Analyzing Variation in Lemur Gut Microbiome Data from the Duke Lemur Center

Brian Janger, Peter Liu, Ryan Mitchell, Qihan Zhao

October 11th, 2023

Lemurs are a broad family of primates containing many different diverse species. Due to their primate status, it is known that lemurs are very genetically similar to humans. According to the American Museum of Natural History, human DNA is, on average, 96% similar to our most distant primate relatives, of which lemurs are included. While lemurs aren't as similar to humans when compared to more close relatives like chimpanzees and bonobos, it is still generally believed that both the *Hominidae* family and the *Lemuroidea* family evolved from a common ancestor over the past 50-60 million years. Because of this, lemurs have become an important species for biological research, and their biological patterns and physical characteristics can be used to find valuable inferences about the biology and evolution of humans. While human biology and taxonomy is out of the scope of this paper, this study aims to analyze lemur gut microbiome data and discover differences in gut microbiomes across species, social group, individual identity, time of day, day, and season.

Specifically, we have a few research questions we would like to answer regarding this case study:

- Accounting for gut microbiome patterns with respect to species and identity, is there a significant difference in lemur gut microbiome patterns depending on the time of day?
- To what extent do lemur gut microbiome patterns vary across days and/or seasons? What patterns can be accounted for by differences in the time the observations were measured?
- Out of all of the predictors, which are the most influential when it comes to variation of lemur gut microbiome patterns?
- Do gut microbiome pattern differences themselves differ between ring-tailed lemurs and sifakas with drastically different gastrointestinal systems?

The data for this study was collected by Dr. Lydia Greene from the Duke Lemur Center, a non-invasive research center affiliated with Duke University which preserves and researches endangered lemur species. The data collected involves 293 different fecal samples from 12 lemurs at the Duke Lemur Center, of which 6 are ring-tailed lemurs and 6 are sifakas. The ring-tailed lemurs and sifakas were both evenly split across 3 different forest enclosures (2 ring-tailed and 2 sifakas per enclosure), and the lemurs' fecal samples were measured in pairs on 3 consecutive days in both the spring and summer seasons. During the days where a lemur was observed, fecal samples were taken and recorded multiple times throughout the day. After cleaning all of the data, the final dataset leaves us with information about the lemur (name, enclosure, and species), the date and time the sample was taken, and 206 measurements of various genomes and families from the bacteria domain. These bacterial measurements are measured as the overall proportion of the sample which contained said bacteria; we will analyze these differences in bacteria in the lemur fecal samples with respect to the aforementioned predictors of time of day, day, season, species, social group, and identity. The data also provided more general metrics about the microbial samples; specifically, we have three different response variables that our models will be fitted for. The first is the Shannon Entropy, which describes the microbial evenness of the sample; higher degrees of bacterial uniformity in the sample are associated with a higher Shannon Entropy. The next is Faith's Phylogenetic Diversity, which describes the overall diversity of the sample; this is similar, but not identical, to Shannon Entropy since more uniform samples are not necessarily more diverse. The last is the number of observed features in the sample, which simply details the number of different bacteria found in the sample. Using these three response variables, our models will attempt to discover which of our predictors have a relevant influence on the variation of lemur gut microbiome patterns.

Studies have already been performed regarding microbial data of lemurs, including a study from last year where Dr. Greene herself was an author. The study found that lemur gut microbial data is sensitive to "multiple scales of environmental differences" and that "non-dietary factors govern some of the variability". While one of our goals is to affirm these findings and account for them in our own analysis, our main goal is to extend the existing knowledge of lemur gut microbiome variation to include factors related to time-series data.

In this case study, we examine the known existing association between lemur gut microbiome bacterial distributions and taxonomic predictors in relation with the possible relevance of time of day, day, and season. This way, the significant associations of the time-related predictor variables can be evaluated while controlling for the known associations of other predictors used in the study. Below, we explore the data to find possible trends between various predictors and our three response variables and examine the most prevalent bacterial families found in both of the lemur species.

Exploratory Data Analysis

Our analysis of lemur gut microbiomes centers around the three aforementioned diversity metrics: observed features, Shannon Entropy, and Faith's Phylogenetic Diversity. To help address our research questions, we are interested in seeing how factors such as species, time of day, season impact these diversity scores. We performed exploratory data analysis surrounding each of the three metrics to uncover initial insights. We were also interested in exploring the bacterial composition data; we perform exploratory data analysis on this as well.

Shannon Entropy

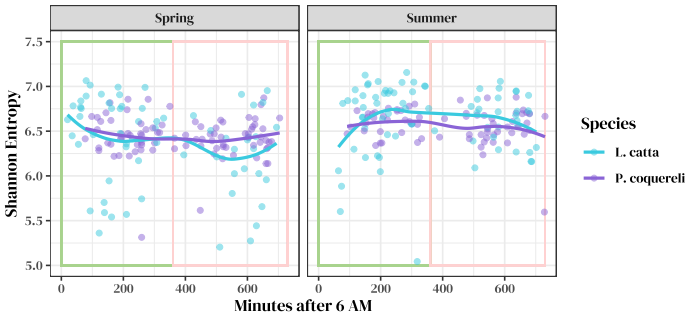


Figure 1: Shannon Entropy Over Time by Season and Species. Green rectangle = AM, Pink rectangle = PM.

