dynamics of this condition and understanding patient outcomes. To achieve this, we could leverage the metagenomic repository Evvy has cultivated to study the genetic profiles of different BV-associated bacterial communities. By leverging these genetic data to determine the metabolic pathways present in these communities, we can identify common functional units that are conserved, regardless of the specific bacterial species or strains involved. This approach allows for the identification of key metabolic processes and interactions that contribute to the dysbiosis observed in BV, providing valuable insights into the underlying mechanisms and potential targets for intervention or therapeutic strategies.

Understanding conserved functional metabolic units in varying BV communities is important for several reasons. Firstly, conserved metabolic units can be used to develop diagnostic tools that can more accurately identify BV based on specific metabolic markers, compared to the standard Nugent and/or Amsel score This can lead to earlier detection and intervention, potentially preventing the progression of BV-related complications. Secondly, by identifying these functional units, we can gain a deeper understanding of the microbial interactions and metabolic pathways that contribute to BV pathogenesis. This knowledge can inform the development of targeted therapies aimed at disrupting these conserved dysbiotic metabolic units, and catalyze restoring a healthy vaginal microbiome.

Translational application

Defining BV Sub-categories & Improving Diagnosis

There are two commonly used diagnostic methods for bacterial vaginosis (BV): the Nugent score, which is a gram stain-based method; and the Amsel score, which is a symptoms-based assessment (Eschenbach et al., 1988). The Nugent score involves assigning a numerical value based on the types of bacteria observed after staining, with higher scores indicating the absence of *Lactobacillus* species (Nugent et al., 1991). However, this scoring system may introduce bias for women of color, as their vaginal microflora naturally tends to be more diverse and less dominated by *Lactobacillus* species (Lokken et al., 2022). On the other hand, the Amsel score relies on assessing three of four specific symptoms to suggest a diagnosis of BV: thin and gray vaginal discharge, vaginal pH higher than 4.5, positive presence of Clue cells (as seen under a microscope), and a positive "Whiff" test where vaginal discharge produces a fishy odor when exposed to hydrogen peroxide (Colonna & Steelman, 2019).

Additional research has been conducted to develop molecular assays for BV diagnosis, building upon the consistent prevalence of *Gardnerella* species as a defining pathogenic feature of this condition (Menard et al., 2008; Schwebke et al., 2014; Shipitsyna et al., 2013). However, the shift towards quantitative diagnostics does not account for the diversity within the diagnosis itself. BV can present as symptomatic or asymptomatic, and there is significant variation in treatment response among individuals. Moreover, the *Gardnerella* genus, which dominates this polymicrobial condition, exhibits a wide range

of metabolic diversity even within species (Khan et al., 2021; Yeoman et al., 2010). Treating all BV cases with a standardized diagnostic and treatment regimen perpetuates high rates of recurrence.

In light of these factors, there is a need to define subtypes of BV based on differential community structures and vaginal environments. By identifying and classifying specific subtypes of BV dysbiosis, we can develop biomarker diagnostics that accurately distinguish the different subtypes, leading to more effective treatment strategies. To start this work, we need to begin recapitulating metagenomic species abundance into metabolic modeling based bacterial communities, in parallel with ordinary differential equation (ODE) based modeling. Using these analyses, we can begin to define community state type categories within BV, and predict differentiating metabolic features of these community groups. These data can be used to develop biomarker diagnostics, which can then be used to investigate treatment response differences between groups. The first step to improving BV treatment is improving the accuracy of BV diagnosis and generating BV subtype definitions works towards this end goal.

Developing Therapeutic Bacterial Consortia

Broad-spectrum antibiotics are the standard treatment for BV) but there is growing interest in developing therapeutic consortia to complement antibiotic regimens. Similar to the gut, antibiotics eradicate the existing vaginal microbiome to eliminate colonized pathogens. However, antibiotics also eliminate "good" bacteria, creating vacant metabolic niches that can be colonized by pathogenic organisms. Therapeutic consortia aim to facilitate the reestablishment of a healthy bacterial community after antibiotic clearance. Current

probiotics have primarily focused on recolonizing the vagina with *Lactobacillus* species. A recent clinical trial of the *Lactobacillus* CTV-05 strain, LACTIN-V, demonstrated reduced BV recurrence when used alongside oral metronidazole antibiotics (Armstrong et al., 2022; Cohen et al., 2020). This exciting breakthrough demonstrates the potential impact of scientifically developed probiotics on BV treatment.

Instead of solely focusing on developing probiotic consortia capable of long-term colonization, an alternative approach could involve probiotic consortia that: 1) occupy metabolic niches to prevent pathogenic recolonization and 2) create a metabolically favorable environment for the restoration of natural vaginal flora post-antibiotic treatment.

This proposed approach leverages metabolic competition and mutualism to inhibit undesirable bacterial growth and promote beneficial bacterial growth. Using competition and mutualism algorithms, as outlined in Chapter 4, we can predict a probiotic consortium that: 1) competes directly with BV-associated bacteria metabolically and 2) produces metabolites that specifically benefit healthy vaginal flora. To identify suitable bacteria, we would employ a BLAST search on generally regarded as safe bacteria, targeting candidate species predicted to consume highly competed-for BV metabolites and that produce metabolites associated with vaginal health based on *in silico* predictions. Once these bacteria are identified, we could investigate host-bacteria interactions using the novel vagina-on-a-chip model. This would allow us to assess host immune response and quantify inflammatory markers. Lastly, we could employ the vagina-on-a-chip BV *in vivo* substitute model to characterize host clearance with different treatment approaches, including

antibiotic-only treatment, antibiotic with probiotic treatment, probiotic-only treatment, and no treatment.

Personalizing BV Treatment Regimens

By expanding the definition of BV to include defined subtypes, developing diagnostic methods to distinguish between these subtypes, and creating new probiotic consortia to reduce rates of BV recurrence more inclusively, we can then progress towards personalized BV treatment regimens. In order to achieve this objective, it is crucial to investigate the impact of race and ethnicity on differential treatment outcomes. This requires the development of treatment response models that incorporate ethnic background and BV community state type as variables, enabling the disentanglement of their respective contributions to treatment outcomes. By developing these models, we can leverage patient metadata along with cervicovaginal fluid metagenomic profiles to determine the optimal treatment approach for each individual patient. Furthermore, as our understanding of the various manifestations of BV improves, there is potential to repurpose existing antibiotics for BV treatment. These advancements will contribute to closing the gender and ethnic health gaps through personalized medicine approaches.

Conclusion

In conclusion, the significant strides I have made through my work have helped pave the way for personalized medicine in women's health. By addressing the unique needs and challenges faced by women, I am proud to have played a part in narrowing the gender health gap. Moving forward, I am committed to continuing my efforts and contributing

further to the betterment of women's health, ensuring that every individual receives the personalized care they deserve.

Supplemental: Understanding COVID-19 Disease Severity - Insights from Metabolic Modeling

The text for this chapter has been previously published as a research article here:

Dillard LR, Wase N, Ramakrishnan G, Park JJ, Sherman NE, Carpenter R, Young M, Donlan AN, Petri W, Papin JA (July 2022) Leveraging Metabolic Modeling and Machine Learning to Identify Metabolic Alterations Associated with COVID-19 Disease Severity. *Metabolomics.* https://doi.org/10.1007/s11306-022-01904-9

Context

Joining Dr. Papin's lab in the spring of 2020 presented challenges. Jumping into my doctorate in the midst of a pandemic was a struggle, with labs closed and unable to generate experimental data. I was fortunate enough to collaborate with Dr. Bill Petri on his work investigating immunological responses to COVID-19, and was able to analyze their rich metabolomics dataset. Through my partnership with Dr. Petri's lab, I was able to provide a computational analysis of the extensive serum metabolomics collected from a cohort of UVA COVID-19 patients. This chapter presents my first graduate level work and initial dive into the intersection of metabolomics and metabolic modeling. The skills I developed through this initial project have provided me with an essential foundation for my final dissertation chapter which again integrates metabolomics and metabolic modeling.

