The Impact of Added Non-Protein Nitrogen on the Rumen Microbiome and Fermentation of Substrates

Mia Carmichael

2019-11-13

# Summary/Abstract

*Write a summary of your project.*

# Introduction (required for part 1)

## General Background Information

The symbiotic relationship between ruminants and their ruminal microbial population allows cattle to occupy an ecological niche in which they have the ability to convert fibrous feedstuffs into higher quality end products via fermentation. Ruminants themselves do not produce fiber degrading enzymes, but they harbor bacteria, fungi, and protozoa that can (Russel, 2001). Ruminal Microbes utilize nitrogen and carbon sources from ingested feedstuff to carryout bilogical processes, and create end products such as volitile fatty acids and microbial crude protein vital to cattle. It is due to this that it is of interest to identify optimal levels and types of non-protein nitrogen and carbon sources that are crucial to the rumen microbiome and cattle to increase efficiency. Previous studies have identified different results of nitrogen and carbon utlization, but few have also analyzed it in conjunction with the analysis of the rumen microbial population.

## Description of data and data source

This will be an in vitro study. Rumen fluid will be collected from cattle at the animal units at UGA. There will be 6 treatment groups and three replicates. We will analyze rumen fluid with corn without added NPN, rumen fluid with corn with added NPN, rumen fluid with bermuda grass and added NPN,rumen fluid with bermuda grass and no NPN, rumen fluid without additional feed or NPN, and rumen fluid without additional feed and with added NPN. We will use whole genome sequencing to aquire the rumen microbial profiles before and after treatments, pH, methane oduction, volitile fatty acid production,and ammonia production.

## Questions/Hypotheses to be addressed

In my project I am trying to answer what impact does added non protein nitrogen have on the rumen microbiome population and its ability to ferment corn and bermuda grass. It will be interesting to see the trends and how the microbial profile and end products change as the nutrient availability changes. In the rumen there are multiple classes of microbes, including fiber degrading, starch degrading, and methane producing bacteria. As the feedstuff changes, the abundance of specific microorganisms will as well, which dictate other outcomes as well, pH, VFA profile, ammonia production, and methane production. The implications of this study can aid in further manipulation of the diet and cow to improve feed efficiency, production costs, and envirornmental outputs.

# Methods and Results

*In most research papers, results and methods are separate. You can combine them here if you find it easier. You are also welcome to structure things such that those are separate sections.*

## Data aquisition

The data described for my next research project is still being gathered, and will be ready for the analysis for this class project. Until then, I will be using data I collected from my last research project, chapter two of my thesis. Data was collected from 12 feedlot steers day one entering the feedlot and day 72, exiting the feedlot.

## Data import and cleaning

Here I have imported and cleaned data from the microbiome (phylum, family, and genus) of High and Low RFI steers, as well as their alpha diversity.

Processing Script: This script loads the raw data, processes and cleans it and saves it as Rds file in the processed\_data folder ` Load needed packages. make sure they are installed.

library(readxl)  
library(readr)  
library(ggplot2)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(here)

## here() starts at /Users/miacarmichael/Desktop/epid8060fall2019/MiaCarmichael-Project

library(visdat)  
library(dplyr)  
library(tidyverse)

## ── Attaching packages ─────────────── tidyverse 1.2.1 ──

## ✔ tibble 2.1.3 ✔ purrr 0.3.2  
## ✔ tidyr 1.0.0 ✔ stringr 1.4.0  
## ✔ tibble 2.1.3 ✔ forcats 0.4.0

## ── Conflicts ────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

#Phylum

Load data, path is relative to project directory. Select a file and read the data into a data frame.

phyladata <- read\_csv(here("data", "raw\_data", "Abundant\_Phyla\_MADA.csv"))

## Parsed with column specification:  
## cols(  
## `Bacterial Phyla` = col\_character(),  
## Beginning = col\_double(),  
## End = col\_double(),  
## `P-value` = col\_double()  
## )

Take a look at the data

head(phyladata)

