PICRUSt_Analysis

Jessica McClintock

2022-09-08

Setup

```
suppressPackageStartupMessages({
  # General use
  library(tidyverse)
  library(knitr)
  library(ggplot2)

  # Data formats
  library(SummarizedExperiment)
  library(ALDEx2)
  })
```

Read in PICRUSt Object

```
picrustSE <- readRDS("/restricted/projectnb/infant-microbiome/work/jessmcc/rsv/PICRUSt/p
icrust2_output/FinalDatPICRUSt2.RDS")
```

Group RSV+ and RSV- Samples

```
pos_samples <- picrustSE[,picrustSE$state == "pos"]
neg_samples <- picrustSE[,picrustSE$state == "neg"]
```

#Run ALDEx2 Analysis

```
#have output that includes pvalues of which pathways are significant between groups
# Round PICRUSt counts so that counts are integers
picrustSE@assays@data@listData[["counts"]] <- round(picrustSE@assays@data@listData[["cou
nts"]], 0)

aldex_results <- aldex(picrustSE@assays@data@listData[["counts"]], picrustSE@colData@lis
tData[["state"]], test="t", effect = TRUE)
```

```
## aldex.clr: generating Monte-Carlo instances and clr values
```
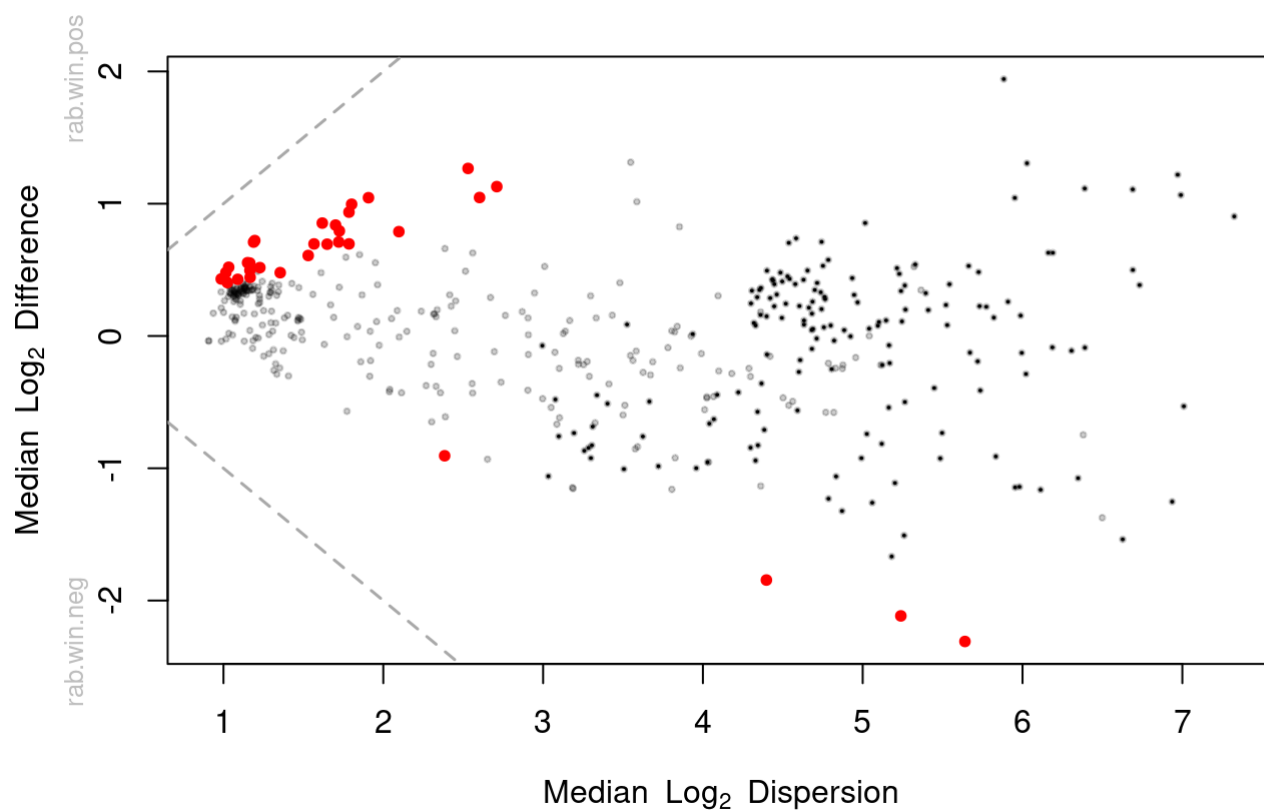
```
## operating in serial mode
```

```
## computing center with all features
```

```
## aldex.ttest: doing t-test
```

```
## aldex.effect: calculating effect sizes
```

#Plot ALDEx2 Effect Plot

```
ALDEx2::aldex.plot(aldex_results, type="MW", test="wilcox", called.cex = 1, cutoff = 0.0
5)
```


#Select Significant Pathways

```r