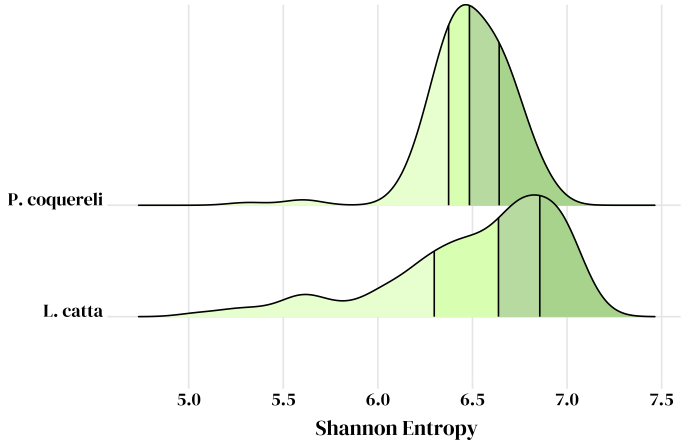


Figure 2: Distributions of Shannon Entropy by Species

Figure 1 aims to visualize the effect of time of day, season, and species on Shannon Entropy score. The scatterplots show that, in general, scores from P. coquereli samples tend to be less variant (or more clustered) than L. catta scores. P. coquereli scores also seem to remain consistent throughout the day, while L. catta shows some peaks and valleys; interestingly, L. catta scores seem to be highest in the early morning during the spring, but lowest in the early morning during the summer. The most interesting insight perhaps comes from comparing the scores themselves; during the spring, P. coquereli scores tend to average higher throughout the day compared to L. catta, but this is reversed during the summer. This suggests there may be an association between season and Shannon Entropy.

Figure 2 hones in on the species-level, giving us more information on how Shannon Entropy is distributed among each lemur classification. Particularly, neither species boasts a normal distribution as both show notable left skews; the first quantile for each species both cover over 60% of that species' range. Overall, L. catta scores skew higher (and have a higher median) than P. coquereli. However, this visualization supports the idea that Shannon Entropy in L. catta is more varied; P. coquereli scores are more heavily concentrated around their median of 6.5.

Faith's Phylogenetic Diversity

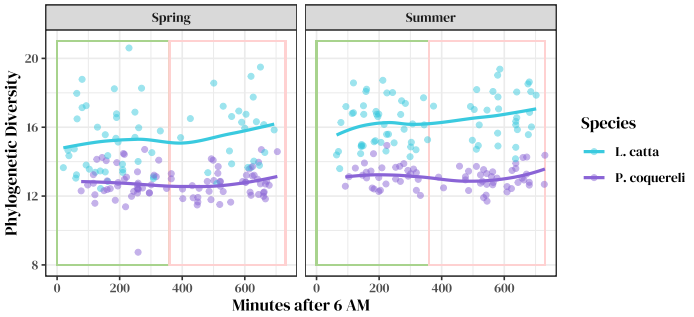


Figure 3: Phylogenetic Diversity Over Time by Season and Species. Green rectangle = AM, Pink rectangle = PM.

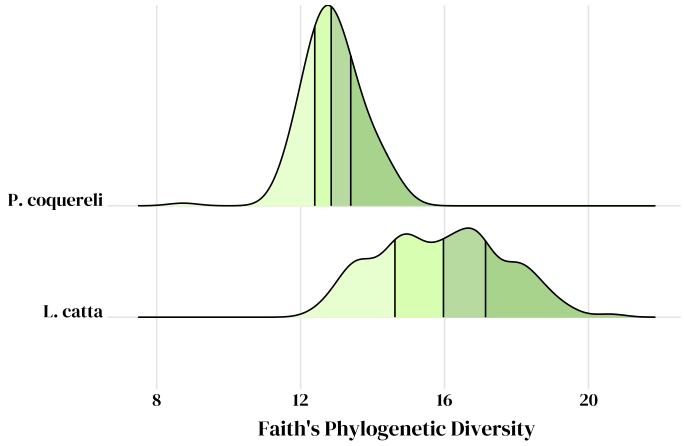


Figure 4: Distributions of Phylogenetic Diversity by Species

Figure 3 aims to visualize the effect of time of day, season, and species on Faith's Phylogenetic Diversity score. Like for Shannon Entropy, the scatterplots show that, in general, scores from P. coquereli samples tend to be less variant (or more clustered) than L. catta scores. Scores from both species seem to remain consistent throughout the day, suggesting that there might not be a strong association between time of day

and Faith PD score. There does not appear to be much of a seasonal difference either. However, there seems to be a clear distinction on the species level, with L. catta scores being consistently higher than P. coquereli scores by a notable margin.

Figure 4 focuses on species. Again, this supports the idea that Faith PD is more variant in L. catta, this time even more notably than for Shannon Entropy; the IQR for P. coquereli is smaller than the range of any quantile of L. catta. Unlike Shannon Entropy, the distributions for both species are approximately normal, with L. catta scores clearly skewing higher. EDA suggests an association exists between species and Faith PD.

Observed Features

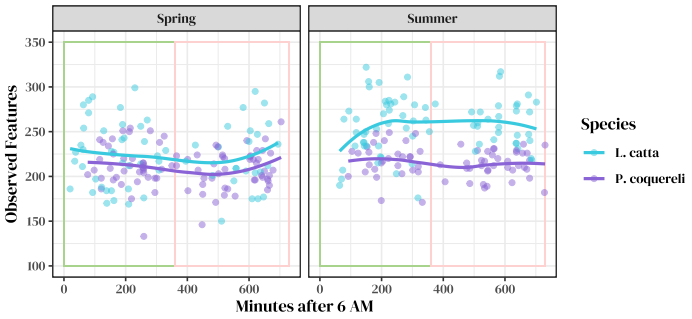


Figure 5: Observed Features Over Time by Season and Species. Green rectangle = AM, Pink rectangle = PM.