Synopsis

Computational metabolic models provide insight into the functional metabolism of a single cell. Through identifying differentially present metabolites in severe versus non-severe COVID-19 we were able to infer systems level metabolic differences in the two disease states. Additionally, by integrating extracellular metabolomics into whole-body metabolic models, we were able to make predictions regarding intracellular functional metabolism across disease severity. These findings have therapeutic applications for mitigating COVID-19 disease progression.

Introduction

The COVID-19 pandemic, caused by SARS-CoV2, has resulted in a significant global health crisis with millions of lives lost and millions more infected (World Health Organization, 2021). While symptoms range from mild to severe, COVID-19 can lead to long-term complications and organ damage (CDC, n.d.). Understanding the host immune response to SARS-CoV2 infection is crucial for developing effective treatments (Walls et al., 2016, 2020; Watanabe et al., 2020). Metabolism and the immune system are closely linked, and identifying metabolic differences between patients could provide insights into immune response variations (Alwarawrah et al., 2018). Previous research has focused on the metabolic byproducts of infection due to changes in host metabolism related to disease severity (Hasan et al., 2021; Páez-Franco et al., 2021; Sindelar et al., 2021; Stukalov et al., 2021; Thomas et al., 2020). Our study aims to identify the functional metabolic roots of alterations associated with COVID-19 disease severity.

To achieve this, we analyzed metabolic signatures of non-acute (patients not requiring hospitalization) and severe COVID-19 patient plasma metabolomes and built associated genome-scale metabolic models to represent these two disease states. Our objective was to identify protective metabolic pathways in non-acute COVID-19 patients and potential pathways whose inhibition could mitigate disease progression in severe COVID-19 patients. We used supervised and unsupervised machine learning approaches to identify key metabolic drivers that predict COVID-19 disease severity and explored potential metabolic mechanisms that link these markers to disease progression. We also analyzed the levels of Interleukin 13 (IL-13), a profibrotic cytokine that has been identified as a driver of COVID-19 severity, relationship to differential metabolites and disease severity (Donlan et al., 2021; Mohning et al., 2019). Our findings characterize the metabolic transition from an innate to adaptive immune response, the signatures of inhibited inflammatory pathways in non-acute COVID-19, and the metabolic byproducts of severe COVID-19 symptoms. This information can be leveraged to develop novel preventative and therapeutic strategies to reduce the burden of the disease.

Methods

Patients

Blood samples were obtained from a cohort of 84 adult patients who tested positive for SARS-CoV2 via PCR at the University of Virginia hospital between April and June 2020. The collected samples were processed to obtain plasma, which was subsequently stored at a temperature of −70°C. 48 out-patient samples were collected and categorized as non-acute COVID-19, while the remaining 36 samples were obtained from patients who

required hospitalization and were classified as severe COVID-19, with some requiring ICU and ventilator support (4 and 25, respectively). All collected samples and subsequent data analysis had previous IRB approval.

Patient plasma preparation

The clinical data for the study participants were retrieved from electronic medical records, and each participant was given a unique identifier to ensure anonymity. The metabolite profiling experiments were conducted at the Biomolecular Analysis Facility in the University of Virginia School of Medicine. In brief, plasma samples were thawed on ice, and 50 µL of plasma was retained for the metabolome analysis. To inactivate any potential viruses, 200 µL of −20 °C methanol was added to the plasma sample and shaken vigorously. The samples were immediately stored at −80 °C until extraction for metabolomics experiments. For extraction, 200 µL of −20 °C methanol was added to each tube, vortexed, and shaken vigorously for 30 min at 4 °C in a temperature-controlled thermal shaker. Next, 200 µL of chloroform and 400 µL of water were added, shaken vigorously, and the top aqueous phase was recovered as a metabolite mixture of diverse chemical nature. Two aliquots of 350 µL each were created, one for the experiment and one as a backup for additional experiments. Each metabolite extract was dried overnight in a speedVac and reconstituted in 60 µL of 0.1% formic acid in water. Prior to sample analysis, the instrument was calibrated using Pierce FlexMix solution. To ensure stable background signal, the LC-MS system was stabilized by running 3–4 wash runs followed by 4 blank runs. A commercial amino acid mixture was used to assess the chromatography quality before running the actual experiment (*Amino Acid Mixtures*, n.d.). To evaluate the chromatography quality, a pooled QC sample was created by removing 10 µL from each

tube and was injected at the beginning and end of the MS sequence run. Additionally, QC samples were injected after every 10 sample injections.

Mass spectrometry preparation and analysis

The AcquireX Intelligent Data Acquisition Workflow was used to generate an exclusion list from a blank run and an inclusion list from an injection of the pooled QC sample, followed by iterative DDA injections. Samples were analyzed using an LC-HRMS system with positive and negative ionization modes, and spectra were acquired over the m/z range 67-1000 in full MS mode. Compounds were detected and processed using Compound Discoverer 3.1, with precursor selection based on MS(n-1) and compounds detected at a mass tolerance of 5 ppm and minimum peak intensity threshold of 500,000. Compound annotations were performed using ddMS2 masses in mzCloud, and statistical differential metabolite discovery was performed using healthy patients as a reference group with a p-value threshold of 0.05 and log2FC threshold of 1 (Sumner et al., 2007). Overall, 680 distinct metabolites were measured in each sample.

Biomarker identification

The data were first processed using a median metabolite value filter, followed by log transformation and autoscaling, to normalize within and between samples. The Mann Whitney U test, a non-parametric statistical test, was used to identify metabolites that were significantly enriched in either the non-acute or severe COVID-19 disease state (Nachar, 2008). To mitigate the false discovery rate, the Benjamini–Hochberg correction was used as the multiple test correction method (Benjamini et al., 1995). Additionally, 124 well-documented pharmaceutical or non-endogenous metabolites were removed to isolate non-

treatment-based metabolic shifts. To account for the disbalance in groups, bootstrapping was employed.

To assess the accuracy of the identified endogenous metabolites in discerning between severe and non-acute COVID-19 status, the supervised learning approach of random forest was utilized (Breiman, 2001). The holdout validation method was applied using 30% holdout, and model performance was also evaluated using receiver operating characteristic curve (ROC) and recall calculation.

Furthermore, the unsupervised learning approach of Nonmetric Multidimensional Scaling (NMDS) and associated Euclidean distance were used to analyze the previously identified endogenous metabolites for significant COVID-19 status clustering. A Permutational Multivariate Analysis of Variance (PERMANOVA) was used to determine whether clustering of the severe and non-acute groups was statistically significant (Anderson, 2017).

Lastly, to investigate the correlation between IL-13 levels and COVID-19 disease severity, the endogenous differential metabolomics data were subset to include the top 25% of patients with the highest IL-13 levels, as well as the bottom 25% patients with the lowest IL-13 levels. An unpaired, one-tailed t-test was run on the selected metabolites to assess significant differences based on IL-13 levels. Bonferroni multiple test correction was applied to all t-test results.