## # A tibble: 6 x 4  
## `Bacterial Phyla` Beginning End `P-value`  
## <chr> <dbl> <dbl> <dbl>  
## 1 Firmicutes 35.0 55.1 0.02  
## 2 Bacteroidetes 48.7 33.0 0.12  
## 3 Actinobacteria 7.21 1.29 0.34  
## 4 TM7 1.81 2.06 0.85  
## 5 Proteobacteria 1.48 2.31 0.43  
## 6 Euryarchaeota 1.28 1.39 0.83

dplyr::glimpse(phyladata)

## Observations: 13  
## Variables: 4  
## $ `Bacterial Phyla` <chr> "Firmicutes", "Bacteroidetes", "Actinobacteria…  
## $ Beginning <dbl> 34.95, 48.73, 7.21, 1.81, 1.48, 1.28, 1.31, 0.…  
## $ End <dbl> 55.14, 33.04, 1.29, 2.06, 2.31, 1.39, 0.81, 0.…  
## $ `P-value` <dbl> 0.02, 0.12, 0.34, 0.85, 0.43, 0.83, 0.42, 0.74…

There are only 11 identified Phylum in this data, but there are currently 13 observations. Remove any observations with NA

phyladata <- drop\_na(phyladata)  
dim(phyladata)

## [1] 11 4

Rename the columes, so when creating a graph things can be plotted easier

names(phyladata)[1:4]<-c("Phyla", "Beginning", "End", "P")

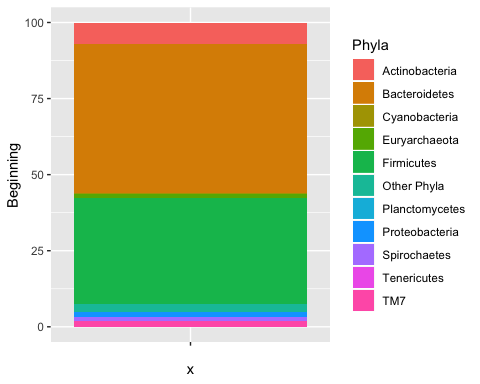
Check name of columns to see what data looks like.

names(phyladata)

## [1] "Phyla" "Beginning" "End" "P"

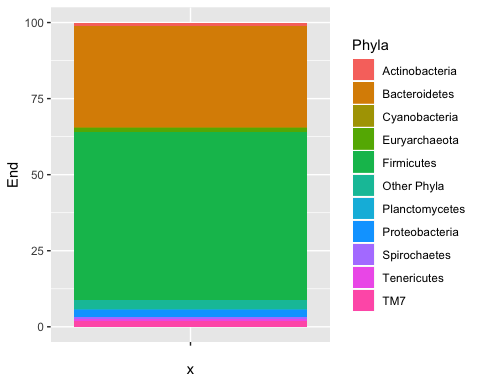
An initial plot is created showing the composition of phyla at the beginning of the feedlot period in High RFI/low efficiency steers. Save plot to a file.

ggplot(phyladata,aes(x="", y=Beginning, fill=Phyla))+ geom\_bar(width = 1, stat = "identity")



A second plot is created showing the composition of families at the end of the feedlot period in High RFI/low efficiency steers.Save the plot to a file.

ggplot(phyladata,aes(x="", y=End, fill=Phyla))+ geom\_bar(width = 1, stat = "identity")



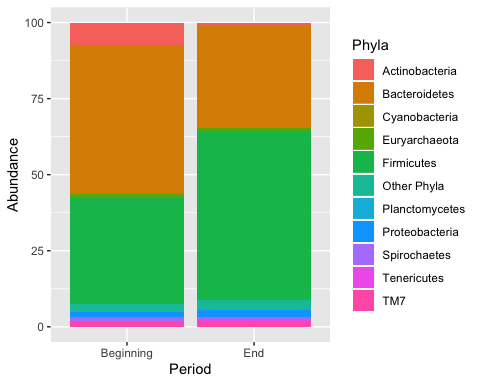
The axis labels aren’t quite right for both graphs. That should be considered when creating them.

It is easier to see these plots side by side. To do this, create subsets to divide by period, “Beginning” and “End”.

