

Main_Figure_6

Micheál Mac Aogáin

2023-11-23

```
#Knitr setup
```

```
knitr::opts_chunk$set(echo = TRUE)
```

```
#Load required R packages
```

```
# Load required R packages
```

```
if (!require("pacman")) install.packages("pacman")
```

```
pacman::p_load(ggplot2, tidyverse, vegan, vcd, knitr, funrar, pheatmap, colorspace, ggpubr, RColorBrewer, re
```

```
#Load and wrangle taxonomic data
```

```
MasterANDGK <- read.csv("../Data/R_input_files/MetaDataWsnfResults_GK_final_R2.csv")
```

```
cohort <- subset(MasterANDGK, is.na(SC_AMR_alt) == FALSE & SC_AMR_alt != 0)
```

```
cohort <- tail(cohort, n = 16)
```

```
cohort <- cohort %>% #clinical variables + amr families
```

```
  as_tibble() %>%
```

```
  select(-48:-373, -395:-714)
```

```
cohort$PseuDOM <- ifelse(cohort$DomBact0th_short == "Pseudomonas", "Pseudomonas", "other")
```

```
cohort$HaemDOM <- ifelse(cohort$DomBact0th_short == "Haemophilus", "Haemophilus", "other")
```

```
cohort$FEVfactor <- cut(cohort$FEV1, breaks = c(0, 30, 50, 70, Inf))
```

```
cohort$AetiologyN3 <- ifelse(cohort$Aetiology_short == "postTB", "postInfect", cohort$Aetiology_short)
```

```
#set levels
```

```
cohort$ExacerbatorState <- factor(cohort$ExacerbatorState, levels = c("NonEx", "Exacerbator", "FreqEx"))
```

```
cohort$Country <- factor(cohort$Country, levels = c("SG", "KL", "DD", "MI"))
```

```
cohort$Aetiology_short <- factor(cohort$Aetiology_short, levels = c("idiopathic", "postInfect", "postTB",
```

```
cohort$DomBact0th_short <- factor(cohort$DomBact0th_short, levels = c("Pseudomonas", "Haemophilus", "Streptococcus",
```

```
cohort$SampleID <- factor(cohort$SampleID, levels = cohort$SampleID[order(cohort$SC_V)])
```

```
cohort$FEVfactor <- fct_rev(cohort$FEVfactor)
```

```
cohort <- cohort %>%
```

```
  gather(Resistome, RPKM, Acridine.dye, Aminocoumarin.antibiotic, Aminoglycoside, Antibacterial.free.fatty
```

```
cohort$PseuDOM <- factor(cohort$PseuDOM, levels = c("Pseudomonas", "other"))
```

```
cohort$HaemDOM <- factor(cohort$HaemDOM, levels = c("Haemophilus", "other"))
```

```
cohort$AetiologyN3 <- factor(cohort$AetiologyN3, levels = c("idiopathic", "postInfect", "other"))
```

```
# cohortMT <- subset(cohort, Matching == "Matched") #Matched data
```

```
# cohortPR <- subset(cohort, Paired == "Pair")
```

```
# cohortTR <- subset(cohort, Trio != "NonTrio")
```

```
# cohortUM <- subset(cohort, Matching == "UnMatched")
```

```
# cohort$CTRL <- ifelse(is.na(cohort$Age), "CTRL", "PATIENT")
```

```
cohort$ReadsNonHuman <- factor(cohort$ReadsNonHuman, levels = c("Pre", "Post"))
```

