VitaminDPaperCode

Ankan Choudhury

2023-04-21

R Markdown

```r
#Figure 2(A)
library(tidyverse)
library(readxl)
library(ggpubr)
sample_metadata <- read_excel("sample-metadata.xlsx") #Contains the metadata
head(sample_metadata)
```

```
## # A tibble: 6 x 29
##   Sample  Sampl~1   Day Subject Cohort Cohor~2   AGE SEX    RACE Vit D~3 Vit D~4
##   <chr>   <chr>   <dbl> <chr>   <chr>  <chr>   <dbl> <chr> <dbl>   <dbl>   <dbl>
## 1 VDMT00~ VDMT00~     1 VDMT001 Treat~ Treatm~    19 F         1    30.9    70.2
## 2 VDMT00~ VDMT00~     7 VDMT001 Treat~ Treatm~    19 F         1    30.9    70.2
## 3 VDMT00~ VDMT00~    14 VDMT001 Treat~ Treatm~    19 F         1    30.9    70.2
## 4 VDMT00~ VDMT00~    78 VDMT001 Treat~ Treatm~    19 F         1    30.9    70.2
## 5 VDMT00~ VDMT00~     1 VDMT002 Treat~ Treatm~    35 M         5    26.3    52.5
## 6 VDMT00~ VDMT00~     7 VDMT002 Treat~ Treatm~    35 M         5    26.3    52.5
## # ... with 18 more variables: Height <dbl>, Wt_pre <dbl>, Wt_post <dbl>,
## #   BMI_pre <dbl>, BMI_post <dbl>, LBM_pre <dbl>, LBM_post <dbl>,
## #   PBF_pre <dbl>, PBF_post <dbl>, Sleep_week <dbl>, Sleep_weekend <dbl>,
## #   Vig_Ex_daily <dbl>, SIT_hr <dbl>, 'Total HEI-2015 Score' <dbl>,
## #   Vit_D_pre_Cat <chr>, Vit_D_post_Cat <chr>, 'Vit_Change%' <dbl>,
## #   'Vit_Change%_Cat' <chr>, and abbreviated variable names 1: Sample_Name,
## #   2: 'Cohort&Day', 3: 'Vit D_pre', 4: 'Vit D_post'
```

```r
VitaminDChange<-sample_metadata%>% #Creating a data frame with the serum Vitamin D changes in each subj
  select(Subject,Cohort,VitD_pre=`Vit D_pre`,VitD_post=`Vit D_post`)%>%
  distinct()%>%
  mutate(VitD_change =VitD_post-VitD_pre,
         VitD_pcchange =VitD_change/VitD_pre*100,
         VitD_pre_cat=if_else(VitD_pre<20,"Deficient (<20 ng/ml)",if_else(VitD_pre<30,"Insufficient (20-
head(VitaminDChange)
```

```
## # A tibble: 6 x 7
##   Subject Cohort    VitD_pre VitD_post VitD_change VitD_pcchange VitD_pre_cat
##   <chr>   <chr>        <dbl>     <dbl>       <dbl>         <dbl> <chr>
## 1 VDMT001 Treatment     30.9      70.2        39.2          127. Sufficient (>3~
## 2 VDMT002 Treatment     26.3      52.5        26.2           99.6 Insufficient (~
```

```
## 3 VDMT003 Treatment      20.9      75.7      54.8        263.    Insufficient (~
## 4 VDMT004 Treatment      28.9      61.9      33.0        114.    Insufficient (~
## 5 VDMT005 Treatment      30.1      49.1      19.0         63.1 Sufficient (>3~
## 6 VDMT006 Placebo        39.6      36.3      -3.34        -8.41 Sufficient (>3~
```

```r
ggplot(VitaminDChange,aes(y=Cohort,x=VitD_change,fill=Cohort))+
  geom_boxplot(color="black",linewidth=1,width=0.25)+
  geom_dotplot(binaxis="x",binwidth = 3,dotsize=1.5,stroke=1,color="black",method = "histodot")+
  guides()+
  labs(y="Cohort",x="Change in serum Vitamin D\nlevels over 12 weeks (ng/mL)")+
  theme(text = element_text(size = 25),
        panel.background = element_rect(fill="white",color="grey",size=1),
        panel.border = element_rect(colour = "black", fill=NA, size=1),
        panel.grid.major= element_blank(),
        axis.text=element_text(size=25,color="black")
  )
```
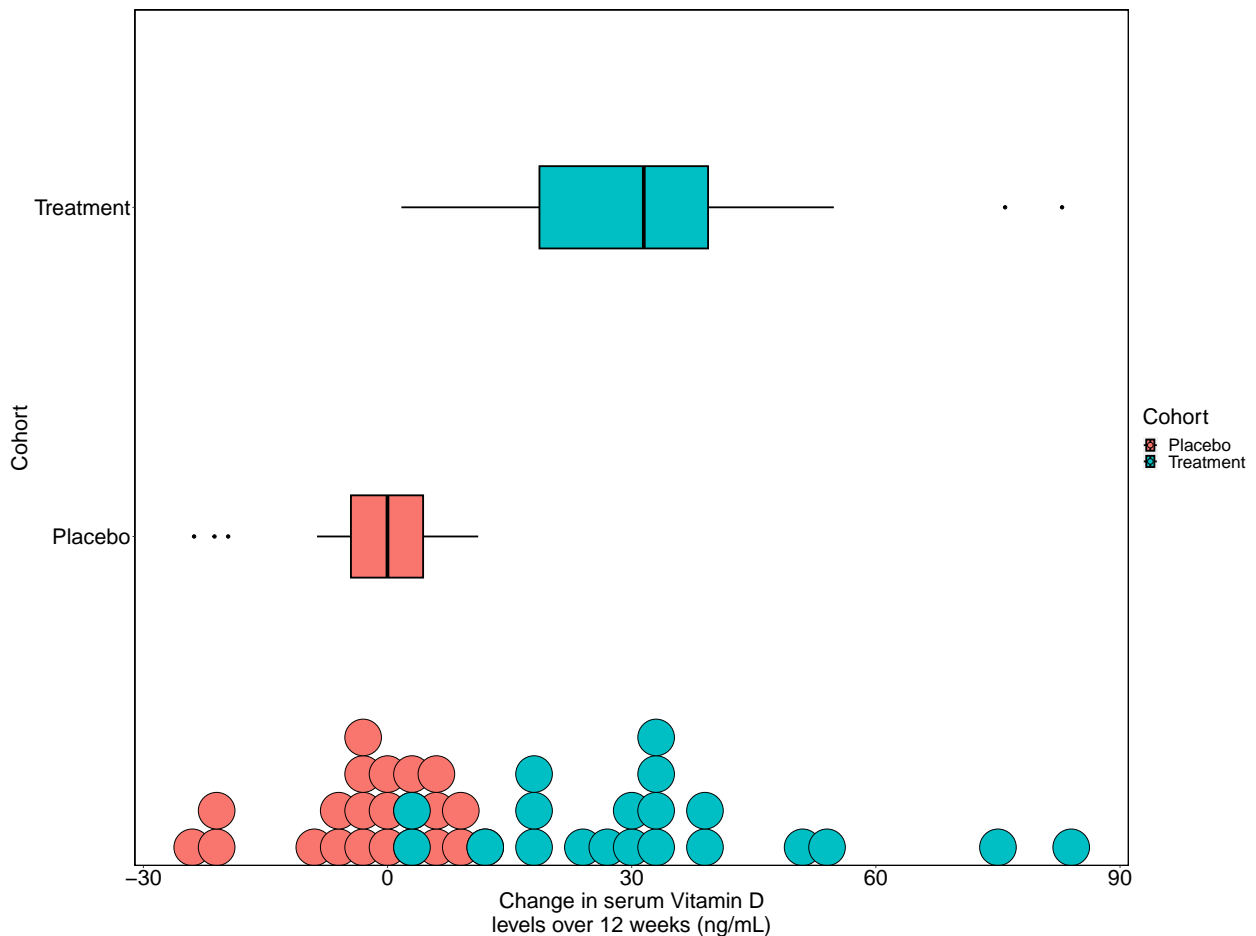


```r
#Figure 2(B)
library(Polychrome) #making  a vector of 50 random colors
set.seed(07181990)
P50 <- createPalette(50, c("#FF0000", "#00FF00", "#0000FF"), range = c(0, 100))
P50 <- sortByHue(P50)
P50 <- as.vector(t(matrix(P50, ncol=10)))
```

```r
names(P50) <- NULL

Genus <- read_excel("level-6.xlsx",
                    sheet = "Genus")%>%
  pivot_longer(-c("index","Sample_Name","Day","Subject","Cohort")) #getting the relative abundance % at
head(Genus)
```

```
## # A tibble: 6 x 7
##   Sample_Name   Day Subject Cohort    index      name                   value
##   <chr>       <dbl> <chr>   <chr>     <chr>      <chr>                   <dbl>
## 1 VDMT001-1       1 VDMT001 Treatment VDMT001-1 unidentified Lachnospirac~  2.49
## 2 VDMT001-1       1 VDMT001 Treatment VDMT001-1 Blautia                  8.10
## 3 VDMT001-1       1 VDMT001 Treatment VDMT001-1 Bacteroides              7.71
## 4 VDMT001-1       1 VDMT001 Treatment VDMT001-1 Faecalibacterium         7.85
## 5 VDMT001-1       1 VDMT001 Treatment VDMT001-1 Bifidobacterium          2.03
## 6 VDMT001-1       1 VDMT001 Treatment VDMT001-1 Subdoligranulum          1.90
```

```r
Genus%>%
  group_by(Cohort,Day,name)%>% #summing up the %RA of each taxon by Cohort and Day and filtering out on
  summarize(sum=sum(value))%>%
  filter(sum>15)%>%
  ggplot(aes(fill=reorder(name,desc(sum)),y=sum,x=factor(Day),width=0.75))+
  geom_bar(position="fill",stat="identity",color="black")+
  scale_y_continuous(labels = scales::percent_format())+
  labs(y="Relative Abundance",x="Day",fill="Genus",title="Changes at Genus Level")+
  guides(fill=guide_legend(ncol=2))+
  facet_grid(~Cohort)+
  scale_fill_manual(values=P50)+
  theme(
    panel.background = element_rect(fill="white"),
    text = element_text(size=25,color="black"),
    panel.border = element_rect(size=1,fill="NA"),
    axis.text=element_text(size=15,color="black"),
    legend.text = element_text(size=15,color="black"),
    legend.title=element_text(size=25,face="bold",color="black"),
    legend.key.size = unit(1, 'cm'))
```
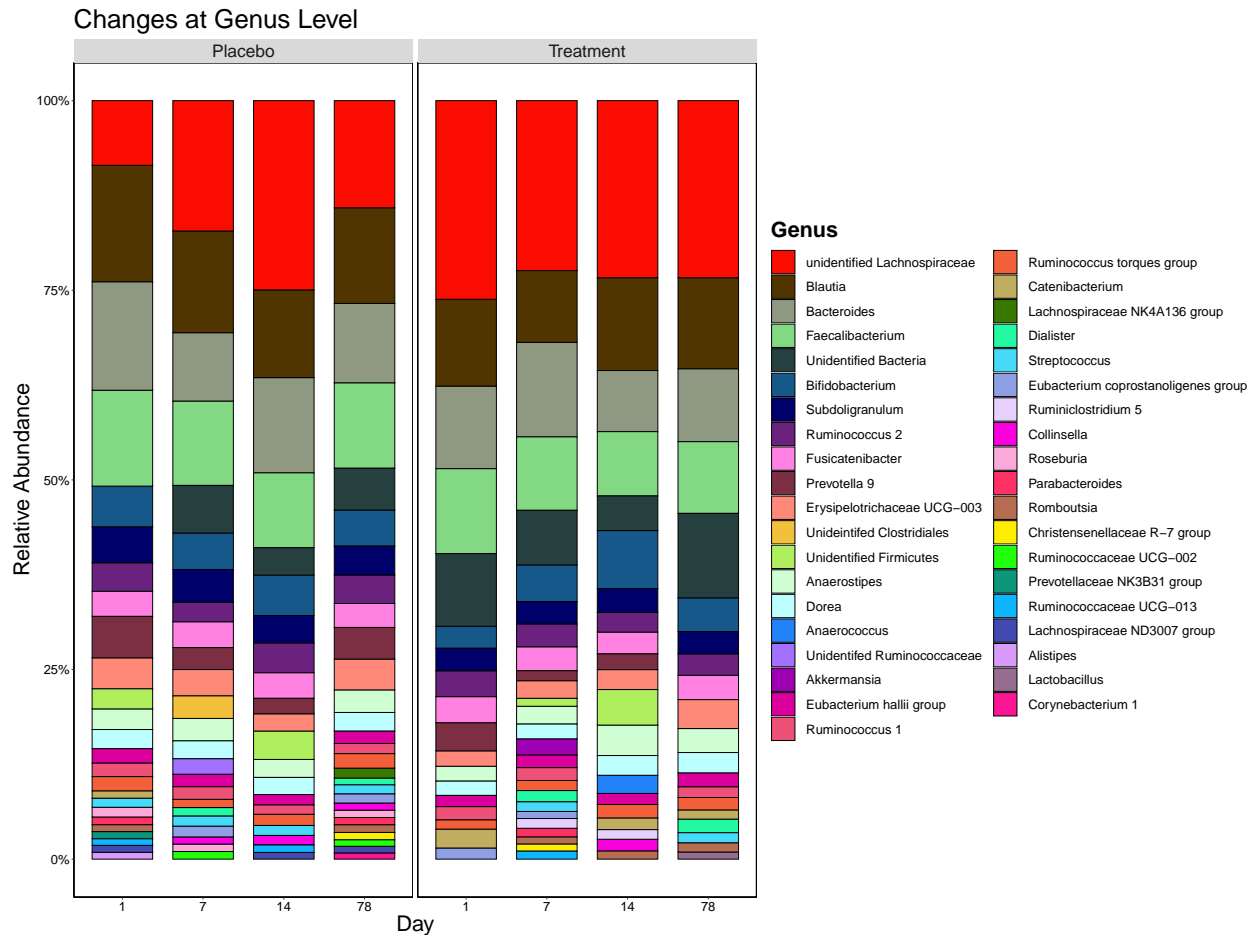
Changes at Genus Level

```
#Figure 3A-D
AlphaVDMT <- read_excel("AlphaVDMT.xlsx",sheet = "Sheet1")# getting the alpha diversity indices generat
head(AlphaVDMT)
```

```
## # A tibble: 6 x 11
##   sampl~1 Subject Cohort   Day shannon faith   OTU HEI   VitD_~2 VitD_~3 VitD_~4
##   <chr>   <chr>   <chr>  <dbl>   <dbl> <dbl> <dbl> <chr> <chr>   <chr>   <chr>
## 1 VDMT02~ VDMT029 Place~     1    6.09  8.88  163. High  29.608~ Low     Insuff~
## 2 VDMT03~ VDMT031 Place~     1    5.05  7.48  133. High  40.311  Standa~ Suffic~
## 3 VDMT03~ VDMT034 Place~     1    6.14  7.87  160. High  43.87   Standa~ Suffic~
## 4 VDMT03~ VDMT037 Place~     1    6.13  8.73  137. High  35.125~ Standa~ Suffic~
## 5 VDMT03~ VDMT038 Place~     1    6.62 13.6   229. High  21.4    Low     Insuff~
## 6 VDMT04~ VDMT040 Place~     1    6.75 10.9   236. High  NA      High    Suffic~
## # ... with abbreviated variable names 1: `sample-id`, 2: VitD_pre,
## #   3: VitD_pre_cat1, 4: VitD_pre_cat
```

```
compare=list(c("Placebo","Treatment"))
compare2=list(c("1","7"),c("7","14"),c("14","78"),c("1","78"))

#Shannon vs Cohort faceted by day
ggplot(data=AlphaVDMT,aes(x=Cohort,y=shannon))+
  geom_boxplot(width=0.35,size=1.5,fill="grey")+
  geom_point(aes(group=Subject,fill=VitD_pre_cat),shape=21,color="black",size=5)+
```
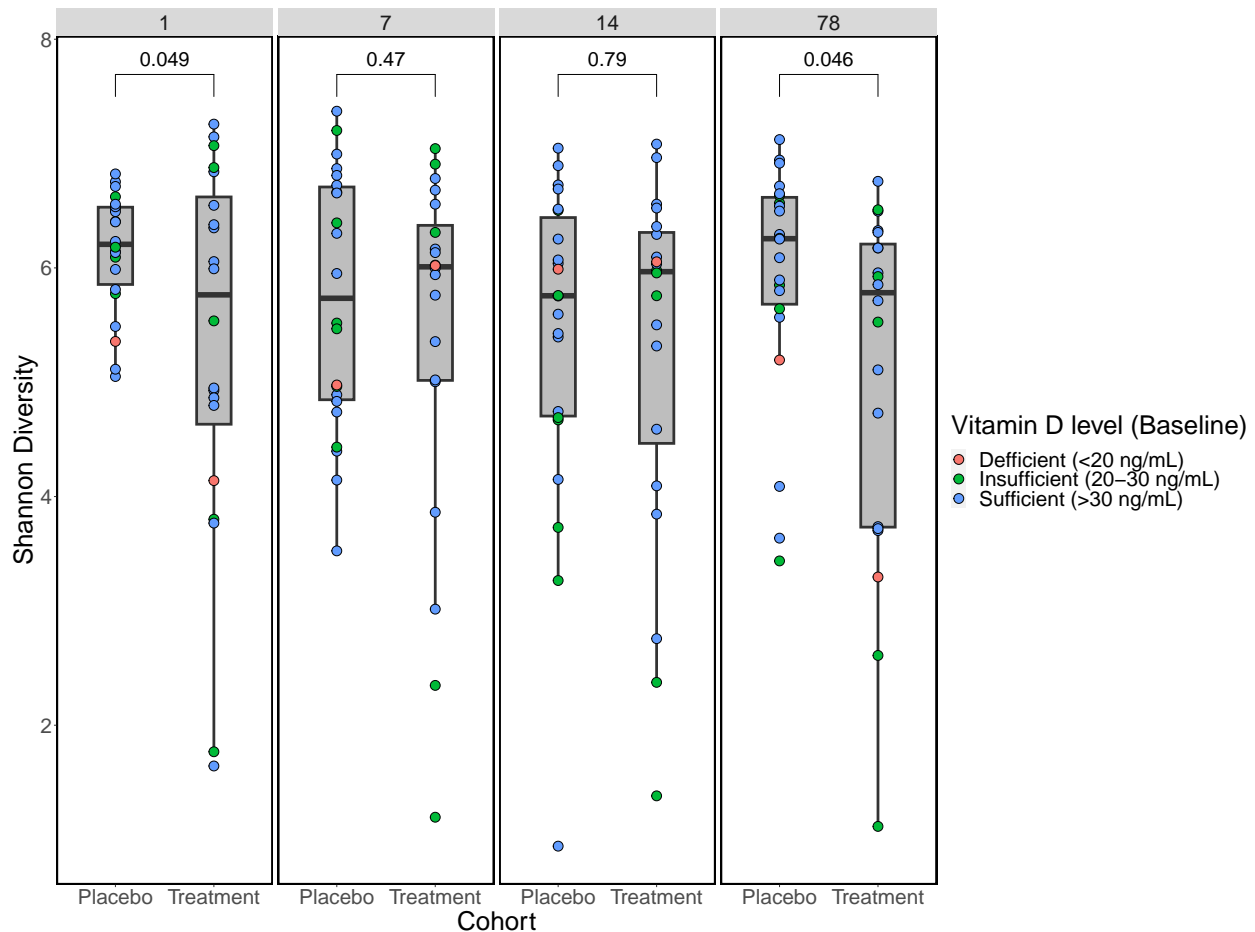
```
labs(x="Cohort",fill="Vitamin D level (Baseline)",y="Shannon Diversity")+
stat_compare_means(comparisons=compare,method="t.test",size=8)+
guides(color="none")+
facet_grid(~factor(Day))+
theme(
  panel.background = element_rect(fill="white"),
  text = element_text(size=30,color="black"),
  panel.border = element_rect(size=2,fill="NA"),
)
```

