

# Pairwise Comparison of Beta Diversity

Jung-Jin Lee

Jan 15, 2020

```
# Metadata
# SampleID: sample identifier
s <- "meta.tsv" %>%
  read_tsv()
dim(s)

## [1] 140 26

# Bacteria counts data
bac_counts <- "bacteria_counts.tsv" %>%
  read_tsv() %>%
  column_to_rownames(var = "Taxa") %>%
  t()
dim(bac_counts) # Each row is a SampleID and each column is a species

## [1] 140 4373

# Bacteria props data
bac_props <- sweep(bac_counts, 1, rowSums(bac_counts), "/")
dim(bac_props)

## [1] 140 4373

# Make a distance object using usedist::dist_make()
dist_beta <- dist_make(bac_counts, bray_curtis)

pairwise_beta_plot <- function(s, dist_beta, color_by) {
  ###=====
  # Step I:
  # Samples in dist_in may not appear in the metadata. Extract only common samples
  ###=====
  common_samples <- intersect(s$SampleID, attributes(dist_beta)$Labels)
  s <- s %>%
    filter(SampleID %in% common_samples)
  dist_beta <- dist_subset(dist_beta, s$SampleID)

  ###=====
  # Step II:
  # Treat color_by as a factor
  ###=====
  color_by <- enquo(color_by)
  if (!is.factor(s %>% pull(!!color_by))) {
    s <- s %>%
      mutate(!!color_by := factor(!!color_by))
  }
}
```

```

color_by_levels = s %>% pull(!color_by) %>% levels()

###=====
# Step III:
# Compute PCoA
###=====
pcoa_result <- ape::pcoa(dist_beta)
pcoa_df <- s %>%
  left_join(pcoa_result$vectors[, c("Axis.1", "Axis.2")] %>%
    as.data.frame() %>%
      rownames_to_column("SampleID"),
    by = "SampleID")
pcoa_pct <- round(pcoa_result$values$Relative_eig*100, 1)

###=====
# Step IV:
# Create pairwise ordination
###=====
sub_df_list <- list()
counter <- 0
for (i in 1:(length(color_by_levels) - 1)) {
  Gr1 <- color_by_levels[i]
  for (j in (i + 1):length(color_by_levels)) {
    Gr2 <- color_by_levels[j]
    counter <- counter + 1
    curr_comparison <- paste0(Gr1, " vs ", Gr2)
    sub_df <- pcoa_df %>%
      filter(!color_by %in% c(Gr1, Gr2)) %>%
      select(!color_by, c("Axis.1", "Axis.2")) %>%
      droplevels()
    sub_df_list[[counter]] <- sub_df %>%
      mutate(Comparison = curr_comparison) %>%
      mutate_if(is.factor, as.character)
  }
}

all_df <- bind_rows(sub_df_list) %>%
  mutate(Comparison = factor(Comparison, levels = unique(.$Comparison)))

g <- all_df %>%
  ggplot(aes(Axis.1, Axis.2)) +
  geom_point(aes(color = !color_by)) +
  theme_bw() +
  theme(aspect.ratio = 1) +
  theme(legend.position = "bottom") +
  facet_wrap(~Comparison)

return(g)
}

pairwise_beta_plot(s, dist_beta, Hospital)

```