Pathway Analysis

Pathway analysis was performed to identify metabolic shifts that are significantly associated with COVID-19 disease state and uncover novel metabolic mechanisms underlying COVID-19 disease progression. This analysis was conducted using the metabolic pathway module of MetaboAnalyst version 5.0 toolbox on the previously selected endogenous metabolites that were enriched in either the severe or non-acute COVID-19 disease state (Xia et al., 2009). The MetaboAnalyst platform utilizes the Kyoto Encyclopedia of Genes and Genomes (KEGG) to run over-representation analysis (ORA) to identify enriched metabolic pathways in one condition compared to a reference (Kanehisa & Goto, 2000). The ORA is based on a calculated enrichment ratio, which compares the actual versus predicted metabolite hits. The significance of these calculated enrichment ratios was determined using a p-value derived from a hypergeometric test with a binomial distribution.

Genome-scale metabolic modeling

We employed a systems-level approach to investigate the functional metabolic mechanisms underlying COVID-19 disease progression. To construct contextualized genome-scale metabolic models (GEMs) of severe and non-acute COVID-19 disease states, we utilized the previously published human whole-body GEM Recon3D as a base model (Brunk et al., 2018; Edwards et al., 2002). Prior to model pruning, we manually matched the identified differential endogenous metabolites to respective model metabolites, and adjusted exchange bounds of associated metabolites to simulate open metabolic exchange for the respective disease states.

After pruning the models using reaction inclusion by parsimony and transcript distribution (RIPTiDe), we used flux balance analysis (FBA) in tandem with GAPSPLIT to generate 500 flux samples from each model to explore potential solution spaces (Keaty et al., 2020). We compared conserved reaction flux values across the models using NMDS followed by a PERMANOVA to determine if the models' flux values were significantly separated. We then identified the top conserved reactions with predictive flux values capable of differentiating the severe versus non-acute models using the supervised learning approach random forest.

Results

Biomarker Identification

The results of the analysis revealed significant differences in the metabolic profiles of severe vs. non-acute COVID-19 samples (**Figure 1**). Specifically, 226 metabolites were identified as significantly different in the severe COVID-19 cases compared to the non-acute cases (Mann Whitney U; $p<0.05$). Of these, 80 metabolites were found to be significantly elevated in non-acute COVID-19 samples and 21 metabolites were significantly elevated in severe COVID-19 samples, after excluding non-endogenous metabolites.

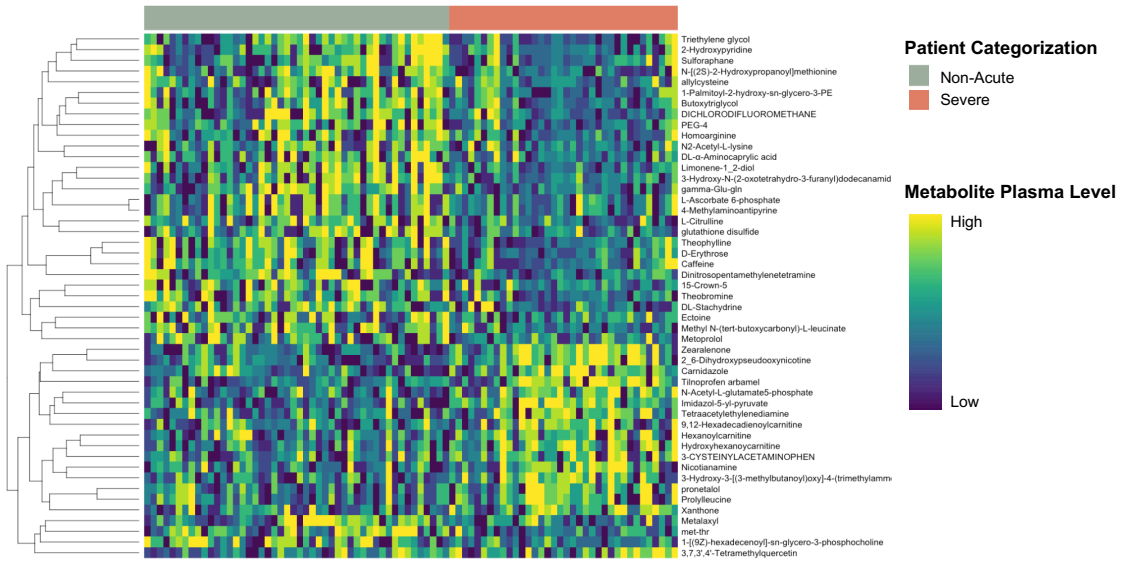


Figure 1: Heatmap of top 50 endogenous and non-endogenous differential metabolites for non-acute and severe COVID-19 patient plasma sample data.

Random forest analysis was conducted using the remaining 101 metabolites, resulting in a model capable of predicting COVID-19 disease severity with an out-of-bag error rate of 7.14%. Notably, the removal of non-endogenous metabolites improved the model classifier's accuracy, as indicated by a reduced out-of-bag error rate of 14.19% when all 680 metabolites were included.

Furthermore, NMDS and associated PERMANOVA revealed a significant grouping ($r^2 = 0.09$, p-value < 0.001) of non-acute COVID-19 samples vs. severe COVID-19 samples based on measured endogenous metabolite levels (**Figure 2; Table 1**).

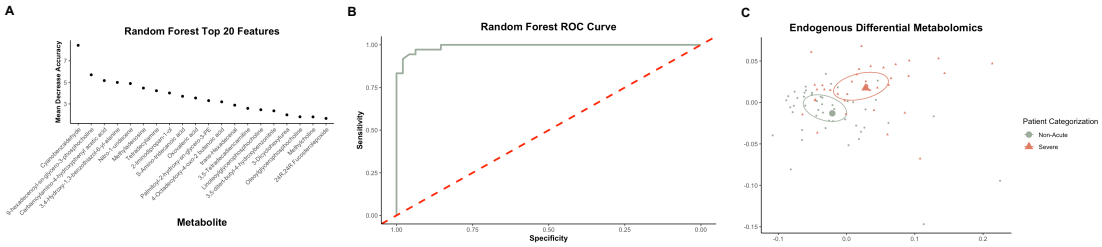


Figure 2: **A)** Top 20 endogenous differential features identified by random forest as most important for predicting COVID-19 disease severity. **B)** Receiver operating curve for random forest generated COVID-19 status model predictability **C)** Non-metric Multi-dimensional Scaling (NMDS) based on all endogenous metabolites identified as significantly different between patient categories. (PERMANOVA: $R^2 = 0.09$, p-value < 0.001)

To investigate the relationship between IL-13 levels, COVID-19 disease severity, and functional metabolism, we integrated IL-13 level data with the previously outlined metabolomics analysis. We first subset the metabolomics data based on the top 25% and bottom 25% IL-13 levels (average IL-13 pg/ml of 36.32 and 5.26, respectively). An unpaired one-tailed t-test analysis was then performed, which revealed 14 metabolites with significantly different levels between the high versus low IL-13 groups, from the previously identified 101 differential metabolites.

