

# RA\_script2

Deepan

7/31/2024

```
library(stringr)
library(ggplot2)
library(FactoMineR)
library(factoextra)
```

```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```
library(reshape)
library(reshape2)
```

```
##
## Attaching package: 'reshape2'

## The following objects are masked from 'package:reshape':
##
##      colsplit, melt, recast
```

```
library(gplots)
```

```
##
## Attaching package: 'gplots'

## The following object is masked from 'package:stats':
##
##      lowess
```

```
library(RColorBrewer)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following object is masked from 'package:reshape':
##
##      rename

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

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

```
library(ggpubr)
library (ANCOMBC)
```

```
###Theme function definitions
```

```
theme_boxp <- function(){

  font <- "Times New Roman"

  theme_minimal() %+replace%

  theme(axis.title.y = element_text(angle = 90, size = 10),
        axis.title.x = element_blank(),
        axis.text.x = element_text(angle = 45, size = 10, color = "black"),
        axis.text.y = element_text(size = 10, color = "black"),

        panel.background = element_rect(fill = "white"),
        axis.line.x = element_line(color = "black"),
        axis.line.y = element_line(color = "black"),

        legend.text = element_text(size = 10, lineheight = 0.6),
        legend.spacing.y = unit(0.6, "mm"),
        legend.key.height = unit(2, "mm"),
        legend.key=element_rect(fill="white", color="black", linewidth = 0),
        legend.title = element_blank(),
        strip.background = element_rect(fill = NA, linetype = "solid", size = 1, colour = "black"),
        strip.text = element_text(size = 10, margin = margin(1,1,1,1)),
        panel.border = element_rect(fill = NA, size = 1),
        panel.spacing.y = unit(1,"line"),
  )
}

theme_pca <- function(){

  font <- "Times New Roman"

  theme_boxp() %+replace%

  theme(legend.position = "none",
        axis.title.y = element_text(angle = 90, size = 12, color = "black"),
        axis.title.x = element_text(size = 12, color = "black"),
        axis.text.x = element_text(size = 10),
        axis.text.y = element_text(size = 10),
  )
}

reg_eq <- function(x,y) {
  m <- lm(y ~ x)
  as.character(
```

```

as.expression(
  substitute(italic(y) == a + b %.% italic(x)*", "~italic(r)^2~"=="~r2,
    list(a = format(coef(m)[1], digits = 4),
      b = format(coef(m)[2], digits = 4),
      r2 = format(summary(m)$r.squared, digits = 3)))
)
}

```

###Defining figure color vectors

```

vec_col_patients <- c( "midnightblue",
  "yellow3", "forestgreen", "chocolate", "plum", "cornflowerblue")

vec_col_type <- c( "darkgoldenrod1", "blue4")

family_cols <- c("Bacteroidaceae"="cyan3", "Prevotellaceae"="coral", "Verrucomicrobiaceae"="#1B9E77", "I

```

###Importing Metadata file

```

metadata <- read.delim("/media/deepan/Deepan/RA_Bottini/sample-metadata-wflownumbers.tsv", header = TRUE)

```

###Generating PCA plots from imported phylo-RPCA, unweighted unifrac and jaccard distance matrices generated with QIIME2 (Bolyen et al., 2019)

```

phylo_rpca <- read.delim("/media/deepan/Deepan/RA_Bottini/final_merged_tables_forwardonly/afterFMT/phylo_rpca.tsv")

phylo_rpca <- merge(phylo_rpca, metadata, by = "sample-ID", all.y = FALSE)

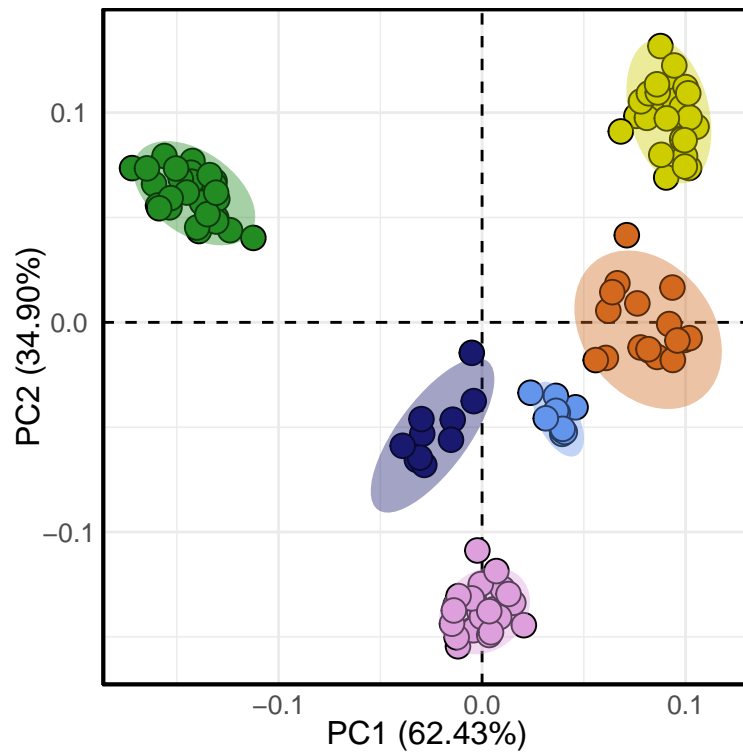
#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Main_Fig1.svg")
ggplot(phylo_rpca, aes(x = PC1, y = PC2, fill = Patient)) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +
  geom_vline(xintercept = 0, linetype = "dashed", color = "black") +
  geom_point(size = 4, pch = 21) +
  stat_ellipse(level = 0.95, geom = "polygon", alpha = 0.4) +
  scale_fill_manual(values = rev(vec_col_patients)) +
  labs(x = "PC1 (62.43%)", y = "PC2 (34.90%)") +
  theme_pca()

```

```

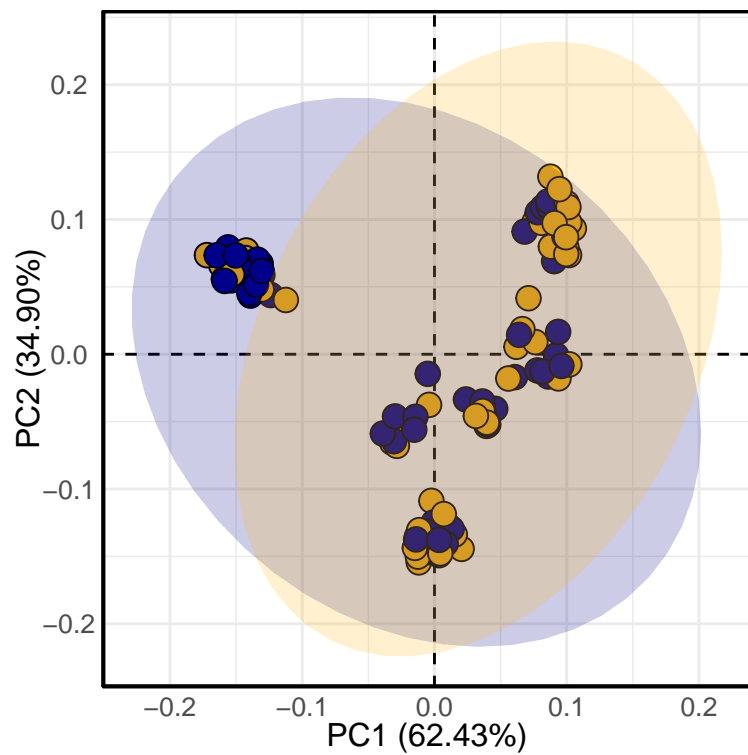
## Warning: The 'size' argument of 'element_rect()' is deprecated as of ggplot2 3.4.0.
## i Please use the 'linewidth' argument instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.

```



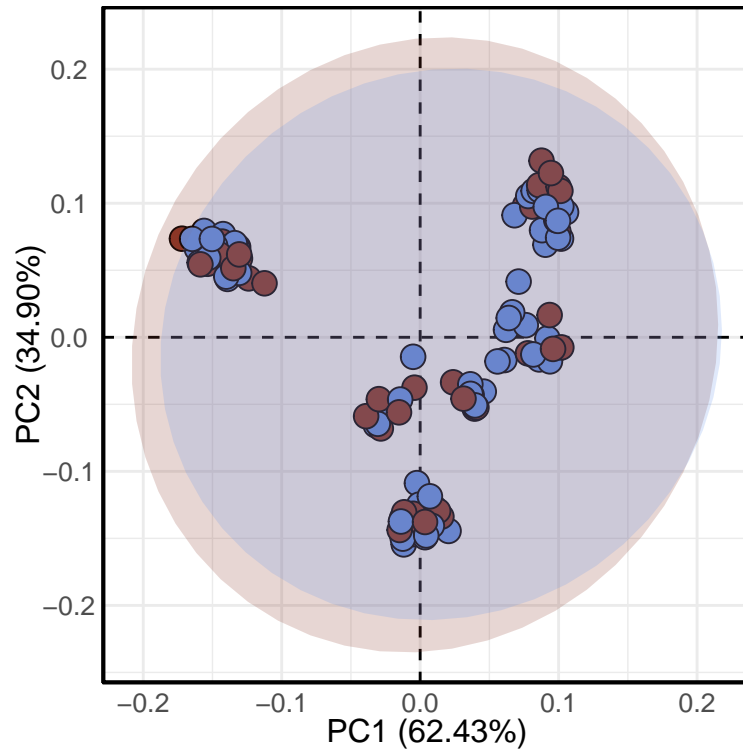
```
#dev.off()
```

