Alternate normalization level analyses

For review only

Kim Dill-McFarland

2018-10-11

Set seed

```
set.seed(8325)
```

Packages

```
#Data manipulation
library(tidyverse)
```

```
## -- Attaching packages ----------------------------------------------- tidyverse 1.2.1 --
```

```
## v ggplot2 3.0.0     v purrr   0.2.5
## v tibble  1.4.2     v dplyr   0.7.6
## v tidyr   0.8.1     v stringr 1.3.1
## v readr   1.1.1     v forcats 0.3.0
```

```
## -- Conflicts -------------------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
#Beta-diversity analyses and plots
library(vegan)
```

```
## Loading required package: permute
```

```
## Loading required package: lattice
```

```
## This is vegan 2.5-2
```

Data

OTU tables

Rows are samples, columns are OTUs, data are given in counts

```
#Fungi
OTUF = read_tsv("data/OTU_table/OTU.Fun.norm1000.txt") %>% rename(Sample=Samle) %>% separate(Sample, c(
```

Metadata by sample

- Animal: animal ID number
- Diet: calf starter (A), corn silage (B), mixture (C)
- AgeGroup: 2 weeks (2w), 4 weeks (4w), 8 weeks (8w), 1 year (1yr), 2 years (2yr)

- AgeExact: exact age in days on date of sampling
- Daysln: natural log of exact age
- TypeAll: feces (F), rumen liquids (L, no number for calves, 1-3 for day on repeated sampling for cows), rumen solids (S, no number for calves, 1-3 for day of consecutive sampling for cows)
- Type: feces (F), rumen liquids (L), rumen solids (S) (simplified to not include numbers for day)
- Sex: males (M), female (F)
- ScourSample: Y/N if sample was taken while calf was experiencing scours
- chao: Chao1 richness estimate
- shannon: Shannon's diversity estimate
- ADGKG: average daily gain in kg from 2 days of age to date of sampling
- Gain: total weight gain in kg relative to 2 day measuremnts
- TotalIntake: total supplement intake (g) from birth up to date of sampling
- ill.scour: total number of days ill with scours up to date of sampling
- ill.resp: total number of days ill with scours up to date of sampling
- ill.ketosis: total number of days ill with ketosis up to date of sampling
- ill.mastitis: total number of days ill with mastitis up to date of sampling
- ill.all: total number of days ill (any recorded illness) up to date of sampling
- MPE: milk production efficiency (energy corrected milk [ECM] / dry matter intake [DMI]) (kg/kg DMI)
- Effic: total energy output efficiency (MJ/kg DMI)

```
#Fungi
metaF = read_tsv("data/metadata/Met.Fun2.txt", col_types = list(Animal = col_character()))
```

Add Diet + Age variable

```
#Fungi
metaF$DietAge = as.factor(paste(metaF$Diet, metaF$AgeGroup, sep=""))
```

Average consecutive 1 & 2yr samples

To remove animal effects in some analyses, average the data from 1 and 2 year samples taken across 3 consecutive days.

OTU tables

Fungi

```
OTUFave = metaF %>%
  # Merge OTU with metadata needed for grouping 3 day replicates in OTU data
  select(Sample, Animal, AgeGroup, Type) %>%
  right_join(OTUF, by="Sample") %>%
  select(-Sample) %>%
  # average 3 day replicates for 1 and 2 yr rumen samples
  group_by(Animal, AgeGroup, Type) %>%
  summarize_if(is.numeric, funs(mean)) %>%
  # round to whole numbers
  mutate_if(is.numeric, funs(round(., 0))) %>%
  ungroup() %>%
  # recreate Sample name as row name
  mutate(Sample = paste(Animal, AgeGroup, sep=".")) %>%
```

```
  mutate(Sample = paste(Sample, Type, sep=".")) %>%
  select(-Animal, -AgeGroup, -Type)
```

Metadata by sample

Fungi

```
metaFave = metaF %>%
  # average 3 day replicates for 1 and 2 yr rumen samples
  group_by(Animal, AgeGroup, Type, Diet) %>%
  summarize_if(is.numeric, funs(mean)) %>%
  ungroup() %>%
  # recreate Sample name as row name
  mutate(Sample = paste(Animal, AgeGroup, sep=".")) %>%
  mutate(Sample = paste(Sample, Type, sep="."))
```