[1] FALSE


```
#FaithPD vs Cohort faceted by day
ggplot(data=AlphaVDMT,aes(x=Cohort,y=faith))+
  geom_boxplot(width=0.35,size=1.5,fill="grey")+
  geom_point(aes(group=Subject,fill=VitD_pre_cat),shape=21,color="black",size=5)+
  labs(x="Cohort",fill="Vitamin D level (Baseline)",y="Faith Phylogenetic Diversity")+
  stat_compare_means(comparisons=compare,method="t.test",size=8)+
  guides(color="none")+
  facet_grid(~factor(Day))+
  theme(
    panel.background = element_rect(fill="white"),
```

```
    text = element_text(size=30,color="black"),
    panel.border = element_rect(size=2,fill="NA"),
  )
```

[1] FALSE


```
#Observed OTUs vs Cohort faceted by day
ggplot(data=AlphaVDMT,aes(x=Cohort,y=OTU))+
  geom_boxplot(width=0.35,size=1.5,fill="grey")+
  geom_point(aes(group=Subject,fill=VitD_pre_cat),shape=21,color="black",size=5)+
  labs(x="Cohort",fill="Vitamin D level (Baseline)",y="Observed OTUs")+
  stat_compare_means(comparisons=compare,method="t.test",size=8)+
  guides(color="none")+
  facet_grid(~factor(Day))+
  theme(
    panel.background = element_rect(fill="white"),
    text = element_text(size=30,color="black"),
    panel.border = element_rect(size=2,fill="NA"),
  )
```
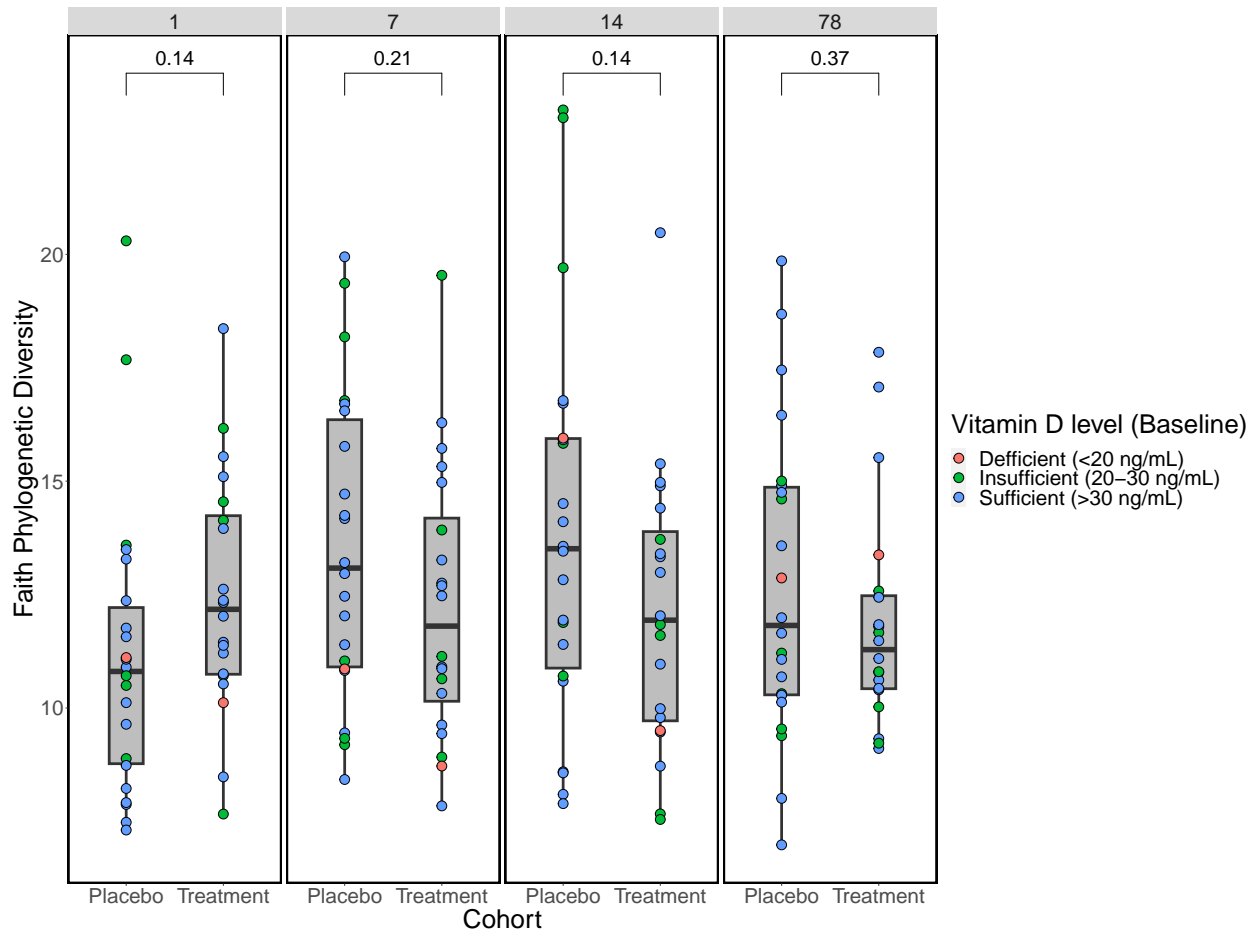
[1] FALSE

```
#Shannon vs Day faceted by Cohort
ggplot(data=AlphaVDMT,aes(x=factor(Day),y=shannon))+
  geom_boxplot(width=0.35,size=1.5,fill="grey")+
  geom_point(aes(group=Subject,fill=VitD_pre_cat),shape=21,color="black",size=5)+
  stat_summary(fun=mean,geom="line",group=1,color="black",size=2)+
  labs(x="Day",fill="Vitamin D level (Baseline)",y="Shannon Diversity")+
  stat_compare_means(comparisons=compare2,method="t.test",paired=TRUE,size=8)+
  guides(fill="none")+
  facet_grid(~Cohort)+
  theme(
    panel.background = element_rect(fill="white"),
    text = element_text(size=30,color="black"),
    panel.border = element_rect(size=2,fill="NA"),
    )
```
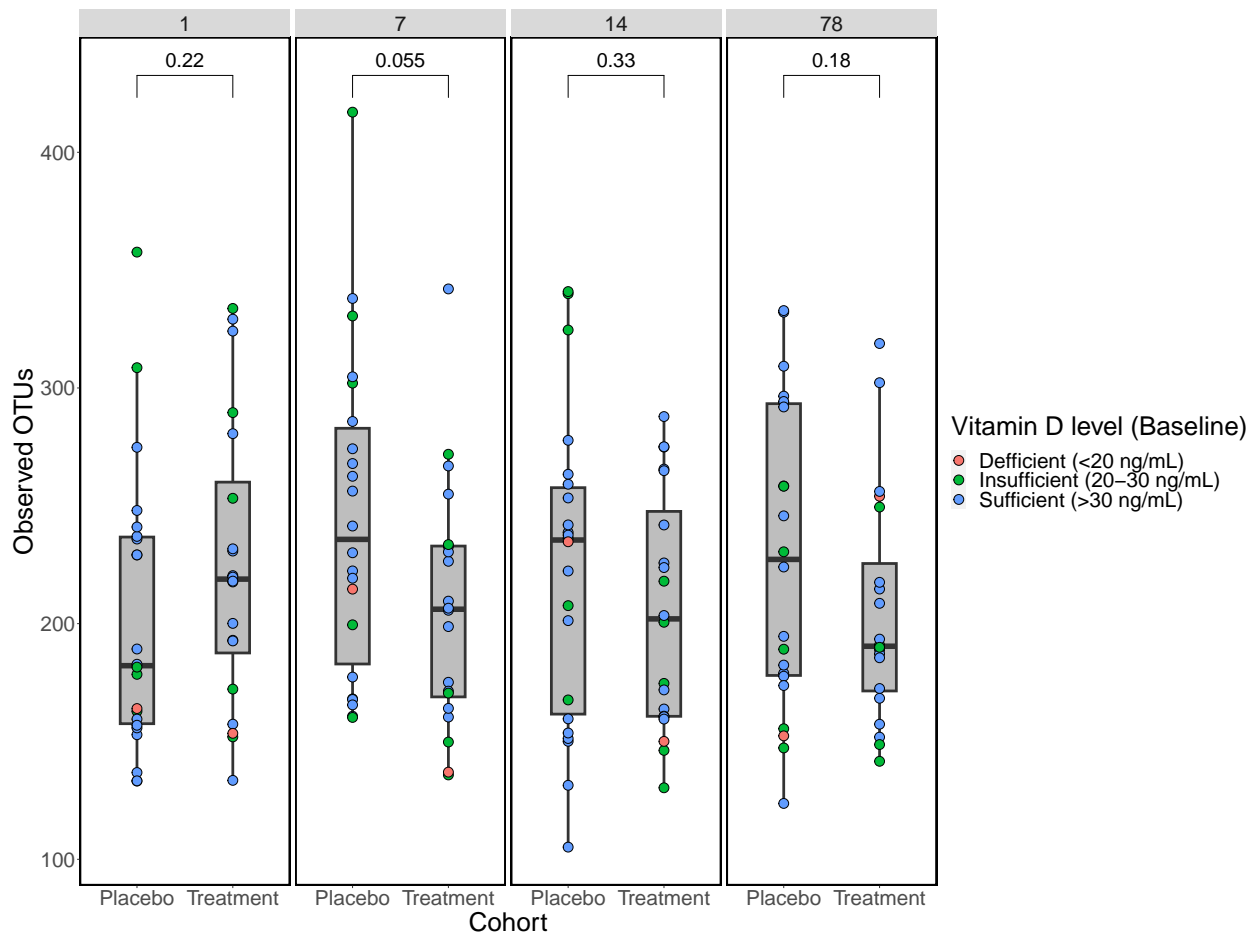
[1] FALSE

```
#Faith PD vs Day faceted by Cohort
ggplot(data=AlphaVDMT,aes(x=factor(Day),y=faith))+
  geom_boxplot(width=0.35,size=1.5,fill="grey")+
  geom_point(aes(group=Subject,fill=VitD_pre_cat),shape=21,color="black",size=5)+
  stat_summary(fun=mean,geom="line",group=1,color="black",size=2)+
  labs(x="Day",fill="Vitamin D level (Baseline)",y="Faith Phylogentic Diversity")+
  stat_compare_means(comparisons=compare2,method="t.test",paired=TRUE,size=8)+
  guides(fill="none")+
  facet_grid(~Cohort)+
  theme(
    panel.background = element_rect(fill="white"),
    text = element_text(size=30,color="black"),
    panel.border = element_rect(size=2,fill="NA"),
  )
```
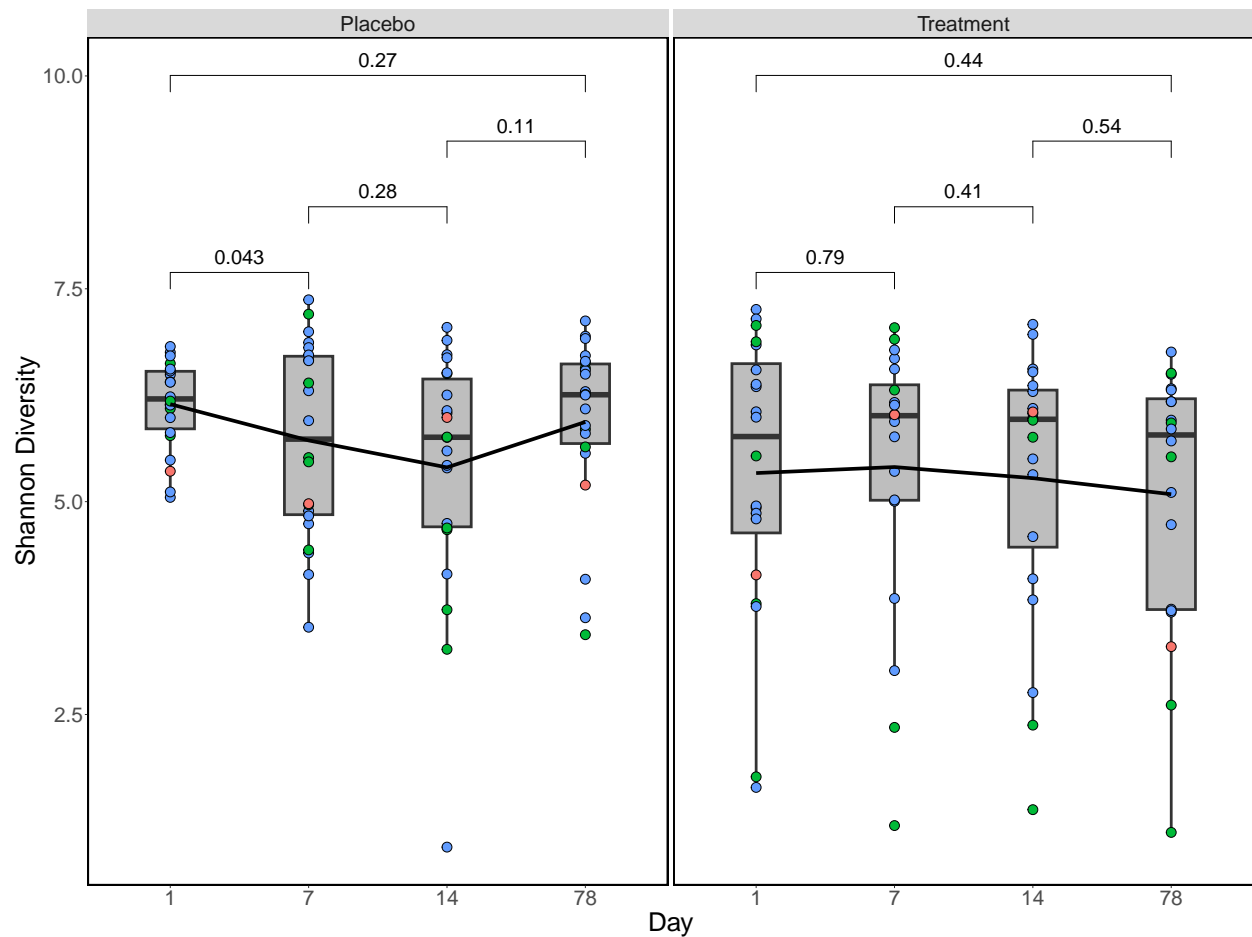
[1] FALSE
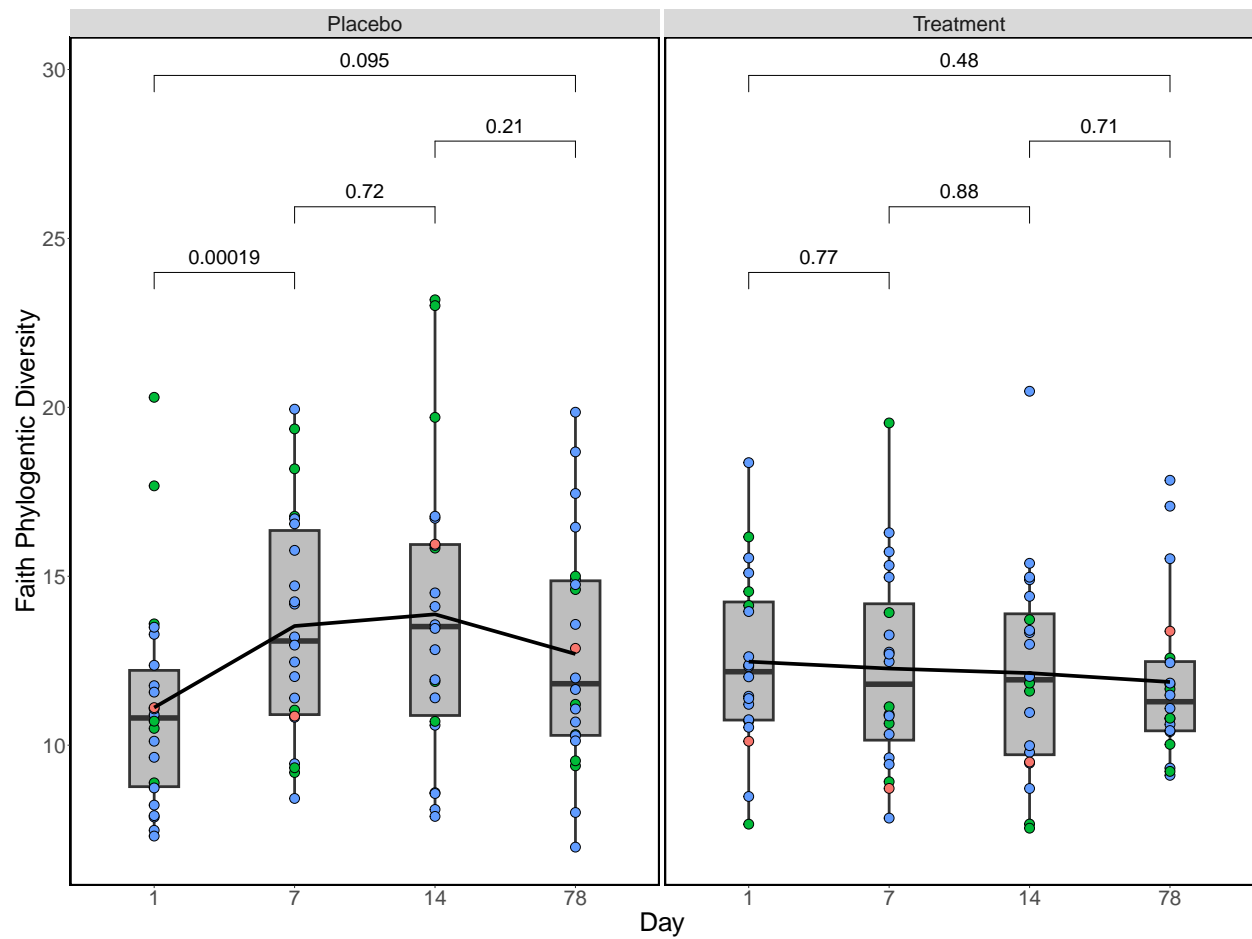
```
#Observed OTUs vs Day faceted by Cohort
ggplot(data=AlphaVDMT,aes(x=factor(Day),y=OTU))+
  geom_boxplot(width=0.35,size=1.5,fill="grey")+
  geom_point(aes(group=Subject,fill=VitD_pre_cat),shape=21,color="black",size=5)+
  stat_summary(fun=mean,geom="line",group=1,color="black",size=2)+
  labs(x="Day",fill="Vitamin D level (Baseline)",y="Observed OTUs")+
  stat_compare_means(comparisons=compare2,method="t.test",paired=TRUE,size=8)+
  guides(fill="none")+
  facet_grid(~Cohort)+
  theme(
    panel.background = element_rect(fill="white"),
    text = element_text(size=30,color="black"),
    panel.border = element_rect(size=2,fill="NA"),
  )
```