```
#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Mai
ggplot(phylo_rpca, aes(x = PC1, y = PC2, fill = MouseType)) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +
  geom_vline(xintercept = 0, linetype = "dashed", color = "black") +
  geom_point(size = 4, pch = 21) +
  stat_ellipse(level = 0.95, geom = "polygon", alpha = 0.2) +
  scale_fill_manual(values = rev(vec_col_type)) +
  labs(x = "PC1 (62.43%)", y = "PC2 (34.90%)") +
  theme_pca()
```



```
#dev.off()
```

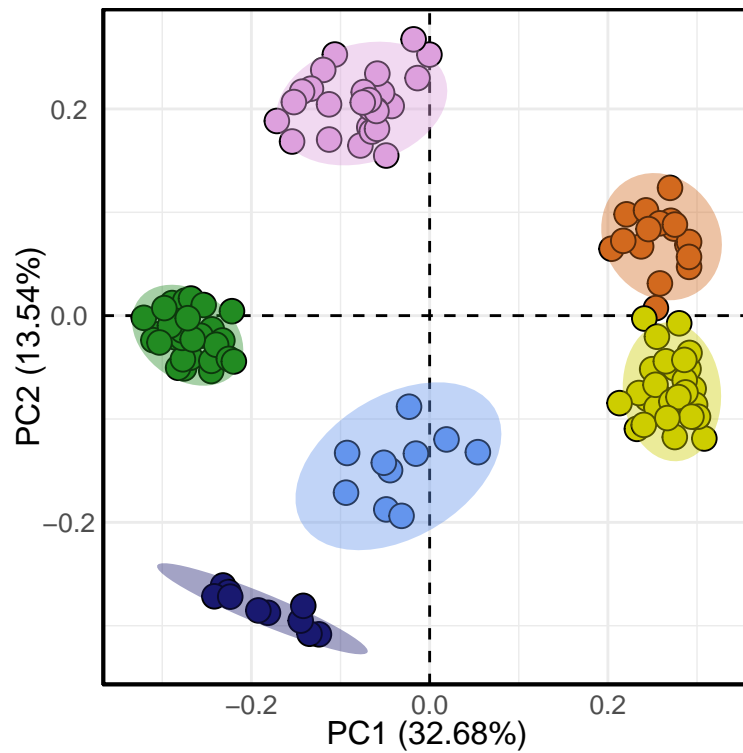
```
#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Mai
ggplot(phylo_rpca, aes(x = PC1, y = PC2, fill = BeforeAfterArthritisInd)) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +
  geom_vline(xintercept = 0, linetype = "dashed", color = "black") +
  geom_point(size = 4, pch = 21) +
  stat_ellipse(level = 0.95, geom = "polygon", alpha = 0.2) +
  scale_fill_manual(values = c("tomato4", "cornflowerblue")) +
  labs(x = "PC1 (62.43%)", y = "PC2 (34.90%)") +
  theme_pca()
```



```
#dev.off()

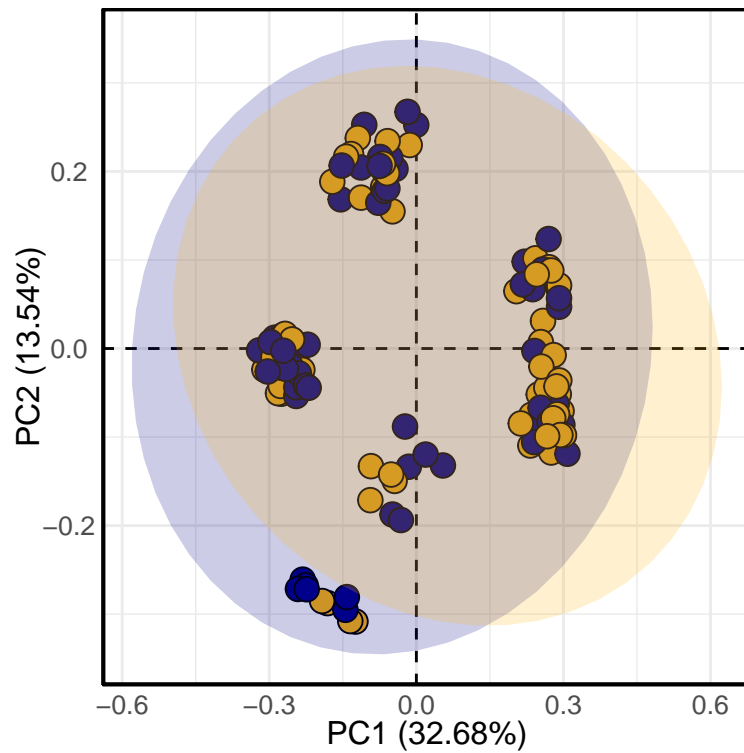
uunifrac <- read.delim("/media/deepan/Deepan/RA_Bottini/final_merged_tables_forwardonly/afterFMT/core-m
uunifrac <- merge(uunifrac, metadata, by = "sample-ID", all.y = FALSE)

#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Supp
ggplot(uunifrac, aes(x = PC1, y = PC2, fill = Patient)) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +
  geom_vline(xintercept = 0, linetype = "dashed", color = "black") +
  geom_point(size = 4, pch = 21) +
  stat_ellipse(level = 0.95, geom = "polygon", alpha = 0.4) +
  scale_fill_manual(values = rev(vec_col_patients)) +
  labs(x = "PC1 (32.68%)", y = "PC2 (13.54%)") +
  theme_pca()
```



```
#dev.off()
```

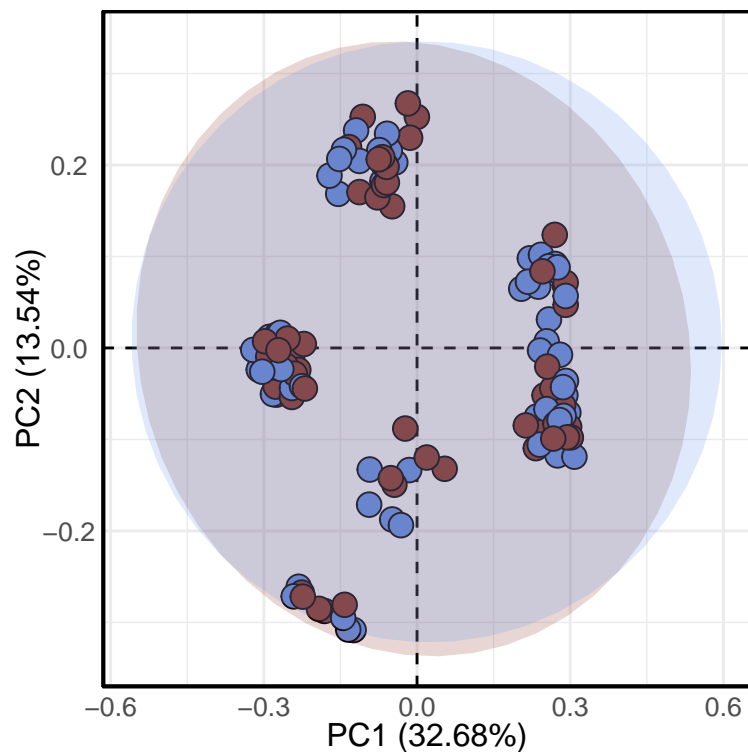
```
#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Supp
ggplot(uunifrac, aes(x = PC1, y = PC2, fill = MouseType)) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +
  geom_vline(xintercept = 0, linetype = "dashed", color = "black") +
  geom_point(size = 4, pch = 21) +
  stat_ellipse(level = 0.95, geom = "polygon", alpha = 0.2) +
  scale_fill_manual(values = rev(vec_col_type)) +
  labs(x = "PC1 (32.68%)", y = "PC2 (13.54%)") +
  theme_pca()
```