BETA-DIVERSITY

Average beta by age

Fungi

Bray-Curtis

```
BC.temp = OTUF %>%
  # Remove non-numeric labels
  select(-Sample,-amplicon) %>%
  vegdist(method="bray") %>%
  # Convert to data frame
  as.matrix() %>%
  as.data.frame() %>%
  # Label columns and rows by sample name
  mutate(Sample=OTUF$Sample) %>%
  rename_if(is.numeric, ~ OTUF$Sample) %>%
  # Add metadata
  full_join(metaF[,c("Sample","Animal","AgeGroup","Type")], by="Sample") %>%
  # Reformat to long
  gather(-Sample, -Animal, -AgeGroup, -Type, key="Sample2", value="BC") %>%
  # Remove self comparisons
  dplyr::filter(BC > 0) %>%
  # Calculate mean and stdev
  group_by(AgeGroup, Type) %>%
  summarize(mean=mean(BC, na.rm=TRUE), sd=sd(BC, na.rm=TRUE)) %>%
  mutate(amplicon="Fun", metric="BC")
```

Jaccard

```
J.temp = OTUF %>%
  # Remove non-numeric labels
  select(-Sample, -amplicon) %>%
  vegdist(method="jaccard") %>%
```

```
  # Convert to data frame
  as.matrix() %>%
  as.data.frame() %>%
  # Label columns and rows by sample name
  mutate(Sample=OTUF$Sample) %>%
  rename_if(is.numeric, ~ OTUF$Sample) %>%
  # Add metadata
  full_join(metaF[,c("Sample","Animal","AgeGroup","Type")], by="Sample") %>%
  # Reformat to long
  gather(-Sample, -Animal, -AgeGroup, -Type, key="Sample2", value="J") %>%
  # Remove self comparisons
  dplyr::filter(J > 0) %>%
  # Calculate mean and stdev
  group_by(AgeGroup, Type) %>%
  summarize(mean=mean(J, na.rm=TRUE), sd=sd(J, na.rm=TRUE)) %>%
  mutate(amplicon="Fun", metric="J")
```

Combine

```
write.table(rbind(BC.temp, J.temp), "results/beta_div/F.beta.txt", row.names=FALSE, sep="\t")
```

Combine all

```
beta.ave = read_tsv("results/beta_div/A.beta.txt") %>%
  rbind(read_tsv("results/beta_div/B.beta.txt")) %>%
  rbind(read_tsv("results/beta_div/F.beta.txt")) %>%
  gather(mean:sd, key=stat, value=value) %>%
  mutate(group=paste(AgeGroup, stat)) %>%
  select(-stat,-AgeGroup) %>%
  group_by(Type, amplicon, metric) %>%
  spread(key=group, value=value)
```

Model 1: Diet and age

Diet * AgeGroup

- Permutational ANOVA (PERMANOVA) followed by pairwise PERMANOVA with FDR correction (Benjamini-Hochberg method) if significant

Fungi

Fecal

Subset data to fecal samples

```
OTU.temp = metaF %>%
  select(Sample, Type, Diet, AgeGroup, DietAge) %>%
  filter(Type == "F") %>%
  left_join(OTUF, by="Sample") %>%
  arrange(Sample)
```

```r
meta.temp = metaF %>%
  filter(Type == "F")%>%
  arrange(Sample)
```

PERMANOVA

Bray-Curtis (structure)

```r
adonis(OTU.temp[, grepl("Otu", colnames(OTU.temp))] ~ Diet*AgeGroup, data=meta.temp, method="bray", perm
```

```
##
## Call:
## adonis(formula = OTU.temp[, grepl("Otu", colnames(OTU.temp))] ~      Diet * AgeGroup, data = meta.ter
##
## Permutation: free
## Number of permutations: 1000
##
## Terms added sequentially (first to last)
##
##               Df SumsOfSqs MeanSqs F.Model      R2   Pr(>F)
## Diet           2    0.6327  0.3164  1.1469 0.02655 0.278721
## AgeGroup       1    5.2769  5.2769 19.1307 0.22140 0.000999 ***
## Diet:AgeGroup  2    0.5468  0.2734  0.9911 0.02294 0.435564
## Residuals     63   17.3777  0.2758         0.72911
## Total         68   23.8342                 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Jaccard (composition)

```r
adonis(OTU.temp[, grepl("Otu", colnames(OTU.temp))] ~ Diet*AgeGroup, data=meta.temp, method="jaccard", p
```

```
##
## Call:
## adonis(formula = OTU.temp[, grepl("Otu", colnames(OTU.temp))] ~      Diet * AgeGroup, data = meta.ter
##
## Permutation: free
## Number of permutations: 1000
##
## Terms added sequentially (first to last)
##
##               Df SumsOfSqs MeanSqs F.Model      R2   Pr(>F)
## Diet           2    0.7183  0.3592  1.0467 0.02645 0.317682
## AgeGroup       1    4.1552  4.1552 12.1091 0.15298 0.000999 ***
## Diet:AgeGroup  2    0.6708  0.3354  0.9774 0.02470 0.485514
## Residuals     63   21.6179  0.3431         0.79588
## Total         68   27.1622                 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Rumen liquids