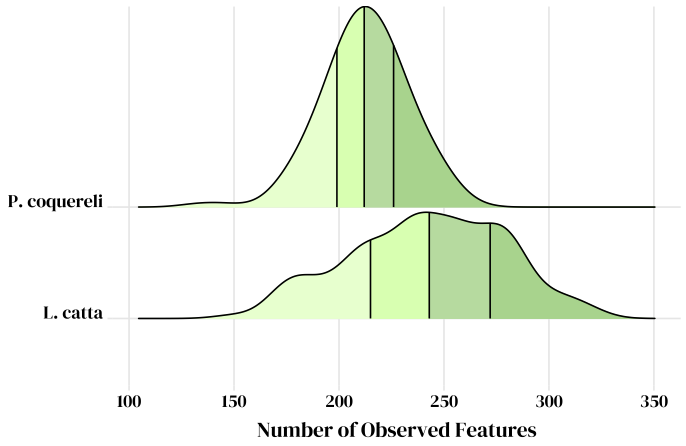


Figure 6: Distributions of Observed Features by Species

Figure 5 aims to visualize the effect of time of day, season, and species on number of observed features. Like for Faith PD, the scatterplots show a clear distinction on the species level, with L. catta scores being consistently higher than P. coquereli scores throughout the day. However, unlike Faith PD, there also appears to be a clear association between season and score, particularly for L. catta lemurs; the difference in scores between the two species is notably exacerbates during the summer, with L. catta scores consistently reaching above 250, a value rarely seen in the spring. Time of day seems to have less clear of an association, but there are some patterns: namely, in the spring, both species see a gradual decline in score from the beginning of the day until the mid-afternoon when it begins to pick up again. In the summer, P. coquereli scores remain relatively consistent, while L. catta sees a gradual increase from the beginning of the day until soon after 9AM. However, these differences from time of day appear minimal and there does not seem to be a stark contrast in observed features in the AM versus the PM for either species.

Figure 6 focuses on species. Once again, the data for L. catta is more varied, with an IQR twice as large as that of P. coquereli. The distributions for both species are approximately normal, and, like Faith PD, scores for L. catta skew higher. However, this difference is not as pronounced as Faith PD.

Bacterial Composition

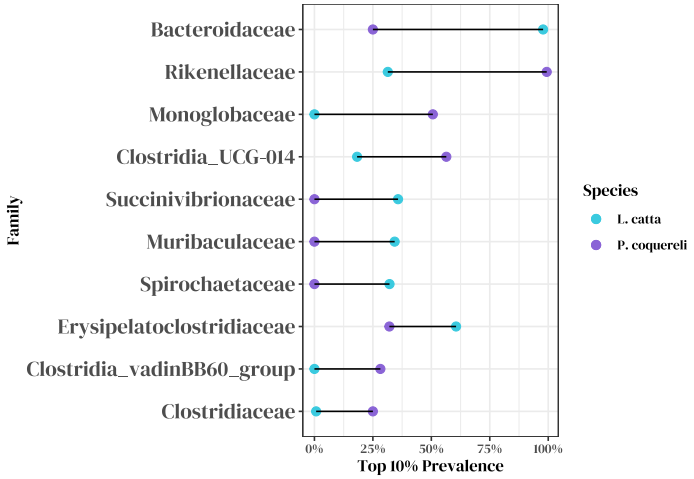


Figure 7: Top 10 families with the highest difference in prevalent appearance between species

We were particularly interested in bacteria that consistently appeared prevalently in one species but perhaps not the other to suggest which bacteria most distinguish L. catta gut microbiomes from P. coquereli gut microbiomes. To decrease the granularity of the data, we extracted the family of each bacteria, reducing the dimensionality to 80. We defined appearing prevalently as being in the top 10% of families in terms of microbial composition for each observation; therefore, we recorded the 8 families with the highest values for each observation. We then calculated the percentage of time each family appeared prevalently in each species by dividing the number of times it was in the top 10% by the total amount of observations for that species. **Figure 7** displays the 10 families with the largest difference in prevalent appearance between P. coquereli and L. catta lemurs. We see that Succinivibrionaceae, Muribaculaceae, Spirochaetaceae, Erysipelatoclostridiaceae, and Bacteroidaceae are notably more prevalent in L. catta, with the latter being the singular family with the highest

difference in prevalence between species. Rikenellaceae, Mongoglobaceae, Clostridia UCG-014, Clostridia VadinBB60, and Clostridiaceae are notably more prevalent in P. coquereli, the latter three all interestingly coming from the Clostridia class.

Methodology

The first step of our methodology was evaluating various aspects of the data to see what models are appropriate. One of our main concerns about the data is the fact that the observations are not fully independent. It is highly likely that observations from the same lemur are highly correlated with one another. Furthermore, there are 6 lemurs of each species (ring-tailed and sifaka) and 3 different enclosures housing 4 lemurs each. Lastly, all of the lemurs were observed at the Duke Lemur Center in Durham, North Carolina, so the lemurs experienced similar climate and weather (although that is out of the scope of this study). We know from previous studies that both environmental factors (enclosure) and biological factors (species and individual identity) have a significant association with lemur gut microbiome patterns, so we knew going into the study that the observations likely have some level of covariance with one another when it comes to predicting the response variables. Despite these independence violations, we elected to not run tests on them because the model we selected (described below) is designed to be robust for dependent observations.

Another part of our data that we needed to consider is the fact that repeated measurements were conducted via a time series; for each lemur, we have many observations at different points in time, in both the spring and summer seasons. Prior to modeling, we were unaware if we could expect a significant difference in gut microbiome patterns of lemurs based on time of day, so we wanted to select a model that could analyze a possible association for time of day, day, and season of the measurement while accounting for the known existing sensitivities of microbiome patterns to biological and environmental factors.