Non-Acute Patient Data

	Total	Average Age	Average BMI	Max Oxygen	Ventilator	Diabetes	Kidney Disease	Heart Disease	Lung Disease	Immunosuppression	Cancer	Remdesivir Treatment
Female	**27**	**56** (15-94)	**33** (22-55)	**10**		**11**	**5**	**7**	**1**	**2**	**2**	**5**
Black	6	72	32	4		4	2	4				
Asian	1	24		0								
Other	10	47	32	10		3	3	1		2	1	
White	10	58	35			4		2	1		1	5
				8								
Male	**21**	**56** (33-84)	**30** (14-62)	**10**		**8**	**5**	**3**	**3**	**1**	**3**	**3**
Black	6	64	30	10		3	1	2	2	1	1	1
Other	11	47	30	5		3	4	1			1	2
White	4	65	24	2		2			1		1	

Severe Patient Data

	Total	Average Age	Average BMI	Max Oxygen	Ventilator	Diabetes	Kidney Disease	Heart Disease	Lung Disease	Immunosuppression	Cancer	Remdesivir Treatment
Female	**10**	**34** (21-55)	**57** (28-75)	**0**	**10**	**5**	**1**	**1**		**1**	**1**	**3**
Black	1	55	60	0	1	1						
Asian	1	33	65	0	1							1
Other	5	35	49	0	5	2					1	1
White	3	27	68	0	3	2	1	1		1		1
Male	**26**	**30** (20-44)	**61** (14-83)	**15**	**20**	**12**	**4**	**4**	**3**	**2**	**1**	**4**
Black	4	35	71	3	3	2	1		1		1	
Other	10	28	46	15	8	5						2
White	12	30	70	4	9	5	3	4	2	2		2

Table 1: Sample metadata outlining patient medical background and COVID-19 related treatment

Two of the 21 metabolites that were significantly elevated in severe COVID-19 samples were found to be significantly different when grouped by IL-13 levels, and both were found to be elevated in IL-13-high patients (**Table 2**). Among the 80 metabolites that were identified as significantly elevated in non-acute COVID-19 samples, 11 metabolite levels were significantly altered when comparing high vs. low IL-13 levels. Notably, all of these

metabolites were found to be significantly higher in low IL-13 level patients, except for L-homocysteic acid, which was reduced in low IL-13 level patients (**Table 2**).

COVID-19 Status	IL-13 Level	Enriched Metabolite
Severe	High	4-imidazolone-5-propanoate
		3-methylglutarylcarnitine
Non-Acute	High	L-homocysteic acid
	Low	(24R,24'R)-fucosterol epoxide
		alanyl-poly(glycerolphosphate)
		Erucamide
		L-Ascorbate 6-phosphate
		N2-acetyl-L-lysine
		11-Nitro-1-undecane
		β –leucine
		n-ribosylhistidine
		Trp-Phe
		trimethylsilyl N,O-bis(trimethylsilyl)serinate

Table 2: Metabolites significantly different (p-value < 0.05) when comparing high vs. low IL-13 patient metabolomics

Pathway Analysis

We conducted pathway analysis to gain insight into the metabolic pathways associated with non-acute and severe COVID-19. Among the list of 80 differentially enriched metabolites associated with non-acute COVID-19 samples, we observed significant associations with tryptophan metabolism (FDR = 0.07), including higher levels of L-tryptophan, melatonin, 5-hydroxy-L-tryptophan, 3-hydroxyanthranilic acid, indoleacetaldehyde, and anthranilate (**Figure 3A**).

Moreover, among the list of 20 differentially enriched metabolites associated with severe COVID-19, we observed a significant association with histidine metabolism (FDR = 0.28), where 4-imidazolone-5-propionic acid, imidazole-pyruvate, and methylimidazoleacetic acid were identified (**Figure 3B**). We also identified several additional metabolites of interest from the list of differential endogenous metabolites. Specifically, we observed significantly higher levels of three acylcarnitines (hexanoylcarnitine, 3-methylglutarylcarnitine, and 9,12-hexadecadienoylcarnitine), L-gamma-glutamyl-L-leucine, and D-galactonate in severe COVID-19 samples compared to non-acute samples (**Figure 3C**). In contrast, 5'-methylthioadenosine was identified as significantly higher in non-acute COVID-19 samples compared to severe cases.

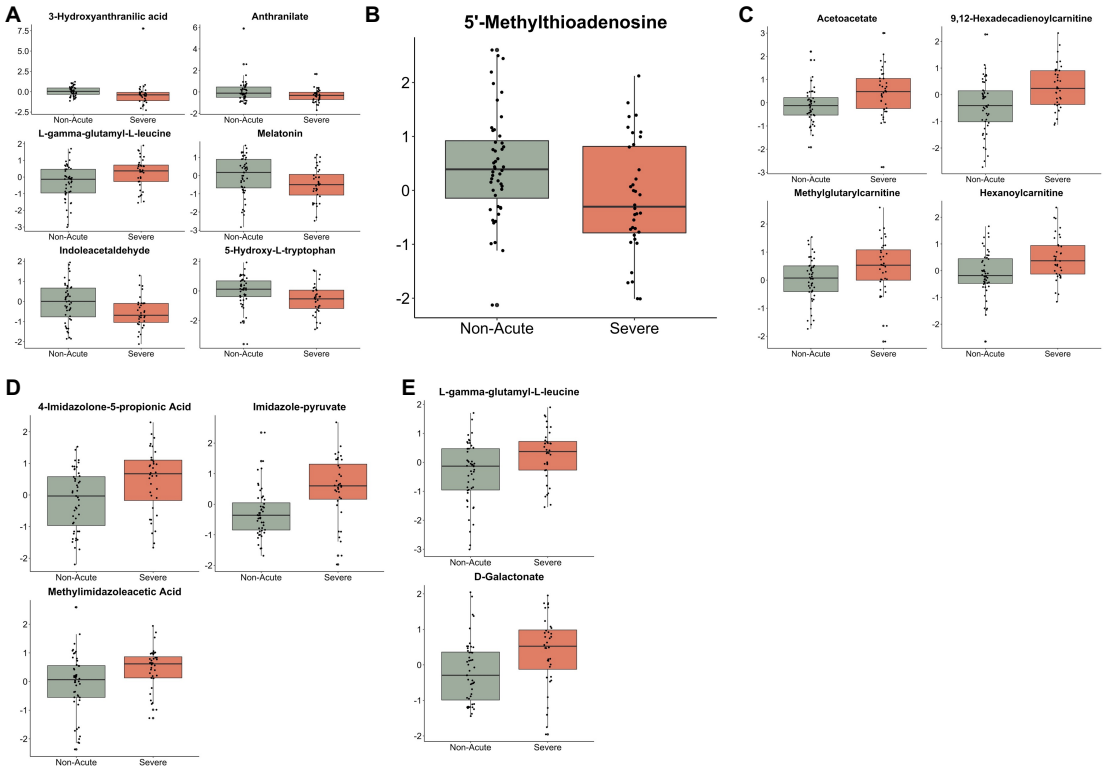


Figure 3: **A)** Patient plasma levels of metabolites involved in tryptophan metabolism and melatonin synthesis (FDR = 0.07) **B)** Patient plasm a levels of the metabolite 5'-methylthioadenosine **C)** Patient plasma levels of metabolites involved in ketone body biosynthesis **D)** Patient plasma levels of metabolites involved in histidine degradation (FDR = 0.28) **E)** Patient plasma levels of metabolites associated with shift in energy source

Genome scale metabolic modeling

The contextualized models revealed differential metabolic pathways between severe and non-acute COVID-19 states. In the severe COVID-19 model, reactions related to glycolysis, gluconeogenesis, and the pentose phosphate pathway were upregulated. Additionally, reactions related to the biosynthesis of purines, pyrimidines, and amino acids were upregulated in the severe model. In contrast, the non-acute COVID-19 model showed downregulation of reactions related to amino acid biosynthesis, fatty acid metabolism, and the pentose phosphate pathway. The non-acute model also showed upregulation of reactions related to the tricarboxylic acid cycle and oxidative phosphorylation.