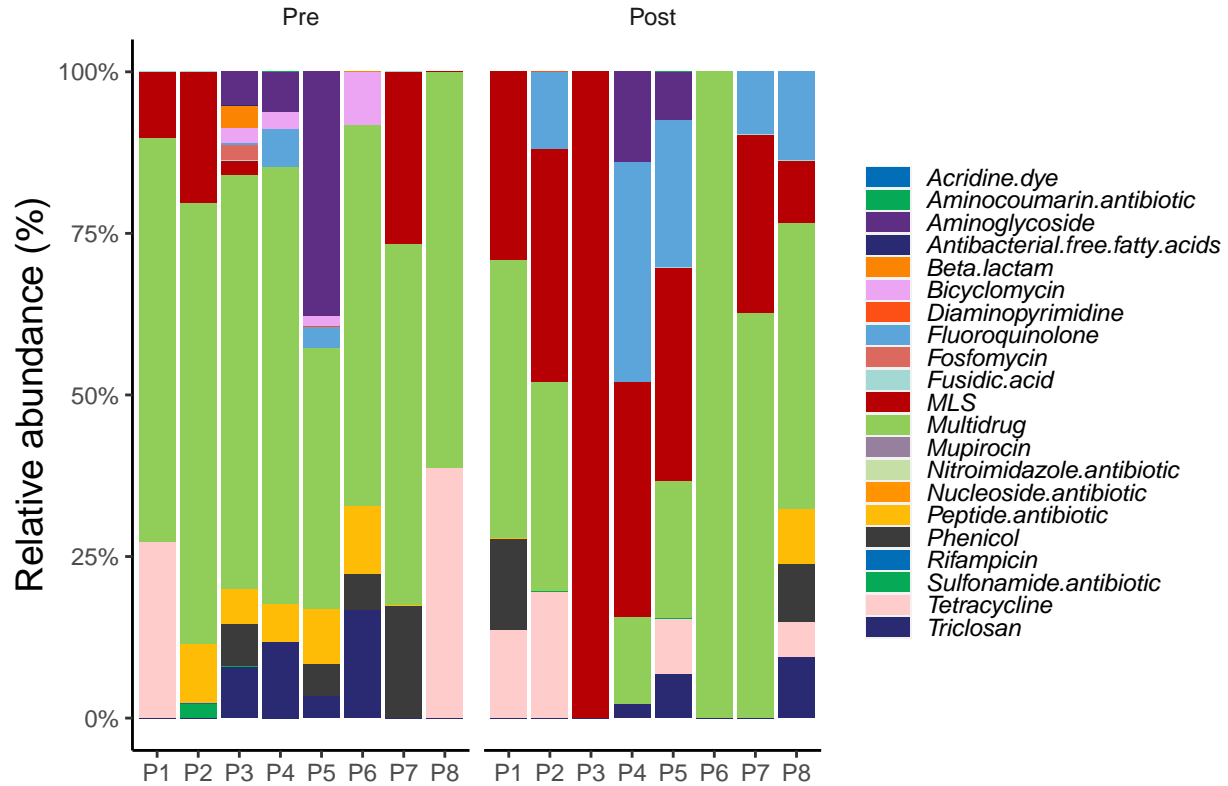
```
cohort$ReadsTrimmed <- factor(cohort$ReadsTrimmed, levels = c("P1", "P2", "P3", "P4", "P5", "P6", "P7", "P8", "P9", "P10", "P11", "P12", "P13", "P14", "P15", "P16", "P17", "P18", "P19", "P20", "P21", "P22", "P23", "P24", "P25", "P26", "P27", "P28", "P29", "P30", "P31", "P32", "P33", "P34", "P35", "P36", "P37", "P38", "P39", "P40", "P41", "P42", "P43", "P44", "P45", "P46", "P47", "P48", "P49", "P50", "P51", "P52", "P53", "P54", "P55", "P56", "P57", "P58", "P59", "P60", "P61", "P62", "P63", "P64", "P65", "P66", "P67", "P68", "P69", "P70", "P71", "P72", "P73", "P74", "P75", "P76", "P77", "P78", "P79", "P80", "P81", "P82", "P83", "P84", "P85", "P86", "P87", "P88", "P89", "P90", "P91", "P92", "P93", "P94", "P95", "P96", "P97", "P98", "P99", "P100"))
```

```
#Figure 6A
```

```

AMR_GK<-ggplot(data=cohort,aes(x=ReadsTrimmed, y=RPKM, fill=Resistome))+
  geom_bar(aes(), stat="identity", position = "fill") +
  scale_fill_manual(values = c("#026EB8", "#06A955", "#5D2E83", "#2A2A73", "#fc8403", "#EBA5F3", "#fc5017", "#
  scale_y_continuous(labels = scales::percent)+
  #scale_x_discrete(labels = c('Non-diseased', 'Bronchiectasis'))+
  theme(legend.position="right",
        #axis.text=element_blank(),
        #axis.title=element_blank(),
        axis.title=element_text(size=14),
        #axis.text.x = element_blank(),
        #axis.text.x = element_text(angle = 90),
        panel.background = element_rect(fill = NA),
        axis.line = element_line(size = 0.5, colour = "black"),
        legend.title = element_blank(),
        legend.text = element_text(face = "italic", size = 9),
        legend.key.height = unit(1.5, "mm"))+
  guides(fill=guide_legend(ncol=1), size = .1)+
  xlab("")+
  ylab("Relative abundance (%)")+
  facet_wrap(~cohort$ReadsNonHuman, scales="free_x")+
  theme(
    strip.background = element_rect(
      color="white", fill="white", size=1, linetype="solid"),
    #strip.text.x = element_blank()
  )
AMR_GK