#Filter for pathways with a significant difference
sig_pathways <- aldex_results %>%
  filter(wi.eBH < 0.05) %>%
  arrange(effect, wi.eBH)

#Add pathway descriptions
path_descriptions <- as.data.frame(picrustSE@elementMetadata@listData[["description"]],
 picrustSE@elementMetadata@listData[["pathway"]])

sig_path_descriptions <- merge(sig_pathways, path_descriptions, by = 0)
sig_path_descriptions
```

```
##                           Row.names  rab.all rab.win.neg rab.win.pos  diff.btw diff.win
## 1    BIOTIN-BIOSYNTHESIS-PWY 5.386322    5.201270    5.895971 0.8534061 1.618390
## 2         DAPLYSINESYN-PWY 5.838353    5.723964    6.131895 0.4787923 1.014847
## 3       FASYN-INITIAL-PWY 5.610629    5.292962    5.812481 0.6960115 1.785401
## 4              FUCCAT-PWY 4.444876    4.171074    4.937940 1.1289643 2.710202
## 5      KDO-NAGLIPASYN-PWY 4.604709    4.345618    5.136526 1.0465327 2.602340
## 6      NAD-BIOSYNTHESIS-II 4.913411    4.584459    5.461877 1.2664139 2.530373
## 7           NAGLIPASYN-PWY 4.900116    4.560289    5.585219 0.9958791 1.801877
## 8            NONMEVIPP-PWY 5.312348    5.002274    5.780315 0.7096197 1.188393
## 9                 PWY-1269 5.032733    4.694203    5.604875 0.9360155 1.785364
## 10                PWY-5097 5.807268    5.676830    6.119703 0.5195598 1.032711
## 11                PWY-5971 5.834528    5.655009    6.129505 0.6081991 1.530221
## 12                PWY-5989 5.666993    