Cryptosporidium-Induced Shifts in the Bacterial Microenvironment of Bovine Intestinal Epithelial Cells

Michael Gage [1]

[1] School of Biomedical Science and Engineering, University of Maine, Orono, ME

Author approval

All authors have approved this draft for submission to Brightspace for grading: yes
The authors will turn this analysis into a manuscript for publication eventually: no

Author contributions

Michael Gage is a graduate student, advised by Dr. Sue Ishaq. Michael is responsible for the data analysis, statistics, creating figures, reading literature, and writing the manuscript. Dr. Ishaq provided the necessary data, coding workflow framework, and contributed to the review and editing of the manuscript.

Keywords (5 - 10): Bacterial, Cattle, Cow, Mammalian, Cell Culture, Intestinal Epithelium, Amplicon Sequencing

Abstract

This study investigates the taxonomic changes resulting from *Cryptosporidium parvum* protozoal infection within bovine intestinal cells. The immediate objective is to indicate microbial shifts in large and small intestine cultures of *Bos taurus* upon pathogen admittance. We hypothesize that the infection of *Cryptosporidium* will significantly alter the bacterial taxa, and that a reduced bacterial diversity in the bovine intestines would increase susceptibility to protozoal infection. Sample collection included the small and large intestines of two cows, and treatments included +/- antibiotics to culture media, and +/- *Cryptosporidium* cysts, followed by DNA sequencing of the bacterial 16S rRNA gene, with sequencing and quality control performed on both forward and reverse reads. Post-analysis uncovered compositional changes associated with *Cryptosporidium* presence, although the findings were not statistically significant. These compositional changes included a decline in microbial diversity in *Cryptosporidium* samples compared to those without infection. However, samples without *Cryptosporidium* show a lower overall abundance of specific predictor taxa compared to infected samples, suggesting a distinct microbial composition. *Cryptosporidium*-infected samples also display increased log abundance levels for a broader range of taxa, particularly for Limosilactobacillus, Veillonella, and Lachnospiraceae. These findings may bolster the understanding of *Cryptosporidium* infection and interactions within bovine gastrointestinal systems in future studies.

Introduction

The bovine gastrointestinal tract hosts a diverse microbial ecosystem critical to cattle health, influencing digestion, nutrient absorption, and immune function [1]. However, this balanced microbial community is vulnerable to disruptions from pathogens such as *Cryptosporidium* spp., a genus of protozoan parasites that target the intestinal epithelium. Infected cattle may suffer from diarrhea, weight loss, reduced growth rates, and other health complications that can have substantial economic impacts on livestock farming [2]. This study seeks to examine how *Cryptosporidium* infection affects the taxonomic composition of the bovine intestinal microbiome, as well as how a low-diversity bacterial community (induced using antibiotics) might increase susceptibility to protozoal infection, which may reveal microbial responses that could inform better management and prevention strategies in the livestock industry. For example, preliminary data suggests that *Cryptosporidium* preferentially infect the ileum of the small intestines, especially in neonate animals, which suggests that low microbial diversity may play a role in microbial ecology and competitive exclusion of pathogens [3].

Our primary objective is to analyze how the introduction of *Cryptosporidium* impacts microbial diversity within the bovine intestine, specifically targeting both the small and large intestinal sections. This study investigates whether specific taxa within the microbiome are consistently impacted by infection and whether certain microbial shifts might correlate with successful pathogen establishment in the host. Understanding these dynamics could offer critical insights into how pathogen-induced microbiome changes influence disease progression and host health [4].

This study is guided by the research question: How does the bacterial microenvironment of the bovine intestine shift following *Cryptosporidium* infection? We hypothesize that pathogenic infection will result in significant microbial alterations. We anticipate that general taxa important for gut health may decrease in relative abundance, while opportunistic or pathogen-associated taxa may thrive in response to infection pressures, reshaping the gut microbial community.

Methods

Experimental Design

In this experiment, samples were collected from the small and large intestines of two different cows (*Bos taurus*). Ten biopsy samples were collected from the small intestine, and ten from the large intestine. Samples were placed individually on well plates and introduced into experimental conditions. Both large and small intestine samples were exposed to the same five experimental conditions. These conditions are as follows: 1) Glycerol, 2) DMEM, 3) DMEM & antibiotics, 4) DMEM & *Crypto*, 5) DMEM, antibiotics & *Crypto*. Two negative control plates were composed of only 1) DMEM, and 2) DMEM with antibiotics (Fig 1). Cells were placed in a 5% CO2 incubator set to 37 degrees Celsius for 6 days. After incubation, DNA extraction, PCR amplifications and 16S rRNA gene paired-end Illumina sequencing were performed to extract data for computational analysis.

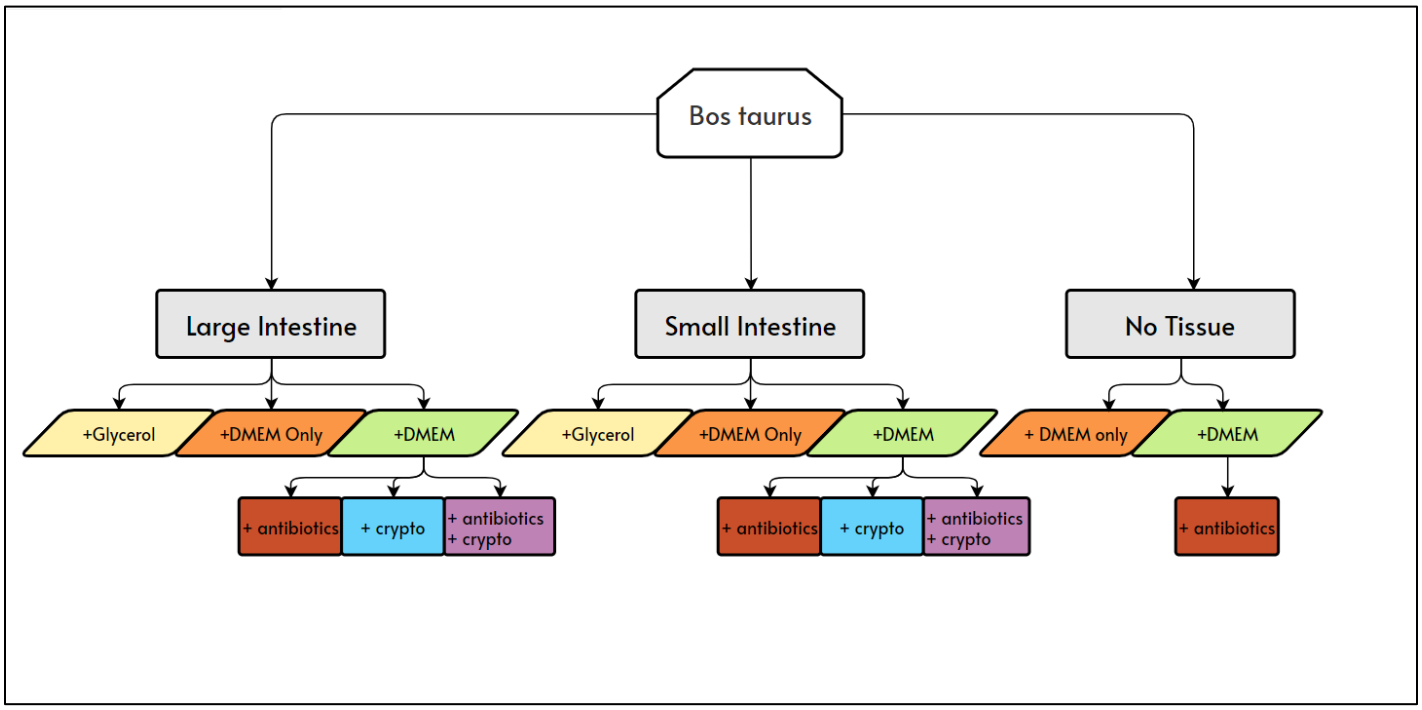

Figure 1. Experimental conditions for Bos taurus large intestine, small intestine, and no tissue controls.