```



#Figure 6B

```
MasterGK <- read.csv("../Data/R_input_files/GK_master_Taxa25_R2.csv") # Top 25 taxa
```

```
AMRGK <- MasterGK %>%
  as_tibble() %>%
  select(1:50)
```

```
AMR_cols <- colnames(AMRGK[14:39])
```

```
AMRGK <- AMRGK %>%
  gather(AMR, RPKM, AMR_cols, -SampleSeqNo, -SputumSampleNo, -TypeSamples, -TypeSamplesA, -TypeSamplesB)
```

```
AMRGK$AMR <- factor(AMRGK$AMR, levels = AMR_cols)
```

```
AMRGK$TypeSamples <- factor(AMRGK$TypeSamples, levels = c("Pre", "Post"))
```

```
AMRGK <- AMRGK[is.na(AMRGK$TypeSamples) != TRUE, , drop = FALSE]
```

```
n <- 56
```

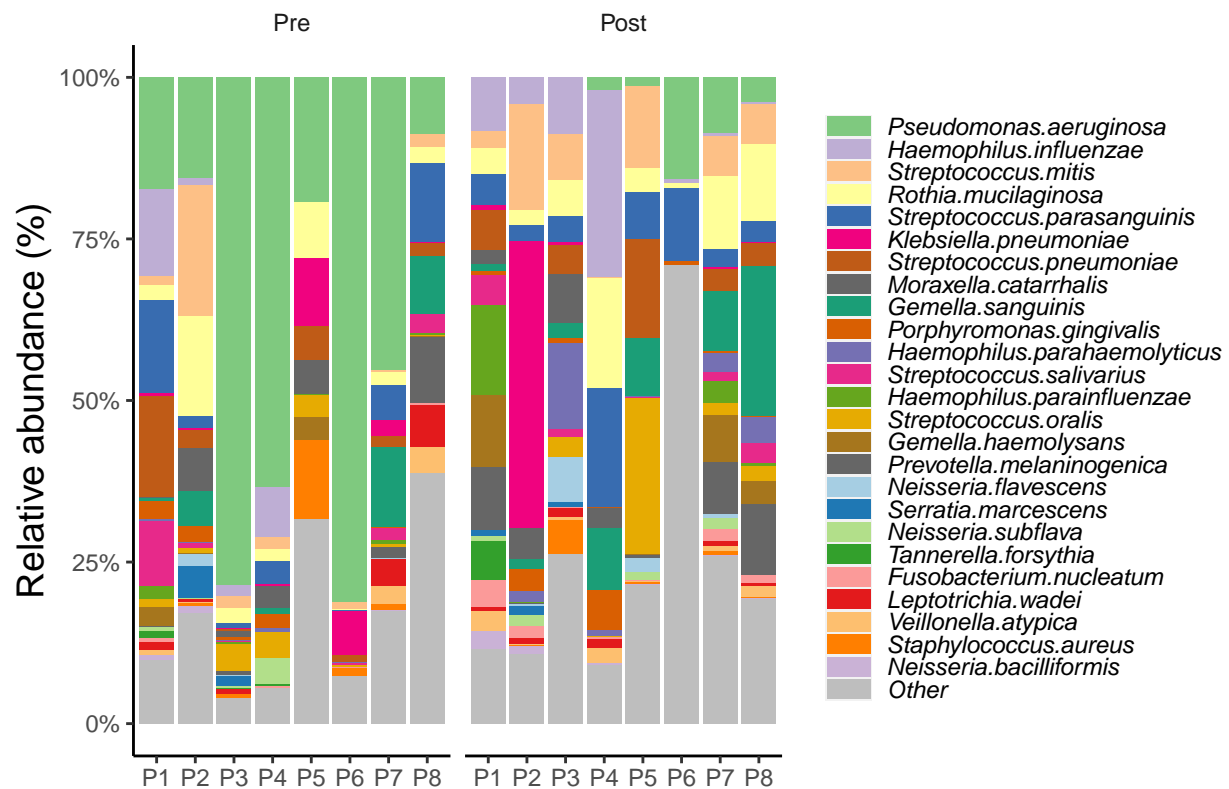
```
qual_col_pals = brewer.pal.info[brewer.pal.info$category == 'qual',]
col_vector = unlist(mapply(brewer.pal, qual_col_pals$maxcolors, rownames(qual_col_pals)))
col_vector_spec <- replace(col_vector, 26, "grey") #
```

```
TAXA_GK <- ggplot(data=AMRGK, aes(x=SputumSampleNo,
  y=RPKM, fill=AMR))+
  scale_fill_manual(values = col_vector_spec) +
  geom_bar(aes(), stat="identity", position = 'fill')+
  scale_y_continuous(labels = scales::percent)+
  #scale_x_discrete(labels = c('IP', 'PI', 'PTB', "other"))+
```

```

theme(legend.position="right",
      #axis.text=element_blank(),
      #axis.title=element_blank(),
      axis.title=element_text(size=14),
      #axis.text.x = element_blank(),
      #axis.text.x = element_text(angle = 90),
      panel.background = element_rect(fill = NA),
      axis.line = element_line(size = 0.5, colour = "black"),
      legend.title = element_blank(),
      legend.text = element_text(face = "italic", size = 9),
      legend.key.height = unit(1.5, "mm"))+
guides(fill=guide_legend(ncol=1), size = .1)+
xlab("")+
ylab("Relative abundance (%)")+
facet_wrap(~AMRGK$TypeSamples, scales="free_x")+
scale_x_discrete(labels = c("P1", "P2", "P3", "P4", "P5", "P6", "P7", "P8")) +
theme(
  strip.background = element_rect(
    color="white", fill="white", size=1, linetype="solid"),
  #strip.text.x = element_blank()
)
TAXA_GK

