Lab3: Antibiotics, microbiome and obesity case study

Study design

- Expose mothers to sub-therapeutic antibiotic treatment (STAT)

<table>
<thead>
<tr>
<th>Age (wk)</th>
<th>0</th>
<th>4</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
<th>26</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body composition (DEXA):</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Microbiome samples:</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>C, I</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hormone levels:</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Control

STAT Penicillin
Data

- Mapping file:
  - mapping.txt
- Metabolic phenotypes
  - phenotype.txt
- OTU table
  - otu_table.txt

Load and normalize the data similarly to Lab1

d = read.table('otu_table.txt', header=T, row.names=1, sep='\t')
taxa.names = d$Consensus.Lineage

d=as.matrix(d[, -dim(d)[2]])
d=scale(d, center=F, scale=colSums(d))
d=t(d)

ff=read.table('mapping.txt', header=T)
ff=ff[order(ff$SampleID),]
In addition load and examine the phenotype data

```r
phen = read.table('phenotype.txt', header=T, sep='\t', row.names=1)
phen = t(phen)
phen
```

Obtain genus-level OTU table using functions from Lab 1

```r
d.genus =
  otu2taxonomy(d, level=7, taxa=taxa.names)

#subset the data to only Fecal samples with mean abundance >1%

d.genus = subset(d.genus, ff$Treatment == 'Fecal')
ff = subset(ff, ff$Treatment == 'Fecal')
```
Performing multiple tests using \texttt{apply(...)}

- Run: \texttt{help(apply)}
- This function allows to automatically apply another function on all rows or columns of a data matrix
- E.g. \texttt{apply(d.genus, 2, mean)}

\textbf{Wilcoxon tests for taxa association with treatment with FDR}

\begin{verbatim}
  d.genus4 = subset(d.genus, ff$Week == 4)
  ff4 = subset(ff, ff$Week == 4)

  run.wilc4 = function(x){
    wilcox.test(x~ff4$Treatment, exact=F)$p.value
  }

  pval4 = apply(d.genus4, 2, run.wilc4)
  fdr4 = p.adjust(pval4, method='fdr')
  plot(pval4, fdr4, xlim=c(0,1), ylim=c(0,1))
\end{verbatim}

Perform the same procedure for week 16 and week 26 samples. At which week there seem to be more taxa associated with the treatment at 10\% fdr?
Kruskal-Wallis Tests for changes in abundance over time

```r
ff$Week = factor(ff$Week, levels = c(4,16,26),
               ordered=T)
# for treated group
d.genusP = subset(d.genus, ff$Treatment == 'P')
ffP = subset(ff, ff$Treatment == 'P')

run.kruskP = function(x){
  kruskal.test(x~ffP$Week)$p.value
}
pvalP = apply(d.genusP, 2, run.kruskP)
fdrP = p.adjust(pvalP)
plot(pvalP, fdrP, xlim=c(0,1), ylim=c(0,1))
```

Run the same procedure for control mice. Which taxa change significantly (at 5% fdr in both treatment groups)? Do boxplots to demonstrate this effect.

Download the solutions file

- lab3.R
Rank correlation tests with metabolic phenotypes

taxon1C = d.genus26[ff26$Treatment=='C', signif.in.both[1]]
taxon2C = d.genus26[ff26$Treatment=='C', signif.in.both[2]]

run.cor1C = function(x){
    cor.test(x, taxon1C, method='kendall', exact=F)$p.value
}
corpval1C = apply(phen[ff26$Treatment=='C',], 2, run.cor1C)
corfdr1C = p.adjust(corpval1C, method='fdr')

run.cor2C = function(x){
    cor.test(x, taxon2C, method='kendall', exact=F)$p.value
}
corpval2C = apply(phen[ff26$Treatment]=='C',], 2, run.cor2C)
corfdr2C = p.adjust(corpval2C, method='fdr')

Do the same for penicillin. Are any of the metabolic variables correlated with the two taxa?