Discussion

Intra- and extra- cellular metabolism are an essential factor in COVID-19 pathology. While prior studies have concentrated on using plasma metabolomics to determine the metabolic impact of COVID-19 on patients and to examine disease progression, our study employs plasma metabolomics data and a genome-scale metabolic model to more comprehensively investigate functional metabolic shifts associated with disease progression. Through metabolomic analysis and computational metabolic modeling, we observed that tryptophan, glutathione, pyrimidine, D-glutamine, and D-glutamate metabolism are significantly enriched in non-acute COVID-19 patients (**Figure 4A**). These metabolic alterations may have physiological implications such as mitigating inflammatory responses, decreasing fibroblast accumulation, increasing vitamin-D levels, and preventing cellular damage. In contrast, severe COVID-19 patients exhibit a significant enrichment in histidine metabolism, which has been previously shown to correlate with increased viral titer and elevated fatty acid oxidation (**Figure 4B**) (Doğan et al., 2021; Kimhofer et al., 2020; López-Hernández et al., 2021; Overmyer et al., 2021; Roberts et al., 2022; Thomas et al., 2020) .

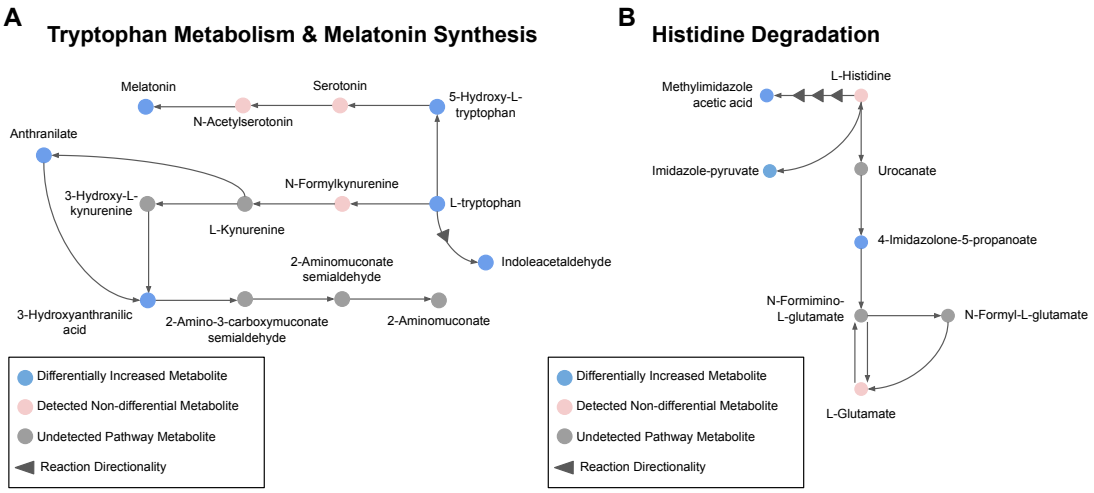


Figure 4: Metabolic pathway p-values are assigned based on how significantly the identified metabolites indicate pathway enrichment. **A)** Metabolic pathways associated with non-severe COVID-19 metabolite predictors. Tryptophan metabolism and melatonin synthesis (FDR = 0.07). **B)** Metabolic pathways associated with severe COVID-19 metabolite predictors. Histidine degradation (FD R= 0.28)

Non-Acute COVID-19 Metabolism

We found that in non-acute COVID-19 patients, tryptophan metabolism is significantly enriched, including the differential enrichment of L-tryptophan, melatonin, 5-hydroxy-L-tryptophan, 3-hydroxyanthranilic acid, indoleacetaldehyde, and anthranilate, as well as Trp-Phe. Tryptophan plays a role in neurotransmitter synthesis, anti-inflammatory pathways, and improved immune response via the gut microbiome (Gao et al., 2018; Krause et al., 2011; L. Wang et al., 2002). Increased tryptophan metabolism in non-acute COVID-19 patients presents a potential inhibitory mechanism for respiratory inflammation, a symptom associated with the cytokine cascade involved in SARS-Cov-2 infection (Dehhaghi et al., 2019).

Differential enrichment of 5′-methylthioadenosine (MTA) in non-acute COVID-19 samples could be a direct result of the body mitigating innate immune system inflammatory

response. Mechanistically, MTA inhibits the pro-inflammatory cytokine TNF-α and increases production of anti-inflammatory IL-10 (Hevia et al., 2004; Veal et al., 2004).

Additionally, L-homocysteic acid is elevated in both non-acute COVID-19 samples and IL-13-high samples. L-homocysteic acid is an endogenous neurotransmitter ligand of N-Methyl-D-aspartic acid, commonly known as NMDA, which plays an integral role in neural plasticity (Jewett & Thapa, 2020; Marshall, 2020; Yirmiya & Goshen, 2011). The high levels of L-homocysteic acid in non-acute COVID-19 and high IL-13 levels suggest a potential mechanism of immune-mediated neural adaptation to disease pathology to avoid neurological symptoms that are sometimes reported in those infected with SARS-Cov2.

Moreover, after investigating functional metabolic changes using a contextualized non-acute COVID-19 metabolic model, three reactions stood out as carrying significantly higher flux compared to the severe COVID-19 metabolic model: nucleoside-diphosphate kinase (DNPK), 5'-nucleotidase (IMP), and adenosine facilitated transport into the cytosol. DNPK plays an integral role in maintaining genomic stability and providing a protective effect in the case of cancer metastasis, a disease that heavily relies on cellular replication (Lacombe et al., 2021). 5'nucleotidase catalyzes cytosolic purine degradation (Ipata & Tozzi, 2006; Pesi et al., 2021). Increased de novo purine synthesis has been linked to an increased inflammatory response in COVID-19 patients (Pieters & Veerman, 1988). 5'nucleotidase activity could be reducing cytosolic purine levels and thus reducing the cytokine response to infection (Xiao et al., 2021; Zhang et al., 2021). Lastly, increased adenosine facilitated transport suggests that those who experience non-acute infection are

removing excess adenosine and, therefore, reducing the subsequent inflammatory response to infection (Singh Patidar et al., 2018).

Severe COVID-19 Metabolism

The results of this study reveal that severe COVID-19 samples exhibit distinct metabolic alterations, including higher levels of imidazole-pyruvate, 4-Imidazolone-5-propanoate, and methylimidazoleacetic acid. The aforementioned metabolites are associated with histidine metabolism, specifically histidine degradation. Histidine is known to play a crucial role in enzyme activation, including the activation of serine proteases which have been shown to facilitate viral entry into host cells and subsequent viral spread (Delattre et al., 2021; Matsuyama et al., 2010; Radisky et al., 2006; Seth et al., 2020). These findings suggest that SARS-CoV2 may exploit the host's metabolism as a means of viral replication and spread.

Furthermore, severe COVID-19 samples were found to have significantly higher levels of acetoacetate, which is associated with ketone body biosynthesis. This increase may be a result of increased fatty acid oxidation, which is also indicated by high levels of acylcarnitines in severe samples (Yan et al., 2019). The observed changes in fatty acid metabolism in severe COVID-19 patients may be linked to the activation of the immune system, as immune cells are known to utilize fatty acids as an energy source during activation and proliferation (NH et al., 2020). Specifically, T cells have been shown to upregulate fatty acid uptake and metabolism upon activation, with inhibition of fatty acid oxidation leading to impaired T cell function and decreased immune response. This suggests a potential interplay between the altered metabolic state observed in severe

COVID-19 patients and the immune response to the virus. However, it is important to note that other factors, such as higher body mass index and increased prevalence of heart disease in severe COVID-19 patients, may also be confounding factors contributing to the observed metabolic alterations (Kang et al., 2018; Mihalik et al., 2010; Ruiz et al., 2017; Smith et al., 2020).

Additionally, severe COVID-19 samples were found to have elevated levels of L-gamma-glutamyl-L-leucine and D-galactonate. The former is associated with incomplete protein breakdown and suggests that muscle tissue may be utilized for emergency energy in severe disease states. The latter is a metabolic byproduct of galactose breakdown and may be indicative of a need for extra energy in the body. However, chronic high levels of D-galactonate can result in adverse health effects, suggesting that further investigation into the potential negative consequences of these metabolic alterations is warranted (Schlueter et al., 2018).