```
#dev.off()
```

```
#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Supp
ggplot(uunifrac, aes(x = PC1, y = PC2, fill = BeforeAfterArthritisInd)) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +
  geom_vline(xintercept = 0, linetype = "dashed", color = "black") +
  geom_point(size = 4, pch = 21) +
  stat_ellipse(level = 0.95, geom = "polygon", alpha = 0.2) +
  scale_fill_manual(values = c("tomato4", "cornflowerblue")) +
  labs(x = "PC1 (32.68%)", y = "PC2 (13.54%)") +
  theme_pca()
```



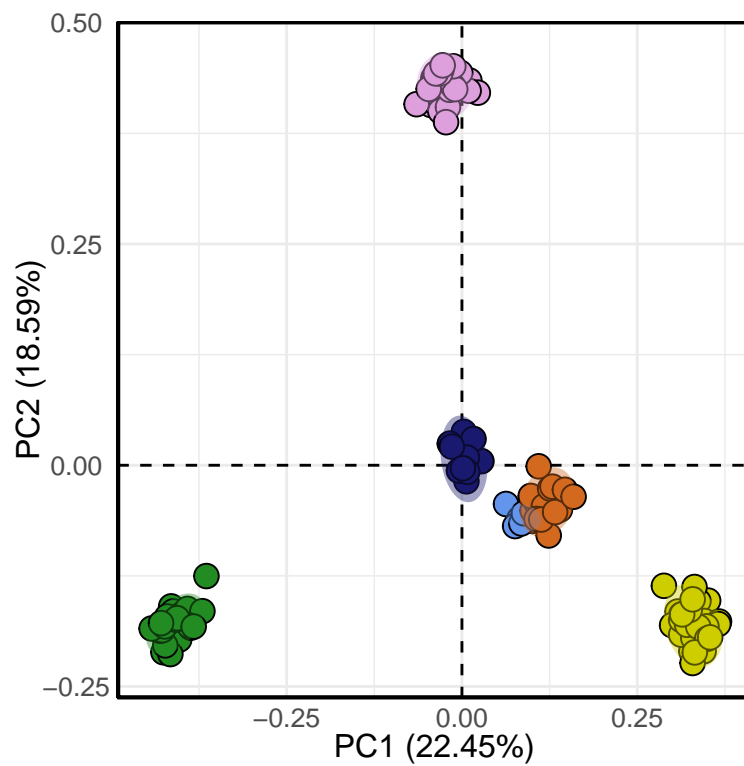


```
#dev.off()

jacc <- read.delim("/media/deepan/Deepan/RA_Bottini/final_merged_tables_forwardonly/afterFMT/core-metri

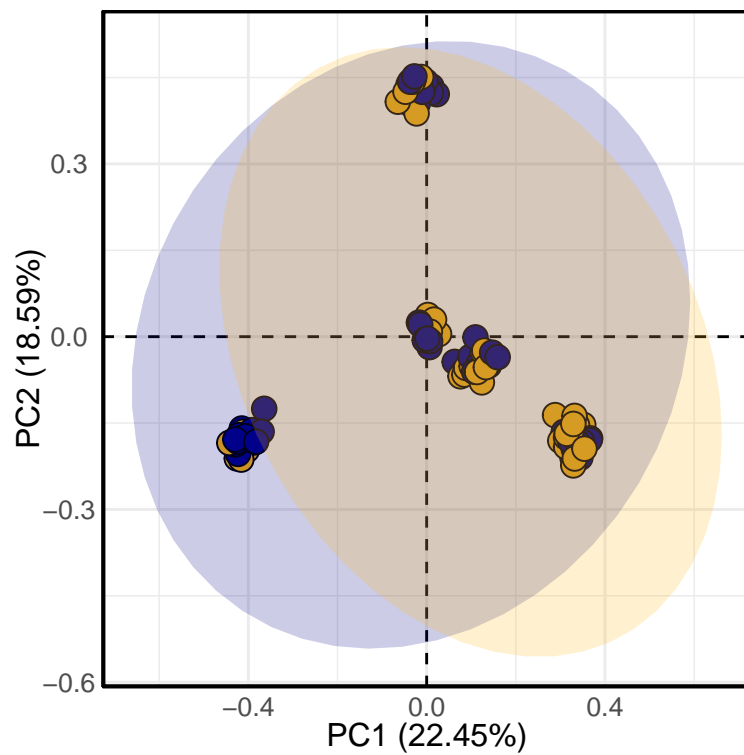
jacc <- merge(jacc, metadata, by = "sample-ID", all.y = FALSE)

#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Supp
ggplot(jacc, aes(x = PC1, y = PC2, fill = Patient)) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +
  geom_vline(xintercept = 0, linetype = "dashed", color = "black") +
  geom_point(size = 4, pch = 21) +
  stat_ellipse(level = 0.95, geom = "polygon", alpha = 0.4) +
  scale_fill_manual(values = rev(vec_col_patients)) +
  labs(x = "PC1 (22.45%)", y = "PC2 (18.59%)") +
  theme_pca()
```



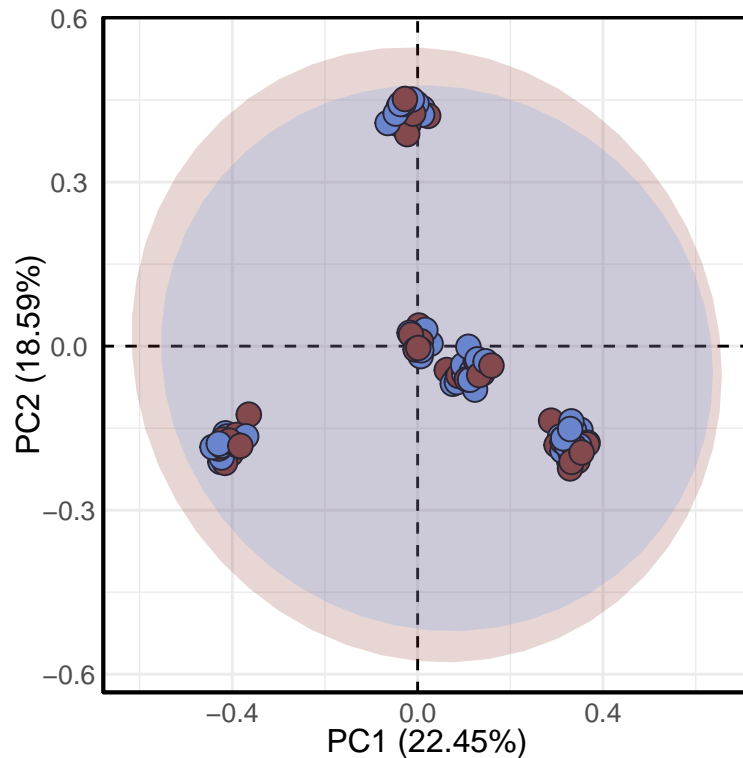
```
#dev.off()
```

```
#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Supp
ggplot(jacc, aes(x = PC1, y = PC2, fill = MouseType)) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +
  geom_vline(xintercept = 0, linetype = "dashed", color = "black") +
  geom_point(size = 4, pch = 21) +
  stat_ellipse(level = 0.95, geom = "polygon", alpha = 0.2) +
  scale_fill_manual(values = rev(vec_col_type)) +
  labs(x = "PC1 (22.45%)", y = "PC2 (18.59%)") +
  theme_pca()
```



```
#dev.off()
```

```
#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Supp
ggplot(jacc, aes(x = PC1, y = PC2, fill = BeforeAfterArthritisInd)) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +
  geom_vline(xintercept = 0, linetype = "dashed", color = "black") +
  geom_point(size = 4, pch = 21) +
  stat_ellipse(level = 0.95, geom = "polygon", alpha = 0.2) +
  scale_fill_manual(values = c("tomato4", "cornflowerblue")) +
  labs(x = "PC1 (22.45%)", y = "PC2 (18.59%)") +
  theme_pca()
```