Filtering and Trimming

Forward and reverse sequencing reads were processed in RStudio to determine their quality and pursue additional trimming parameters as necessary. The R packages ggplot2, devtools, dada2, and Rqc were used for the quality assessment and trimming of the data [5, 6, 7, 8]. The data contained 23 samples and 3,610,171 paired forward/reverse reads, each 300 bases long. The forward read quality score dropped below 35 at approximately cycle 175. The reverse read quality score dropped below 35 at cycle 200. The forward and reverse reads were used, as the majority of bases had an acceptable quality score. The first 10 bases were trimmed with the trimLeft parameter to exclude poor quality bases and avoid inclusion of the primers. For the trimRight parameter, 25 bases were removed from both the forward and reverse. For both reads, the quality score began to significantly decline around 250-275 bases. At around 275-300 bases, the score fell below 25, so 25 bases were trimmed from both right ends. The maxEE parameter was set for both reads: 2 for forward, and 3 for reverse. maxN was set to 0 (as dada2 would not allow for ambiguous bases), and rm.phix was set to TRUE to exclude any phix DNA used as a positive control (ensuring that only sample data was included in the assessment). Using these parameters, a filtered plot for forward and reverse reads was produced. The quality scores in these filtered aggregate plots did not fall below 25, so it was determined that the quality parameters were defined correctly (Fig 2). The total raw read count was 3,610,171, and the total filtered read count was 2,745,757. The largest number of raw reads was 296,378 for sample “40”, and the smallest was 54,609 for sample “37”. For filtered reads, the largest number was 237,245 for sample “40”, and the smallest was 47,109 for sample “39”.

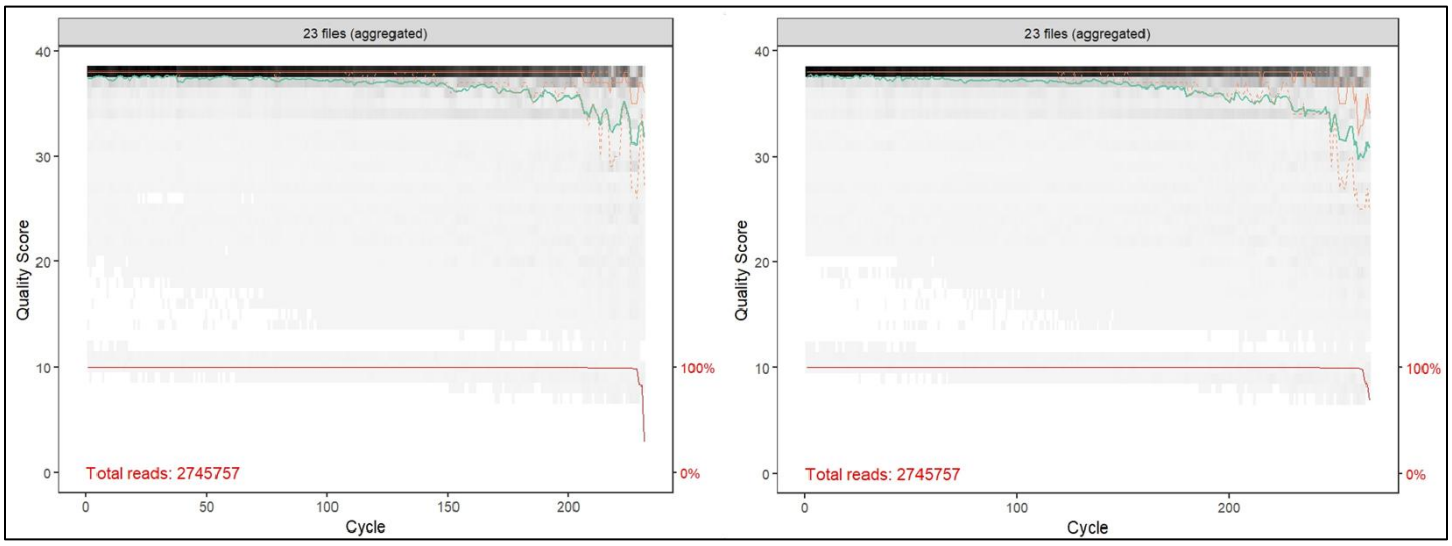

Figure 2. Filtered aggregate reads for forward (left) and reverse (right) showing a right cutoff at a quality score of approximately 25.

For error rate picking and the creation of sequence variants, the R package dada2 was used [7]. The number of bases used for error rate picking was 27,737,502 for forward reads, and 12,505,257 for reverse reads. Error profiles were analyzed for both forward and reverse reads. Both showed a similar trend in expected base change error frequency relative to the estimated error in the data, as expected (Fig 3).

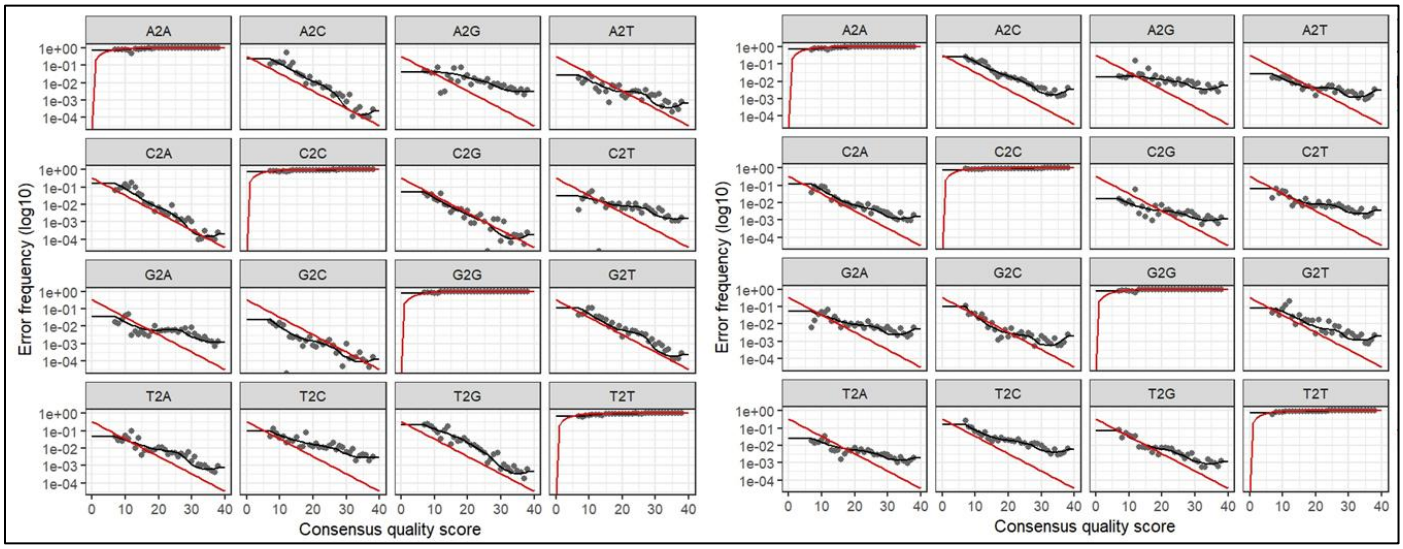

Figure 3. Error frequency of forward and reverse reads. A similar trend is shown in expected base change error frequency relative to the estimated error.

Filtered forward and reverse reads were merged, generating a total of 9,962 sequence variants in 23 samples retained after filtering. Chimeric sequences were then identified and filtered out of the merged sequence table. Out of 9,962 SVs, 6,578 chimeras were found and filtered. The resulting table included 23 samples and 3,384 SVs. At 66%, the volume of chimeric sequences developed during PCR amplification was abnormally high. A summary of quality filtering and trimming stages was visualized as a form of workflow verification using the tidyr and ggplot2 packages [9, 5] (Fig 4).