[1] FALSE
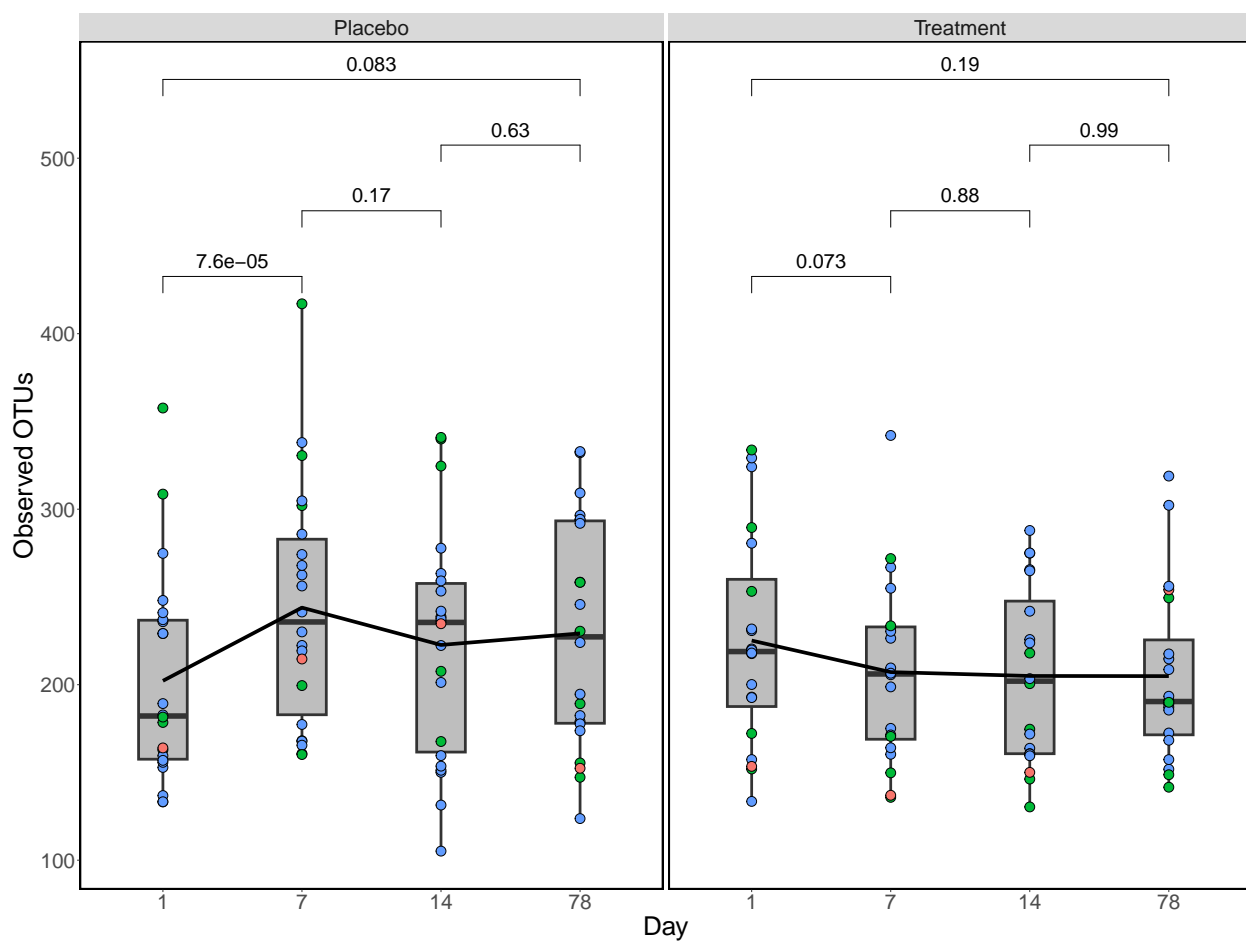
```
#Alternative figures with Alpha diversity indices calculated straight from QIIME2 derived feature table

#Importing rooted phylogenetic tree created by QIIME2
tree<-ape::as.phylo(ape::read.tree("tree.nwk"))

#importing the feature table created by DADA2 in QIIME2
library(biomformat)
reads<-read_biom("feature-table.biom")
view(t(as.matrix(biom_data(reads))))

#Converting the feature table into a dataframe with samples as rows
OTUtable<-as.data.frame(t(as.matrix(biom_data(reads))))%>%
  rownames_to_column("Sample")%>%
  pivot_longer(-Sample)
head(OTUtable)
```

```
## # A tibble: 6 x 3
##   Sample    name                             value
##   <chr>     <chr>                            <dbl>
## 1 VDMT001-1 ea723c21e394dc811a79b4bb80bbda44  1063
## 2 VDMT001-1 26f49ebf316637bc067ab0edf346b36e  1310
## 3 VDMT001-1 14cb11641b9c5483ba2f98e38d6a951b     0
## 4 VDMT001-1 dbb30eb7756e8d9e6c742bd6eed556fa   811
```

```
## 5 VDMT001-1 fc868901fa4106f07c09f5779ae5757d    755
## 6 VDMT001-1 a4c5be8e22fc71dbee3442ffcebe9b97    904
```

```r
#Summarizing the number reads in each sample as a dataframe and finding the sample with the least numbe
OTUcount<-OTUtable%>%
  group_by(Sample)%>%
  summarise(n=sum(value))%>%
  arrange(n)
head(OTUcount)
```

```
## # A tibble: 6 x 2
##   Sample          n
##   <chr>       <dbl>
## 1 VDMT036-14 25371
## 2 VDMT037-1  26364
## 3 VDMT010-1  26455
## 4 VDMT011-14 26605
## 5 VDMT014-14 27146
## 6 VDMT011-15 28769
```

```r
#Converting OTUtable dataframe into a count matrix as inputs for the distance matrix calculations
OTUmatrix<-OTUtable%>%
  pivot_wider(names_from = "name",values_from="value")%>%
  column_to_rownames("Sample")%>%
  as.matrix

library(picante)
#calculating the alpha diversity indices using the OTUs derived from the feature table and the phylogen
richness<-specnumber(OTUmatrix)%>% #Calculating richness/Observed OTU
  as.data.frame()%>%
  rownames_to_column("Sample")%>%
  rename_with(~"OTU",2)
head(richness)
```

```
##        Sample OTU
## 1  VDMT001-1 329
## 2 VDMT001-14 272
## 3 VDMT001-15 216
## 4  VDMT001-7 270
## 5  VDMT002-1 311
## 6 VDMT002-14 209
```

```r
shannon<-diversity(OTUmatrix)%>% #Calculating Shannon Diversity
  as.data.frame()%>%
  rownames_to_column("Sample")%>%
  rename_with(~"shannon",2)
head(shannon)
```

```
##        Sample  shannon
## 1  VDMT001-1 5.041706
## 2 VDMT001-14 2.848287
## 3 VDMT001-15 4.508551
```

```
## 4  VDMT001-7 4.181596
## 5  VDMT002-1 2.648678
## 6 VDMT002-14 1.653942
```

```r
faith<-pd(OTUmatrix,tree)%>% #Calculating Faith phylogenetic Diversity
  rownames_to_column("Sample")%>%
  rename_with(~"faithPD",2)
head(faith)
```

```
##        Sample  faithPD  SR
## 1  VDMT001-1 15.35811 329
## 2 VDMT001-14 15.48439 272
## 3 VDMT001-15 11.93431 216
## 4  VDMT001-7 16.27767 270
## 5  VDMT002-1 15.46386 311
## 6 VDMT002-14 12.12556 209
```

```r
#joining all the indices together and merging with the metadata and the Vitamin D category variable
AlphaVDMT2<-inner_join(richness,shannon,by="Sample")%>%
  inner_join(faith,by="Sample")%>%
  inner_join(OTUcount,by="Sample")%>%
  select(-SR)%>%
  pivot_longer(-c(Sample,n))%>%
  rename_with(~"Alpha",3)%>%
  inner_join(sample_metadata,by="Sample")%>%
  inner_join(VitaminDChange[,c("Subject","VitD_pre_cat")],by="Subject")
head(AlphaVDMT2)
```

```
## # A tibble: 6 x 33
##   Sample         n Alpha  value Sampl~1   Day Subject Cohort Cohor~2   AGE SEX
##   <chr>      <dbl> <chr>  <dbl> <chr>   <dbl> <chr>   <chr>  <chr>   <dbl> <chr>
## 1 VDMT001-1  58173 OTU    329   VDMT00~     1 VDMT001 Treat~ Treatm~    19 F
## 2 VDMT001-1  58173 shan~   5.04 VDMT00~     1 VDMT001 Treat~ Treatm~    19 F
## 3 VDMT001-1  58173 fait~  15.4  VDMT00~     1 VDMT001 Treat~ Treatm~    19 F
## 4 VDMT001-14 73429 OTU    272   VDMT00~    14 VDMT001 Treat~ Treatm~    19 F
## 5 VDMT001-14 73429 shan~   2.85 VDMT00~    14 VDMT001 Treat~ Treatm~    19 F
## 6 VDMT001-14 73429 fait~  15.5  VDMT00~    14 VDMT001 Treat~ Treatm~    19 F
## # ... with 22 more variables: RACE <dbl>, `Vit D_pre` <dbl>,
## #   `Vit D_post` <dbl>, Height <dbl>, Wt_pre <dbl>, Wt_post <dbl>,
## #   BMI_pre <dbl>, BMI_post <dbl>, LBM_pre <dbl>, LBM_post <dbl>,
## #   PBF_pre <dbl>, PBF_post <dbl>, Sleep_week <dbl>, Sleep_weekend <dbl>,
## #   Vig_Ex_daily <dbl>, SIT_hr <dbl>, `Total HEI-2015 Score` <dbl>,
## #   Vit_D_pre_Cat <chr>, Vit_D_post_Cat <chr>, `Vit_Change%` <dbl>,
## #   `Vit_Change%_Cat` <chr>, VitD_pre_cat <chr>, and abbreviated variable ...
```

```r
alphaList<-c("Shannon Diversity","Faith Phylogenetic\nDiversity","Observed OTUs")
names(alphaList)<-c("shannon","faithPD","OTU")

#Alpha indices vs Day faceted by Cohort
ggplot(AlphaVDMT2,aes(x=factor(Day),y=value))+
  geom_boxplot(width=0.35,size=1.5,fill="grey")+
  geom_point(aes(group=Subject,fill=VitD_pre_cat),shape=21,color="black",size=5)+
```

```
  stat_compare_means(comparisons=compare2,method="wilcox.test",paired=TRUE,size=4,bracket.size = 1,step
  stat_summary(group=1,fun=mean,geom="line",color="black",size=2)+
  facet_grid(Alpha~Cohort,scales="free_y",labeller=labeller(Alpha=alphaList))+
  theme_classic()+
  labs(x="Day")+
  ylab(NULL)+
  guides(fill="none")+
  theme(
    text=element_text(size=25,face="bold"),
    panel.border = element_rect(linewidth=1,color="black",fill=NA),
    strip.background = element_rect(fill="grey"),
    strip.background.y=element_blank(),
    strip.placement = "outside")
```

[1] FALSE


```
#Alpha indices vs Cohort faceted by Day
ggplot(AlphaVDMT2,aes(x=Cohort,y=value))+
  geom_boxplot(width=0.35,size=1.5,fill="grey")+
  geom_point(aes(group=Subject,fill=VitD_pre_cat),shape=21,color="black",size=5)+
  stat_compare_means(comparisons=compare,method="wilcox.test",size=4,bracket.size = 1,step.increase = 0
  facet_grid(Alpha~factor(Day),scales="free_y",labeller=labeller(Alpha=alphaList))+
```
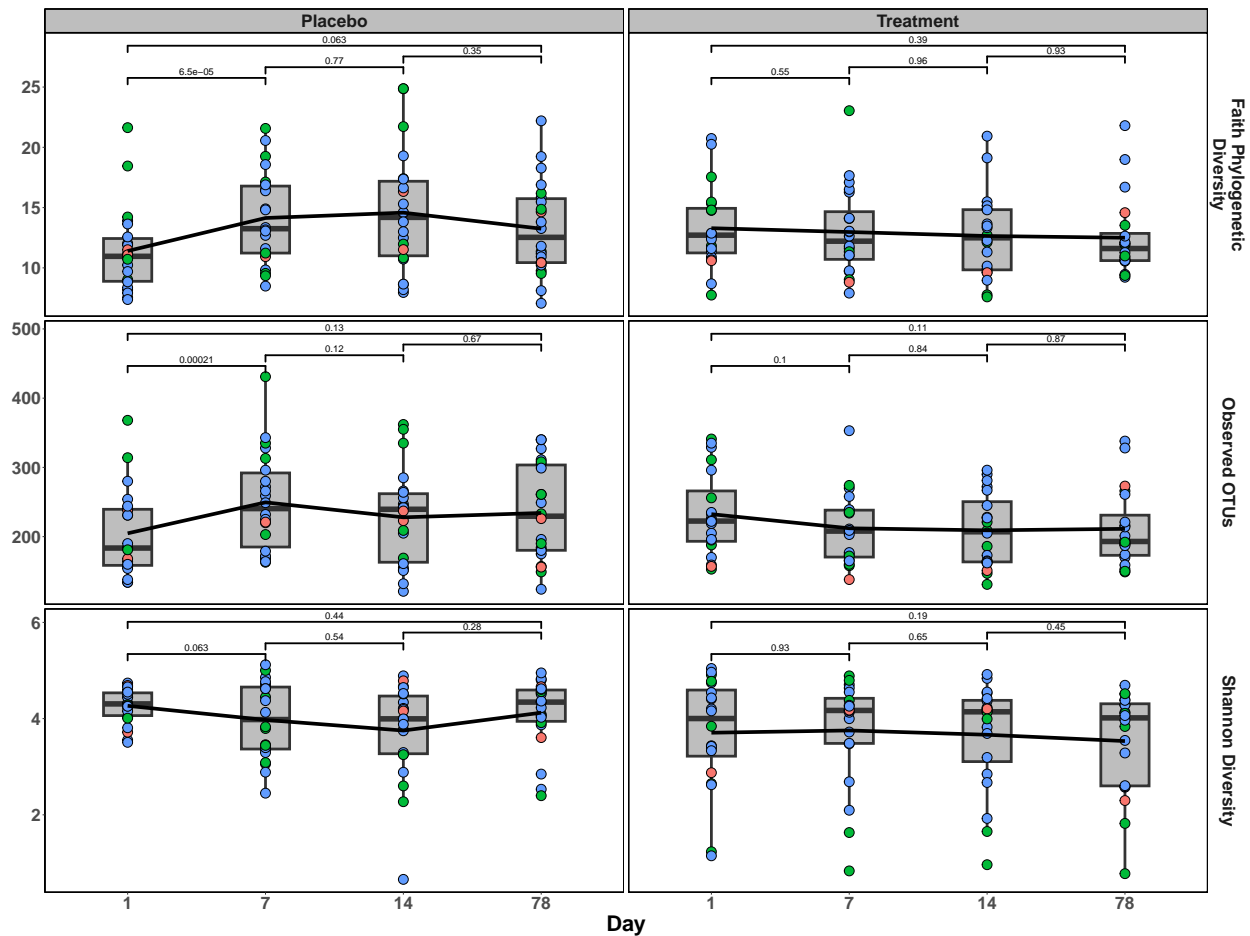
```
  theme_classic()+
  labs(x="Day")+
  ylab(NULL)+
  theme(
    text=element_text(size=25,face="bold"),
    panel.border = element_rect(linewidth=1,color="black",fill=NA),
    strip.background = element_rect(fill="grey"),
    strip.background.y=element_blank(),
    strip.placement = "outside")
```

[1] FALSE
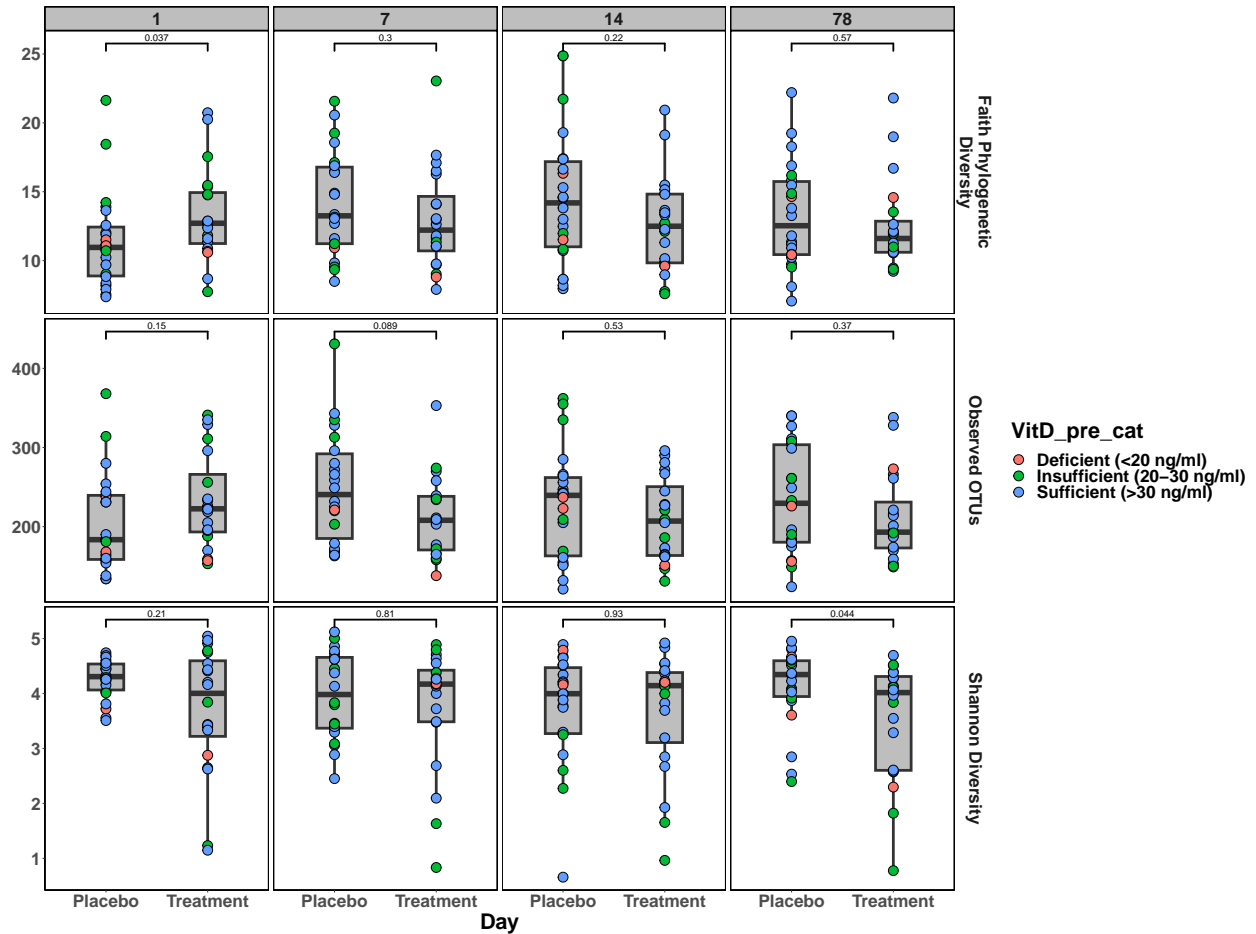

```
#Figure 4
library(vegan)
#Calculating Aitchison's Distance from the OTUMatrix and making a dataframe containing the Day and Coho
aitch_dist<-vegdist(OTUmatrix,method="aitchison",pseudocount=1)%>%
  as.matrix()
head(aitch_dist)
```

```
##             VDMT001-1 VDMT001-14 VDMT001-15 VDMT001-7 VDMT002-1 VDMT002-14
## VDMT001-1    0.00000    53.44140    64.98187   55.57226  72.36703   63.22714
## VDMT001-14  53.44140     0.00000    49.30497   55.86624  62.23656   50.76014
```