Subset data to liquid samples

```r
OTU.temp = metaF %>%
  select(Sample, Type, Diet, AgeGroup, DietAge) %>%
  filter(Type == "L") %>%
  left_join(OTUF, by="Sample") %>%
  arrange(Sample)

meta.temp = metaF %>%
  filter(Type == "L") %>%
  arrange(Sample)
```

PERMANOVA

Bray-Curtis (structure)

```r
adonis(OTU.temp[, grepl("Otu", colnames(OTU.temp))] ~ Diet*AgeGroup, data=meta.temp, method="bray", perm
```

```
##
## Call:
## adonis(formula = OTU.temp[, grepl("Otu", colnames(OTU.temp))] ~      Diet * AgeGroup, data = meta.ter
##
## Permutation: free
## Number of permutations: 1000
##
## Terms added sequentially (first to last)
##
##               Df SumsOfSqs MeanSqs F.Model      R2   Pr(>F)
## Diet           2    0.6185 0.30927  1.9769 0.03278 0.014985 *
## AgeGroup       2    5.4603 2.73015 17.4513 0.28933 0.000999 ***
## Diet:AgeGroup  4    1.6860 0.42151  2.6943 0.08934 0.000999 ***
## Residuals     71   11.1076 0.15644         0.58856
## Total         79   18.8724                 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Jaccard (composition)

```r
adonis(OTU.temp[, grepl("Otu", colnames(OTU.temp))] ~ Diet*AgeGroup, data=meta.temp, method="jaccard", p
```

```
##
## Call:
## adonis(formula = OTU.temp[, grepl("Otu", colnames(OTU.temp))] ~      Diet * AgeGroup, data = meta.ter
##
## Permutation: free
## Number of permutations: 1000
##
## Terms added sequentially (first to last)
##
##               Df SumsOfSqs MeanSqs F.Model      R2   Pr(>F)
## Diet           2    0.9589 0.47946  2.1178 0.03903 0.004995 **
## AgeGroup       2    5.5086 2.75429 12.1661 0.22423 0.000999 ***
## Diet:AgeGroup  4    2.0249 0.50622  2.2361 0.08243 0.000999 ***
## Residuals     71   16.0738 0.22639         0.65431
## Total         79   24.5661                 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Rumen solids

Subset data to solid samples

```
OTU.temp = metaF %>%
  select(Sample, Type, Diet, AgeGroup, DietAge) %>%
  filter(Type == "S") %>%
  left_join(OTUF, by="Sample") %>%
  arrange(Sample)

meta.temp = metaF %>%
  filter(Type == "S") %>%
  arrange(Sample)
```

PERMANOVA

Bray-Curtis (structure)

```
adonis(OTU.temp[, grepl("Otu", colnames(OTU.temp))] ~ Diet*AgeGroup, data=meta.temp, method="bray", perm
```

```
##
## Call:
## adonis(formula = OTU.temp[, grepl("Otu", colnames(OTU.temp))] ~      Diet * AgeGroup, data = meta.ter
##
## Permutation: free
## Number of permutations: 1000
##
## Terms added sequentially (first to last)
##
##               Df SumsOfSqs MeanSqs F.Model      R2   Pr(>F)
## Diet           2    0.4556 0.22781  1.4172 0.02294 0.136863
## AgeGroup       2    5.8531 2.92654 18.2055 0.29466 0.000999 ***
## Diet:AgeGroup  4    1.9813 0.49533  3.0814 0.09974 0.000999 ***
## Residuals     72   11.5740 0.16075         0.58266
## Total         80   19.8640                 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Jaccard (composition)

```
adonis(OTU.temp[, grepl("Otu", colnames(OTU.temp))] ~ Diet*AgeGroup, data=meta.temp, method="jaccard", p
```

```
##
## Call:
## adonis(formula = OTU.temp[, grepl("Otu", colnames(OTU.temp))] ~      Diet * AgeGroup, data = meta.ter
##
## Permutation: free
## Number of permutations: 1000
##
## Terms added sequentially (first to last)
##
##               Df SumsOfSqs MeanSqs F.Model      R2   Pr(>F)
## Diet           2    0.8176 0.40880  1.8140 0.03234 0.011988 *
## AgeGroup       2    5.9442 2.97211 13.1882 0.23510 0.000999 ***
## Diet:AgeGroup  4    2.2964 0.57410  2.5475 0.09082 0.000999 ***
## Residuals     72   16.2259 0.22536         0.64174
```