#Create datasets   
  
Period <- c(rep("Beginning",11), rep("End",11))  
  
Phyla <- rep(c("Firmicutes", "Bacteroidetes",   
"Actinobacteria", "TM7", "Proteobacteria", "Euryarchaeota", "Spirochaetes", "Cyanobacteria",   
"Tenericutes", "Planctomycetes", "Other Phyla"),2)   
  
Abundance <- c(34.95, 48.73, 7.21, 1.81, 1.48, 1.28, 1.31, 0.39, 0.16, 0.21, 2.46, 55.14, 33.04, 1.29, 2.06, 2.31, 1.39, 0.81, 0.33, 0.46, 0.27, 2.91)  
  
  
   
#Combine datasets   
phylacomb <- data.frame(Period, Phyla, Abundance)  
  
#Remove missing data  
phylacomb <- drop\_na(phylacomb)  
dim(phylacomb)

## [1] 22 3

#Graph combined dataset  
ggplot(phylacomb, aes(fill=Phyla, y=Abundance, x=Period)) +  
 geom\_bar(stat = "identity")

 The graph above displays the composition of phyla over the course of the feedlot period in High RFI/low efficiency steers Repeat for Family and Genus

#Family Load data, path is relative to project directory. Select a file and read the data into a data frame.

familydata <- read\_csv(here("data", "raw\_data", "Abundant\_FamiliesHighRFI\_MADA.csv"))

## Parsed with column specification:  
## cols(  
## Family = col\_character(),  
## Beginning = col\_double(),  
## End = col\_double(),  
## `P-VALUE` = col\_double()  
## )

Take a look at the data

head(familydata)

## # A tibble: 6 x 4  
## Family Beginning End `P-VALUE`  
## <chr> <dbl> <dbl> <dbl>  
## 1 Ruminococcaceae 26.0 40.3 0.17   
## 2 Prevotellaceae\* 36.9 14.3 0.05   
## 3 Order Clostridiales 5.93 8.40 0.08   
## 4 Order Bacteroidales 5.94 7.49 0.15   
## 5 Lachnospiraceae 5.19 7.70 0.290  
## 6 RF16 1.94 1.36 0.14

dplyr::glimpse(familydata)

## Observations: 18  
## Variables: 4  
## $ Family <chr> "Ruminococcaceae", "Prevotellaceae\*", "Order Clostridi…  
## $ Beginning <dbl> 25.9771285, 36.9109508, 5.9340251, 5.9409425, 5.189480…  
## $ End <dbl> 40.3397858, 14.3071369, 8.4018357, 7.4860880, 7.701011…  
## $ `P-VALUE` <dbl> 0.17, 0.05, 0.08, 0.15, 0.29, 0.14, 0.01, 0.09, 0.49, …

There are only 18 identified Phylum in this data, but there are currently observations. Remove any observations with NA

familydata <- drop\_na(familydata)  
dim(familydata)

## [1] 18 4

Rename the columes, so when creating a graph things can be plotted easier

names(familydata)[1:4]<-c("Family", "Beginning", "End", "P")

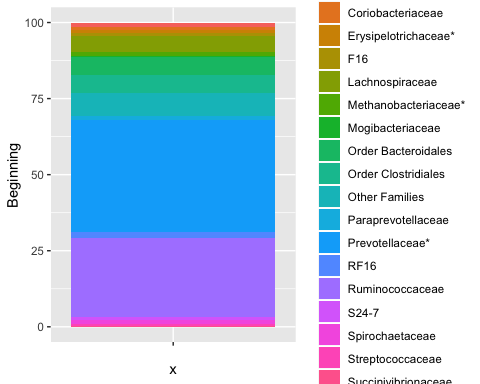
Check name of columns to see what data looks like.

names(familydata)

## [1] "Family" "Beginning" "End" "P"

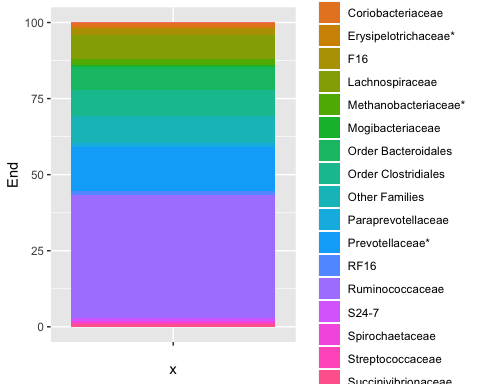
An initial plot is created showing the composition of families at the beginning of the feedlot period in High RFI/low efficiency steers. Save plot to a file.

ggplot(familydata,aes(x="", y=Beginning, fill=Family))+ geom\_bar(width = 1, stat = "identity")



A second plot is created showing the composition of families at the end of the feedlot period in High RFI/low efficiency steers.Save the plot to a file.