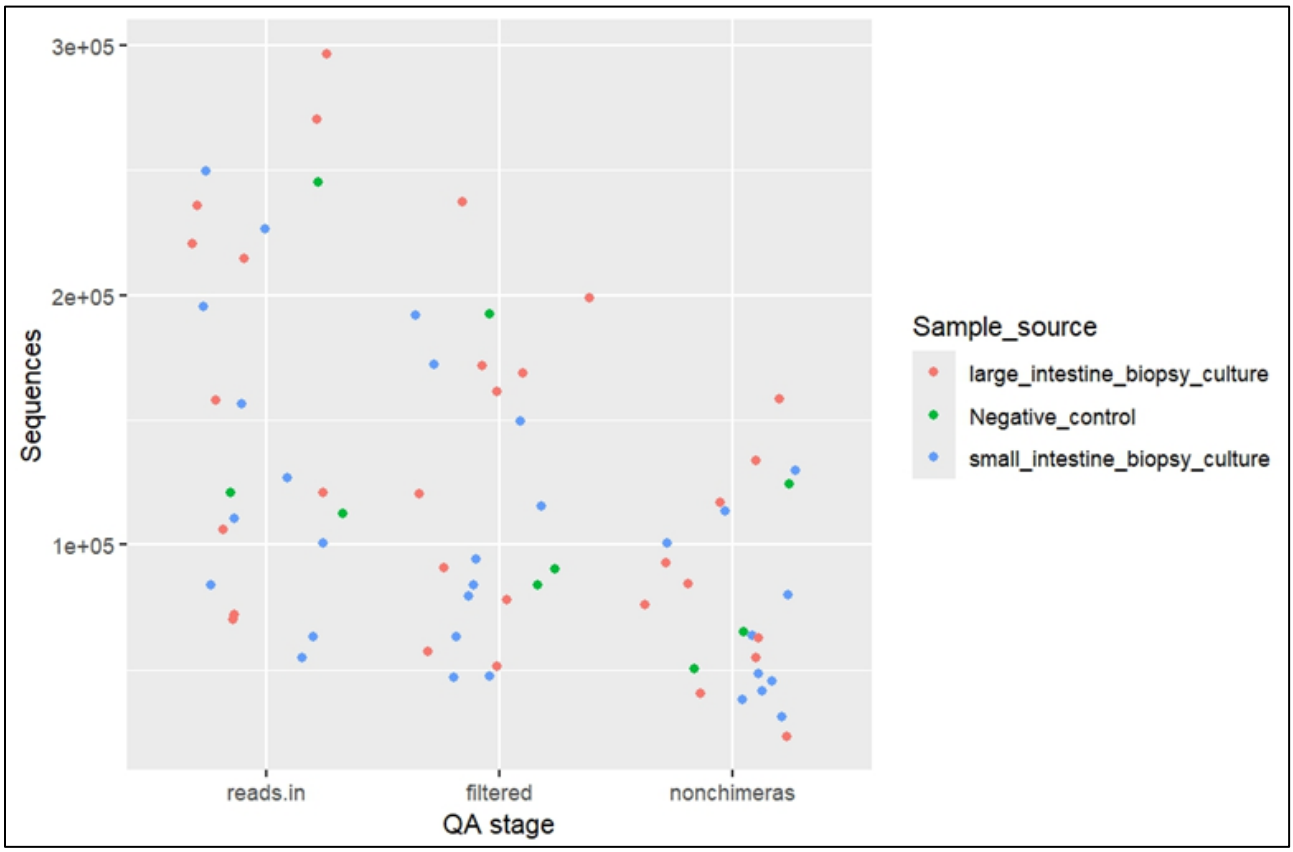

Figure 4. QA stages from reads.in, to initial filtering and removal of chimeric sequences. Samples colored by anatomical source.

Contaminants were then identified and removed from the dataset, as well as negative controls. This step was performed using the Ishaq Decontamination Method in RStudio, removing 375 taxa identified in control samples. After these taxa were removed, the resulting data table contained 20 samples and 3,009 SVs.

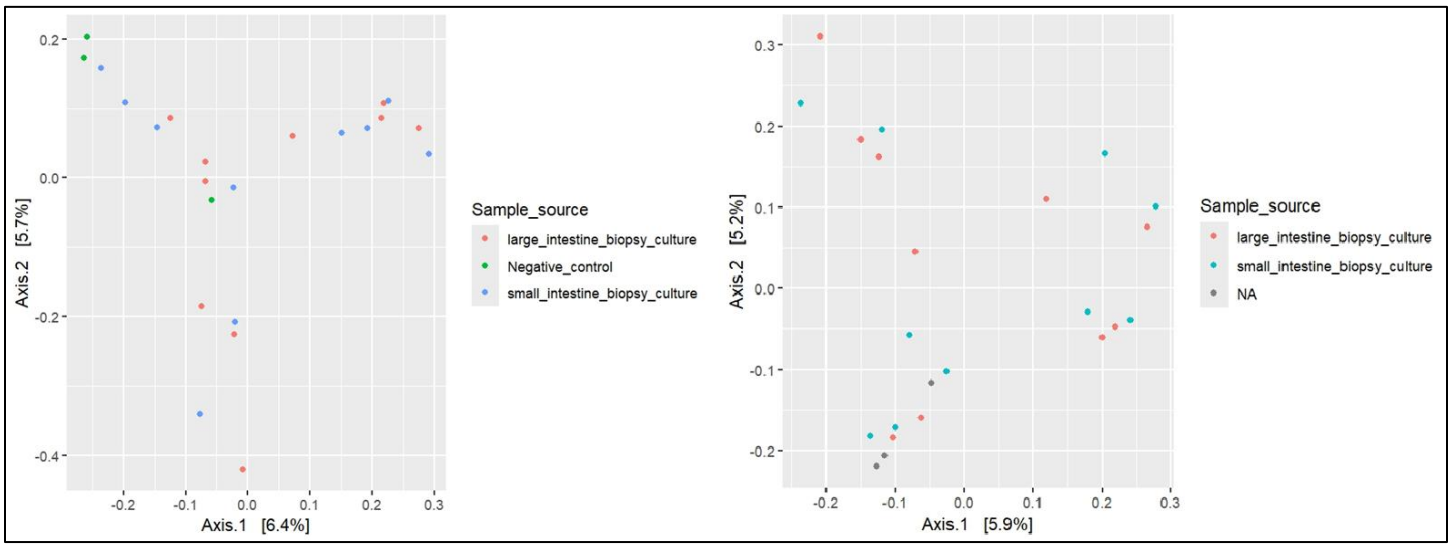

Figure 5. Samples before (left) and after (right) contaminants and negative controls were removed.

Taxonomy Assignment

To assign taxonomy, "Silva 138.1 prokaryotic SSU taxonomic training data formatted for DADA2" (file name "silva_nr99_v138.1_train_set.fa.gz") was selected and used [10]. Using the dplyr package, mitochondria and chloroplast taxa were removed [11]. The removal of these taxa resulted in the remaining 20 samples and 2,136 SVs.

Rarefaction

Rarefaction of the data was performed to normalize SVs to even depth across all samples. Using the vegan R package, 3 samples and 1,059 SVs were removed [12]. The original data collected for this experiment was small in volume, so the rarefaction cutoff value (sample.size) was set to 2,000. This allowed for minimal loss of collected samples, however, this parameter does inflate the Kurtosis value described later.

Visual and Statistical Analysis

Bacterial richness was analyzed in samples with and without *Crypto*, showing the variance across samples in a violin plot. This alpha diversity measure was completed using the ggplot2 and ggsignif R packages to run statistical analysis and visualization [5, 13]. The phyloseq R package was then used to determine the richness, evenness, and Shannon Diversity metrics of the data [14]. A kurtosis value was determined via the PerformanceAnalytics package, which was also used to run the Shapiro-Wilk Normality test to determine alpha diversity distribution [15].

The Mann-Whitney statistical test was performed using lme4, lmerTest, and emmeans packages [15, 16, 17]. This test was necessary due to the normal distribution of richness, evenness, and Shannon Diversity metrics. Alpha diversity was measured again and organized by phylum present in *Crypto* and non-*Crypto* samples.