Time Series

To examine the time series data, we wanted to evaluate the general trends of the data within days and between days using a statistical test. Primarily, we wished to observe if there were any general trends with our response variables "in general" depending on the time of day or day the study was conducted. To do this, we elected to perform a series of Augmented Dickey-Fuller (ADF) tests on our three response variables (Faith's PD, Shannon entropy, and number of observed features) to test the stationarity of the data (specifically, if the mean or variance of the three response variables changes throughout days or between days.) The results of the ADF tests on a within-day scale showed strong indication that throughout the day, regardless of the day the observation was measured or the individual identity of the lemur, the mean and variance of all three response variables remain stationary. While this is a crude analysis, it is good background information to have when examining the possible associations of time with gut microbiome patterns when considering other predictors. All of the ADF tests were performed with a k-lag value of 25.

Table 1: p-values for ADF Tests Within Days

Faith's PD	Shannon Entropy	Observed Features
0.044	0.047	0.037

We can see from **Table 1** that all of the p-values are significant using a significance level of $\alpha = 0.05$, indicating that across the board (meaning across all seasons and individual lemurs), the distributions of all three response variables remain fairly constant throughout the day.

Table 2: p-values for ADF Tests Between Days

Faith's PD	Shannon Entropy	Observed Features
0.044	0.047	0.037

Table 2 shows the results for the between-days ADF tests. As we can see, when not accounting for other predictors, the distributions of our three response variables between days are also deemed stationary by the ADF test when using a k-lag value of 25. This is interesting considering we saw different distributions for the three response variables across seasons (which is a large gap, or lag, in days) for the L. catta species during our exploratory data analysis. However, overall, the distribution of the three response variables across days appeared stationary. While the ADF tests show general stationarity of all three response variables during days and across days, the tests themselves fail to distinguish between the differences in gut microbiome distributions associated with a different individual identity, lemur enclosure, and/or season. Our models will take these predictors into account and attempt to discover if the time of the observation (fixed effect) is associated with differences in lemur gut microbiome patterns while accounting for the individual identity of the lemur (random effect). It is possible that the aggregated data over all lemurs used for the ADF tests does not fully capture potential stationarity of the response variables due to variation between lemurs, and our model will take this concern into account when implemented.

Between-Group Correlation

To discover potential correlation between groups of lemurs, we chose to group the observations into 24 categories where each category indicates a specific lemur in one of the two seasons (this was done due to noticeable differences during exploratory data analysis involving L. catta lemurs between seasons). Using these groupings, we examined intraclass and interclass correlation of these groups using the intraclass correlation coefficient (ICC) and Spearman's correlation coefficient, respectively. Spearman's coefficient was selected to

evaluate possible monotonic nonlinearity for the three response variables (i.e. nonlinear relationships between classes). The results of the ICC indicated that for Faith's Phylogenetic Diversity, there was significant correlation within groups; this lines up with the findings of **Figure 3**, since the distributions for that response variable appear more spread out than the other response variables. For the other two response variables, insignificant intraclass correlation was observed. The Spearman's correlation matrices for the 24 groups in all 3 response variables did find a few instances of statistically significant correlation between the groups, but most of these supposed correlations were between groups involving two different lemurs (and often two different seasons). It is notable that after computing p-values to determine statistical significance of all of the pairwise correlations for the groups, Nikos and Sophia were the two lemurs that showed statistically significant correlation between the two groups involving themselves (e.g. when measuring the pairwise correlation of Faith's Phylogenetic Diversity scores, Nikos and Sophia were the only two lemurs to have significantly correlated response values across season). Both of these lemurs are L. catta lemurs, revealing that none of the P. coquereli lemurs exhibited statistically significant correlation of the response variables across seasons. While this is an indicator that the response variables being evaluated are correlated for these specific pairs of groups, the actual level of correlation for most pairs appears to be somewhat low (all of the off-diagonal pairwise correlations for all of the response variables were at or below an absolute correlation of 0.5).

Linear Mixed Effects Model

To investigate the remaining fixed effects, which are time of day and season, we implement three separate linear mixed models, each focusing on one of the microbiome diversity metrics as response variable: Faith's phylogenetic diversity, Shannon entropy, and the number of observed features of the different bacteria in the fecal sample. We are motivated to use linear mixed models because this dataset contains dependent observations at the individual animal level–namely, a lemur's gut microbiome diversity at one time point can be associated with its gut microbiome diversity in the future–for which linear mixed models can be more suitable than using dummy variables to encode for each individual lemur. Therefore, in our linear mixed model, we treat individual identity of each lemur as our random effect. Dr. Greene's prior work indicated that we should expect different patterns when examining microbiome composition in fecal sample based on species. We decided to first build a base model including species as the fixed effect and individual identity of the lemur as random effect. This separation of random and fixed effects automatically addresses the issue of non-independent observations. The model can be specified as:

$$
\begin{aligned}
Y_i &= a_i + b_i * \text{Species} + \epsilon_i \\
a_i &= \alpha_0 + u_i \\
b_i &= \beta_0 + v_i \\
\epsilon_i &\sim N(0, \sigma^2)
\end{aligned}
$$

Here we choose to only incorporate species as the fixed effect, since species and Group variables are correlated (social groups Beatrice, Gertrude and Gisela belong to P. coquereli lemurs whereas Alena, Sophia and Brigitta belong to L. catta lemurs), and incoporating a two-level species will make our model output look cleaner, especially after we expand the model to add more fixed effects terms. In the model, a_i indicates the random effect from each individual, with α_0 being the mean for the reference species and the reference group, and β_0 being the change towards the mean for the specific level group, compared to the reference group. σ^2 refers to the within-individual-lemur variance, while between-lemur variance is broken down into u_i and v_i. In addition, u_i and v_i are assumed, by convention, to have a bivariate normal distribution, with each centered around 0 with their own variance and covariance.