5.428626    6.009527 0.7938297 1.725133
## 13                PWY-6282 5.748505    5.489181    6.044548 0.7102605 1.721162
## 14                PWY-6467 4.933951    4.484461    5.609206 1.0451808 1.907328
## 15                PWY-6519 5.485049    5.247978    5.955945 0.8383233 1.701405
## 16                PWY-7539 5.937436    5.747794    6.375354 0.5538256 1.151010
## 17                PWY-7560 5.311656    5.002191    5.778669 0.7212208 1.195910
## 18                PWY-7664 5.810215    5.610144    6.136857 0.6958158 1.566600
## 19                PWYG-321 5.837844    5.591490    6.128445 0.6938195 1.648655
## 20            RIBOSYN2-PWY 5.841550    5.661319    6.215660 0.5513168 1.164577
##       effect   overlap       we.ep       we.eBH        wi.ep       wi.eBH
## 1  0.4929811 0.2781250 1.121933e-04 0.013574869 9.843094e-05 0.006997105
## 2  0.4326429 0.3184375 2.966091e-04 0.014228959 1.424768e-03 0.034049801
## 3  0.3790364 0.3061543 7.958000e-04 0.020044427 1.140688e-03 0.029610625
## 4  0.3624188 0.3214620 1.379609e-03 0.029944252 2.035688e-03 0.044559306
## 5  0.3359745 0.3099032 8.584646e-03 0.071659793 1.027326e-03 0.028083500
## 6  0.4524258 0.2827242 2.365367e-04 0.013648764 1.497103e-04 0.008581840
## 7  0.4952275 0.2461731 3.036937e-04 0.014158395 1.623424e-05 0.003075624
## 8  0.5463095 0.2618750 2.302171e-05 0.005463957 5.663170e-05 0.005385683
## 9  0.4949001 0.2553125 