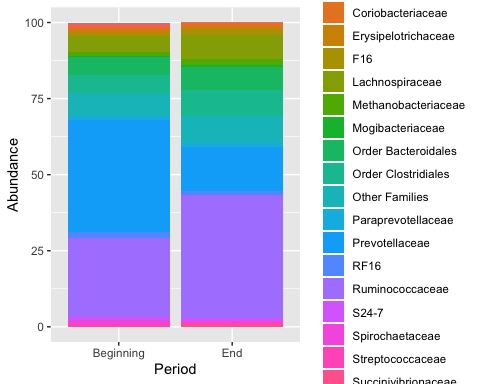
ggplot(familydata,aes(x="", y=End, fill=Family))+ geom\_bar(width = 1, stat = "identity")



The axis labels aren’t quite right for both graphs. That should be considered when creating them.

It is easier to see these plots side by side. To do this, create subsets to divide by period, “Beginning” and “End”.

#Create datasets   
  
Period <- c(rep("Beginning",18), rep("End",18))  
  
Family <- rep(c("Ruminococcaceae", "Prevotellaceae", "Order Clostridiales", "Order Bacteroidales", "Lachnospiraceae", "RF16", "Methanobacteriaceae", "F16", "Paraprevotellaceae", "S24-7", "Coriobacteriaceae", "Erysipelotrichaceae", "Bifidobacteriaceae", "Spirochaetaceae", "Streptococcaceae", "Mogibacteriaceae", "Succinivibrionaceae", "Other Families"),2)   
  
Abundance <- c(25.97712848, 36.91095082, 5.934025064, 5.940942506, 5.189480948, 1.937621255, 1.235079234, 1.049483052, 1.214503215, 0.956874771, 0.860283125, 1.158815891, 1.451193422, 1.135264464, 0.802975589, 0.476774167, 0.255410691, 7.513193298, 40.33978584, 14.30713689, 8.401835665, 7.486088, 7.701011161, 1.357351894, 2.004700597, 2.111417947, 1.348409159, 1.177308762, 1.058027748, 0.733590889, 0.311866263, 0.47494151, 0.461737838, 0.773816017, 0.916878216, 9.034095609)  
   
#Combine datasets   
familycomb <- data.frame(Period, Family, Abundance)  
  
#Graph combined dataset  
ggplot(familycomb, aes(fill=Family, y=Abundance, x=Period)) +  
 geom\_bar(stat = "identity")

 The graph above displays the abundance of families over the course of the feedlot period in High RFI/low efficiency steers

#Genus For my dataset, there was litte variation over the course of the feedlot amongst genra. Howver there were noticible changes amongst methanogen populations.

Load data, path is relative to project directory. Select a file and read the data into a data frame.

genusdata <- read\_csv(here("data", "raw\_data", "Genera\_MADA.csv"))

## Parsed with column specification:  
## cols(  
## `HIGH-RFI` = col\_character(),  
## Beginning = col\_double(),  
## End = col\_double(),  
## `P-VALUE` = col\_double()  
## )

Take a look at the data

head(genusdata)

## # A tibble: 2 x 4  
## `HIGH-RFI` Beginning End `P-VALUE`  
## <chr> <dbl> <dbl> <dbl>  
## 1 Methanobrevibacter 1.2 2 0.01  
## 2 Methanosphaera 0.01 0.03 0.09

dplyr::glimpse(genusdata)

## Observations: 2  
## Variables: 4  
## $ `HIGH-RFI` <chr> "Methanobrevibacter", "Methanosphaera"  
## $ Beginning <dbl> 1.20, 0.01  
## $ End <dbl> 2.00, 0.03  
## $ `P-VALUE` <dbl> 0.01, 0.09

Remove any observations with NA

genusdata <- drop\_na(genusdata)  
dim(genusdata)

## [1] 2 4

Rename the columes, so when creating a graph things can be plotted easier

names(genusdata)[1:4]<-c("Genus", "Beginning", "End", "P")