Moreover, the genome-scale metabolic model utilized in this study was able to identify three reactions that were not identified through metabolomics alone. These reactions were found to be associated with the transport of phenylalanine, asparagine, and methionine, as well as isoleucine, tryptophan, and tyrosine. Notably, high plasma levels of phenylalanine and tyrosine have previously been identified as biomarkers of increased disease severity in COVID-19, indicating that the metabolic model utilized in this study is recapitulating previously identified markers of severe disease (Luporini et al., 2021; Shi et al., 2021). Furthermore, increased fumarate:thiosulfate antiport in the mitochondria, which was

identified through the metabolic model, suggests reduced access to oxygen (Spinelli et al., 2021)what .

Conclusion

In conclusion, the present study revealed enriched metabolites associated with tryptophan metabolism in non-acute COVID-19 samples, which suggested a metabolic picture of mitigated innate immune system inflammatory response and prevention of immunopathology-related lung damage (**Figure 4A- B**). Additionally, our results indicate an increased prevalence of histidine- and ketone-related metabolism in severe COVID-19 samples, which may provide a potential mechanistic insight into musculoskeletal degeneration-induced muscular weakness and the host's metabolism hijacked by SARS-CoV2 infection to increase viral replication and invasion. The non-acute COVID-19 samples demonstrated a reduced inflammatory response and a metabolic safety net to inhibit immunopathology-related lung damage, whereas severe COVID-19 samples' metabolomes indicated a takeover by the virus, resulting in a metabolic environment conducive to increased disease severity. The findings of this study could be leveraged to improve and advance COVID-19 disease treatment and lead to further investigation into metabolic predictors of disease progression.

References

Alwarawrah, Y., Kiernan, K., & MacIver, N. J. (2018). Changes in nutritional status impact immune cell metabolism and function. *Frontiers in Immunology*, *9*(MAY), 1. https://doi.org/10.3389/fimmu.2018.01055

Amino Acid Mixtures. (n.d.). Retrieved March 14, 2022, from https://www.promega.com/products/protein-expression/cell-free-protein-expression/amino-acid-mixtures/?catNum=L4461

Anderson, M. J. (2017). Permutational Multivariate Analysis of Variance (PERMANOVA). In *Wiley StatsRef: Statistics Reference Online* (pp. 1–15). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118445112.stat07841

Benjamini, Y., Hochberg, Y., & Benjamini, Yoav, H. Y. (1995). Benjamini and Y FDR.pdf. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brunk, E., Sahoo, S., Zielinski, D. C., Altunkaya, A., Dräger, A., Mih, N., Gatto, F., Nilsson, A., Preciat Gonzalez, G. A., Aurich, M. K., Prlic, A., Sastry, A., Danielsdottir, A. D., Heinken, A., Noronha, A., Rose, P. W., Burley, S. K., Fleming, R. M. T., Nielsen, J., … Palsson, B. O. (2018). Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nature Biotechnology 2018 36:3*, *36*(3), 272–281. https://doi.org/10.1038/nbt.4072

CDC. (n.d.). *Symptoms of Coronavirus*. Retrieved March 3, 2021, from https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html

Dehhaghi, M., Kazemi Shariat Panahi, H., & Guillemin, G. J. (2019). Microorganisms, Tryptophan Metabolism, and Kynurenine Pathway: A Complex Interconnected Loop Influencing Human Health Status. *International Journal of Tryptophan Research*, *12*. https://doi.org/10.1177/1178646919852996

Delattre, H., Sasidharan, K., & Soyer, O. S. (2021). Inhibiting the reproduction of SARS-CoV-2 through perturbations in human lung cell metabolic network. *Life Science Alliance*, *4*(1). https://doi.org/10.26508/LSA.202000869

Doğan, H. O., Şenol, O., Bolat, S., Yıldız, Ş. N., Büyüktuna, S. A., Sarıismailoğlu, R., Doğan, K., Hasbek, M., & Hekim, S. N. (2021). Understanding the pathophysiological changes via untargeted metabolomics in COVID-19 patients. *Journal of Medical Virology*, *93*(4), 2340–2349. https://doi.org/10.1002/JMV.26716

Donlan, A. N., Sutherland, T. E., Marie, C., Preissner, S., Bradley, B. T., Carpenter, R. M., Sturek, J. M., Ma, J. Z., Moreau, G. B., Donowitz, J. R., Buck, G. A., Serrano, M. G., Burgess, S. L., Abhyankar, M. M., Mura, C., Bourne, P. E., Preissner, R., Young, M. K., Lyons, G. R., … Petri, W. A. (2021). IL-13 is a driver of COVID-19 severity. *JCI Insight*, *6*(15). https://doi.org/10.1172/JCI.INSIGHT.150107

Edwards, J. S., Covert, M., & Palsson, B. (2002). Metabolic modelling of microbes: The flux-balance approach. *Environmental Microbiology*, *4*(3), 133–140. https://doi.org/10.1046/j.1462-2920.2002.00282.x

Gao, J., Xu, K., Liu, H., Liu, G., Bai, M., Peng, C., Li, T., & Yin, Y. (2018). Impact of the gut microbiota on intestinal immunity mediated by tryptophan metabolism.

Frontiers in Cellular and Infection Microbiology, *8*(FEB). https://doi.org/10.3389/fcimb.2018.00013

Hasan, M. R., Suleiman, M., & Pérez-López, A. (2021). Metabolomics in the Diagnosis and Prognosis of COVID-19. *Frontiers in Genetics*, *12*. https://doi.org/10.3389/FGENE.2021.721556

Hevia, H., Varela-Rey, M., Corrales, F. J., Berasain, C., Martínez-Chantar, M. L., Latasa, M. U., Lu, S. C., Mato, J. M., García-Trevijano, E. R., & Avila, M. A. (2004). 5′-Methylthioadenosine Modulates the Inflammatory Response to Endotoxin in Mice and in Rat Hepatocytes. *Hepatology*, *39*(4), 1088–1098. https://doi.org/10.1002/hep.20154

Ipata, P. L., & Tozzi, M. G. (2006). Recent advances in structure and function of cytosolic IMP-GMP specific 5′nucleotidase II (cN-II). *Purinergic Signalling*, *2*(4), 669. https://doi.org/10.1007/S11302-006-9009-Z

Jewett, B. E., & Thapa, B. (2020). Physiology, NMDA Receptor. *StatPearls*. https://www.ncbi.nlm.nih.gov/books/NBK519495/

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, *28*(1), 27–30. https://doi.org/10.1093/nar/28.1.27

Kang, M., Yoo, H. J., Kim, M., Kim, M., & Lee, J. H. (2018). Metabolomics identifies increases in the acylcarnitine profiles in the plasma of overweight subjects in response to mild weight loss: A randomized, controlled design study. *Lipids in Health and Disease*, *17*(1), 1–13. https://doi.org/10.1186/S12944-018-0887-1/FIGURES/3

Keaty, T. C., Keaty, T. C., Jensen, P. A., Jensen, P. A., & Jensen, P. A. (2020). Gapsplit: Efficient random sampling for non-convex constraint-based models. *Bioinformatics (Oxford, England)*, *36*(8), 2623–2625. https://doi.org/10.1093/BIOINFORMATICS/BTZ971