14

```
## VDMT001-15   64.98187     49.30497      0.00000    71.69744      49.81025      44.57682
## VDMT001-7    55.57226     55.86624     71.69744     0.00000      82.77886      70.36695
## VDMT002-1    72.36703     62.23656     49.81025    82.77886       0.00000      55.70898
## VDMT002-14   63.22714     50.76014     44.57682    70.36695      55.70898       0.00000
##            VDMT002-15 VDMT002-7 VDMT003-1 VDMT003-14 VDMT003-15 VDMT003-7
## VDMT001-1    71.64264     65.23350     71.04660     64.62125      64.82486     92.48807
## VDMT001-14   62.86742     50.88778     59.64580     52.40588      52.50022     83.93266
## VDMT001-15   53.70056     43.86062     53.15558     44.83659      45.15743     79.91686
## VDMT001-7    79.58116     68.54502     71.25063     68.35783      67.14247     86.07685
## VDMT002-1    50.00346     58.76545     66.02136     55.10366      57.27083     89.48802
## VDMT002-14   54.41395     40.55185     52.40850     41.45427      42.63877     75.47700
##            VDMT004-1 VDMT004-14 VDMT004-15 VDMT004-7 VDMT005-1 VDMT005-14
## VDMT001-1    70.29881     64.18932     67.02360     69.76877      68.95788     70.58240
## VDMT001-14   63.05892     56.93227     56.24870     63.61394      57.97478     62.95514
## VDMT001-15   59.89675     57.15233     45.94497     62.53248      45.76293     54.87428
## VDMT001-7    76.27230     70.88163     69.36052     70.63132      76.80300     76.83832
## VDMT002-1    64.72547     71.72821     67.28714     70.53597      53.18941     62.72026
## VDMT002-14   61.73649     57.06228     54.05021     62.12500      50.77385     52.74627
##            VDMT005-15 VDMT005-7 VDMT006-1 VDMT006-14 VDMT006-15 VDMT006-7
## VDMT001-1    70.20933     71.93835     75.75978     72.08185      71.52647     72.05023
## VDMT001-14   58.22564     59.14706     67.04156     61.83132      62.76443     62.33301
## VDMT001-15   46.66390     47.29433     63.49566     54.82858      56.25166     56.99881
## VDMT001-7    75.81205     77.23733     81.98738     80.16833      80.52208     82.61876
## VDMT002-1    55.80673     57.49010     69.69298     62.56306      62.48282     63.03522
## VDMT002-14   49.00647     52.02687     56.79622     51.70387      53.91423     53.13879
##            VDMT007-1 VDMT007-14 VDMT007-15 VDMT007-7 VDMT008-1 VDMT008-14
## VDMT001-1    61.10747     68.15722     68.98143     67.09802      70.55630     71.45628
## VDMT001-14   56.45082     55.30508     55.48345     62.30608      59.29178     63.03636
## VDMT001-15   58.28049     46.36959     40.76029     62.28324      51.47331     56.60867
## VDMT001-7    68.03024     74.37611     74.09667     73.91529      79.33818     79.88068
## VDMT002-1    66.73642     60.30245     59.64712     69.03148      52.69715     61.56373
## VDMT002-14   60.06466     46.98049     47.96132     63.11764      54.89595     52.79712
##            VDMT008-15 VDMT008-7 VDMT009-1 VDMT009-14 VDMT009-15 VDMT009-7
## VDMT001-1    75.68556     72.24780     70.27554     68.94302      73.08818     70.92756
## VDMT001-14   68.27084     63.49208     61.77769     61.36369      62.24852     61.64055
## VDMT001-15   61.52239     57.37686     57.19124     54.21180      55.51108     55.02127
## VDMT001-7    81.73835     80.05750     78.35633     78.96951      79.32995     80.48585
## VDMT002-1    68.04246     62.94807     62.71060     62.59652      65.00187     63.65379
## VDMT002-14   57.64113     54.36705     51.26773     51.97269      54.60765     51.48356
##            VDMT010-1 VDMT010-14 VDMT010-15 VDMT010-7 VDMT011-1 VDMT011-14
## VDMT001-1    66.64478     69.77209     69.26791     71.06472      68.30534     72.60384
## VDMT001-14   54.71456     56.69964     57.80240     59.93206      55.91129     61.72740
## VDMT001-15   45.49928     49.70182     50.09004     51.81417      46.23853     52.57883
## VDMT001-7    71.64130     73.98343     72.90470     75.81235      74.16850     76.22083
## VDMT002-1    59.45896     62.76692     63.74365     65.12144      54.14737     52.59338
## VDMT002-14   43.67090     46.06341     47.25288     48.83142      50.88346     51.96525
##            VDMT011-15 VDMT011-7 VDMT012-1 VDMT012-14 VDMT012-15 VDMT012-7
## VDMT001-1    70.35370     67.18566     73.56260     72.83125      75.20739     72.37291
## VDMT001-14   58.49867     54.64865     64.77201     62.47782      62.33426     60.54639
## VDMT001-15   53.20504     44.36186     60.90099     55.67861      54.62937     53.95547
## VDMT001-7    70.79610     74.27885     78.64088     75.92159      71.91407     70.51183
## VDMT002-1    65.51744     52.70603     68.14752     65.71383      69.62827     69.83490
## VDMT002-14   51.55955     49.37076     55.10045     52.76410      53.14570     54.11432
##            VDMT013-1 VDMT013-14 VDMT013-15 VDMT013-7 VDMT014-1 VDMT014-14
```

```
## VDMT001-1    67.93493    66.77555    72.20355    70.31923    72.28024    72.31260
## VDMT001-14   57.09268    57.34197    62.48348    60.73454    60.35135    62.01244
## VDMT001-15   48.85291    53.48258    55.70960    53.02904    51.98753    57.33336
## VDMT001-7    73.66948    69.62431    75.71377    76.14432    74.21879    72.20426
## VDMT002-1    61.58712    68.18237    65.81515    67.54556    63.87922    69.02157
## VDMT002-14   48.81886    55.46803    55.01426    56.02776    52.78002    56.25833
##              VDMT014-15 VDMT014-7 VDMT015-1 VDMT015-14 VDMT015-15 VDMT015-7
## VDMT001-1    69.79025    65.63685    70.02278    68.56508    70.54311    67.87712
## VDMT001-14   58.59476    53.12194    59.88888    57.52279    59.73789    57.76443
## VDMT001-15   53.15860    48.48177    52.68607    51.39912    51.74184    51.97432
## VDMT001-7    70.75818    74.73929    74.69469    74.11913    75.19605    73.89509
## VDMT002-1    65.87638    54.81373    63.71823    63.03368    63.22642    62.87808
## VDMT002-14   51.94189    45.35431    50.39305    49.12432    49.05964    50.19028
##              VDMT016-1 VDMT016-14 VDMT016-15 VDMT016-7 VDMT017-1 VDMT017-14
## VDMT001-1    94.39278    72.81678    72.15385    73.56408    74.62694    68.27006
## VDMT001-14   85.58209    61.63542    59.31512    61.64228    64.55945    61.82468
## VDMT001-15   81.61765    50.30456    45.29165    50.25659    54.09584    61.42150
## VDMT001-7    87.78865    73.77762    72.79110    73.81096    80.07923    69.41248
## VDMT002-1    90.97272    66.69492    65.06653    68.34884    53.53806    71.61167
## VDMT002-14   77.47375    53.34317    51.22807    55.07422    55.33976    62.44965
##              VDMT017-15 VDMT017-7 VDMT018-1 VDMT018-14 VDMT018-15 VDMT018-7
## VDMT001-1    73.46786    70.28620    71.86772    70.49995    74.59949    75.69803
## VDMT001-14   63.64933    58.66549    62.53170    61.72534    65.58284    66.97782
## VDMT001-15   58.56836    51.01841    54.98166    53.93013    58.37693    59.90769
## VDMT001-7    72.78493    78.91383    76.99743    74.93720    79.13155    79.67086
## VDMT002-1    68.97710    59.76699    63.91575    65.82167    68.76929    70.46224
## VDMT002-14   56.29576    50.31353    52.83774    50.89845    56.96101    57.63465
##              VDMT019-1 VDMT019-14 VDMT019-15 VDMT019-7 VDMT020-1 VDMT020-14
## VDMT001-1    73.80605    72.83601    74.08024    71.42463    68.59039    67.83681
## VDMT001-14   64.38790    61.68097    62.89931    58.80176    55.22076    58.12988
## VDMT001-15   53.91927    53.40312    53.47294    51.66606    47.18304    52.94845
## VDMT001-7    76.12446    74.50190    76.12885    75.86006    71.82219    73.38550
## VDMT002-1    64.72570    65.04651    64.52207    65.08441    58.93526    56.49348
## VDMT002-14   53.67472    51.64611    52.93885    49.57750    45.58085    52.23328
##              VDMT020-15 VDMT020-7 VDMT022-1 VDMT022-14 VDMT022-15 VDMT022-7
## VDMT001-1    71.68371    70.42636    72.97919    70.09970    72.15039    72.20432
## VDMT001-14   61.60350    57.46898    62.54234    60.32976    62.20120    60.35116
## VDMT001-15   54.37147    46.53949    56.67448    52.17434    55.25336    52.92206
## VDMT001-7    74.36903    75.95772    71.80783    73.66314    74.72978    72.36148
## VDMT002-1    62.77848    55.62330    69.43743    65.38452    64.34910    67.12638
## VDMT002-14   53.29261    50.13721    55.63697    49.96061    57.01138    52.60302
##              VDMT023-1 VDMT023-14 VDMT023-15 VDMT023-7 VDMT024-1 VDMT024-14
## VDMT001-1    77.67480    74.45793    78.95918    77.73366    80.07725    87.93670
## VDMT001-14   67.90775    66.51645    69.61945    67.26131    71.81425    78.78910
## VDMT001-15   66.05295    62.35848    65.08565    61.18837    65.28046    74.85811
## VDMT001-7    72.73274    71.37779    74.98394    75.64643    86.55850    81.53359
## VDMT002-1    77.34631    74.64975    76.57072    73.41221    58.22224    84.48206
## VDMT002-14   65.94284    60.64275    63.99594    60.79444    65.80553    70.39945
##              VDMT024-15 VDMT024-7 VDMT025-1 VDMT025-14 VDMT025-15 VDMT025-7
## VDMT001-1    76.15775    77.73005    72.24393    67.16616    61.39137    71.79937
## VDMT001-14   67.30470    67.98237    65.54264    60.62893    62.27257    62.16091
## VDMT001-15   59.83827    57.19343    54.87508    53.93713    60.07938    52.80673
## VDMT001-7    80.38003    86.75938    80.01348    76.17256    77.46746    78.01210
## VDMT002-1    53.18566    51.01223    48.60265    61.39475    64.91329    51.28563
```

```
## VDMT002-14     60.85526   60.26503   56.43250   51.43901   58.79734   54.65677
##             VDMT026-1 VDMT026-14 VDMT026-15 VDMT026-7 VDMT027-1 VDMT027-14
## VDMT001-1     65.91776   68.57380   68.11390   71.85952   75.28604   70.70863
## VDMT001-14    62.86062   56.71865   58.72391   71.32316   65.77281   61.42770
## VDMT001-15    66.52672   52.87343   52.29790   74.01991   61.40996   54.97815
## VDMT001-7     73.21353   76.77695   77.14883   80.67596   78.59874   77.98733
## VDMT002-1     74.80058   57.81898   57.14363   80.42348   71.74974   62.44552
## VDMT002-14    66.00610   52.12999   50.39747   74.32216   61.34449   53.50141
##             VDMT027-15 VDMT027-7 VDMT028-1 VDMT028-14 VDMT028-15 VDMT028-7
## VDMT001-1     70.95172   72.84624   74.40653   70.53329   74.14950   65.70429
## VDMT001-14    62.34613   63.98115   64.93441   59.82407   64.71710   54.88920
## VDMT001-15    55.86090   59.82135   59.71447   53.25640   57.73956   44.84579
## VDMT001-7     79.38378   78.25907   82.41825   78.08238   82.83874   70.95324
## VDMT002-1     62.75242   69.04696   65.20562   61.19063   65.25078   56.98681
## VDMT002-14    55.87426   59.65708   58.33763   52.39451   59.00965   43.26575
##             VDMT029-1 VDMT029-14 VDMT029-15 VDMT029-7 VDMT030-1 VDMT030-14
## VDMT001-1     71.15222   67.88733   69.99626   72.70768   71.20091   67.65529
## VDMT001-14    61.26787   55.28685   59.68716   63.66980   62.37369   56.25930
## VDMT001-15    53.63939   47.48082   52.50329   55.66240   55.50926   50.13520
## VDMT001-7     73.96507   72.59172   73.52516   76.67424   77.93728   74.97580
## VDMT002-1     64.94994   56.80006   62.57943   65.07376   61.75862   58.86252
## VDMT002-14    51.10550   47.56146   49.92448   54.46362   53.85866   47.00651
##             VDMT030-15 VDMT030-7 VDMT031-1 VDMT031-14 VDMT031-15 VDMT031-7
## VDMT001-1     72.57916   76.82440   68.92053   70.99571   71.00668   65.53063
## VDMT001-14    62.91602   67.67447   57.22325   60.11907   60.82988   52.74288
## VDMT001-15    56.34814   62.92524   50.84251   55.91791   53.48570   46.32988
## VDMT001-7     78.73683   80.81777   71.73491   75.49330   81.20075   73.26783
## VDMT002-1     62.78941   69.15167   59.08476   62.32592   57.73597   53.40349
## VDMT002-14    56.59476   61.91062   47.50738   53.01890   52.17488   42.73868
##             VDMT032-1 VDMT032-14 VDMT032-15 VDMT032-7 VDMT033-1 VDMT033-14
## VDMT001-1     76.00478   76.17101   77.22000   78.21161   72.65494   70.20355
## VDMT001-14    65.22810   65.09176   66.36945   67.77206   62.86787   59.80601
## VDMT001-15    58.23362   57.94727   62.08440   61.58654   53.97770   57.77391
## VDMT001-7     73.31819   75.71808   74.88683   77.11032   75.06416   70.62296
## VDMT002-1     72.42400   70.93901   75.40078   75.71258   67.25306   68.02345
## VDMT002-14    57.63676   57.58937   60.86162   61.83906   53.83018   56.02240
##             VDMT033-15 VDMT033-7 VDMT034-1 VDMT034-14 VDMT034-15 VDMT034-7
## VDMT001-1     74.04850   75.60114   70.14290   71.11400   71.19539   72.14771
## VDMT001-14    61.71872   65.63545   61.85457   62.02926   62.45727   63.21214
## VDMT001-15    54.61034   56.58782   56.28855   55.70132   55.92510   57.90492
## VDMT001-7     75.34991   79.94689   73.66060   74.03815   77.93882   75.75017
## VDMT002-1     67.17764   67.25444   66.57190   65.27235   64.20816   66.19562
## VDMT002-14    53.12433   57.17723   52.88703   51.68735   53.22952   53.36735
##             VDMT035-1 VDMT035-14 VDMT035-15 VDMT035-7 VDMT036-1 VDMT036-14
## VDMT001-1     71.85031   69.93648   70.80661   68.20214   76.32250   75.38709
## VDMT001-14    61.18607   61.94592   61.14322   59.91321   66.24960   65.08952
## VDMT001-15    54.85572   57.26499   59.15779   57.84934   57.87035   57.26443
## VDMT001-7     74.86716   71.82701   71.54474   69.16287   77.12736   77.09434
## VDMT002-1     65.34611   68.48642   69.11483   70.37891   70.92402   68.85244
## VDMT002-14    53.50147   55.22135   56.13772   56.80585   56.70880   54.70124
##             VDMT036-15 VDMT036-7 VDMT037-1 VDMT037-14 VDMT037-15 VDMT037-7
## VDMT001-1     76.24920   77.52966   69.91980   68.38566   76.65580   69.90977
## VDMT001-14    67.16982   67.83347   60.08020   60.36135   68.34169   61.52428
## VDMT001-15    59.76485   60.53538   52.43500   56.00327   61.24155   56.29467
```

```
## VDMT001-7     78.90288   78.46222   71.18594   70.76115    85.39869   75.94879
## VDMT002-1     70.19898   71.54083   64.42143   66.85629    64.07651   62.60082
## VDMT002-14    57.79640   58.47906   50.02179   55.36604    60.19754   53.00459
##              VDMT038-1 VDMT038-14 VDMT038-15 VDMT038-7 VDMT040-1 VDMT040-14
## VDMT001-1     74.17259   68.26737   73.32625   79.01420   78.95231   75.10687
## VDMT001-14    66.13066   58.97363   64.60298   72.89480   71.25480   66.23771
## VDMT001-15    60.42568   55.22444   58.31822   69.98040   63.52142   58.08599
## VDMT001-7     78.86590   73.32254   78.61290   85.15417   83.81383   79.98607
## VDMT002-1     69.34775   64.09435   65.80370   76.39861   70.76758   66.37691
## VDMT002-14    56.80973   52.66690   55.25532   66.53686   62.73763   57.67502
##              VDMT040-15 VDMT040-7 VDMT041-1 VDMT041-14 VDMT041-15 VDMT041-7
## VDMT001-1     76.63164   75.35017   75.09659   73.47276   83.26888   79.20435
## VDMT001-14    67.58802   65.72411   66.35698   64.03902   76.44250   71.64742
## VDMT001-15    60.81896   59.25459   59.82214   55.39981   70.80143   66.91261
## VDMT001-7     81.95238   80.16718   79.45834   79.20111   89.47267   83.44241
## VDMT002-1     67.09245   67.20856   68.41766   62.78068   76.37502   72.86411
## VDMT002-14    59.21430   57.74011   59.13004   57.26936   71.62106   68.36932
##              VDMT042-1 VDMT042-14 VDMT042-15 VDMT042-7 VDMT043-1 VDMT043-14
## VDMT001-1     78.23779   71.93104   74.67489   79.30455   79.41975   78.54515
## VDMT001-14    68.15303   60.31561   67.98126   69.87917   71.71659   70.32443
## VDMT001-15    63.12968   53.86098   64.37815   65.64618   64.85513   63.88848
## VDMT001-7     75.61535   80.21961   77.34676   76.69665   77.96873   76.03958
## VDMT002-1     76.65476   58.13368   75.39253   79.98365   76.52294   76.64803
## VDMT002-14    61.99779   54.21439   63.63090   64.61148   64.74779   64.52696
##              VDMT043-15 VDMT043-7 VDMT044-1 VDMT044-14 VDMT044-15 VDMT044-7
## VDMT001-1     77.83355   79.10447   78.44589   82.60289   82.87745   85.70308
## VDMT001-14    70.39096   68.18048   69.88766   73.34565   73.59502   76.76296
## VDMT001-15    69.90671   61.66783   63.09065   66.12845   67.48285   69.77862
## VDMT001-7     70.97653   75.81353   80.16830   85.45986   87.77151   87.94563
## VDMT002-1     81.97091   73.15743   68.16434   70.75420   70.52598   73.16778
## VDMT002-14    69.75371   62.16099   64.03316   66.25538   67.58867   70.67659
```

```r
#Converting the matrix to data frame and pivoting it lengthwise
aitch_dist_df<-aitch_dist%>%
  as.data.frame()%>%
  rownames_to_column("Sample")%>%
  pivot_longer(-Sample)%>%
  filter(name<Sample)%>%
  inner_join(sample_metadata[,c("Sample","Day","Cohort")],by="Sample")%>% #joining Subject, Day and Coh
  merge(sample_metadata[,c("Sample","Day","Cohort")],by.x="name",by.y="Sample")

head(aitch_dist_df)
```

```
##        name        Sample    value Day.x  Cohort.x Day.y  Cohort.y
## 1 VDMT001-1 VDMT001-14 53.44140    14 Treatment     1 Treatment
## 2 VDMT001-1 VDMT001-15 64.98187    78 Treatment     1 Treatment
## 3 VDMT001-1  VDMT022-7 72.20432     7 Treatment     1 Treatment
## 4 VDMT001-1  VDMT001-7 55.57226     7 Treatment     1 Treatment
## 5 VDMT001-1 VDMT034-15 71.19539    78   Placebo     1 Treatment
## 6 VDMT001-1 VDMT043-14 78.54515    14 Treatment     1 Treatment
```

```r
Cohorts<-c("Within Placebo group","Within Treatment group")
names(Cohorts)<-c("Placebo","Treatment")
```