```
#dev.off()
```

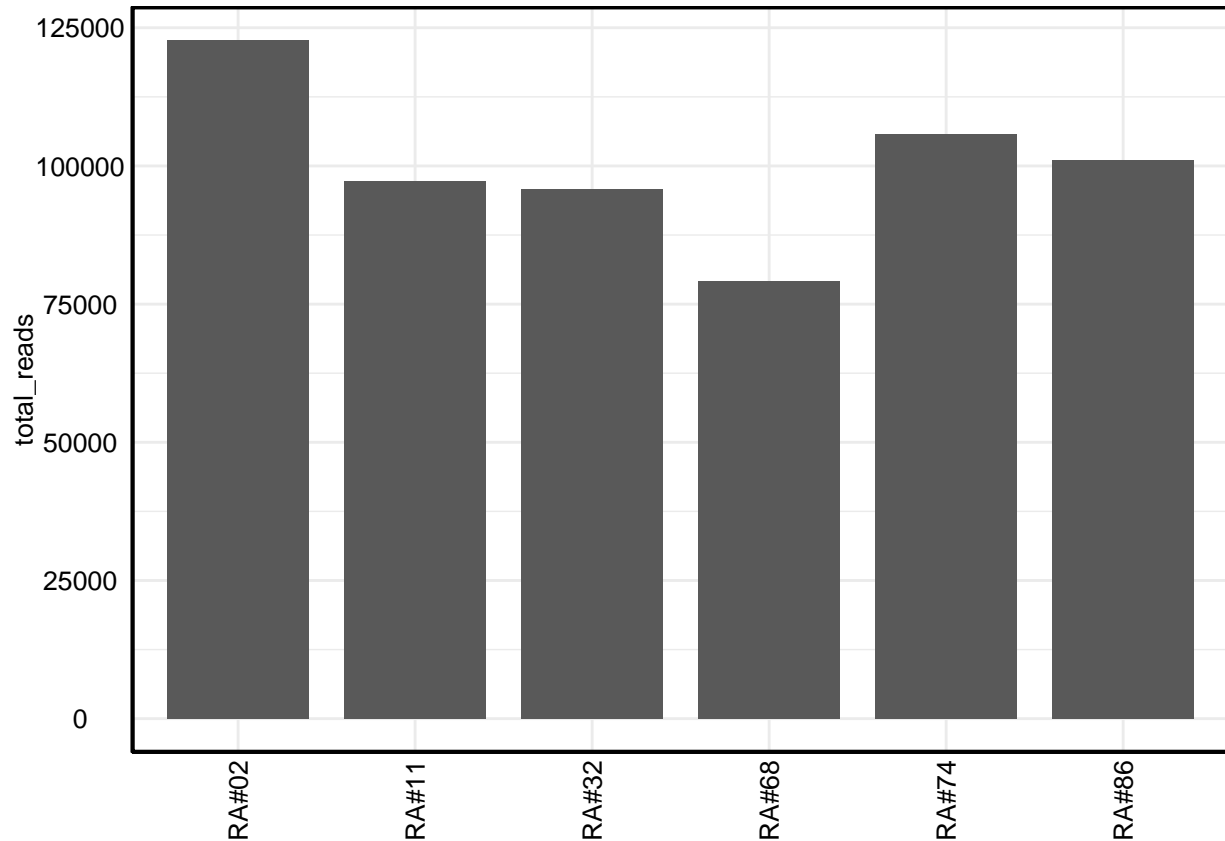
```
#### Stacked barplots of original fecal samples family level
```

```
fecal <- as.data.frame(t(read.delim("/media/deepan/Deepan/RA_Bottini/fecal_samples/level-6.csv", header
colnames(fecal) <- fecal[1,]
fecal <- fecal[c(-1, -nrow(fecal)),]
fecal$Taxa <- rownames(fecal)
fecal <- fecal[, c(ncol(fecal), 1:(ncol(fecal)-1))]
rownames(fecal) <- c()
fecal[,2:ncol(fecal)] <- sapply(fecal[,2:ncol(fecal)], as.numeric)
colnames(fecal) <- c(colnames(fecal)[1], str_replace_all(colnames(fecal[,2:ncol(fecal)]), "RA0", "RA#"))

fecal <- fecal[, c(1, 7, 5, 4, 3, 6, 2)]

fecal$Taxa <- str_replace_all(fecal$Taxa, "\\[", "")
fecal$Taxa <- str_replace_all(fecal$Taxa, "\\]", "")

total_reads <- as.data.frame(apply(fecal[, -1], 2, sum))
colnames(total_reads) <- "total_reads"
total_reads$`sample-ID` <- rownames(total_reads)
ggplot(total_reads, aes(x = `sample-ID`, y = total_reads)) + geom_bar(stat = "identity", width = 0.8) +
```



```
fecal$mean <- apply(fecal[,-1], 1, function(x) mean(x))
fecal <- fecal[order(fecal$mean, decreasing = TRUE),]

fecal <- fecal[fecal$mean > 2.5,]
fecal <- fecal[-nrow(fecal),]

fecal$final_tax <- str_split(str_split(fecal$Taxa, ";f_", simplify= TRUE)[,2], ";", simplify = TRUE)[,]
fecal <- fecal[, c(1, ncol(fecal), 2:(ncol(fecal)-1))]

fecal <- fecal[, -ncol(fecal)]

fecal$final_tax[15:nrow(fecal)] <- "other"
fecal$final_tax <- str_replace(fecal$final_tax, "S24-7", "Muribaculaceae")

fecal_family <- aggregate(fecal[,c(-1, -2)], by = list(fecal$final_tax), sum)
fecal_family$Group.1 <- ordered(fecal_family$Group.1, levels = c(unique(fecal$final_tax)))

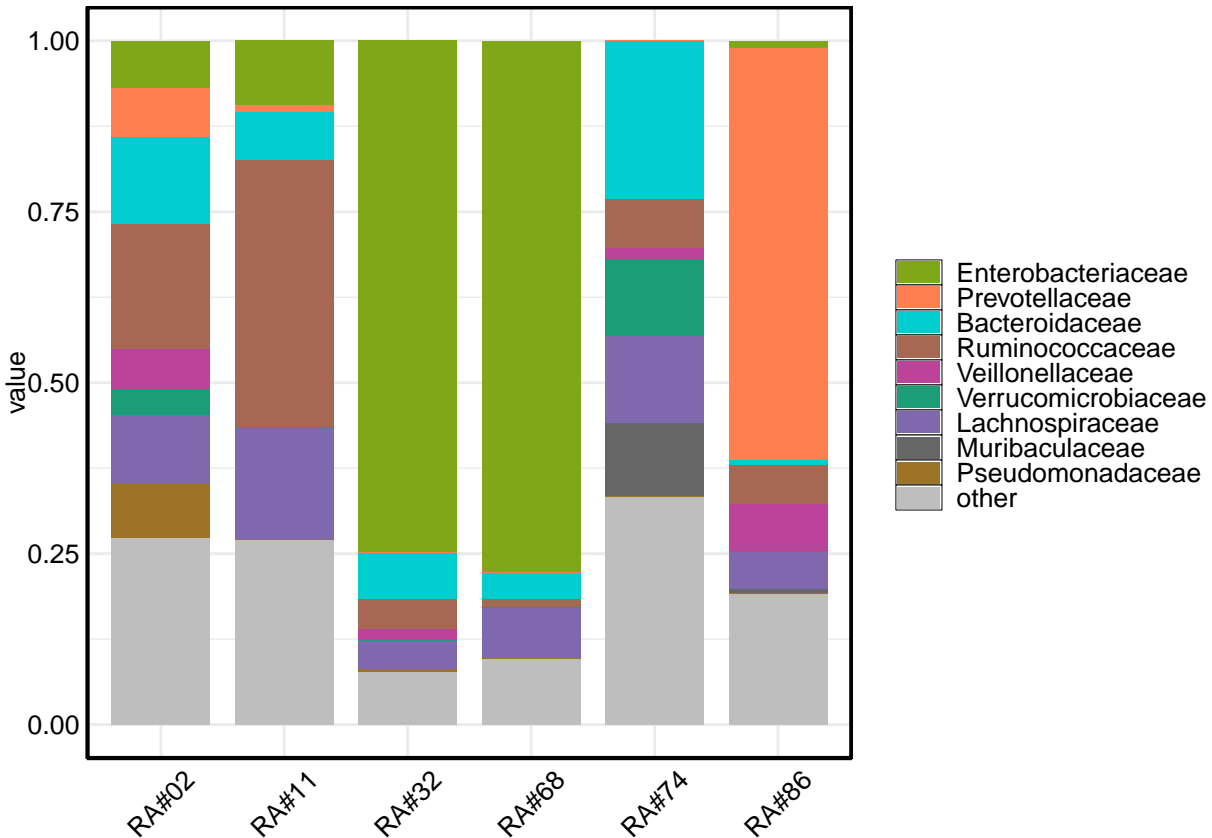
fecal$final_tax <- str_split(str_split(fecal$Taxa, ";g_", simplify= TRUE)[,2], ";", simplify = TRUE)[,]
fecal$final_tax <- ifelse(fecal$final_tax == "", "other", fecal$final_tax)
fecal_genus <- aggregate(fecal[,c(-1, -2)], by = list(fecal$final_tax), sum)

fecal_genus[,-1] <- apply(fecal_genus[,-1], 2, function(x) x/sum(x))
```

###Generating patient fecal sample barplot (Main Fig 4d)

```
#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Mai
ggplot(melt(fecal_family), aes(x = variable, y = value, fill = Group.1)) + geom_bar(stat = "identity",
scale_fill_manual(values = family_cols) +
  theme(text = element_text(size = 15), axis.text.x = element_text(angle = 90)) + theme_boxp()
```

