African Elephant Microbiome Stats

Joe Gunn

2/21/2019

AFRICAN ELEPHANT MICROBIOME - BETA DIVERSITY

Purpose: Here we are assessing Beta diversity (differences in microbial community composition) between African Elephant species and among African Elephant groups using PERMANOVA. Groups include all combinations of diet and habitat so that all combinations can be compared. We also include beta dispersion analyses for African Elephant species, diet, and habitat to determine whether there are differences in within-group variance. We visualized beta diversity using NMDS.

Data used:

```
  TABLE output from QIIME: all OTUs detected across African Elephant samples, taxonomic level, and raw abundance
for each individual <- 8,248 OTUs (after rarefaction)
```

Libraries Needed for Analysis

Metadata

Get Rarefied OTU abundance data

```
## `set.seed(5)` was used to initialize repeatable random subsampling.
```

```
## Please record this for your records so others can reproduce.
```

```
## Try `set.seed(5)); .Random.seed` for the full vector
```

```
## ...
```

```
## 8180TUs were removed because they are no longer
## present in any sample after random subsampling
```

```
## ...
```

Prepare and clean data for two datasets

```
#Prepare data at the OTU level for analysis with both datasets (just forest non-croparders from both species, and d just africans)
asm_physeq <- as.data.frame(otu_table(asm_physeq))
asm_physeq_t <- t(asm_physeq)
asm_physeq <- as.data.frame(asm_physeq_t)
asm_physeq <- asm_physeq %>%%
  rownames_to_column("sample_id")

asm_physeq <- merge(metadata, asm_physeq, by = "sample_id")
asm_physeq_forest_ncr <- asm_physeq %>%
  filter(Habitat == "Forest") %>%
  filter(Raider == "No")

asm_physeq_africana <- asm_physeq %>%
  filter(Species == "africana")
```

Prepare Data for NMDS for African elephant Species, Diet, and Habitat

```
#Make NMDS for all samples and both datasets together
asm_physeq_all_data <- asm_physeq[,-c(2:11)]
asm_physeq_all_data <- column_to_rownames(asm_physeq_all_data, "sample_id")

asm_all_veg <- decostand(asm_physeq_all_data, "total")
asm_all_veg <- vegdist(asm_all_veg, "bray")
asm_all_pcoa <- pcoa(asm_all_veg)

asm_all_pcs <- asm_all_pcoa$vectors
asm_all_pcs <- as.data.frame(asm_all_pcs)
asm_all_pcs <- asm_all_pcs %>%
  rownames_to_column("sample_id")

asm_all_pcs <- merge(metadata, asm_all_pcs, by = "sample_id")
asm_physeq_all_ado1 <- column_to_rownames(asm_physeq, "sample_id")

#Run NMDS
vare.mds0_all <- isoNDS(asm_all_veg, trace = 0)
vare.mds_all <- metaNDS(asm_physeq_all_data, trace = FALSE, autotransform = F)
score_all <- scores(vare.mds_all)
score_all <- as.data.frame(score_all)
score_all <- score_all %>%
  rownames_to_column("sample_id")

merge_all <- merge(metadata, score_all, by = "sample_id")
```

Run PERMANOVA analysis for African elephant Species, Diet, and Habitat

Prepare data for PERMANOVA FOR SPECIES

```
asm_veg_data <- asm_physeq_forest_ncr[,-c(2:11)]
asm_veg_data <- column_to_rownames(asm_veg_data, "sample_id")

#Run Distance Matrix on Abundance
asm_veg <- decostand(asm_veg_data, "total") #standardize abundance data by dividing each number by the total numb
er of samples (individuals in the dataset)
asm_veg <- vegdist(asm_veg, "bray") #calculate Bray-curtis distances between samples
asm_pcoa <- pcoa(asm_veg) #calculates principal components

#Prepare PCOA (unused in manuscript)
asm_pcs <- asm_pcoa$vectors
asm_pcs <- as.data.frame(asm_pcs)
asm_pcs <- asm_pcs %>%
  rownames_to_column("sample_id")

asm_pcs <- merge(metadata, asm_pcs, by = "sample_id")
asm_physeq_ado1 <- column_to_rownames(asm_physeq_forest_ncr, "sample_id")
```