```



#Figure 6C

```

Master <-read.csv("../Data/R_input_files/MetaDataWsnfResults_GK_final_R2.csv") %>%
  #Master <-read.csv("R_input_files/MetaDataWsnfResults_3.5temp_TJ.csv") %>% # v 3.3 includes raw read
  as_tibble()

AMR_diversity <- Master %>%
  subset(is.na(SC_AMR_alt)==FALSE) %>%
  as_tibble() %>%
  select(1:1,395:673) #for genes
#select(1:1,374:394) #for amr drug class
NAMES_list <- AMR_diversity$SampleID
main_data <- AMR_diversity[AMR_diversity$SampleID %in% NAMES_list, ]
AMR_diversity<-as.matrix(AMR_diversity)
rownames(AMR_diversity) <- AMR_diversity[,1]
AMR_diversity = as.data.frame(subset(AMR_diversity, select = -c(SampleID) ))
AMR_diversity[] <- lapply(AMR_diversity, as.numeric)
AMR_diversity<-AMR_diversity[row.names(AMR_diversity) != "TBS672", , drop = FALSE]

isZero <- base::rowSums(AMR_diversity) == 0
AMR_diversity<-AMR_diversity[!isZero,]

MasterVIZ = Master
MasterVIZ <- MasterVIZ[ MasterVIZ$SampleID %in% row.names(AMR_diversity), ]

diversityttt<-as.data.frame(diversity(AMR_diversity, index = "shannon"))

vegdist(AMR_diversity, "bray")-> Mbiome_PCoA
as.matrix(Mbiome_PCoA)->Mbiome_PCoA
BrayCurtMbiome=cmdscale(Mbiome_PCoA)
#ordiplot (BrayCurtMbiome, display = 'species', type = 'text')