18

```
#Boxplot by filtering the data to keep only the distances between the same group in each Day, Figure 4A
aitch_dist_df%>%
  filter(Day.x==Day.y&Cohort.x==Cohort.y)%>%
  ggplot(aes(x=factor(Day.x),y=value))+
  geom_boxplot(width=0.35,size=1.5,fill="grey")+
  geom_jitter(width=0.05)+
  stat_summary(fun=mean,geom="line",group=1,color="black",size=2)+
  labs(x="Day",y="Aitchison's Distance")+
  ylim(20,130)+
  stat_compare_means(comparisons=compare2,method="wilcox.test",size=7.5)+
  facet_grid(~Cohort.x,labeller=labeller(Cohort.x=Cohorts))+
  theme(
    panel.background = element_rect(fill="white"),
    text = element_text(size=30,color="black"),
    panel.border = element_rect(size=2,fill="NA"),
  )
```

[1] FALSE
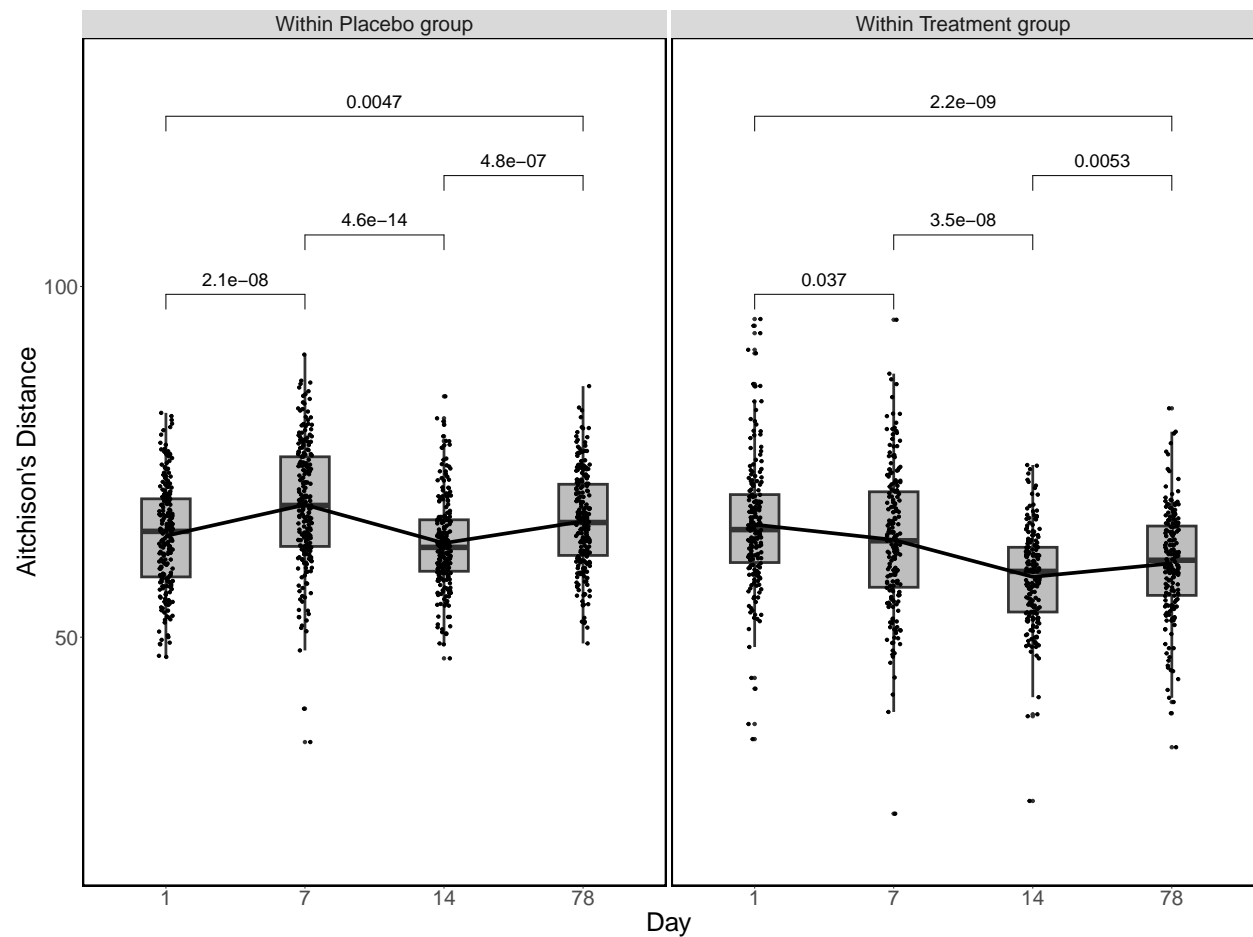

```
#Figure 4C
library(ggrepel)
```

```
#Performing PCOA based on the aitchison's distances using CMD scale
aitchpcoa<-aitch_dist%>%
  cmdscale(k=2,eig=TRUE,add=TRUE)

#Extracting the top 2 axis explaining most of the variations based on the eigen values
aitchpcoaaxispc<-round(100*aitchpcoa$eig/sum(aitchpcoa$eig),2)
view(aitchpcoaaxispc[1:10])

#Converting the PCOA into a dataframe with Axis coordinates for each sample and obtaining Subject, DAy
aitchpcoaplot<-as.data.frame(aitchpcoa$points)%>%
  rename_with(~c("Axis1","Axis2"),c(1,2))%>%
  rownames_to_column("Sample")%>%
  inner_join(sample_metadata[,c("Sample","Day","Cohort","Subject")],by="Sample")%>%
  mutate(index=substr(Sample,6,10))
head(aitchpcoaplot)
```

```
##       Sample      Axis1      Axis2 Day     Cohort Subject index
## 1  VDMT001-1  -9.177468 14.072476   1 Treatment VDMT001  01-1
## 2 VDMT001-14  -5.808833  8.909106  14 Treatment VDMT001 01-14
## 3 VDMT001-15  -5.808286  2.075926  78 Treatment VDMT001 01-15
## 4  VDMT001-7  11.645910 16.954911   7 Treatment VDMT001  01-7
## 5  VDMT002-1 -22.588411  7.263327   1 Treatment VDMT002  02-1
## 6 VDMT002-14  -4.885393 -3.012433  14 Treatment VDMT002 02-14
```

```
#Filtering the previous data frame to only contain samples from Days 7,14 and 78 for PERMANOVA analysis
aitch_permanova<-aitch_dist%>%
  as.data.frame()%>%
  rownames_to_column("Sample")%>%
  inner_join(aitchpcoaplot,by="Sample")%>%
  filter(Day>1)
head(aitch_permanova)
```

```
##        Sample VDMT001-1 VDMT001-14 VDMT001-15 VDMT001-7 VDMT002-1 VDMT002-14
## 1 VDMT001-14  53.44140    0.00000   49.30497  55.86624  62.23656   50.76014
## 2 VDMT001-15  64.98187   49.30497    0.00000  71.69744  49.81025   44.57682
## 3  VDMT001-7  55.57226   55.86624   71.69744   0.00000  82.77886   70.36695
## 4 VDMT002-14  63.22714   50.76014   44.57682  70.36695  55.70898    0.00000
## 5 VDMT002-15  71.64264   62.86742   53.70056  79.58116  50.00346   54.41395
## 6  VDMT002-7  65.23350   50.88778   43.86062  68.54502  58.76545   40.55185
##    VDMT002-15 VDMT002-7 VDMT003-1 VDMT003-14 VDMT003-15 VDMT003-7 VDMT004-1
## 1    62.86742  50.88778  59.64580   52.40588   52.50022  83.93266  63.05892
## 2    53.70056  43.86062  53.15558   44.83659   45.15743  79.91686  59.89675
## 3    79.58116  68.54502  71.25063   68.35783   67.14247  86.07685  76.27230
## 4    54.41395  40.55185  52.40850   41.45427   42.63877  75.47700  61.73649
## 5     0.00000  56.73932  61.78785   53.21141   54.55117  85.32046  63.85983
## 6    56.73932   0.00000  50.85272   42.34990   40.80507  71.99301  63.72877
##    VDMT004-14 VDMT004-15 VDMT004-7 VDMT005-1 VDMT005-14 VDMT005-15 VDMT005-7
## 1    56.93227   56.24870  63.61394  57.97478   62.95514   58.22564  59.14706
## 2    57.15233   45.94497  62.53248  45.76293   54.87428   46.66390  47.29433
## 3    70.88163   69.36052  70.63132  76.80300   76.83832   75.81205  77.23733
## 4    57.06228   54.05021  62.12500  50.77385   52.74627   49.00647  52.02687
## 5    69.05975   66.69610  68.05283  61.98481   60.27043   59.78628  62.54681
## 6    57.51781   50.30825  62.22572  50.17384   55.29332   48.54354  49.22674
```

```
##      VDMT006-1 VDMT006-14 VDMT006-15 VDMT006-7 VDMT007-1 VDMT007-14 VDMT007-15
## 1    67.04156   61.83132   62.76443  62.33301  56.45082   55.30508   55.48345
## 2    63.49566   54.82858   56.25166  56.99881  58.28049   46.36959   40.76029
## 3    81.98738   80.16833   80.52208  82.61876  68.03024   74.37611   74.09667
## 4    56.79622   51.70387   53.91423  53.13879  60.06466   46.98049   47.96132
## 5    64.86184   58.72526   60.48950  60.91409  65.39126   60.74391   59.73162
## 6    59.16867   52.99256   57.48935  58.55270  62.04504   46.21672   45.98810
##      VDMT007-7 VDMT008-1 VDMT008-14 VDMT008-15 VDMT008-7 VDMT009-1 VDMT009-14
## 1    62.30608   59.29178   63.03636  68.27084   63.49208  61.77769   61.36369
## 2    62.28324   51.47331   56.60867  61.52239   57.37686  57.19124   54.21180
## 3    73.91529   79.33818   79.88068  81.73835   80.05750  78.35633   78.96951
## 4    63.11764   54.89595   52.79712  57.64113   54.36705  51.26773   51.97269
## 5    67.89329   60.76402   62.43475  66.33346   64.16613  62.23387   62.93309
## 6    65.62631   58.29341   58.51995  60.04886   58.84792  54.01590   55.53143
##      VDMT009-15 VDMT009-7 VDMT010-1 VDMT010-14 VDMT010-15 VDMT010-7 VDMT011-1
## 1    62.24852   61.64055   54.71456  56.69964   57.80240  59.93206   55.91129
## 2    55.51108   55.02127   45.49928  49.70182   50.09004  51.81417   46.23853
## 3    79.32995   80.48585   71.64130  73.98343   72.90470  75.81235   74.16850
## 4    54.60765   51.48356   43.67090  46.06341   47.25288  48.83142   50.88346
## 5    66.94655   62.73628   57.67169  60.55219   60.86943  63.66938   62.49643
## 6    56.40143   55.20513   42.94584  44.35368   44.73593  49.00794   49.31901
##      VDMT011-14 VDMT011-15 VDMT011-7 VDMT012-1 VDMT012-14 VDMT012-15 VDMT012-7
## 1    61.72740   58.49867   54.64865  64.77201  62.47782   62.33426   60.54639
## 2    52.57883   53.20504   44.36186  60.90099  55.67861   54.62937   53.95547
## 3    76.22083   70.79610   74.27885  78.64088  75.92159   71.91407   70.51183
## 4    51.96525   51.55955   49.37076  55.10045  52.76410   53.14570   54.11432
## 5    44.97898   63.42072   59.86617  62.99273  61.62092   68.42770   68.02405
## 6    52.86132   47.81458   47.71225  55.43263  50.18700   50.97436   50.95514
##      VDMT013-1 VDMT013-14 VDMT013-15 VDMT013-7 VDMT014-1 VDMT014-14 VDMT014-15
## 1    57.09268   57.34197   62.48348  60.73454  60.35135   62.01244   58.59476
## 2    48.85291   53.48258   55.70960  53.02904  51.98753   57.33336   53.15860
## 3    73.66948   69.62431   75.71377  76.14432  74.21879   72.20426   70.75818
## 4    48.81886   55.46803   55.01426  56.02776  52.78002   56.25833   51.94189
## 5    61.81156   67.31584   64.29056  66.01845  60.57803   66.42331   64.33358
## 6    47.24509   55.25475   52.60729  53.97102  49.38659   52.11346   50.04527
##      VDMT014-7 VDMT015-1 VDMT015-14 VDMT015-15 VDMT015-7 VDMT016-1 VDMT016-14
## 1    53.12194   59.88888   57.52279  59.73789  57.76443   85.58209   61.63542
## 2    48.48177   52.68607   51.39912  51.74184  51.97432   81.61765   50.30456
## 3    74.73929   74.69469   74.11913  75.19605  73.89509   87.78865   73.77762
## 4    45.35431   50.39305   49.12432  49.05964  50.19028   77.47375   53.34317
## 5    56.44705   61.62346   62.61674  61.86250  58.76734   86.96263   65.85298
## 6    48.50754   50.68731   49.62300  49.84242  52.11325   73.99604   50.86457
##      VDMT016-15 VDMT016-7 VDMT017-1 VDMT017-14 VDMT017-15 VDMT017-7 VDMT018-1
## 1    59.31512   61.64228   64.55945  61.82468  63.64933   58.66549   62.53170
## 2    45.29165   50.25659   54.09584  61.42150  58.56836   51.01841   54.98166
## 3    72.79110   73.81096   80.07923  69.41248  72.78493   78.91383   76.99743
## 4    51.22807   55.07422   55.33976  62.44965  56.29576   50.31353   52.83774
## 5    64.41601   68.20483   41.45143  69.67007  67.72107   63.85625   62.28038
## 6    47.09997   51.42521   56.06519  62.69244  55.66120   51.11046   53.10720
##      VDMT018-14 VDMT018-15 VDMT018-7 VDMT019-1 VDMT019-14 VDMT019-15 VDMT019-7
## 1    61.72534   65.58284   66.97782  64.38790  61.68097   62.89931   58.80176
## 2    53.93013   58.37693   59.90769  53.91927  53.40312   53.47294   51.66606
## 3    74.93720   79.13155   79.67086  76.12446  74.50190   76.12885   75.86006
## 4    50.89845   56.96101   57.63465  53.67472  51.64611   52.93885   49.57750
```

```
## 5     63.48031     67.61207   69.38239   64.01423   62.17771    63.11695   63.55255
## 6     49.98067     55.54953   56.78054   51.85470   50.38040    50.72934   48.09054
##     VDMT020-1 VDMT020-14 VDMT020-15 VDMT020-7 VDMT022-1 VDMT022-14 VDMT022-15
## 1   55.22076     58.12988   61.60350   57.46898   62.54234    60.32976   62.20120
## 2   47.18304     52.94845   54.37147   46.53949   56.67448    52.17434   55.25336
## 3   71.82219     73.38550   74.36903   75.95772   71.80783    73.66314   74.72978
## 4   45.58085     52.23328   53.29261   50.13721   55.63697    49.96061   57.01138
## 5   57.37710     50.94489   61.11051   61.70394   66.15073    63.18280   64.70113
## 6   44.47176     54.35972   53.48655   47.52646   52.42535    48.76119   56.69002
##    VDMT022-7 VDMT023-1 VDMT023-14 VDMT023-15 VDMT023-7 VDMT024-1 VDMT024-14
## 1  60.35116   67.90775   66.51645   69.61945   67.26131   71.81425   78.78910
## 2  52.92206   66.05295   62.35848   65.08565   61.18837   65.28046   74.85811
## 3  72.36148   72.73274   71.37779   74.98394   75.64643   86.55850   81.53359
## 4  52.60302   65.94284   60.64275   63.99594   60.79444   65.80553   70.39945
## 5  65.90079   74.76375   70.98498   71.45136   70.49189   45.45026   80.67658
## 6  49.62614   63.09997   59.40537   60.91070   58.62345   66.22173   67.16785
##    VDMT024-15 VDMT024-7 VDMT025-1 VDMT025-14 VDMT025-15 VDMT025-7 VDMT026-1
## 1  67.30470   67.98237   65.54264   60.62893   62.27257   62.16091   62.86062
## 2  59.83827   57.19343   54.87508   53.93713   60.07938   52.80673   66.52672
## 3  80.38003   86.75938   80.01348   76.17256   77.46746   78.01210   73.21353
## 4  60.85526   60.26503   56.43250   51.43901   58.79734   54.65677   66.00610
## 5  46.84090   55.51343   45.09782   59.55671   66.05748   50.52556   73.94756
## 6  61.77920   62.93319   56.78992   49.81640   58.97886   55.45555   69.10706
##    VDMT026-14 VDMT026-15 VDMT026-7 VDMT027-1 VDMT027-14 VDMT027-15 VDMT027-7
## 1  56.71865   58.72391   71.32316   65.77281   61.42770   62.34613   63.98115
## 2  52.87343   52.29790   74.01991   61.40996   54.97815   55.86090   59.82135
## 3  76.77695   77.14883   80.67596   78.59874   77.98733   79.38378   78.25907
## 4  52.12999   50.39747   74.32216   61.34449   53.50141   55.87426   59.65708
## 5  58.29100   58.75527   80.22168   70.27459   62.11532   63.81173   67.30897
## 6  57.71949   54.61797   76.59222   58.51405   55.31966   59.70852   57.35388
##    VDMT028-1 VDMT028-14 VDMT028-15 VDMT028-7 VDMT029-1 VDMT029-14 VDMT029-15
## 1  64.93441   59.82407   64.71710   54.88920   61.26787   55.28685   59.68716
## 2  59.71447   53.25640   57.73956   44.84579   53.63939   47.48082   52.50329
## 3  82.41825   78.08238   82.83874   70.95324   73.96507   72.59172   73.52516
## 4  58.33763   52.39451   59.00965   43.26575   51.10550   47.56146   49.92448
## 5  66.89681   61.85384   66.84823   55.79775   62.39943   57.30974   61.42947
## 6  59.57386   56.41552   59.84162   44.25831   51.97924   49.14082   48.92847
##    VDMT029-7 VDMT030-1 VDMT030-14 VDMT030-15 VDMT030-7 VDMT031-1 VDMT031-14
## 1  63.66980   62.37369   56.25930   62.91602   67.67447   57.22325   60.11907
## 2  55.66240   55.50926   50.13520   56.34814   62.92524   50.84251   55.91791
## 3  76.67424   77.93728   74.97580   78.73683   80.81777   71.73491   75.49330
## 4  54.46362   53.85866   47.00651   56.59476   61.91062   47.50738   53.01890
## 5  63.16643   60.30556   57.70531   61.64968   67.36788   56.05139   61.33356
## 6  52.75444   51.01355   48.23104   56.87392   61.65630   47.91465   53.29820
##    VDMT031-15 VDMT031-7 VDMT032-1 VDMT032-14 VDMT032-15 VDMT032-7 VDMT033-1
## 1  60.82988   52.74288   65.22810   65.09176   66.36945   67.77206   62.86787
## 2  53.48570   46.32988   58.23362   57.94727   62.08440   61.58654   53.97770
## 3  81.20075   73.26783   73.31819   75.71808   74.88683   77.11032   75.06416
## 4  52.17488   42.73868   57.63676   57.58937   60.86162   61.83906   53.83018
## 5  60.88767   55.30967   70.69907   69.41614   74.39068   74.47017   64.76799
## 6  57.71530   48.23587   53.66426   53.34617   57.69048   58.02379   50.59764
##    VDMT033-14 VDMT033-15 VDMT033-7 VDMT034-1 VDMT034-14 VDMT034-15 VDMT034-7
## 1  59.80601   61.71872   65.63545   61.85457   62.02926   62.45727   63.21214
## 2  57.77391   54.61034   56.58782   56.28855   55.70132   55.92510   57.90492
```

```
## 3    70.62296    75.34991    79.94689    73.66060    74.03815    77.93882    75.75017
## 4    56.02240    53.12433    57.17723    52.88703    51.68735    53.22952    53.36735
## 5    68.17181    66.09334    66.72238    65.25610    64.16128    63.82313    65.10094
## 6    54.56722    51.04695    56.16325    50.64533    50.73176    52.38956    52.80771
##    VDMT035-1 VDMT035-14 VDMT035-15 VDMT035-7 VDMT036-1 VDMT036-14 VDMT036-15
## 1   61.18607    61.94592    61.14322    59.91321    66.24960    65.08952    67.16982
## 2   54.85572    57.26499    59.15779    57.84934    57.87035    57.26443    59.76485
## 3   74.86716    71.82701    71.54474    69.16287    77.12736    77.09434    78.90288
## 4   53.50147    55.22135    56.13772    56.80585    56.70880    54.70124    57.79640
## 5   64.32371    65.42646    66.28083    68.80138    67.80905    65.92002    68.45796
## 6   52.95927    52.91448    56.38031    56.54906    54.94250    52.94846    56.79782
##    VDMT036-7 VDMT037-1 VDMT037-14 VDMT037-15 VDMT037-7 VDMT038-1 VDMT038-14
## 1   67.83347    60.08020    60.36135    68.34169    61.52428    66.13066    58.97363
## 2   60.53538    52.43500    56.00327    61.24155    56.29467    60.42568    55.22444
## 3   78.46222    71.18594    70.76115    85.39869    75.94879    78.86590    73.32254
## 4   58.47906    50.02179    55.36604    60.19754    53.00459    56.80973    52.66690
## 5   68.67151    63.79254    66.69675    64.75287    63.02811    65.92703    64.64887
## 6   57.05506    48.64898    55.59941    63.71485    52.65967    56.16093    52.64756
##    VDMT038-15 VDMT038-7 VDMT040-1 VDMT040-14 VDMT040-15 VDMT040-7 VDMT041-1
## 1    64.60298    72.89480    71.25480    66.23771    67.58802    65.72411    66.35698
## 2    58.31822    69.98040    63.52142    58.08599    60.81896    59.25459    59.82214
## 3    78.61290    85.15417    83.81383    79.98607    81.95238    80.16718    79.45834
## 4    55.25532    66.53686    62.73763    57.67502    59.21430    57.74011    59.13004
## 5    64.45033    74.85901    68.49209    63.10305    64.47395    65.88282    66.26561
## 6    55.27359    67.54419    61.69609    57.60419    58.74646    57.17364    58.94889
##    VDMT041-14 VDMT041-15 VDMT041-7 VDMT042-1 VDMT042-14 VDMT042-15 VDMT042-7
## 1    64.03902    76.44250    71.64742    68.15303    60.31561    67.98126    69.87917
## 2    55.39981    70.80143    66.91261    63.12968    53.86098    64.37815    65.64618
## 3    79.20111    89.47267    83.44241    75.61535    80.21961    77.34676    76.69665
## 4    57.26936    71.62106    68.36932    61.99779    54.21439    63.63090    64.61148
## 5    63.14823    76.78077    71.02488    74.36302    61.39524    74.88997    76.89981
## 6    57.27930    71.78801    69.03238    59.20186    58.29438    63.48613    63.37043
##    VDMT043-1 VDMT043-14 VDMT043-15 VDMT043-7 VDMT044-1 VDMT044-14 VDMT044-15
## 1   71.71659    70.32443    70.39096    68.18048    69.88766    73.34565    73.59502
## 2   64.85513    63.88848    69.90671    61.66783    63.09065    66.12845    67.48285
## 3   77.96873    76.03958    70.97653    75.81353    80.16830    85.45986    87.77151
## 4   64.74779    64.52696    69.75371    62.16099    64.03316    66.25538    67.58867
## 5   75.81491    75.85662    79.03709    72.52406    67.61850    69.83530    70.87394
## 6   62.52023    62.28142    67.98777    60.05456    63.74778    65.39005    66.87135
##    VDMT044-7      Axis1     Axis2 Day    Cohort Subject index
## 1   76.76296  -5.808833  8.909106  14 Treatment VDMT001 01-14
## 2   69.77862  -5.808286  2.075926  78 Treatment VDMT001 01-15
## 3   87.94563  11.645910 16.954911   7 Treatment VDMT001  01-7
## 4   70.67659  -4.885393 -3.012433  14 Treatment VDMT002 02-14
## 5   73.00442 -18.172379  7.494989  78 Treatment VDMT002 02-15
## 6   70.68407   3.266790 -4.261459   7 Treatment VDMT002  02-7
```

```r
#Converting the previously obtained dataframe into a distance objcet and performing PERMANOVA analysis
aitch_permanova_dist<-as.dist(select(aitch_permanova,all_of(aitch_permanova[["Sample"]])))
aitch_adonis<-adonis2(aitch_permanova_dist~Cohort,data=aitch_permanova,permutations=10000)

#Ploting the PCOA distribution of samples from Days 7,14 and 78 along with the PERMANOVA stats
aitchpcoaplot%>%
  filter(Day>1)%>%
```

```
ggplot(aes(x=Axis1,y=Axis2,fill=Cohort))+
geom_point(color="black",shape=21,stroke=1.5,alpha=0.75,size=7.5)+
scale_fill_manual(values=c("#416DD4",  "#FFBC33"))+
scale_color_manual(values=c("#416DD4",  "#FFBC33"))+
labs(x=paste("PCoA Axis 1 (",aitchpcoaaxispc[1],"%)",sep=""),
     y=paste("PCoA Axis 2 (",aitchpcoaaxispc[2],"%)",sep=""),
     title="PCoA based on Aitchison Distance of samples from\n Days 7, 14 and 84")+
geom_text_repel(aes(label=index),max.overlaps=200, hjust = "right",force=1,size=6)+
stat_ellipse(aes(color=Cohort),type="t",size=1,linetype=2)+
annotate(geom="text",x=20,y=25,label=paste("PERMANOVA\nF-statistics =",round(aitch_adonis["Cohort","F
theme(text = element_text(size = 25,face="bold"),
      panel.background = element_rect(fill="white",color="grey",linewidth=2),
      panel.border = element_rect(colour = "black", fill=NA, size=2),
      panel.grid.major= element_blank(),
      axis.text=element_text(size=25,color="black"),
      legend.key.size = unit(10, "mm"))
```