Changes in taxonomy between samples without and with *Crypto* were determined using Feature Prediction (Differential Abundance) with Random Forest. The following R packages were used: phyloseq, vegan, plyr, dplyr, magrittr, scales, grid, reshape2, knitr, randomForest, rfPermute, ggplot2, RColorBrewer [14, 12, 19, 11, 20, 21, 22, 23, 24, 25, 26, 5, 27]. The output produced a random forest plot showing predictor taxa in the two factorial groups.

Additional data visualizations and statistical tests were performed to evaluate beta diversity. These analyses were completed using the ggplot2, viridis, and vegan packages [5, 28, 12]. A permanova test was completed on Crypto_added x Samples_source metadata, as well as Crypto_added x Antibiotics metadata in RStudio. An NMDS beta ordination plot (with associated stress values) was developed to determine microbial diversity. The vegan package was then utilized to perform a beta dispersion test and an Anova test on the beta dispersion data.

Phyloseq, vegan, and ggplot2 packages were utilized to conduct a constrained analysis of principal coordinates or distance-based RDA (db-RDA) for non-linear relationships [14, 12, 5]. This model included both Crypto_added and Sample_source variable in the analysis. A subsequent ANOVA test was performed to determine the efficacy of the model and the data.

An ecological core microbiome model was performed to determine the prevalence of specific bacterial sequence variants (SVs) in *Bos taurus* (cow) intestinal samples. R packages ggplot2, RColorBrewer, viridis, dplyr, and microbiome were utilized to execute the model and develop a heatmap of results [5, 27, 28, 11, 29].

An R file with the detailed workflow necessary for computational methods can be found in the “Supplementary Materials” section.

Results

Alpha Diversity

Data frames were developed for richness, evenness, and Shannon Diversity, then merged into one consolidated data frame. This data was then used to develop three histograms: one for richness, one for evenness, and one for Shannon Diversity (Fig 6).

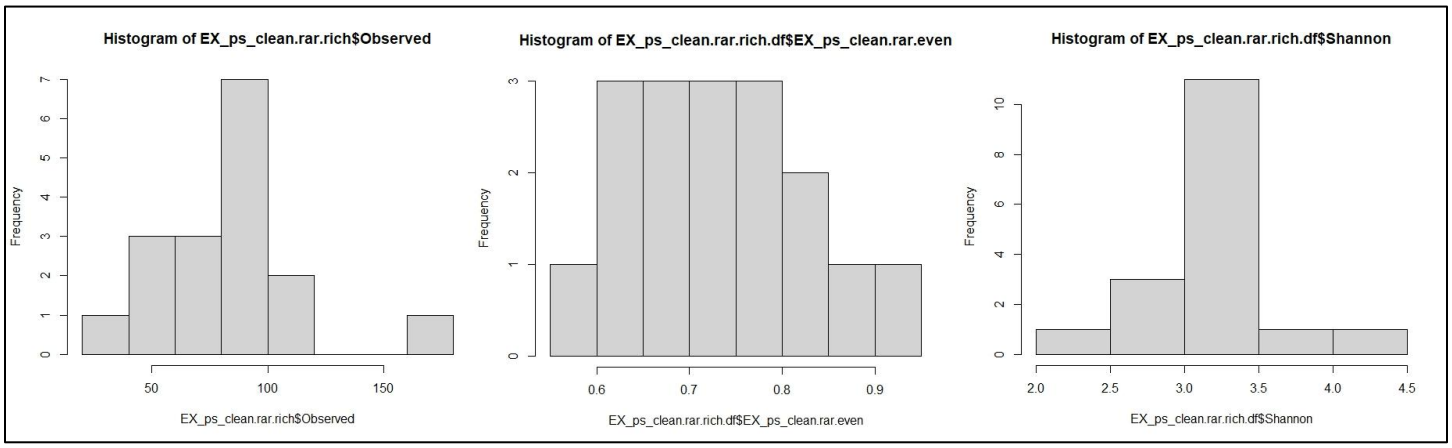

Figure 6. Histograms representing Richness (left), Evenness (center), and Shannon Diversity (right) metrics. All show slight left skew, with Evenness displaying near-platykurtic behavior.

All three show relative left skew, with the evenness histogram displaying near-platykurtic shape. To further analyze these results, three kurtosis values were calculated: richness = 2.18, evenness = -0.82, and Shannon = 1.59 (Table 1). Of these three values, richness and Shannon are very high, indicating several potential outliers in the data. This is a direct result of the rarefaction value being set to 2000. This was set to avoid losing too much data from a minimal sample pool. The Shapiro-Wilk Normality test was performed to determine the statistical distribution of the data. In all three cases (richness, evenness, and Shannon Diversity), it is determined that data is near normal distribution (W statistic near 1) despite the abnormal kurtosis values. The p-values for all three metrics are above a value of 0.05, indicating that we cannot conclude that the data significantly deviates from a normal distribution. (Table 1).

Table 1. Kurtosis and Shapiro-Wilk statistical values for metrics Richness, Evenness, and Shannon Diversity.

Metric	Kurtosis value	W Statistic (Shapiro-Wilk)	p-value (Shapiro-Wilk)
Richness	2.18	0.90	0.07
Evenness	-0.82	0.97	0.77
Shannon	1.59	0.92	0.15

After determining the normal distribution of data, a Mann-Whitney statistical test for pairwise categorical factors was performed. The p-value for all three metrics was determined to be insignificant, showing lack of evidence to claim differences in distribution between *Crypto*. and non-*Crypto* groups.

Table 2. Results of the Mann-Whitney test for Richness, Evenness, and Shannon Diversity Metrics.

Metric	U Statistic (W) (Mann-Whitney)	p-value (Mann-Whitney)
Richness	17	0.12
Evenness	48	0.15
Shannon	39	0.59

In Fig. 7, observed richness and Shannon Diversity metrics are displayed in a violin plot, showing bacterial richness in samples with and without *Cryptosporidium*. It may be observed that the bacterial richness across samples decreases in diversity when *Crypto*. is present. This result assumes that the invasion of *Crypto*. alters the biodiversity in bovine intestinal cells.

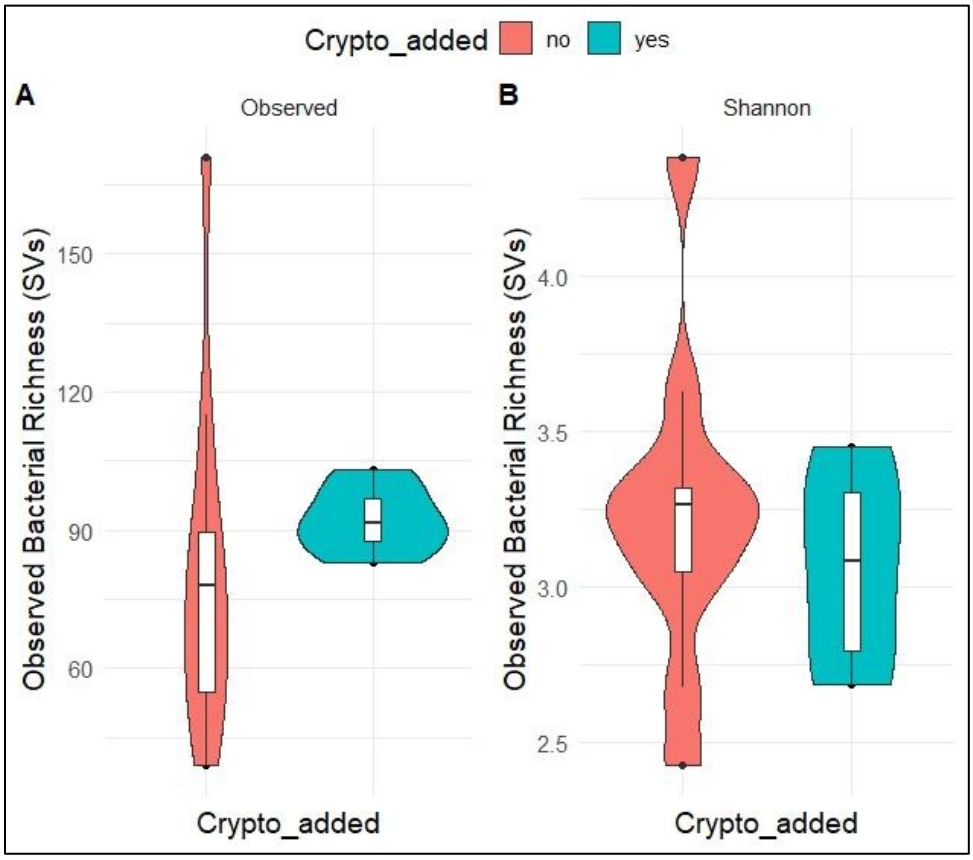

Figure 7. Observed and Shannon Diversity bacterial richness present in Crypto. and non-Crypto. samples.