```
## Using Group.1 as id variables
```



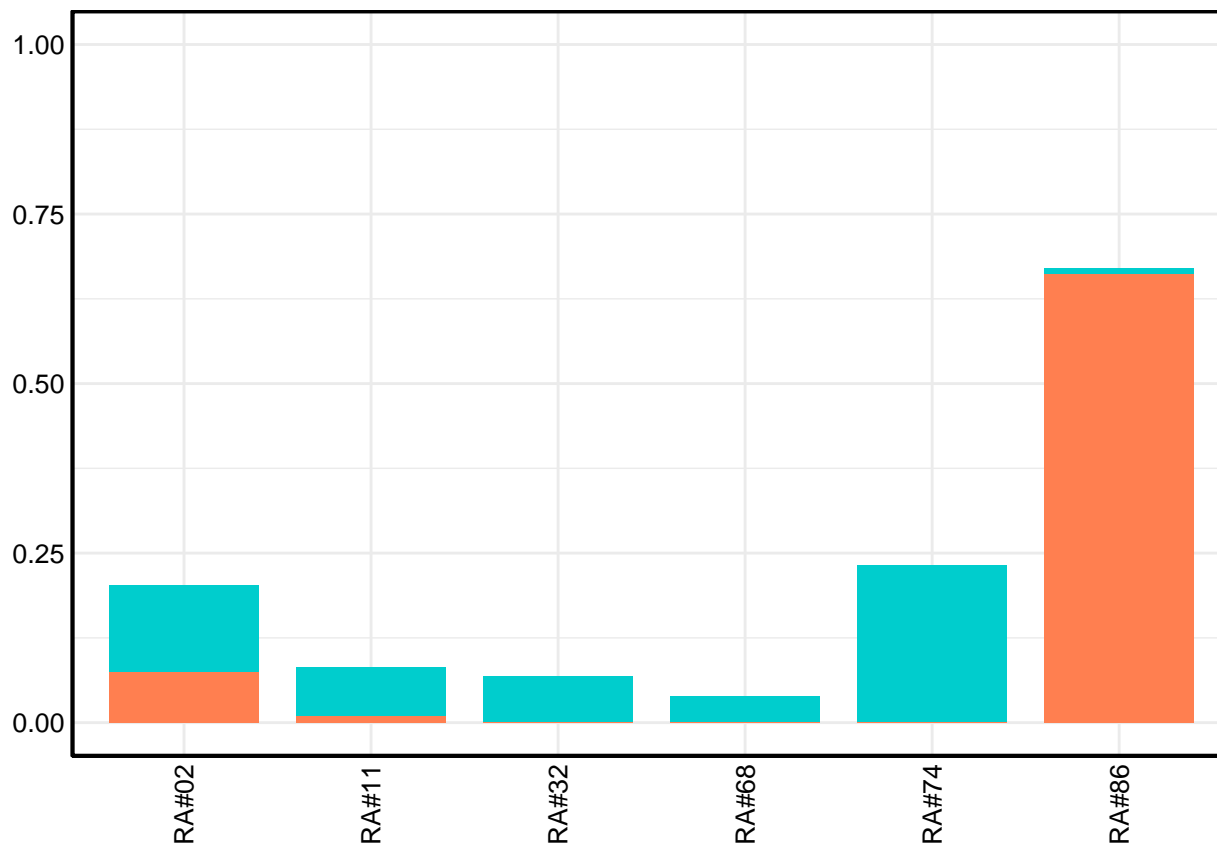
```
#dev.off()
```

```
###Ratio of Prevotellaceae to Bacteroides (Main Fig 4e)
```

```
temp <- fecal_genus[str_detect(fecal_genus$Group.1, "Prevotella|Bacteroides"),]
#temp <- aggregate(temp[,c(-1, -2)], by = list(temp$final_tax), sum)

#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Mai
ggplot(melt(temp), aes(x = variable, y = value, fill = Group.1)) + geom_bar(stat = "identity",width = 0
scale_fill_manual(values = c("Bacteroides"="cyan3", "Prevotella"="coral")) +
scale_y_continuous(limits = c(0, 1)) +
  theme_boxp() + theme(text = element_text(size = 15), axis.text.x = element_text(angle = 90), axis.ti
```

```
## Using Group.1 as id variables
```



```
#dev.off()
```

```
####Stacked barplots for all mice fecal samples
```

```
data <- as.data.frame(t(read.delim("/media/deepan/Deepan/RA_Bottini/final_merged_tables_forwardonly/lev
colnames(data) <- data[1,]
data <- data[-1,]
data$Taxa <- rownames(data)
data <- data[, c(ncol(data), 1:(ncol(data)-1))]
rownames(data) <- c()
data[,2:ncol(data)] <- sapply(data[,2:ncol(data)], as.numeric)
```

```
data[data$Taxa == "k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Pr
```

```
##
## 62 k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella
## B-410b B-411b B-402b B-403c B-406b B-405b B-407c B-404c B-403b B-408c B-409c
## 62 0 2 0 0 7 14 2 8 4 2 0
## B-409b B-402c B-411c B-408b B-410c B-405c B-407b B-406c B-404b B-RA032-297-2
## 62 5 0 8 4 2 2 12 5 4 0
## B-RA032-267-2 B-RA074-011-3 B-RA011-296-2 B-RA011-034-2 B-RA032-411-2
## 62 0 0 0 0 0
## B-RA068-281b-2 B-RA011-281-2 B-RA068-293-2 B-RA011-031-3 B-RA011-014-2
## 62 0 0 0 0 0
## B-RA068-279-3 B-RA032-298-2 B-RA074-019-2 B-RA032-412-3 B-RA011-034-3
```

```

## 62      0      0      0      0      0
##      B-RA011-284-3 B-RA068-292-2 B-RA068-277-2 B-RA032-296-2 B-RA011-018-3
## 62      0      0      0      0      0
##      B-RA068-279-2 B-RA068-295-2 B-RA068-282b-3 B-RA068-294-2 B-RA068-280b-2
## 62      0      0      0      0      0
##      B-RA074-025-2 B-RA074-040-2 B-RA068-293-3 B-RA032-268-2 B-RA032-413-2
## 62      0      0      0      0      0
##      B-RA011-296-3 B-RA011-283-3 B-RA068-277b-2 B-RA068-281b-3 B-RA032-265-2
## 62      0      0      0      0      0
##      B-RA068-291-2 B-RA074-228-3 B-RA074-229-2 B-RA074-011-2 B-RA068-280b-3
## 62      0      0      0      0      0
##      B-RA074-099-2 B-RA074-407-3 B-RA011-283-2 B-RA011-298-3 B-RA032-414-2
## 62      0      0      0      0      0
##      B-RA011-014-3 B-RA032-411-3 B-RA074-227-2 B-RA074-017-2 B-RA011-282-2
## 62      0      0      0      0      0
##      B-RA032-410-2 B-RA074-408-3 B-RA011-300-3 B-RA068-291-3 B-RA074-406-3
## 62      0      0      0      0      0
##      B-RA074-040-3 B-RA074-226-3 B-RA074-407-2 B-RA068-290-2 B-RA068-278-2
## 62      0      0      0      0      0
##      B-RA074-409-3 B-RA011-018-2 B-RA074-408-2 B-RA074-227-3 B-RA068-277b-3
## 62      0      0      0      0      0
##      B-RA032-413-3 B-RA011-299-2 B-RA068-278b-3 B-RA011-297-2 B-RA011-281-3
## 62      0      0      0      0      0
##      B-RA074-036-2 B-RA068-279b-3 B-RA074-406-2 B-RA011-039-3 B-RA074-017-3
## 62      0      0      0      0      0
##      B-RA074-229-3 B-RA068-279b-2 B-RA011-282-3 B-RA011-039-2 B-RA074-226-2
## 62      0      0      0      0      0
##      B-RA011-284-2 B-RA032-414-3 B-RA011-031-2 B-RA011-299-3 B-RA068-278-3
## 62      0      0      0      0      0
##      B-RA032-410-3 B-RA068-290-3 B-RA068-294-3 B-RA074-036-3 B-RA074-099-3
## 62      0      0      0      0      0
##      B-RA011-298-2 B-RA032-412-2 B-RA074-228-2 B-RA068-278b-2 B-RA068-295-3
## 62      0      0      0      0      0
##      B-RA011-300-2 B-RA032-269-2 B-RA068-277-3 B-RA068-292-3 B-RA074-409-2
## 62      0      0      0      0      0
##      B-RA074-019-3 B-RA068-282b-2 B-RA011-297-3 B-RA074-025-3
## 62      0      0      0      0

```

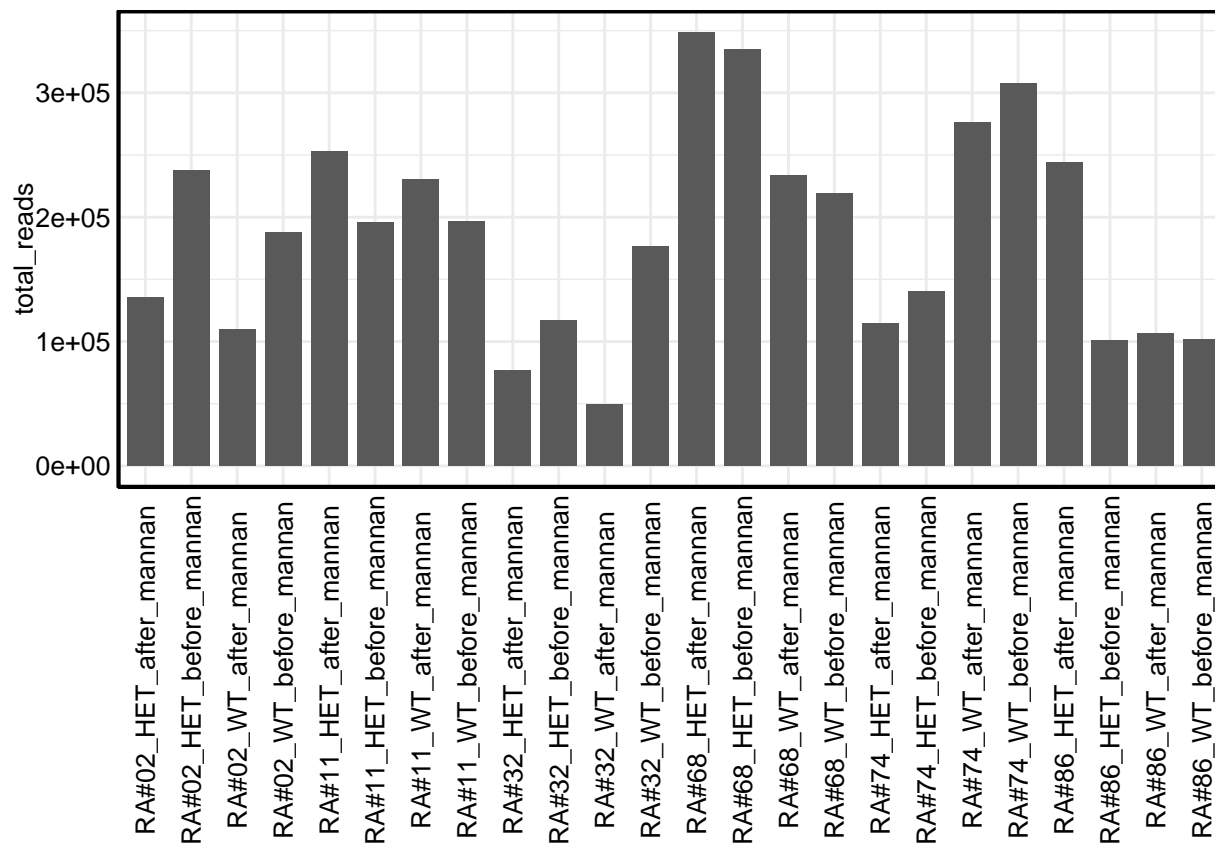
```