```
## Total         80   25.2841                 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 2: Milk production efficiency

2yr samples only, averaged 3 consecutive rumen samples to avoid animal effects
Effic + MPE

- Permutational ANOVA (PERMANOVA) followed by pairwise PERMANOVA with FDR correction (Benjamini-Hochberg method) if significant

Fungi

Fecal

Subset data to 2yr fecal samples

```
## Metadata
meta.temp = metaFave %>%
  filter(Type == "F" & AgeGroup == "2yr")
```

```
OTU.temp = metaFave %>%
  # Combine with metadata
  select(Sample, AgeGroup, Type) %>%
  right_join(OTUFave, by="Sample") %>%
  # Filter to only 2 yr fecal samples
  filter(Type == "F" & AgeGroup == "2yr") %>%
  # Remove metadata variables
  select(-Sample, -AgeGroup, -Type)
```

PERMANOVA
Bray-Curtis (structure)

```
adonis(OTU.temp ~ Effic+MPE, data=meta.temp, method="bray", permutations=1000)
```

```
##
## Call:
## adonis(formula = OTU.temp ~ Effic + MPE, data = meta.temp, permutations = 1000,      method = "bray"
##
## Permutation: free
## Number of permutations: 1000
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## Effic      1    0.3371 0.33706 0.91268 0.02879 0.5504
## MPE        1    0.2899 0.28988 0.78492 0.02476 0.7592
## Residuals 30   11.0793 0.36931         0.94644
## Total     32   11.7062                 1.00000
```

Jaccard (composition)

```
adonis(OTU.temp ~ Effic+MPE, data=meta.temp, method="jaccard", permutations=1000)
```

```
##
## Call:
## adonis(formula = OTU.temp ~ Effic + MPE, data = meta.temp, permutations = 1000,      method = "jacca
##
## Permutation: free
## Number of permutations: 1000
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## Effic      1    0.4230 0.42303 1.00605 0.03154 0.4346
## MPE        1    0.3733 0.37327 0.88771 0.02783 0.6943
## Residuals 30   12.6146 0.42049         0.94062
## Total     32   13.4109                 1.00000
```

Rumen liquids

Subset data to 2yr rumen liquid samples

```
## Metadata
meta.temp = metaFave %>%
  # Efficieny
  filter(Type == "L" & AgeGroup == "2yr")
```

```
OTU.temp = metaFave %>%
  # Combine with metadata
  select(Sample, AgeGroup, Type) %>%
  right_join(OTUFave, by="Sample") %>%
  # Filter to only 2 yr fecal samples
  filter(Type == "L" & AgeGroup == "2yr") %>%
  # Remove metadata variables
  select(-Sample, -AgeGroup, -Type)
```

PERMANOVA

Bray-Curtis (structure)

```
adonis(OTU.temp ~ Effic+MPE, data=meta.temp, method="bray", permutations=1000)
```

```
##
## Call:
## adonis(formula = OTU.temp ~ Effic + MPE, data = meta.temp, permutations = 1000,      method = "bray"
##
## Permutation: free
## Number of permutations: 1000
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs  MeanSqs F.Model      R2 Pr(>F)
## Effic      1   0.16574 0.165739  1.8008 0.14848 0.1049
## MPE        1   0.12216 0.122161  1.3273 0.10944 0.2388
## Residuals  9   0.82833 0.092037         0.74208
## Total     11   1.11623                 1.00000
```

Jaccard (composition)

```
adonis(OTU.temp ~ Effic+MPE, data=meta.temp, method="jaccard", permutations=1000)
```

```
##
## Call:
## adonis(formula = OTU.temp ~ Effic + MPE, data = meta.temp, permutations = 1000,      method = "jacca
##
## Permutation: free
## Number of permutations: 1000
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## Effic      1   0.22738 0.22738  1.3758 0.11785 0.1588
## MPE        1   0.21452 0.21452  1.2980 0.11119 0.2418
## Residuals  9   1.48742 0.16527         0.77096
## Total     11   1.92932                 1.00000
```

Rumen solids

Subset data to 2yr rumen solid samples

```
## Metadata
meta.temp = metaFave %>%
  # Efficieny
  filter(Type == "S" & AgeGroup == "2yr")
```

```
OTU.temp = metaFave %>%
  # Combine with metadata
  select(Sample, AgeGroup, Type) %>%
  right_join(OTUFave, by="Sample") %>%
  # Filter to only 2 yr fecal samples
  filter(Type == "S" & AgeGroup == "2yr") %>%
  # Remove metadata variables
  select(-Sample, -AgeGroup, -Type)
```