Kimhofer, T., Lodge, S., Whiley, L., Gray, N., Loo, R. L., Lawler, N. G., Nitschke, P., Bong, S. H., Morrison, D. L., Begum, S., Richards, T., Yeap, B. B., Smith, C., Smith, K. G. C., Holmes, E., & Nicholson, J. K. (2020). Integrative Modeling of Quantitative Plasma Lipoprotein, Metabolic, and Amino Acid Data Reveals a Multiorgan Pathological Signature of SARS-CoV-2 Infection. *Journal of Proteome Research*, *19*(11), 4442–4454. https://doi.org/10.1021/ACS.JPROTEOME.0C00519/ASSET/IMAGES/LARGE/PR0C00519_0003.JPEG

Krause, D., Suh, H. S., Tarassishin, L., Cui, Q. L., Durafourt, B. A., Choi, N., Bauman, A., Cosenza-Nashat, M., Antel, J. P., Zhao, M. L., & Lee, S. C. (2011). The tryptophan metabolite 3-hydroxyanthranilic acid plays anti-inflammatory and neuroprotective roles during inflammation: Role of hemeoxygenase-1. *American Journal of Pathology*, *179*(3), 1360–1372. https://doi.org/10.1016/j.ajpath.2011.05.048

Lacombe, M. L., Lamarche, F., De Wever, O., Padilla-Benavides, T., Carlson, A., Khan, I., Huna, A., Vacher, S., Calmel, C., Desbourdes, C., Cottet-Rousselle, C., Hininger-Favier, I., Attia, S., Nawrocki-Raby, B., Raingeaud, J., Machon, C., Guitton, J., Le Gall, M., Clary, G., … Boissan, M. (2021). The mitochondrially-

localized nucleoside diphosphate kinase D (NME4) is a novel metastasis suppressor. *BMC Biology*, *19*(1). https://doi.org/10.1186/S12915-021-01155-5

López-Hernández, Y., Monárrez-Espino, J., Oostdam, A. S. H. van, Delgado, J. E. C., Zhang, L., Zheng, J., Valdez, J. J. O., Mandal, R., González, F. de L. O., Moreno, J. C. B., Trejo-Medinilla, F. M., López, J. A., Moreno, J. A. E., & Wishart, D. S. (2021). Targeted metabolomics identifies high performing diagnostic and prognostic biomarkers for COVID-19. *Scientific Reports 2021 11:1*, *11*(1), 1–13. https://doi.org/10.1038/s41598-021-94171-y

Luporini, R. L., Pott-Junior, H., Di Medeiros Leal, M. C. B., Castro, A., Ferreira, A. G., Cominetti, M. R., & de Freitas Anibal, F. (2021). Phenylalanine and COVID-19: Tracking disease severity markers. *International Immunopharmacology*, *101*(Pt A). https://doi.org/10.1016/J.INTIMP.2021.108313

Marshall, M. (2020). How COVID-19 can damage the brain. *Nature*, *585*(7825), 342–343. https://doi.org/10.1038/D41586-020-02599-5

Matsuyama, S., Nagata, N., Shirato, K., Kawase, M., Takeda, M., & Taguchi, F. (2010). Efficient Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein by the Transmembrane Protease TMPRSS2. *Journal of Virology*, *84*(24), 12658–12664. https://doi.org/10.1128/jvi.01542-10

Mihalik, S. J., Goodpaster, B. H., Kelley, D. E., Chace, D. H., Vockley, J., Toledo, F. G. S., & Delany, J. P. (2010). Increased Levels of Plasma Acylcarnitines in Obesity and Type 2 Diabetes and Identification of a Marker of Glucolipotoxicity. *Obesity (Silver Spring, Md.)*, *18*(9), 1695. https://doi.org/10.1038/OBY.2009.510

Mohning, M. P., Downey, G. P., Cosgrove, G. P., & Redente, E. F. (2019). Mechanisms of Fibrosis. *Idiopathic Pulmonary Fibrosis*, 9–31. https://doi.org/10.1016/B978-0-323-54431-3.00003-2

Nachar, N. (2008). The Mann-Whitney U: A Test for Assessing Whether Two Independent Samples Come from the Same Distribution. *Tutorials in Quantitative Methods for Psychology*, *4*(1), 13–20. https://doi.org/10.20982/tqmp.04.1.p013

NH, K., KJ, S., AL, H., MM, C., RK, A., YH, L., KA, S., MR, G., D, J., S, L., DH, S., D, F.-P., J, L., FB, H., WS, G., VA, N., EA, O., JH, K., CM, P., … CH, L. (2020). Interleukin-13 drives metabolic conditioning of muscle to endurance exercise. *Science (New York, N.Y.)*, *368*(6490). https://doi.org/10.1126/SCIENCE.AAT3987

Overmyer, K. A., Shishkova, E., Miller, I. J., Balnis, J., Bernstein, M. N., Peters-Clarke, T. M., Meyer, J. G., Quan, Q., Muehlbauer, L. K., Trujillo, E. A., He, Y., Chopra, A., Chieng, H. C., Tiwari, A., Judson, M. A., Paulson, B., Brademan, D. R., Zhu, Y., Serrano, L. R., … Jaitovich, A. (2021). Large-Scale Multi-omic Analysis of COVID-19 Severity. *Cell Systems*, *12*(1), 23-40.e7. https://doi.org/10.1016/J.CELS.2020.10.003

Páez-Franco, J. C., Torres-Ruiz, J., Sosa-Hernández, V. A., Cervantes-Díaz, R., Romero-Ramírez, S., Pérez-Fragoso, A., Meza-Sánchez, D. E., Germán-Acacio, J. M., Maravillas-Montero, J. L., Mejía-Domínguez, N. R., Ponce-de-León, A., Ulloa-Aguirre, A., Gómez-Martín, D., & Llorente, L. (2021). Metabolomics analysis reveals a modified amino acid metabolism that correlates with altered oxygen

homeostasis in COVID-19 patients. *Scientific Reports 2021 11:1*, *11*(1), 1–12. https://doi.org/10.1038/s41598-021-85788-0

Pesi, R., Allegrini, S., Balestri, F., Garcia-gil, M., Cividini, F., Colombaioni, L., Jordheim, L. P., Camici, M., & Tozzi, M. G. (2021). Cytosolic 5'-Nucleotidase II Is a Sensor of Energy Charge and Oxidative Stress: A Possible Function as Metabolic Regulator. *Cells*, *10*(1), 1–13. https://doi.org/10.3390/CELLS10010182

Pieters, R., & Veerman, A. J. P. (1988). The role of 5'nucleotidase in therapy-resistance of childhood leukemia. *Medical Hypotheses*, *27*(1), 77–80. https://doi.org/10.1016/0306-9877(88)90088-6

Radisky, E. S., Lee, J. M., Lu, C. J. K., & Koshland, D. E. (2006). Insights into the serine protease mechanism from atomic resolution structures of trypsin reaction intermediates. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(18), 6835–6840. https://doi.org/10.1073/pnas.0601910103

Roberts, I., Wright Muelas, M., Taylor, J. M., Davison, A. S., Xu, Y., Grixti, J. M., Gotts, N., Sorokin, A., Goodacre, R., & Kell, D. B. (2022). Untargeted metabolomics of COVID-19 patient serum reveals potential prognostic markers of both severity and outcome. *Metabolomics*, *18*(1), 1–19. https://doi.org/10.1007/S11306-021-01859-3/FIGURES/4

Ruiz, M., Labarthe, F., Fortier, A., Bouchard, B., Legault, J. T., Bolduc, V., Rigal, O., Chen, J., Ducharme, A., Crawford, P. A., Tardif, J. C., & Des Rosiers, C. (2017). Circulating acylcarnitine profile in human heart failure: A surrogate of fatty acid metabolic dysregulation in mitochondria and beyond. *American Journal of*