5.231867e-04 0.016898697 1.982344e-05 0.003092822
## 10 0.4690758 0.3084375 1.496853e-04 0.013574869 8.581246e-04 0.026303427
## 11 0.3687973 0.3036551 3.716913e-03 0.061792523 7.579198e-04 0.024936339
## 12 0.4276664 0.2840625 2.479880e-04 0.013686526 2.459445e-04 0.012461032
## 13 0.3979168 0.2942830 4.945335e-04 0.016759369 4.599249e-04 0.020452584
## 14 0.4734786 0.2471103 1.678507e-04 0.013623691 1.553915e-05 0.003075624
## 15 0.4472396 0.2796875 2.111264e-04 0.013633748 1.128769e-04 0.007330908
## 16 0.4571878 0.3093750 4.538462e-04 0.016078471 1.099011e-03 0.029011276
## 17 0.5533461 0.2696033 2.324453e-05 0.005463957 5.858455e-05 0.005393040
## 18 0.4148392 0.3009375 6.115849e-04 0.017660275 6.053351e-04 0.022901019
## 19 0.3970832 0.3002187 6.179010e-04 0.017696848 6.292616e-04 0.023262265
## 20 0.4422152 0.3120900 4.120428e-04 0.015889292 1.075352e-03 0.028677474
##                  picrustSE@elementMetadata@listData[["description"]]
## 1                                              biotin biosynthesis I
## 2                                            L-lysine biosynthesis I
## 3            superpathway of fatty acid biosynthesis initiation (E. coli)
## 4                                                   fucose degradation
## 5                           superpathway of (Kdo)2-lipid A biosynthesis
## 6                                                  NAD salvage pathway II
## 