To investigate potential associations of time of day and season with gut microbiome patterns, we will adopt a nested-model approach and add these two covariates on top of the base model. Our EDA scatter plots also suggest that the association due to species may be different across time of day and season, hence interaction terms between species/season and species/time of day are added as well. Our samples span from spring to summer, meaning season covariates will be two-level. We also decide to use two-level AM/PM categorization for time of day. Unlike in our EDA, we didn't choose *minutes after 6AM* to operationalize time of day. Doing so would automatically assume that the response has a monotonic increasing/decreasing association with time elapsed; as we will have a single coefficient estimate for time of day, and our EDA doesn't suggest the presence of a monotonic association, we decided to use this binary metric instead. Finally, we elected not to include the date in our model due to the fact that each individual lemur is always measured in a period of three consecutive days. Since there is a large time gap between the spring and summer measurements, including the day of the observation would hurt model interpretation more than it would help.

Our nested model can therefore be specified as:

$$
\begin{aligned}
Y_{ij} &= a_i + b_i * \text{Species} + c_i * \text{AMPM} + d_i * \text{Season} + \epsilon_{ij} \\
a_i &= \alpha_0 + u_i \\
b_i &= \beta_0 + v_i \\
c_i &= c_0 + c_1 * \text{Species} + k_i \\
d_i &= d_0 + d_1 * \text{Species} + q_i \\
\epsilon_{ij} &\sim N(0, \sigma^2)
\end{aligned}
$$

and the error terms associated with each term, namely u_i, v_i, k_i, p_i are assumed to be multivariate normally distributed with mean centered at zero, and each has their own variance and covariance with other error terms.

Then, we perform chi-square difference tests on the two models to investigate if there is a statistically significant improvement from our base model to our new model, with time of day and season included as new fixed effects.

Results

Linear Mixed Effects Model Output

As described in our methodology, three separate linear mixed effects models are fitted for Shannon Entropy, observed features and Faith's Phylogenetic Diversity. The models are fitted using the `lme4` package in R.

Table 3: Fixed Effects for Shannon Base Model

	Estimate	Std..Error	t.value
Intercept	6.526	0.092	70.618
P.coquereli	-0.031	0.101	-0.308

Table 4: Fixed Effects for Features Base Model

	Estimate	Std..Error	t.value
Intercept	242.680	6.256	38.792
P.coquereli	-30.122	7.145	-4.216

Table 5: Fixed Effects for Faith Base Model

	Estimate	Std..Error	t.value
Intercept	16.037	0.526	30.493
P.coquereli	-3.116	0.558	-5.586

Table 3 shows the coefficients of the fixed effects for our base model, with species and individual identity of the lemur. As expected, lemur species is associated with changes in both observed features and phylogenetic diversity (both with $|t - value| > 2$) in the base model, while it is insignificant for Shannon Entropy. We can also draw conclusions on the variability within and between individual lemurs from the model output. Using our Shannon Entropy base model as an example, within-individual variance (σ^2) is 0.1, while between-lemur variance is only 0.04 ($\sigma^2_{u_i}$). However, it is worth noting that the associated standard deviations for these variance estimations are quite large (0.22 and 0.29 respectively), so the interpretations for within-individual and between-individual variability for the three base models should be treated with caution.

With the base model set up, we proceed to build three additional models with time of day, season, and interactions of time of day and season with species included. While we will interpret the Shannon Entropy model in detail, we will point out only the significant results in the observed features and phylogenetic diversity model (as they are all interpreted similarly). All models are interpreted controlling for the random effect from individuals to avoid repetition. For the observed features model, we recognized that the model would yield estimates that give trailing digits despite the response variable being integer-only. For the purpose of understanding trends of associations between explanatory variables of interest and the response variable, we believe it is still informative to interpret the estimated model coefficients.

Table 6: Fixed Effects for Shannon Nested Model

	Estimate	Std..Error	t.value
Intercept	6.430	0.104	61.853
Summer	0.252	0.110	2.295
P.coquereli	0.025	0.119	0.212
PM	-0.154	0.063	-2.432
P.coquereli * PM	0.116	0.087	1.342
P.coquereli * Summer	-0.105	0.153	-0.682

Table 7: Fixed Effects for Features Nested Model

	Estimate	Std..Error	t.value
Intercept	224.094	8.851	25.319
Summer	31.649	9.799	3.230
P.coquereli	-12.287	12.721	-0.966
PM	-1.285	4.680	-0.275
P.coquereli * PM	-1.893	6.362	-0.298
P.coquereli * Summer	-26.022	13.690	-1.901

Table 8: Fixed Effects for Faith Nested Model

	Estimate	Std..Error	t.value
Intercept	15.136	0.671	22.562
Summer	1.216	0.502	2.423
P.coquereli	-2.375	0.879	-2.703
PM	0.281	0.152	1.845
P.coquereli * PM	-0.286	0.206	-1.391
P.coquereli * Summer	-0.840	0.705	-1.191

- $\hat{\alpha}_0 = 6.430$: We expect the Shannon Entropy mean for L.catta species in the spring measured during the AM to be 6.430.
- $\hat{\beta}_0 = 0.025$: We expect Shannon Entropy to be 0.025 higher on average for P. coquereli than L. catta, after controlling for time of day and season.
- $\hat{c}_0 = -0.154$: We expect Shannon Entropy to be 0.154 lower on average in the PM than the AM, after controlling for species and season.
- $\hat{d}_0 = 0.252$: We expect Shannon Entropy for L. catta to be 0.252 higher on average in the summer than in the spring, after controlling for time of day.
- $\hat{c}_1 = 0.116$: For P. coquereli species, we expect Shannon Entropy to be 0.116 higher on average when measured in the PM than the AM, after controlling for season. In other words, the estimated Shannon Entropy measured in PM for P. coquereli is on average $|0.025 + 0.116 - 0.154| \approx 0.01$ lower than the Shannon Entropy measured in PM for L. catta, controlling for season.