Physiology. Heart and Circulatory Physiology, *313*(4), 768–781. https://doi.org/10.1152/AJPHEART.00820.2016

Schlueter, R. J., Al-Akwaa, F. M., Benny, P. A., Gurary, A., Xie, G., Jia, W., Chun, S. J., Chern, I., & Garmire, L. X. (2018). Pre-pregnant obesity of mothers in a multi-ethnic cohort is associated with cord blood metabolomic changes in offspring. *BioRxiv*. https://doi.org/10.1101/264374

Seth, S., Batra, J., & Srinivasan, S. (2020). COVID-19: Targeting Proteases in Viral Invasion and Host Immune Response. *Frontiers in Molecular Biosciences*, *7*, 215. https://doi.org/10.3389/fmolb.2020.00215

Shi, D., Yan, R., Lv, L., Jiang, H., Lu, Y., Sheng, J., Xie, J., Wu, W., Xia, J., Xu, K., Gu, S., Chen, Y., Huang, C., Guo, J., Du, Y., & Li, L. (2021). The serum metabolome of COVID-19 patients is distinctive and predictive. *Metabolism: Clinical and Experimental*, *118*. https://doi.org/10.1016/J.METABOL.2021.154739

Sindelar, M., Stancliffe, E., Schwaiger-Haber, M., Anbukumar, D. S., Adkins-Travis, K., Goss, C. W., O'Halloran, J. A., Mudd, P. A., Liu, W.-C., Albrecht, R. A., García-Sastre, A., Shriver, L. P., & Patti, G. J. (2021). Longitudinal metabolomics of human plasma reveals prognostic markers of COVID-19 disease severity. *Cell Reports. Medicine*, *2*(8), 100369. https://doi.org/10.1016/j.xcrm.2021.100369

Singh Patidar, B., Meena, A., Kumar, M., Menon, B., Rohil, V., & Kumar Bansal, S. (2018). Adenosine Metabolism in COPD: A Study on Adenosine Levels, 5'-Nucleotidase, Adenosine Deaminase and Its Isoenzymes Activity in Serum, Lymphocytes and Erythrocytes. *COPD*, *15*(6), 559–571. https://doi.org/10.1080/15412555.2018.1537365

Smith, E., Fernandez, C., Melander, O., & Ottosson, F. (2020). Altered acylcarnitine metabolism is associated with an increased risk of atrial fibrillation. *Journal of the American Heart Association*, *9*(21). https://doi.org/10.1161/JAHA.120.016737

Spinelli, J. B., Rosen, P. C., Sprenger, H. G., Puszynska, A. M., Mann, J. L., Roessler, J. M., Cangelosi, A. L., Henne, A., Condon, K. J., Zhang, T., Kunchok, T., Lewis, C. A., Chandel, N. S., & Sabatini, D. M. (2021). Fumarate is a terminal electron acceptor in the mammalian electron transport chain. *Science (New York, N.Y.)*, *374*(6572), 1227–1237. https://doi.org/10.1126/SCIENCE.ABI7495

Stukalov, A., Girault, V., Grass, V., Karayel, O., Bergant, V., Urban, C., Haas, D. A., Huang, Y., Oubraham, L., Wang, A., Hamad, M. S., Piras, A., Hansen, F. M., Tanzer, M. C., Paron, I., Zinzula, L., Enghleitner, T., Reinecke, M., Lavacca, T. M., … Pichlmair, A. (2021). Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV. *Nature*, 1–7. https://doi.org/10.1038/s41586-021-03493-4

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., Fan, T. W.-M., Fiehn, O., Goodacre, R., Griffin, J. L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A. N., Lindon, J. C., Marriott, P., Nicholls, A. W., … Viant, M. R. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics*, *3*(3), 211–221. https://doi.org/10.1007/s11306-007-0082-2

Thomas, T., Stefanoni, D., Reisz, J. A., Nemkov, T., Bertolone, L., Francis, R. O., Hudson, K. E., Zimring, J. C., Hansen, K. C., Hod, E. A., Spitalnik, S. L., & D'Alessandro, A. (2020). COVID-19 infection alters kynurenine and fatty acid

metabolism, correlating with IL-6 levels and renal status. *JCI Insight*, *5*(14). https://doi.org/10.1172/JCI.INSIGHT.140327

Veal, N., Hsieh, C. L., Xiong, S., Mato, J. M., Lu, S., & Tsukamoto, H. (2004). Inhibition of lipopolysaccharide-stimulated TNF-α promoter activity by S-adenosylmethionine and 5′-methylthioadenosine. *American Journal of Physiology - Gastrointestinal and Liver Physiology*, *287*(2 50-2). https://doi.org/10.1152/ajpgi.00316.2003

Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., & Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*, *181*(2), 281-292.e6. https://doi.org/10.1016/j.cell.2020.02.058

Walls, A. C., Tortorici, M. A., Frenz, B., Snijder, J., Li, W., Rey, F. A., DiMaio, F., Bosch, B. J., & Veesler, D. (2016). Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nature Structural and Molecular Biology*, *23*(10), 899–905. https://doi.org/10.1038/nsmb.3293

Wang, L., Erlandsen, H., Haavik, J., Knappskog, P. M., & Stevens, R. C. (2002). Three-dimensional structure of human tryptophan hydroxylase and its implications for the biosynthesis of the neurotransmitters serotonin and melatonin. *Biochemistry*, *41*(42), 12569–12574. https://doi.org/10.1021/bi026561f

Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S., & Crispin, M. (2020). Site-specific glycan analysis of the SARS-CoV-2 spike. *Science*, *369*(6501), 330–333. https://doi.org/10.1126/science.abb9983

World Health Organization. (2021, December 12). *Coronavirus (COVID-19) Dashboard With Vaccination Data*. https://covid19.who.int/

Xia, J., Psychogios, N., Young, N., & Wishart, D. S. (2009). MetaboAnalyst: A web server for metabolomic data analysis and interpretation. *Nucleic Acids Research*, *37*(SUPPL. 2). https://doi.org/10.1093/nar/gkp356

Xiao, N., Nie, M., Pang, H., Wang, B., Hu, J., Meng, X., Li, K., Ran, X., Long, Q., Deng, H., Chen, N., Li, S., Tang, N., Huang, A., & Hu, Z. (2021). Integrated cytokine and metabolite analysis reveals immunometabolic reprogramming in COVID-19 patients with therapeutic implications. *Nature Communications*, *12*(1). https://doi.org/10.1038/S41467-021-21907-9

Yan, B., Chu, H., Yang, D., Sze, K.-H., Lai, P.-M., Yuan, S., Shuai, H., Wang, Y., Kao, R. Y.-T., Chan, J. F.-W., & Yuen, K.-Y. (2019). Characterization of the Lipidomic Profile of Human Coronavirus-Infected Cells: Implications for Lipid Metabolism Remodeling upon Coronavirus Replication. *Viruses*, *11*(1), 73. https://doi.org/10.3390/v11010073

Yirmiya, R., & Goshen, I. (2011). Immune modulation of learning, memory, neural plasticity and neurogenesis. *Brain, Behavior, and Immunity*, *25*(2), 181–213. https://doi.org/10.1016/J.BBI.2010.10.015

Zhang, Y., Guo, R., Kim, S. H., Shah, H., Zhang, S., Liang, J. H., Fang, Y., Gentili, M., Leary, C. N. O., Elledge, S. J., Hung, D. T., Mootha, V. K., & Gewurz, B. E. (2021). SARS-CoV-2 hijacks folate and one-carbon metabolism for viral replication. *Nature Communications*, *12*(1). https://doi.org/10.1038/S41467-021-21903-Z