7                                                  lipid IVA biosynthesis
## 8                                    methylerythritol phosphate pathway I
## 9                       CMP-3-deoxy-D-manno-octulosonate biosynthesis I
```

```
## 10                                                          L-lysine biosynthesis VI
## 11                              palmitate biosynthesis II (bacteria and plants)
## 12                               stearate biosynthesis II (bacteria and plants)
## 13                   palmitoleate biosynthesis I (from (5Z)-dodec-5-enoate)
## 14                                     Kdo transfer to lipid IVA III (Chlamydia)
## 15                                       8-amino-7-oxononanoate biosynthesis I
## 16 6-hydroxymethyl-dihydropterin diphosphate biosynthesis III (Chlamydia)
## 17                                       methylerythritol phosphate pathway II
## 18                                             oleate biosynthesis IV (anaerobic)
## 19                                                         mycolate biosynthesis
## 20                               flavin biosynthesis I (bacteria and plants)
```

#Boxplot of significant pathways

```
#Select Significant Pathways to plot
sigPathNames <- row.names(sig_pathways)

countsToPlot <- assays(picrustSE)$counts[rownames(picrustSE@assays@data@listData[["count
s"]]) %in% sigPathNames, ]

countsToPlot <- t(countsToPlot) %>% as.data.frame()