Beta Diversity

The NMDS beta diversity ordination plot (Fig 8) shows broader microbial diversity in samples without *Crypto*. as opposed to those with *Crypto*. These results may confirm the introduction of *Crypto*. does in fact alter the microbial diversity in samples. During the production of the NMDS plot, stress values were gathered for all 21 runs, averaging at 0.16 (between an acceptable value of 0.15 - 0.2). This value indicates that there is a reasonably accurate representation of relationships among samples in this ordination.

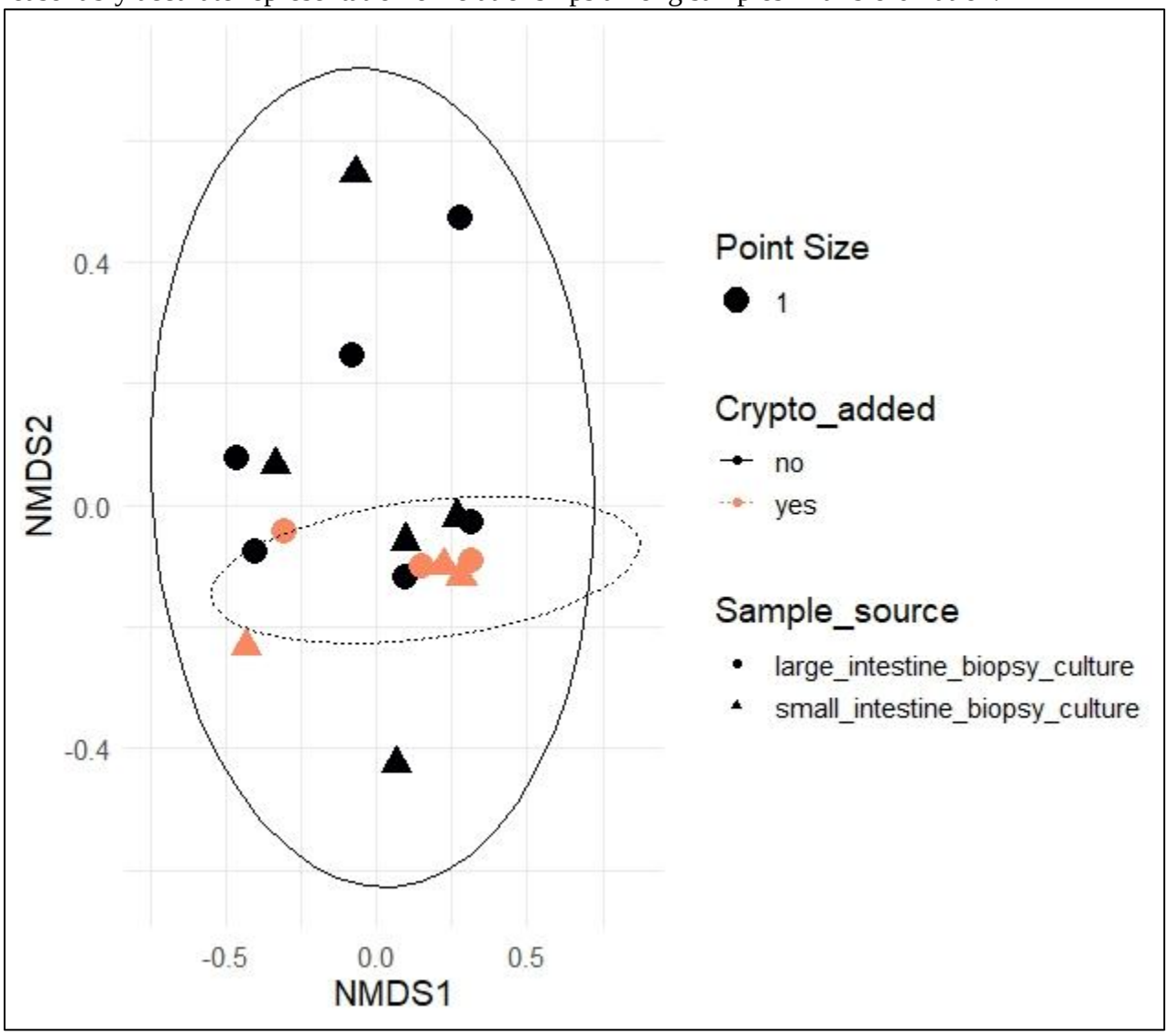

Figure 8. NMDS beta ordination plot delineating *Crypto* groups and sample source. Samples with and without *Crypto* maintain varying microbial diversity, and sample source follows a random distribution.

The permANOVA (Permutational Multivariate Analysis of Variance) statistical model detected differences in the multivariate distances between groups based on factors *Crypto*_added and Sample_source. For these two factors, an R-squared value of 0.17 and p-value of 0.91 were obtained. These two factors did not significantly account for the variation in samples, so the factors Crypto_added and Antibiotics were selected. The R-squared value was 0.22 and the p-value was 0.007. These results still do not account for much of the sample variance, yet are significant. It is determined that the presence of antibiotics in samples may primarily alter microbial variance over *Crypto*. infection in this study.

An additional beta dispersion test was performed to determine the homogeneity of group dispersions. The average distance to median in samples with *Crypto*. is 0.59 and 0.65 without *Crypto*. This indicates that samples in groups with *Crypto*. infection are less dispersed than those without, resulting in slightly more homogeneity. A subsequent beta dispersion ANOVA test showed a p-value of 0.02, showing a significant dispersion in groups with and without *Crypto*. This result may suggest that the results from the permANOVA test could be influenced by unequal group dispersion.

Taxonomy

Sample data was grouped by phylum to determine the presence of certain phyla in samples with and without *Crypto*. invasion. In Fig. 9, certain phyla are present in samples without *Crypto*., yet lack presence in those with *Crypto*. invasion. This supports prior alpha diversity measures indicating that *Crypto*. may alter the biodiversity in bovine intestinal cells.