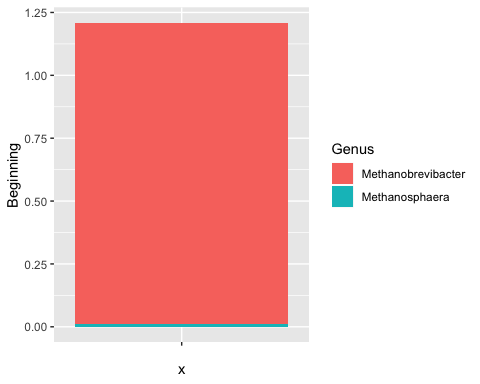
Check name of columns to see what data looks like.

names(genusdata)

## [1] "Genus" "Beginning" "End" "P"

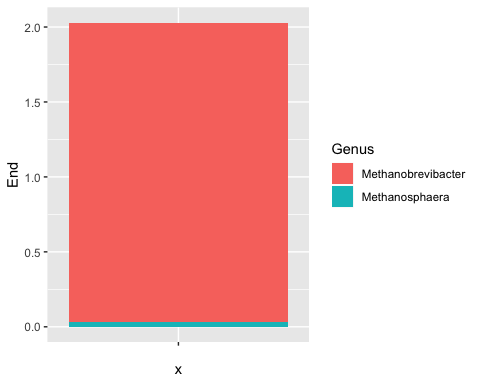
An initial plot is created showing the composition of two genera at the beginning of the feedlot period in High RFI/low efficiency steers. Save plot to a file.

ggplot(genusdata,aes(x="", y=Beginning, fill=Genus))+ geom\_bar(width = 1, stat = "identity")



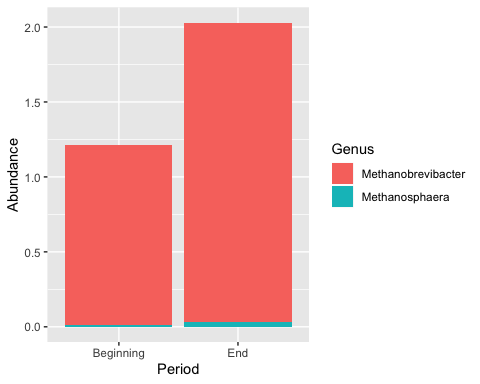
A second plot is created showing the composition of two genera at the end of the feedlot period in High RFI/low efficiency steers.Save the plot to a file.

ggplot(genusdata,aes(x="", y=End, fill=Genus))+ geom\_bar(width = 1, stat = "identity")



It is easier to see these plots side by side. To do this, create subsets to divide by period, “Beginning” and “End”.

Period <- c(rep("Beginning",2), rep("End",2))  
  
Genus <- rep(c("Methanobrevibacter",  
"Methanosphaera"),2)   
  
Abundance <- c(1.2, 0.01, 2.0, 0.03)  
   
#Combine datasets   
genuscomb <- data.frame(Period, Genus, Abundance)  
  
#Graph combined dataset  
ggplot(genuscomb, aes(fill=Genus, y=Abundance, x=Period)) +  
 geom\_bar(stat = "identity")



## Univariate analysis

*Use a combination of text/tables/figures to explore and describe your data. You should produce plots or tables or other summary quantities for most of your variables. You definitely need to do it for the important variables, i.e. if you have main exposure or outcome variables, those need to be explored. Depending on the total number of variables in your dataset, explore all or some of the others.*

## Bivariate analysis

*Create plots or tables and compute simple statistics (e.g. t-tests, simple regression model with 1 predictor, etc.) to look for associations between your outcome(s) and each individual predictor variable*

## Full analysis

*Use one or several suitable statistical/machine learning methods to analyze your data and to produce meaningful figures, tables, etc. This might again be code that is best placed in one or several separate R scripts that need to be well documented. You can then load the results produced by this code*

# Discussion

## Summary and Interpretation

*Summarize what you did, what you found and what it means.*

## Strengths and Limitations

*Discuss what you perceive as strengths and limitations of your analysis.*

## Conclusions

*What are the main take-home messages?*

*Include citations in your Rmd file using bibtex, the list of references will automatically be placed at the end*

# References