data$Taxa <- str_replace_all(data$Taxa, "\\[", "")
data$Taxa <- str_replace_all(data$Taxa, "\\]", "")

total_reads <- as.data.frame(apply(data[, -1], 2, sum))
colnames(total_reads) <- "total_reads"
total_reads$`sample-ID` <- rownames(total_reads)
total_reads <- merge(total_reads, metadata, by = "sample-ID")
ggplot(total_reads, aes(x = PatientTypebeforeafterRA, y = total_reads)) + geom_bar(stat = "identity", w

```





```

data$mean <- apply(data[,-1], 1, function(x) mean(x))
data <- data[order(data$mean, decreasing = TRUE),]

tax_names <- str_split(data$Taxa, ";", simplify = TRUE)

data$final_tax <- str_split(str_split(data$Taxa, ";f_", simplify= TRUE)[,2], ";", simplify = TRUE)[,1]
data$final_tax[data$final_tax == ""] <- "other"

data <- data[, c(1, ncol(data), 2:(ncol(data)-1))]

data <- data[, -ncol(data)]

data$final_tax[25:nrow(data)] <- "other"
data$final_tax <- str_replace(data$final_tax, "S24-7", "Muribaculaceae")

data_family <- aggregate(data[,c(-1, -2)], by = list(data$final_tax), sum)
data_family$Group.1 <- ordered(data_family$Group.1, levels = c(unique(data$final_tax)))

data_family$Group.1 <- ordered(data_family$Group.1, levels = c(unique(data$final_tax)[unique(data$final_tax) != "other"]))

data$final_tax <- str_split(str_split(data$Taxa, ";g_", simplify= TRUE)[,2], ";", simplify = TRUE)[,1]
data$final_tax <- ifelse(data$final_tax == "", "other", data$final_tax)
data_genus <- aggregate(data[,c(-1, -2)], by = list(data$final_tax), sum)

data_genus[,-1] <- apply(data_genus[,-1], 2, function(x) x/sum(x))

```

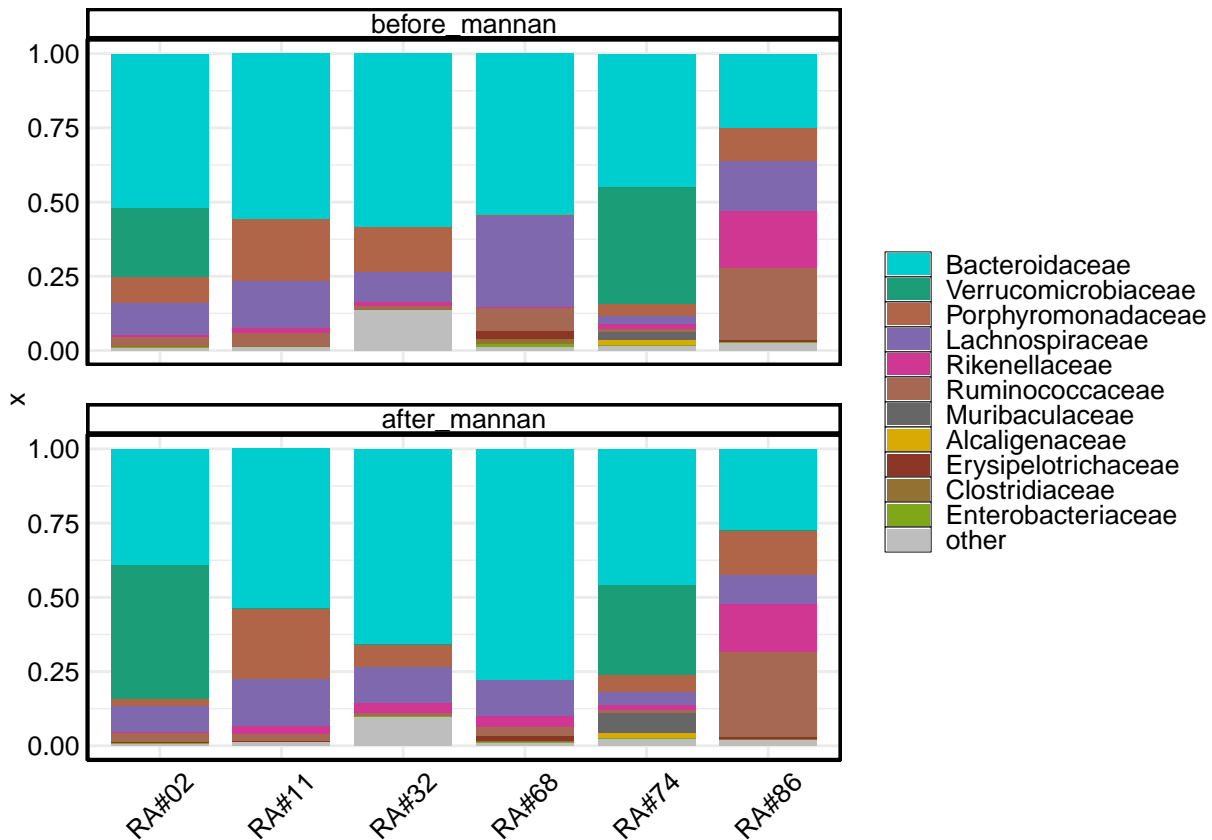
### Generating bar plots of all mice fecal samples (Supp Fig 4)

```
temp <- melt(data_family)
```

## Using Group.1 as id variables

```
colnames(temp) <- c("Group.1", "sample-ID", "value")
temp <- merge(temp, metadata, by = "sample-ID", all.y = FALSE)
temp$BeforeAfterArthritisInd <- ordered(temp$BeforeAfterArthritisInd, levels = c("before_mannan", "after_mannan"))
temp <- temp[temp$MouseType == "HET",]
temp <- aggregate(temp[,c(3)], by = list(temp$PatientbeforeafterRA, temp$Group.1, temp$Patient, temp$PatientbeforeafterRA), FUN = sum)