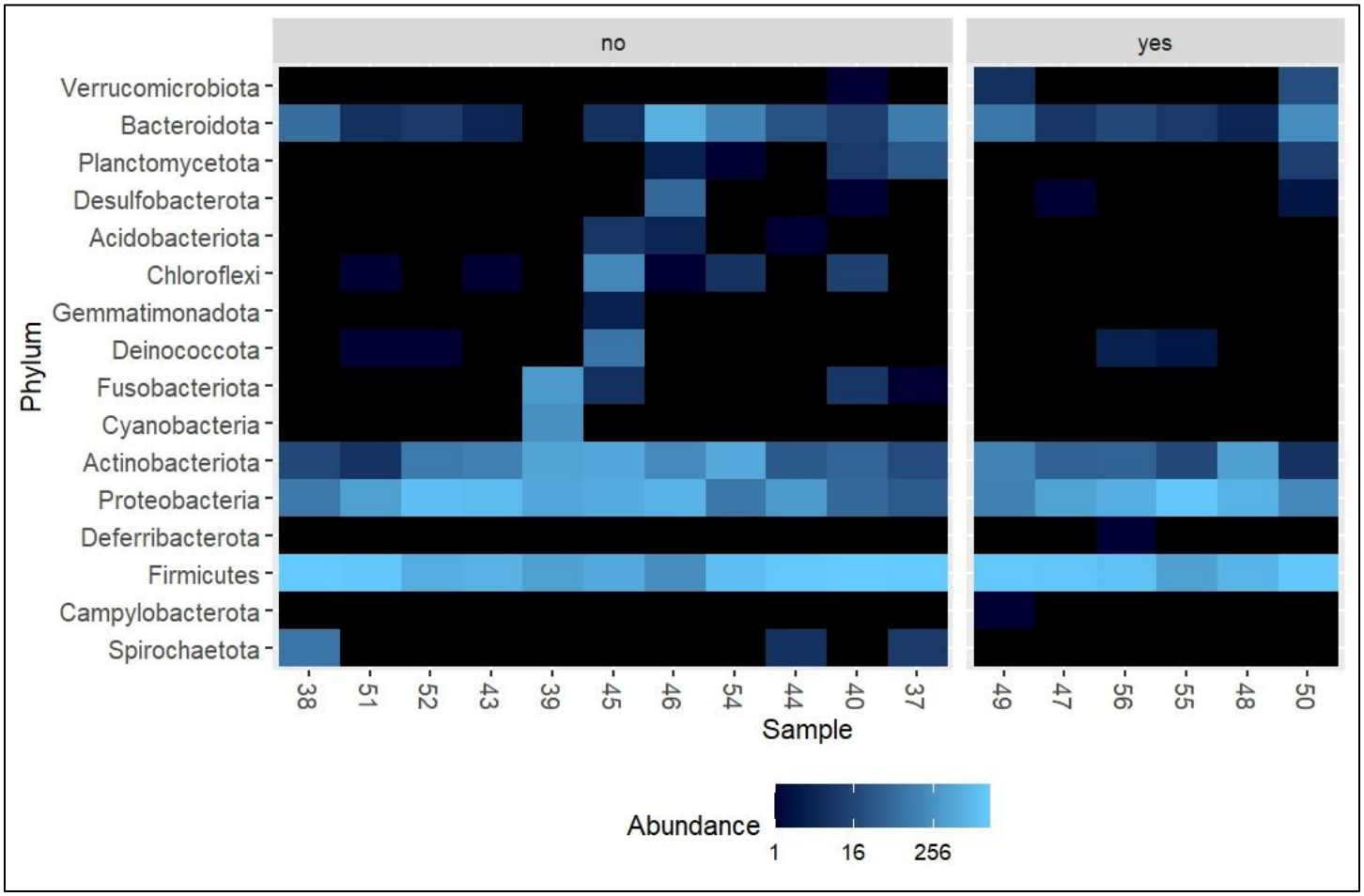

Figure 9. Heatmap showing phyla distribution in samples with and without *Crypto*. The abundance of certain phyla is altered with the introduction of *Crypto*.

Differential Abundance

The Random Forest model was constructed to classify samples based on the presence of *Cryptosporidium* ("With_Crypto") and its absence ("Without_Crypto"). The model achieved an overall classification accuracy of 60%, with a 95% confidence interval (CI) of 36.1% to 80.9%. For "Without_Crypto" samples, the model accurately classified all 12 samples, resulting in 100% accuracy for this group. For "With_Crypto" samples, the model did not correctly classify any of the 8 samples, leading to a 0% accuracy in this group. Of the 2,136 SVs present in 20 total samples, only 10 were selected at a p-value at or below 0.05 as predictor taxa in the Random Forest model. The heatmap in Fig. 10 shows the log abundance of the 10 predictor taxa (SVs) for each sample, with samples split by group.

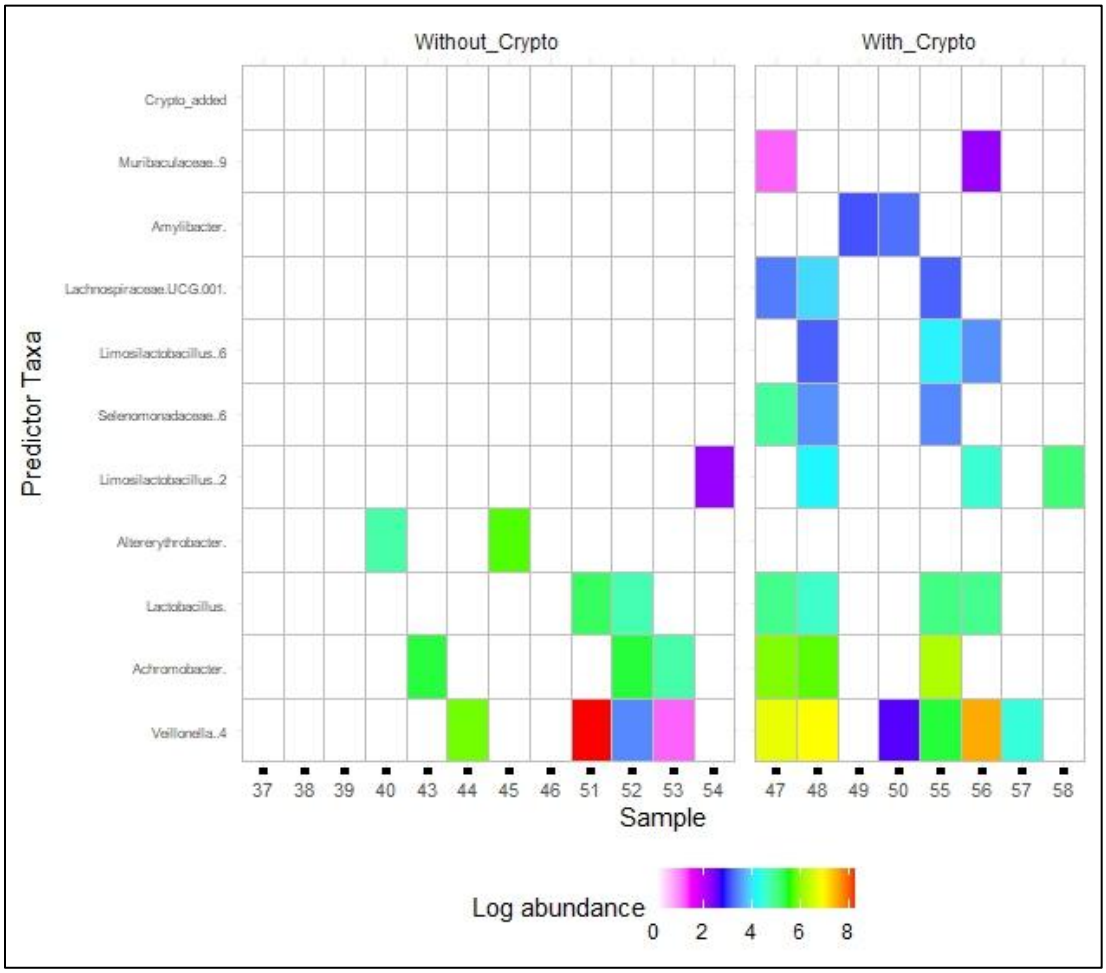

Figure 10. Heatmap displaying the log abundance of 10 predictor taxa (SVs) with and without *Cryptosporidium*.

Each cell in the heatmap represents the abundance of a particular taxon (y-axis) within a sample (x-axis). The colors represent log abundance levels, with a gradient scale at the bottom for reference. The "Without_Crypto" samples show a lower overall abundance of most predictor taxa compared to "With_Crypto" samples, suggesting a distinct microbial composition. The “With_Crypto” samples display increased log abundance levels for a broader range of taxa, particularly for Limosilactobacillus, Veillonella, and Lachnospiraceae. This may indicate a microbial community shift associated with the presence of *Cryptosporidium.*

The model's inability to classify "With_Crypto" samples may indicate limitations in the predictive features or overlapping microbial profiles between groups. Additional predictive features or larger sample sizes might improve classification performance.

Constrained Analysis

Distance-based redundancy analysis (db-RDA) for non-linear data was conducted as a constrained ordination to determine the effect of combined explanatory variables (*Crypto* presence, Sample Type) on the response variable (community composition) (Fig. 11). An ANOVA test was performed to determine the significance of this model, and the results were non-significant compared to the standard 0.05 significance threshold (P = 0.997).