- $\hat{d}_1 = -0.105$: For P. coquereli species, we expect Shannon Entropy to be 0.105 lower on average when measured in summer than the spring, after controlling for time of day. This means that, on average, the estimated Shannon Entropy measured in summer for P.coquereli is $|0.025 + 0.252 - 0.105| \approx 0.17$ higher than the Shannon Entropy measured in summer for L. catta, controlling for time of day.

In our model with observed features as response variable, we noticed this interesting coefficient sign change compared to the shannon entropy model:

- $\hat{\beta}_0 = -12.28$: The estimated observed features is 12.28 lower on average for P. coquereli than L. catta, after controlling for time of day and season.

In our model with Faith's PD as the response variable, we noticed these interesting coeffcient sign changes compared to the shannon entropy model:

- $\hat{\beta}_0 = -2.37$: The estimated Faith's PD is 2.37 lower on average for P. coquereli than L. catta, after controlling for time of day and season.
- $\hat{c}_0 = 0.28$: The estimated Faith's PD is 0.28 on average higher in PM than AM, after controlling for species and season.
- $\hat{c}_1 = -0.28$: For P. coquereli species, the estimated Faith's PD is 0.28 lower when measured in PM than AM, after controlling for season, which is similar in terms of magnitude to the association described by $\hat{c}_0$. In other words, the estimated Faith's PD measured in PM for P. coquereli is on average $|-2.37 + 0.28 - 0.28| = 2.37$ lower than the Faith's PD measured in PM for L. catta, controlling for season.

Since our new models are nested model compared to our base model, which only includes species and individual lemur ID, we carried out a Chi-square difference test to see whether the new model performs significantly better than the base model.

Judging from the p-values, we can conclude that our new models for all three response variables perform significantly better than the base model. All the three new models experience significant drops in AIC, BIC and deviance (except for the new shannon entropy model, which actually experiences a slight increase in BIC). Considering that BIC also penalize for overly complicated models, this comparison suggests evidence that the models with predictors other than species may be better at capturing underlying relationships within our current dataset.

Discussion

Linear Mixed Model Interpretation

It's worth noting that the t values associated with the fixed effects in our new models are mostly between 1-2, and their p values are not provided with them as the underlying distribution for these fixed effect coefficients are unknown for a linear mixed effect model. Therefore, we can't conclude the significance of these individual effects from a traditional confidence level perspective; however, these fixed coefficients still suggest the trend of how response variable changes as our predictor varies. Our model diagnostics and interpretation can be found in the Appendix: overall, we didn't observe significant deviation that questions the choice of Linear Mixed Models as our inference tool for this report. It's also worth noting that we only have a total of 293 observations for all 12 lemurs, with each lemur only accounting for twenty or so data points: it's likely that if more fecal samples were collected for these lemurs, we would have larger statistical power in determining our coefficients with reduced standard deviation, as their associated t values would be larger as well.

In summary, we learned from the models that we would expect Shannon Entropy and Observed Features to go down on average when measured in later times of the day because of the negative association due to time of day, controlling for species and season. Comparing the two seasons, we would expect all three markers to be higher on average when measured in summer compared to spring, meaning the lemurs would have a more diverse and balanced gut microbiome overall in summer, controlling for species and time of day. In terms of the magnitude of coefficient values, association due to season is also the "strongest" variable in our Shannon Entropy and Observed Features Nested Models, after controlling for time of day and species. In the Result section, we also interpreted when we would expect P.coquereli to have a more proportionally balanced microbiome than L.catta on average, controlling for season. Our model estimates provide support for the potential value of investigating relevant biological factors in lemurs, as well as the environmental variables related to their habitats, which can vary with time of day and season. Such an inquiry may lead to a more comprehensive understanding of the complex relationship between these factors and the lemur microbiome.

In future research, a valuable avenue to explore is geospatial analysis, which can provide critical insights into the interplay between the physical environment and lemur gut microbiome diversity. This analysis involves the use of geographic information systems (GIS) to map enclosure locations and assess the impact of landscape characteristics on microbiome composition. By considering factors such as elevation, vegetation type, and proximity to water sources, we can elucidate how specific geographic features may be linked to observed variations in microbial communities. The integration of geospatial analysis will enhance our understanding of the ecological factors influencing lemur health and microbiota, offering a more holistic view of their habitat and its microbial dynamics.

Appendix

Data Dictionary

- **Shannon Entropy:** Describes the microbial evenness of the taken sample - higher degrees of uniformity among the bacterial data is associated with a higher Shannon Entropy
- **Faith's Phylogenetic Diversity:** Describes the overall diversity of the sample - similar to Shannon Entropy, the more uniform the bacterial data appears, the more diverse the sample, and consequently the higher the Faith's Phylogenetic Diversity Score
- **Observed Features:** The number of different bacteria observed in the taken sample

Variables included in the random effects models are defined below:

- **Season:** Defines the season the sample was taken in - either spring or summer
- **Animal.ID:** Defines the individual identity of the lemur via their name - only 12 unique lemurs measured
- **minutes_after_6am:** Defines the number of minutes after 6AM the sample was taken on the given day
- **AMPM:** Defines the two time categories derived from minutes_after_6am - AM (6:00am-11:59am) and PM (12:00pm-6:20pm)
- **Species:** Defines the species of the sampled lemur - either L. catta or P. coquereli

Methodology Test Information

Table 9: Significant Correlation Between Groups (Faith PD)