#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Supp Fig 4")
ggplot(temp, aes(x = Group.3, y = x, fill = Group.2)) + geom_bar(stat = "identity", position = "fill", width = 0.8) +
  scale_fill_manual(values = family_cols) +
  facet_wrap(~Group.5, scales = "free_y", nrow = 2) +
  theme(text = element_text(size = 15), axis.text.x = element_text(angle = 90)) + theme_boxp()
```



```
#dev.off()
```

```
temp <- melt(data_family)
```

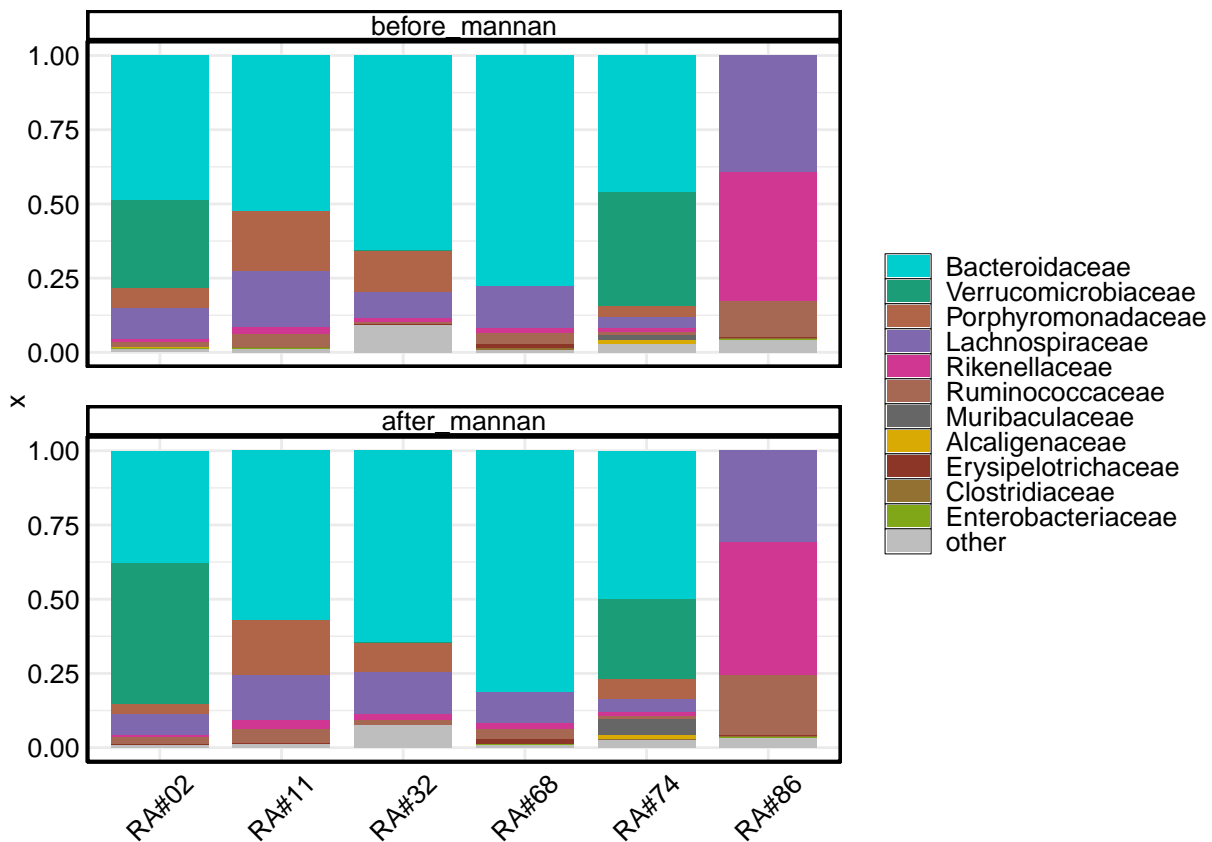
## Using Group.1 as id variables

```

colnames(temp) <- c("Group.1", "sample-ID", "value")
temp <- merge(temp, metadata, by = "sample-ID", all.y = FALSE)
temp$BeforeAfterArthritisInd <- ordered(temp$BeforeAfterArthritisInd, levels = c("before_mannan", "after_mannan"))
temp <- temp[temp$MouseType == "WT",]
temp <- aggregate(temp[,c(3)], by = list(temp$PatientbeforeafterRA, temp$Group.1, temp$Patient, temp$PatientbeforeafterRA), FUN = sum)

#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Supp Fig 4.svg")
ggplot(temp, aes(x = Group.3, y = x, fill = Group.2)) + geom_bar(stat = "identity", position = "fill", width = 0.8) +
  scale_fill_manual(values = family_cols) +
  facet_wrap(~Group.5, scales = "free_y", nrow = 2) +
  theme(text = element_text(size = 15), axis.text.x = element_text(angle = 90)) + theme_boxp()

```



```
#dev.off()
```

```
###ratio of Bacteroides to Prevotella (Supp Fig 4)
```

```

temp <- data_genus[str_detect(data_genus$Group.1, "Prevotella|Bacteroides"),]
temp <- melt(temp)

```

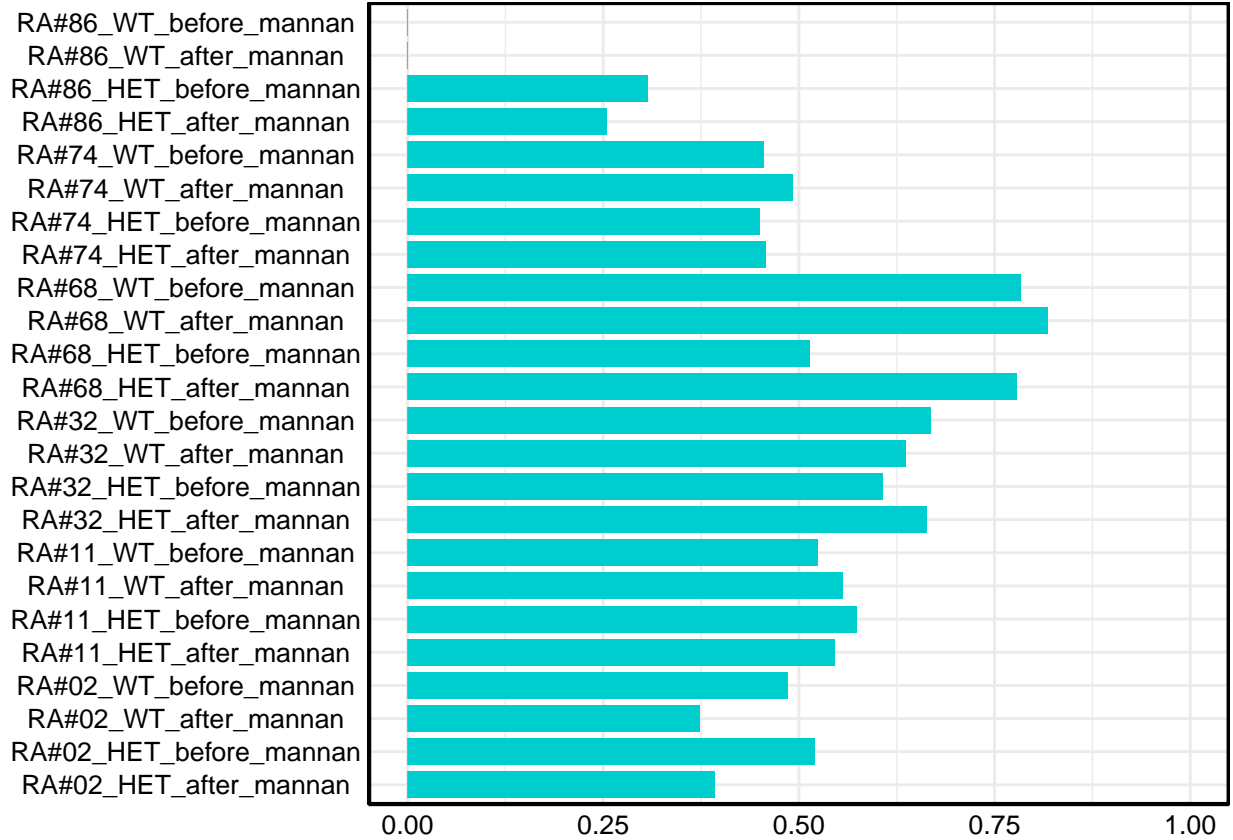
```
## Using Group.1 as id variables
```

```
colnames(temp)[2] <- "sample-ID"
```

```
temp <- merge(temp, metadata[, c(1, 4, 12, 13:ncol(metadata))], by = "sample-ID")

temp <- aggregate(temp[, c(3)], by = list(temp$PatientTypebeforeafterRA, temp$Group.1), mean)