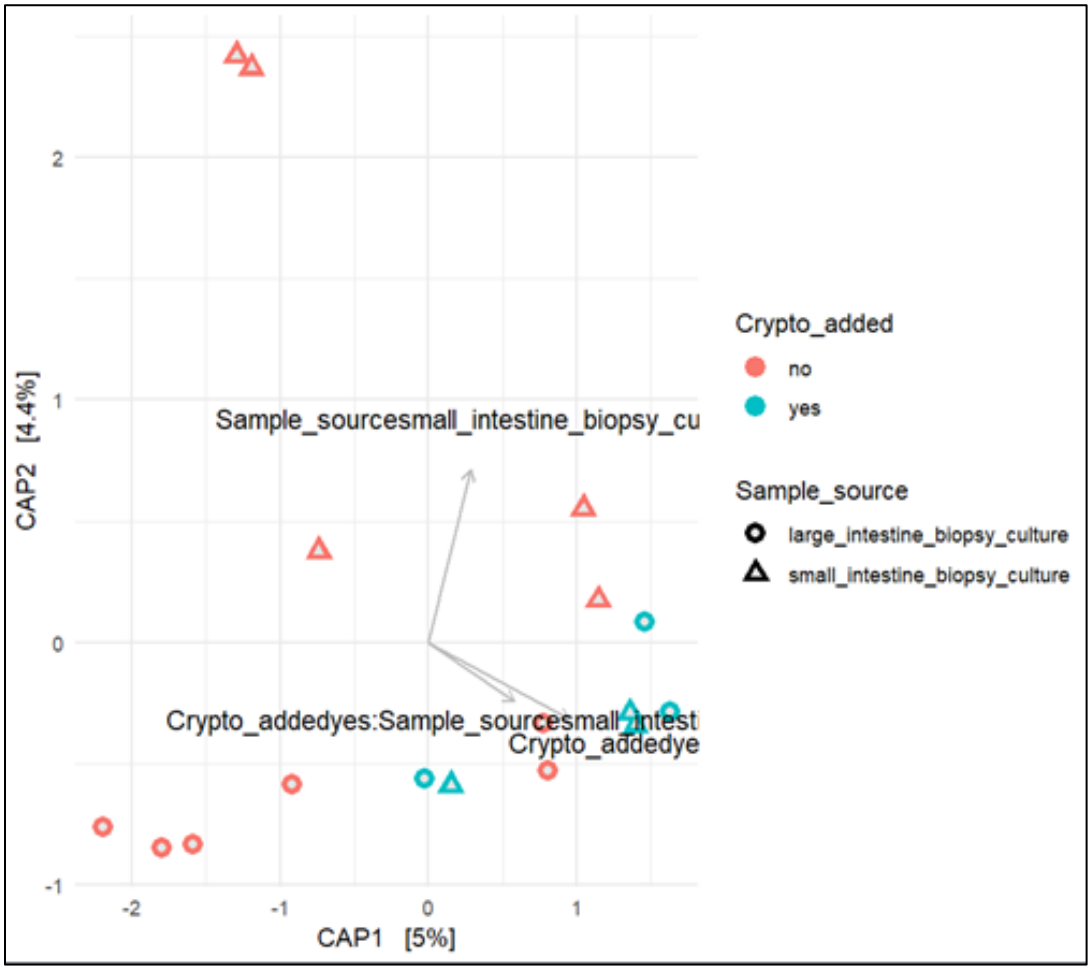

Figure 11. db-RDA plot exhibiting variables Crypto_added and Sample_source. The axes represent the constrained ordination space derived from the explanatory variables. CAP1 explains 5% of the variance, and CAP2 explains 4.4%.

Ecology

An additional core microbiome analysis was performed to determine the SV's that have a minimum prevalence set to 60% or higher in samples. This model was visualized in a heatmap, only including SV's in the >=60% threshold (Figure 12). *Clostridium sensu stricto 1* is consistently detected at high prevalence even at lower detection thresholds, suggesting it is both abundant and widespread in the intestinal samples. *Escherichia-Shigella* has a moderately high prevalence but drops off at higher detection thresholds, indicating it is less abundant in some samples. *Bifidobacterium* and *Veillonella* are less prevalent at higher thresholds, suggesting these taxa are present but at relatively low abundances. This model was conducted to determine the baseline bacterial community members across all samples without considering the effect of *Crypto* on these communities.

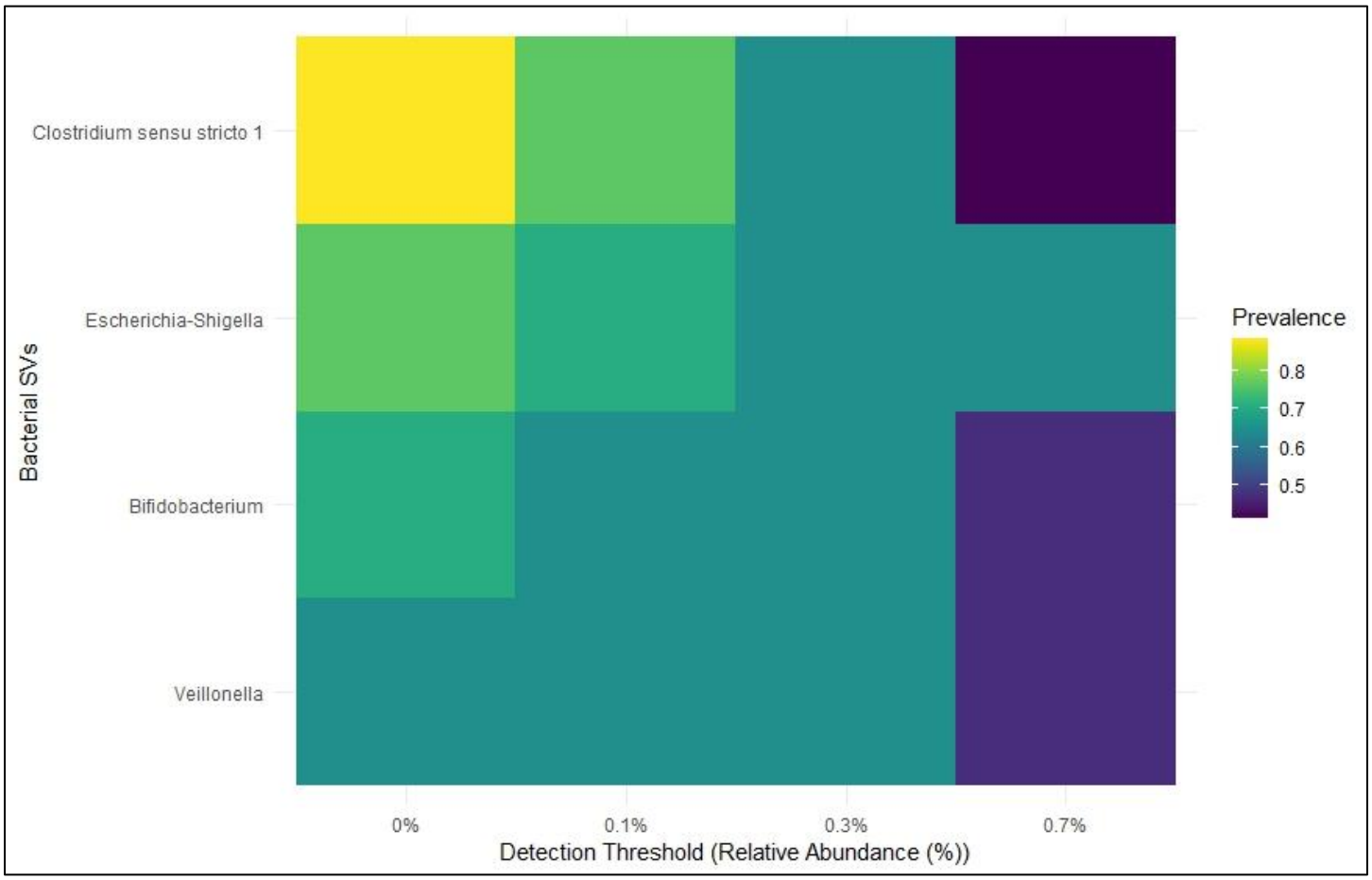

Figure 12. Core microbiome model heatmap representing the prevalence of specific bacterial sequence variants (SVs) in Bos taurus (cow) intestinal samples.