Pair 1	Pair 2	Correlation
Alena summer	Rupert summer	0.0210
Brigitta summer	Elliot summer	0.0431
Brigitta summer	Gisela summer	0.0462
Elliot spring	Gertrude spring	0.0345
Elliot spring	Gertrude summer	0.0080
Elliot spring	Nikos summer	0.0093
Elliot spring	Stewart summer	0.0115
Elliot summer	Nikos spring	0.0467
Gertrude spring	Remus spring	0.0436
Gertrude spring	Remus summer	0.0152
Gisela spring	Randy spring	0.0020
Gisela spring	Sophia summer	0.0141
Nikos summer	Rupert summer	0.0481
Nikos summer	Stewart summer	0.0170
Randy summer	Sophia summer	0.0064
Rupert summer	Stewart summer	0.0003
Sophia summer	Stewart spring	0.0294

Table 10: Significant Correlation Between Groups (Shannon Entropy)

Pair 1	Pair 2	Correlation
Alena spring	Beatrice spring	0.0118
Alena summer	Gertrude summer	0.0344
Alena summer	Remus summer	0.0121
Beatrice spring	Elliot summer	0.0234
Beatrice summer	Sophia spring	0.0070
Brigitta spring	Gisela spring	0.0017
Brigitta spring	Rupert spring	0.0060
Brigitta spring	Stewart summer	0.0147
Brigitta summer	Remus spring	0.0262
Elliot spring	Rupert summer	0.0056
Elliot summer	Nikos summer	0.0093
Elliot summer	Randy summer	0.0172
Gertrude spring	Remus spring	0.0218
Gertrude spring	Stewart spring	0.0092
Gertrude summer	Rupert spring	0.0060
Gertrude summer	Sophia spring	0.0234
Gertrude summer	Stewart summer	0.0004
Nikos spring	Randy spring	0.0049
Rupert spring	Sophia spring	0.0195
Rupert spring	Stewart summer	0.0001
Rupert summer	Stewart summer	0.0067
Sophia spring	Stewart summer	0.0087
Stewart spring	Stewart summer	0.0485

Table 11: Significant Correlation Between Groups (Observed Features)

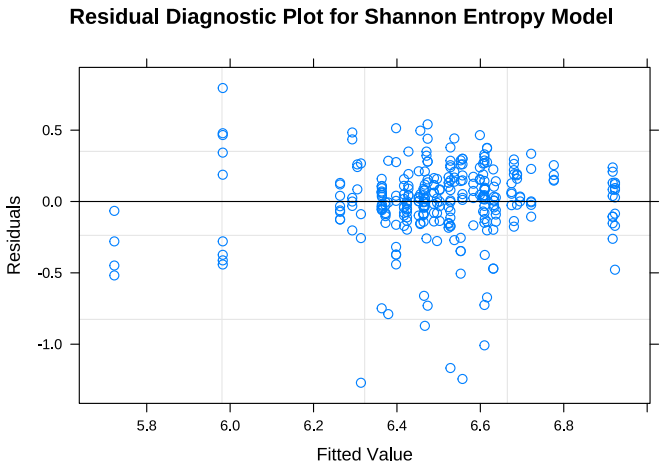
Pair 1	Pair 2	Correlation
Alena spring	Gisela summer	0.0189
Alena spring	Randy summer	0.0236
Alena spring	Remus summer	0.0039
Beatrice summer	Elliot spring	0.0288
Brigitta summer	Gisela summer	0.0457
Brigitta summer	Nikos summer	0.0007
Elliot summer	Randy summer	0.0396
Gertrude spring	Remus summer	0.0257
Gertrude summer	Stewart summer	0.0180
Gisela summer	Remus summer	0.0446
Nikos summer	Stewart summer	0.0295
Randy spring	Rupert summer	0.0258

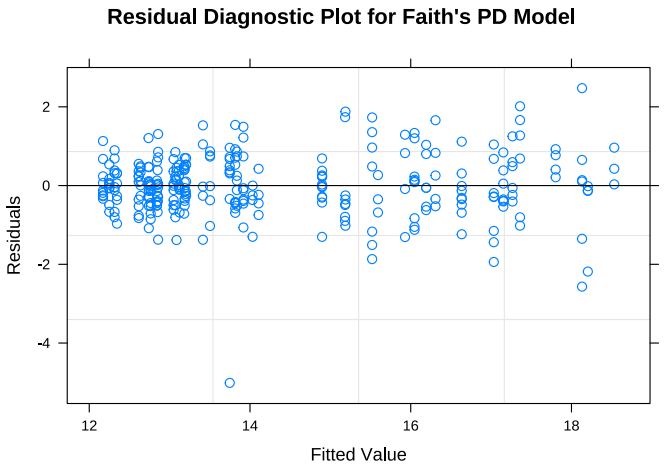
We can see from the above graphs that although the correlation was deemed to be statistically significant between the two groups, none of the correlations between the (lemur, season) pairs are actually strong. In fact, we see next to no correlation between nearly all of the significant correlative pairs.