PERMANOVA

Bray-Curtis (structure)

```
adonis(OTU.temp ~ Effic+MPE, data=meta.temp, method="bray", permutations=1000)
```

```
##
## Call:
## adonis(formula = OTU.temp ~ Effic + MPE, data = meta.temp, permutations = 1000,      method = "bray"
##
## Permutation: free
## Number of permutations: 1000
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## Effic      1   0.20197 0.20197  1.9313 0.15923 0.1079
## MPE        1   0.12524 0.12524  1.1975 0.09873 0.3177
## Residuals  9   0.94121 0.10458         0.74203
## Total     11   1.26842                 1.00000
```

Jaccard (composition)

```
adonis(OTU.temp ~ Effic+MPE, data=meta.temp, method="jaccard", permutations=1000)
```

```
##
## Call:
## adonis(formula = OTU.temp ~ Effic + MPE, data = meta.temp, permutations = 1000,      method = "jacca
##
## Permutation: free
## Number of permutations: 1000
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## Effic      1   0.25101 0.25101  1.3863 0.11948 0.1818
## MPE        1   0.22025 0.22025  1.2164 0.10484 0.3027
## Residuals  9   1.62960 0.18107         0.77568
## Total     11   2.10086                 1.00000
```

ALPHA-DIVERSITY

Model 1: Diet and age

Diet * AgeGroup

- ANOVA

Fungi

Fecal

Shannon's diversity

ANOVA

```
metaF %>%
  filter(Type == "F") %>%
  aov(shannon ~ Diet*AgeGroup, .) %>%
  summary()
```

```
##               Df Sum Sq Mean Sq F value   Pr(>F)
## Diet           2  1.052   0.526   1.542   0.2219
## AgeGroup       1  8.517   8.517  24.963 4.92e-06 ***
## Diet:AgeGroup  2  2.404   1.202   3.523   0.0354 *
## Residuals     63 21.493   0.341
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Rumen liquids

Shannon's diversity

ANOVA

```
metaF %>%
  filter(Type == "L") %>%
  aov(shannon ~ Diet*AgeGroup, .) %>%
  summary()
```

```
##                Df Sum Sq Mean Sq F value  Pr(>F)
## Diet            2  0.327  0.1633   1.504  0.2292
## AgeGroup        2  3.255  1.6277  14.994 3.7e-06 ***
## Diet:AgeGroup   4  1.251  0.3128   2.881  0.0286 *
## Residuals      71  7.707  0.1086
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Rumen solids

Shannon's diversity

ANOVA

```
metaF %>%
  filter(Type == "S") %>%
  aov(shannon ~ Diet*AgeGroup, .) %>%
  summary()
```

```
##                Df Sum Sq Mean Sq F value   Pr(>F)
## Diet            2  0.026  0.0131   0.069 0.933258
## AgeGroup        2  2.981  1.4904   7.879 0.000805 ***
## Diet:AgeGroup   4  1.073  0.2683   1.418 0.236575
## Residuals      72 13.620  0.1892
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 2: Milk production efficiency

2yr samples only, averaged 3 consecutive rumen samples to avoid animal effects
Effic + MPE

- ANOVA followed by pairwise TukeyHSD if significant

Fungi

Fecal

Shannon's diversity

ANOVA

```
metaFave %>%
  filter(Type == "F" & AgeGroup == "2yr") %>%
  aov(shannon ~ Effic+MPE, .) %>%
  summary()
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Effic        1  0.769  0.7687   1.286  0.266
## MPE          1  0.190  0.1897   0.317  0.577
## Residuals   30 17.931  0.5977
```

Rumen liquids

Shannon's diversity

ANOVA

```
metaFave %>%
  filter(Type == "L" & AgeGroup == "2yr") %>%
  aov(shannon ~ Effic+MPE, .) %>%
  summary()
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Effic        1 0.0525 0.05251   1.272  0.289
## MPE          1 0.0000 0.00001   0.000  0.988
## Residuals    9 0.3715 0.04128
```

Rumen solids

Shannon's diversity

ANOVA

```
metaFave %>%
  filter(Type == "S" & AgeGroup == "2yr") %>%
  aov(shannon ~ Effic+MPE, .) %>%
  summary()
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Effic        1 0.0639 0.06390   1.345  0.276
## MPE          1 0.0082 0.00823   0.173  0.687
## Residuals    9 0.4275 0.04750
```