Prepare data for PERMANOVA FOR DIET and HABITAT

```
asm_veg2_data <- asm_physeq_africana[,-c(2:11)]
asm_veg2_data <- column_to_rownames(asm_veg2_data, "sample_id")

#Run Distance Matrix on Abundance
asm_veg2 <- decostand(asm_veg2_data, "total")
asm_veg2 <- vegdist(asm_veg2, "bray")
asm_pcoa2 <- pcoa(asm_veg2)

#Prepare PCOA
asm_pcs2 <- asm_pcoa2$vectors
asm_pcs2 <- as.data.frame(asm_pcs2)
asm_pcs2 <- asm_pcs2 %>%
  rownames_to_column("sample_id")

asm_pcs2 <- merge(metadata, asm_pcs2, by = "sample_id")
asm_physeq_ado2 <- column_to_rownames(asm_physeq_africana, "sample_id")
```

Beta Dispersion Anlysis

```
#Beta Dispersion Analysis for African Elephant species
species_beta <- factor(asm_physeq_ado1[,2])
beta_analysis <- betadisper(asm_veg, species_beta, type = c("centroid"))

anova(beta_analysis) #F = 0.4346, F = 0.6405
```

```
## Analysis of Variance Table
##
## Response: Distances
##           Df  Sum Sq  Mean Sq F value Pr(>F)
## Groups     1 0.006622 0.006622  0.4405 0.4346
## Residuals 17 0.175765 0.0103391
```

```
#Graph Beta Analysis for Species
  # beta_vectors <- beta_analysis$vectors
  # beta_vectors <- as.data.frame(beta_vectors)
  # beta_vectors <- beta_vectors %>% rownames_to_column("sample_id")
  # beta_vectors <- merge(metadata, beta_vectors, by = "sample_id")
  # beta_graph <- ggplot(beta_vectors, aes(x = PCoA1, y = PCoA2, color = Species)) + geom_point()

#Beta Dispersion Analysis for Diet and Habitat
diet_beta <- factor(asm_physeq_ado2[,4])
habitat_beta <- factor(asm_physeq_ado2[,5])
beta_analysis2 <- betadisper(asm_veg2, diet_beta, type = c("centroid"))
beta_analysis3 <-betadisper(asm_veg2, habitat_beta, type = c("centroid"))

anova(beta_analysis2) #F = 0.1224, F = 2.5136
```

```
## Analysis of Variance Table
##
## Response: Distances
##           Df  Sum Sq  Mean Sq F value Pr(>F)
## Groups     1 0.009585 0.0095850  2.5136 0.1224
## Residuals 33 0.125837 0.0038132
```

```
anova(beta_analysis3) #F = 0.1022, F = 2.8266
```

```
## Analysis of Variance Table
##
## Response: Distances
##           Df  Sum Sq  Mean Sq F value Pr(>F)
## Groups     1 0.015145 0.0151454  2.8266 0.1022
## Residuals 33 0.176823 0.0053583
```

Run PERMANOVA

```
ado_species <- adonis(asm_veg ~ Species, asm_physeq_ado1, Strata = Sex, distance = "bray", permutations = 9999)
  #p-value = 0.001
  #F-model = 3.527

ado_diet_hab_interaction <- adonis(asm_veg2 ~ Raider*Habitat, Strata = Sex, asm_physeq_ado2, distance = "bray", p
ermutations = 9999)
  #p-value = 0.0974
  #F-model value = 1.369

#####Not a significant interaction, so run diet and habitat separately with sex as strata and infer main effects#
#######

ado_diet <- adonis(asm_veg2 ~ Raider, asm_physeq_ado2, Strata = Sex, distance = "bray", permutations = 9999)
  #p-value = 0.011
  #F-model value = 1.861

ado_habitat <- adonis(asm_veg2 ~ Habitat, asm_physeq_ado2, Strata = Sex, distance = "bray", permutations = 9999)
  #p-value = 0.0001
  #F-model value = 2.874
```