PCoA based on Aitchison Distance of samples from Days 7, 14 and 84
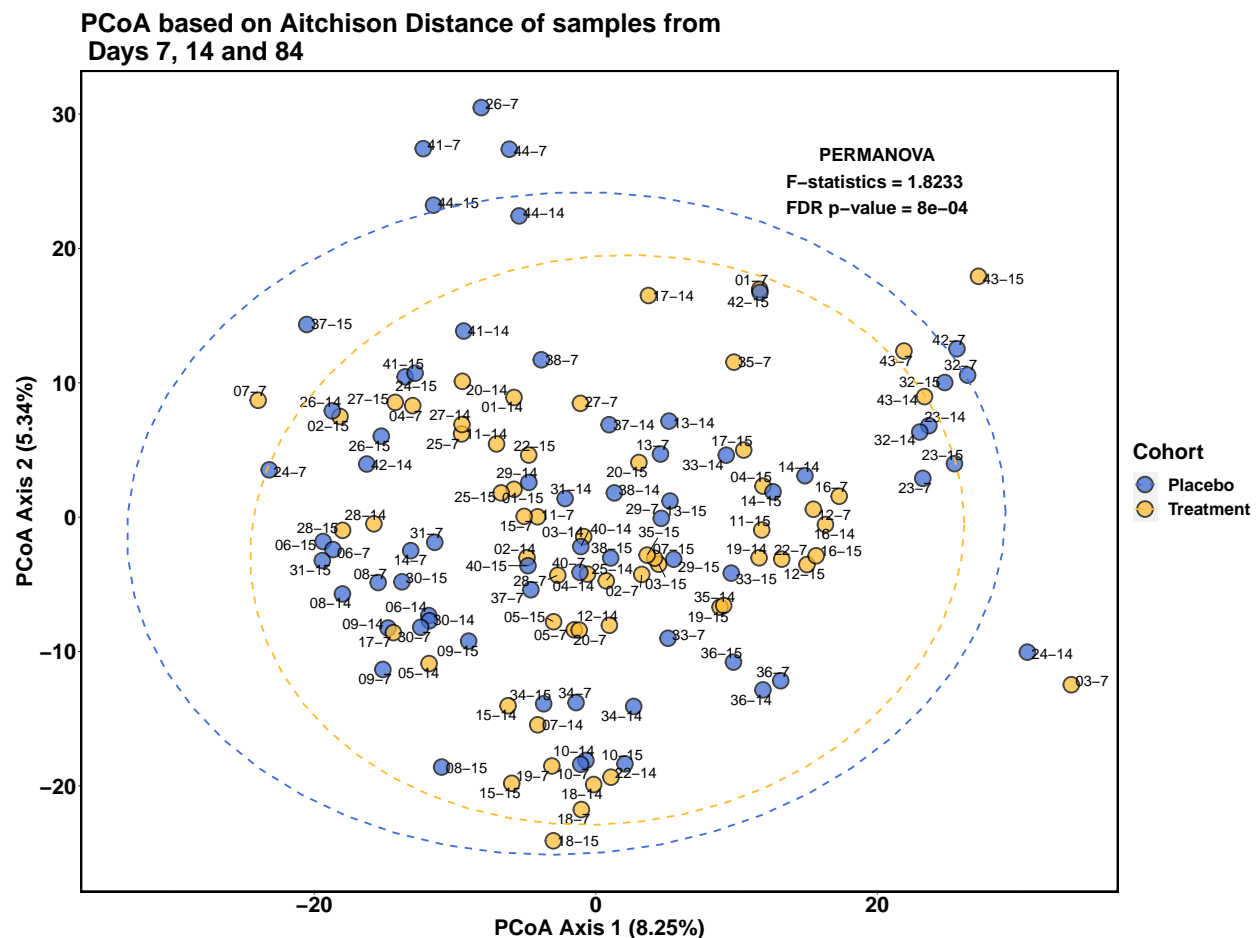
```
library(tidyr)
#Figure 4D
#Plotting the centroids (mean and Median) of the axis coordinates of the samples form Day 7,14 and 78
aitchpcoaplot%>%
  filter(Day>1)%>%
  group_by(Subject)%>%
```

```
summarize(median_Axis1=median(Axis1),
          median_Axis2=median(Axis2),
          mean_Axis1=mean(Axis1),
          mean_Axis2=mean(Axis2))%>%
inner_join(VitaminDChange[,c("Cohort","Subject")],by="Subject")%>%
pivot_longer(-c("Subject","Cohort"))%>%
separate(name, sep = "_", into = c("Centroid", "Axes"))%>%
pivot_wider(values_from="value",names_from="Axes")%>%
mutate(index=substr(Subject,6,7))%>%
ggplot(aes(x=Axis1,y=Axis2,fill=Centroid))+
geom_point(color="black",shape=21,stroke=1.5,alpha=0.75,size=7.5)+
labs(x=paste("PCoA Axis 1 (",aitchpcoaaxispc[1],"%)",sep=""),
     y=paste("PCoA Axis 2 (",aitchpcoaaxispc[2],"%)",sep=""),
     title="PCoA based on Aitchison Distance of samples from\n Days 7, 14 and 84")+
geom_text_repel(aes(label=index),max.overlaps=200, hjust = "right",force=1,size=6)+
stat_ellipse(aes(color=Centroid),type="t",size=1,linetype=2)+
facet_grid(~Cohort)+
theme(text = element_text(size = 25,face="bold"),
      panel.background = element_rect(fill="white",color="grey",linewidth=2),
      panel.border = element_rect(colour = "black", fill=NA, size=2),
      panel.grid.major= element_blank(),
      axis.text=element_text(size=25,color="black"),
      legend.key.size = unit(10, "mm"))
```


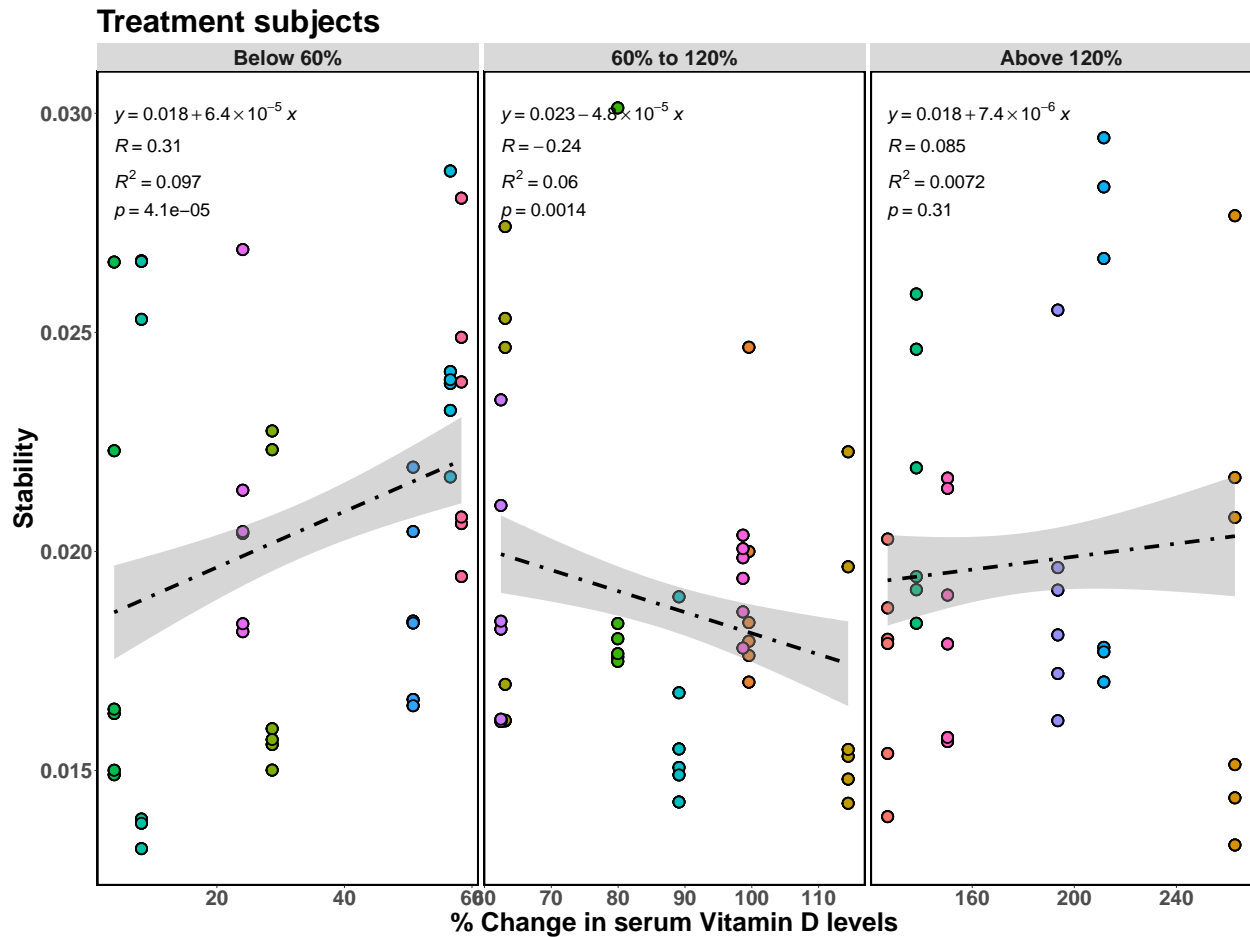**PCoA based on Aitchison Distance of samples from Days 7, 14 and 84**
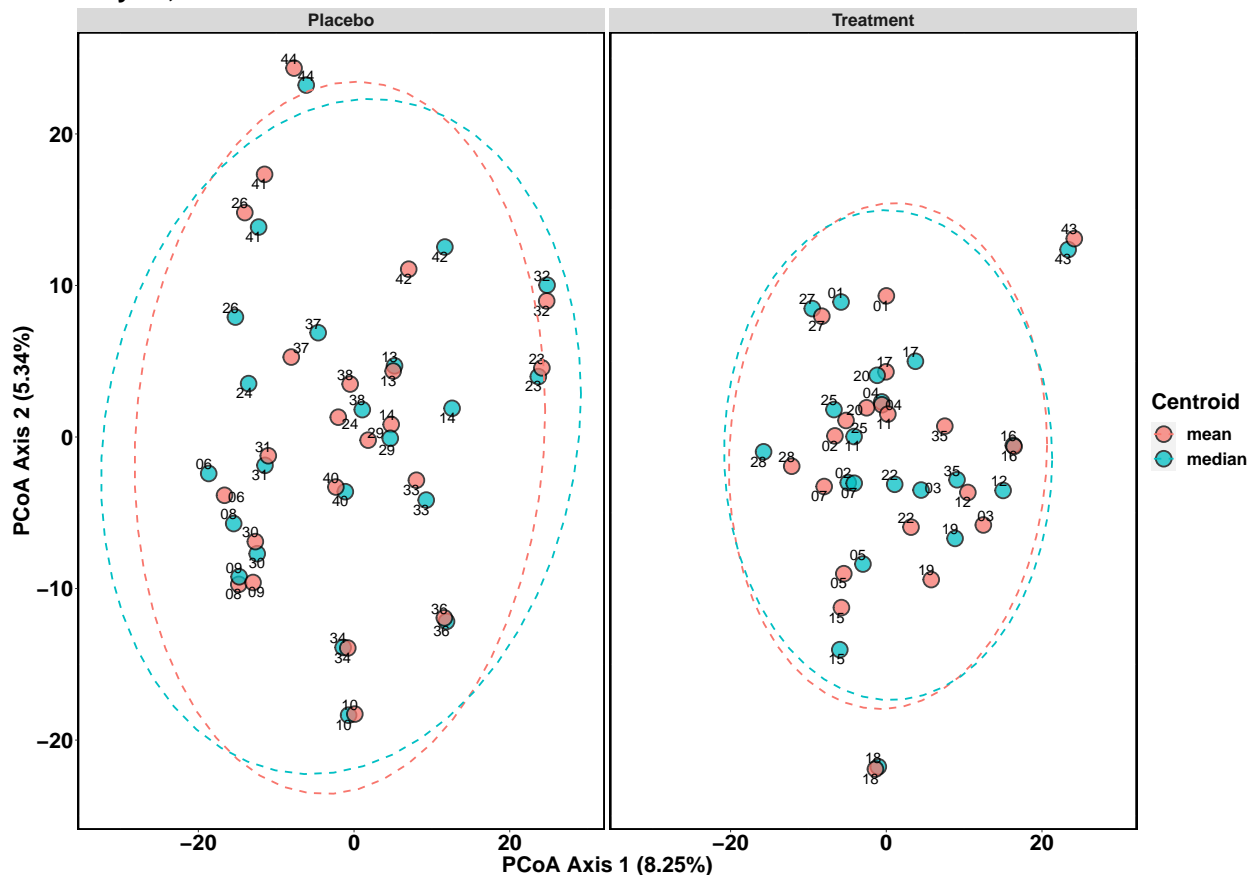
```r
#Figure 4E
#Filtering the dataframe with the aitchison's distance to only contain distances between the same subje
aitch_dist_df2<-aitch_dist%>%
  as.data.frame()%>%
  rownames_to_column("Sample")%>%
  pivot_longer(-Sample)%>%
  filter(name<Sample)%>%
  inner_join(sample_metadata[,c("Sample","Day","Subject")],by="Sample")%>%
  merge(sample_metadata[,c("Sample","Day","Subject")],by.x="name",by.y="Sample")%>%
  filter(Subject.x==Subject.y,Day.x!=Day.y)

#Plotting the stability (inverse of aitchison's distance) for every subject against the % change in the
VitaminDChange%>%
  inner_join(sample_metadata[,c("Sample","Subject")],by="Subject")%>%
  merge(aitch_dist_df2,by.x="Subject",by.y="Subject.x")%>%
  mutate(Stability=1/value)%>%
  filter(Cohort=="Treatment")%>%
  mutate(Change_class=ifelse(VitD_pcchange<60,"Below 60%",ifelse(VitD_pcchange<120,"60% to 120%","Above
  ggplot(aes(x=VitD_pcchange,y=Stability))+
  geom_point(aes(fill=Subject),shape=21,size=5,color="black",stroke=1.5)+
  geom_smooth(aes(group=Change_class),method="lm",formula=y ~ x,size=2,color="black",linetype=4)+
  labs(x="% Change in serum Vitamin D levels", y="Stability",title="Treatment subjects")+
  guides(fill="none",color="none")+
  facet_grid(~factor(Change_class,levels=c("Below 60%","60% to 120%","Above 120%")),scales="free_x")+
  stat_regline_equation(aes(group=Change_class),label.y=0.03,size=7.5,color="black")+
  stat_cor(aes(group=Change_class,label=..r.label..),label.y=0.02925,size=7.5,color="black")+
  stat_cor(aes(group=Change_class,label=..rr.label..),label.y=0.0285,size=7.5,color="black")+
  stat_cor(aes(group=Change_class,label=..p.label..),label.y=0.02775,size=7.5,color="black")+
  theme(
    panel.background = element_rect(fill="white"),
    text = element_text(size=30,color="black",face="bold"),
    panel.border = element_rect(size=2,fill="NA"),
  )
```
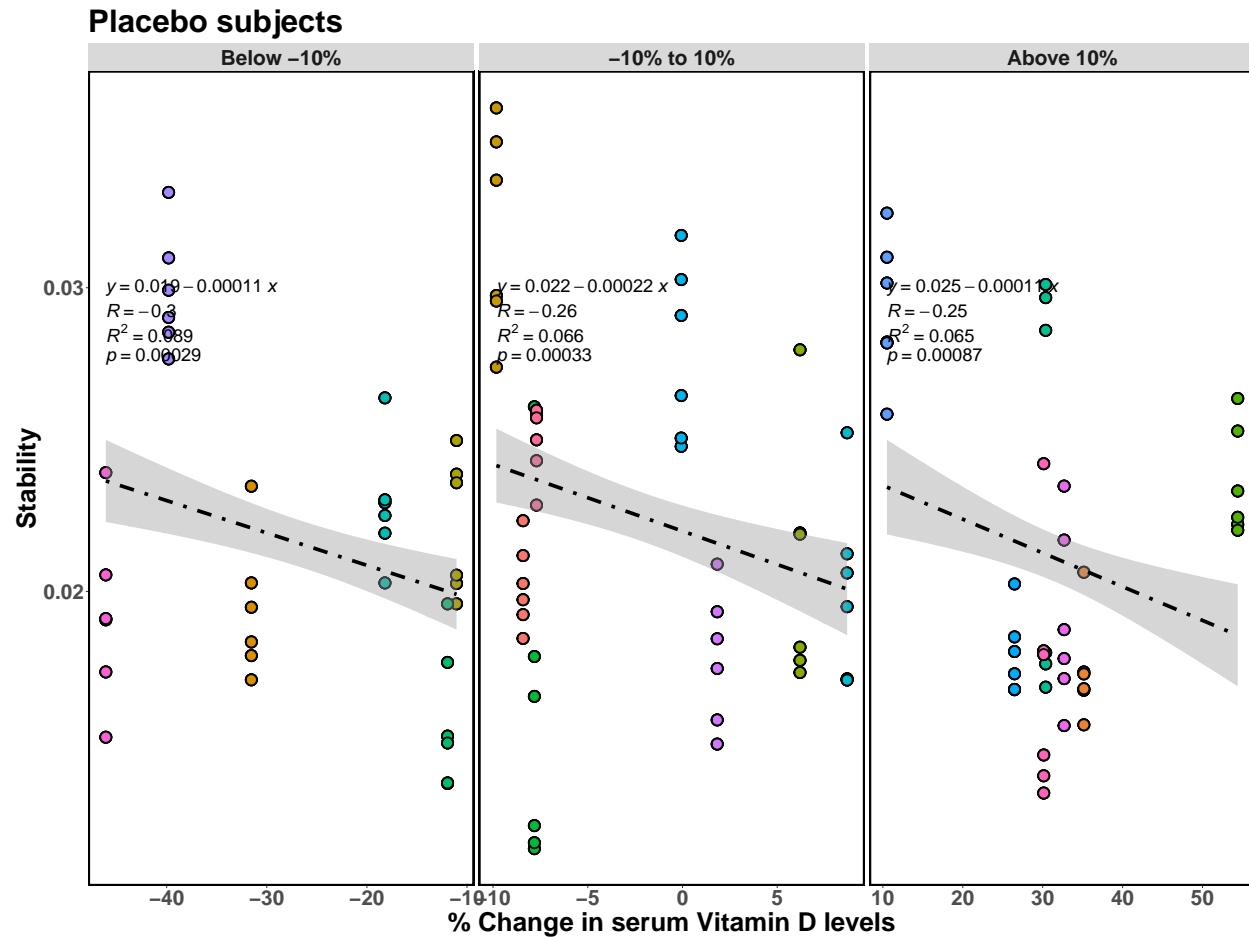
Treatment subjects