Discussion

This study aimed to determine whether the introduction and invasion of Cryptosporidium parvum would adversely alter the bacterial taxa represented in bovine intestinal cells. We hypothesized that the infection of infection of Cryptosporidium would significantly alter the bacterial taxa, and that a reduced bacterial diversity in the bovine intestines would increase susceptibility to protozoal infection.

Results support the hypothesis that the introduction of Cryptosporidium does alter the microenvironment of cells, however, these results were not statistically significant. This suggests that alterations to the experiment may be necessary. Increasing the sample size in repeated studies may be the most essential step in producing reliable, significant results. In this study, many outliers were maintained in the data to conserve volume. As an example, the rarefaction value was set to 2000 (instead of a standard 5000) during the alpha diversity analysis step. In a future study, a sufficient sample size would warrant the filtering and trimming out of such data when appropriate.

During this study, several statistical tests were performed. Most resulted in highly insignificant p values, indicating a strong rejection of the null hypothesis in most cases. This occurred when accounting for Crypto_added and Sample_source variables, however, p-values did become significant when accounting for the addition of Antibiotics in sample groups. For the permANOVA test, The R-squared value was 0.22 and the p-value was 0.007. These results still do not account for much of the sample variance, yet are significant. It was determined that the presence of antibiotics in samples may primarily alter microbial variance over Crypto. infection in this study. This was highly anticipated as the study was conducted, due to the poor sample size.

For future consideration, a similar study may be performed with additional factors, variables, and a much larger sample size. Although the results of this study were not significant, they tend to agree with prior similar studies. *Cryptosporidium* has been observed in a previous study finding that infection in mice leads to a significant reduction in microbial diversity [30].

In our taxonomic analysis of *Crypto* and non-*Crypto* infected specimen, *Proteobacteria* and *Firmicutes* did not show a decrease in abundance like many other taxa. We observe that these taxa are either unaffected by or benefit from the introduction of *Cryptosporidium parvum*. A similar study on mice found that *Crypto* infection resulted in a marked increase in *Proteobacteria* and *Firmicutes* [31].

Due to the similar findings between this study and past research, it would be wise to revisit this analysis with new data rather than reject the results of this study altogether.

Acknowledgements

The authors acknowledge the University of Maine AVS 454/554 class for their collective advice, peer review, and training on amplicon data analysis.

The authors acknowledge investigators Emily Pierce, Sue Ishaq, and Sarah Hosler for the collection and distribution of the data used in this analysis.

References

1. McGavin, M. C., & H. F. (2015). The role of the microbiome in gastrointestinal health and disease. In Veterinary Clinics of North America: Food Animal Practice (Vol. 31, pp. 205-215). Elsevier.
2. O'Handley, R. M., & R. W. (2004). Cryptosporidium and Giardia in cattle: A review. Veterinary Parasitology, 125(1-2), 1-18. DOI: 10.1016/j.vetpar.2004.08.006.
3. Ichikawa-Seki, M., Motooka, D., Kinami, A., Murakoshi, F., Takahashi, Y., Aita, J., Hayashi, K., Tashibu, A., Nakamura, S., Iida, T., Horii, T., & Nishikawa, Y. (2019). Specific increase of fusobacterium in the faecal microbiota of neonatal calves infected with cryptosporidium parvum. Scientific Reports, 9(1). https://doi.org/10.1038/s41598-019-48969-6
4. Wang, L., Zhang, H., & Xu, Z. (2021). Interaction between Cryptosporidium spp. and gut microbiota: A review. *Frontiers in Microbiology*, 12, 678084. DOI: 10.3389/fmicb.2021.678084.
5. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag. https://ggplot2.tidyverse.org

6. Wickham, H., & R. H. (2021). devtools: Tools to Make Developing R Packages Easier. R package version 2.4.3. https://CRAN.R-project.org/package=devtools
7. Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581-583. DOI: 10.1038/nmeth.3869.
8. Gascuel, O., & G. D. (2018). rqc: A Collection of Functions for Quality Control of Data Frames. R package version 0.1.2. https://CRAN.R-project.org/package=rqc
9. Wickham, H. (2021). tidyr: Tidy Messy Data. R package version 1.1.4. https://CRAN.R-project.org/package=tidyr
10. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schneider, J., & B. F. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 41(D1), D590-D596. DOI: 10.1093/nar/gks1219. https://www.arb-silva.de/
11. Wickham, H. (2021). dplyr: A Grammar of Data Manipulation. R package version 1.0.7. https://CRAN.R-project.org/package=dplyr
12. Oksanen, J., Guillaume, F., Blake, S., & Peter, R. (2020). vegan: Community Ecology Package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan
13. Ahlmann-Eltze, C. (2020). ggsignif: Significance Brackets for 'ggplot2'. R package version 0.6.1. https://CRAN.R-project.org/package=ggsignif
14. McMurdie, P. J., & Holmes, S. P. (2013). 14q: An R Package for Handling and Analyzing High-Throughput Microbiome Census Data. *PLoS ONE*, 8(4), e61217. DOI: 10.1371/journal.pone.0061217.
15. Peterson, B. G., & Carl, P. (2021). PerformanceAnalytics: Econometric Tools for Performance and Risk Analysis. R package version 2.1.2. https://CRAN.R-project.org/package=PerformanceAnalytics
16. Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear Mixed-Effects Models Using 'Eigen' and S4. R package version 1.1-12. https://CRAN.R-project.org/package=lme4
17. Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest: Tests in Linear Mixed Effects Models. R package version 3.1-0. https://CRAN.R-project.org/package=lmerTest
18. Lenth, R. V. (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.7.2. https://CRAN.R-project.org/package=emmeans
19. Wickham, H. (2011). plyr: Tools for Splitting, Applying and Combining Data. R package version 1.8.6. https://CRAN.R-project.org/package=plyr
20. Bache, S. M., & Wickham, H. (2014). magrittr: A Forward-Pipe Operator for R. R package version 1.5. https://CRAN.R-project.org/package=magrittr
21. Wickham, H. (2021). scales: Scale Functions for Visualization. R package version 1.1.1. https://CRAN.R-project.org/package=scales
22. Chang, W. (2021). grid: Graphics in R. R package version 4.1.0. https://CRAN.R-project.org/package=grid
23. Wickham, H. (2007). reshape: A Reshaping Package. R package version 1.4.4. https://CRAN.R-project.org/package=reshape2
24. Yihui, X. (2021). knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.33. https://CRAN.R-project.org/package=knitr
25. Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, 2(3), 18-22. https://CRAN.R-project.org/package=randomForest
26. Wright, M. N., & Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, 77(1), 1-17. DOI: 10.18637/jss.v077.i01.
27. Neuwirth, E. (2014). RColorBrewer: ColorBrewer Palettes. R package version 1.1-2. https://CRAN.R-project.org/package=RColorBrewer

28. Garnier, S. (2018). viridis: Default Color Maps from 'matplotlib'. R package version 0.5.1. https://CRAN.R-project.org/package=viridis
29. Lahti, L., Shetty, S., et al. (2017–2020). Tools for microbiome analysis in R. R Package Version 1.28.0. Bioconductor. http://microbiome.github.io/microbiome
30. Charania, R., Wade, B. E., McNair, N. N., & Mead, J. R. (2020). Changes in the microbiome of cryptosporidium-infected mice correlate to differences in susceptibility and infection levels. *Microorganisms*, *8*(6), 879. https://doi.org/10.3390/microorganisms8060879
31. Oliveira, B. C., & Widmer, G. (2018). Probiotic product enhances susceptibility of mice to cryptosporidiosis. *Applied and Environmental Microbiology*, *84*(21). https://doi.org/10.1128/aem.01408-18

Supplementary Material

RStudio Code File
Cryptosporidium_parvum_bos_taurus_R_workflow.R