NMDS for Abundance data

```
###PLOT NMDS BY SPECIES
vare.mds0 <- isoNDS(asm_veg, trace = 0)
vare.mds <- metaNDS(asm_veg_data, trace = FALSE, autotransform = F)
score <- scores(vare.mds)
score <- as.data.frame(score)
score <- score %>%
  rownames_to_column("sample_id")

merge <- merge(metadata, score, by = "sample_id")

levels(merge$Species) <- c("L. africana", "L. cyclotis")

centroids <- aggregate(cbind(NMDS1, NMDS2) ~ Species, merge, mean)
gg <- merge(merge, aggregate(cbind(mean.x = NMDS1, mean.y = NMDS2) ~ Species, merge, mean),by = "Species")


species_perm <- ggplot(gg, aes(NMDS1, NMDS2, color = Species)) +
  geom_point(size = 2, show.legend = F) +
  geom_point(aes(x = mean.x, y = mean.y), size = 2, show.legend = F) +
  geom_segment(aes(x = mean.x, y = mean.y, xend = NMDS1, yend = NMDS2), show.legend = F) +
  theme_set(theme_cowplot(12)) +
  theme(legend.position=c(0.03,0.85)) +
  scale_color_manual(values = c("red2", "blue2")) +
  labs(color = "Elephant Species", title = "Beta Diversity: Split Models", y = "NMDS 2") +
  theme(axis.title.x = element_blank()) + theme(legend.text = element_text(face = "italic")) +
  theme(plot.title = element_text(size = 20, hjust = 0.5, face = "bold"))

#PLOT NMDS BY DIET
vare.mds1 <- isoNDS(asm_veg2, trace = 0)
vare.mds2 <- metaNDS(asm_veg2_data, trace = FALSE, autotransform = F)
score_diet <- scores(vare.mds2)
score_diet <- as.data.frame(score_diet)
score_diet <- score_diet %>%
  rownames_to_column("sample_id")

merge_diet <- merge(metadata, score_diet, by = "sample_id")

#Plot
levels(merge_diet$Raider) <- c("Non-Crop-Raider", "Crop-Raider")
centroids_diet <- aggregate(cbind(NMDS1, NMDS2) ~ Raider, merge_diet, mean)
gg_diet <- merge(merge_diet, aggregate(cbind(mean.x = NMDS1, mean.y = NMDS2) ~ Raider, merge_diet, mean), by = "R
aider")

diet_perm <- ggplot(gg_diet, aes(NMDS1, NMDS2, color = Raider)) +
  geom_point(size = 2, show.legend = F) +
  geom_point(aes(x = mean.x, y = mean.y), size = 2, show.legend = F) +
  theme_set(theme_cowplot(12)) +
  geom_segment(aes(x = mean.x, y = mean.y, xend = NMDS1, yend = NMDS2), show.legend = F) +
  theme(legend.position = c(0.01,0.85)) +
  scale_color_manual(values = c("midnightblue", "deeppink1")) +
  theme(axis.title.x = element_blank()) +
  labs(y = "NMDS 2", color = "Elephant Diet")

#PLOT NMDS BY HABITAT
vare.mds_hab <- isoNDS(asm_veg2, trace = 0)
vare.mds_hab2 <- metaNDS(asm_veg2_data, trace = FALSE, autotransform = F)
score_hab <- scores(vare.mds_hab2)
score_hab <- as.data.frame(score_hab)
score_hab <- score_hab %>% rownames_to_column("sample_id")
merge_hab <- merge(metadata, score_hab, by = "sample_id")

#Plot
centroids_hab <- aggregate(cbind(NMDS1, NMDS2) ~ Habitat, merge_hab, mean)
gg_hab <- merge(merge_hab, aggregate(cbind(mean.x = NMDS1, mean.y = NMDS2) ~ Habitat, merge_hab, mean), by = "Hab
itat")

habitat_perm <- ggplot(gg_hab, aes(NMDS1, NMDS2, color = Habitat)) +
  geom_point(size = 2, show.legend = F) +
  geom_point(aes(x = mean.x, y = mean.y), size = 2,  show.legend = F) +
  geom_segment(aes(x = mean.x, y = mean.y, xend = NMDS1, yend = NMDS2), show.legend = F) +
  theme_set(theme_cowplot(12)) +
  theme(legend.position = c(0.01,0.85)) +
  scale_color_manual(values = c("forestgreen", "coral1")) + labs(color = "Elephant Habitat", x = "NMDS 1", y = "N
MDS 2")


#PLOT EVERYTHING TOGETHER
pdf("/Users/jogunn/Desktop/Grad_School_Stuff/Research/Projects/Elephant_Microbiome/Attempt_2/visualization/beta_
diversity_figures/beta_diversity.pdf", width = 6, height = 12)

plot_grid(species_perm, diet_perm, habitat_perm, nrow = 3, labels = c("A", "B", "C"))

dev.off()
```

```
## quartz_off_screen
##                 2
```