BCords<-scores(BrayCurtMbiome)
BCords<-(as.data.frame(t(BCords)))
BCords<-as.data.frame(t(BCords))

MasterVIZ$Dim1<-BCords$Dim1
MasterVIZ$Dim2<-BCords$Dim2

MasterVIZ$Country <- factor(MasterVIZ$Country, levels = c("SG", "KL", "DD", "MI"))
MasterVIZ$Aetiology_short<- factor(MasterVIZ$Aetiology_short, levels=c("idiopathic", "postInfect", "pos

MasterVIZ$PrePost<-ifelse(MasterVIZ$ReadsNonHuman == "Pre","Pre",NA)
MasterVIZ$PrePost<-ifelse(MasterVIZ$ReadsNonHuman == "Post","Post", MasterVIZ$PrePost)
MasterVIZ$PrePost<-ifelse(is.na(MasterVIZ$PrePost) == TRUE, MasterVIZ$SC_AMR_alt, MasterVIZ$PrePost)

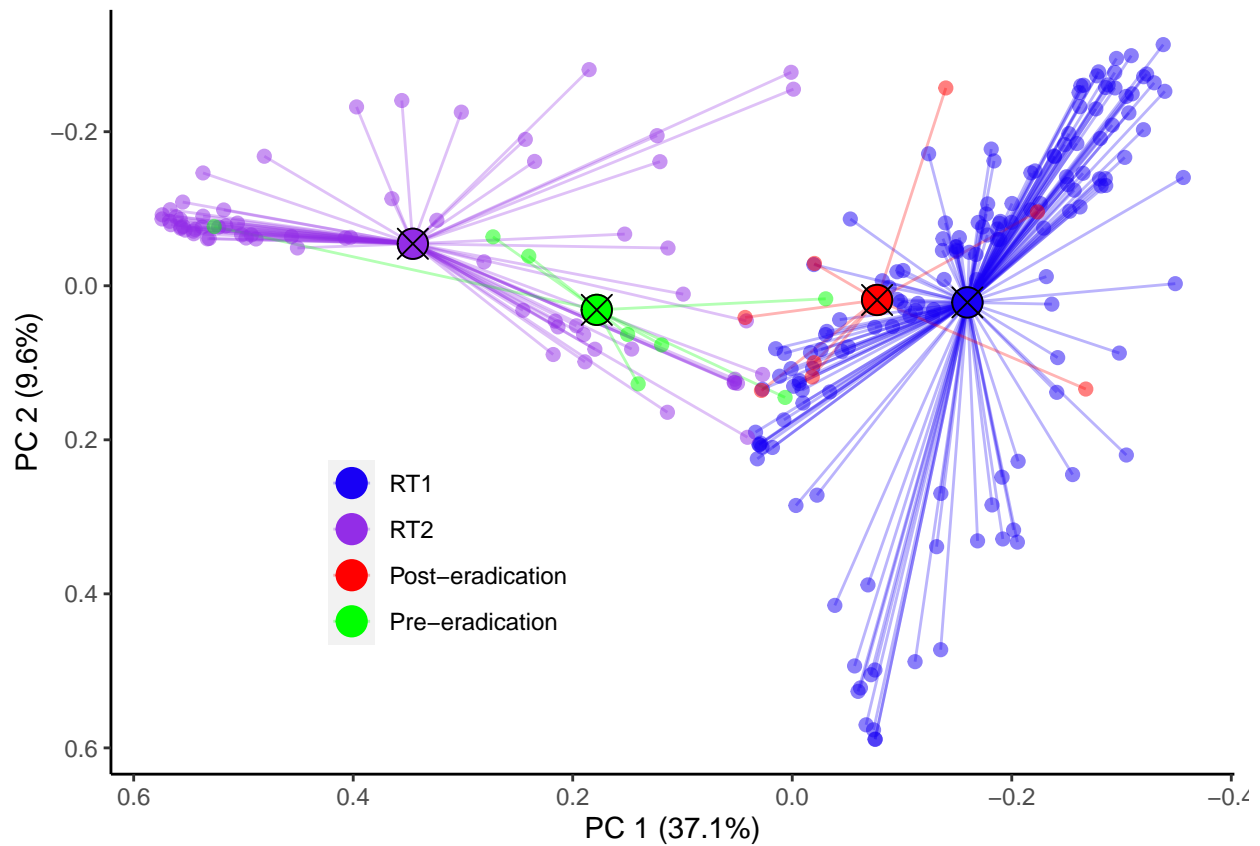
gg <- data.frame(cluster=factor(MasterVIZ$PrePost), x=MasterVIZ$Dim1, y=MasterVIZ$Dim2, grp=MasterVIZ$S
# calculate group centroid locations
centroids <- aggregate(cbind(x,y)~cluster,data=gg,mean)
# merge centroid locations into ggplot dataframe
gg <- merge(gg,centroids,by="cluster",suffixes=c("", ".centroid"))
# generate star plot...
PA_erd_PCA<-ggplot(gg) +
  #scale_col_manual(values=c(16, 16, 16,16))+
  scale_linetype_identity() +
  geom_segment(aes(x=x.centroid, y=y.centroid, xend=x, yend=y, colour = cluster),alpha = 0.3)+