| Below 60% | 60% to 120% | Above 120% |

$y = 0.018 + 6.4 \times 10^{-5}\,x$
$R = 0.31$
$R^2 = 0.097$
$p = 4.1e{-}05$

$y = 0.023 - 4.8 \times 10^{-5}\,x$
$R = -0.24$
$R^2 = 0.06$
$p = 0.0014$

$y = 0.018 + 7.4 \times 10^{-6}\,x$
$R = 0.085$
$R^2 = 0.0072$
$p = 0.31$

Stability

% Change in serum Vitamin D levels

```
#Supplementary figure
VitaminDChange%>%
  inner_join(sample_metadata[,c("Sample","Subject")],by="Subject")%>%
  merge(aitch_dist_df2,by.x="Subject",by.y="Subject.x")%>%
  mutate(Stability=1/value)%>%
  filter(Cohort=="Placebo",Subject!="VDMT040")%>%
  mutate(Change_class=if_else(VitD_pcchange<(-10),"Below -10%",if_else(VitD_pcchange<(10),"-10% to 10%"
  ggplot(aes(x=VitD_pcchange,y=Stability))+
  geom_point(aes(fill=Subject),shape=21,size=5,color="black",stroke=1.5)+
  geom_smooth(aes(group=Change_class),method="lm",formula=y ~ x,size=2,color="black",linetype=4)+
  labs(x="% Change in serum Vitamin D levels", y="Stability",title="Placebo subjects")+
  guides(fill="none",color="none")+
  facet_grid(~factor(Change_class,levels=c("Below -10%","-10% to 10%","Above 10%")),scales="free_x")+
  stat_regline_equation(aes(group=Change_class),label.y=0.03,size=7.5,color="black")+
  stat_cor(aes(group=Change_class,label=..r.label..),label.y=0.02925,size=7.5,color="black")+
  stat_cor(aes(group=Change_class,label=..rr.label..),label.y=0.0285,size=7.5,color="black")+
  stat_cor(aes(group=Change_class,label=..p.label..),label.y=0.02775,size=7.5,color="black")+
  theme(
    panel.background = element_rect(fill="white"),
    text = element_text(size=30,color="black",face="bold"),
    panel.border = element_rect(size=2,fill="NA"),
  )
```

Placebo subjects

	Below −10%	−10% to 10%	Above 10%

$y = 0.019 - 0.00011\,x$
$R = -0.3$
$R^2 = 0.089$
$p = 0.00029$

$y = 0.022 - 0.00022\,x$
$R = -0.26$
$R^2 = 0.066$
$p = 0.00033$

$y = 0.025 - 0.00011\,x$
$R = -0.25$
$R^2 = 0.065$
$p = 0.00087$

Stability (y-axis)

% Change in serum Vitamin D levels (x-axis)

```
#Figure 5 A-C
library(pheatmap)
#Getting the taxa with significantly different abundabce compared to Day 1 for the Placebo Group and th
PlaceboChangeOnlySigs <- read_excel("Top100taxaVDMT.xlsx",
                         sheet = "PlaceboTaxaChangeOnlySigs")
PlaceboChangeOnlySigspval <- read_excel("Top100taxaVDMT.xlsx",
                      sheet = "PlaceboTaxaChangeOnlySigspval2")
PlaceboChangeOnlySigs2<-as.matrix(PlaceboChangeOnlySigs[,c(2:30)])
rownames(PlaceboChangeOnlySigs2)<-PlaceboChangeOnlySigs$Day
PlaceboChangeOnlySigspval2<-PlaceboChangeOnlySigspval[,c(2:30)]
PlaceboChangeOnlySigspval2<-replace(PlaceboChangeOnlySigspval2,is.na(PlaceboChangeOnlySigspval2)," ")
rownames(PlaceboChangeOnlySigspval2)<-PlaceboChangeOnlySigspval$Day

#Plotting a heatmap with differentially abundant taxa
pheatmap(PlaceboChangeOnlySigs2,display_num=PlaceboChangeOnlySigspval2,
      color=colorRampPalette(c("#416DD4", "#565E6A", "#FFBC33"))(100),scale="none",
      cluster_rows = FALSE,show_rownames=TRUE,
      border_color=NA,
      fontsize_number=25,
      fontsize_row=25,
      fontsize_col= 25,
      number_color="white",
      cellheight=100,
```
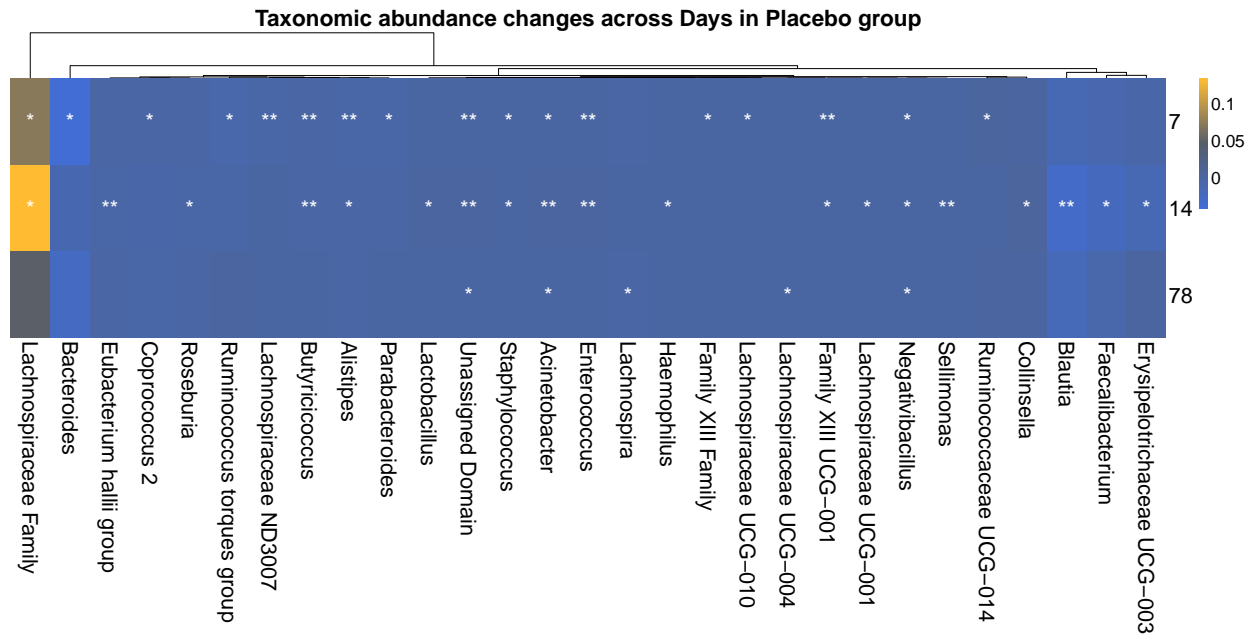
```
                fontsize=20,
                main="Taxonomic abundance changes across Days in Placebo group")
```


**Taxonomic abundance changes across Days in Placebo group**

```
#Getting the taxa with significantly different abundabce compared to Day 1 for the Treatment Group and
TreatmentChangeOnlySigs <- read_excel("Top100taxaVDMT.xlsx",
                                    sheet = "TreatmentTaxaChangeOnlySigs")
TreatmentChangeOnlySigspval <- read_excel("Top100taxaVDMT.xlsx",
                                        sheet = "TreatmentTaxaChangeOnlySigpval2")
TreatmentChangeOnlySigs2<-as.matrix(TreatmentChangeOnlySigs[,c(2:34)])
rownames(TreatmentChangeOnlySigs2)<-TreatmentChangeOnlySigs$Day
TreatmentChangeOnlySigspval2<-TreatmentChangeOnlySigspval[,c(2:34)]
TreatmentChangeOnlySigspval2<-replace(TreatmentChangeOnlySigspval2,is.na(TreatmentChangeOnlySigspval2),
rownames(TreatmentChangeOnlySigspval2)<-TreatmentChangeOnlySigspval$Day

#Plotting a heatmap with differentially abundant taxa
pheatmap(TreatmentChangeOnlySigs2,display_num=TreatmentChangeOnlySigspval2,
        color=colorRampPalette(c("#416DD4", "#565E6A", "#FFBC33"))(100),scale="none",
        cluster_rows = FALSE,show_rownames=TRUE,
        border_color=NA,
        fontsize_number=25,
        fontsize_row=25,
        fontsize_col= 20,
        number_color="white",
```

```
          cellheight=100,
          fontsize=20,
          main="Taxonomic abundance changes across Days in Treatment group")
```

Taxonomic abundance changes across Days in Treatment group
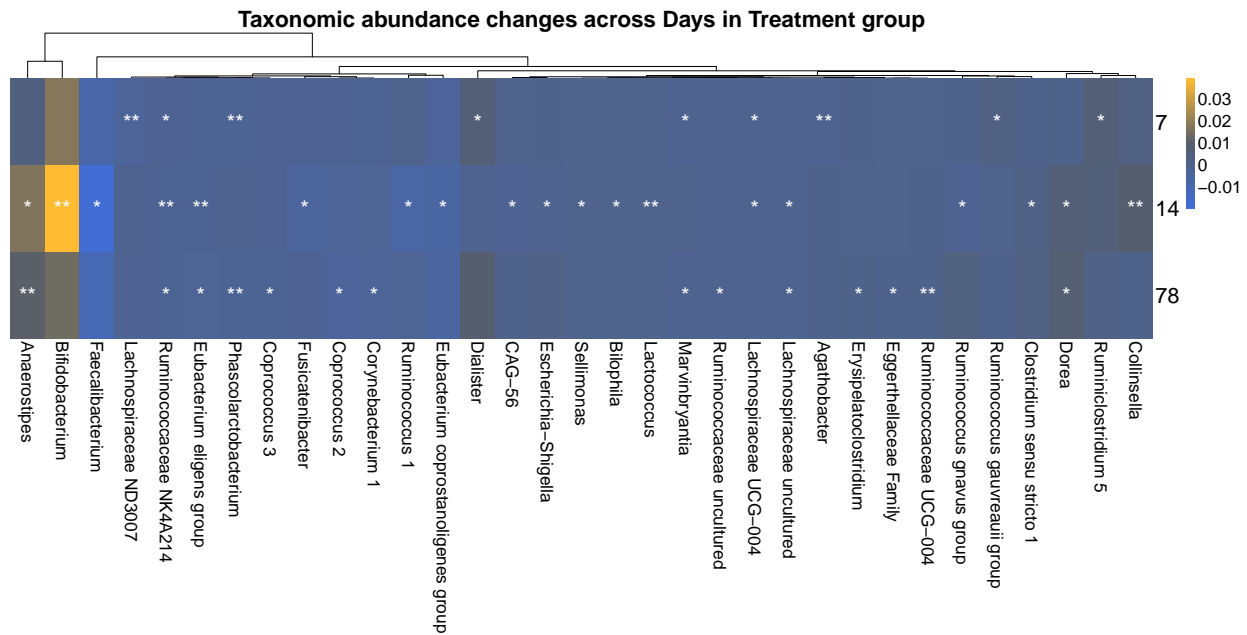

```
#Getting the top 5 taxa that significantly chnaged in abundance compared to Day 1 for Treatment group
TreatmentTaxaChange <- read_excel("Top100taxaVDMT.xlsx",sheet = "TreatmentTaxaChange2")%>%
  mutate(Change2=ifelse(Change>0,"Increase","Decrease"))
ggplot(TreatmentTaxaChange,aes(x=reorder(taxa,Change),y=Change))+
  geom_bar(stat="identity",color="black",aes(fill=Change2))+
  facet_grid(~Day,scales="free_x")+
  labs(x="Genus",y="Change in Relative Abundance",title="Top 5 taxa that changed significantly in abunda
  theme(
    panel.background = element_rect(fill="white"),
    text = element_text(size=25,color="black"),
    panel.border = element_rect(size=2,fill="NA"),
    axis.text.x=element_text(angle=90,hjust=1,vjust=1,color="black")
  )
```