samples <- rownames(countsToPlot)
countsToPlot <- cbind(samples, colData(picrustSE)$state, countsToPlot)
colnames(countsToPlot)[2] <- "state"

for (item in 1:length(countsToPlot$samples)){
  if(countsToPlot$state[item] == "pos"){
    countsToPlot$samples[item] <- paste0("+", countsToPlot$samples[item])
  }else if(countsToPlot$state[item] == "neg"){
    countsToPlot$samples[item] <- paste0("-", countsToPlot$samples[item])
  }
}

#Create Tidy Data
 countsToPlot <- countsToPlot %>%
  pivot_longer(cols = !c("state", "samples") , names_to = "pathway", values_to = "Abunda
nce")
#
# #plot in a box plot
ggplot(countsToPlot, aes(x=state, y=Abundance, fill = state)) +
  geom_boxplot()+
  facet_wrap(~pathway)
```
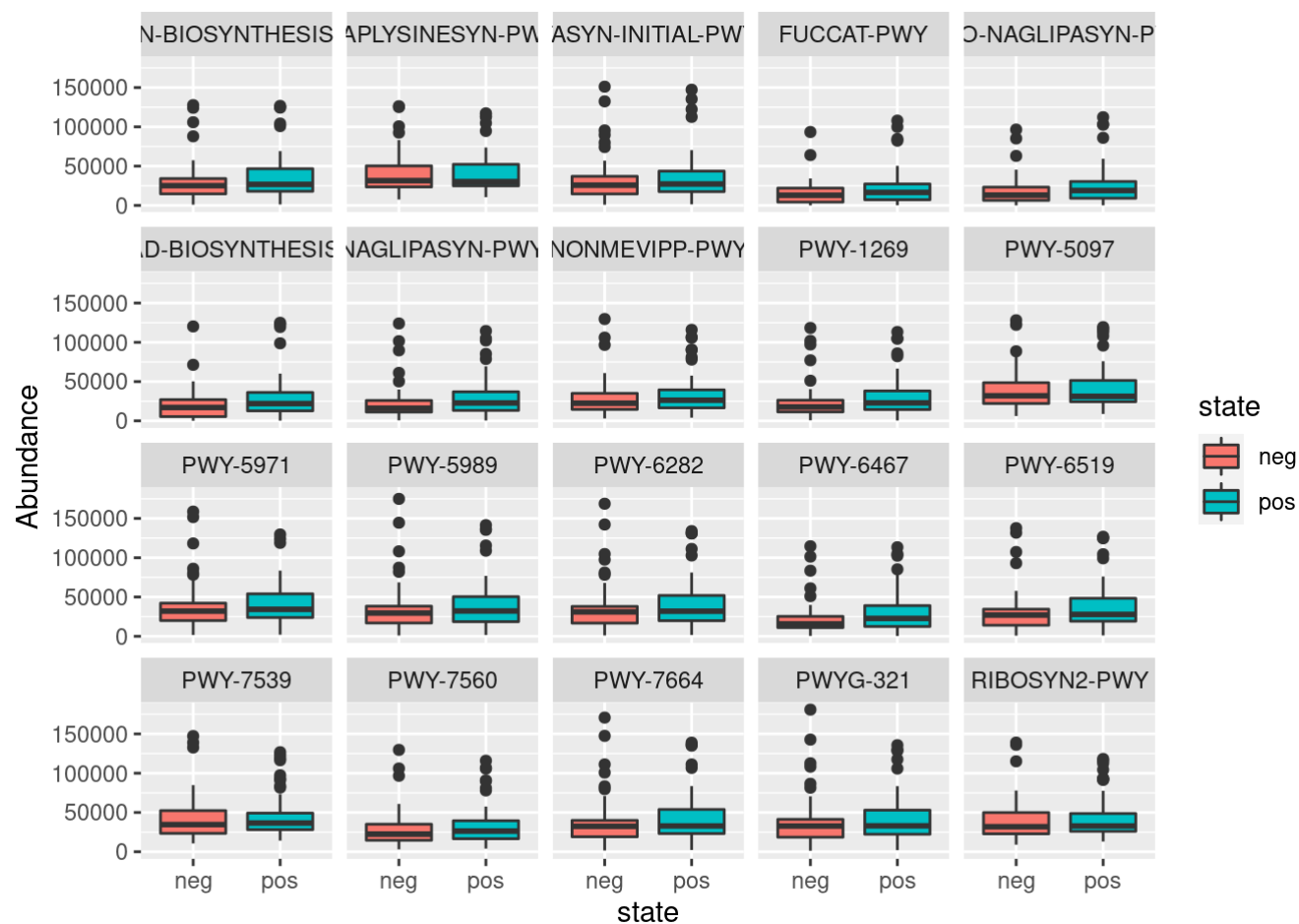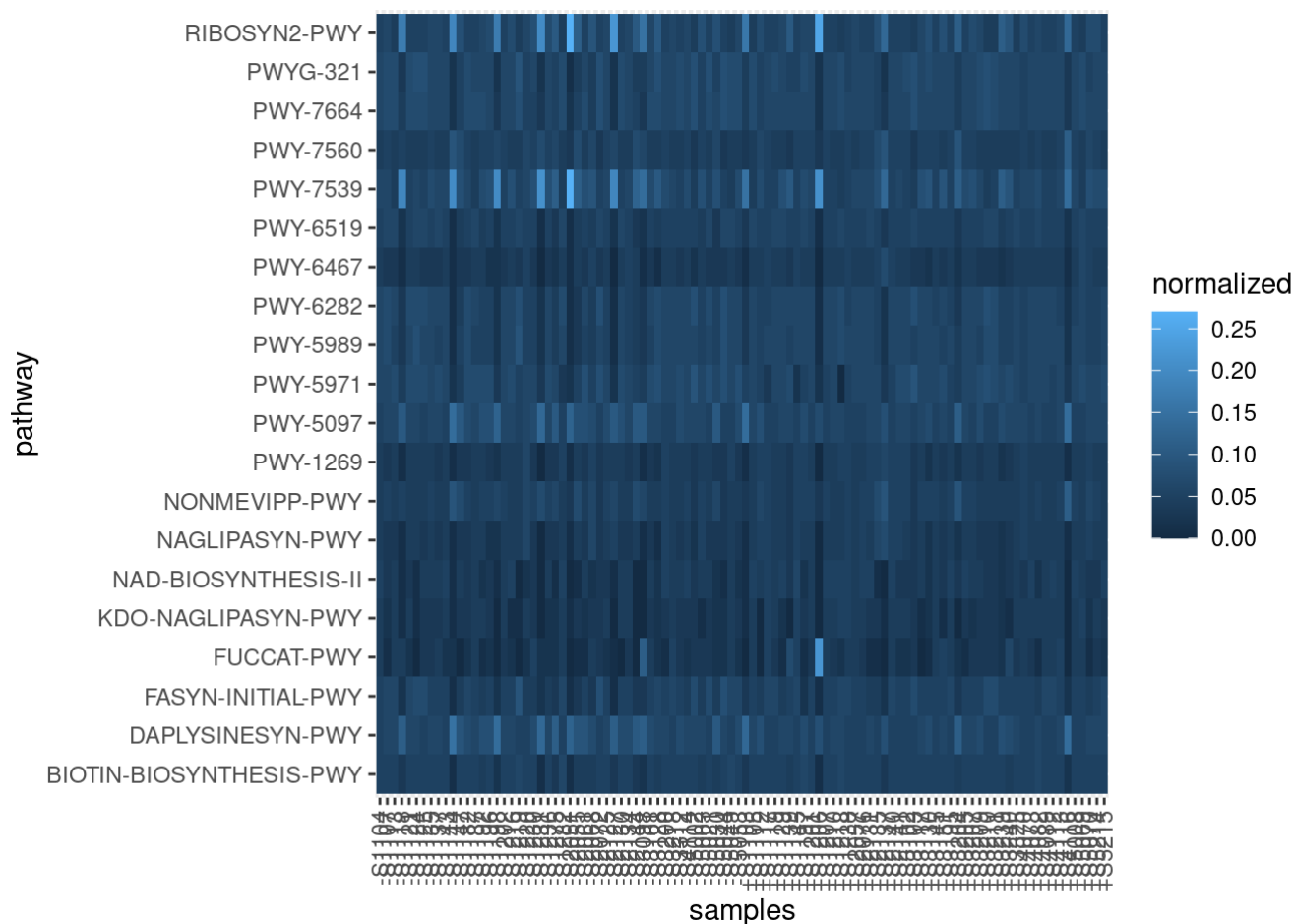
#Heatmap

```
countsToPlot <- countsToPlot %>%
  group_by(samples) %>%
  mutate(summation = sum(Abundance)) %>%
  mutate(normalized = round((Abundance/summation), digits = 2))

ggplot(countsToPlot, aes(samples, pathway)) +
  geom_tile(aes(fill = normalized))+
  scale_x_discrete(guide = guide_axis(angle = 90))
```
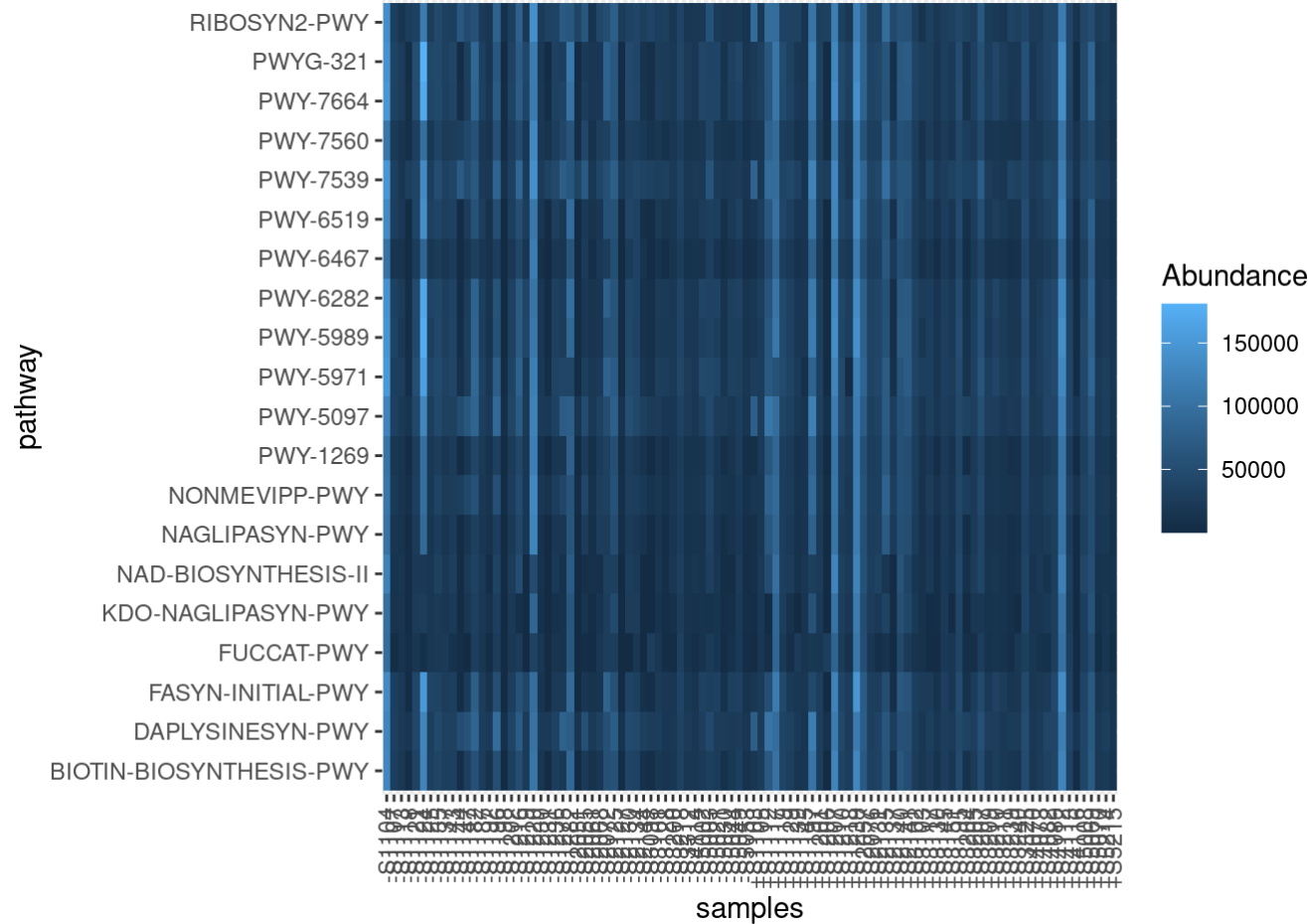
```
#scale_fill_discrete(name = "Abundance")

#group_by(~state)
```

```
ggplot(countsToPlot, aes(samples, pathway)) +
  geom_tile(aes(fill = Abundance))+
  scale_x_discrete(guide = guide_axis(angle = 90))
```

```
#scale_fill_discrete(name = "Abundance")
```