```

```

geom_point(aes(x=x,y=y, colour = cluster), size = 2, alpha = 0.5) + #can add ",shape = shape" in aes
#geom_point(aes(x=x,y=y, colour = cluster, shape = shape), size = 2) +
geom_point(data=centroids, aes(x=x, y=y, color=cluster), size=5) +
geom_point(data=centroids, aes(x=x, y=y, color=cluster), size=5, shape = 13, colour = "black") +
scale_colour_manual(values = c("#1800F5", "#932DE7", "red", "green"), labels = c("RT1", "RT2", "Post-eradication", "Pre-eradication"),
labs(colour="",
      x = "PC 1 (37.1%)", y = "PC 2 (9.6%)")+
theme(legend.position=c(0.3,0.3),
      legend.title = element_blank(),
      axis.line = element_line(size = 0.5, colour = "black"),
      panel.background = element_rect(fill = NA),
)+
scale_y_reverse()+
scale_x_reverse()+
guides(colour = guide_legend(reverse = FALSE))
PA_erd_PCA

```



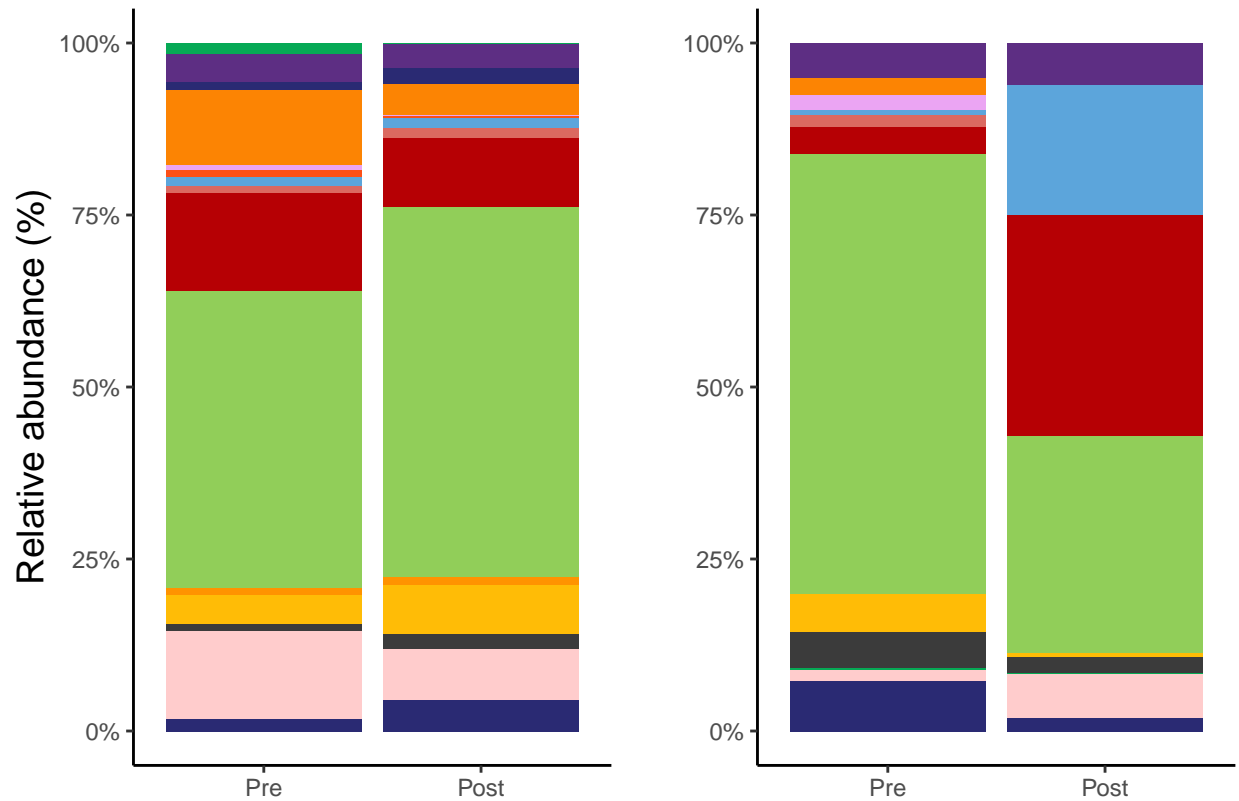
```

# checkEig<-capscale(AMR_diversity ~1)
# Eig <-eigenvals(checkEig)
# print(Eig[1:2] / sum(Eig))

Adonis_ERD<-tail(AMR_diversity, 16)
Adonis_ERD_meta<-tail(MasterVIZ, 16)

adonis2(Adonis_ERD~ReadsNonHuman , data = Adonis_ERD_meta, method = "bray",permutations=9999, strata = A

```

```
adonis2(Adonis_ERD~ReadsNonHuman , data = Adonis_ERD_meta, method = "bray",permutations=9999, strata = A
```

```
## Permutation test for adonis under reduced model
```

```
## Terms added sequentially (first to last)
```

```
## Blocks: strata
```

```
## Permutation: free
```

```
## Number of permutations: 255
```

```
##
```

```
## adonis2(formula = Adonis_ERD ~ ReadsNonHuman, data = Adonis_ERD_meta, permutations = 9999, method =
```

```
##          Df SumOfSqs      R2      F Pr(>F)
```

```
## ReadsNonHuman  1   0.6139 0.09332 1.4409 0.01562 *
```

```
## Residual      14   5.9643 0.90668
```

```
## Total         15   6.5782 1.00000
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```