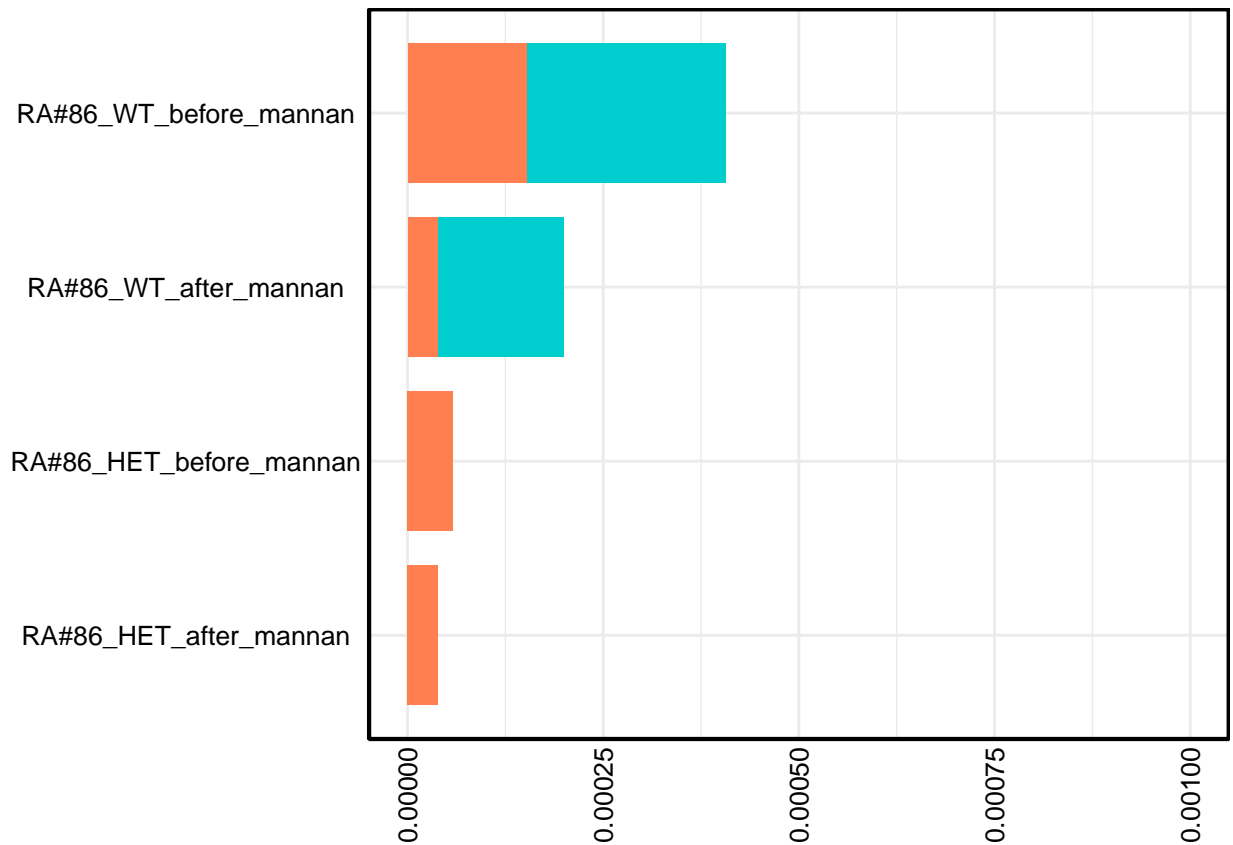
#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Mai
ggplot(temp, aes(x = Group.1, y = x, fill = Group.2)) + geom_bar(stat = "identity", width = 0.8) +
  scale_fill_manual(values = c("Bacteroides"="cyan3", "Prevotella"="coral")) +
  coord_flip() +
  ylim(0,1) +
  theme_boxp() + theme(text = element_text(size = 15), axis.text.x = element_text(angle = 0), axis.tit
```



```
#dev.off()

#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Mai
ggplot(temp[str_detect(temp$Group.1, "RA#86"),], aes(x = Group.1, y = x, fill = Group.2)) + geom_bar(sta
  scale_fill_manual(values = c("Bacteroides"="cyan3", "Prevotella"="coral")) +
  coord_flip() +
  ylim(NA, 0.001) +
  theme_boxp() + theme(text = element_text(size = 15), axis.text.x = element_text(angle = 90), axis.tit
```

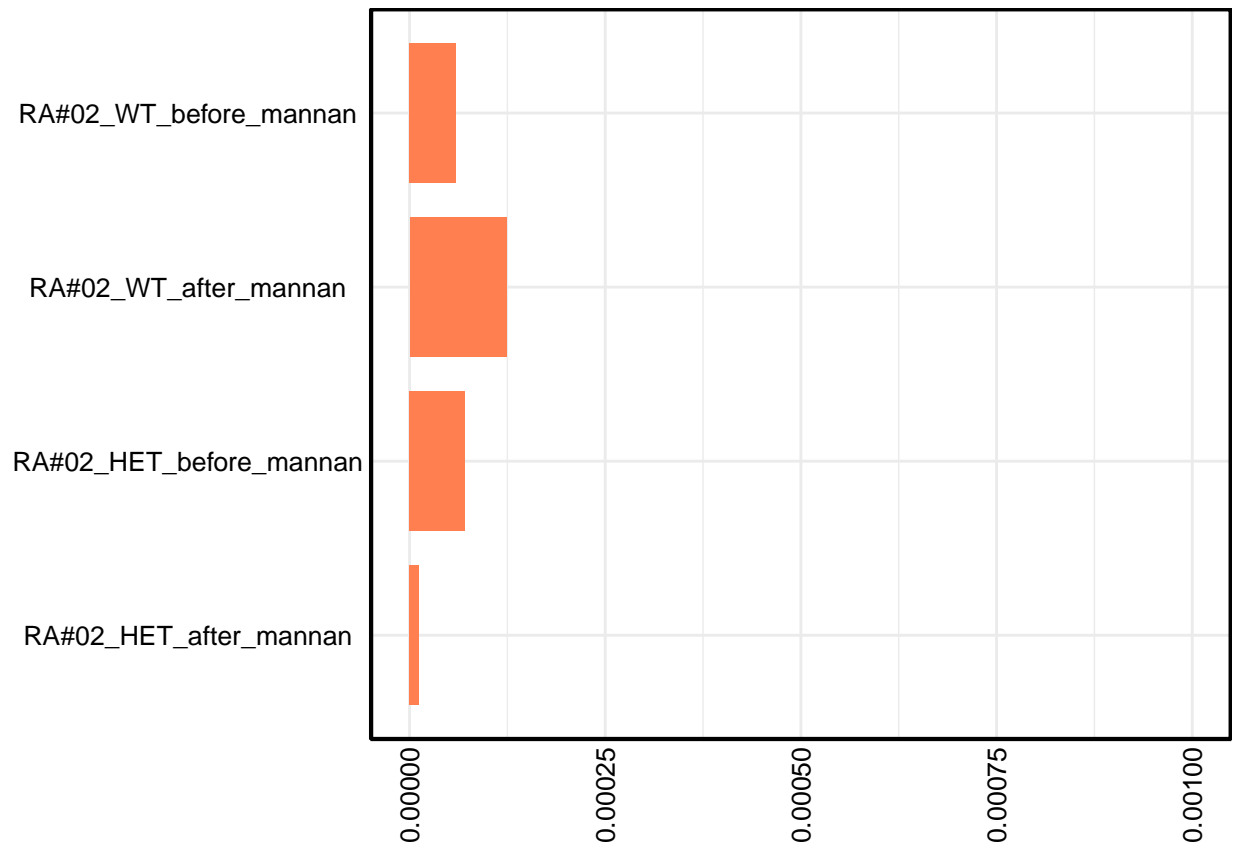
```
## Warning: Removed 2 rows containing missing values ('position_stack()').
```



```
#dev.off()

#svg("/home/deepan/Dropbox/Deepan-Only/UCSD/Zengler Lab/MiscellaneousExpts/RA-Bottini/Ppaer-figures/Main")
ggplot(temp[str_detect(temp$Group.1, "RA#02"),], aes(x = Group.1, y = x, fill = Group.2)) + geom_bar(stat = "identity") +
  scale_fill_manual(values = c("Bacteroides"="cyan3", "Prevotella"="coral")) +
  coord_flip() +
  ylim(NA, 0.001) +
  theme_boxp() + theme(text = element_text(size = 15), axis.text.x = element_text(angle = 90), axis.ticks.x = element_text(size = 10))
```

```
## Warning: Removed 4 rows containing missing values ('position_stack()').
```



*#dev.off()*