Top 5 taxa that changed significantly in abundance in Treatment group
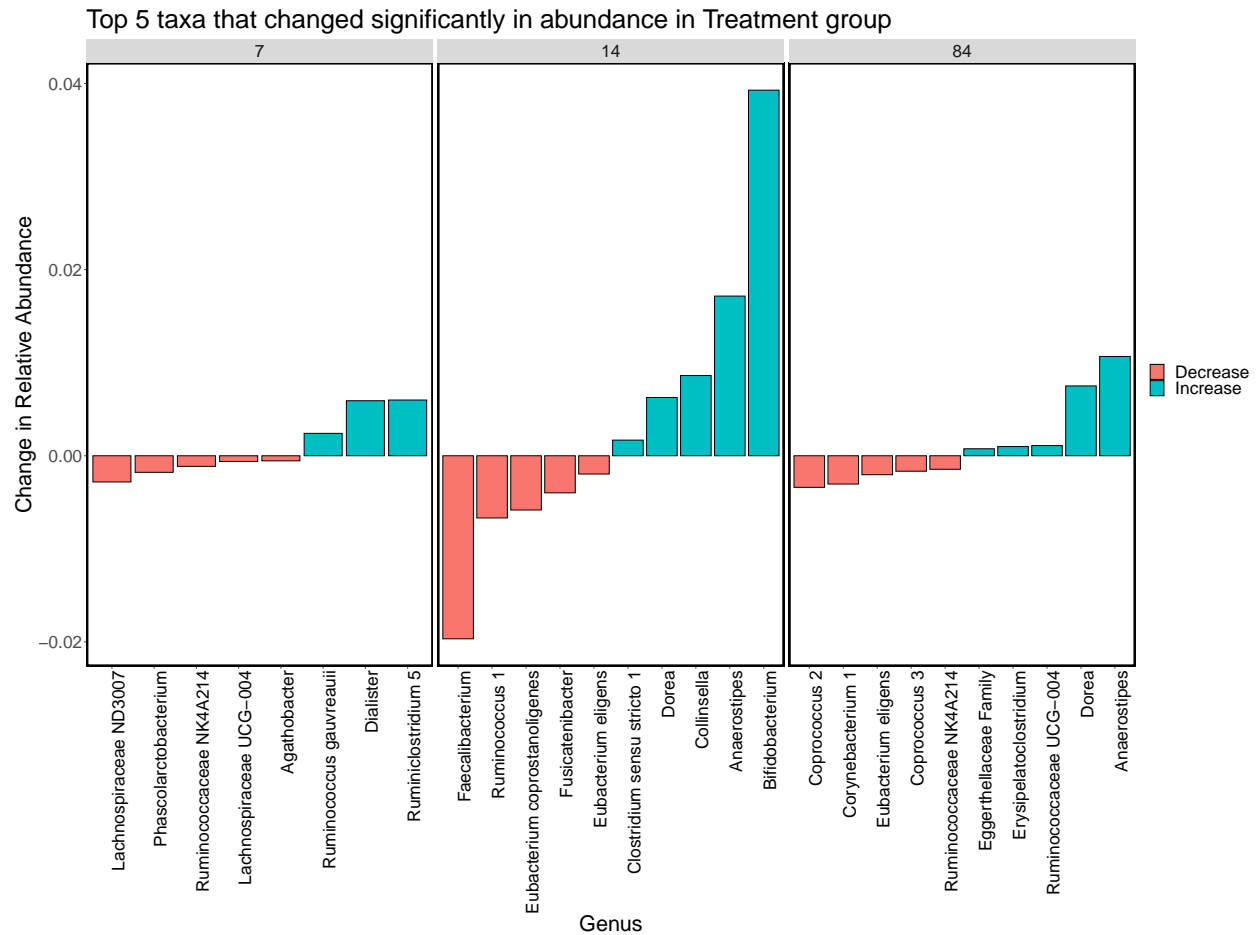

```
#Regression analysis of individual taxa with VitaminD change and Figure 6A-C
library(broom)
GenusVitChange<-read_excel("level-6.xlsx",sheet = "Genus2")%>%                    #getting abundance data of
  pivot_longer(-c("index","Sample_Name","Day","Subject","Cohort"))%>%
  inner_join(VitaminDChange[,c("Subject","VitD_pcchange")],by="Subject")%>% #joining the Taxa abundance
  group_by(Cohort,name)%>%
  mutate(meanAbundance=mean(value))%>%                         #filtering only those taxa that has a mean abundanc
  filter(meanAbundance>0.1)%>%
  mutate(ChangeCat=ifelse(VitD_pcchange>=50,">50%","<50%")) #creating a category variable that tells wh

Sig_taxa<-GenusVitChange %>%      #figuring out the taxa that changed their abundance significantly after
  filter(Cohort=="Treatment",Day>1)%>%
  nest(data=-name)%>%
  mutate(test=map(.x=data,~aov(value~ChangeCat,data=.x)%>%tidy))%>%
  unnest(test)%>%
  mutate(FDR=p.adjust(p.value,method="BH"))%>%
  filter(FDR<0.01)

#Plotting the abundance of those taxa as Figure 6A
GenusVitChange %>%
  inner_join(Sig_taxa,by="name")%>%
  filter(Cohort=="Treatment"&Day>1)%>%
  ggplot(aes(x=value,y=ChangeCat,color=ChangeCat,fill=ChangeCat))+
```

```
geom_boxplot(width=.35,color="black",size=2)+
geom_jitter(shape=21,size=7.5,stroke=1.5,color="black",position=position_jitterdodge(dodge.width=1,ji
facet_wrap(~name)+
xlab("log of Relative Abundance")+ylab("Change in Vitamin D")+
guides(color="none",fill="none")+
scale_x_log10()+
theme_classic()+
theme(text=element_text(size=20,color="black",face="bold"),
      strip.text=element_text(size=20,color="black",face="bold.italic"),
      axis.text = element_text(size=20,color="black",face="bold"))
```
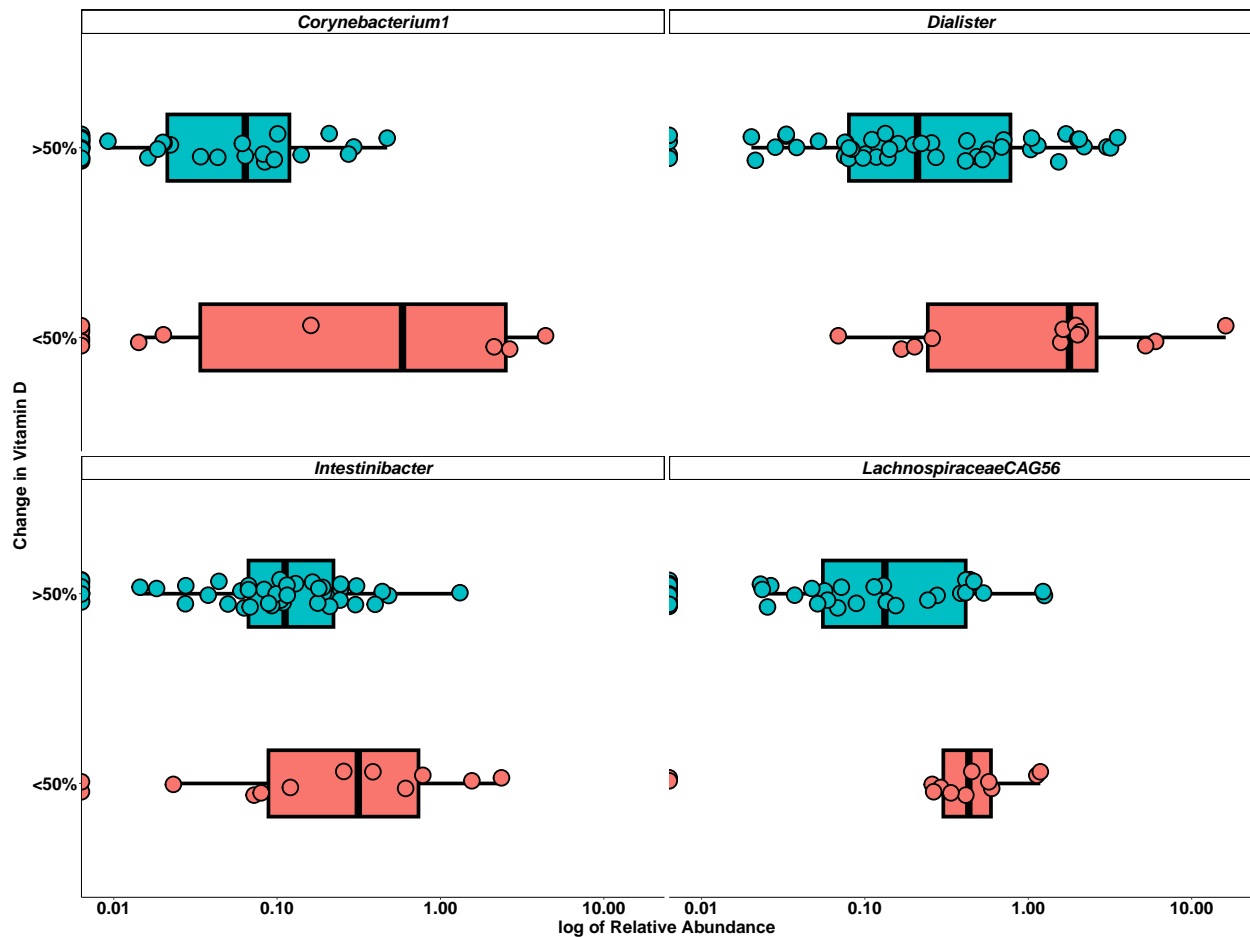


```
#Regression analysis to see whether the abundance of these taxa on Day 1 are a significant predictor of
GenusRegression<-read_excel("level-6.xlsx",sheet = "Genus2")%>% #getting the data
  inner_join(VitaminDChange[,c("Subject","VitD_pre","VitD_post","VitD_pcchange")],by="Subject")%>% #joi
  filter(Day==1&Cohort=="Treatment") %>%
  mutate(ChangeCat=ifelse(VitD_pcchange>=50,">50%","<50%"))

Regression<-lm(VitD_pcchange~VitD_pre+LachnospiraceaeCAG56+Dialister+Corynebacterium1+Intestinibacter,da
summary(Regression)
```

```
##
## Call:
```

```
## lm(formula = VitD_pcchange ~ VitD_pre + LachnospiraceaeCAG56 +
##     Dialister + Corynebacterium1 + Intestinibacter, data = GenusRegression)
##
## Residuals:
##     Min     1Q  Median     3Q     Max
## -90.289 -38.982  -1.883  19.029 112.495
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          185.4326    40.7449   4.551 0.000453 ***
## VitD_pre              -1.8388     0.9245  -1.989 0.066603 .
## LachnospiraceaeCAG56 -119.3113    68.0875  -1.752 0.101581
## Dialister             22.3105    19.6427   1.136 0.275104
## Corynebacterium1      14.0410    16.5517   0.848 0.410534
## Intestinibacter      -51.0897    47.4803  -1.076 0.300113
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 57.62 on 14 degrees of freedom
## Multiple R-squared:  0.4847, Adjusted R-squared:  0.3006
## F-statistic: 2.634 on 5 and 14 DF,  p-value: 0.07019
```

```
Regression<-lm(VitD_pcchange~VitD_pre+LachnospiraceaeCAG56+Dialister+Intestinibacter,data=GenusRegressi
summary(Regression)
```

```
##
## Call:
## lm(formula = VitD_pcchange ~ VitD_pre + LachnospiraceaeCAG56 +
##     Dialister + Intestinibacter, data = GenusRegression)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -88.89 -31.02  -8.49  21.00 109.75
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          193.4068    39.2735   4.925 0.000183 ***
## VitD_pre              -2.0152     0.8923  -2.258 0.039244 *
## LachnospiraceaeCAG56 -77.3097    46.2996  -1.670 0.115695
## Dialister             14.0736    16.9147   0.832 0.418440
## Intestinibacter      -47.7602    46.8735  -1.019 0.324396
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 57.08 on 15 degrees of freedom
## Multiple R-squared:  0.4582, Adjusted R-squared:  0.3137
## F-statistic: 3.171 on 4 and 15 DF,  p-value: 0.04476
```

```
library(car)
avPlots(Regression, col="red",pch=10,lwd=5,main="%Change in VitaminD against selected bacteria and basel
```

%Change in VitaminD against selected bacteria and baseline Vitamin D level
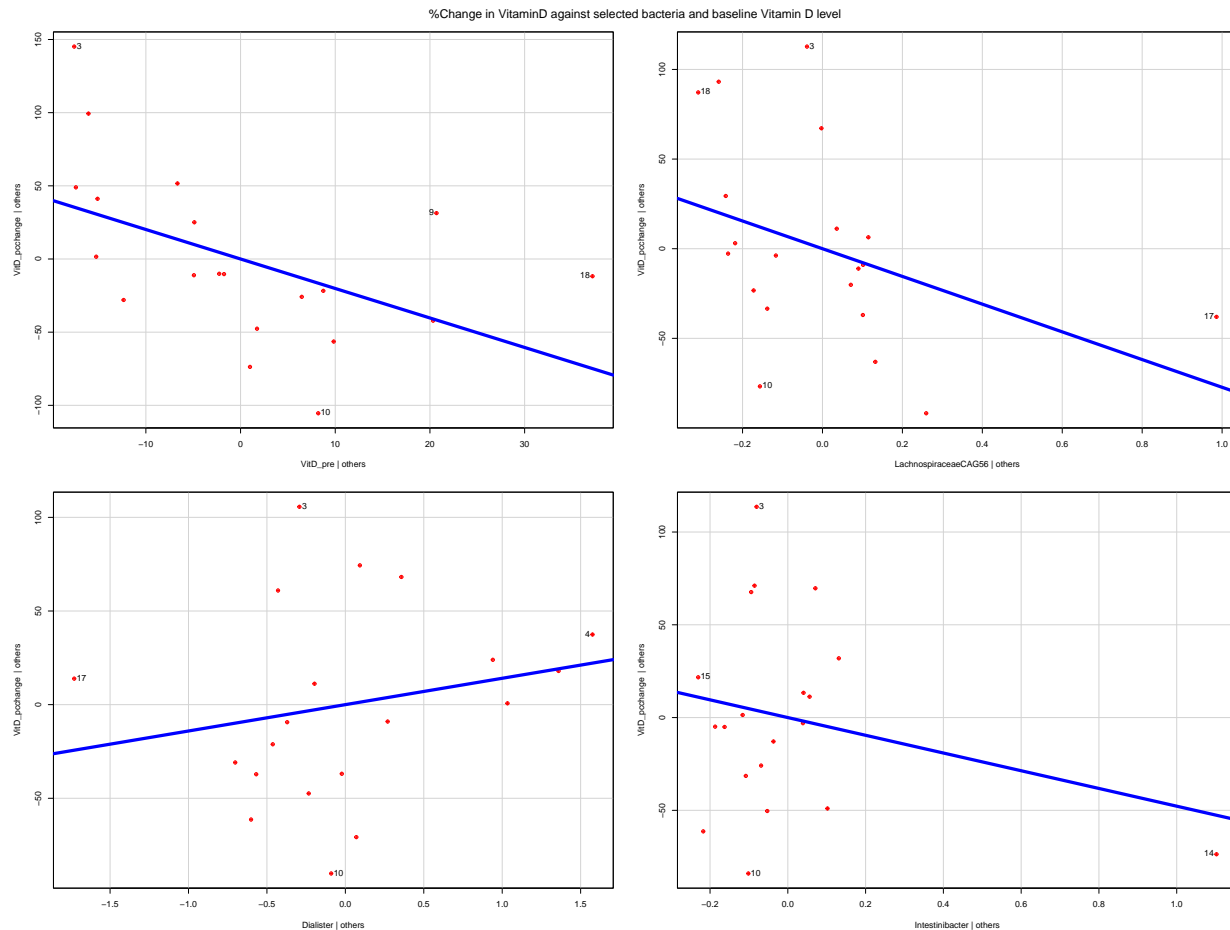
```
#Random forest analysis and Figure 6B&C

library(randomForest)
library(ROCR)

RandomForest<-read_excel("level-6.xlsx",sheet = "Genus2")%>%                    #getting abundance data of th
  inner_join(VitaminDChange[,c("Subject","VitD_pcchange")],by="Subject")%>% #joining the Taxa abundance
  mutate(ChangeCat=ifelse(VitD_pcchange>=50,">50%","<50%"))%>%
  filter(Day>1&Cohort=="Treatment") %>%
  select(-c(1,2,3,4,5,446))

set.seed(07181990)
rf<-randomForest(as.factor(ChangeCat)~.,data=RandomForest,ntree=10000,type="classification",proximity=TR

Imp<-as.data.frame(importance(rf))
Imp$variable<-rownames(Imp)
Imp<-Imp[order(Imp$MeanDecreaseGini,decreasing=TRUE),]%>%
  filter(MeanDecreaseGini>0.2)
Imp%>% #Figure 6B
  head(n=20)%>%
  ggplot(aes(x=MeanDecreaseGini*10,y=reorder(variable,MeanDecreaseGini)))+
  geom_bar(stat="identity",aes(fill=-MeanDecreaseGini*10),color="black")+
  labs(title="Top taxa contributing to the Random Forest model in\npredicting change of Vitamin D >50%"
```
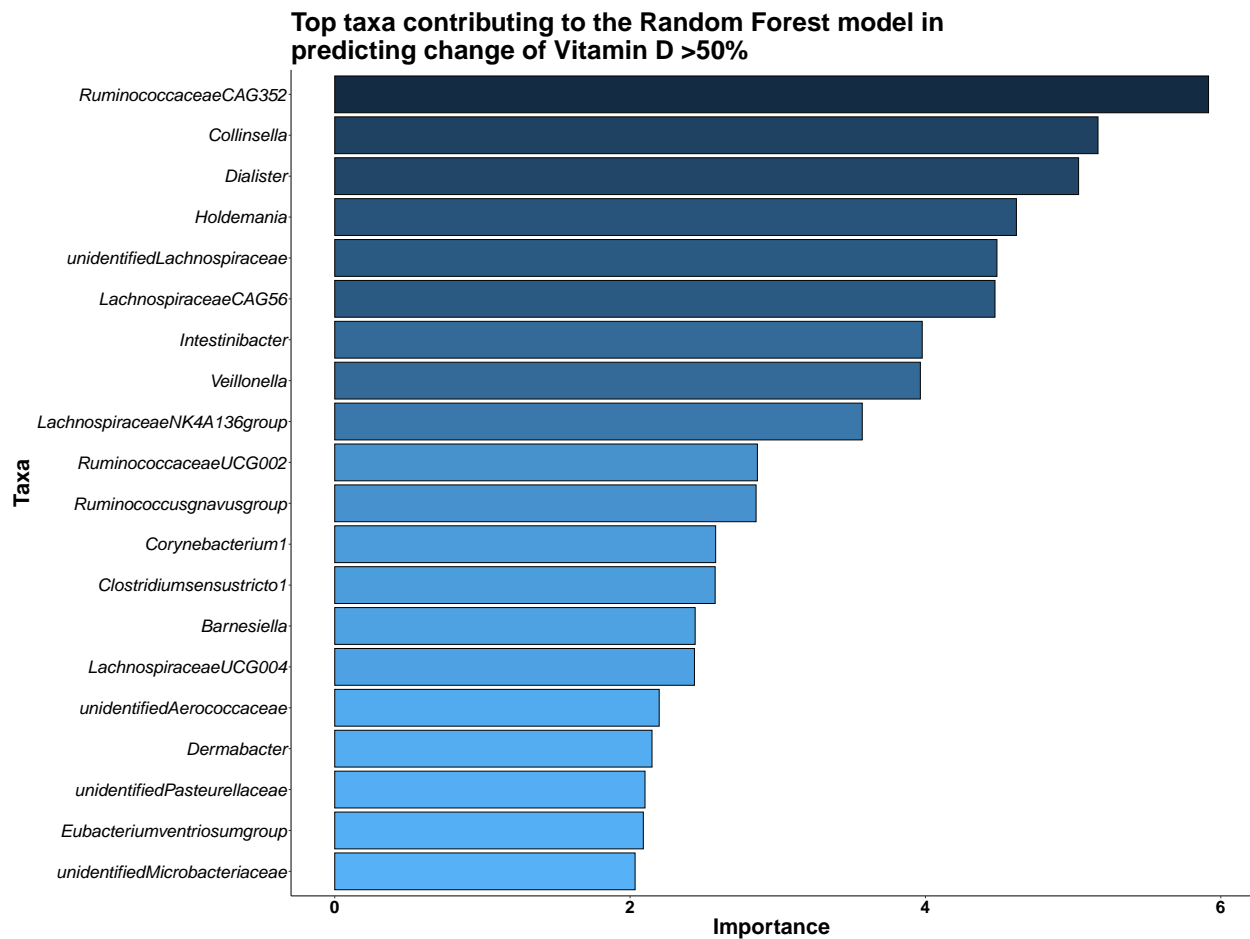
```
        x="Importance",y="Taxa",fill="")+
  guides(fill="none")+
  theme_classic()+
  theme(text=element_text(size=25,face="bold"),
        axis.text.y = element_text(size=20,face="italic",color="black"),
        legend.key.size=unit(1,"cm"),
        axis.text.x=element_text(size=20,face="bold",color="black"),)
```

Top taxa contributing to the Random Forest model in predicting change of Vitamin D >50%


```
RandomForest<-RandomForest%>%
  select(ChangeCat,Imp$variable)
rf<-randomForest(as.factor(ChangeCat)~.,data=RandomForest,ntree=10000,type="classification",proximity=T
rf
```

```
##
## Call:
##  randomForest(formula = as.factor(ChangeCat) ~ ., data = RandomForest,     ntree = 10000, type = "cl
##                Type of random forest: classification
##                      Number of trees: 10000
## No. of variables tried at each split: 4
##
##          OOB estimate of  error rate: 11.67%
## Confusion matrix:
##       <50% >50% class.error
```

```
## <50%     6     6  0.50000000
## >50%     1    47  0.02083333
```

```
pred1=predict(rf,type = "prob")
perf = prediction(pred1[,2], RandomForest$ChangeCat)
# 1. Area under curve
auc = performance(perf, "auc")
auc
```

```
## A performance instance
##   'Area under the ROC curve'
```

```
# 2. True Positive and Negative Rate
pred3 = performance(perf, "tpr","fpr")
auc_ROCR <- performance(perf, measure = "auc")
auc_ROCR@y.values[[1]]
```

```
## [1] 0.9444444
```

```
pred2<-as.data.frame(cbind(pred3@x.values[[1]],pred3@y.values[[1]]))
# 3. Plotting the ROC curve- Figure 6C
ggplot(pred2,aes(x=V1,y=V2))+
  geom_line(color="red",size=3)+
  geom_abline(slope=1,size=2,color="black",linetype=2)+
  geom_text(x=0.5,y=0.75,label=paste("AUCROC = ",round(auc_ROCR@y.values[[1]],3)),size=7.5,color="black"
  labs(title="ROC of the Random Forest Model",x="False Positive rate",y="True Positive rate")+
  theme_classic()+
  theme(text=element_text(size=25,face="bold"),
        axis.text= element_text(size=20,face="bold",color="black"),
        panel.border=element_rect(linewidth=2,color="black",fill=NA))
```

ROC of the Random Forest Model

AUCROC = 0.944