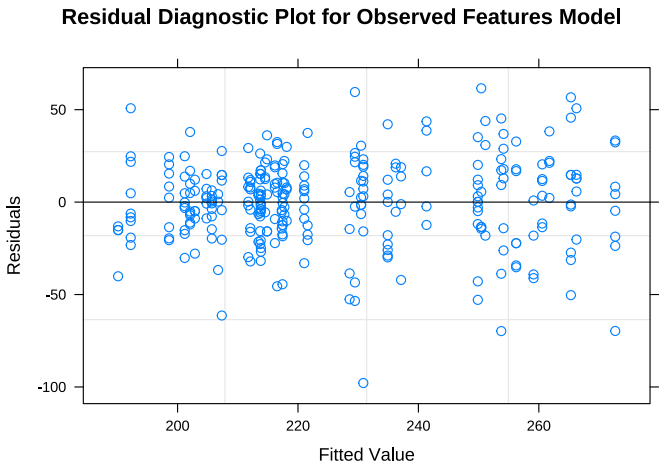
Model Diagnostics Plots

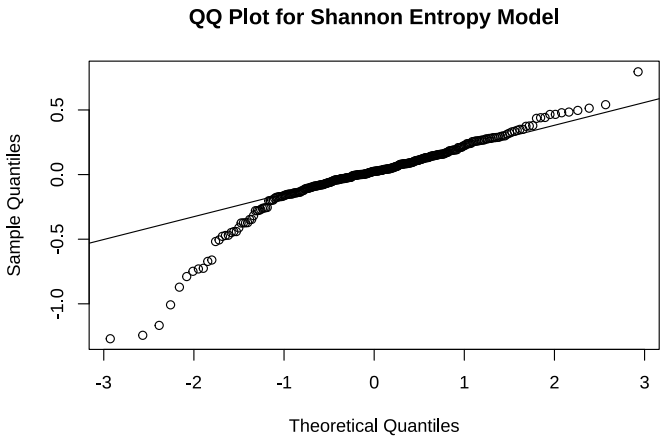
Based on the plot where we graph residuals vs fitted value, there is no considerable sign of heteroscedasticity in both

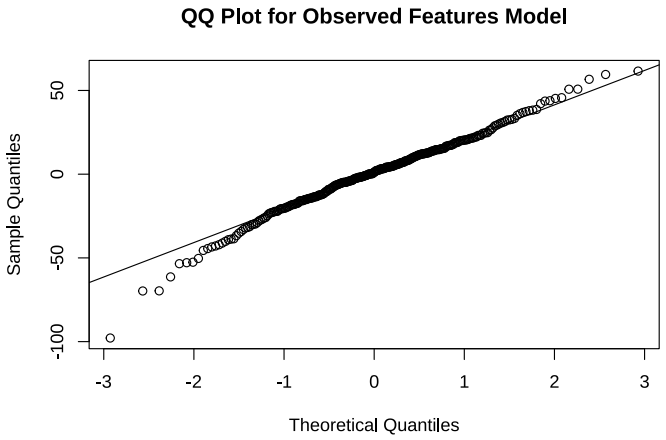
Shannon Entropy Model and Features Model. However, we did notice a slight fanning structure in the residual plot of Faith's PD Model.In terms of residual normality, although all three models have falling tails on both ends, there is no substantial deviation across all models.

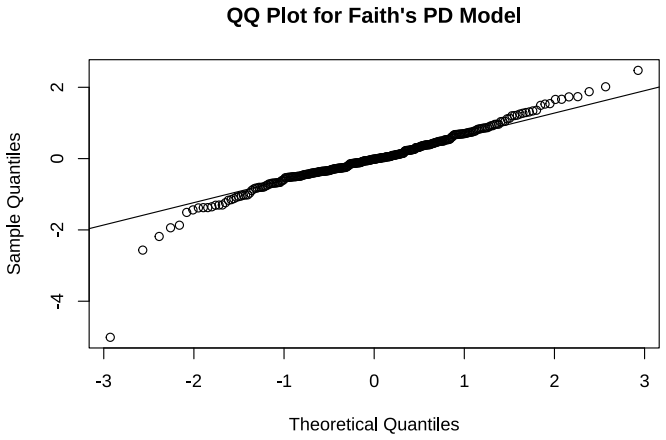












Sources Cited

1. https://bookdown.org/roback/bookdown-BeyondMLR/ch-multilevelintro.html
2. https://www.amnh.org/exhibitions/permanent/human-origins/understanding-our-past/living-primates#
3. Bornbusch, S.L., Greene, L.K., Rahobilalaina, S. et al. *Gut microbiota of ring-tailed lemurs (Lemur catta) vary across natural and captive populations and correlate with environmental microbiota.* anim microbiome 4, 29 (2022). https://doi.org/10.1186/s42523-022-00176-x

Response to Comments

Introduction

- In the introduction, we added background information on the three response variables (Shannon Entropy, Faith's PD, and the Number of Observed Features) in response to your comment about a lack of clarity regarding the meaning of these variables - this information was originally in the Appendix, but it is now in the introduction.
- You commented that causal statements are a weak point of the paper. We removed a research question in the introduction that would require causal statements to explain (specifically, the question regarding whether the response variable in question is affected by one variable more than another). In our opinion, removing this question from the study strengthens our analysis by reducing the scope of the study.

Methodology

- We reworded sentences in both the introduction and methodology (page 1-3) to avoid implying causal statements, which was a repeated concern shown in the comments.
- You noted that we should include fewer decimal points in the reporting of our results. We reformatted our tables and rounded output to 3 significant digits after the decimal point to be consistent with the formatting style in the result section.
- We fixed our LaTeX notation and added text explanations of some terms that are not directly explained. Hopefully, this makes it more clear that we are treating individual lemur as the random effect in our mixed model. This point is elaborated below.
- Your main comment about our methodology was regarding the use of a two-level random effects hierarchy of individual lemur nested within species. While we tried to see where we used this approach, we failed to find it - we found that our random effects only had one level, which was the individual identity of the lemur. Therefore, we elected to make no changes to our study from this comment.

Results and Discussion

- You noted that while our discussion was strong, it could have flowed better by organizing the text appropriately. We moved our summary of our findings after the discussion of limitation in the discussion section so that it now ends with takeaways from this report as well as potential future direction, instead of ending on a note where we criticize our own paper's findings. This organization leaves the reader with a reminder of the findings of our study.
- We cleaned some table columns in the appendix for general better report appearance. This was not done in response to any particular comment but some aspects of the appendix still required improvement after our intial submission.